

MP5_doubleDQN

May 6, 2024

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.

```
[ ]: # you will be prompted with a window asking to grant permissions
from google.colab import drive
drive.mount("/content/drive")
```

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

```
[ ]: # fill in the path in your Google Drive in the string below. Note: do not
↳escape slashes or spaces
import os
datadir = "/content/drive/MyDrive/CS444/assignment5/assignment5_materials"
if not os.path.exists(datadir):
    !ln -s "/content/drive/MyDrive/CS444/assignment5/assignment5_materials"
    ↳$datadir
os.chdir(datadir)
!pwd
```

/content/drive/MyDrive/CS444/assignment5/assignment5_materials

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

Requirement already satisfied: gym in /usr/local/lib/python3.10/dist-packages (0.25.2)

Requirement already satisfied: pyvirtualdisplay in /usr/local/lib/python3.10/dist-packages (3.0)

Requirement already satisfied: numpy>=1.18.0 in /usr/local/lib/python3.10/dist-packages (from gym) (1.25.2)

Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from gym) (2.2.1)

Requirement already satisfied: gym-notices>=0.0.4 in /usr/local/lib/python3.10/dist-packages (from gym) (0.0.8)

Reading package lists... Done

Building dependency tree... Done
Reading state information... Done
E: Unable to locate package python-opengl

```
[ ]: !pip3 install --upgrade setuptools --user
!pip3 install ez_setup
!pip3 install gym[atari]
!pip3 install gym[accept-rom-license]
```

Requirement already satisfied: setuptools in /root/.local/lib/python3.10/site-packages (69.5.1)

Collecting ez_setup

Using cached ez_setup-0.9.tar.gz (6.6 kB)

Preparing metadata (setup.py) ... done

Building wheels for collected packages: ez_setup

Building wheel for ez_setup (setup.py) ... done

Created wheel for ez_setup: filename=ez_setup-0.9-py3-none-any.whl size=10995 sha256=cad5c9b467baac038ed7f971a347c1a20689c66aa090df38bf1f97a95c97039e

Stored in directory: /root/.cache/pip/wheels/7a/d6/77/8f495e85fb7df23d41c328b9ea3cf0d9e83631b20bba479293

Successfully built ez_setup

Installing collected packages: ez_setup

Successfully installed ez_setup-0.9

Requirement already satisfied: gym[atari] in /usr/local/lib/python3.10/dist-packages (0.25.2)

Requirement already satisfied: numpy>=1.18.0 in /usr/local/lib/python3.10/dist-packages (from gym[atari]) (1.25.2)

Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from gym[atari]) (2.2.1)

Requirement already satisfied: gym-notices>=0.0.4 in /usr/local/lib/python3.10/dist-packages (from gym[atari]) (0.0.8)

Requirement already satisfied: ale-py~=0.7.5 in /usr/local/lib/python3.10/dist-packages (from gym[atari]) (0.7.5)

Requirement already satisfied: importlib-resources in /usr/local/lib/python3.10/dist-packages (from ale-py~=0.7.5->gym[atari]) (6.4.0)

Requirement already satisfied: gym[accept-rom-license] in /usr/local/lib/python3.10/dist-packages (0.25.2)

Requirement already satisfied: numpy>=1.18.0 in /usr/local/lib/python3.10/dist-packages (from gym[accept-rom-license]) (1.25.2)

Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from gym[accept-rom-license]) (2.2.1)

Requirement already satisfied: gym-notices>=0.0.4 in /usr/local/lib/python3.10/dist-packages (from gym[accept-rom-license]) (0.0.8)

Collecting autorom[accept-rom-license]~=0.4.2 (from gym[accept-rom-license])

Downloading AutoROM-0.4.2-py3-none-any.whl (16 kB)

Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license]) (8.1.7)

Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-

```
packages (from autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license])
(2.31.0)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages
(from autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license]) (4.66.2)
Collecting AutoROM.accept-rom-license (from autorom[accept-rom-
license]~=0.4.2->gym[accept-rom-license])
  Downloading AutoROM.accept-rom-license-0.6.1.tar.gz (434 kB)
    434.7/434.7
```

```
kB 10.8 MB/s eta 0:00:00
```

```
  Installing build dependencies ... done
  Getting requirements to build wheel ... done
  Preparing metadata (pyproject.toml) ... done
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.10/dist-packages (from requests->autorom[accept-rom-
license]~=0.4.2->gym[accept-rom-license]) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-
packages (from requests->autorom[accept-rom-license]~=0.4.2->gym[accept-rom-
license]) (3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.10/dist-packages (from requests->autorom[accept-rom-
license]~=0.4.2->gym[accept-rom-license]) (2.0.7)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.10/dist-packages (from requests->autorom[accept-rom-
license]~=0.4.2->gym[accept-rom-license]) (2024.2.2)
Building wheels for collected packages: AutoROM.accept-rom-license
  Building wheel for AutoROM.accept-rom-license (pyproject.toml) ...
done
  Created wheel for AutoROM.accept-rom-license:
filename=AutoROM.accept_rom_license-0.6.1-py3-none-any.whl size=446659
sha256=a04f0363cfae61ad8f1e8eba9c170783c11cc170881210b45ee5e0e23f5bf1e6
  Stored in directory: /root/.cache/pip/wheels/6b/1b/ef/a43ff1a2f1736d5711faa1ba
4c1f61be1131b8899e6a057811
Successfully built AutoROM.accept-rom-license
Installing collected packages: AutoROM.accept-rom-license, autorom
Successfully installed AutoROM.accept-rom-license-0.6.1 autorom-0.4.2
```

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

```
[ ]: %matplotlib inline

import sys
import gym
import torch
import pylab
import random
```

```

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

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

```

```

/usr/local/lib/python3.10/dist-packages/skimage/util/dtype.py:27:
DeprecationWarning: `np.bool8` is a deprecated alias for `np.bool_`.
(Deprecated NumPy 1.24)
  np.bool8: (False, True),

```

1.1 Understanding the environment

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

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

```

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

```

```

/usr/local/lib/python3.10/dist-packages/gym/core.py:317: DeprecationWarning:
WARN: Initializing wrapper in old step API which returns one bool instead
of two. It is recommended to set `new_step_api=True` to use new step API. This
will be the default behaviour in future.
  deprecation(
/usr/local/lib/python3.10/dist-
packages/gym/wrappers/step_api_compatibility.py:39: DeprecationWarning:
WARN: Initializing environment in old step API which returns one bool
instead of two. It is recommended to set `new_step_api=True` to use new step
API. This will be the default behaviour in future.
  deprecation(

```

```

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

```

```

/usr/local/lib/python3.10/dist-packages/gym/utils/passive_env_checker.py:227:
DeprecationWarning: WARN: Core environment is written in old step API which
returns one bool instead of two. It is recommended to rewrite the environment
with new step API.
  logger.deprecation(
/usr/local/lib/python3.10/dist-packages/gym/utils/passive_env_checker.py:233:
DeprecationWarning: `np.bool8` is a deprecated alias for `np.bool_`.
(Deprecated NumPy 1.24)
  if not isinstance(done, (bool, np.bool8)):

```

1.2 Creating a DQN Agent

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

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

Frame : Number of frames processed in total.

Memory Size : The current size of the replay memory.

```

[ ]: double_dqn = 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
    state = env.reset()

```

```

next_state = state
life = number_lives

get_init_state(history, state, HISTORY_SIZE)

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 = torch.tensor([[0]]).cuda()
    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['lives'])

    life = info['lives']
    r = reward

    # Store the transition in memory
    agent.memory.push(deepcopy(frame_next_state), action.cpu(), r,
    ↪terminal_state)

    # Start training after random sample generation
    if (frame >= train_frame): # You can set train_frame to a lower value
    ↪while testing your starts training earlier
        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_double_dqn.
↳pth")

            best_eval_reward = np.mean(evaluation_reward)

```

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283:

DeprecationWarning: `should_run_async` will not call `transform_cell` automatically in the future. Please pass the result to `transformed_cell` argument and any exception that happen during the transform in `preprocessing_exc_tuple` in IPython 7.17 and above.

```

    and should_run_async(code)

```

```

episode: 0   score: 0.0   memory length: 124   epsilon: 1.0   steps: 124   lr:
0.0001      evaluation reward: 0.0
episode: 1   score: 1.0   memory length: 297   epsilon: 1.0   steps: 173   lr:
0.0001      evaluation reward: 0.5
episode: 2   score: 0.0   memory length: 421   epsilon: 1.0   steps: 124   lr:
0.0001      evaluation reward: 0.3333333333333333
episode: 3   score: 4.0   memory length: 719   epsilon: 1.0   steps: 298   lr:
0.0001      evaluation reward: 1.25
episode: 4   score: 5.0   memory length: 1035   epsilon: 1.0   steps: 316
lr: 0.0001   evaluation reward: 2.0
episode: 5   score: 3.0   memory length: 1301   epsilon: 1.0   steps: 266
lr: 0.0001   evaluation reward: 2.1666666666666665
episode: 6   score: 4.0   memory length: 1576   epsilon: 1.0   steps: 275
lr: 0.0001   evaluation reward: 2.4285714285714284
episode: 7   score: 1.0   memory length: 1746   epsilon: 1.0   steps: 170
lr: 0.0001   evaluation reward: 2.25
episode: 8   score: 3.0   memory length: 1995   epsilon: 1.0   steps: 249
lr: 0.0001   evaluation reward: 2.3333333333333335
episode: 9   score: 1.0   memory length: 2165   epsilon: 1.0   steps: 170
lr: 0.0001   evaluation reward: 2.2
episode: 10  score: 4.0   memory length: 2461   epsilon: 1.0   steps: 296
lr: 0.0001   evaluation reward: 2.3636363636363638

```

episode: 11	score: 2.0	memory length: 2663	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 2.3333333333333335			
episode: 12	score: 1.0	memory length: 2832	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 2.230769230769231			
episode: 13	score: 0.0	memory length: 2956	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 2.0714285714285716			
episode: 14	score: 2.0	memory length: 3154	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 2.066666666666667			
episode: 15	score: 2.0	memory length: 3353	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 2.0625			
episode: 16	score: 4.0	memory length: 3666	epsilon: 1.0	steps: 313
lr: 0.0001	evaluation reward: 2.176470588235294			
episode: 17	score: 1.0	memory length: 3818	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 2.111111111111111			
episode: 18	score: 2.0	memory length: 4019	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 2.1052631578947367			
episode: 19	score: 1.0	memory length: 4189	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 2.05			
episode: 20	score: 0.0	memory length: 4313	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.9523809523809523			
episode: 21	score: 0.0	memory length: 4436	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.8636363636363635			
episode: 22	score: 1.0	memory length: 4588	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.826086956521739			
episode: 23	score: 4.0	memory length: 4885	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.916666666666667			
episode: 24	score: 0.0	memory length: 5009	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.84			
episode: 25	score: 3.0	memory length: 5240	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.8846153846153846			
episode: 26	score: 2.0	memory length: 5423	epsilon: 1.0	steps: 183
lr: 0.0001	evaluation reward: 1.8888888888888888			
episode: 27	score: 1.0	memory length: 5593	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.8571428571428572			
episode: 28	score: 1.0	memory length: 5765	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.8275862068965518			
episode: 29	score: 2.0	memory length: 5984	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.8333333333333333			
episode: 30	score: 2.0	memory length: 6166	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.8387096774193548			
episode: 31	score: 1.0	memory length: 6335	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.8125			
episode: 32	score: 3.0	memory length: 6563	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.8484848484848484			
episode: 33	score: 1.0	memory length: 6714	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.8235294117647058			
episode: 34	score: 3.0	memory length: 6944	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.8571428571428572			

episode: 35	score: 2.0	memory length: 7163	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.8611111111111112			
episode: 36	score: 2.0	memory length: 7362	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.864864864864865			
episode: 37	score: 0.0	memory length: 7486	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.8157894736842106			
episode: 38	score: 0.0	memory length: 7610	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.7692307692307692			
episode: 39	score: 2.0	memory length: 7809	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.775			
episode: 40	score: 1.0	memory length: 7961	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.7560975609756098			
episode: 41	score: 1.0	memory length: 8134	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.7380952380952381			
episode: 42	score: 3.0	memory length: 8383	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.7674418604651163			
episode: 43	score: 2.0	memory length: 8585	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.7727272727272727			
episode: 44	score: 1.0	memory length: 8757	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.7555555555555555			
episode: 45	score: 3.0	memory length: 9007	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.7826086956521738			
episode: 46	score: 0.0	memory length: 9130	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7446808510638299			
episode: 47	score: 0.0	memory length: 9253	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7083333333333333			
episode: 48	score: 1.0	memory length: 9422	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.6938775510204083			
episode: 49	score: 0.0	memory length: 9546	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.66			
episode: 50	score: 0.0	memory length: 9669	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6274509803921569			
episode: 51	score: 0.0	memory length: 9792	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5961538461538463			
episode: 52	score: 2.0	memory length: 9995	epsilon: 1.0	steps: 203
lr: 0.0001	evaluation reward: 1.6037735849056605			
episode: 53	score: 0.0	memory length: 10118	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5740740740740742			
episode: 54	score: 0.0	memory length: 10242	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5454545454545454			
episode: 55	score: 2.0	memory length: 10443	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.5535714285714286			
episode: 56	score: 1.0	memory length: 10613	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.543859649122807			
episode: 57	score: 2.0	memory length: 10830	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.5517241379310345			
episode: 58	score: 0.0	memory length: 10954	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5254237288135593			

episode: 59	score: 2.0	memory length: 11153	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.5333333333333334			
episode: 60	score: 2.0	memory length: 11353	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.540983606557377			
episode: 61	score: 1.0	memory length: 11523	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.532258064516129			
episode: 62	score: 1.0	memory length: 11693	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.5238095238095237			
episode: 63	score: 1.0	memory length: 11863	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.515625			
episode: 64	score: 0.0	memory length: 11986	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4923076923076923			
episode: 65	score: 1.0	memory length: 12157	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.4848484848484849			
episode: 66	score: 3.0	memory length: 12405	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.507462686567164			
episode: 67	score: 4.0	memory length: 12694	epsilon: 1.0	steps: 289
lr: 0.0001	evaluation reward: 1.5441176470588236			
episode: 68	score: 1.0	memory length: 12864	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.536231884057971			
episode: 69	score: 1.0	memory length: 13015	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.5285714285714285			
episode: 70	score: 0.0	memory length: 13138	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5070422535211268			
episode: 71	score: 2.0	memory length: 13357	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.5138888888888888			
episode: 72	score: 0.0	memory length: 13481	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.4931506849315068			
episode: 73	score: 1.0	memory length: 13652	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.4864864864864864			
episode: 74	score: 6.0	memory length: 14034	epsilon: 1.0	steps: 382
lr: 0.0001	evaluation reward: 1.5466666666666666			
episode: 75	score: 1.0	memory length: 14206	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.5394736842105263			
episode: 76	score: 2.0	memory length: 14422	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.5454545454545454			
episode: 77	score: 3.0	memory length: 14689	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.564102564102564			
episode: 78	score: 2.0	memory length: 14891	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.5696202531645569			
episode: 79	score: 1.0	memory length: 15043	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.5625			
episode: 80	score: 1.0	memory length: 15214	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.5555555555555556			
episode: 81	score: 2.0	memory length: 15412	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5609756097560976			
episode: 82	score: 2.0	memory length: 15630	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.5662650602409638			

episode: 83	score: 1.0	memory length: 15799	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.5595238095238095			
episode: 84	score: 1.0	memory length: 15971	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.5529411764705883			
episode: 85	score: 2.0	memory length: 16193	epsilon: 1.0	steps: 222
lr: 0.0001	evaluation reward: 1.558139534883721			
episode: 86	score: 0.0	memory length: 16317	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5402298850574712			
episode: 87	score: 1.0	memory length: 16487	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.5340909090909092			
episode: 88	score: 2.0	memory length: 16707	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.5393258426966292			
episode: 89	score: 0.0	memory length: 16831	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5222222222222221			
episode: 90	score: 1.0	memory length: 17001	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.5164835164835164			
episode: 91	score: 2.0	memory length: 17200	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.5217391304347827			
episode: 92	score: 3.0	memory length: 17447	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.5376344086021505			
episode: 93	score: 2.0	memory length: 17648	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.5425531914893618			
episode: 94	score: 2.0	memory length: 17847	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.5473684210526315			
episode: 95	score: 0.0	memory length: 17971	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.53125			
episode: 96	score: 3.0	memory length: 18202	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.5463917525773196			
episode: 97	score: 1.0	memory length: 18374	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.5408163265306123			
episode: 98	score: 0.0	memory length: 18497	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5252525252525253			
episode: 99	score: 5.0	memory length: 18862	epsilon: 1.0	steps: 365
lr: 0.0001	evaluation reward: 1.56			
episode: 100	score: 0.0	memory length: 18986	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.56			
episode: 101	score: 1.0	memory length: 19156	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.56			
episode: 102	score: 1.0	memory length: 19325	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.57			
episode: 103	score: 1.0	memory length: 19498	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.54			
episode: 104	score: 2.0	memory length: 19718	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.51			
episode: 105	score: 4.0	memory length: 20036	epsilon: 1.0	steps: 318
lr: 0.0001	evaluation reward: 1.52			
episode: 106	score: 2.0	memory length: 20255	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.5			

episode: 107	score: 2.0	memory length: 20477	epsilon: 1.0	steps: 222
lr: 0.0001	evaluation reward: 1.51			
episode: 108	score: 9.0	memory length: 20821	epsilon: 1.0	steps: 344
lr: 0.0001	evaluation reward: 1.57			
episode: 109	score: 3.0	memory length: 21035	epsilon: 1.0	steps: 214
lr: 0.0001	evaluation reward: 1.59			
episode: 110	score: 1.0	memory length: 21208	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.56			
episode: 111	score: 1.0	memory length: 21381	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.55			
episode: 112	score: 2.0	memory length: 21565	epsilon: 1.0	steps: 184
lr: 0.0001	evaluation reward: 1.56			
episode: 113	score: 1.0	memory length: 21717	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.57			
episode: 114	score: 0.0	memory length: 21841	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.55			
episode: 115	score: 0.0	memory length: 21965	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.53			
episode: 116	score: 2.0	memory length: 22163	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.51			
episode: 117	score: 2.0	memory length: 22364	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.52			
episode: 118	score: 3.0	memory length: 22590	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.53			
episode: 119	score: 0.0	memory length: 22713	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52			
episode: 120	score: 2.0	memory length: 22912	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.54			
episode: 121	score: 0.0	memory length: 23036	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54			
episode: 122	score: 0.0	memory length: 23160	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.53			
episode: 123	score: 2.0	memory length: 23359	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.51			
episode: 124	score: 0.0	memory length: 23483	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.51			
episode: 125	score: 2.0	memory length: 23704	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.5			
episode: 126	score: 2.0	memory length: 23902	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5			
episode: 127	score: 4.0	memory length: 24180	epsilon: 1.0	steps: 278
lr: 0.0001	evaluation reward: 1.53			
episode: 128	score: 0.0	memory length: 24304	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.52			
episode: 129	score: 2.0	memory length: 24503	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.52			
episode: 130	score: 1.0	memory length: 24655	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.51			

episode: 131	score: 1.0	memory length: 24807	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.51			
episode: 132	score: 2.0	memory length: 25005	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5			
episode: 133	score: 3.0	memory length: 25273	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.52			
episode: 134	score: 2.0	memory length: 25491	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.51			
episode: 135	score: 0.0	memory length: 25614	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49			
episode: 136	score: 0.0	memory length: 25738	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.47			
episode: 137	score: 0.0	memory length: 25862	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.47			
episode: 138	score: 3.0	memory length: 26109	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.5			
episode: 139	score: 2.0	memory length: 26308	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.5			
episode: 140	score: 2.0	memory length: 26507	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.51			
episode: 141	score: 0.0	memory length: 26630	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5			
episode: 142	score: 2.0	memory length: 26848	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.49			
episode: 143	score: 3.0	memory length: 27097	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.5			
episode: 144	score: 0.0	memory length: 27220	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49			
episode: 145	score: 0.0	memory length: 27343	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46			
episode: 146	score: 2.0	memory length: 27525	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.48			
episode: 147	score: 3.0	memory length: 27771	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.51			
episode: 148	score: 0.0	memory length: 27895	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5			
episode: 149	score: 2.0	memory length: 28096	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.52			
episode: 150	score: 0.0	memory length: 28219	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52			
episode: 151	score: 0.0	memory length: 28342	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52			
episode: 152	score: 0.0	memory length: 28466	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5			
episode: 153	score: 5.0	memory length: 28791	epsilon: 1.0	steps: 325
lr: 0.0001	evaluation reward: 1.55			
episode: 154	score: 2.0	memory length: 28990	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.57			

episode: 155	score: 1.0	memory length: 29159	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.56			
episode: 156	score: 1.0	memory length: 29329	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.56			
episode: 157	score: 0.0	memory length: 29453	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54			
episode: 158	score: 2.0	memory length: 29651	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.56			
episode: 159	score: 1.0	memory length: 29802	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.55			
episode: 160	score: 2.0	memory length: 30001	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.55			
episode: 161	score: 1.0	memory length: 30170	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.55			
episode: 162	score: 3.0	memory length: 30400	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.57			
episode: 163	score: 3.0	memory length: 30652	epsilon: 1.0	steps: 252
lr: 0.0001	evaluation reward: 1.59			
episode: 164	score: 4.0	memory length: 30928	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.63			
episode: 165	score: 2.0	memory length: 31127	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.64			
episode: 166	score: 2.0	memory length: 31345	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.63			
episode: 167	score: 2.0	memory length: 31562	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.61			
episode: 168	score: 1.0	memory length: 31714	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.61			
episode: 169	score: 1.0	memory length: 31885	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.61			
episode: 170	score: 2.0	memory length: 32083	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.63			
episode: 171	score: 0.0	memory length: 32207	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.61			
episode: 172	score: 3.0	memory length: 32451	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.64			
episode: 173	score: 0.0	memory length: 32575	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.63			
episode: 174	score: 3.0	memory length: 32804	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.6			
episode: 175	score: 2.0	memory length: 33022	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.61			
episode: 176	score: 5.0	memory length: 33362	epsilon: 1.0	steps: 340
lr: 0.0001	evaluation reward: 1.64			
episode: 177	score: 5.0	memory length: 33684	epsilon: 1.0	steps: 322
lr: 0.0001	evaluation reward: 1.66			
episode: 178	score: 2.0	memory length: 33883	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.66			

episode: 179	score: 3.0	memory length: 34114	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.68			
episode: 180	score: 0.0	memory length: 34238	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.67			
episode: 181	score: 1.0	memory length: 34408	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.66			
episode: 182	score: 0.0	memory length: 34532	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.64			
episode: 183	score: 1.0	memory length: 34703	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.64			
episode: 184	score: 2.0	memory length: 34901	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.65			
episode: 185	score: 1.0	memory length: 35053	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.64			
episode: 186	score: 3.0	memory length: 35297	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.67			
episode: 187	score: 3.0	memory length: 35527	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.69			
episode: 188	score: 0.0	memory length: 35651	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.67			
episode: 189	score: 1.0	memory length: 35823	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.68			
episode: 190	score: 1.0	memory length: 35992	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.68			
episode: 191	score: 1.0	memory length: 36162	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.67			
episode: 192	score: 0.0	memory length: 36285	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64			
episode: 193	score: 3.0	memory length: 36530	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.65			
episode: 194	score: 2.0	memory length: 36749	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.65			
episode: 195	score: 1.0	memory length: 36921	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.66			
episode: 196	score: 4.0	memory length: 37234	epsilon: 1.0	steps: 313
lr: 0.0001	evaluation reward: 1.67			
episode: 197	score: 1.0	memory length: 37386	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.67			
episode: 198	score: 3.0	memory length: 37634	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.7			
episode: 199	score: 0.0	memory length: 37758	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.65			
episode: 200	score: 1.0	memory length: 37929	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.66			
episode: 201	score: 2.0	memory length: 38145	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.67			
episode: 202	score: 0.0	memory length: 38268	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.66			

episode: 203	score: 3.0	memory length: 38500	epsilon: 1.0	steps: 232
lr: 0.0001	evaluation reward: 1.68			
episode: 204	score: 2.0	memory length: 38682	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.68			
episode: 205	score: 1.0	memory length: 38851	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.65			
episode: 206	score: 0.0	memory length: 38975	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.63			
episode: 207	score: 0.0	memory length: 39099	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.61			
episode: 208	score: 2.0	memory length: 39318	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.54			
episode: 209	score: 1.0	memory length: 39488	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.52			
episode: 210	score: 2.0	memory length: 39686	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.53			
episode: 211	score: 1.0	memory length: 39855	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.53			
episode: 212	score: 0.0	memory length: 39978	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51			
episode: 213	score: 4.0	memory length: 40275	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.54			
episode: 214	score: 5.0	memory length: 40593	epsilon: 1.0	steps: 318
lr: 0.0001	evaluation reward: 1.59			
episode: 215	score: 5.0	memory length: 40924	epsilon: 1.0	steps: 331
lr: 0.0001	evaluation reward: 1.64			
episode: 216	score: 2.0	memory length: 41143	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.64			
episode: 217	score: 1.0	memory length: 41313	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.63			
episode: 218	score: 1.0	memory length: 41486	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.61			
episode: 219	score: 0.0	memory length: 41610	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.61			
episode: 220	score: 0.0	memory length: 41733	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59			
episode: 221	score: 0.0	memory length: 41856	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59			
episode: 222	score: 0.0	memory length: 41980	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.59			
episode: 223	score: 4.0	memory length: 42298	epsilon: 1.0	steps: 318
lr: 0.0001	evaluation reward: 1.61			
episode: 224	score: 2.0	memory length: 42518	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.63			
episode: 225	score: 1.0	memory length: 42670	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.62			
episode: 226	score: 0.0	memory length: 42794	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.6			

episode: 227	score: 3.0	memory length: 43025	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.59			
episode: 228	score: 2.0	memory length: 43224	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.61			
episode: 229	score: 0.0	memory length: 43348	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.59			
episode: 230	score: 5.0	memory length: 43692	epsilon: 1.0	steps: 344
lr: 0.0001	evaluation reward: 1.63			
episode: 231	score: 1.0	memory length: 43864	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.63			
episode: 232	score: 0.0	memory length: 43988	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.61			
episode: 233	score: 0.0	memory length: 44111	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.58			
episode: 234	score: 4.0	memory length: 44424	epsilon: 1.0	steps: 313
lr: 0.0001	evaluation reward: 1.6			
episode: 235	score: 1.0	memory length: 44596	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.61			
episode: 236	score: 3.0	memory length: 44846	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.64			
episode: 237	score: 2.0	memory length: 45062	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.66			
episode: 238	score: 0.0	memory length: 45186	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.63			
episode: 239	score: 3.0	memory length: 45456	epsilon: 1.0	steps: 270
lr: 0.0001	evaluation reward: 1.64			
episode: 240	score: 0.0	memory length: 45579	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62			
episode: 241	score: 1.0	memory length: 45749	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.63			
episode: 242	score: 0.0	memory length: 45873	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.61			
episode: 243	score: 1.0	memory length: 46025	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.59			
episode: 244	score: 2.0	memory length: 46206	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.61			
episode: 245	score: 0.0	memory length: 46329	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61			
episode: 246	score: 3.0	memory length: 46576	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.62			
episode: 247	score: 1.0	memory length: 46747	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.6			
episode: 248	score: 2.0	memory length: 46929	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.62			
episode: 249	score: 2.0	memory length: 47128	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.62			
episode: 250	score: 1.0	memory length: 47301	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.63			

episode: 251	score: 3.0	memory length: 47549	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.66			
episode: 252	score: 1.0	memory length: 47700	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.67			
episode: 253	score: 4.0	memory length: 47968	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.66			
episode: 254	score: 0.0	memory length: 48091	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64			
episode: 255	score: 1.0	memory length: 48242	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.64			
episode: 256	score: 2.0	memory length: 48440	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.65			
episode: 257	score: 1.0	memory length: 48609	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.66			
episode: 258	score: 3.0	memory length: 48856	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.67			
episode: 259	score: 1.0	memory length: 49025	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.67			
episode: 260	score: 0.0	memory length: 49148	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65			
episode: 261	score: 2.0	memory length: 49348	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.66			
episode: 262	score: 1.0	memory length: 49518	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.64			
episode: 263	score: 2.0	memory length: 49716	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.63			
episode: 264	score: 2.0	memory length: 49934	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.61			
episode: 265	score: 1.0	memory length: 50086	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.6			
episode: 266	score: 3.0	memory length: 50331	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.61			
episode: 267	score: 7.0	memory length: 50736	epsilon: 1.0	steps: 405
lr: 0.0001	evaluation reward: 1.66			
episode: 268	score: 2.0	memory length: 50937	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.67			
episode: 269	score: 3.0	memory length: 51185	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.69			
episode: 270	score: 4.0	memory length: 51500	epsilon: 1.0	steps: 315
lr: 0.0001	evaluation reward: 1.71			
episode: 271	score: 2.0	memory length: 51701	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.73			
episode: 272	score: 0.0	memory length: 51824	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7			
episode: 273	score: 0.0	memory length: 51947	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7			
episode: 274	score: 0.0	memory length: 52070	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67			

episode: 275	score: 1.0	memory length: 52221	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.66			
episode: 276	score: 1.0	memory length: 52390	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.62			
episode: 277	score: 1.0	memory length: 52542	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.58			
episode: 278	score: 1.0	memory length: 52694	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.57			
episode: 279	score: 0.0	memory length: 52818	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54			
episode: 280	score: 0.0	memory length: 52942	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54			
episode: 281	score: 0.0	memory length: 53066	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.53			
episode: 282	score: 1.0	memory length: 53237	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.54			
episode: 283	score: 3.0	memory length: 53485	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.56			
episode: 284	score: 2.0	memory length: 53705	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.56			
episode: 285	score: 0.0	memory length: 53829	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.55			
episode: 286	score: 0.0	memory length: 53953	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.52			
episode: 287	score: 1.0	memory length: 54105	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.5			
episode: 288	score: 2.0	memory length: 54288	epsilon: 1.0	steps: 183
lr: 0.0001	evaluation reward: 1.52			
episode: 289	score: 3.0	memory length: 54534	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.54			
episode: 290	score: 4.0	memory length: 54832	epsilon: 1.0	steps: 298
lr: 0.0001	evaluation reward: 1.57			
episode: 291	score: 0.0	memory length: 54956	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.56			
episode: 292	score: 4.0	memory length: 55254	epsilon: 1.0	steps: 298
lr: 0.0001	evaluation reward: 1.6			
episode: 293	score: 3.0	memory length: 55500	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.6			
episode: 294	score: 2.0	memory length: 55698	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6			
episode: 295	score: 2.0	memory length: 55918	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.61			
episode: 296	score: 0.0	memory length: 56041	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57			
episode: 297	score: 1.0	memory length: 56211	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.57			
episode: 298	score: 1.0	memory length: 56383	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.55			

episode: 299	score: 0.0	memory length: 56507	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.55			
episode: 300	score: 2.0	memory length: 56705	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.56			
episode: 301	score: 2.0	memory length: 56906	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.56			
episode: 302	score: 1.0	memory length: 57078	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.57			
episode: 303	score: 1.0	memory length: 57247	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.55			
episode: 304	score: 3.0	memory length: 57476	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.56			
episode: 305	score: 0.0	memory length: 57600	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.55			
episode: 306	score: 1.0	memory length: 57770	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.56			
episode: 307	score: 0.0	memory length: 57894	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.56			
episode: 308	score: 0.0	memory length: 58018	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54			
episode: 309	score: 2.0	memory length: 58204	epsilon: 1.0	steps: 186
lr: 0.0001	evaluation reward: 1.55			
episode: 310	score: 3.0	memory length: 58452	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.56			
episode: 311	score: 2.0	memory length: 58635	epsilon: 1.0	steps: 183
lr: 0.0001	evaluation reward: 1.57			
episode: 312	score: 2.0	memory length: 58834	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.59			
episode: 313	score: 3.0	memory length: 59082	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.58			
episode: 314	score: 0.0	memory length: 59206	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.53			
episode: 315	score: 1.0	memory length: 59376	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.49			
episode: 316	score: 1.0	memory length: 59545	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.48			
episode: 317	score: 0.0	memory length: 59669	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.47			
episode: 318	score: 0.0	memory length: 59793	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.46			
episode: 319	score: 1.0	memory length: 59966	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.47			
episode: 320	score: 0.0	memory length: 60089	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 321	score: 3.0	memory length: 60340	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.5			
episode: 322	score: 1.0	memory length: 60509	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.51			

episode: 323	score: 0.0	memory length: 60633	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.47			
episode: 324	score: 0.0	memory length: 60757	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.45			
episode: 325	score: 2.0	memory length: 60979	epsilon: 1.0	steps: 222
lr: 0.0001	evaluation reward: 1.46			
episode: 326	score: 1.0	memory length: 61132	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.47			
episode: 327	score: 3.0	memory length: 61398	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.47			
episode: 328	score: 2.0	memory length: 61596	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47			
episode: 329	score: 1.0	memory length: 61749	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.48			
episode: 330	score: 3.0	memory length: 62016	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.46			
episode: 331	score: 0.0	memory length: 62140	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.45			
episode: 332	score: 0.0	memory length: 62263	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45			
episode: 333	score: 2.0	memory length: 62461	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47			
episode: 334	score: 2.0	memory length: 62680	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.45			
episode: 335	score: 0.0	memory length: 62803	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 336	score: 0.0	memory length: 62926	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41			
episode: 337	score: 0.0	memory length: 63050	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.39			
episode: 338	score: 0.0	memory length: 63173	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39			
episode: 339	score: 5.0	memory length: 63483	epsilon: 1.0	steps: 310
lr: 0.0001	evaluation reward: 1.41			
episode: 340	score: 0.0	memory length: 63607	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.41			
episode: 341	score: 2.0	memory length: 63824	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.42			
episode: 342	score: 1.0	memory length: 63975	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.43			
episode: 343	score: 2.0	memory length: 64174	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.44			
episode: 344	score: 2.0	memory length: 64375	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.44			
episode: 345	score: 0.0	memory length: 64499	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.44			
episode: 346	score: 3.0	memory length: 64765	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.44			

episode: 347	score: 1.0	memory length: 64935	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.44			
episode: 348	score: 3.0	memory length: 65187	epsilon: 1.0	steps: 252
lr: 0.0001	evaluation reward: 1.45			
episode: 349	score: 2.0	memory length: 65385	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45			
episode: 350	score: 3.0	memory length: 65611	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.47			
episode: 351	score: 2.0	memory length: 65827	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.46			
episode: 352	score: 1.0	memory length: 65979	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.46			
episode: 353	score: 0.0	memory length: 66103	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.42			
episode: 354	score: 1.0	memory length: 66273	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.43			
episode: 355	score: 2.0	memory length: 66476	epsilon: 1.0	steps: 203
lr: 0.0001	evaluation reward: 1.44			
episode: 356	score: 2.0	memory length: 66697	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.44			
episode: 357	score: 1.0	memory length: 66866	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.44			
episode: 358	score: 1.0	memory length: 67037	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.42			
episode: 359	score: 2.0	memory length: 67257	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.43			
episode: 360	score: 0.0	memory length: 67381	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.43			
episode: 361	score: 2.0	memory length: 67582	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.43			
episode: 362	score: 0.0	memory length: 67706	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.42			
episode: 363	score: 5.0	memory length: 68036	epsilon: 1.0	steps: 330
lr: 0.0001	evaluation reward: 1.45			
episode: 364	score: 2.0	memory length: 68235	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.45			
episode: 365	score: 3.0	memory length: 68499	epsilon: 1.0	steps: 264
lr: 0.0001	evaluation reward: 1.47			
episode: 366	score: 1.0	memory length: 68651	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.45			
episode: 367	score: 6.0	memory length: 69020	epsilon: 1.0	steps: 369
lr: 0.0001	evaluation reward: 1.44			
episode: 368	score: 0.0	memory length: 69144	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.42			
episode: 369	score: 4.0	memory length: 69421	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.43			
episode: 370	score: 2.0	memory length: 69619	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.41			

episode: 371	score: 1.0	memory length: 69789	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.4			
episode: 372	score: 0.0	memory length: 69913	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.4			
episode: 373	score: 5.0	memory length: 70239	epsilon: 1.0	steps: 326
lr: 0.0001	evaluation reward: 1.45			
episode: 374	score: 2.0	memory length: 70443	epsilon: 1.0	steps: 204
lr: 0.0001	evaluation reward: 1.47			
episode: 375	score: 2.0	memory length: 70660	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.48			
episode: 376	score: 0.0	memory length: 70783	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 377	score: 0.0	memory length: 70906	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46			
episode: 378	score: 1.0	memory length: 71075	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.46			
episode: 379	score: 1.0	memory length: 71245	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.47			
episode: 380	score: 2.0	memory length: 71465	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.49			
episode: 381	score: 3.0	memory length: 71710	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.52			
episode: 382	score: 0.0	memory length: 71833	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51			
episode: 383	score: 3.0	memory length: 72100	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.51			
episode: 384	score: 4.0	memory length: 72415	epsilon: 1.0	steps: 315
lr: 0.0001	evaluation reward: 1.53			
episode: 385	score: 1.0	memory length: 72567	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.54			
episode: 386	score: 0.0	memory length: 72690	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.54			
episode: 387	score: 2.0	memory length: 72910	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.55			
episode: 388	score: 2.0	memory length: 73133	epsilon: 1.0	steps: 223
lr: 0.0001	evaluation reward: 1.55			
episode: 389	score: 3.0	memory length: 73344	epsilon: 1.0	steps: 211
lr: 0.0001	evaluation reward: 1.55			
episode: 390	score: 3.0	memory length: 73571	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.54			
episode: 391	score: 1.0	memory length: 73741	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.55			
episode: 392	score: 2.0	memory length: 73940	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.53			
episode: 393	score: 3.0	memory length: 74188	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.53			
episode: 394	score: 2.0	memory length: 74386	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.53			

episode: 395	score: 2.0	memory length: 74586	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.53			
episode: 396	score: 1.0	memory length: 74757	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.54			
episode: 397	score: 4.0	memory length: 75059	epsilon: 1.0	steps: 302
lr: 0.0001	evaluation reward: 1.57			
episode: 398	score: 1.0	memory length: 75230	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.57			
episode: 399	score: 3.0	memory length: 75478	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.6			
episode: 400	score: 0.0	memory length: 75602	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.58			
episode: 401	score: 5.0	memory length: 75919	epsilon: 1.0	steps: 317
lr: 0.0001	evaluation reward: 1.61			
episode: 402	score: 1.0	memory length: 76089	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.61			
episode: 403	score: 0.0	memory length: 76212	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6			
episode: 404	score: 3.0	memory length: 76460	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.6			
episode: 405	score: 2.0	memory length: 76679	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.62			
episode: 406	score: 0.0	memory length: 76803	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.61			
episode: 407	score: 1.0	memory length: 76955	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.62			
episode: 408	score: 1.0	memory length: 77125	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.63			
episode: 409	score: 2.0	memory length: 77342	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.63			
episode: 410	score: 1.0	memory length: 77511	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.61			
episode: 411	score: 0.0	memory length: 77634	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59			
episode: 412	score: 1.0	memory length: 77805	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.58			
episode: 413	score: 0.0	memory length: 77929	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.55			
episode: 414	score: 1.0	memory length: 78100	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.56			
episode: 415	score: 4.0	memory length: 78400	epsilon: 1.0	steps: 300
lr: 0.0001	evaluation reward: 1.59			
episode: 416	score: 2.0	memory length: 78599	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.6			
episode: 417	score: 0.0	memory length: 78722	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6			
episode: 418	score: 4.0	memory length: 79018	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.64			

episode: 419	score: 1.0	memory length: 79188	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.64			
episode: 420	score: 4.0	memory length: 79466	epsilon: 1.0	steps: 278
lr: 0.0001	evaluation reward: 1.68			
episode: 421	score: 1.0	memory length: 79618	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.66			
episode: 422	score: 0.0	memory length: 79742	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.65			
episode: 423	score: 1.0	memory length: 79893	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.66			
episode: 424	score: 3.0	memory length: 80141	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.69			
episode: 425	score: 2.0	memory length: 80340	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.69			
episode: 426	score: 7.0	memory length: 80780	epsilon: 1.0	steps: 440
lr: 0.0001	evaluation reward: 1.75			
episode: 427	score: 2.0	memory length: 80978	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.74			
episode: 428	score: 2.0	memory length: 81176	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.74			
episode: 429	score: 6.0	memory length: 81551	epsilon: 1.0	steps: 375
lr: 0.0001	evaluation reward: 1.79			
episode: 430	score: 3.0	memory length: 81799	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.79			
episode: 431	score: 1.0	memory length: 81969	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.8			
episode: 432	score: 1.0	memory length: 82138	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.81			
episode: 433	score: 4.0	memory length: 82414	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.83			
episode: 434	score: 2.0	memory length: 82613	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.83			
episode: 435	score: 0.0	memory length: 82737	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.83			
episode: 436	score: 0.0	memory length: 82861	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.83			
episode: 437	score: 0.0	memory length: 82985	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.83			
episode: 438	score: 1.0	memory length: 83154	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.84			
episode: 439	score: 3.0	memory length: 83402	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.82			
episode: 440	score: 1.0	memory length: 83572	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.83			
episode: 441	score: 3.0	memory length: 83798	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.84			
episode: 442	score: 3.0	memory length: 84051	epsilon: 1.0	steps: 253
lr: 0.0001	evaluation reward: 1.86			

episode: 443	score: 4.0	memory length: 84349	epsilon: 1.0	steps: 298
lr: 0.0001	evaluation reward: 1.88			
episode: 444	score: 0.0	memory length: 84473	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.86			
episode: 445	score: 1.0	memory length: 84625	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.87			
episode: 446	score: 0.0	memory length: 84748	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.84			
episode: 447	score: 2.0	memory length: 84947	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.85			
episode: 448	score: 1.0	memory length: 85101	epsilon: 1.0	steps: 154
lr: 0.0001	evaluation reward: 1.83			
episode: 449	score: 0.0	memory length: 85225	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.81			
episode: 450	score: 0.0	memory length: 85349	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.78			
episode: 451	score: 0.0	memory length: 85472	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.76			
episode: 452	score: 2.0	memory length: 85670	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.77			
episode: 453	score: 0.0	memory length: 85794	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.77			
episode: 454	score: 1.0	memory length: 85964	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.77			
episode: 455	score: 1.0	memory length: 86117	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.76			
episode: 456	score: 0.0	memory length: 86241	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.74			
episode: 457	score: 1.0	memory length: 86412	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.74			
episode: 458	score: 0.0	memory length: 86536	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.73			
episode: 459	score: 1.0	memory length: 86705	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.72			
episode: 460	score: 3.0	memory length: 86952	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.75			
episode: 461	score: 3.0	memory length: 87196	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.76			
episode: 462	score: 1.0	memory length: 87365	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.77			
episode: 463	score: 0.0	memory length: 87489	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.72			
episode: 464	score: 1.0	memory length: 87659	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.71			
episode: 465	score: 1.0	memory length: 87811	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.69			
episode: 466	score: 1.0	memory length: 87981	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.69			

episode: 467	score: 4.0	memory length: 88258	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.67			
episode: 468	score: 2.0	memory length: 88477	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.69			
episode: 469	score: 0.0	memory length: 88600	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65			
episode: 470	score: 1.0	memory length: 88751	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.64			
episode: 471	score: 1.0	memory length: 88922	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.64			
episode: 472	score: 2.0	memory length: 89106	epsilon: 1.0	steps: 184
lr: 0.0001	evaluation reward: 1.66			
episode: 473	score: 0.0	memory length: 89230	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.61			
episode: 474	score: 0.0	memory length: 89353	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59			
episode: 475	score: 2.0	memory length: 89552	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.59			
episode: 476	score: 2.0	memory length: 89732	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.61			
episode: 477	score: 2.0	memory length: 89931	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.63			
episode: 478	score: 0.0	memory length: 90055	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.62			
episode: 479	score: 2.0	memory length: 90272	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.63			
episode: 480	score: 1.0	memory length: 90442	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.62			
episode: 481	score: 1.0	memory length: 90612	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.6			
episode: 482	score: 4.0	memory length: 90910	epsilon: 1.0	steps: 298
lr: 0.0001	evaluation reward: 1.64			
episode: 483	score: 4.0	memory length: 91205	epsilon: 1.0	steps: 295
lr: 0.0001	evaluation reward: 1.65			
episode: 484	score: 1.0	memory length: 91375	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.62			
episode: 485	score: 3.0	memory length: 91606	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.64			
episode: 486	score: 3.0	memory length: 91874	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.67			
episode: 487	score: 3.0	memory length: 92125	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.68			
episode: 488	score: 1.0	memory length: 92277	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.67			
episode: 489	score: 0.0	memory length: 92400	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64			
episode: 490	score: 0.0	memory length: 92523	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61			

episode: 491	score: 4.0	memory length: 92820	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.64			
episode: 492	score: 1.0	memory length: 92990	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.63			
episode: 493	score: 2.0	memory length: 93188	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.62			
episode: 494	score: 4.0	memory length: 93463	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.64			
episode: 495	score: 1.0	memory length: 93633	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.63			
episode: 496	score: 1.0	memory length: 93803	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.63			
episode: 497	score: 1.0	memory length: 93975	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.6			
episode: 498	score: 2.0	memory length: 94174	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.61			
episode: 499	score: 1.0	memory length: 94344	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.59			
episode: 500	score: 0.0	memory length: 94467	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59			
episode: 501	score: 0.0	memory length: 94591	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54			
episode: 502	score: 1.0	memory length: 94742	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.54			
episode: 503	score: 2.0	memory length: 94940	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.56			
episode: 504	score: 4.0	memory length: 95238	epsilon: 1.0	steps: 298
lr: 0.0001	evaluation reward: 1.57			
episode: 505	score: 2.0	memory length: 95436	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.57			
episode: 506	score: 0.0	memory length: 95559	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57			
episode: 507	score: 0.0	memory length: 95683	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.56			
episode: 508	score: 0.0	memory length: 95807	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.55			
episode: 509	score: 2.0	memory length: 96005	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.55			
episode: 510	score: 3.0	memory length: 96250	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.57			
episode: 511	score: 1.0	memory length: 96401	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.58			
episode: 512	score: 0.0	memory length: 96525	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.57			
episode: 513	score: 1.0	memory length: 96694	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.58			
episode: 514	score: 1.0	memory length: 96865	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.58			

episode: 515	score: 1.0	memory length: 97035	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.55			
episode: 516	score: 0.0	memory length: 97158	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53			
episode: 517	score: 3.0	memory length: 97426	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.56			
episode: 518	score: 2.0	memory length: 97645	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.54			
episode: 519	score: 3.0	memory length: 97894	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.56			
episode: 520	score: 3.0	memory length: 98144	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.55			
episode: 521	score: 2.0	memory length: 98343	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.56			
episode: 522	score: 2.0	memory length: 98561	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.58			
episode: 523	score: 5.0	memory length: 98908	epsilon: 1.0	steps: 347
lr: 0.0001	evaluation reward: 1.62			
episode: 524	score: 3.0	memory length: 99174	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.62			
episode: 525	score: 2.0	memory length: 99375	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.62			
episode: 526	score: 2.0	memory length: 99576	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.57			
episode: 527	score: 1.0	memory length: 99747	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.56			
episode: 528	score: 2.0	memory length: 99946	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.56			
episode: 529	score: 8.0	memory length: 100406	epsilon: 0.9991941400000175	
steps: 460	lr: 0.0001	evaluation reward: 1.58		
episode: 530	score: 4.0	memory length: 100664	epsilon: 0.9986833000000286	
steps: 258	lr: 0.0001	evaluation reward: 1.59		
episode: 531	score: 0.0	memory length: 100788	epsilon: 0.9984377800000339	
steps: 124	lr: 0.0001	evaluation reward: 1.58		
episode: 532	score: 0.0	memory length: 100912	epsilon: 0.9981922600000392	
steps: 124	lr: 0.0001	evaluation reward: 1.57		
episode: 533	score: 1.0	memory length: 101082	epsilon: 0.9978556600000466	
steps: 170	lr: 0.0001	evaluation reward: 1.54		
episode: 534	score: 2.0	memory length: 101281	epsilon: 0.9974616400000551	
steps: 199	lr: 0.0001	evaluation reward: 1.54		
episode: 535	score: 4.0	memory length: 101599	epsilon: 0.9968320000000688	
steps: 318	lr: 0.0001	evaluation reward: 1.58		
episode: 536	score: 5.0	memory length: 101945	epsilon: 0.9961469200000836	
steps: 346	lr: 0.0001	evaluation reward: 1.63		
episode: 537	score: 1.0	memory length: 102114	epsilon: 0.9958123000000909	
steps: 169	lr: 0.0001	evaluation reward: 1.64		
episode: 538	score: 1.0	memory length: 102284	epsilon: 0.9954757000000982	
steps: 170	lr: 0.0001	evaluation reward: 1.64		

episode: 539	score: 1.0	memory length: 102436	epsilon: 0.9951747400001048
steps: 152	lr: 0.0001	evaluation reward: 1.62	
episode: 540	score: 0.0	memory length: 102560	epsilon: 0.9949292200001101
steps: 124	lr: 0.0001	evaluation reward: 1.61	
episode: 541	score: 2.0	memory length: 102759	epsilon: 0.9945352000001186
steps: 199	lr: 0.0001	evaluation reward: 1.6	
episode: 542	score: 2.0	memory length: 102940	epsilon: 0.9941768200001264
steps: 181	lr: 0.0001	evaluation reward: 1.59	
episode: 543	score: 0.0	memory length: 103064	epsilon: 0.9939313000001317
steps: 124	lr: 0.0001	evaluation reward: 1.55	
episode: 544	score: 2.0	memory length: 103283	epsilon: 0.9934976800001412
steps: 219	lr: 0.0001	evaluation reward: 1.57	
episode: 545	score: 0.0	memory length: 103407	epsilon: 0.9932521600001465
steps: 124	lr: 0.0001	evaluation reward: 1.56	
episode: 546	score: 0.0	memory length: 103530	epsilon: 0.9930086200001518
steps: 123	lr: 0.0001	evaluation reward: 1.56	
episode: 547	score: 2.0	memory length: 103748	epsilon: 0.9925769800001611
steps: 218	lr: 0.0001	evaluation reward: 1.56	
episode: 548	score: 1.0	memory length: 103900	epsilon: 0.9922760200001677
steps: 152	lr: 0.0001	evaluation reward: 1.56	
episode: 549	score: 1.0	memory length: 104072	epsilon: 0.9919354600001751
steps: 172	lr: 0.0001	evaluation reward: 1.57	
episode: 550	score: 0.0	memory length: 104196	epsilon: 0.9916899400001804
steps: 124	lr: 0.0001	evaluation reward: 1.57	
episode: 551	score: 0.0	memory length: 104319	epsilon: 0.9914464000001857
steps: 123	lr: 0.0001	evaluation reward: 1.57	
episode: 552	score: 3.0	memory length: 104551	epsilon: 0.9909870400001957
steps: 232	lr: 0.0001	evaluation reward: 1.58	
episode: 553	score: 0.0	memory length: 104675	epsilon: 0.990741520000201
steps: 124	lr: 0.0001	evaluation reward: 1.58	
episode: 554	score: 3.0	memory length: 104944	epsilon: 0.9902089000002126
steps: 269	lr: 0.0001	evaluation reward: 1.6	
episode: 555	score: 3.0	memory length: 105191	epsilon: 0.9897198400002232
steps: 247	lr: 0.0001	evaluation reward: 1.62	
episode: 556	score: 1.0	memory length: 105361	epsilon: 0.9893832400002305
steps: 170	lr: 0.0001	evaluation reward: 1.63	
episode: 557	score: 2.0	memory length: 105560	epsilon: 0.988989220000239
steps: 199	lr: 0.0001	evaluation reward: 1.64	
episode: 558	score: 3.0	memory length: 105808	epsilon: 0.9884981800002497
steps: 248	lr: 0.0001	evaluation reward: 1.67	
episode: 559	score: 3.0	memory length: 106038	epsilon: 0.9880427800002596
steps: 230	lr: 0.0001	evaluation reward: 1.69	
episode: 560	score: 1.0	memory length: 106211	epsilon: 0.987700240000267
steps: 173	lr: 0.0001	evaluation reward: 1.67	
episode: 561	score: 3.0	memory length: 106460	epsilon: 0.9872072200002777
steps: 249	lr: 0.0001	evaluation reward: 1.67	
episode: 562	score: 2.0	memory length: 106658	epsilon: 0.9868151800002862
steps: 198	lr: 0.0001	evaluation reward: 1.68	

episode: 563	score: 1.0	memory length: 106810	epsilon: 0.9865142200002928
steps: 152	lr: 0.0001	evaluation reward: 1.69	
episode: 564	score: 2.0	memory length: 107008	epsilon: 0.9861221800003013
steps: 198	lr: 0.0001	evaluation reward: 1.7	
episode: 565	score: 0.0	memory length: 107132	epsilon: 0.9858766600003066
steps: 124	lr: 0.0001	evaluation reward: 1.69	
episode: 566	score: 2.0	memory length: 107331	epsilon: 0.9854826400003152
steps: 199	lr: 0.0001	evaluation reward: 1.7	
episode: 567	score: 4.0	memory length: 107615	epsilon: 0.9849203200003274
steps: 284	lr: 0.0001	evaluation reward: 1.7	
episode: 568	score: 2.0	memory length: 107813	epsilon: 0.9845282800003359
steps: 198	lr: 0.0001	evaluation reward: 1.7	
episode: 569	score: 0.0	memory length: 107937	epsilon: 0.9842827600003412
steps: 124	lr: 0.0001	evaluation reward: 1.7	
episode: 570	score: 1.0	memory length: 108088	epsilon: 0.9839837800003477
steps: 151	lr: 0.0001	evaluation reward: 1.7	
episode: 571	score: 1.0	memory length: 108258	epsilon: 0.983647180000355
steps: 170	lr: 0.0001	evaluation reward: 1.7	
episode: 572	score: 0.0	memory length: 108381	epsilon: 0.9834036400003603
steps: 123	lr: 0.0001	evaluation reward: 1.68	
episode: 573	score: 1.0	memory length: 108533	epsilon: 0.9831026800003668
steps: 152	lr: 0.0001	evaluation reward: 1.69	
episode: 574	score: 0.0	memory length: 108657	epsilon: 0.9828571600003722
steps: 124	lr: 0.0001	evaluation reward: 1.69	
episode: 575	score: 0.0	memory length: 108780	epsilon: 0.9826136200003774
steps: 123	lr: 0.0001	evaluation reward: 1.67	
episode: 576	score: 0.0	memory length: 108904	epsilon: 0.9823681000003828
steps: 124	lr: 0.0001	evaluation reward: 1.65	
episode: 577	score: 2.0	memory length: 109103	epsilon: 0.9819740800003913
steps: 199	lr: 0.0001	evaluation reward: 1.65	
episode: 578	score: 3.0	memory length: 109332	epsilon: 0.9815206600004012
steps: 229	lr: 0.0001	evaluation reward: 1.68	
episode: 579	score: 0.0	memory length: 109456	epsilon: 0.9812751400004065
steps: 124	lr: 0.0001	evaluation reward: 1.66	
episode: 580	score: 2.0	memory length: 109657	epsilon: 0.9808771600004151
steps: 201	lr: 0.0001	evaluation reward: 1.67	
episode: 581	score: 1.0	memory length: 109830	epsilon: 0.9805346200004226
steps: 173	lr: 0.0001	evaluation reward: 1.67	
episode: 582	score: 2.0	memory length: 110049	epsilon: 0.980101000000432
steps: 219	lr: 0.0001	evaluation reward: 1.65	
episode: 583	score: 1.0	memory length: 110219	epsilon: 0.9797644000004393
steps: 170	lr: 0.0001	evaluation reward: 1.62	
episode: 584	score: 1.0	memory length: 110371	epsilon: 0.9794634400004458
steps: 152	lr: 0.0001	evaluation reward: 1.62	
episode: 585	score: 0.0	memory length: 110495	epsilon: 0.9792179200004512
steps: 124	lr: 0.0001	evaluation reward: 1.59	
episode: 586	score: 2.0	memory length: 110713	epsilon: 0.9787862800004605
steps: 218	lr: 0.0001	evaluation reward: 1.58	

episode: 587	score: 1.0	memory length: 110885	epsilon: 0.9784457200004679
steps: 172	lr: 0.0001	evaluation reward: 1.56	
episode: 588	score: 3.0	memory length: 111153	epsilon: 0.9779150800004794
steps: 268	lr: 0.0001	evaluation reward: 1.58	
episode: 589	score: 3.0	memory length: 111365	epsilon: 0.9774953200004886
steps: 212	lr: 0.0001	evaluation reward: 1.61	
episode: 590	score: 2.0	memory length: 111583	epsilon: 0.9770636800004979
steps: 218	lr: 0.0001	evaluation reward: 1.63	
episode: 591	score: 0.0	memory length: 111707	epsilon: 0.9768181600005033
steps: 124	lr: 0.0001	evaluation reward: 1.59	
episode: 592	score: 0.0	memory length: 111830	epsilon: 0.9765746200005085
steps: 123	lr: 0.0001	evaluation reward: 1.58	
episode: 593	score: 0.0	memory length: 111954	epsilon: 0.9763291000005139
steps: 124	lr: 0.0001	evaluation reward: 1.56	
episode: 594	score: 1.0	memory length: 112124	epsilon: 0.9759925000005212
steps: 170	lr: 0.0001	evaluation reward: 1.53	
episode: 595	score: 2.0	memory length: 112323	epsilon: 0.9755984800005297
steps: 199	lr: 0.0001	evaluation reward: 1.54	
episode: 596	score: 2.0	memory length: 112522	epsilon: 0.9752044600005383
steps: 199	lr: 0.0001	evaluation reward: 1.55	
episode: 597	score: 1.0	memory length: 112695	epsilon: 0.9748619200005457
steps: 173	lr: 0.0001	evaluation reward: 1.55	
episode: 598	score: 1.0	memory length: 112846	epsilon: 0.9745629400005522
steps: 151	lr: 0.0001	evaluation reward: 1.54	
episode: 599	score: 0.0	memory length: 112970	epsilon: 0.9743174200005575
steps: 124	lr: 0.0001	evaluation reward: 1.53	
episode: 600	score: 0.0	memory length: 113094	epsilon: 0.9740719000005629
steps: 124	lr: 0.0001	evaluation reward: 1.53	
episode: 601	score: 0.0	memory length: 113218	epsilon: 0.9738263800005682
steps: 124	lr: 0.0001	evaluation reward: 1.53	
episode: 602	score: 0.0	memory length: 113342	epsilon: 0.9735808600005735
steps: 124	lr: 0.0001	evaluation reward: 1.52	
episode: 603	score: 0.0	memory length: 113466	epsilon: 0.9733353400005789
steps: 124	lr: 0.0001	evaluation reward: 1.5	
episode: 604	score: 0.0	memory length: 113590	epsilon: 0.9730898200005842
steps: 124	lr: 0.0001	evaluation reward: 1.46	
episode: 605	score: 3.0	memory length: 113817	epsilon: 0.972640360000594
steps: 227	lr: 0.0001	evaluation reward: 1.47	
episode: 606	score: 1.0	memory length: 113988	epsilon: 0.9723017800006013
steps: 171	lr: 0.0001	evaluation reward: 1.48	
episode: 607	score: 0.0	memory length: 114112	epsilon: 0.9720562600006066
steps: 124	lr: 0.0001	evaluation reward: 1.48	
episode: 608	score: 2.0	memory length: 114331	epsilon: 0.971622640000616
steps: 219	lr: 0.0001	evaluation reward: 1.5	
episode: 609	score: 1.0	memory length: 114484	epsilon: 0.9713197000006226
steps: 153	lr: 0.0001	evaluation reward: 1.49	
episode: 610	score: 1.0	memory length: 114654	epsilon: 0.9709831000006299
steps: 170	lr: 0.0001	evaluation reward: 1.47	

episode: 611	score: 0.0	memory length: 114778	epsilon: 0.9707375800006353
steps: 124	lr: 0.0001	evaluation reward: 1.46	
episode: 612	score: 0.0	memory length: 114902	epsilon: 0.9704920600006406
steps: 124	lr: 0.0001	evaluation reward: 1.46	
episode: 613	score: 0.0	memory length: 115025	epsilon: 0.9702485200006459
steps: 123	lr: 0.0001	evaluation reward: 1.45	
episode: 614	score: 8.0	memory length: 115445	epsilon: 0.9694169200006639
steps: 420	lr: 0.0001	evaluation reward: 1.52	
episode: 615	score: 0.0	memory length: 115568	epsilon: 0.9691733800006692
steps: 123	lr: 0.0001	evaluation reward: 1.51	
episode: 616	score: 0.0	memory length: 115692	epsilon: 0.9689278600006745
steps: 124	lr: 0.0001	evaluation reward: 1.51	
episode: 617	score: 1.0	memory length: 115863	epsilon: 0.9685892800006819
steps: 171	lr: 0.0001	evaluation reward: 1.49	
episode: 618	score: 2.0	memory length: 116062	epsilon: 0.9681952600006904
steps: 199	lr: 0.0001	evaluation reward: 1.49	
episode: 619	score: 1.0	memory length: 116234	epsilon: 0.9678547000006978
steps: 172	lr: 0.0001	evaluation reward: 1.47	
episode: 620	score: 1.0	memory length: 116385	epsilon: 0.9675557200007043
steps: 151	lr: 0.0001	evaluation reward: 1.45	
episode: 621	score: 2.0	memory length: 116570	epsilon: 0.9671894200007123
steps: 185	lr: 0.0001	evaluation reward: 1.45	
episode: 622	score: 1.0	memory length: 116739	epsilon: 0.9668548000007195
steps: 169	lr: 0.0001	evaluation reward: 1.44	
episode: 623	score: 2.0	memory length: 116955	epsilon: 0.9664271200007288
steps: 216	lr: 0.0001	evaluation reward: 1.41	
episode: 624	score: 2.0	memory length: 117142	epsilon: 0.9660568600007369
steps: 187	lr: 0.0001	evaluation reward: 1.4	
episode: 625	score: 0.0	memory length: 117266	epsilon: 0.9658113400007422
steps: 124	lr: 0.0001	evaluation reward: 1.38	
episode: 626	score: 2.0	memory length: 117464	epsilon: 0.9654193000007507
steps: 198	lr: 0.0001	evaluation reward: 1.38	
episode: 627	score: 4.0	memory length: 117735	epsilon: 0.9648827200007624
steps: 271	lr: 0.0001	evaluation reward: 1.41	
episode: 628	score: 1.0	memory length: 117887	epsilon: 0.9645817600007689
steps: 152	lr: 0.0001	evaluation reward: 1.4	
episode: 629	score: 1.0	memory length: 118040	epsilon: 0.9642788200007755
steps: 153	lr: 0.0001	evaluation reward: 1.33	
episode: 630	score: 5.0	memory length: 118405	epsilon: 0.9635561200007912
steps: 365	lr: 0.0001	evaluation reward: 1.34	
episode: 631	score: 0.0	memory length: 118529	epsilon: 0.9633106000007965
steps: 124	lr: 0.0001	evaluation reward: 1.34	
episode: 632	score: 1.0	memory length: 118698	epsilon: 0.9629759800008038
steps: 169	lr: 0.0001	evaluation reward: 1.35	
episode: 633	score: 0.0	memory length: 118821	epsilon: 0.962732440000809
steps: 123	lr: 0.0001	evaluation reward: 1.34	
episode: 634	score: 2.0	memory length: 119040	epsilon: 0.9622988200008185
steps: 219	lr: 0.0001	evaluation reward: 1.34	

episode: 635	score: 2.0	memory length: 119259	epsilon: 0.9618652000008279
steps: 219	lr: 0.0001	evaluation reward: 1.32	
episode: 636	score: 0.0	memory length: 119383	epsilon: 0.9616196800008332
steps: 124	lr: 0.0001	evaluation reward: 1.27	
episode: 637	score: 2.0	memory length: 119599	epsilon: 0.9611920000008425
steps: 216	lr: 0.0001	evaluation reward: 1.28	
episode: 638	score: 0.0	memory length: 119723	epsilon: 0.9609464800008478
steps: 124	lr: 0.0001	evaluation reward: 1.27	
episode: 639	score: 2.0	memory length: 119941	epsilon: 0.9605148400008572
steps: 218	lr: 0.0001	evaluation reward: 1.28	
episode: 640	score: 3.0	memory length: 120208	epsilon: 0.9599861800008687
steps: 267	lr: 0.0001	evaluation reward: 1.31	
episode: 641	score: 1.0	memory length: 120378	epsilon: 0.959649580000876
steps: 170	lr: 0.0001	evaluation reward: 1.3	
episode: 642	score: 1.0	memory length: 120529	epsilon: 0.9593506000008825
steps: 151	lr: 0.0001	evaluation reward: 1.29	
episode: 643	score: 0.0	memory length: 120652	epsilon: 0.9591070600008877
steps: 123	lr: 0.0001	evaluation reward: 1.29	
episode: 644	score: 1.0	memory length: 120822	epsilon: 0.958770460000895
steps: 170	lr: 0.0001	evaluation reward: 1.28	
episode: 645	score: 2.0	memory length: 121041	epsilon: 0.9583368400009045
steps: 219	lr: 0.0001	evaluation reward: 1.3	
episode: 646	score: 1.0	memory length: 121212	epsilon: 0.9579982600009118
steps: 171	lr: 0.0001	evaluation reward: 1.31	
episode: 647	score: 0.0	memory length: 121336	epsilon: 0.9577527400009171
steps: 124	lr: 0.0001	evaluation reward: 1.29	
episode: 648	score: 0.0	memory length: 121459	epsilon: 0.9575092000009224
steps: 123	lr: 0.0001	evaluation reward: 1.28	
episode: 649	score: 1.0	memory length: 121611	epsilon: 0.957208240000929
steps: 152	lr: 0.0001	evaluation reward: 1.28	
episode: 650	score: 2.0	memory length: 121829	epsilon: 0.9567766000009383
steps: 218	lr: 0.0001	evaluation reward: 1.3	
episode: 651	score: 3.0	memory length: 122095	epsilon: 0.9562499200009498
steps: 266	lr: 0.0001	evaluation reward: 1.33	
episode: 652	score: 0.0	memory length: 122219	epsilon: 0.9560044000009551
steps: 124	lr: 0.0001	evaluation reward: 1.3	
episode: 653	score: 0.0	memory length: 122343	epsilon: 0.9557588800009604
steps: 124	lr: 0.0001	evaluation reward: 1.3	
episode: 654	score: 0.0	memory length: 122466	epsilon: 0.9555153400009657
steps: 123	lr: 0.0001	evaluation reward: 1.27	
episode: 655	score: 2.0	memory length: 122667	epsilon: 0.9551173600009744
steps: 201	lr: 0.0001	evaluation reward: 1.26	
episode: 656	score: 1.0	memory length: 122840	epsilon: 0.9547748200009818
steps: 173	lr: 0.0001	evaluation reward: 1.26	
episode: 657	score: 2.0	memory length: 123039	epsilon: 0.9543808000009903
steps: 199	lr: 0.0001	evaluation reward: 1.26	
episode: 658	score: 2.0	memory length: 123239	epsilon: 0.953984800000999
steps: 200	lr: 0.0001	evaluation reward: 1.25	

episode: 659	score: 0.0	memory length: 123363	epsilon: 0.9537392800010043
steps: 124	lr: 0.0001	evaluation reward: 1.22	
episode: 660	score: 3.0	memory length: 123591	epsilon: 0.9532878400010141
steps: 228	lr: 0.0001	evaluation reward: 1.24	
episode: 661	score: 0.0	memory length: 123714	epsilon: 0.9530443000010194
steps: 123	lr: 0.0001	evaluation reward: 1.21	
episode: 662	score: 0.0	memory length: 123838	epsilon: 0.9527987800010247
steps: 124	lr: 0.0001	evaluation reward: 1.19	
episode: 663	score: 2.0	memory length: 124057	epsilon: 0.9523651600010341
steps: 219	lr: 0.0001	evaluation reward: 1.2	
episode: 664	score: 0.0	memory length: 124180	epsilon: 0.9521216200010394
steps: 123	lr: 0.0001	evaluation reward: 1.18	
episode: 665	score: 1.0	memory length: 124332	epsilon: 0.9518206600010459
steps: 152	lr: 0.0001	evaluation reward: 1.19	
episode: 666	score: 0.0	memory length: 124456	epsilon: 0.9515751400010513
steps: 124	lr: 0.0001	evaluation reward: 1.17	
episode: 667	score: 2.0	memory length: 124657	epsilon: 0.9511771600010599
steps: 201	lr: 0.0001	evaluation reward: 1.15	
episode: 668	score: 0.0	memory length: 124780	epsilon: 0.9509336200010652
steps: 123	lr: 0.0001	evaluation reward: 1.13	
episode: 669	score: 0.0	memory length: 124904	epsilon: 0.9506881000010705
steps: 124	lr: 0.0001	evaluation reward: 1.13	
episode: 670	score: 2.0	memory length: 125103	epsilon: 0.9502940800010791
steps: 199	lr: 0.0001	evaluation reward: 1.14	
episode: 671	score: 2.0	memory length: 125283	epsilon: 0.9499376800010868
steps: 180	lr: 0.0001	evaluation reward: 1.15	
episode: 672	score: 2.0	memory length: 125505	epsilon: 0.9494981200010963
steps: 222	lr: 0.0001	evaluation reward: 1.17	
episode: 673	score: 0.0	memory length: 125629	epsilon: 0.9492526000011017
steps: 124	lr: 0.0001	evaluation reward: 1.16	
episode: 674	score: 1.0	memory length: 125801	epsilon: 0.9489120400011091
steps: 172	lr: 0.0001	evaluation reward: 1.17	
episode: 675	score: 2.0	memory length: 125999	epsilon: 0.9485200000011176
steps: 198	lr: 0.0001	evaluation reward: 1.19	
episode: 676	score: 1.0	memory length: 126171	epsilon: 0.9481794400011125
steps: 172	lr: 0.0001	evaluation reward: 1.2	
episode: 677	score: 0.0	memory length: 126294	epsilon: 0.9479359000011303
steps: 123	lr: 0.0001	evaluation reward: 1.18	
episode: 678	score: 0.0	memory length: 126418	epsilon: 0.9476903800011356
steps: 124	lr: 0.0001	evaluation reward: 1.15	
episode: 679	score: 3.0	memory length: 126665	epsilon: 0.9472013200011462
steps: 247	lr: 0.0001	evaluation reward: 1.18	
episode: 680	score: 2.0	memory length: 126866	epsilon: 0.9468033400011548
steps: 201	lr: 0.0001	evaluation reward: 1.18	
episode: 681	score: 0.0	memory length: 126989	epsilon: 0.9465598000011601
steps: 123	lr: 0.0001	evaluation reward: 1.17	
episode: 682	score: 1.0	memory length: 127160	epsilon: 0.9462212200011675
steps: 171	lr: 0.0001	evaluation reward: 1.16	

episode: 683	score: 0.0	memory length: 127284	epsilon: 0.9459757000011728
steps: 124	lr: 0.0001	evaluation reward: 1.15	
episode: 684	score: 3.0	memory length: 127515	epsilon: 0.9455183200011827
steps: 231	lr: 0.0001	evaluation reward: 1.17	
episode: 685	score: 1.0	memory length: 127685	epsilon: 0.94518172000119
steps: 170	lr: 0.0001	evaluation reward: 1.18	
episode: 686	score: 2.0	memory length: 127884	epsilon: 0.9447877000011986
steps: 199	lr: 0.0001	evaluation reward: 1.18	
episode: 687	score: 2.0	memory length: 128083	epsilon: 0.9443936800012072
steps: 199	lr: 0.0001	evaluation reward: 1.19	
episode: 688	score: 0.0	memory length: 128206	epsilon: 0.9441501400012124
steps: 123	lr: 0.0001	evaluation reward: 1.16	
episode: 689	score: 1.0	memory length: 128376	epsilon: 0.9438135400012198
steps: 170	lr: 0.0001	evaluation reward: 1.14	
episode: 690	score: 2.0	memory length: 128599	epsilon: 0.9433720000012293
steps: 223	lr: 0.0001	evaluation reward: 1.14	
episode: 691	score: 1.0	memory length: 128768	epsilon: 0.9430373800012366
steps: 169	lr: 0.0001	evaluation reward: 1.15	
episode: 692	score: 0.0	memory length: 128892	epsilon: 0.9427918600012419
steps: 124	lr: 0.0001	evaluation reward: 1.15	
episode: 693	score: 2.0	memory length: 129090	epsilon: 0.9423998200012504
steps: 198	lr: 0.0001	evaluation reward: 1.17	
episode: 694	score: 3.0	memory length: 129337	epsilon: 0.9419107600012611
steps: 247	lr: 0.0001	evaluation reward: 1.19	
episode: 695	score: 2.0	memory length: 129556	epsilon: 0.9414771400012705
steps: 219	lr: 0.0001	evaluation reward: 1.19	
episode: 696	score: 1.0	memory length: 129708	epsilon: 0.941176180001277
steps: 152	lr: 0.0001	evaluation reward: 1.18	
episode: 697	score: 2.0	memory length: 129906	epsilon: 0.9407841400012855
steps: 198	lr: 0.0001	evaluation reward: 1.19	
episode: 698	score: 4.0	memory length: 130178	epsilon: 0.9402455800012972
steps: 272	lr: 0.0001	evaluation reward: 1.22	
episode: 699	score: 2.0	memory length: 130379	epsilon: 0.9398476000013058
steps: 201	lr: 0.0001	evaluation reward: 1.24	
episode: 700	score: 4.0	memory length: 130658	epsilon: 0.9392951800013178
steps: 279	lr: 0.0001	evaluation reward: 1.28	
episode: 701	score: 1.0	memory length: 130810	epsilon: 0.9389942200013244
steps: 152	lr: 0.0001	evaluation reward: 1.29	
episode: 702	score: 0.0	memory length: 130933	epsilon: 0.9387506800013297
steps: 123	lr: 0.0001	evaluation reward: 1.29	
episode: 703	score: 3.0	memory length: 131198	epsilon: 0.938225980001341
steps: 265	lr: 0.0001	evaluation reward: 1.32	
episode: 704	score: 1.0	memory length: 131369	epsilon: 0.9378874000013484
steps: 171	lr: 0.0001	evaluation reward: 1.33	
episode: 705	score: 1.0	memory length: 131539	epsilon: 0.9375508000013557
steps: 170	lr: 0.0001	evaluation reward: 1.31	
episode: 706	score: 3.0	memory length: 131786	epsilon: 0.9370617400013663
steps: 247	lr: 0.0001	evaluation reward: 1.33	

episode: 707	score: 3.0	memory length: 132017	epsilon: 0.9366043600013763
steps: 231	lr: 0.0001	evaluation reward: 1.36	
episode: 708	score: 2.0	memory length: 132215	epsilon: 0.9362123200013848
steps: 198	lr: 0.0001	evaluation reward: 1.36	
episode: 709	score: 0.0	memory length: 132339	epsilon: 0.9359668000013901
steps: 124	lr: 0.0001	evaluation reward: 1.35	
episode: 710	score: 1.0	memory length: 132510	epsilon: 0.9356282200013974
steps: 171	lr: 0.0001	evaluation reward: 1.35	
episode: 711	score: 3.0	memory length: 132740	epsilon: 0.9351728200014073
steps: 230	lr: 0.0001	evaluation reward: 1.38	
episode: 712	score: 5.0	memory length: 133039	epsilon: 0.9345808000014202
steps: 299	lr: 0.0001	evaluation reward: 1.43	
episode: 713	score: 3.0	memory length: 133307	epsilon: 0.9340501600014317
steps: 268	lr: 0.0001	evaluation reward: 1.46	
episode: 714	score: 0.0	memory length: 133431	epsilon: 0.933804640001437
steps: 124	lr: 0.0001	evaluation reward: 1.38	
episode: 715	score: 3.0	memory length: 133677	epsilon: 0.9333175600014476
steps: 246	lr: 0.0001	evaluation reward: 1.41	
episode: 716	score: 5.0	memory length: 134004	epsilon: 0.9326701000014617
steps: 327	lr: 0.0001	evaluation reward: 1.46	
episode: 717	score: 0.0	memory length: 134128	epsilon: 0.932424580001467
steps: 124	lr: 0.0001	evaluation reward: 1.45	
episode: 718	score: 1.0	memory length: 134298	epsilon: 0.9320879800014743
steps: 170	lr: 0.0001	evaluation reward: 1.44	
episode: 719	score: 0.0	memory length: 134421	epsilon: 0.9318444400014796
steps: 123	lr: 0.0001	evaluation reward: 1.43	
episode: 720	score: 0.0	memory length: 134544	epsilon: 0.9316009000014849
steps: 123	lr: 0.0001	evaluation reward: 1.42	
episode: 721	score: 2.0	memory length: 134748	epsilon: 0.9311969800014936
steps: 204	lr: 0.0001	evaluation reward: 1.42	
episode: 722	score: 2.0	memory length: 134968	epsilon: 0.9307613800015031
steps: 220	lr: 0.0001	evaluation reward: 1.43	
episode: 723	score: 1.0	memory length: 135139	epsilon: 0.9304228000015105
steps: 171	lr: 0.0001	evaluation reward: 1.42	
episode: 724	score: 0.0	memory length: 135263	epsilon: 0.9301772800015158
steps: 124	lr: 0.0001	evaluation reward: 1.4	
episode: 725	score: 1.0	memory length: 135433	epsilon: 0.9298406800015231
steps: 170	lr: 0.0001	evaluation reward: 1.41	
episode: 726	score: 4.0	memory length: 135730	epsilon: 0.9292526200015359
steps: 297	lr: 0.0001	evaluation reward: 1.43	
episode: 727	score: 1.0	memory length: 135882	epsilon: 0.9289516600015424
steps: 152	lr: 0.0001	evaluation reward: 1.4	
episode: 728	score: 2.0	memory length: 136099	epsilon: 0.9285220000015517
steps: 217	lr: 0.0001	evaluation reward: 1.41	
episode: 729	score: 0.0	memory length: 136223	epsilon: 0.928276480001557
steps: 124	lr: 0.0001	evaluation reward: 1.4	
episode: 730	score: 1.0	memory length: 136376	epsilon: 0.9279735400015636
steps: 153	lr: 0.0001	evaluation reward: 1.36	

episode: 731	score: 0.0	memory length: 136500	epsilon: 0.927728020001569
steps: 124	lr: 0.0001	evaluation reward: 1.36	
episode: 732	score: 1.0	memory length: 136652	epsilon: 0.9274270600015755
steps: 152	lr: 0.0001	evaluation reward: 1.36	
episode: 733	score: 1.0	memory length: 136823	epsilon: 0.9270884800015828
steps: 171	lr: 0.0001	evaluation reward: 1.37	
episode: 734	score: 2.0	memory length: 137022	epsilon: 0.9266944600015914
steps: 199	lr: 0.0001	evaluation reward: 1.37	
episode: 735	score: 5.0	memory length: 137346	epsilon: 0.9260529400016053
steps: 324	lr: 0.0001	evaluation reward: 1.4	
episode: 736	score: 2.0	memory length: 137547	epsilon: 0.925654960001614
steps: 201	lr: 0.0001	evaluation reward: 1.42	
episode: 737	score: 2.0	memory length: 137746	epsilon: 0.9252609400016225
steps: 199	lr: 0.0001	evaluation reward: 1.42	
episode: 738	score: 0.0	memory length: 137870	epsilon: 0.9250154200016278
steps: 124	lr: 0.0001	evaluation reward: 1.42	
episode: 739	score: 3.0	memory length: 138097	epsilon: 0.9245659600016376
steps: 227	lr: 0.0001	evaluation reward: 1.43	
episode: 740	score: 1.0	memory length: 138267	epsilon: 0.9242293600016449
steps: 170	lr: 0.0001	evaluation reward: 1.41	
episode: 741	score: 2.0	memory length: 138466	epsilon: 0.9238353400016535
steps: 199	lr: 0.0001	evaluation reward: 1.42	
episode: 742	score: 0.0	memory length: 138590	epsilon: 0.9235898200016588
steps: 124	lr: 0.0001	evaluation reward: 1.41	
episode: 743	score: 2.0	memory length: 138771	epsilon: 0.9232314400016666
steps: 181	lr: 0.0001	evaluation reward: 1.43	
episode: 744	score: 4.0	memory length: 139067	epsilon: 0.9226453600016793
steps: 296	lr: 0.0001	evaluation reward: 1.46	
episode: 745	score: 7.0	memory length: 139368	epsilon: 0.9220493800016922
steps: 301	lr: 0.0001	evaluation reward: 1.51	
episode: 746	score: 2.0	memory length: 139567	epsilon: 0.9216553600017008
steps: 199	lr: 0.0001	evaluation reward: 1.52	
episode: 747	score: 1.0	memory length: 139736	epsilon: 0.921320740001708
steps: 169	lr: 0.0001	evaluation reward: 1.53	
episode: 748	score: 1.0	memory length: 139887	epsilon: 0.9210217600017145
steps: 151	lr: 0.0001	evaluation reward: 1.54	
episode: 749	score: 0.0	memory length: 140010	epsilon: 0.9207782200017198
steps: 123	lr: 0.0001	evaluation reward: 1.53	
episode: 750	score: 3.0	memory length: 140255	epsilon: 0.9202931200017304
steps: 245	lr: 0.0001	evaluation reward: 1.54	
episode: 751	score: 2.0	memory length: 140473	epsilon: 0.9198614800017397
steps: 218	lr: 0.0001	evaluation reward: 1.53	
episode: 752	score: 4.0	memory length: 140750	epsilon: 0.9193130200017516
steps: 277	lr: 0.0001	evaluation reward: 1.57	
episode: 753	score: 4.0	memory length: 141049	epsilon: 0.9187210000017645
steps: 299	lr: 0.0001	evaluation reward: 1.61	
episode: 754	score: 2.0	memory length: 141248	epsilon: 0.918326980001773
steps: 199	lr: 0.0001	evaluation reward: 1.63	

episode: 755	score: 3.0	memory length: 141480	epsilon: 0.917867620001783
steps: 232	lr: 0.0001	evaluation reward: 1.64	
episode: 756	score: 4.0	memory length: 141797	epsilon: 0.9172399600017966
steps: 317	lr: 0.0001	evaluation reward: 1.67	
episode: 757	score: 3.0	memory length: 142045	epsilon: 0.9167489200018073
steps: 248	lr: 0.0001	evaluation reward: 1.68	
episode: 758	score: 2.0	memory length: 142243	epsilon: 0.9163568800018158
steps: 198	lr: 0.0001	evaluation reward: 1.68	
episode: 759	score: 2.0	memory length: 142441	epsilon: 0.9159648400018243
steps: 198	lr: 0.0001	evaluation reward: 1.7	
episode: 760	score: 3.0	memory length: 142712	epsilon: 0.915428260001836
steps: 271	lr: 0.0001	evaluation reward: 1.7	
episode: 761	score: 2.0	memory length: 142911	epsilon: 0.9150342400018445
steps: 199	lr: 0.0001	evaluation reward: 1.72	
episode: 762	score: 0.0	memory length: 143035	epsilon: 0.9147887200018499
steps: 124	lr: 0.0001	evaluation reward: 1.72	
episode: 763	score: 1.0	memory length: 143186	epsilon: 0.9144897400018563
steps: 151	lr: 0.0001	evaluation reward: 1.71	
episode: 764	score: 2.0	memory length: 143366	epsilon: 0.9141333400018641
steps: 180	lr: 0.0001	evaluation reward: 1.73	
episode: 765	score: 0.0	memory length: 143489	epsilon: 0.9138898000018694
steps: 123	lr: 0.0001	evaluation reward: 1.72	
episode: 766	score: 1.0	memory length: 143643	epsilon: 0.913584880001876
steps: 154	lr: 0.0001	evaluation reward: 1.73	
episode: 767	score: 2.0	memory length: 143862	epsilon: 0.9131512600018854
steps: 219	lr: 0.0001	evaluation reward: 1.73	
episode: 768	score: 2.0	memory length: 144060	epsilon: 0.9127592200018939
steps: 198	lr: 0.0001	evaluation reward: 1.75	
episode: 769	score: 2.0	memory length: 144240	epsilon: 0.9124028200019016
steps: 180	lr: 0.0001	evaluation reward: 1.77	
episode: 770	score: 0.0	memory length: 144364	epsilon: 0.912157300001907
steps: 124	lr: 0.0001	evaluation reward: 1.75	
episode: 771	score: 2.0	memory length: 144583	epsilon: 0.9117236800019164
steps: 219	lr: 0.0001	evaluation reward: 1.75	
episode: 772	score: 2.0	memory length: 144782	epsilon: 0.911329660001925
steps: 199	lr: 0.0001	evaluation reward: 1.75	
episode: 773	score: 3.0	memory length: 145049	epsilon: 0.9108010000019364
steps: 267	lr: 0.0001	evaluation reward: 1.78	
episode: 774	score: 0.0	memory length: 145172	epsilon: 0.9105574600019417
steps: 123	lr: 0.0001	evaluation reward: 1.77	
episode: 775	score: 0.0	memory length: 145296	epsilon: 0.910311940001947
steps: 124	lr: 0.0001	evaluation reward: 1.75	
episode: 776	score: 0.0	memory length: 145419	epsilon: 0.9100684000019523
steps: 123	lr: 0.0001	evaluation reward: 1.74	
episode: 777	score: 0.0	memory length: 145543	epsilon: 0.9098228800019577
steps: 124	lr: 0.0001	evaluation reward: 1.74	
episode: 778	score: 3.0	memory length: 145770	epsilon: 0.9093734200019674
steps: 227	lr: 0.0001	evaluation reward: 1.77	

episode: 779	score: 0.0	memory length: 145894	epsilon: 0.9091279000019727
steps: 124	lr: 0.0001	evaluation reward: 1.74	
episode: 780	score: 0.0	memory length: 146018	epsilon: 0.9088823800019781
steps: 124	lr: 0.0001	evaluation reward: 1.72	
episode: 781	score: 2.0	memory length: 146217	epsilon: 0.9084883600019866
steps: 199	lr: 0.0001	evaluation reward: 1.74	
episode: 782	score: 0.0	memory length: 146340	epsilon: 0.9082448200019919
steps: 123	lr: 0.0001	evaluation reward: 1.73	
episode: 783	score: 2.0	memory length: 146538	epsilon: 0.9078527800020004
steps: 198	lr: 0.0001	evaluation reward: 1.75	
episode: 784	score: 2.0	memory length: 146737	epsilon: 0.907458760002009
steps: 199	lr: 0.0001	evaluation reward: 1.74	
episode: 785	score: 2.0	memory length: 146935	epsilon: 0.9070667200020175
steps: 198	lr: 0.0001	evaluation reward: 1.75	
episode: 786	score: 0.0	memory length: 147059	epsilon: 0.9068212000020228
steps: 124	lr: 0.0001	evaluation reward: 1.73	
episode: 787	score: 3.0	memory length: 147286	epsilon: 0.9063717400020326
steps: 227	lr: 0.0001	evaluation reward: 1.74	
episode: 788	score: 2.0	memory length: 147485	epsilon: 0.9059777200020411
steps: 199	lr: 0.0001	evaluation reward: 1.76	
episode: 789	score: 3.0	memory length: 147731	epsilon: 0.9054906400020517
steps: 246	lr: 0.0001	evaluation reward: 1.78	
episode: 790	score: 1.0	memory length: 147900	epsilon: 0.905156020002059
steps: 169	lr: 0.0001	evaluation reward: 1.77	
episode: 791	score: 0.0	memory length: 148024	epsilon: 0.9049105000020643
steps: 124	lr: 0.0001	evaluation reward: 1.76	
episode: 792	score: 0.0	memory length: 148147	epsilon: 0.9046669600020696
steps: 123	lr: 0.0001	evaluation reward: 1.76	
episode: 793	score: 1.0	memory length: 148319	epsilon: 0.904326400002077
steps: 172	lr: 0.0001	evaluation reward: 1.75	
episode: 794	score: 2.0	memory length: 148500	epsilon: 0.9039680200020848
steps: 181	lr: 0.0001	evaluation reward: 1.74	
episode: 795	score: 3.0	memory length: 148770	epsilon: 0.9034334200020964
steps: 270	lr: 0.0001	evaluation reward: 1.75	
episode: 796	score: 4.0	memory length: 149088	epsilon: 0.90280378000211
steps: 318	lr: 0.0001	evaluation reward: 1.78	
episode: 797	score: 0.0	memory length: 149211	epsilon: 0.9025602400021153
steps: 123	lr: 0.0001	evaluation reward: 1.76	
episode: 798	score: 0.0	memory length: 149335	epsilon: 0.9023147200021207
steps: 124	lr: 0.0001	evaluation reward: 1.72	
episode: 799	score: 3.0	memory length: 149561	epsilon: 0.9018672400021304
steps: 226	lr: 0.0001	evaluation reward: 1.73	
episode: 800	score: 1.0	memory length: 149734	epsilon: 0.9015247000021378
steps: 173	lr: 0.0001	evaluation reward: 1.7	
episode: 801	score: 3.0	memory length: 149961	epsilon: 0.9010752400021476
steps: 227	lr: 0.0001	evaluation reward: 1.72	
episode: 802	score: 3.0	memory length: 150230	epsilon: 0.9005426200021591
steps: 269	lr: 0.0001	evaluation reward: 1.75	

episode: 803	score: 0.0	memory length: 150354	epsilon: 0.9002971000021645
steps: 124	lr: 0.0001	evaluation reward: 1.72	
episode: 804	score: 2.0	memory length: 150557	epsilon: 0.8998951600021732
steps: 203	lr: 0.0001	evaluation reward: 1.73	
episode: 805	score: 3.0	memory length: 150807	epsilon: 0.8994001600021839
steps: 250	lr: 0.0001	evaluation reward: 1.75	
episode: 806	score: 4.0	memory length: 151104	epsilon: 0.8988121000021967
steps: 297	lr: 0.0001	evaluation reward: 1.76	
episode: 807	score: 0.0	memory length: 151228	epsilon: 0.898566580002202
steps: 124	lr: 0.0001	evaluation reward: 1.73	
episode: 808	score: 2.0	memory length: 151449	epsilon: 0.8981290000022115
steps: 221	lr: 0.0001	evaluation reward: 1.73	
episode: 809	score: 1.0	memory length: 151622	epsilon: 0.897786460002219
steps: 173	lr: 0.0001	evaluation reward: 1.74	
episode: 810	score: 2.0	memory length: 151803	epsilon: 0.8974280800022267
steps: 181	lr: 0.0001	evaluation reward: 1.75	
episode: 811	score: 3.0	memory length: 152070	epsilon: 0.8968994200022382
steps: 267	lr: 0.0001	evaluation reward: 1.75	
episode: 812	score: 2.0	memory length: 152269	epsilon: 0.8965054000022468
steps: 199	lr: 0.0001	evaluation reward: 1.72	
episode: 813	score: 0.0	memory length: 152393	epsilon: 0.8962598800022521
steps: 124	lr: 0.0001	evaluation reward: 1.69	
episode: 814	score: 0.0	memory length: 152516	epsilon: 0.8960163400022574
steps: 123	lr: 0.0001	evaluation reward: 1.69	
episode: 815	score: 1.0	memory length: 152668	epsilon: 0.8957153800022639
steps: 152	lr: 0.0001	evaluation reward: 1.67	
episode: 816	score: 1.0	memory length: 152840	epsilon: 0.8953748200022713
steps: 172	lr: 0.0001	evaluation reward: 1.63	
episode: 817	score: 0.0	memory length: 152964	epsilon: 0.8951293000022766
steps: 124	lr: 0.0001	evaluation reward: 1.63	
episode: 818	score: 3.0	memory length: 153231	epsilon: 0.8946006400022881
steps: 267	lr: 0.0001	evaluation reward: 1.65	
episode: 819	score: 2.0	memory length: 153450	epsilon: 0.8941670200022975
steps: 219	lr: 0.0001	evaluation reward: 1.67	
episode: 820	score: 6.0	memory length: 153817	epsilon: 0.8934403600023133
steps: 367	lr: 0.0001	evaluation reward: 1.73	
episode: 821	score: 3.0	memory length: 154085	epsilon: 0.8929097200023248
steps: 268	lr: 0.0001	evaluation reward: 1.74	
episode: 822	score: 4.0	memory length: 154384	epsilon: 0.8923177000023377
steps: 299	lr: 0.0001	evaluation reward: 1.76	
episode: 823	score: 1.0	memory length: 154554	epsilon: 0.891981100002345
steps: 170	lr: 0.0001	evaluation reward: 1.76	
episode: 824	score: 0.0	memory length: 154678	epsilon: 0.8917355800023503
steps: 124	lr: 0.0001	evaluation reward: 1.76	
episode: 825	score: 2.0	memory length: 154877	epsilon: 0.8913415600023589
steps: 199	lr: 0.0001	evaluation reward: 1.77	
episode: 826	score: 1.0	memory length: 155049	epsilon: 0.8910010000023663
steps: 172	lr: 0.0001	evaluation reward: 1.74	

episode: 827	score: 1.0	memory length: 155201	epsilon: 0.8907000400023728
steps: 152	lr: 0.0001	evaluation reward: 1.74	
episode: 828	score: 4.0	memory length: 155495	epsilon: 0.8901179200023854
steps: 294	lr: 0.0001	evaluation reward: 1.76	
episode: 829	score: 1.0	memory length: 155647	epsilon: 0.889816960002392
steps: 152	lr: 0.0001	evaluation reward: 1.77	
episode: 830	score: 0.0	memory length: 155771	epsilon: 0.8895714400023973
steps: 124	lr: 0.0001	evaluation reward: 1.76	
episode: 831	score: 4.0	memory length: 156048	epsilon: 0.8890229800024092
steps: 277	lr: 0.0001	evaluation reward: 1.8	
episode: 832	score: 0.0	memory length: 156172	epsilon: 0.8887774600024145
steps: 124	lr: 0.0001	evaluation reward: 1.79	
episode: 833	score: 0.0	memory length: 156296	epsilon: 0.8885319400024199
steps: 124	lr: 0.0001	evaluation reward: 1.78	
episode: 834	score: 0.0	memory length: 156420	epsilon: 0.8882864200024252
steps: 124	lr: 0.0001	evaluation reward: 1.76	
episode: 835	score: 2.0	memory length: 156603	epsilon: 0.8879240800024331
steps: 183	lr: 0.0001	evaluation reward: 1.73	
episode: 836	score: 1.0	memory length: 156772	epsilon: 0.8875894600024403
steps: 169	lr: 0.0001	evaluation reward: 1.72	
episode: 837	score: 3.0	memory length: 156998	epsilon: 0.88714198000245
steps: 226	lr: 0.0001	evaluation reward: 1.73	
episode: 838	score: 0.0	memory length: 157122	epsilon: 0.8868964600024554
steps: 124	lr: 0.0001	evaluation reward: 1.73	
episode: 839	score: 2.0	memory length: 157320	epsilon: 0.8865044200024639
steps: 198	lr: 0.0001	evaluation reward: 1.72	
episode: 840	score: 1.0	memory length: 157472	epsilon: 0.8862034600024704
steps: 152	lr: 0.0001	evaluation reward: 1.72	
episode: 841	score: 2.0	memory length: 157653	epsilon: 0.8858450800024782
steps: 181	lr: 0.0001	evaluation reward: 1.72	
episode: 842	score: 5.0	memory length: 157979	epsilon: 0.8851996000024922
steps: 326	lr: 0.0001	evaluation reward: 1.77	
episode: 843	score: 3.0	memory length: 158225	epsilon: 0.8847125200025028
steps: 246	lr: 0.0001	evaluation reward: 1.78	
episode: 844	score: 1.0	memory length: 158395	epsilon: 0.8843759200025101
steps: 170	lr: 0.0001	evaluation reward: 1.75	
episode: 845	score: 0.0	memory length: 158518	epsilon: 0.8841323800025154
steps: 123	lr: 0.0001	evaluation reward: 1.68	
episode: 846	score: 2.0	memory length: 158720	epsilon: 0.883732420002524
steps: 202	lr: 0.0001	evaluation reward: 1.68	
episode: 847	score: 1.0	memory length: 158892	epsilon: 0.8833918600025314
steps: 172	lr: 0.0001	evaluation reward: 1.68	
episode: 848	score: 0.0	memory length: 159016	epsilon: 0.8831463400025368
steps: 124	lr: 0.0001	evaluation reward: 1.67	
episode: 849	score: 1.0	memory length: 159186	epsilon: 0.8828097400025441
steps: 170	lr: 0.0001	evaluation reward: 1.68	
episode: 850	score: 2.0	memory length: 159384	epsilon: 0.8824177000025526
steps: 198	lr: 0.0001	evaluation reward: 1.67	

episode: 851	score: 7.0	memory length: 159684	epsilon: 0.8818237000025655
steps: 300	lr: 0.0001	evaluation reward: 1.72	
episode: 852	score: 3.0	memory length: 159935	epsilon: 0.8813267200025763
steps: 251	lr: 0.0001	evaluation reward: 1.71	
episode: 853	score: 2.0	memory length: 160134	epsilon: 0.8809327000025848
steps: 199	lr: 0.0001	evaluation reward: 1.69	
episode: 854	score: 0.0	memory length: 160257	epsilon: 0.8806891600025901
steps: 123	lr: 0.0001	evaluation reward: 1.67	
episode: 855	score: 1.0	memory length: 160429	epsilon: 0.8803486000025975
steps: 172	lr: 0.0001	evaluation reward: 1.65	
episode: 856	score: 5.0	memory length: 160709	epsilon: 0.8797942000026095
steps: 280	lr: 0.0001	evaluation reward: 1.66	
episode: 857	score: 2.0	memory length: 160928	epsilon: 0.879360580002619
steps: 219	lr: 0.0001	evaluation reward: 1.65	
episode: 858	score: 2.0	memory length: 161147	epsilon: 0.8789269600026284
steps: 219	lr: 0.0001	evaluation reward: 1.65	
episode: 859	score: 0.0	memory length: 161271	epsilon: 0.8786814400026337
steps: 124	lr: 0.0001	evaluation reward: 1.63	
episode: 860	score: 0.0	memory length: 161395	epsilon: 0.878435920002639
steps: 124	lr: 0.0001	evaluation reward: 1.6	
episode: 861	score: 1.0	memory length: 161548	epsilon: 0.8781329800026456
steps: 153	lr: 0.0001	evaluation reward: 1.59	
episode: 862	score: 2.0	memory length: 161746	epsilon: 0.8777409400026541
steps: 198	lr: 0.0001	evaluation reward: 1.61	
episode: 863	score: 2.0	memory length: 161945	epsilon: 0.8773469200026627
steps: 199	lr: 0.0001	evaluation reward: 1.62	
episode: 864	score: 2.0	memory length: 162146	epsilon: 0.8769489400026713
steps: 201	lr: 0.0001	evaluation reward: 1.62	
episode: 865	score: 0.0	memory length: 162270	epsilon: 0.8767034200026766
steps: 124	lr: 0.0001	evaluation reward: 1.62	
episode: 866	score: 3.0	memory length: 162535	epsilon: 0.876178720002688
steps: 265	lr: 0.0001	evaluation reward: 1.64	
episode: 867	score: 4.0	memory length: 162832	epsilon: 0.8755906600027008
steps: 297	lr: 0.0001	evaluation reward: 1.66	
episode: 868	score: 1.0	memory length: 163004	epsilon: 0.8752501000027082
steps: 172	lr: 0.0001	evaluation reward: 1.65	
episode: 869	score: 0.0	memory length: 163128	epsilon: 0.8750045800027135
steps: 124	lr: 0.0001	evaluation reward: 1.63	
episode: 870	score: 2.0	memory length: 163347	epsilon: 0.8745709600027229
steps: 219	lr: 0.0001	evaluation reward: 1.65	
episode: 871	score: 3.0	memory length: 163577	epsilon: 0.8741155600027328
steps: 230	lr: 0.0001	evaluation reward: 1.66	
episode: 872	score: 2.0	memory length: 163776	epsilon: 0.8737215400027414
steps: 199	lr: 0.0001	evaluation reward: 1.66	
episode: 873	score: 2.0	memory length: 163975	epsilon: 0.8733275200027499
steps: 199	lr: 0.0001	evaluation reward: 1.65	
episode: 874	score: 2.0	memory length: 164174	epsilon: 0.8729335000027585
steps: 199	lr: 0.0001	evaluation reward: 1.67	

episode: 875	score: 7.0	memory length: 164475	epsilon: 0.8723375200027714
steps: 301	lr: 0.0001	evaluation reward: 1.74	
episode: 876	score: 0.0	memory length: 164599	epsilon: 0.8720920000027768
steps: 124	lr: 0.0001	evaluation reward: 1.74	
episode: 877	score: 0.0	memory length: 164723	epsilon: 0.8718464800027821
steps: 124	lr: 0.0001	evaluation reward: 1.74	
episode: 878	score: 2.0	memory length: 164924	epsilon: 0.8714485000027907
steps: 201	lr: 0.0001	evaluation reward: 1.73	
episode: 879	score: 4.0	memory length: 165180	epsilon: 0.8709416200028017
steps: 256	lr: 0.0001	evaluation reward: 1.77	
episode: 880	score: 1.0	memory length: 165332	epsilon: 0.8706406600028083
steps: 152	lr: 0.0001	evaluation reward: 1.78	
episode: 881	score: 3.0	memory length: 165596	epsilon: 0.8701179400028196
steps: 264	lr: 0.0001	evaluation reward: 1.79	
episode: 882	score: 3.0	memory length: 165824	epsilon: 0.8696665000028294
steps: 228	lr: 0.0001	evaluation reward: 1.82	
episode: 883	score: 0.0	memory length: 165947	epsilon: 0.8694229600028347
steps: 123	lr: 0.0001	evaluation reward: 1.8	
episode: 884	score: 1.0	memory length: 166116	epsilon: 0.869088340002842
steps: 169	lr: 0.0001	evaluation reward: 1.79	
episode: 885	score: 2.0	memory length: 166318	epsilon: 0.8686883800028506
steps: 202	lr: 0.0001	evaluation reward: 1.79	
episode: 886	score: 5.0	memory length: 166621	epsilon: 0.8680884400028637
steps: 303	lr: 0.0001	evaluation reward: 1.84	
episode: 887	score: 2.0	memory length: 166820	epsilon: 0.8676944200028722
steps: 199	lr: 0.0001	evaluation reward: 1.83	
episode: 888	score: 5.0	memory length: 167110	epsilon: 0.8671202200028847
steps: 290	lr: 0.0001	evaluation reward: 1.86	
episode: 889	score: 1.0	memory length: 167263	epsilon: 0.8668172800028913
steps: 153	lr: 0.0001	evaluation reward: 1.84	
episode: 890	score: 3.0	memory length: 167527	epsilon: 0.8662945600029026
steps: 264	lr: 0.0001	evaluation reward: 1.86	
episode: 891	score: 4.0	memory length: 167805	epsilon: 0.8657441200029146
steps: 278	lr: 0.0001	evaluation reward: 1.9	
episode: 892	score: 2.0	memory length: 168022	epsilon: 0.8653144600029239
steps: 217	lr: 0.0001	evaluation reward: 1.92	
episode: 893	score: 4.0	memory length: 168299	epsilon: 0.8647660000029358
steps: 277	lr: 0.0001	evaluation reward: 1.95	
episode: 894	score: 4.0	memory length: 168580	epsilon: 0.8642096200029479
steps: 281	lr: 0.0001	evaluation reward: 1.97	
episode: 895	score: 2.0	memory length: 168778	epsilon: 0.8638175800029564
steps: 198	lr: 0.0001	evaluation reward: 1.96	
episode: 896	score: 4.0	memory length: 169076	epsilon: 0.8632275400029692
steps: 298	lr: 0.0001	evaluation reward: 1.96	
episode: 897	score: 3.0	memory length: 169322	epsilon: 0.8627404600029798
steps: 246	lr: 0.0001	evaluation reward: 1.99	
episode: 898	score: 0.0	memory length: 169446	epsilon: 0.8624949400029851
steps: 124	lr: 0.0001	evaluation reward: 1.99	

episode: 899	score: 4.0	memory length: 169764	epsilon: 0.8618653000029988
steps: 318	lr: 0.0001	evaluation reward: 2.0	
episode: 900	score: 1.0	memory length: 169916	epsilon: 0.8615643400030053
steps: 152	lr: 0.0001	evaluation reward: 2.0	
episode: 901	score: 1.0	memory length: 170086	epsilon: 0.8612277400030126
steps: 170	lr: 0.0001	evaluation reward: 1.98	
episode: 902	score: 1.0	memory length: 170238	epsilon: 0.8609267800030191
steps: 152	lr: 0.0001	evaluation reward: 1.96	
episode: 903	score: 1.0	memory length: 170391	epsilon: 0.8606238400030257
steps: 153	lr: 0.0001	evaluation reward: 1.97	
episode: 904	score: 3.0	memory length: 170639	epsilon: 0.8601328000030364
steps: 248	lr: 0.0001	evaluation reward: 1.98	
episode: 905	score: 2.0	memory length: 170838	epsilon: 0.8597387800030449
steps: 199	lr: 0.0001	evaluation reward: 1.97	
episode: 906	score: 2.0	memory length: 171037	epsilon: 0.8593447600030535
steps: 199	lr: 0.0001	evaluation reward: 1.95	
episode: 907	score: 1.0	memory length: 171208	epsilon: 0.8590061800030608
steps: 171	lr: 0.0001	evaluation reward: 1.96	
episode: 908	score: 0.0	memory length: 171332	epsilon: 0.8587606600030662
steps: 124	lr: 0.0001	evaluation reward: 1.94	
episode: 909	score: 1.0	memory length: 171483	epsilon: 0.8584616800030727
steps: 151	lr: 0.0001	evaluation reward: 1.94	
episode: 910	score: 2.0	memory length: 171703	epsilon: 0.8580260800030821
steps: 220	lr: 0.0001	evaluation reward: 1.94	
episode: 911	score: 3.0	memory length: 171950	epsilon: 0.8575370200030927
steps: 247	lr: 0.0001	evaluation reward: 1.94	
episode: 912	score: 2.0	memory length: 172131	epsilon: 0.8571786400031005
steps: 181	lr: 0.0001	evaluation reward: 1.94	
episode: 913	score: 6.0	memory length: 172486	epsilon: 0.8564757400031158
steps: 355	lr: 0.0001	evaluation reward: 2.0	
episode: 914	score: 2.0	memory length: 172703	epsilon: 0.8560460800031251
steps: 217	lr: 0.0001	evaluation reward: 2.02	
episode: 915	score: 3.0	memory length: 172930	epsilon: 0.8555966200031349
steps: 227	lr: 0.0001	evaluation reward: 2.04	
episode: 916	score: 4.0	memory length: 173190	epsilon: 0.855081820003146
steps: 260	lr: 0.0001	evaluation reward: 2.07	
episode: 917	score: 1.0	memory length: 173362	epsilon: 0.8547412600031534
steps: 172	lr: 0.0001	evaluation reward: 2.08	
episode: 918	score: 1.0	memory length: 173535	epsilon: 0.8543987200031609
steps: 173	lr: 0.0001	evaluation reward: 2.06	
episode: 919	score: 2.0	memory length: 173735	epsilon: 0.8540027200031695
steps: 200	lr: 0.0001	evaluation reward: 2.06	
episode: 920	score: 2.0	memory length: 173934	epsilon: 0.853608700003178
steps: 199	lr: 0.0001	evaluation reward: 2.02	
episode: 921	score: 5.0	memory length: 174300	epsilon: 0.8528840200031937
steps: 366	lr: 0.0001	evaluation reward: 2.04	
episode: 922	score: 1.0	memory length: 174472	epsilon: 0.8525434600032011
steps: 172	lr: 0.0001	evaluation reward: 2.01	

episode: 923	score: 2.0	memory length: 174673	epsilon: 0.8521454800032098
steps: 201	lr: 0.0001	evaluation reward: 2.02	
episode: 924	score: 3.0	memory length: 174920	epsilon: 0.8516564200032204
steps: 247	lr: 0.0001	evaluation reward: 2.05	
episode: 925	score: 2.0	memory length: 175119	epsilon: 0.8512624000032229
steps: 199	lr: 0.0001	evaluation reward: 2.05	
episode: 926	score: 1.0	memory length: 175270	epsilon: 0.8509634200032354
steps: 151	lr: 0.0001	evaluation reward: 2.05	
episode: 927	score: 0.0	memory length: 175394	epsilon: 0.8507179000032408
steps: 124	lr: 0.0001	evaluation reward: 2.04	
episode: 928	score: 1.0	memory length: 175563	epsilon: 0.850383280003248
steps: 169	lr: 0.0001	evaluation reward: 2.01	
episode: 929	score: 1.0	memory length: 175715	epsilon: 0.8500823200032546
steps: 152	lr: 0.0001	evaluation reward: 2.01	
episode: 930	score: 0.0	memory length: 175838	epsilon: 0.8498387800032599
steps: 123	lr: 0.0001	evaluation reward: 2.01	
episode: 931	score: 1.0	memory length: 175990	epsilon: 0.8495378200032664
steps: 152	lr: 0.0001	evaluation reward: 1.98	
episode: 932	score: 1.0	memory length: 176160	epsilon: 0.8492012200032737
steps: 170	lr: 0.0001	evaluation reward: 1.99	
episode: 933	score: 3.0	memory length: 176405	epsilon: 0.8487161200032842
steps: 245	lr: 0.0001	evaluation reward: 2.02	
episode: 934	score: 5.0	memory length: 176732	epsilon: 0.8480686600032983
steps: 327	lr: 0.0001	evaluation reward: 2.07	
episode: 935	score: 1.0	memory length: 176883	epsilon: 0.8477696800033048
steps: 151	lr: 0.0001	evaluation reward: 2.06	
episode: 936	score: 2.0	memory length: 177082	epsilon: 0.8473756600033133
steps: 199	lr: 0.0001	evaluation reward: 2.07	
episode: 937	score: 1.0	memory length: 177251	epsilon: 0.8470410400033206
steps: 169	lr: 0.0001	evaluation reward: 2.05	
episode: 938	score: 4.0	memory length: 177512	epsilon: 0.8465242600033318
steps: 261	lr: 0.0001	evaluation reward: 2.09	
episode: 939	score: 0.0	memory length: 177636	epsilon: 0.8462787400033371
steps: 124	lr: 0.0001	evaluation reward: 2.07	
episode: 940	score: 4.0	memory length: 177894	epsilon: 0.8457679000033482
steps: 258	lr: 0.0001	evaluation reward: 2.1	
episode: 941	score: 5.0	memory length: 178216	epsilon: 0.8451303400033621
steps: 322	lr: 0.0001	evaluation reward: 2.13	
episode: 942	score: 2.0	memory length: 178433	epsilon: 0.8447006800033714
steps: 217	lr: 0.0001	evaluation reward: 2.1	
episode: 943	score: 3.0	memory length: 178660	epsilon: 0.8442512200033812
steps: 227	lr: 0.0001	evaluation reward: 2.1	
episode: 944	score: 1.0	memory length: 178812	epsilon: 0.8439502600033877
steps: 152	lr: 0.0001	evaluation reward: 2.1	
episode: 945	score: 2.0	memory length: 178995	epsilon: 0.8435879200033956
steps: 183	lr: 0.0001	evaluation reward: 2.12	
episode: 946	score: 5.0	memory length: 179341	epsilon: 0.8429028400034104
steps: 346	lr: 0.0001	evaluation reward: 2.15	

episode: 947	score: 2.0	memory length: 179558	epsilon: 0.8424731800034198
steps: 217	lr: 0.0001	evaluation reward: 2.16	
episode: 948	score: 4.0	memory length: 179840	epsilon: 0.8419148200034319
steps: 282	lr: 0.0001	evaluation reward: 2.2	
episode: 949	score: 2.0	memory length: 180058	epsilon: 0.8414831800034412
steps: 218	lr: 0.0001	evaluation reward: 2.21	
episode: 950	score: 2.0	memory length: 180239	epsilon: 0.841124800003449
steps: 181	lr: 0.0001	evaluation reward: 2.21	
episode: 951	score: 2.0	memory length: 180439	epsilon: 0.8407288000034576
steps: 200	lr: 0.0001	evaluation reward: 2.16	
episode: 952	score: 4.0	memory length: 180715	epsilon: 0.8401823200034695
steps: 276	lr: 0.0001	evaluation reward: 2.17	
episode: 953	score: 1.0	memory length: 180884	epsilon: 0.8398477000034767
steps: 169	lr: 0.0001	evaluation reward: 2.16	
episode: 954	score: 1.0	memory length: 181036	epsilon: 0.8395467400034833
steps: 152	lr: 0.0001	evaluation reward: 2.17	
episode: 955	score: 4.0	memory length: 181349	epsilon: 0.8389270000034967
steps: 313	lr: 0.0001	evaluation reward: 2.2	
episode: 956	score: 2.0	memory length: 181548	epsilon: 0.8385329800035053
steps: 199	lr: 0.0001	evaluation reward: 2.17	
episode: 957	score: 3.0	memory length: 181777	epsilon: 0.8380795600035151
steps: 229	lr: 0.0001	evaluation reward: 2.18	
episode: 958	score: 2.0	memory length: 181975	epsilon: 0.8376875200035236
steps: 198	lr: 0.0001	evaluation reward: 2.18	
episode: 959	score: 2.0	memory length: 182156	epsilon: 0.8373291400035314
steps: 181	lr: 0.0001	evaluation reward: 2.2	
episode: 960	score: 1.0	memory length: 182326	epsilon: 0.8369925400035387
steps: 170	lr: 0.0001	evaluation reward: 2.21	
episode: 961	score: 2.0	memory length: 182526	epsilon: 0.8365965400035473
steps: 200	lr: 0.0001	evaluation reward: 2.22	
episode: 962	score: 1.0	memory length: 182677	epsilon: 0.8362975600035538
steps: 151	lr: 0.0001	evaluation reward: 2.21	
episode: 963	score: 3.0	memory length: 182904	epsilon: 0.8358481000035636
steps: 227	lr: 0.0001	evaluation reward: 2.22	
episode: 964	score: 0.0	memory length: 183027	epsilon: 0.8356045600035689
steps: 123	lr: 0.0001	evaluation reward: 2.2	
episode: 965	score: 1.0	memory length: 183179	epsilon: 0.8353036000035754
steps: 152	lr: 0.0001	evaluation reward: 2.21	
episode: 966	score: 0.0	memory length: 183302	epsilon: 0.8350600600035807
steps: 123	lr: 0.0001	evaluation reward: 2.18	
episode: 967	score: 2.0	memory length: 183485	epsilon: 0.8346977200035886
steps: 183	lr: 0.0001	evaluation reward: 2.16	
episode: 968	score: 2.0	memory length: 183683	epsilon: 0.8343056800035971
steps: 198	lr: 0.0001	evaluation reward: 2.17	
episode: 969	score: 0.0	memory length: 183807	epsilon: 0.8340601600036024
steps: 124	lr: 0.0001	evaluation reward: 2.17	
episode: 970	score: 4.0	memory length: 184066	epsilon: 0.8335473400036135
steps: 259	lr: 0.0001	evaluation reward: 2.19	

episode: 971	score: 1.0	memory length: 184218	epsilon: 0.8332463800036201
steps: 152	lr: 0.0001	evaluation reward: 2.17	
episode: 972	score: 4.0	memory length: 184499	epsilon: 0.8326900000036321
steps: 281	lr: 0.0001	evaluation reward: 2.19	
episode: 973	score: 1.0	memory length: 184671	epsilon: 0.8323494400036395
steps: 172	lr: 0.0001	evaluation reward: 2.18	
episode: 974	score: 2.0	memory length: 184870	epsilon: 0.8319554200036481
steps: 199	lr: 0.0001	evaluation reward: 2.18	
episode: 975	score: 2.0	memory length: 185071	epsilon: 0.8315574400036567
steps: 201	lr: 0.0001	evaluation reward: 2.13	
episode: 976	score: 2.0	memory length: 185291	epsilon: 0.8311218400036662
steps: 220	lr: 0.0001	evaluation reward: 2.15	
episode: 977	score: 0.0	memory length: 185415	epsilon: 0.8308763200036715
steps: 124	lr: 0.0001	evaluation reward: 2.15	
episode: 978	score: 4.0	memory length: 185730	epsilon: 0.830252620003685
steps: 315	lr: 0.0001	evaluation reward: 2.17	
episode: 979	score: 2.0	memory length: 185928	epsilon: 0.8298605800036936
steps: 198	lr: 0.0001	evaluation reward: 2.15	
episode: 980	score: 3.0	memory length: 186173	epsilon: 0.8293754800037041
steps: 245	lr: 0.0001	evaluation reward: 2.17	
episode: 981	score: 3.0	memory length: 186423	epsilon: 0.8288804800037148
steps: 250	lr: 0.0001	evaluation reward: 2.17	
episode: 982	score: 1.0	memory length: 186594	epsilon: 0.8285419000037222
steps: 171	lr: 0.0001	evaluation reward: 2.15	
episode: 983	score: 3.0	memory length: 186830	epsilon: 0.8280746200037323
steps: 236	lr: 0.0001	evaluation reward: 2.18	
episode: 984	score: 3.0	memory length: 187074	epsilon: 0.8275915000037428
steps: 244	lr: 0.0001	evaluation reward: 2.2	
episode: 985	score: 1.0	memory length: 187225	epsilon: 0.8272925200037493
steps: 151	lr: 0.0001	evaluation reward: 2.19	
episode: 986	score: 5.0	memory length: 187549	epsilon: 0.8266510000037632
steps: 324	lr: 0.0001	evaluation reward: 2.19	
episode: 987	score: 2.0	memory length: 187747	epsilon: 0.8262589600037717
steps: 198	lr: 0.0001	evaluation reward: 2.19	
episode: 988	score: 2.0	memory length: 187946	epsilon: 0.8258649400037803
steps: 199	lr: 0.0001	evaluation reward: 2.16	
episode: 989	score: 1.0	memory length: 188118	epsilon: 0.8255243800037877
steps: 172	lr: 0.0001	evaluation reward: 2.16	
episode: 990	score: 3.0	memory length: 188364	epsilon: 0.8250373000037983
steps: 246	lr: 0.0001	evaluation reward: 2.16	
episode: 991	score: 2.0	memory length: 188563	epsilon: 0.8246432800038068
steps: 199	lr: 0.0001	evaluation reward: 2.14	
episode: 992	score: 3.0	memory length: 188809	epsilon: 0.8241562000038174
steps: 246	lr: 0.0001	evaluation reward: 2.15	
episode: 993	score: 3.0	memory length: 189056	epsilon: 0.823667140003828
steps: 247	lr: 0.0001	evaluation reward: 2.14	
episode: 994	score: 3.0	memory length: 189328	epsilon: 0.8231285800038397
steps: 272	lr: 0.0001	evaluation reward: 2.13	

episode: 995	score: 3.0	memory length: 189579	epsilon: 0.8226316000038505
steps: 251	lr: 0.0001	evaluation reward: 2.14	
episode: 996	score: 4.0	memory length: 189853	epsilon: 0.8220890800038623
steps: 274	lr: 0.0001	evaluation reward: 2.14	
episode: 997	score: 3.0	memory length: 190086	epsilon: 0.8216277400038723
steps: 233	lr: 0.0001	evaluation reward: 2.14	
episode: 998	score: 3.0	memory length: 190316	epsilon: 0.8211723400038822
steps: 230	lr: 0.0001	evaluation reward: 2.17	
episode: 999	score: 4.0	memory length: 190611	epsilon: 0.8205882400038949
steps: 295	lr: 0.0001	evaluation reward: 2.17	
episode: 1000	score: 5.0	memory length: 190959	epsilon: 0.8198992000039098
steps: 348	lr: 0.0001	evaluation reward: 2.21	
episode: 1001	score: 5.0	memory length: 191306	epsilon: 0.8192121400039247
steps: 347	lr: 0.0001	evaluation reward: 2.25	
episode: 1002	score: 5.0	memory length: 191597	epsilon: 0.8186359600039372
steps: 291	lr: 0.0001	evaluation reward: 2.29	
episode: 1003	score: 2.0	memory length: 191800	epsilon: 0.818234020003946
steps: 203	lr: 0.0001	evaluation reward: 2.3	
episode: 1004	score: 1.0	memory length: 191952	epsilon: 0.8179330600039525
steps: 152	lr: 0.0001	evaluation reward: 2.28	
episode: 1005	score: 3.0	memory length: 192178	epsilon: 0.8174855800039622
steps: 226	lr: 0.0001	evaluation reward: 2.29	
episode: 1006	score: 1.0	memory length: 192350	epsilon: 0.8171450200039696
steps: 172	lr: 0.0001	evaluation reward: 2.28	
episode: 1007	score: 2.0	memory length: 192533	epsilon: 0.8167826800039775
steps: 183	lr: 0.0001	evaluation reward: 2.29	
episode: 1008	score: 0.0	memory length: 192657	epsilon: 0.8165371600039828
steps: 124	lr: 0.0001	evaluation reward: 2.29	
episode: 1009	score: 2.0	memory length: 192858	epsilon: 0.8161391800039914
steps: 201	lr: 0.0001	evaluation reward: 2.3	
episode: 1010	score: 1.0	memory length: 193030	epsilon: 0.8157986200039988
steps: 172	lr: 0.0001	evaluation reward: 2.29	
episode: 1011	score: 3.0	memory length: 193296	epsilon: 0.8152719400040103
steps: 266	lr: 0.0001	evaluation reward: 2.29	
episode: 1012	score: 2.0	memory length: 193496	epsilon: 0.8148759400040189
steps: 200	lr: 0.0001	evaluation reward: 2.29	
episode: 1013	score: 2.0	memory length: 193697	epsilon: 0.8144779600040275
steps: 201	lr: 0.0001	evaluation reward: 2.25	
episode: 1014	score: 4.0	memory length: 193971	epsilon: 0.8139354400040393
steps: 274	lr: 0.0001	evaluation reward: 2.27	
episode: 1015	score: 2.0	memory length: 194170	epsilon: 0.8135414200040478
steps: 199	lr: 0.0001	evaluation reward: 2.26	
episode: 1016	score: 0.0	memory length: 194293	epsilon: 0.8132978800040531
steps: 123	lr: 0.0001	evaluation reward: 2.22	
episode: 1017	score: 2.0	memory length: 194475	epsilon: 0.8129375200040609
steps: 182	lr: 0.0001	evaluation reward: 2.23	
episode: 1018	score: 1.0	memory length: 194646	epsilon: 0.8125989400040683
steps: 171	lr: 0.0001	evaluation reward: 2.23	

episode: 1019	score: 0.0	memory length: 194770	epsilon: 0.8123534200040736
steps: 124	lr: 0.0001	evaluation reward: 2.21	
episode: 1020	score: 2.0	memory length: 194969	epsilon: 0.8119594000040822
steps: 199	lr: 0.0001	evaluation reward: 2.21	
episode: 1021	score: 2.0	memory length: 195190	epsilon: 0.8115218200040917
steps: 221	lr: 0.0001	evaluation reward: 2.18	
episode: 1022	score: 0.0	memory length: 195314	epsilon: 0.811276300004097
steps: 124	lr: 0.0001	evaluation reward: 2.17	
episode: 1023	score: 0.0	memory length: 195438	epsilon: 0.8110307800041023
steps: 124	lr: 0.0001	evaluation reward: 2.15	
episode: 1024	score: 2.0	memory length: 195621	epsilon: 0.8106684400041102
steps: 183	lr: 0.0001	evaluation reward: 2.14	
episode: 1025	score: 3.0	memory length: 195866	epsilon: 0.8101833400041207
steps: 245	lr: 0.0001	evaluation reward: 2.15	
episode: 1026	score: 2.0	memory length: 196070	epsilon: 0.8097794200041295
steps: 204	lr: 0.0001	evaluation reward: 2.16	
episode: 1027	score: 5.0	memory length: 196377	epsilon: 0.8091715600041427
steps: 307	lr: 0.0001	evaluation reward: 2.21	
episode: 1028	score: 2.0	memory length: 196558	epsilon: 0.8088131800041505
steps: 181	lr: 0.0001	evaluation reward: 2.22	
episode: 1029	score: 2.0	memory length: 196776	epsilon: 0.8083815400041598
steps: 218	lr: 0.0001	evaluation reward: 2.23	
episode: 1030	score: 0.0	memory length: 196900	epsilon: 0.8081360200041652
steps: 124	lr: 0.0001	evaluation reward: 2.23	
episode: 1031	score: 0.0	memory length: 197024	epsilon: 0.8078905000041705
steps: 124	lr: 0.0001	evaluation reward: 2.22	
episode: 1032	score: 2.0	memory length: 197244	epsilon: 0.80745490000418
steps: 220	lr: 0.0001	evaluation reward: 2.23	
episode: 1033	score: 2.0	memory length: 197445	epsilon: 0.8070569200041886
steps: 201	lr: 0.0001	evaluation reward: 2.22	
episode: 1034	score: 4.0	memory length: 197721	epsilon: 0.8065104400042005
steps: 276	lr: 0.0001	evaluation reward: 2.21	
episode: 1035	score: 4.0	memory length: 197996	epsilon: 0.8059659400042123
steps: 275	lr: 0.0001	evaluation reward: 2.24	
episode: 1036	score: 3.0	memory length: 198226	epsilon: 0.8055105400042222
steps: 230	lr: 0.0001	evaluation reward: 2.25	
episode: 1037	score: 2.0	memory length: 198408	epsilon: 0.80515018000423
steps: 182	lr: 0.0001	evaluation reward: 2.26	
episode: 1038	score: 3.0	memory length: 198637	epsilon: 0.8046967600042398
steps: 229	lr: 0.0001	evaluation reward: 2.25	
episode: 1039	score: 2.0	memory length: 198838	epsilon: 0.8042987800042485
steps: 201	lr: 0.0001	evaluation reward: 2.27	
episode: 1040	score: 2.0	memory length: 199018	epsilon: 0.8039423800042562
steps: 180	lr: 0.0001	evaluation reward: 2.25	
episode: 1041	score: 0.0	memory length: 199141	epsilon: 0.8036988400042615
steps: 123	lr: 0.0001	evaluation reward: 2.2	
episode: 1042	score: 3.0	memory length: 199367	epsilon: 0.8032513600042712
steps: 226	lr: 0.0001	evaluation reward: 2.21	

episode: 1043	score: 0.0	memory length: 199490	epsilon: 0.8030078200042765
steps: 123	lr: 0.0001	evaluation reward: 2.18	
episode: 1044	score: 2.0	memory length: 199689	epsilon: 0.8026138000042851
steps: 199	lr: 0.0001	evaluation reward: 2.19	
episode: 1045	score: 3.0	memory length: 199934	epsilon: 0.8021287000042956
steps: 245	lr: 0.0001	evaluation reward: 2.2	
episode: 1046	score: 1.0	memory length: 200106	epsilon: 0.801788140004303
steps: 172	lr: 4e-05	evaluation reward: 2.16	
episode: 1047	score: 1.0	memory length: 200257	epsilon: 0.8014891600043095
steps: 151	lr: 4e-05	evaluation reward: 2.15	
episode: 1048	score: 3.0	memory length: 200522	epsilon: 0.8009644600043209
steps: 265	lr: 4e-05	evaluation reward: 2.14	
episode: 1049	score: 2.0	memory length: 200721	epsilon: 0.8005704400043294
steps: 199	lr: 4e-05	evaluation reward: 2.14	
episode: 1050	score: 1.0	memory length: 200873	epsilon: 0.800269480004336
steps: 152	lr: 4e-05	evaluation reward: 2.13	
episode: 1051	score: 1.0	memory length: 201025	epsilon: 0.7999685200043425
steps: 152	lr: 4e-05	evaluation reward: 2.12	
episode: 1052	score: 5.0	memory length: 201323	epsilon: 0.7993784800043553
steps: 298	lr: 4e-05	evaluation reward: 2.13	
episode: 1053	score: 1.0	memory length: 201493	epsilon: 0.7990418800043626
steps: 170	lr: 4e-05	evaluation reward: 2.13	
episode: 1054	score: 2.0	memory length: 201694	epsilon: 0.7986439000043712
steps: 201	lr: 4e-05	evaluation reward: 2.14	
episode: 1055	score: 0.0	memory length: 201818	epsilon: 0.7983983800043766
steps: 124	lr: 4e-05	evaluation reward: 2.1	
episode: 1056	score: 3.0	memory length: 202065	epsilon: 0.7979093200043872
steps: 247	lr: 4e-05	evaluation reward: 2.11	
episode: 1057	score: 1.0	memory length: 202217	epsilon: 0.7976083600043937
steps: 152	lr: 4e-05	evaluation reward: 2.09	
episode: 1058	score: 5.0	memory length: 202528	epsilon: 0.7969925800044071
steps: 311	lr: 4e-05	evaluation reward: 2.12	
episode: 1059	score: 4.0	memory length: 202804	epsilon: 0.796446100004419
steps: 276	lr: 4e-05	evaluation reward: 2.14	
episode: 1060	score: 2.0	memory length: 203004	epsilon: 0.7960501000044276
steps: 200	lr: 4e-05	evaluation reward: 2.15	
episode: 1061	score: 2.0	memory length: 203203	epsilon: 0.7956560800044361
steps: 199	lr: 4e-05	evaluation reward: 2.15	
episode: 1062	score: 4.0	memory length: 203497	epsilon: 0.7950739600044487
steps: 294	lr: 4e-05	evaluation reward: 2.18	
episode: 1063	score: 3.0	memory length: 203747	epsilon: 0.7945789600044595
steps: 250	lr: 4e-05	evaluation reward: 2.18	
episode: 1064	score: 0.0	memory length: 203871	epsilon: 0.7943334400044648
steps: 124	lr: 4e-05	evaluation reward: 2.18	
episode: 1065	score: 2.0	memory length: 204092	epsilon: 0.7938958600044743
steps: 221	lr: 4e-05	evaluation reward: 2.19	
episode: 1066	score: 1.0	memory length: 204244	epsilon: 0.7935949000044809
steps: 152	lr: 4e-05	evaluation reward: 2.2	

episode: 1067	score: 4.0	memory length: 204512	epsilon: 0.7930642600044924
steps: 268	lr: 4e-05	evaluation reward: 2.22	
episode: 1068	score: 1.0	memory length: 204681	epsilon: 0.7927296400044996
steps: 169	lr: 4e-05	evaluation reward: 2.21	
episode: 1069	score: 4.0	memory length: 204960	epsilon: 0.7921772200045116
steps: 279	lr: 4e-05	evaluation reward: 2.25	
episode: 1070	score: 4.0	memory length: 205241	epsilon: 0.7916208400045237
steps: 281	lr: 4e-05	evaluation reward: 2.25	
episode: 1071	score: 4.0	memory length: 205537	epsilon: 0.7910347600045364
steps: 296	lr: 4e-05	evaluation reward: 2.28	
episode: 1072	score: 7.0	memory length: 205939	epsilon: 0.7902388000045537
steps: 402	lr: 4e-05	evaluation reward: 2.31	
episode: 1073	score: 5.0	memory length: 206267	epsilon: 0.7895893600045678
steps: 328	lr: 4e-05	evaluation reward: 2.35	
episode: 1074	score: 1.0	memory length: 206436	epsilon: 0.7892547400045751
steps: 169	lr: 4e-05	evaluation reward: 2.34	
episode: 1075	score: 1.0	memory length: 206589	epsilon: 0.7889518000045816
steps: 153	lr: 4e-05	evaluation reward: 2.33	
episode: 1076	score: 3.0	memory length: 206841	epsilon: 0.7884528400045925
steps: 252	lr: 4e-05	evaluation reward: 2.34	
episode: 1077	score: 2.0	memory length: 207041	epsilon: 0.7880568400046011
steps: 200	lr: 4e-05	evaluation reward: 2.36	
episode: 1078	score: 5.0	memory length: 207364	epsilon: 0.787417300004615
steps: 323	lr: 4e-05	evaluation reward: 2.37	
episode: 1079	score: 3.0	memory length: 207612	epsilon: 0.7869262600046256
steps: 248	lr: 4e-05	evaluation reward: 2.38	
episode: 1080	score: 4.0	memory length: 207903	epsilon: 0.7863500800046381
steps: 291	lr: 4e-05	evaluation reward: 2.39	
episode: 1081	score: 0.0	memory length: 208027	epsilon: 0.7861045600046435
steps: 124	lr: 4e-05	evaluation reward: 2.36	
episode: 1082	score: 1.0	memory length: 208179	epsilon: 0.78580360000465
steps: 152	lr: 4e-05	evaluation reward: 2.36	
episode: 1083	score: 2.0	memory length: 208361	epsilon: 0.7854432400046578
steps: 182	lr: 4e-05	evaluation reward: 2.35	
episode: 1084	score: 3.0	memory length: 208608	epsilon: 0.7849541800046684
steps: 247	lr: 4e-05	evaluation reward: 2.35	
episode: 1085	score: 1.0	memory length: 208761	epsilon: 0.784651240004675
steps: 153	lr: 4e-05	evaluation reward: 2.35	
episode: 1086	score: 1.0	memory length: 208913	epsilon: 0.7843502800046815
steps: 152	lr: 4e-05	evaluation reward: 2.31	
episode: 1087	score: 2.0	memory length: 209115	epsilon: 0.7839503200046902
steps: 202	lr: 4e-05	evaluation reward: 2.31	
episode: 1088	score: 3.0	memory length: 209361	epsilon: 0.7834632400047008
steps: 246	lr: 4e-05	evaluation reward: 2.32	
episode: 1089	score: 0.0	memory length: 209485	epsilon: 0.7832177200047061
steps: 124	lr: 4e-05	evaluation reward: 2.31	
episode: 1090	score: 0.0	memory length: 209609	epsilon: 0.7829722000047115
steps: 124	lr: 4e-05	evaluation reward: 2.28	

episode: 1091	score: 4.0	memory length: 209905	epsilon: 0.7823861200047242
steps: 296	lr: 4e-05	evaluation reward: 2.3	
episode: 1092	score: 4.0	memory length: 210182	epsilon: 0.7818376600047361
steps: 277	lr: 4e-05	evaluation reward: 2.31	
episode: 1093	score: 2.0	memory length: 210381	epsilon: 0.7814436400047446
steps: 199	lr: 4e-05	evaluation reward: 2.3	
episode: 1094	score: 1.0	memory length: 210532	epsilon: 0.7811446600047511
steps: 151	lr: 4e-05	evaluation reward: 2.28	
episode: 1095	score: 2.0	memory length: 210731	epsilon: 0.7807506400047597
steps: 199	lr: 4e-05	evaluation reward: 2.27	
episode: 1096	score: 4.0	memory length: 211007	epsilon: 0.7802041600047716
steps: 276	lr: 4e-05	evaluation reward: 2.27	
episode: 1097	score: 2.0	memory length: 211211	epsilon: 0.7798002400047803
steps: 204	lr: 4e-05	evaluation reward: 2.26	
episode: 1098	score: 4.0	memory length: 211508	epsilon: 0.7792121800047931
steps: 297	lr: 4e-05	evaluation reward: 2.27	
episode: 1099	score: 4.0	memory length: 211803	epsilon: 0.7786280800048058
steps: 295	lr: 4e-05	evaluation reward: 2.27	
episode: 1100	score: 3.0	memory length: 212070	epsilon: 0.7780994200048172
steps: 267	lr: 4e-05	evaluation reward: 2.25	
episode: 1101	score: 4.0	memory length: 212351	epsilon: 0.7775430400048293
steps: 281	lr: 4e-05	evaluation reward: 2.24	
episode: 1102	score: 1.0	memory length: 212521	epsilon: 0.7772064400048366
steps: 170	lr: 4e-05	evaluation reward: 2.2	
episode: 1103	score: 1.0	memory length: 212672	epsilon: 0.7769074600048431
steps: 151	lr: 4e-05	evaluation reward: 2.19	
episode: 1104	score: 1.0	memory length: 212824	epsilon: 0.7766065000048497
steps: 152	lr: 4e-05	evaluation reward: 2.19	
episode: 1105	score: 3.0	memory length: 213038	epsilon: 0.7761827800048589
steps: 214	lr: 4e-05	evaluation reward: 2.19	
episode: 1106	score: 2.0	memory length: 213237	epsilon: 0.7757887600048674
steps: 199	lr: 4e-05	evaluation reward: 2.2	
episode: 1107	score: 2.0	memory length: 213457	epsilon: 0.7753531600048769
steps: 220	lr: 4e-05	evaluation reward: 2.2	
episode: 1108	score: 1.0	memory length: 213609	epsilon: 0.7750522000048834
steps: 152	lr: 4e-05	evaluation reward: 2.21	
episode: 1109	score: 2.0	memory length: 213825	epsilon: 0.7746245200048927
steps: 216	lr: 4e-05	evaluation reward: 2.21	
episode: 1110	score: 3.0	memory length: 214039	epsilon: 0.7742008000049019
steps: 214	lr: 4e-05	evaluation reward: 2.23	
episode: 1111	score: 10.0	memory length: 214452	epsilon:
0.7733830600049196	steps: 413	lr: 4e-05	evaluation reward: 2.3
episode: 1112	score: 3.0	memory length: 214685	epsilon: 0.7729217200049296
steps: 233	lr: 4e-05	evaluation reward: 2.31	
episode: 1113	score: 8.0	memory length: 215121	epsilon: 0.7720584400049484
steps: 436	lr: 4e-05	evaluation reward: 2.37	
episode: 1114	score: 2.0	memory length: 215322	epsilon: 0.771660460004957
steps: 201	lr: 4e-05	evaluation reward: 2.35	

episode: 1115	score: 1.0	memory length: 215492	epsilon: 0.7713238600049643
steps: 170	lr: 4e-05	evaluation reward: 2.34	
episode: 1116	score: 3.0	memory length: 215740	epsilon: 0.770832820004975
steps: 248	lr: 4e-05	evaluation reward: 2.37	
episode: 1117	score: 4.0	memory length: 216035	epsilon: 0.7702487200049877
steps: 295	lr: 4e-05	evaluation reward: 2.39	
episode: 1118	score: 1.0	memory length: 216187	epsilon: 0.7699477600049942
steps: 152	lr: 4e-05	evaluation reward: 2.39	
episode: 1119	score: 2.0	memory length: 216404	epsilon: 0.7695181000050035
steps: 217	lr: 4e-05	evaluation reward: 2.41	
episode: 1120	score: 4.0	memory length: 216681	epsilon: 0.7689696400050154
steps: 277	lr: 4e-05	evaluation reward: 2.43	
episode: 1121	score: 1.0	memory length: 216832	epsilon: 0.7686706600050219
steps: 151	lr: 4e-05	evaluation reward: 2.42	
episode: 1122	score: 2.0	memory length: 217031	epsilon: 0.7682766400050305
steps: 199	lr: 4e-05	evaluation reward: 2.44	
episode: 1123	score: 2.0	memory length: 217230	epsilon: 0.767882620005039
steps: 199	lr: 4e-05	evaluation reward: 2.46	
episode: 1124	score: 5.0	memory length: 217576	epsilon: 0.7671975400050539
steps: 346	lr: 4e-05	evaluation reward: 2.49	
episode: 1125	score: 6.0	memory length: 217953	epsilon: 0.7664510800050701
steps: 377	lr: 4e-05	evaluation reward: 2.52	
episode: 1126	score: 2.0	memory length: 218155	epsilon: 0.7660511200050788
steps: 202	lr: 4e-05	evaluation reward: 2.52	
episode: 1127	score: 7.0	memory length: 218545	epsilon: 0.7652789200050956
steps: 390	lr: 4e-05	evaluation reward: 2.54	
episode: 1128	score: 3.0	memory length: 218793	epsilon: 0.7647878800051062
steps: 248	lr: 4e-05	evaluation reward: 2.55	
episode: 1129	score: 3.0	memory length: 219005	epsilon: 0.7643681200051153
steps: 212	lr: 4e-05	evaluation reward: 2.56	
episode: 1130	score: 2.0	memory length: 219208	epsilon: 0.7639661800051241
steps: 203	lr: 4e-05	evaluation reward: 2.58	
episode: 1131	score: 2.0	memory length: 219406	epsilon: 0.7635741400051326
steps: 198	lr: 4e-05	evaluation reward: 2.6	
episode: 1132	score: 3.0	memory length: 219657	epsilon: 0.7630771600051434
steps: 251	lr: 4e-05	evaluation reward: 2.61	
episode: 1133	score: 4.0	memory length: 219952	epsilon: 0.762493060005156
steps: 295	lr: 4e-05	evaluation reward: 2.63	
episode: 1134	score: 2.0	memory length: 220171	epsilon: 0.7620594400051655
steps: 219	lr: 4e-05	evaluation reward: 2.61	
episode: 1135	score: 3.0	memory length: 220440	epsilon: 0.761526820005177
steps: 269	lr: 4e-05	evaluation reward: 2.6	
episode: 1136	score: 1.0	memory length: 220611	epsilon: 0.7611882400051844
steps: 171	lr: 4e-05	evaluation reward: 2.58	
episode: 1137	score: 2.0	memory length: 220792	epsilon: 0.7608298600051921
steps: 181	lr: 4e-05	evaluation reward: 2.58	
episode: 1138	score: 4.0	memory length: 221068	epsilon: 0.760283380005204
steps: 276	lr: 4e-05	evaluation reward: 2.59	

episode: 1139	score: 1.0	memory length: 221241	epsilon: 0.7599408400052114
steps: 173	lr: 4e-05	evaluation reward: 2.58	
episode: 1140	score: 4.0	memory length: 221500	epsilon: 0.7594280200052226
steps: 259	lr: 4e-05	evaluation reward: 2.6	
episode: 1141	score: 4.0	memory length: 221761	epsilon: 0.7589112400052338
steps: 261	lr: 4e-05	evaluation reward: 2.64	
episode: 1142	score: 6.0	memory length: 222124	epsilon: 0.7581925000052494
steps: 363	lr: 4e-05	evaluation reward: 2.67	
episode: 1143	score: 3.0	memory length: 222353	epsilon: 0.7577390800052592
steps: 229	lr: 4e-05	evaluation reward: 2.7	
episode: 1144	score: 3.0	memory length: 222584	epsilon: 0.7572817000052692
steps: 231	lr: 4e-05	evaluation reward: 2.71	
episode: 1145	score: 2.0	memory length: 222803	epsilon: 0.7568480800052786
steps: 219	lr: 4e-05	evaluation reward: 2.7	
episode: 1146	score: 3.0	memory length: 223049	epsilon: 0.7563610000052892
steps: 246	lr: 4e-05	evaluation reward: 2.72	
episode: 1147	score: 3.0	memory length: 223317	epsilon: 0.7558303600053007
steps: 268	lr: 4e-05	evaluation reward: 2.74	
episode: 1148	score: 7.0	memory length: 223691	epsilon: 0.7550898400053168
steps: 374	lr: 4e-05	evaluation reward: 2.78	
episode: 1149	score: 3.0	memory length: 223921	epsilon: 0.7546344400053266
steps: 230	lr: 4e-05	evaluation reward: 2.79	
episode: 1150	score: 2.0	memory length: 224140	epsilon: 0.7542008200053361
steps: 219	lr: 4e-05	evaluation reward: 2.8	
episode: 1151	score: 1.0	memory length: 224312	epsilon: 0.7538602600053435
steps: 172	lr: 4e-05	evaluation reward: 2.8	
episode: 1152	score: 4.0	memory length: 224607	epsilon: 0.7532761600053561
steps: 295	lr: 4e-05	evaluation reward: 2.79	
episode: 1153	score: 0.0	memory length: 224731	epsilon: 0.7530306400053615
steps: 124	lr: 4e-05	evaluation reward: 2.78	
episode: 1154	score: 4.0	memory length: 225011	epsilon: 0.7524762400053735
steps: 280	lr: 4e-05	evaluation reward: 2.8	
episode: 1155	score: 3.0	memory length: 225256	epsilon: 0.751991140005384
steps: 245	lr: 4e-05	evaluation reward: 2.83	
episode: 1156	score: 3.0	memory length: 225521	epsilon: 0.7514664400053954
steps: 265	lr: 4e-05	evaluation reward: 2.83	
episode: 1157	score: 3.0	memory length: 225751	epsilon: 0.7510110400054053
steps: 230	lr: 4e-05	evaluation reward: 2.85	
episode: 1158	score: 2.0	memory length: 225934	epsilon: 0.7506487000054132
steps: 183	lr: 4e-05	evaluation reward: 2.82	
episode: 1159	score: 3.0	memory length: 226165	epsilon: 0.7501913200054231
steps: 231	lr: 4e-05	evaluation reward: 2.81	
episode: 1160	score: 5.0	memory length: 226512	epsilon: 0.749504260005438
steps: 347	lr: 4e-05	evaluation reward: 2.84	
episode: 1161	score: 1.0	memory length: 226684	epsilon: 0.7491637000054454
steps: 172	lr: 4e-05	evaluation reward: 2.83	
episode: 1162	score: 1.0	memory length: 226853	epsilon: 0.7488290800054527
steps: 169	lr: 4e-05	evaluation reward: 2.8	

episode: 1163	score: 1.0	memory length: 227005	epsilon: 0.7485281200054592
steps: 152	lr: 4e-05	evaluation reward: 2.78	
episode: 1164	score: 7.0	memory length: 227422	epsilon: 0.7477024600054771
steps: 417	lr: 4e-05	evaluation reward: 2.85	
episode: 1165	score: 3.0	memory length: 227657	epsilon: 0.7472371600054872
steps: 235	lr: 4e-05	evaluation reward: 2.86	
episode: 1166	score: 3.0	memory length: 227905	epsilon: 0.7467461200054979
steps: 248	lr: 4e-05	evaluation reward: 2.88	
episode: 1167	score: 3.0	memory length: 228117	epsilon: 0.7463263600055507
steps: 212	lr: 4e-05	evaluation reward: 2.87	
episode: 1168	score: 3.0	memory length: 228345	epsilon: 0.7458749200055168
steps: 228	lr: 4e-05	evaluation reward: 2.89	
episode: 1169	score: 4.0	memory length: 228605	epsilon: 0.7453601200055528
steps: 260	lr: 4e-05	evaluation reward: 2.89	
episode: 1170	score: 4.0	memory length: 228863	epsilon: 0.7448492800055391
steps: 258	lr: 4e-05	evaluation reward: 2.89	
episode: 1171	score: 7.0	memory length: 229251	epsilon: 0.7440810400055557
steps: 388	lr: 4e-05	evaluation reward: 2.92	
episode: 1172	score: 2.0	memory length: 229434	epsilon: 0.7437187000055636
steps: 183	lr: 4e-05	evaluation reward: 2.87	
episode: 1173	score: 6.0	memory length: 229848	epsilon: 0.7428989800055814
steps: 414	lr: 4e-05	evaluation reward: 2.88	
episode: 1174	score: 2.0	memory length: 230047	epsilon: 0.742504960005559
steps: 199	lr: 4e-05	evaluation reward: 2.89	
episode: 1175	score: 3.0	memory length: 230258	epsilon: 0.7420871800055599
steps: 211	lr: 4e-05	evaluation reward: 2.91	
episode: 1176	score: 5.0	memory length: 230581	epsilon: 0.7414476400056129
steps: 323	lr: 4e-05	evaluation reward: 2.93	
episode: 1177	score: 4.0	memory length: 230875	epsilon: 0.7408655200056256
steps: 294	lr: 4e-05	evaluation reward: 2.95	
episode: 1178	score: 2.0	memory length: 231056	epsilon: 0.7405071400056333
steps: 181	lr: 4e-05	evaluation reward: 2.92	
episode: 1179	score: 5.0	memory length: 231400	epsilon: 0.7398260200056481
steps: 344	lr: 4e-05	evaluation reward: 2.94	
episode: 1180	score: 5.0	memory length: 231748	epsilon: 0.7391369800056631
steps: 348	lr: 4e-05	evaluation reward: 2.95	
episode: 1181	score: 5.0	memory length: 232091	epsilon: 0.7384578400056778
steps: 343	lr: 4e-05	evaluation reward: 3.0	
episode: 1182	score: 2.0	memory length: 232311	epsilon: 0.7380222400056873
steps: 220	lr: 4e-05	evaluation reward: 3.01	
episode: 1183	score: 3.0	memory length: 232559	epsilon: 0.7375312000056979
steps: 248	lr: 4e-05	evaluation reward: 3.02	
episode: 1184	score: 4.0	memory length: 232835	epsilon: 0.7369847200057098
steps: 276	lr: 4e-05	evaluation reward: 3.03	
episode: 1185	score: 1.0	memory length: 232987	epsilon: 0.7366837600057163
steps: 152	lr: 4e-05	evaluation reward: 3.03	
episode: 1186	score: 5.0	memory length: 233283	epsilon: 0.7360976800057291
steps: 296	lr: 4e-05	evaluation reward: 3.07	

episode: 1187	score: 3.0	memory length: 233531	epsilon: 0.7356066400057397
steps: 248	lr: 4e-05	evaluation reward: 3.08	
episode: 1188	score: 2.0	memory length: 233711	epsilon: 0.7352502400057475
steps: 180	lr: 4e-05	evaluation reward: 3.07	
episode: 1189	score: 4.0	memory length: 234010	epsilon: 0.7346582200057603
steps: 299	lr: 4e-05	evaluation reward: 3.11	
episode: 1190	score: 2.0	memory length: 234229	epsilon: 0.7342246000057697
steps: 219	lr: 4e-05	evaluation reward: 3.13	
episode: 1191	score: 2.0	memory length: 234430	epsilon: 0.7338266200057784
steps: 201	lr: 4e-05	evaluation reward: 3.11	
episode: 1192	score: 5.0	memory length: 234741	epsilon: 0.7332108400057917
steps: 311	lr: 4e-05	evaluation reward: 3.12	
episode: 1193	score: 3.0	memory length: 234975	epsilon: 0.7327475200058018
steps: 234	lr: 4e-05	evaluation reward: 3.13	
episode: 1194	score: 2.0	memory length: 235157	epsilon: 0.7323871600058096
steps: 182	lr: 4e-05	evaluation reward: 3.14	
episode: 1195	score: 3.0	memory length: 235404	epsilon: 0.7318981000058202
steps: 247	lr: 4e-05	evaluation reward: 3.15	
episode: 1196	score: 2.0	memory length: 235605	epsilon: 0.7315001200058289
steps: 201	lr: 4e-05	evaluation reward: 3.13	
episode: 1197	score: 0.0	memory length: 235729	epsilon: 0.7312546000058342
steps: 124	lr: 4e-05	evaluation reward: 3.11	
episode: 1198	score: 0.0	memory length: 235853	epsilon: 0.7310090800058395
steps: 124	lr: 4e-05	evaluation reward: 3.07	
episode: 1199	score: 1.0	memory length: 236005	epsilon: 0.7307081200058461
steps: 152	lr: 4e-05	evaluation reward: 3.04	
episode: 1200	score: 2.0	memory length: 236204	epsilon: 0.7303141000058546
steps: 199	lr: 4e-05	evaluation reward: 3.03	
episode: 1201	score: 5.0	memory length: 236549	epsilon: 0.7296310000058694
steps: 345	lr: 4e-05	evaluation reward: 3.04	
episode: 1202	score: 4.0	memory length: 236841	epsilon: 0.729052840005882
steps: 292	lr: 4e-05	evaluation reward: 3.07	
episode: 1203	score: 1.0	memory length: 237010	epsilon: 0.7287182200058893
steps: 169	lr: 4e-05	evaluation reward: 3.07	
episode: 1204	score: 6.0	memory length: 237317	epsilon: 0.7281103600059025
steps: 307	lr: 4e-05	evaluation reward: 3.12	
episode: 1205	score: 2.0	memory length: 237516	epsilon: 0.727716340005911
steps: 199	lr: 4e-05	evaluation reward: 3.11	
episode: 1206	score: 5.0	memory length: 237803	epsilon: 0.7271480800059233
steps: 287	lr: 4e-05	evaluation reward: 3.14	
episode: 1207	score: 3.0	memory length: 238030	epsilon: 0.7266986200059331
steps: 227	lr: 4e-05	evaluation reward: 3.15	
episode: 1208	score: 2.0	memory length: 238228	epsilon: 0.7263065800059416
steps: 198	lr: 4e-05	evaluation reward: 3.16	
episode: 1209	score: 8.0	memory length: 238663	epsilon: 0.7254452800059603
steps: 435	lr: 4e-05	evaluation reward: 3.22	
episode: 1210	score: 7.0	memory length: 239024	epsilon: 0.7247305000059758
steps: 361	lr: 4e-05	evaluation reward: 3.26	

episode: 1211	score: 2.0	memory length: 239242	epsilon: 0.7242988600059852
steps: 218	lr: 4e-05	evaluation reward: 3.18	
episode: 1212	score: 3.0	memory length: 239469	epsilon: 0.723849400005995
steps: 227	lr: 4e-05	evaluation reward: 3.18	
episode: 1213	score: 3.0	memory length: 239714	epsilon: 0.7233643000060055
steps: 245	lr: 4e-05	evaluation reward: 3.13	
episode: 1214	score: 2.0	memory length: 239930	epsilon: 0.7229366200060148
steps: 216	lr: 4e-05	evaluation reward: 3.13	
episode: 1215	score: 0.0	memory length: 240054	epsilon: 0.7226911000060201
steps: 124	lr: 4e-05	evaluation reward: 3.12	
episode: 1216	score: 3.0	memory length: 240300	epsilon: 0.7222040200060307
steps: 246	lr: 4e-05	evaluation reward: 3.12	
episode: 1217	score: 4.0	memory length: 240543	epsilon: 0.7217228800060411
steps: 243	lr: 4e-05	evaluation reward: 3.12	
episode: 1218	score: 5.0	memory length: 240894	epsilon: 0.7210279000060562
steps: 351	lr: 4e-05	evaluation reward: 3.16	
episode: 1219	score: 4.0	memory length: 241172	epsilon: 0.7204774600060682
steps: 278	lr: 4e-05	evaluation reward: 3.18	
episode: 1220	score: 6.0	memory length: 241542	epsilon: 0.7197448600060841
steps: 370	lr: 4e-05	evaluation reward: 3.2	
episode: 1221	score: 4.0	memory length: 241823	epsilon: 0.7191884800060961
steps: 281	lr: 4e-05	evaluation reward: 3.23	
episode: 1222	score: 4.0	memory length: 242117	epsilon: 0.7186063600061088
steps: 294	lr: 4e-05	evaluation reward: 3.25	
episode: 1223	score: 3.0	memory length: 242366	epsilon: 0.7181133400061195
steps: 249	lr: 4e-05	evaluation reward: 3.26	
episode: 1224	score: 4.0	memory length: 242642	epsilon: 0.7175668600061313
steps: 276	lr: 4e-05	evaluation reward: 3.25	
episode: 1225	score: 3.0	memory length: 242893	epsilon: 0.7170698800061421
steps: 251	lr: 4e-05	evaluation reward: 3.22	
episode: 1226	score: 5.0	memory length: 243199	epsilon: 0.7164640000061553
steps: 306	lr: 4e-05	evaluation reward: 3.25	
episode: 1227	score: 4.0	memory length: 243497	epsilon: 0.7158739600061681
steps: 298	lr: 4e-05	evaluation reward: 3.22	
episode: 1228	score: 3.0	memory length: 243730	epsilon: 0.7154126200061781
steps: 233	lr: 4e-05	evaluation reward: 3.22	
episode: 1229	score: 6.0	memory length: 244106	epsilon: 0.7146681400061943
steps: 376	lr: 4e-05	evaluation reward: 3.25	
episode: 1230	score: 2.0	memory length: 244289	epsilon: 0.7143058000062021
steps: 183	lr: 4e-05	evaluation reward: 3.25	
episode: 1231	score: 2.0	memory length: 244488	epsilon: 0.7139117800062107
steps: 199	lr: 4e-05	evaluation reward: 3.25	
episode: 1232	score: 2.0	memory length: 244707	epsilon: 0.7134781600062201
steps: 219	lr: 4e-05	evaluation reward: 3.24	
episode: 1233	score: 8.0	memory length: 245128	epsilon: 0.7126445800062382
steps: 421	lr: 4e-05	evaluation reward: 3.28	
episode: 1234	score: 4.0	memory length: 245426	epsilon: 0.712054540006251
steps: 298	lr: 4e-05	evaluation reward: 3.3	

episode: 1235	score: 3.0	memory length: 245674	epsilon: 0.7115635000062617
steps: 248	lr: 4e-05	evaluation reward: 3.3	
episode: 1236	score: 4.0	memory length: 245969	epsilon: 0.7109794000062744
steps: 295	lr: 4e-05	evaluation reward: 3.33	
episode: 1237	score: 5.0	memory length: 246260	epsilon: 0.7104032200062869
steps: 291	lr: 4e-05	evaluation reward: 3.36	
episode: 1238	score: 3.0	memory length: 246487	epsilon: 0.7099537600062966
steps: 227	lr: 4e-05	evaluation reward: 3.35	
episode: 1239	score: 2.0	memory length: 246685	epsilon: 0.7095617200063051
steps: 198	lr: 4e-05	evaluation reward: 3.36	
episode: 1240	score: 3.0	memory length: 246912	epsilon: 0.7091122600063149
steps: 227	lr: 4e-05	evaluation reward: 3.35	
episode: 1241	score: 2.0	memory length: 247114	epsilon: 0.7087123000063236
steps: 202	lr: 4e-05	evaluation reward: 3.33	
episode: 1242	score: 2.0	memory length: 247295	epsilon: 0.7083539200063314
steps: 181	lr: 4e-05	evaluation reward: 3.29	
episode: 1243	score: 2.0	memory length: 247494	epsilon: 0.7079599000063399
steps: 199	lr: 4e-05	evaluation reward: 3.28	
episode: 1244	score: 2.0	memory length: 247711	epsilon: 0.7075302400063492
steps: 217	lr: 4e-05	evaluation reward: 3.27	
episode: 1245	score: 4.0	memory length: 248005	epsilon: 0.7069481200063619
steps: 294	lr: 4e-05	evaluation reward: 3.29	
episode: 1246	score: 3.0	memory length: 248253	epsilon: 0.7064570800063725
steps: 248	lr: 4e-05	evaluation reward: 3.29	
episode: 1247	score: 4.0	memory length: 248531	epsilon: 0.7059066400063845
steps: 278	lr: 4e-05	evaluation reward: 3.3	
episode: 1248	score: 6.0	memory length: 248893	epsilon: 0.7051898800064
steps: 362	lr: 4e-05	evaluation reward: 3.29	
episode: 1249	score: 2.0	memory length: 249091	epsilon: 0.7047978400064085
steps: 198	lr: 4e-05	evaluation reward: 3.28	
episode: 1250	score: 4.0	memory length: 249333	epsilon: 0.704318680006419
steps: 242	lr: 4e-05	evaluation reward: 3.3	
episode: 1251	score: 1.0	memory length: 249485	epsilon: 0.7040177200064255
steps: 152	lr: 4e-05	evaluation reward: 3.3	
episode: 1252	score: 5.0	memory length: 249828	epsilon: 0.7033385800064402
steps: 343	lr: 4e-05	evaluation reward: 3.31	
episode: 1253	score: 6.0	memory length: 250183	epsilon: 0.7026356800064555
steps: 355	lr: 4e-05	evaluation reward: 3.37	
episode: 1254	score: 2.0	memory length: 250383	epsilon: 0.7022396800064641
steps: 200	lr: 4e-05	evaluation reward: 3.35	
episode: 1255	score: 3.0	memory length: 250609	epsilon: 0.7017922000064738
steps: 226	lr: 4e-05	evaluation reward: 3.35	
episode: 1256	score: 5.0	memory length: 250899	epsilon: 0.7012180000064863
steps: 290	lr: 4e-05	evaluation reward: 3.37	
episode: 1257	score: 3.0	memory length: 251146	epsilon: 0.7007289400064969
steps: 247	lr: 4e-05	evaluation reward: 3.37	
episode: 1258	score: 4.0	memory length: 251422	epsilon: 0.7001824600065087
steps: 276	lr: 4e-05	evaluation reward: 3.39	

episode: 1259	score: 2.0	memory length: 251607	epsilon: 0.6998161600065167
steps: 185	lr: 4e-05	evaluation reward: 3.38	
episode: 1260	score: 3.0	memory length: 251840	epsilon: 0.6993548200065267
steps: 233	lr: 4e-05	evaluation reward: 3.36	
episode: 1261	score: 2.0	memory length: 252039	epsilon: 0.6989608000065353
steps: 199	lr: 4e-05	evaluation reward: 3.37	
episode: 1262	score: 3.0	memory length: 252290	epsilon: 0.698463820006546
steps: 251	lr: 4e-05	evaluation reward: 3.39	
episode: 1263	score: 4.0	memory length: 252534	epsilon: 0.6979807000065565
steps: 244	lr: 4e-05	evaluation reward: 3.42	
episode: 1264	score: 2.0	memory length: 252735	epsilon: 0.6975827200065652
steps: 201	lr: 4e-05	evaluation reward: 3.37	
episode: 1265	score: 0.0	memory length: 252858	epsilon: 0.6973391800065705
steps: 123	lr: 4e-05	evaluation reward: 3.34	
episode: 1266	score: 2.0	memory length: 253041	epsilon: 0.6969768400065783
steps: 183	lr: 4e-05	evaluation reward: 3.33	
episode: 1267	score: 2.0	memory length: 253239	epsilon: 0.6965848000065868
steps: 198	lr: 4e-05	evaluation reward: 3.32	
episode: 1268	score: 1.0	memory length: 253392	epsilon: 0.6962818600065934
steps: 153	lr: 4e-05	evaluation reward: 3.3	
episode: 1269	score: 2.0	memory length: 253590	epsilon: 0.6958898200066019
steps: 198	lr: 4e-05	evaluation reward: 3.28	
episode: 1270	score: 0.0	memory length: 253714	epsilon: 0.6956443000066073
steps: 124	lr: 4e-05	evaluation reward: 3.24	
episode: 1271	score: 4.0	memory length: 254028	epsilon: 0.6950225800066208
steps: 314	lr: 4e-05	evaluation reward: 3.21	
episode: 1272	score: 3.0	memory length: 254275	epsilon: 0.6945335200066314
steps: 247	lr: 4e-05	evaluation reward: 3.22	
episode: 1273	score: 2.0	memory length: 254474	epsilon: 0.6941395000066399
steps: 199	lr: 4e-05	evaluation reward: 3.18	
episode: 1274	score: 6.0	memory length: 254875	epsilon: 0.6933455200066572
steps: 401	lr: 4e-05	evaluation reward: 3.22	
episode: 1275	score: 1.0	memory length: 255045	epsilon: 0.6930089200066645
steps: 170	lr: 4e-05	evaluation reward: 3.2	
episode: 1276	score: 7.0	memory length: 255425	epsilon: 0.6922565200066808
steps: 380	lr: 4e-05	evaluation reward: 3.22	
episode: 1277	score: 3.0	memory length: 255655	epsilon: 0.6918011200066907
steps: 230	lr: 4e-05	evaluation reward: 3.21	
episode: 1278	score: 3.0	memory length: 255884	epsilon: 0.6913477000067005
steps: 229	lr: 4e-05	evaluation reward: 3.22	
episode: 1279	score: 1.0	memory length: 256036	epsilon: 0.6910467400067071
steps: 152	lr: 4e-05	evaluation reward: 3.18	
episode: 1280	score: 2.0	memory length: 256235	epsilon: 0.6906527200067156
steps: 199	lr: 4e-05	evaluation reward: 3.15	
episode: 1281	score: 2.0	memory length: 256435	epsilon: 0.6902567200067242
steps: 200	lr: 4e-05	evaluation reward: 3.12	
episode: 1282	score: 5.0	memory length: 256743	epsilon: 0.6896468800067375
steps: 308	lr: 4e-05	evaluation reward: 3.15	

episode: 1283	score: 2.0	memory length: 256926	epsilon: 0.6892845400067453
steps: 183	lr: 4e-05	evaluation reward: 3.14	
episode: 1284	score: 7.0	memory length: 257322	epsilon: 0.6885004600067623
steps: 396	lr: 4e-05	evaluation reward: 3.17	
episode: 1285	score: 9.0	memory length: 257742	epsilon: 0.6876688600067804
steps: 420	lr: 4e-05	evaluation reward: 3.25	
episode: 1286	score: 3.0	memory length: 257974	epsilon: 0.6872095000067904
steps: 232	lr: 4e-05	evaluation reward: 3.23	
episode: 1287	score: 3.0	memory length: 258188	epsilon: 0.6867857800067996
steps: 214	lr: 4e-05	evaluation reward: 3.23	
episode: 1288	score: 2.0	memory length: 258387	epsilon: 0.6863917600068081
steps: 199	lr: 4e-05	evaluation reward: 3.23	
episode: 1289	score: 3.0	memory length: 258598	epsilon: 0.6859739800068172
steps: 211	lr: 4e-05	evaluation reward: 3.22	
episode: 1290	score: 2.0	memory length: 258781	epsilon: 0.6856116400068251
steps: 183	lr: 4e-05	evaluation reward: 3.22	
episode: 1291	score: 2.0	memory length: 259000	epsilon: 0.6851780200068345
steps: 219	lr: 4e-05	evaluation reward: 3.22	
episode: 1292	score: 2.0	memory length: 259181	epsilon: 0.6848196400068423
steps: 181	lr: 4e-05	evaluation reward: 3.19	
episode: 1293	score: 0.0	memory length: 259305	epsilon: 0.6845741200068476
steps: 124	lr: 4e-05	evaluation reward: 3.16	
episode: 1294	score: 6.0	memory length: 259701	epsilon: 0.6837900400068646
steps: 396	lr: 4e-05	evaluation reward: 3.2	
episode: 1295	score: 2.0	memory length: 259882	epsilon: 0.6834316600068724
steps: 181	lr: 4e-05	evaluation reward: 3.19	
episode: 1296	score: 2.0	memory length: 260081	epsilon: 0.6830376400068809
steps: 199	lr: 4e-05	evaluation reward: 3.19	
episode: 1297	score: 6.0	memory length: 260392	epsilon: 0.6824218600068943
steps: 311	lr: 4e-05	evaluation reward: 3.25	
episode: 1298	score: 8.0	memory length: 260808	epsilon: 0.6815981800069122
steps: 416	lr: 4e-05	evaluation reward: 3.33	
episode: 1299	score: 6.0	memory length: 261144	epsilon: 0.6809329000069266
steps: 336	lr: 4e-05	evaluation reward: 3.38	
episode: 1300	score: 4.0	memory length: 261419	epsilon: 0.6803884000069385
steps: 275	lr: 4e-05	evaluation reward: 3.4	
episode: 1301	score: 1.0	memory length: 261571	epsilon: 0.680087440006945
steps: 152	lr: 4e-05	evaluation reward: 3.36	
episode: 1302	score: 2.0	memory length: 261752	epsilon: 0.6797290600069528
steps: 181	lr: 4e-05	evaluation reward: 3.34	
episode: 1303	score: 5.0	memory length: 262059	epsilon: 0.679121200006966
steps: 307	lr: 4e-05	evaluation reward: 3.38	
episode: 1304	score: 3.0	memory length: 262288	epsilon: 0.6786677800069758
steps: 229	lr: 4e-05	evaluation reward: 3.35	
episode: 1305	score: 6.0	memory length: 262649	epsilon: 0.6779530000069913
steps: 361	lr: 4e-05	evaluation reward: 3.39	
episode: 1306	score: 2.0	memory length: 262848	epsilon: 0.6775589800069999
steps: 199	lr: 4e-05	evaluation reward: 3.36	

episode: 1307	score: 3.0	memory length: 263095	epsilon: 0.67706992000070105
steps: 247	lr: 4e-05	evaluation reward: 3.36	
episode: 1308	score: 3.0	memory length: 263322	epsilon: 0.67662046000070203
steps: 227	lr: 4e-05	evaluation reward: 3.37	
episode: 1309	score: 4.0	memory length: 263609	epsilon: 0.67605220000070326
steps: 287	lr: 4e-05	evaluation reward: 3.33	
episode: 1310	score: 1.0	memory length: 263780	epsilon: 0.67571362000070399
steps: 171	lr: 4e-05	evaluation reward: 3.27	
episode: 1311	score: 1.0	memory length: 263949	epsilon: 0.67537900000070472
steps: 169	lr: 4e-05	evaluation reward: 3.26	
episode: 1312	score: 3.0	memory length: 264176	epsilon: 0.6749295400007057
steps: 227	lr: 4e-05	evaluation reward: 3.26	
episode: 1313	score: 3.0	memory length: 264408	epsilon: 0.67447018000070669
steps: 232	lr: 4e-05	evaluation reward: 3.26	
episode: 1314	score: 3.0	memory length: 264620	epsilon: 0.6740504200007076
steps: 212	lr: 4e-05	evaluation reward: 3.27	
episode: 1315	score: 2.0	memory length: 264802	epsilon: 0.67369006000070839
steps: 182	lr: 4e-05	evaluation reward: 3.29	
episode: 1316	score: 5.0	memory length: 265113	epsilon: 0.67307428000070972
steps: 311	lr: 4e-05	evaluation reward: 3.31	
episode: 1317	score: 2.0	memory length: 265332	epsilon: 0.67264066000071066
steps: 219	lr: 4e-05	evaluation reward: 3.29	
episode: 1318	score: 3.0	memory length: 265559	epsilon: 0.67219120000071164
steps: 227	lr: 4e-05	evaluation reward: 3.27	
episode: 1319	score: 3.0	memory length: 265786	epsilon: 0.67174174000071262
steps: 227	lr: 4e-05	evaluation reward: 3.26	
episode: 1320	score: 6.0	memory length: 266111	epsilon: 0.67109824000071401
steps: 325	lr: 4e-05	evaluation reward: 3.26	
episode: 1321	score: 4.0	memory length: 266389	epsilon: 0.67054780000071521
steps: 278	lr: 4e-05	evaluation reward: 3.26	
episode: 1322	score: 4.0	memory length: 266703	epsilon: 0.66992608000071656
steps: 314	lr: 4e-05	evaluation reward: 3.26	
episode: 1323	score: 4.0	memory length: 267018	epsilon: 0.66930238000071791
steps: 315	lr: 4e-05	evaluation reward: 3.27	
episode: 1324	score: 5.0	memory length: 267342	epsilon: 0.6686608600007193
steps: 324	lr: 4e-05	evaluation reward: 3.28	
episode: 1325	score: 2.0	memory length: 267561	epsilon: 0.66822724000072025
steps: 219	lr: 4e-05	evaluation reward: 3.27	
episode: 1326	score: 5.0	memory length: 267889	epsilon: 0.66757780000072166
steps: 328	lr: 4e-05	evaluation reward: 3.27	
episode: 1327	score: 6.0	memory length: 268263	epsilon: 0.66683728000072326
steps: 374	lr: 4e-05	evaluation reward: 3.29	
episode: 1328	score: 2.0	memory length: 268480	epsilon: 0.6664076200007242
steps: 217	lr: 4e-05	evaluation reward: 3.28	
episode: 1329	score: 5.0	memory length: 268806	epsilon: 0.6657621400007256
steps: 326	lr: 4e-05	evaluation reward: 3.27	
episode: 1330	score: 4.0	memory length: 269086	epsilon: 0.6652077400007268
steps: 280	lr: 4e-05	evaluation reward: 3.29	

episode: 1331	score: 6.0	memory length: 269480	epsilon: 0.664427620007285
steps: 394	lr: 4e-05	evaluation reward: 3.33	
episode: 1332	score: 6.0	memory length: 269847	epsilon: 0.6637009600073007
steps: 367	lr: 4e-05	evaluation reward: 3.37	
episode: 1333	score: 2.0	memory length: 270030	epsilon: 0.6633386200073086
steps: 183	lr: 4e-05	evaluation reward: 3.31	
episode: 1334	score: 3.0	memory length: 270275	epsilon: 0.6628535200073191
steps: 245	lr: 4e-05	evaluation reward: 3.3	
episode: 1335	score: 2.0	memory length: 270494	epsilon: 0.6624199000073285
steps: 219	lr: 4e-05	evaluation reward: 3.29	
episode: 1336	score: 6.0	memory length: 270822	epsilon: 0.6617704600073426
steps: 328	lr: 4e-05	evaluation reward: 3.31	
episode: 1337	score: 6.0	memory length: 271192	epsilon: 0.6610378600073585
steps: 370	lr: 4e-05	evaluation reward: 3.32	
episode: 1338	score: 6.0	memory length: 271530	epsilon: 0.6603686200073731
steps: 338	lr: 4e-05	evaluation reward: 3.35	
episode: 1339	score: 3.0	memory length: 271762	epsilon: 0.659909260007383
steps: 232	lr: 4e-05	evaluation reward: 3.36	
episode: 1340	score: 1.0	memory length: 271914	epsilon: 0.6596083000073896
steps: 152	lr: 4e-05	evaluation reward: 3.34	
episode: 1341	score: 3.0	memory length: 272128	epsilon: 0.6591845800073988
steps: 214	lr: 4e-05	evaluation reward: 3.35	
episode: 1342	score: 4.0	memory length: 272451	epsilon: 0.6585450400074127
steps: 323	lr: 4e-05	evaluation reward: 3.37	
episode: 1343	score: 8.0	memory length: 272889	epsilon: 0.6576778000074315
steps: 438	lr: 4e-05	evaluation reward: 3.43	
episode: 1344	score: 2.0	memory length: 273090	epsilon: 0.6572798200074401
steps: 201	lr: 4e-05	evaluation reward: 3.43	
episode: 1345	score: 2.0	memory length: 273289	epsilon: 0.6568858000074487
steps: 199	lr: 4e-05	evaluation reward: 3.41	
episode: 1346	score: 3.0	memory length: 273537	epsilon: 0.6563947600074593
steps: 248	lr: 4e-05	evaluation reward: 3.41	
episode: 1347	score: 4.0	memory length: 273812	epsilon: 0.6558502600074712
steps: 275	lr: 4e-05	evaluation reward: 3.41	
episode: 1348	score: 5.0	memory length: 274137	epsilon: 0.6552067600074851
steps: 325	lr: 4e-05	evaluation reward: 3.4	
episode: 1349	score: 6.0	memory length: 274516	epsilon: 0.6544563400075014
steps: 379	lr: 4e-05	evaluation reward: 3.44	
episode: 1350	score: 4.0	memory length: 274793	epsilon: 0.6539078800075133
steps: 277	lr: 4e-05	evaluation reward: 3.44	
episode: 1351	score: 1.0	memory length: 274964	epsilon: 0.6535693000075207
steps: 171	lr: 4e-05	evaluation reward: 3.44	
episode: 1352	score: 7.0	memory length: 275353	epsilon: 0.6527990800075374
steps: 389	lr: 4e-05	evaluation reward: 3.46	
episode: 1353	score: 6.0	memory length: 275698	epsilon: 0.6521159800075522
steps: 345	lr: 4e-05	evaluation reward: 3.46	
episode: 1354	score: 3.0	memory length: 275928	epsilon: 0.6516605800075621
steps: 230	lr: 4e-05	evaluation reward: 3.47	

episode: 1355	score: 2.0	memory length: 276109	epsilon: 0.6513022000075699
steps: 181	lr: 4e-05	evaluation reward: 3.46	
episode: 1356	score: 2.0	memory length: 276327	epsilon: 0.6508705600075793
steps: 218	lr: 4e-05	evaluation reward: 3.43	
episode: 1357	score: 3.0	memory length: 276556	epsilon: 0.6504171400075891
steps: 229	lr: 4e-05	evaluation reward: 3.43	
episode: 1358	score: 2.0	memory length: 276759	epsilon: 0.6500152000075978
steps: 203	lr: 4e-05	evaluation reward: 3.41	
episode: 1359	score: 2.0	memory length: 276960	epsilon: 0.6496172200076065
steps: 201	lr: 4e-05	evaluation reward: 3.41	
episode: 1360	score: 1.0	memory length: 277112	epsilon: 0.649316260007613
steps: 152	lr: 4e-05	evaluation reward: 3.39	
episode: 1361	score: 3.0	memory length: 277342	epsilon: 0.6488608600076229
steps: 230	lr: 4e-05	evaluation reward: 3.4	
episode: 1362	score: 5.0	memory length: 277670	epsilon: 0.648211420007637
steps: 328	lr: 4e-05	evaluation reward: 3.42	
episode: 1363	score: 8.0	memory length: 278078	epsilon: 0.6474035800076545
steps: 408	lr: 4e-05	evaluation reward: 3.46	
episode: 1364	score: 5.0	memory length: 278401	epsilon: 0.6467640400076684
steps: 323	lr: 4e-05	evaluation reward: 3.49	
episode: 1365	score: 1.0	memory length: 278552	epsilon: 0.6464650600076749
steps: 151	lr: 4e-05	evaluation reward: 3.5	
episode: 1366	score: 4.0	memory length: 278851	epsilon: 0.6458730400076877
steps: 299	lr: 4e-05	evaluation reward: 3.52	
episode: 1367	score: 3.0	memory length: 279098	epsilon: 0.6453839800076984
steps: 247	lr: 4e-05	evaluation reward: 3.53	
episode: 1368	score: 1.0	memory length: 279249	epsilon: 0.6450850000077049
steps: 151	lr: 4e-05	evaluation reward: 3.53	
episode: 1369	score: 2.0	memory length: 279468	epsilon: 0.6446513800077143
steps: 219	lr: 4e-05	evaluation reward: 3.53	
episode: 1370	score: 1.0	memory length: 279638	epsilon: 0.6443147800077216
steps: 170	lr: 4e-05	evaluation reward: 3.54	
episode: 1371	score: 3.0	memory length: 279866	epsilon: 0.6438633400077314
steps: 228	lr: 4e-05	evaluation reward: 3.53	
episode: 1372	score: 4.0	memory length: 280124	epsilon: 0.6433525000077425
steps: 258	lr: 4e-05	evaluation reward: 3.54	
episode: 1373	score: 2.0	memory length: 280306	epsilon: 0.6429921400077503
steps: 182	lr: 4e-05	evaluation reward: 3.54	
episode: 1374	score: 7.0	memory length: 280727	epsilon: 0.6421585600077684
steps: 421	lr: 4e-05	evaluation reward: 3.55	
episode: 1375	score: 2.0	memory length: 280915	epsilon: 0.6417863200077765
steps: 188	lr: 4e-05	evaluation reward: 3.56	
episode: 1376	score: 3.0	memory length: 281146	epsilon: 0.6413289400077864
steps: 231	lr: 4e-05	evaluation reward: 3.52	
episode: 1377	score: 5.0	memory length: 281458	epsilon: 0.6407111800077998
steps: 312	lr: 4e-05	evaluation reward: 3.54	
episode: 1378	score: 7.0	memory length: 281883	epsilon: 0.6398696800078181
steps: 425	lr: 4e-05	evaluation reward: 3.58	

episode: 1379	score: 4.0	memory length: 282142	epsilon: 0.6393568600078292
steps: 259	lr: 4e-05	evaluation reward: 3.61	
episode: 1380	score: 3.0	memory length: 282351	epsilon: 0.6389430400078382
steps: 209	lr: 4e-05	evaluation reward: 3.62	
episode: 1381	score: 4.0	memory length: 282629	epsilon: 0.6383926000078501
steps: 278	lr: 4e-05	evaluation reward: 3.64	
episode: 1382	score: 1.0	memory length: 282780	epsilon: 0.6380936200078566
steps: 151	lr: 4e-05	evaluation reward: 3.6	
episode: 1383	score: 5.0	memory length: 283126	epsilon: 0.6374085400078715
steps: 346	lr: 4e-05	evaluation reward: 3.63	
episode: 1384	score: 4.0	memory length: 283393	epsilon: 0.636879880007883
steps: 267	lr: 4e-05	evaluation reward: 3.6	
episode: 1385	score: 4.0	memory length: 283656	epsilon: 0.6363591400078943
steps: 263	lr: 4e-05	evaluation reward: 3.55	
episode: 1386	score: 6.0	memory length: 284028	epsilon: 0.6356225800079103
steps: 372	lr: 4e-05	evaluation reward: 3.58	
episode: 1387	score: 5.0	memory length: 284374	epsilon: 0.6349375000079251
steps: 346	lr: 4e-05	evaluation reward: 3.6	
episode: 1388	score: 4.0	memory length: 284654	epsilon: 0.6343831000079372
steps: 280	lr: 4e-05	evaluation reward: 3.62	
episode: 1389	score: 2.0	memory length: 284873	epsilon: 0.6339494800079466
steps: 219	lr: 4e-05	evaluation reward: 3.61	
episode: 1390	score: 2.0	memory length: 285092	epsilon: 0.633515860007956
steps: 219	lr: 4e-05	evaluation reward: 3.61	
episode: 1391	score: 5.0	memory length: 285414	epsilon: 0.6328783000079699
steps: 322	lr: 4e-05	evaluation reward: 3.64	
episode: 1392	score: 3.0	memory length: 285662	epsilon: 0.6323872600079805
steps: 248	lr: 4e-05	evaluation reward: 3.65	
episode: 1393	score: 6.0	memory length: 286001	epsilon: 0.6317160400079951
steps: 339	lr: 4e-05	evaluation reward: 3.71	
episode: 1394	score: 4.0	memory length: 286260	epsilon: 0.6312032200080062
steps: 259	lr: 4e-05	evaluation reward: 3.69	
episode: 1395	score: 3.0	memory length: 286490	epsilon: 0.6307478200080161
steps: 230	lr: 4e-05	evaluation reward: 3.7	
episode: 1396	score: 4.0	memory length: 286783	epsilon: 0.6301676800080287
steps: 293	lr: 4e-05	evaluation reward: 3.72	
episode: 1397	score: 5.0	memory length: 287110	epsilon: 0.6295202200080428
steps: 327	lr: 4e-05	evaluation reward: 3.71	
episode: 1398	score: 2.0	memory length: 287310	epsilon: 0.6291242200080513
steps: 200	lr: 4e-05	evaluation reward: 3.65	
episode: 1399	score: 7.0	memory length: 287733	epsilon: 0.6282866800080695
steps: 423	lr: 4e-05	evaluation reward: 3.66	
episode: 1400	score: 3.0	memory length: 287984	epsilon: 0.6277897000080803
steps: 251	lr: 4e-05	evaluation reward: 3.65	
episode: 1401	score: 4.0	memory length: 288242	epsilon: 0.6272788600080914
steps: 258	lr: 4e-05	evaluation reward: 3.68	
episode: 1402	score: 2.0	memory length: 288424	epsilon: 0.6269185000080992
steps: 182	lr: 4e-05	evaluation reward: 3.68	

episode: 1403	score: 4.0	memory length: 288702	epsilon: 0.6263680600081112
steps: 278	lr: 4e-05	evaluation reward: 3.67	
episode: 1404	score: 4.0	memory length: 288980	epsilon: 0.6258176200081231
steps: 278	lr: 4e-05	evaluation reward: 3.68	
episode: 1405	score: 2.0	memory length: 289181	epsilon: 0.6254196400081318
steps: 201	lr: 4e-05	evaluation reward: 3.64	
episode: 1406	score: 2.0	memory length: 289385	epsilon: 0.6250157200081405
steps: 204	lr: 4e-05	evaluation reward: 3.64	
episode: 1407	score: 7.0	memory length: 289810	epsilon: 0.6241742200081588
steps: 425	lr: 4e-05	evaluation reward: 3.68	
episode: 1408	score: 1.0	memory length: 289962	epsilon: 0.6238732600081653
steps: 152	lr: 4e-05	evaluation reward: 3.66	
episode: 1409	score: 1.0	memory length: 290132	epsilon: 0.6235366600081726
steps: 170	lr: 4e-05	evaluation reward: 3.63	
episode: 1410	score: 4.0	memory length: 290430	epsilon: 0.6229466200081855
steps: 298	lr: 4e-05	evaluation reward: 3.66	
episode: 1411	score: 5.0	memory length: 290739	epsilon: 0.6223348000081987
steps: 309	lr: 4e-05	evaluation reward: 3.7	
episode: 1412	score: 6.0	memory length: 291118	epsilon: 0.621584380008215
steps: 379	lr: 4e-05	evaluation reward: 3.73	
episode: 1413	score: 4.0	memory length: 291380	epsilon: 0.6210656200082263
steps: 262	lr: 4e-05	evaluation reward: 3.74	
episode: 1414	score: 5.0	memory length: 291710	epsilon: 0.6204122200082405
steps: 330	lr: 4e-05	evaluation reward: 3.76	
episode: 1415	score: 4.0	memory length: 292007	epsilon: 0.6198241600082532
steps: 297	lr: 4e-05	evaluation reward: 3.78	
episode: 1416	score: 4.0	memory length: 292284	epsilon: 0.6192757000082652
steps: 277	lr: 4e-05	evaluation reward: 3.77	
episode: 1417	score: 0.0	memory length: 292408	epsilon: 0.6190301800082705
steps: 124	lr: 4e-05	evaluation reward: 3.75	
episode: 1418	score: 4.0	memory length: 292668	epsilon: 0.6185153800082817
steps: 260	lr: 4e-05	evaluation reward: 3.76	
episode: 1419	score: 3.0	memory length: 292920	epsilon: 0.6180164200082925
steps: 252	lr: 4e-05	evaluation reward: 3.76	
episode: 1420	score: 9.0	memory length: 293310	epsilon: 0.6172442200083093
steps: 390	lr: 4e-05	evaluation reward: 3.79	
episode: 1421	score: 3.0	memory length: 293524	epsilon: 0.6168205000083185
steps: 214	lr: 4e-05	evaluation reward: 3.78	
episode: 1422	score: 4.0	memory length: 293800	epsilon: 0.6162740200083303
steps: 276	lr: 4e-05	evaluation reward: 3.78	
episode: 1423	score: 5.0	memory length: 294104	epsilon: 0.6156721000083434
steps: 304	lr: 4e-05	evaluation reward: 3.79	
episode: 1424	score: 5.0	memory length: 294413	epsilon: 0.6150602800083567
steps: 309	lr: 4e-05	evaluation reward: 3.79	
episode: 1425	score: 4.0	memory length: 294674	epsilon: 0.6145435000083679
steps: 261	lr: 4e-05	evaluation reward: 3.81	
episode: 1426	score: 3.0	memory length: 294903	epsilon: 0.6140900800083777
steps: 229	lr: 4e-05	evaluation reward: 3.79	

episode: 1427 score: 12.0 memory length: 295413 epsilon:
 0.6130802800083996 steps: 510 lr: 4e-05 evaluation reward: 3.85
 episode: 1428 score: 5.0 memory length: 295743 epsilon: 0.6124268800084138
 steps: 330 lr: 4e-05 evaluation reward: 3.88
 episode: 1429 score: 2.0 memory length: 295944 epsilon: 0.6120289000084225
 steps: 201 lr: 4e-05 evaluation reward: 3.85
 episode: 1430 score: 3.0 memory length: 296155 epsilon: 0.6116111200084315
 steps: 211 lr: 4e-05 evaluation reward: 3.84
 episode: 1431 score: 3.0 memory length: 296403 epsilon: 0.6111200800084422
 steps: 248 lr: 4e-05 evaluation reward: 3.81
 episode: 1432 score: 6.0 memory length: 296769 epsilon: 0.6103954000084579
 steps: 366 lr: 4e-05 evaluation reward: 3.81
 episode: 1433 score: 3.0 memory length: 297035 epsilon: 0.6098687200084694
 steps: 266 lr: 4e-05 evaluation reward: 3.82
 episode: 1434 score: 1.0 memory length: 297186 epsilon: 0.6095697400084759
 steps: 151 lr: 4e-05 evaluation reward: 3.8
 episode: 1435 score: 4.0 memory length: 297465 epsilon: 0.6090173200084879
 steps: 279 lr: 4e-05 evaluation reward: 3.82
 episode: 1436 score: 3.0 memory length: 297694 epsilon: 0.6085639000084977
 steps: 229 lr: 4e-05 evaluation reward: 3.79
 episode: 1437 score: 4.0 memory length: 298014 epsilon: 0.6079303000085114
 steps: 320 lr: 4e-05 evaluation reward: 3.77
 episode: 1438 score: 7.0 memory length: 298444 epsilon: 0.6070789000085299
 steps: 430 lr: 4e-05 evaluation reward: 3.78
 episode: 1439 score: 5.0 memory length: 298778 epsilon: 0.6064175800085443
 steps: 334 lr: 4e-05 evaluation reward: 3.8
 episode: 1440 score: 8.0 memory length: 299221 epsilon: 0.6055404400085633
 steps: 443 lr: 4e-05 evaluation reward: 3.87
 episode: 1441 score: 3.0 memory length: 299450 epsilon: 0.6050870200085732
 steps: 229 lr: 4e-05 evaluation reward: 3.87
 episode: 1442 score: 1.0 memory length: 299602 epsilon: 0.6047860600085797
 steps: 152 lr: 4e-05 evaluation reward: 3.84
 episode: 1443 score: 2.0 memory length: 299802 epsilon: 0.6043900600085883
 steps: 200 lr: 4e-05 evaluation reward: 3.78
 episode: 1444 score: 4.0 memory length: 300081 epsilon: 0.6038376400086003
 steps: 279 lr: 1.6000000000000003e-05 evaluation reward: 3.8
 episode: 1445 score: 6.0 memory length: 300420 epsilon: 0.6031664200086149
 steps: 339 lr: 1.6000000000000003e-05 evaluation reward: 3.84
 episode: 1446 score: 6.0 memory length: 300794 epsilon: 0.6024259000086309
 steps: 374 lr: 1.6000000000000003e-05 evaluation reward: 3.87
 episode: 1447 score: 4.0 memory length: 301071 epsilon: 0.6018774400086428
 steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 3.87
 episode: 1448 score: 5.0 memory length: 301384 epsilon: 0.6012577000086563
 steps: 313 lr: 1.6000000000000003e-05 evaluation reward: 3.87
 episode: 1449 score: 2.0 memory length: 301565 epsilon: 0.6008993200086641
 steps: 181 lr: 1.6000000000000003e-05 evaluation reward: 3.83
 episode: 1450 score: 5.0 memory length: 301878 epsilon: 0.6002795800086775
 steps: 313 lr: 1.6000000000000003e-05 evaluation reward: 3.84

episode: 1451 score: 2.0 memory length: 302059 epsilon: 0.5999212000086853
 steps: 181 lr: 1.6000000000000003e-05 evaluation reward: 3.85
 episode: 1452 score: 5.0 memory length: 302405 epsilon: 0.5992361200087002
 steps: 346 lr: 1.6000000000000003e-05 evaluation reward: 3.83
 episode: 1453 score: 2.0 memory length: 302605 epsilon: 0.5988401200087088
 steps: 200 lr: 1.6000000000000003e-05 evaluation reward: 3.79
 episode: 1454 score: 5.0 memory length: 302924 epsilon: 0.5982085000087225
 steps: 319 lr: 1.6000000000000003e-05 evaluation reward: 3.81
 episode: 1455 score: 3.0 memory length: 303154 epsilon: 0.5977531000087324
 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 3.82
 episode: 1456 score: 2.0 memory length: 303336 epsilon: 0.5973927400087402
 steps: 182 lr: 1.6000000000000003e-05 evaluation reward: 3.82
 episode: 1457 score: 4.0 memory length: 303594 epsilon: 0.5968819000087513
 steps: 258 lr: 1.6000000000000003e-05 evaluation reward: 3.83
 episode: 1458 score: 2.0 memory length: 303795 epsilon: 0.5964839200087599
 steps: 201 lr: 1.6000000000000003e-05 evaluation reward: 3.83
 episode: 1459 score: 4.0 memory length: 304056 epsilon: 0.5959671400087712
 steps: 261 lr: 1.6000000000000003e-05 evaluation reward: 3.85
 episode: 1460 score: 3.0 memory length: 304266 epsilon: 0.5955513400087802
 steps: 210 lr: 1.6000000000000003e-05 evaluation reward: 3.87
 episode: 1461 score: 5.0 memory length: 304615 epsilon: 0.5948603200087952
 steps: 349 lr: 1.6000000000000003e-05 evaluation reward: 3.89
 episode: 1462 score: 4.0 memory length: 304910 epsilon: 0.5942762200088079
 steps: 295 lr: 1.6000000000000003e-05 evaluation reward: 3.88
 episode: 1463 score: 1.0 memory length: 305061 epsilon: 0.5939772400088144
 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 3.81
 episode: 1464 score: 7.0 memory length: 305466 epsilon: 0.5931753400088318
 steps: 405 lr: 1.6000000000000003e-05 evaluation reward: 3.83
 episode: 1465 score: 3.0 memory length: 305676 epsilon: 0.5927595400088408
 steps: 210 lr: 1.6000000000000003e-05 evaluation reward: 3.85
 episode: 1466 score: 0.0 memory length: 305799 epsilon: 0.5925160000088461
 steps: 123 lr: 1.6000000000000003e-05 evaluation reward: 3.81
 episode: 1467 score: 4.0 memory length: 306061 epsilon: 0.5919972400088573
 steps: 262 lr: 1.6000000000000003e-05 evaluation reward: 3.82
 episode: 1468 score: 5.0 memory length: 306391 epsilon: 0.5913438400088715
 steps: 330 lr: 1.6000000000000003e-05 evaluation reward: 3.86
 episode: 1469 score: 4.0 memory length: 306647 epsilon: 0.5908369600088825
 steps: 256 lr: 1.6000000000000003e-05 evaluation reward: 3.88
 episode: 1470 score: 4.0 memory length: 306908 epsilon: 0.5903201800088937
 steps: 261 lr: 1.6000000000000003e-05 evaluation reward: 3.91
 episode: 1471 score: 7.0 memory length: 307277 epsilon: 0.5895895600089096
 steps: 369 lr: 1.6000000000000003e-05 evaluation reward: 3.95
 episode: 1472 score: 4.0 memory length: 307536 epsilon: 0.5890767400089207
 steps: 259 lr: 1.6000000000000003e-05 evaluation reward: 3.95
 episode: 1473 score: 4.0 memory length: 307795 epsilon: 0.5885639200089319
 steps: 259 lr: 1.6000000000000003e-05 evaluation reward: 3.97
 episode: 1474 score: 2.0 memory length: 308013 epsilon: 0.5881322800089412
 steps: 218 lr: 1.6000000000000003e-05 evaluation reward: 3.92

episode: 1475 score: 12.0 memory length: 308496 epsilon: 0.587175940008962
 steps: 483 lr: 1.6000000000000003e-05 evaluation reward: 4.02
 episode: 1476 score: 0.0 memory length: 308620 epsilon: 0.5869304200089673
 steps: 124 lr: 1.6000000000000003e-05 evaluation reward: 3.99
 episode: 1477 score: 4.0 memory length: 308877 epsilon: 0.5864215600089784
 steps: 257 lr: 1.6000000000000003e-05 evaluation reward: 3.98
 episode: 1478 score: 6.0 memory length: 309269 epsilon: 0.5856454000089952
 steps: 392 lr: 1.6000000000000003e-05 evaluation reward: 3.97
 episode: 1479 score: 4.0 memory length: 309549 epsilon: 0.5850910000090073
 steps: 280 lr: 1.6000000000000003e-05 evaluation reward: 3.97
 episode: 1480 score: 6.0 memory length: 309868 epsilon: 0.584459380009021
 steps: 319 lr: 1.6000000000000003e-05 evaluation reward: 4.0
 episode: 1481 score: 3.0 memory length: 310116 epsilon: 0.5839683400090316
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 3.99
 episode: 1482 score: 5.0 memory length: 310428 epsilon: 0.583350580009045
 steps: 312 lr: 1.6000000000000003e-05 evaluation reward: 4.03
 episode: 1483 score: 5.0 memory length: 310752 epsilon: 0.582709060009059
 steps: 324 lr: 1.6000000000000003e-05 evaluation reward: 4.03
 episode: 1484 score: 7.0 memory length: 311134 epsilon: 0.5819527000090754
 steps: 382 lr: 1.6000000000000003e-05 evaluation reward: 4.06
 episode: 1485 score: 3.0 memory length: 311364 epsilon: 0.5814973000090853
 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 4.05
 episode: 1486 score: 3.0 memory length: 311574 epsilon: 0.5810815000090943
 steps: 210 lr: 1.6000000000000003e-05 evaluation reward: 4.02
 episode: 1487 score: 4.0 memory length: 311832 epsilon: 0.5805706600091054
 steps: 258 lr: 1.6000000000000003e-05 evaluation reward: 4.01
 episode: 1488 score: 3.0 memory length: 312062 epsilon: 0.5801152600091153
 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 4.0
 episode: 1489 score: 2.0 memory length: 312261 epsilon: 0.5797212400091238
 steps: 199 lr: 1.6000000000000003e-05 evaluation reward: 4.0
 episode: 1490 score: 9.0 memory length: 312763 epsilon: 0.5787272800091454
 steps: 502 lr: 1.6000000000000003e-05 evaluation reward: 4.07
 episode: 1491 score: 4.0 memory length: 313061 epsilon: 0.5781372400091582
 steps: 298 lr: 1.6000000000000003e-05 evaluation reward: 4.06
 episode: 1492 score: 1.0 memory length: 313233 epsilon: 0.5777966800091656
 steps: 172 lr: 1.6000000000000003e-05 evaluation reward: 4.04
 episode: 1493 score: 9.0 memory length: 313687 epsilon: 0.5768977600091851
 steps: 454 lr: 1.6000000000000003e-05 evaluation reward: 4.07
 episode: 1494 score: 3.0 memory length: 313900 epsilon: 0.5764760200091943
 steps: 213 lr: 1.6000000000000003e-05 evaluation reward: 4.06
 episode: 1495 score: 5.0 memory length: 314204 epsilon: 0.5758741000092074
 steps: 304 lr: 1.6000000000000003e-05 evaluation reward: 4.08
 episode: 1496 score: 5.0 memory length: 314530 epsilon: 0.5752286200092214
 steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 4.09
 episode: 1497 score: 3.0 memory length: 314797 epsilon: 0.5746999600092328
 steps: 267 lr: 1.6000000000000003e-05 evaluation reward: 4.07
 episode: 1498 score: 6.0 memory length: 315152 epsilon: 0.5739970600092481
 steps: 355 lr: 1.6000000000000003e-05 evaluation reward: 4.11

episode: 1499 score: 2.0 memory length: 315332 epsilon: 0.5736406600092558
 steps: 180 lr: 1.6000000000000003e-05 evaluation reward: 4.06
 episode: 1500 score: 4.0 memory length: 315609 epsilon: 0.5730922000092677
 steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 4.07
 episode: 1501 score: 6.0 memory length: 315965 epsilon: 0.572387320009283
 steps: 356 lr: 1.6000000000000003e-05 evaluation reward: 4.09
 episode: 1502 score: 3.0 memory length: 316173 epsilon: 0.571975480009292
 steps: 208 lr: 1.6000000000000003e-05 evaluation reward: 4.1
 episode: 1503 score: 2.0 memory length: 316391 epsilon: 0.5715438400093014
 steps: 218 lr: 1.6000000000000003e-05 evaluation reward: 4.08
 episode: 1504 score: 6.0 memory length: 316785 epsilon: 0.5707637200093183
 steps: 394 lr: 1.6000000000000003e-05 evaluation reward: 4.1
 episode: 1505 score: 9.0 memory length: 317322 epsilon: 0.5697004600093414
 steps: 537 lr: 1.6000000000000003e-05 evaluation reward: 4.17
 episode: 1506 score: 3.0 memory length: 317568 epsilon: 0.569213380009352
 steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 4.18
 episode: 1507 score: 5.0 memory length: 317867 epsilon: 0.5686213600093648
 steps: 299 lr: 1.6000000000000003e-05 evaluation reward: 4.16
 episode: 1508 score: 3.0 memory length: 318138 epsilon: 0.5680847800093765
 steps: 271 lr: 1.6000000000000003e-05 evaluation reward: 4.18
 episode: 1509 score: 8.0 memory length: 318546 epsilon: 0.567276940009394
 steps: 408 lr: 1.6000000000000003e-05 evaluation reward: 4.25
 episode: 1510 score: 5.0 memory length: 318835 epsilon: 0.5667047200094064
 steps: 289 lr: 1.6000000000000003e-05 evaluation reward: 4.26
 episode: 1511 score: 5.0 memory length: 319159 epsilon: 0.5660632000094203
 steps: 324 lr: 1.6000000000000003e-05 evaluation reward: 4.26
 episode: 1512 score: 5.0 memory length: 319449 epsilon: 0.5654890000094328
 steps: 290 lr: 1.6000000000000003e-05 evaluation reward: 4.25
 episode: 1513 score: 4.0 memory length: 319750 epsilon: 0.5648930200094457
 steps: 301 lr: 1.6000000000000003e-05 evaluation reward: 4.25
 episode: 1514 score: 4.0 memory length: 320028 epsilon: 0.5643425800094577
 steps: 278 lr: 1.6000000000000003e-05 evaluation reward: 4.24
 episode: 1515 score: 10.0 memory length: 320524 epsilon: 0.563360500009479
 steps: 496 lr: 1.6000000000000003e-05 evaluation reward: 4.3
 episode: 1516 score: 6.0 memory length: 320844 epsilon: 0.5627269000094928
 steps: 320 lr: 1.6000000000000003e-05 evaluation reward: 4.32
 episode: 1517 score: 6.0 memory length: 321217 epsilon: 0.5619883600095088
 steps: 373 lr: 1.6000000000000003e-05 evaluation reward: 4.38
 episode: 1518 score: 8.0 memory length: 321573 epsilon: 0.5612834800095241
 steps: 356 lr: 1.6000000000000003e-05 evaluation reward: 4.42
 episode: 1519 score: 3.0 memory length: 321784 epsilon: 0.5608657000095332
 steps: 211 lr: 1.6000000000000003e-05 evaluation reward: 4.42
 episode: 1520 score: 5.0 memory length: 322126 epsilon: 0.5601885400095479
 steps: 342 lr: 1.6000000000000003e-05 evaluation reward: 4.38
 episode: 1521 score: 9.0 memory length: 322618 epsilon: 0.559214380009569
 steps: 492 lr: 1.6000000000000003e-05 evaluation reward: 4.44
 episode: 1522 score: 6.0 memory length: 322976 epsilon: 0.5585055400095844
 steps: 358 lr: 1.6000000000000003e-05 evaluation reward: 4.46

episode: 1523 score: 4.0 memory length: 323236 epsilon: 0.5579907400095956
 steps: 260 lr: 1.6000000000000003e-05 evaluation reward: 4.45
 episode: 1524 score: 3.0 memory length: 323448 epsilon: 0.5575709800096047
 steps: 212 lr: 1.6000000000000003e-05 evaluation reward: 4.43
 episode: 1525 score: 4.0 memory length: 323688 epsilon: 0.557095780009615
 steps: 240 lr: 1.6000000000000003e-05 evaluation reward: 4.43
 episode: 1526 score: 4.0 memory length: 323949 epsilon: 0.5565790000096262
 steps: 261 lr: 1.6000000000000003e-05 evaluation reward: 4.44
 episode: 1527 score: 6.0 memory length: 324304 epsilon: 0.5558761000096415
 steps: 355 lr: 1.6000000000000003e-05 evaluation reward: 4.38
 episode: 1528 score: 2.0 memory length: 324503 epsilon: 0.55548208000965
 steps: 199 lr: 1.6000000000000003e-05 evaluation reward: 4.35
 episode: 1529 score: 8.0 memory length: 324942 epsilon: 0.5546128600096689
 steps: 439 lr: 1.6000000000000003e-05 evaluation reward: 4.41
 episode: 1530 score: 5.0 memory length: 325248 epsilon: 0.5540069800096821
 steps: 306 lr: 1.6000000000000003e-05 evaluation reward: 4.43
 episode: 1531 score: 7.0 memory length: 325673 epsilon: 0.5531654800097003
 steps: 425 lr: 1.6000000000000003e-05 evaluation reward: 4.47
 episode: 1532 score: 4.0 memory length: 325972 epsilon: 0.5525734600097132
 steps: 299 lr: 1.6000000000000003e-05 evaluation reward: 4.45
 episode: 1533 score: 3.0 memory length: 326185 epsilon: 0.5521517200097223
 steps: 213 lr: 1.6000000000000003e-05 evaluation reward: 4.45
 episode: 1534 score: 7.0 memory length: 326548 epsilon: 0.551432980009738
 steps: 363 lr: 1.6000000000000003e-05 evaluation reward: 4.51
 episode: 1535 score: 5.0 memory length: 326909 epsilon: 0.5507182000097535
 steps: 361 lr: 1.6000000000000003e-05 evaluation reward: 4.52
 episode: 1536 score: 5.0 memory length: 327221 epsilon: 0.5501004400097669
 steps: 312 lr: 1.6000000000000003e-05 evaluation reward: 4.54
 episode: 1537 score: 7.0 memory length: 327617 epsilon: 0.5493163600097839
 steps: 396 lr: 1.6000000000000003e-05 evaluation reward: 4.57
 episode: 1538 score: 3.0 memory length: 327831 epsilon: 0.5488926400097931
 steps: 214 lr: 1.6000000000000003e-05 evaluation reward: 4.53
 episode: 1539 score: 5.0 memory length: 328128 epsilon: 0.5483045800098059
 steps: 297 lr: 1.6000000000000003e-05 evaluation reward: 4.53
 episode: 1540 score: 11.0 memory length: 328718 epsilon:
 0.5471363800098312 steps: 590 lr: 1.6000000000000003e-05 evaluation
 reward: 4.56
 episode: 1541 score: 3.0 memory length: 328948 epsilon: 0.5466809800098411
 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 4.56
 episode: 1542 score: 4.0 memory length: 329207 epsilon: 0.5461681600098522
 steps: 259 lr: 1.6000000000000003e-05 evaluation reward: 4.59
 episode: 1543 score: 5.0 memory length: 329552 epsilon: 0.5454850600098671
 steps: 345 lr: 1.6000000000000003e-05 evaluation reward: 4.62
 episode: 1544 score: 4.0 memory length: 329846 epsilon: 0.5449029400098797
 steps: 294 lr: 1.6000000000000003e-05 evaluation reward: 4.62
 episode: 1545 score: 3.0 memory length: 330057 epsilon: 0.5444851600098888
 steps: 211 lr: 1.6000000000000003e-05 evaluation reward: 4.59
 episode: 1546 score: 5.0 memory length: 330354 epsilon: 0.5438971000099015

```

steps: 297    lr: 1.6000000000000003e-05    evaluation reward: 4.58
episode: 1547  score: 4.0    memory length: 330631    epsilon: 0.5433486400099135
steps: 277    lr: 1.6000000000000003e-05    evaluation reward: 4.58
episode: 1548  score: 3.0    memory length: 330857    epsilon: 0.5429011600099232
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 4.56
episode: 1549  score: 3.0    memory length: 331106    epsilon: 0.5424081400099339
steps: 249    lr: 1.6000000000000003e-05    evaluation reward: 4.57
episode: 1550  score: 6.0    memory length: 331498    epsilon: 0.5416319800099507
steps: 392    lr: 1.6000000000000003e-05    evaluation reward: 4.58
episode: 1551  score: 5.0    memory length: 331807    epsilon: 0.541020160009964
steps: 309    lr: 1.6000000000000003e-05    evaluation reward: 4.61
episode: 1552  score: 3.0    memory length: 332018    epsilon: 0.5406023800099731
steps: 211    lr: 1.6000000000000003e-05    evaluation reward: 4.59
episode: 1553  score: 4.0    memory length: 332275    epsilon: 0.5400935200099841
steps: 257    lr: 1.6000000000000003e-05    evaluation reward: 4.61
episode: 1554  score: 6.0    memory length: 332615    epsilon: 0.5394203200099987
steps: 340    lr: 1.6000000000000003e-05    evaluation reward: 4.62
episode: 1555  score: 5.0    memory length: 332922    epsilon: 0.5388124600100119
steps: 307    lr: 1.6000000000000003e-05    evaluation reward: 4.64
episode: 1556  score: 4.0    memory length: 333163    epsilon: 0.5383352800100223
steps: 241    lr: 1.6000000000000003e-05    evaluation reward: 4.66
episode: 1557  score: 8.0    memory length: 333558    epsilon: 0.5375531800100393
steps: 395    lr: 1.6000000000000003e-05    evaluation reward: 4.7
episode: 1558  score: 4.0    memory length: 333841    epsilon: 0.5369928400100514
steps: 283    lr: 1.6000000000000003e-05    evaluation reward: 4.72
episode: 1559  score: 6.0    memory length: 334214    epsilon: 0.5362543000100675
steps: 373    lr: 1.6000000000000003e-05    evaluation reward: 4.74
episode: 1560  score: 3.0    memory length: 334426    epsilon: 0.5358345400100766
steps: 212    lr: 1.6000000000000003e-05    evaluation reward: 4.74
episode: 1561  score: 6.0    memory length: 334748    epsilon: 0.5351969800100904
steps: 322    lr: 1.6000000000000003e-05    evaluation reward: 4.75
episode: 1562  score: 4.0    memory length: 335008    epsilon: 0.5346821800101016
steps: 260    lr: 1.6000000000000003e-05    evaluation reward: 4.75
episode: 1563  score: 4.0    memory length: 335266    epsilon: 0.5341713400101127
steps: 258    lr: 1.6000000000000003e-05    evaluation reward: 4.78
episode: 1564  score: 7.0    memory length: 335653    epsilon: 0.5334050800101293
steps: 387    lr: 1.6000000000000003e-05    evaluation reward: 4.78
episode: 1565  score: 9.0    memory length: 336122    epsilon: 0.5324764600101495
steps: 469    lr: 1.6000000000000003e-05    evaluation reward: 4.84
episode: 1566  score: 2.0    memory length: 336302    epsilon: 0.5321200600101572
steps: 180    lr: 1.6000000000000003e-05    evaluation reward: 4.86
episode: 1567  score: 12.0   memory length: 336882    epsilon:
0.5309716600101821    steps: 580    lr: 1.6000000000000003e-05    evaluation
reward: 4.94
episode: 1568  score: 8.0    memory length: 337304    epsilon: 0.5301361000102003
steps: 422    lr: 1.6000000000000003e-05    evaluation reward: 4.97
episode: 1569  score: 4.0    memory length: 337582    epsilon: 0.5295856600102122
steps: 278    lr: 1.6000000000000003e-05    evaluation reward: 4.97

```


episode: 1570 score: 4.0 memory length: 337842 epsilon: 0.5290708600102234
 steps: 260 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1571 score: 3.0 memory length: 338056 epsilon: 0.5286471400102326
 steps: 214 lr: 1.6000000000000003e-05 evaluation reward: 4.93
 episode: 1572 score: 3.0 memory length: 338306 epsilon: 0.5281521400102434
 steps: 250 lr: 1.6000000000000003e-05 evaluation reward: 4.92
 episode: 1573 score: 7.0 memory length: 338663 epsilon: 0.5274452800102587
 steps: 357 lr: 1.6000000000000003e-05 evaluation reward: 4.95
 episode: 1574 score: 8.0 memory length: 339073 epsilon: 0.5266334800102763
 steps: 410 lr: 1.6000000000000003e-05 evaluation reward: 5.01
 episode: 1575 score: 6.0 memory length: 339447 epsilon: 0.5258929600102924
 steps: 374 lr: 1.6000000000000003e-05 evaluation reward: 4.95
 episode: 1576 score: 5.0 memory length: 339742 epsilon: 0.5253088600103051
 steps: 295 lr: 1.6000000000000003e-05 evaluation reward: 5.0
 episode: 1577 score: 9.0 memory length: 340211 epsilon: 0.5243802400103252
 steps: 469 lr: 1.6000000000000003e-05 evaluation reward: 5.05
 episode: 1578 score: 5.0 memory length: 340510 epsilon: 0.5237882200103381
 steps: 299 lr: 1.6000000000000003e-05 evaluation reward: 5.04
 episode: 1579 score: 5.0 memory length: 340808 epsilon: 0.5231981800103509
 steps: 298 lr: 1.6000000000000003e-05 evaluation reward: 5.05
 episode: 1580 score: 5.0 memory length: 341136 epsilon: 0.522548740010365
 steps: 328 lr: 1.6000000000000003e-05 evaluation reward: 5.04
 episode: 1581 score: 1.0 memory length: 341289 epsilon: 0.5222458000103716
 steps: 153 lr: 1.6000000000000003e-05 evaluation reward: 5.02
 episode: 1582 score: 5.0 memory length: 341595 epsilon: 0.5216399200103847
 steps: 306 lr: 1.6000000000000003e-05 evaluation reward: 5.02
 episode: 1583 score: 3.0 memory length: 341842 epsilon: 0.5211508600103953
 steps: 247 lr: 1.6000000000000003e-05 evaluation reward: 5.0
 episode: 1584 score: 2.0 memory length: 342045 epsilon: 0.5207489200104041
 steps: 203 lr: 1.6000000000000003e-05 evaluation reward: 4.95
 episode: 1585 score: 3.0 memory length: 342259 epsilon: 0.5203252000104133
 steps: 214 lr: 1.6000000000000003e-05 evaluation reward: 4.95
 episode: 1586 score: 7.0 memory length: 342634 epsilon: 0.5195827000104294
 steps: 375 lr: 1.6000000000000003e-05 evaluation reward: 4.99
 episode: 1587 score: 7.0 memory length: 343027 epsilon: 0.5188045600104463
 steps: 393 lr: 1.6000000000000003e-05 evaluation reward: 5.02
 episode: 1588 score: 5.0 memory length: 343363 epsilon: 0.5181392800104607
 steps: 336 lr: 1.6000000000000003e-05 evaluation reward: 5.04
 episode: 1589 score: 8.0 memory length: 343771 epsilon: 0.5173314400104783
 steps: 408 lr: 1.6000000000000003e-05 evaluation reward: 5.1
 episode: 1590 score: 9.0 memory length: 344210 epsilon: 0.5164622200104971
 steps: 439 lr: 1.6000000000000003e-05 evaluation reward: 5.1
 episode: 1591 score: 2.0 memory length: 344411 epsilon: 0.5160642400105058
 steps: 201 lr: 1.6000000000000003e-05 evaluation reward: 5.08
 episode: 1592 score: 3.0 memory length: 344661 epsilon: 0.5155692400105165
 steps: 250 lr: 1.6000000000000003e-05 evaluation reward: 5.1
 episode: 1593 score: 3.0 memory length: 344874 epsilon: 0.5151475000105257
 steps: 213 lr: 1.6000000000000003e-05 evaluation reward: 5.04

episode: 1594 score: 0.0 memory length: 344998 epsilon: 0.514901980010531
 steps: 124 lr: 1.6000000000000003e-05 evaluation reward: 5.01
 episode: 1595 score: 7.0 memory length: 345391 epsilon: 0.5141238400105479
 steps: 393 lr: 1.6000000000000003e-05 evaluation reward: 5.03
 episode: 1596 score: 7.0 memory length: 345769 epsilon: 0.5133754000105641
 steps: 378 lr: 1.6000000000000003e-05 evaluation reward: 5.05
 episode: 1597 score: 6.0 memory length: 346149 epsilon: 0.5126230000105805
 steps: 380 lr: 1.6000000000000003e-05 evaluation reward: 5.08
 episode: 1598 score: 6.0 memory length: 346471 epsilon: 0.5119854400105943
 steps: 322 lr: 1.6000000000000003e-05 evaluation reward: 5.08
 episode: 1599 score: 7.0 memory length: 346898 epsilon: 0.5111399800106127
 steps: 427 lr: 1.6000000000000003e-05 evaluation reward: 5.13
 episode: 1600 score: 3.0 memory length: 347112 epsilon: 0.5107162600106219
 steps: 214 lr: 1.6000000000000003e-05 evaluation reward: 5.12
 episode: 1601 score: 4.0 memory length: 347355 epsilon: 0.5102351200106323
 steps: 243 lr: 1.6000000000000003e-05 evaluation reward: 5.1
 episode: 1602 score: 5.0 memory length: 347645 epsilon: 0.5096609200106448
 steps: 290 lr: 1.6000000000000003e-05 evaluation reward: 5.12
 episode: 1603 score: 7.0 memory length: 348019 epsilon: 0.5089204000106609
 steps: 374 lr: 1.6000000000000003e-05 evaluation reward: 5.17
 episode: 1604 score: 2.0 memory length: 348217 epsilon: 0.5085283600106694
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 5.13
 episode: 1605 score: 5.0 memory length: 348516 epsilon: 0.5079363400106822
 steps: 299 lr: 1.6000000000000003e-05 evaluation reward: 5.09
 episode: 1606 score: 3.0 memory length: 348728 epsilon: 0.5075165800106913
 steps: 212 lr: 1.6000000000000003e-05 evaluation reward: 5.09
 episode: 1607 score: 10.0 memory length: 349259 epsilon:
 0.5064652000107142 steps: 531 lr: 1.6000000000000003e-05 evaluation
 reward: 5.14
 episode: 1608 score: 6.0 memory length: 349621 epsilon: 0.5057484400107297
 steps: 362 lr: 1.6000000000000003e-05 evaluation reward: 5.17
 episode: 1609 score: 5.0 memory length: 349907 epsilon: 0.505182160010742
 steps: 286 lr: 1.6000000000000003e-05 evaluation reward: 5.14
 episode: 1610 score: 9.0 memory length: 350420 epsilon: 0.5041664200107641
 steps: 513 lr: 1.6000000000000003e-05 evaluation reward: 5.18
 episode: 1611 score: 3.0 memory length: 350690 epsilon: 0.5036318200107757
 steps: 270 lr: 1.6000000000000003e-05 evaluation reward: 5.16
 episode: 1612 score: 3.0 memory length: 350958 epsilon: 0.5031011800107872
 steps: 268 lr: 1.6000000000000003e-05 evaluation reward: 5.14
 episode: 1613 score: 4.0 memory length: 351218 epsilon: 0.5025863800107984
 steps: 260 lr: 1.6000000000000003e-05 evaluation reward: 5.14
 episode: 1614 score: 5.0 memory length: 351523 epsilon: 0.5019824800108115
 steps: 305 lr: 1.6000000000000003e-05 evaluation reward: 5.15
 episode: 1615 score: 6.0 memory length: 351865 epsilon: 0.5013053200108262
 steps: 342 lr: 1.6000000000000003e-05 evaluation reward: 5.11
 episode: 1616 score: 1.0 memory length: 352017 epsilon: 0.5010043600108327
 steps: 152 lr: 1.6000000000000003e-05 evaluation reward: 5.06
 episode: 1617 score: 5.0 memory length: 352330 epsilon: 0.5003846200108462

steps: 313 lr: 1.6000000000000003e-05 evaluation reward: 5.05
 episode: 1618 score: 7.0 memory length: 352718 epsilon:
 0.49961638001085207 steps: 388 lr: 1.6000000000000003e-05 evaluation
 reward: 5.04
 episode: 1619 score: 4.0 memory length: 352982 epsilon:
 0.49909366001084876 steps: 264 lr: 1.6000000000000003e-05 evaluation
 reward: 5.05
 episode: 1620 score: 2.0 memory length: 353162 epsilon: 0.4987372600108465
 steps: 180 lr: 1.6000000000000003e-05 evaluation reward: 5.02
 episode: 1621 score: 4.0 memory length: 353422 epsilon:
 0.49822246001084325 steps: 260 lr: 1.6000000000000003e-05 evaluation
 reward: 4.97
 episode: 1622 score: 5.0 memory length: 353696 epsilon: 0.4976799400108398
 steps: 274 lr: 1.6000000000000003e-05 evaluation reward: 4.96
 episode: 1623 score: 4.0 memory length: 353939 epsilon:
 0.49719880001083677 steps: 243 lr: 1.6000000000000003e-05 evaluation
 reward: 4.96
 episode: 1624 score: 4.0 memory length: 354235 epsilon:
 0.49661272001083306 steps: 296 lr: 1.6000000000000003e-05 evaluation
 reward: 4.97
 episode: 1625 score: 3.0 memory length: 354467 epsilon:
 0.49615336001083016 steps: 232 lr: 1.6000000000000003e-05 evaluation
 reward: 4.96
 episode: 1626 score: 1.0 memory length: 354619 epsilon:
 0.49585240001082825 steps: 152 lr: 1.6000000000000003e-05 evaluation
 reward: 4.93
 episode: 1627 score: 9.0 memory length: 355138 epsilon:
 0.49482478001082175 steps: 519 lr: 1.6000000000000003e-05 evaluation
 reward: 4.96
 episode: 1628 score: 9.0 memory length: 355619 epsilon: 0.4938724000108157
 steps: 481 lr: 1.6000000000000003e-05 evaluation reward: 5.03
 episode: 1629 score: 6.0 memory length: 355984 epsilon:
 0.49314970001081115 steps: 365 lr: 1.6000000000000003e-05 evaluation
 reward: 5.01
 episode: 1630 score: 2.0 memory length: 356182 epsilon: 0.4927576600108087
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 4.98
 episode: 1631 score: 7.0 memory length: 356604 epsilon: 0.4919221000108034
 steps: 422 lr: 1.6000000000000003e-05 evaluation reward: 4.98
 episode: 1632 score: 4.0 memory length: 356889 epsilon: 0.4913578000107998
 steps: 285 lr: 1.6000000000000003e-05 evaluation reward: 4.98
 episode: 1633 score: 7.0 memory length: 357267 epsilon: 0.4906093600107951
 steps: 378 lr: 1.6000000000000003e-05 evaluation reward: 5.02
 episode: 1634 score: 4.0 memory length: 357562 epsilon: 0.4900252600107914
 steps: 295 lr: 1.6000000000000003e-05 evaluation reward: 4.99
 episode: 1635 score: 3.0 memory length: 357792 epsilon: 0.4895698600107885
 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1636 score: 2.0 memory length: 357991 epsilon: 0.489175840010786
 steps: 199 lr: 1.6000000000000003e-05 evaluation reward: 4.94

episode: 1637 score: 5.0 memory length: 358305 epsilon: 0.4885541200107821
 steps: 314 lr: 1.6000000000000003e-05 evaluation reward: 4.92
 episode: 1638 score: 8.0 memory length: 358746 epsilon:
 0.48768094001077655 steps: 441 lr: 1.6000000000000003e-05 evaluation
 reward: 4.97
 episode: 1639 score: 5.0 memory length: 359087 epsilon: 0.4870057600107723
 steps: 341 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1640 score: 7.0 memory length: 359483 epsilon: 0.4862216800107673
 steps: 396 lr: 1.6000000000000003e-05 evaluation reward: 4.93
 episode: 1641 score: 6.0 memory length: 359842 epsilon: 0.4855108600107628
 steps: 359 lr: 1.6000000000000003e-05 evaluation reward: 4.96
 episode: 1642 score: 5.0 memory length: 360135 epsilon:
 0.48493072001075915 steps: 293 lr: 1.6000000000000003e-05 evaluation
 reward: 4.97
 episode: 1643 score: 8.0 memory length: 360541 epsilon:
 0.48412684001075407 steps: 406 lr: 1.6000000000000003e-05 evaluation
 reward: 5.0
 episode: 1644 score: 7.0 memory length: 360935 epsilon:
 0.48334672001074913 steps: 394 lr: 1.6000000000000003e-05 evaluation
 reward: 5.03
 episode: 1645 score: 4.0 memory length: 361192 epsilon: 0.4828378600107459
 steps: 257 lr: 1.6000000000000003e-05 evaluation reward: 5.04
 episode: 1646 score: 6.0 memory length: 361542 epsilon: 0.4821448600107415
 steps: 350 lr: 1.6000000000000003e-05 evaluation reward: 5.05
 episode: 1647 score: 4.0 memory length: 361840 epsilon: 0.4815548200107378
 steps: 298 lr: 1.6000000000000003e-05 evaluation reward: 5.05
 episode: 1648 score: 3.0 memory length: 362067 epsilon:
 0.48110536001073495 steps: 227 lr: 1.6000000000000003e-05 evaluation
 reward: 5.05
 episode: 1649 score: 4.0 memory length: 362327 epsilon: 0.4805905600107317
 steps: 260 lr: 1.6000000000000003e-05 evaluation reward: 5.06
 episode: 1650 score: 9.0 memory length: 362801 epsilon:
 0.47965204001072576 steps: 474 lr: 1.6000000000000003e-05 evaluation
 reward: 5.09
 episode: 1651 score: 11.0 memory length: 363345 epsilon:
 0.47857492001071894 steps: 544 lr: 1.6000000000000003e-05 evaluation
 reward: 5.15
 episode: 1652 score: 5.0 memory length: 363639 epsilon:
 0.47799280001071526 steps: 294 lr: 1.6000000000000003e-05 evaluation
 reward: 5.17
 episode: 1653 score: 6.0 memory length: 363978 epsilon: 0.477321580010711
 steps: 339 lr: 1.6000000000000003e-05 evaluation reward: 5.19
 episode: 1654 score: 6.0 memory length: 364322 epsilon: 0.4766404600107067
 steps: 344 lr: 1.6000000000000003e-05 evaluation reward: 5.19
 episode: 1655 score: 5.0 memory length: 364634 epsilon: 0.4760227000107028
 steps: 312 lr: 1.6000000000000003e-05 evaluation reward: 5.19
 episode: 1656 score: 6.0 memory length: 364975 epsilon: 0.4753475200106985
 steps: 341 lr: 1.6000000000000003e-05 evaluation reward: 5.21

episode: 1657 score: 11.0 memory length: 365524 epsilon:
 0.47426050001069164 steps: 549 lr: 1.6000000000000003e-05 evaluation
 reward: 5.24
 episode: 1658 score: 4.0 memory length: 365825 epsilon: 0.4736645200106879
 steps: 301 lr: 1.6000000000000003e-05 evaluation reward: 5.24
 episode: 1659 score: 7.0 memory length: 366213 epsilon: 0.472896280010683
 steps: 388 lr: 1.6000000000000003e-05 evaluation reward: 5.25
 episode: 1660 score: 7.0 memory length: 366596 epsilon: 0.4721379400106782
 steps: 383 lr: 1.6000000000000003e-05 evaluation reward: 5.29
 episode: 1661 score: 5.0 memory length: 366924 epsilon: 0.4714885000106741
 steps: 328 lr: 1.6000000000000003e-05 evaluation reward: 5.28
 episode: 1662 score: 7.0 memory length: 367277 epsilon: 0.4707895600106697
 steps: 353 lr: 1.6000000000000003e-05 evaluation reward: 5.31
 episode: 1663 score: 7.0 memory length: 367652 epsilon: 0.470047060010665
 steps: 375 lr: 1.6000000000000003e-05 evaluation reward: 5.34
 episode: 1664 score: 4.0 memory length: 367916 epsilon: 0.4695243400106617
 steps: 264 lr: 1.6000000000000003e-05 evaluation reward: 5.31
 episode: 1665 score: 7.0 memory length: 368284 epsilon:
 0.46879570001065707 steps: 368 lr: 1.6000000000000003e-05 evaluation
 reward: 5.29
 episode: 1666 score: 3.0 memory length: 368511 epsilon: 0.4683462400106542
 steps: 227 lr: 1.6000000000000003e-05 evaluation reward: 5.3
 episode: 1667 score: 2.0 memory length: 368694 epsilon:
 0.46798390001065193 steps: 183 lr: 1.6000000000000003e-05 evaluation
 reward: 5.2
 episode: 1668 score: 6.0 memory length: 369057 epsilon: 0.4672651600106474
 steps: 363 lr: 1.6000000000000003e-05 evaluation reward: 5.18
 episode: 1669 score: 6.0 memory length: 369412 epsilon:
 0.46656226001064294 steps: 355 lr: 1.6000000000000003e-05 evaluation
 reward: 5.2
 episode: 1670 score: 6.0 memory length: 369788 epsilon: 0.4658177800106382
 steps: 376 lr: 1.6000000000000003e-05 evaluation reward: 5.22
 episode: 1671 score: 10.0 memory length: 370339 epsilon:
 0.4647268000106313 steps: 551 lr: 1.6000000000000003e-05 evaluation
 reward: 5.29
 episode: 1672 score: 3.0 memory length: 370568 epsilon:
 0.46427338001062846 steps: 229 lr: 1.6000000000000003e-05 evaluation
 reward: 5.29
 episode: 1673 score: 5.0 memory length: 370859 epsilon: 0.4636972000106248
 steps: 291 lr: 1.6000000000000003e-05 evaluation reward: 5.27
 episode: 1674 score: 6.0 memory length: 371203 epsilon: 0.4630160800106205
 steps: 344 lr: 1.6000000000000003e-05 evaluation reward: 5.25
 episode: 1675 score: 6.0 memory length: 371580 epsilon: 0.4622696200106158
 steps: 377 lr: 1.6000000000000003e-05 evaluation reward: 5.25
 episode: 1676 score: 5.0 memory length: 371921 epsilon: 0.4615944400106115
 steps: 341 lr: 1.6000000000000003e-05 evaluation reward: 5.25
 episode: 1677 score: 7.0 memory length: 372277 epsilon:
 0.46088956001060705 steps: 356 lr: 1.6000000000000003e-05 evaluation

```

reward: 5.23
episode: 1678  score: 9.0  memory length: 372744  epsilon: 0.4599649000106012
steps: 467  lr: 1.6000000000000003e-05  evaluation reward: 5.27
episode: 1679  score: 5.0  memory length: 373068  epsilon:
0.45932338001059714  steps: 324  lr: 1.6000000000000003e-05  evaluation
reward: 5.27
episode: 1680  score: 3.0  memory length: 373281  epsilon:
0.45890164001059447  steps: 213  lr: 1.6000000000000003e-05  evaluation
reward: 5.25
episode: 1681  score: 4.0  memory length: 373523  epsilon:
0.45842248001059144  steps: 242  lr: 1.6000000000000003e-05  evaluation
reward: 5.28
episode: 1682  score: 6.0  memory length: 373920  epsilon:
0.45763642001058646  steps: 397  lr: 1.6000000000000003e-05  evaluation
reward: 5.29
episode: 1683  score: 3.0  memory length: 374151  epsilon:
0.45717904001058357  steps: 231  lr: 1.6000000000000003e-05  evaluation
reward: 5.29
episode: 1684  score: 3.0  memory length: 374361  epsilon:
0.45676324001058094  steps: 210  lr: 1.6000000000000003e-05  evaluation
reward: 5.3
episode: 1685  score: 3.0  memory length: 374591  epsilon:
0.45630784001057806  steps: 230  lr: 1.6000000000000003e-05  evaluation
reward: 5.3
episode: 1686  score: 4.0  memory length: 374872  epsilon:
0.45575146001057454  steps: 281  lr: 1.6000000000000003e-05  evaluation
reward: 5.27
episode: 1687  score: 11.0  memory length: 375462  epsilon:
0.45458326001056715  steps: 590  lr: 1.6000000000000003e-05  evaluation
reward: 5.31
episode: 1688  score: 4.0  memory length: 375742  epsilon:
0.45402886001056364  steps: 280  lr: 1.6000000000000003e-05  evaluation
reward: 5.3
episode: 1689  score: 7.0  memory length: 376164  epsilon:
0.45319330001055835  steps: 422  lr: 1.6000000000000003e-05  evaluation
reward: 5.29
episode: 1690  score: 5.0  memory length: 376492  epsilon:
0.45254386001055424  steps: 328  lr: 1.6000000000000003e-05  evaluation
reward: 5.25
episode: 1691  score: 8.0  memory length: 376937  epsilon:
0.45166276001054867  steps: 445  lr: 1.6000000000000003e-05  evaluation
reward: 5.31
episode: 1692  score: 13.0  memory length: 377434  epsilon:
0.45067870001054244  steps: 497  lr: 1.6000000000000003e-05  evaluation
reward: 5.41
episode: 1693  score: 6.0  memory length: 377766  epsilon: 0.4500213400105383
steps: 332  lr: 1.6000000000000003e-05  evaluation reward: 5.44
episode: 1694  score: 3.0  memory length: 377994  epsilon:

```

0.44956990001053543 steps: 228 lr: 1.6000000000000003e-05 evaluation
 reward: 5.47
 episode: 1695 score: 5.0 memory length: 378318 epsilon:
 0.44892838001053137 steps: 324 lr: 1.6000000000000003e-05 evaluation
 reward: 5.45
 episode: 1696 score: 5.0 memory length: 378623 epsilon:
 0.44832448001052755 steps: 305 lr: 1.6000000000000003e-05 evaluation
 reward: 5.43
 episode: 1697 score: 2.0 memory length: 378841 epsilon: 0.4478928400105248
 steps: 218 lr: 1.6000000000000003e-05 evaluation reward: 5.39
 episode: 1698 score: 8.0 memory length: 379261 epsilon:
 0.44706124001051956 steps: 420 lr: 1.6000000000000003e-05 evaluation
 reward: 5.41
 episode: 1699 score: 5.0 memory length: 379605 epsilon:
 0.44638012001051525 steps: 344 lr: 1.6000000000000003e-05 evaluation
 reward: 5.39
 episode: 1700 score: 6.0 memory length: 379963 epsilon:
 0.44567128001051076 steps: 358 lr: 1.6000000000000003e-05 evaluation
 reward: 5.42
 episode: 1701 score: 5.0 memory length: 380291 epsilon:
 0.44502184001050665 steps: 328 lr: 1.6000000000000003e-05 evaluation
 reward: 5.43
 episode: 1702 score: 5.0 memory length: 380581 epsilon: 0.444447640010503
 steps: 290 lr: 1.6000000000000003e-05 evaluation reward: 5.43
 episode: 1703 score: 5.0 memory length: 380893 epsilon: 0.4438298800104991
 steps: 312 lr: 1.6000000000000003e-05 evaluation reward: 5.41
 episode: 1704 score: 5.0 memory length: 381188 epsilon: 0.4432457800104954
 steps: 295 lr: 1.6000000000000003e-05 evaluation reward: 5.44
 episode: 1705 score: 8.0 memory length: 381608 epsilon:
 0.44241418001049015 steps: 420 lr: 1.6000000000000003e-05 evaluation
 reward: 5.47
 episode: 1706 score: 7.0 memory length: 381991 epsilon:
 0.44165584001048536 steps: 383 lr: 1.6000000000000003e-05 evaluation
 reward: 5.51
 episode: 1707 score: 9.0 memory length: 382337 epsilon: 0.440970760010481
 steps: 346 lr: 1.6000000000000003e-05 evaluation reward: 5.5
 episode: 1708 score: 7.0 memory length: 382744 epsilon: 0.4401649000104759
 steps: 407 lr: 1.6000000000000003e-05 evaluation reward: 5.51
 episode: 1709 score: 9.0 memory length: 383233 epsilon: 0.4391966800104698
 steps: 489 lr: 1.6000000000000003e-05 evaluation reward: 5.55
 episode: 1710 score: 4.0 memory length: 383473 epsilon: 0.4387214800104668
 steps: 240 lr: 1.6000000000000003e-05 evaluation reward: 5.5
 episode: 1711 score: 3.0 memory length: 383704 epsilon: 0.4382641000104639
 steps: 231 lr: 1.6000000000000003e-05 evaluation reward: 5.5
 episode: 1712 score: 4.0 memory length: 383986 epsilon:
 0.43770574001046036 steps: 282 lr: 1.6000000000000003e-05 evaluation
 reward: 5.51
 episode: 1713 score: 8.0 memory length: 384389 epsilon: 0.4369078000104553

```

steps: 403    lr: 1.6000000000000003e-05    evaluation reward: 5.55
episode: 1714  score: 15.0  memory length: 384812  epsilon: 0.43607026001045
steps: 423    lr: 1.6000000000000003e-05    evaluation reward: 5.65
episode: 1715  score: 5.0  memory length: 385105  epsilon:
0.43549012001044635  steps: 293  lr: 1.6000000000000003e-05    evaluation
reward: 5.64
episode: 1716  score: 7.0  memory length: 385549  epsilon: 0.4346110000104408
steps: 444    lr: 1.6000000000000003e-05    evaluation reward: 5.7
episode: 1717  score: 4.0  memory length: 385809  epsilon: 0.4340962000104375
steps: 260    lr: 1.6000000000000003e-05    evaluation reward: 5.69
episode: 1718  score: 4.0  memory length: 386103  epsilon:
0.43351408001043384  steps: 294  lr: 1.6000000000000003e-05    evaluation
reward: 5.66
episode: 1719  score: 3.0  memory length: 386354  epsilon: 0.4330171000104307
steps: 251    lr: 1.6000000000000003e-05    evaluation reward: 5.65
episode: 1720  score: 4.0  memory length: 386634  epsilon: 0.4324627000104272
steps: 280    lr: 1.6000000000000003e-05    evaluation reward: 5.67
episode: 1721  score: 5.0  memory length: 386962  epsilon: 0.4318132600104231
steps: 328    lr: 1.6000000000000003e-05    evaluation reward: 5.68
episode: 1722  score: 6.0  memory length: 387319  epsilon: 0.4311064000104186
steps: 357    lr: 1.6000000000000003e-05    evaluation reward: 5.69
episode: 1723  score: 4.0  memory length: 387580  epsilon:
0.43058962001041534  steps: 261  lr: 1.6000000000000003e-05    evaluation
reward: 5.69
episode: 1724  score: 7.0  memory length: 387948  epsilon:
0.42986098001041073  steps: 368  lr: 1.6000000000000003e-05    evaluation
reward: 5.72
episode: 1725  score: 4.0  memory length: 388193  epsilon:
0.42937588001040766  steps: 245  lr: 1.6000000000000003e-05    evaluation
reward: 5.73
episode: 1726  score: 3.0  memory length: 388426  epsilon:
0.42891454001040474  steps: 233  lr: 1.6000000000000003e-05    evaluation
reward: 5.75
episode: 1727  score: 4.0  memory length: 388721  epsilon:
0.42833044001040105  steps: 295  lr: 1.6000000000000003e-05    evaluation
reward: 5.7
episode: 1728  score: 6.0  memory length: 389081  epsilon:
0.42761764001039654  steps: 360  lr: 1.6000000000000003e-05    evaluation
reward: 5.67
episode: 1729  score: 6.0  memory length: 389440  epsilon:
0.42690682001039204  steps: 359  lr: 1.6000000000000003e-05    evaluation
reward: 5.67
episode: 1730  score: 5.0  memory length: 389714  epsilon: 0.4263643000103886
steps: 274    lr: 1.6000000000000003e-05    evaluation reward: 5.7
episode: 1731  score: 5.0  memory length: 390021  epsilon:
0.42575644001038476  steps: 307  lr: 1.6000000000000003e-05    evaluation
reward: 5.68
episode: 1732  score: 3.0  memory length: 390235  epsilon: 0.4253327200103821

```


steps: 214 lr: 1.6000000000000003e-05 evaluation reward: 5.67
 episode: 1733 score: 7.0 memory length: 390630 epsilon:
 0.42455062001037713 steps: 395 lr: 1.6000000000000003e-05 evaluation
 reward: 5.67
 episode: 1734 score: 4.0 memory length: 390889 epsilon: 0.4240378000103739
 steps: 259 lr: 1.6000000000000003e-05 evaluation reward: 5.67
 episode: 1735 score: 8.0 memory length: 391333 epsilon: 0.4231586800103683
 steps: 444 lr: 1.6000000000000003e-05 evaluation reward: 5.72
 episode: 1736 score: 4.0 memory length: 391576 epsilon: 0.4226775400103653
 steps: 243 lr: 1.6000000000000003e-05 evaluation reward: 5.74
 episode: 1737 score: 6.0 memory length: 391906 epsilon:
 0.42202414001036115 steps: 330 lr: 1.6000000000000003e-05 evaluation
 reward: 5.75
 episode: 1738 score: 7.0 memory length: 392311 epsilon: 0.4212222400103561
 steps: 405 lr: 1.6000000000000003e-05 evaluation reward: 5.74
 episode: 1739 score: 7.0 memory length: 392698 epsilon: 0.4204559800103512
 steps: 387 lr: 1.6000000000000003e-05 evaluation reward: 5.76
 episode: 1740 score: 7.0 memory length: 393101 epsilon: 0.4196580400103462
 steps: 403 lr: 1.6000000000000003e-05 evaluation reward: 5.76
 episode: 1741 score: 7.0 memory length: 393511 epsilon:
 0.41884624001034104 steps: 410 lr: 1.6000000000000003e-05 evaluation
 reward: 5.77
 episode: 1742 score: 9.0 memory length: 393912 epsilon: 0.418052260010336
 steps: 401 lr: 1.6000000000000003e-05 evaluation reward: 5.81
 episode: 1743 score: 9.0 memory length: 394434 epsilon: 0.4170187000103295
 steps: 522 lr: 1.6000000000000003e-05 evaluation reward: 5.82
 episode: 1744 score: 4.0 memory length: 394731 epsilon:
 0.41643064001032576 steps: 297 lr: 1.6000000000000003e-05 evaluation
 reward: 5.79
 episode: 1745 score: 5.0 memory length: 395058 epsilon:
 0.41578318001032166 steps: 327 lr: 1.6000000000000003e-05 evaluation
 reward: 5.8
 episode: 1746 score: 5.0 memory length: 395385 epsilon:
 0.41513572001031757 steps: 327 lr: 1.6000000000000003e-05 evaluation
 reward: 5.79
 episode: 1747 score: 8.0 memory length: 395830 epsilon: 0.414254620010312
 steps: 445 lr: 1.6000000000000003e-05 evaluation reward: 5.83
 episode: 1748 score: 6.0 memory length: 396189 epsilon: 0.4135438000103075
 steps: 359 lr: 1.6000000000000003e-05 evaluation reward: 5.86
 episode: 1749 score: 3.0 memory length: 396402 epsilon: 0.4131220600103048
 steps: 213 lr: 1.6000000000000003e-05 evaluation reward: 5.85
 episode: 1750 score: 6.0 memory length: 396725 epsilon: 0.4124825200103008
 steps: 323 lr: 1.6000000000000003e-05 evaluation reward: 5.82
 episode: 1751 score: 4.0 memory length: 397006 epsilon:
 0.41192614001029726 steps: 281 lr: 1.6000000000000003e-05 evaluation
 reward: 5.75
 episode: 1752 score: 2.0 memory length: 397207 epsilon:
 0.41152816001029474 steps: 201 lr: 1.6000000000000003e-05 evaluation

```

reward: 5.72
episode: 1753  score: 6.0  memory length: 397607  epsilon:
0.41073616001028973  steps: 400  lr: 1.6000000000000003e-05  evaluation
reward: 5.72
episode: 1754  score: 6.0  memory length: 397934  epsilon:
0.41008870001028563  steps: 327  lr: 1.6000000000000003e-05  evaluation
reward: 5.72
episode: 1755  score: 4.0  memory length: 398190  epsilon: 0.4095818200102824
steps: 256  lr: 1.6000000000000003e-05  evaluation reward: 5.71
episode: 1756  score: 7.0  memory length: 398549  epsilon:
0.40887100001027793  steps: 359  lr: 1.6000000000000003e-05  evaluation
reward: 5.72
episode: 1757  score: 9.0  memory length: 399063  epsilon: 0.4078532800102715
steps: 514  lr: 1.6000000000000003e-05  evaluation reward: 5.7
episode: 1758  score: 12.0  memory length: 399495  epsilon:
0.4069979200102661  steps: 432  lr: 1.6000000000000003e-05  evaluation
reward: 5.78
episode: 1759  score: 9.0  memory length: 399967  epsilon:
0.40606336001026017  steps: 472  lr: 1.6000000000000003e-05  evaluation
reward: 5.8
episode: 1760  score: 10.0  memory length: 400507  epsilon:
0.4049941600102534  steps: 540  lr: 6.4000000000000001e-06  evaluation
reward: 5.83
episode: 1761  score: 2.0  memory length: 400689  epsilon: 0.4046338000102511
steps: 182  lr: 6.4000000000000001e-06  evaluation reward: 5.8
episode: 1762  score: 6.0  memory length: 401067  epsilon: 0.4038853600102464
steps: 378  lr: 6.4000000000000001e-06  evaluation reward: 5.79
episode: 1763  score: 6.0  memory length: 401390  epsilon:
0.40324582001024234  steps: 323  lr: 6.4000000000000001e-06  evaluation
reward: 5.78
episode: 1764  score: 8.0  memory length: 401818  epsilon: 0.402398380010237
steps: 428  lr: 6.4000000000000001e-06  evaluation reward: 5.82
episode: 1765  score: 6.0  memory length: 402159  epsilon: 0.4017232000102327
steps: 341  lr: 6.4000000000000001e-06  evaluation reward: 5.81
episode: 1766  score: 20.0  memory length: 402842  epsilon:
0.40037086001022415  steps: 683  lr: 6.4000000000000001e-06  evaluation
reward: 5.98
episode: 1767  score: 5.0  memory length: 403122  epsilon:
0.39981646001022064  steps: 280  lr: 6.4000000000000001e-06  evaluation
reward: 6.01
episode: 1768  score: 5.0  memory length: 403414  epsilon: 0.399238300010217
steps: 292  lr: 6.4000000000000001e-06  evaluation reward: 6.0
episode: 1769  score: 9.0  memory length: 403904  epsilon:
0.39826810001021085  steps: 490  lr: 6.4000000000000001e-06  evaluation
reward: 6.03
episode: 1770  score: 8.0  memory length: 404339  epsilon: 0.3974068000102054
steps: 435  lr: 6.4000000000000001e-06  evaluation reward: 6.05
episode: 1771  score: 9.0  memory length: 404846  epsilon:

```

0.39640294001019905 steps: 507 lr: 6.4000000000000001e-06 evaluation
 reward: 6.04
 episode: 1772 score: 9.0 memory length: 405303 epsilon: 0.3954980800101933
 steps: 457 lr: 6.4000000000000001e-06 evaluation reward: 6.1
 episode: 1773 score: 6.0 memory length: 405669 epsilon:
 0.39477340001018874 steps: 366 lr: 6.4000000000000001e-06 evaluation
 reward: 6.11
 episode: 1774 score: 5.0 memory length: 405983 epsilon: 0.3941516800101848
 steps: 314 lr: 6.4000000000000001e-06 evaluation reward: 6.1
 episode: 1775 score: 16.0 memory length: 406646 epsilon:
 0.3928389400101765 steps: 663 lr: 6.4000000000000001e-06 evaluation
 reward: 6.2
 episode: 1776 score: 6.0 memory length: 406970 epsilon:
 0.39219742001017244 steps: 324 lr: 6.4000000000000001e-06 evaluation
 reward: 6.21
 episode: 1777 score: 8.0 memory length: 407398 epsilon: 0.3913499800101671
 steps: 428 lr: 6.4000000000000001e-06 evaluation reward: 6.22
 episode: 1778 score: 9.0 memory length: 407832 epsilon:
 0.39049066001016164 steps: 434 lr: 6.4000000000000001e-06 evaluation
 reward: 6.22
 episode: 1779 score: 4.0 memory length: 408131 epsilon: 0.3898986400101579
 steps: 299 lr: 6.4000000000000001e-06 evaluation reward: 6.21
 episode: 1780 score: 9.0 memory length: 408441 epsilon: 0.389284840010154
 steps: 310 lr: 6.4000000000000001e-06 evaluation reward: 6.27
 episode: 1781 score: 6.0 memory length: 408780 epsilon:
 0.38861362001014976 steps: 339 lr: 6.4000000000000001e-06 evaluation
 reward: 6.29
 episode: 1782 score: 3.0 memory length: 408991 epsilon: 0.3881958400101471
 steps: 211 lr: 6.4000000000000001e-06 evaluation reward: 6.26
 episode: 1783 score: 8.0 memory length: 409434 epsilon:
 0.38731870001014157 steps: 443 lr: 6.4000000000000001e-06 evaluation
 reward: 6.31
 episode: 1784 score: 4.0 memory length: 409716 epsilon:
 0.38676034001013804 steps: 282 lr: 6.4000000000000001e-06 evaluation
 reward: 6.32
 episode: 1785 score: 6.0 memory length: 410076 epsilon: 0.3860475400101335
 steps: 360 lr: 6.4000000000000001e-06 evaluation reward: 6.35
 episode: 1786 score: 6.0 memory length: 410402 epsilon:
 0.38540206001012944 steps: 326 lr: 6.4000000000000001e-06 evaluation
 reward: 6.37
 episode: 1787 score: 4.0 memory length: 410681 epsilon:
 0.38484964001012595 steps: 279 lr: 6.4000000000000001e-06 evaluation
 reward: 6.3
 episode: 1788 score: 6.0 memory length: 411043 epsilon: 0.3841328800101214
 steps: 362 lr: 6.4000000000000001e-06 evaluation reward: 6.32
 episode: 1789 score: 6.0 memory length: 411386 epsilon: 0.3834537400101171
 steps: 343 lr: 6.4000000000000001e-06 evaluation reward: 6.31
 episode: 1790 score: 6.0 memory length: 411747 epsilon: 0.3827389600101126

```

steps: 361    lr: 6.4000000000000001e-06    evaluation reward: 6.32
episode: 1791  score: 5.0  memory length: 412059  epsilon: 0.3821212000101087
steps: 312    lr: 6.4000000000000001e-06    evaluation reward: 6.29
episode: 1792  score: 4.0  memory length: 412319  epsilon:
0.38160640001010543  steps: 260    lr: 6.4000000000000001e-06    evaluation
reward: 6.2
episode: 1793  score: 4.0  memory length: 412562  epsilon: 0.3811252600101024
steps: 243    lr: 6.4000000000000001e-06    evaluation reward: 6.18
episode: 1794  score: 8.0  memory length: 412964  epsilon:
0.38032930001009735  steps: 402    lr: 6.4000000000000001e-06    evaluation
reward: 6.23
episode: 1795  score: 10.0  memory length: 413469  epsilon: 0.379329400010091
steps: 505    lr: 6.4000000000000001e-06    evaluation reward: 6.28
episode: 1796  score: 7.0  memory length: 413836  epsilon: 0.3786027400100864
steps: 367    lr: 6.4000000000000001e-06    evaluation reward: 6.3
episode: 1797  score: 9.0  memory length: 414310  epsilon: 0.3776642200100805
steps: 474    lr: 6.4000000000000001e-06    evaluation reward: 6.37
episode: 1798  score: 6.0  memory length: 414686  epsilon: 0.3769197400100758
steps: 376    lr: 6.4000000000000001e-06    evaluation reward: 6.35
episode: 1799  score: 19.0  memory length: 415333  epsilon:
0.37563868001006767  steps: 647    lr: 6.4000000000000001e-06    evaluation
reward: 6.49
episode: 1800  score: 6.0  memory length: 415693  epsilon:
0.37492588001006316  steps: 360    lr: 6.4000000000000001e-06    evaluation
reward: 6.49
episode: 1801  score: 6.0  memory length: 416048  epsilon: 0.3742229800100587
steps: 355    lr: 6.4000000000000001e-06    evaluation reward: 6.5
episode: 1802  score: 6.0  memory length: 416397  epsilon:
0.37353196001005434  steps: 349    lr: 6.4000000000000001e-06    evaluation
reward: 6.51
episode: 1803  score: 9.0  memory length: 416854  epsilon: 0.3726271000100486
steps: 457    lr: 6.4000000000000001e-06    evaluation reward: 6.55
episode: 1804  score: 5.0  memory length: 417159  epsilon: 0.3720232000100448
steps: 305    lr: 6.4000000000000001e-06    evaluation reward: 6.55
episode: 1805  score: 5.0  memory length: 417441  epsilon:
0.37146484001004126  steps: 282    lr: 6.4000000000000001e-06    evaluation
reward: 6.52
episode: 1806  score: 5.0  memory length: 417735  epsilon: 0.3708827200100376
steps: 294    lr: 6.4000000000000001e-06    evaluation reward: 6.5
episode: 1807  score: 7.0  memory length: 418097  epsilon:
0.37016596001003305  steps: 362    lr: 6.4000000000000001e-06    evaluation
reward: 6.48
episode: 1808  score: 6.0  memory length: 418429  epsilon: 0.3695086000100289
steps: 332    lr: 6.4000000000000001e-06    evaluation reward: 6.47
episode: 1809  score: 6.0  memory length: 418804  epsilon: 0.3687661000100242
steps: 375    lr: 6.4000000000000001e-06    evaluation reward: 6.44
episode: 1810  score: 4.0  memory length: 419064  epsilon:
0.36825130001002093  steps: 260    lr: 6.4000000000000001e-06    evaluation

```

```

reward: 6.44
episode: 1811  score: 6.0  memory length: 419424  epsilon: 0.3675385000100164
steps: 360  lr: 6.4000000000000001e-06  evaluation reward: 6.47
episode: 1812  score: 5.0  memory length: 419706  epsilon: 0.3669801400100129
steps: 282  lr: 6.4000000000000001e-06  evaluation reward: 6.48
episode: 1813  score: 1.0  memory length: 419858  epsilon: 0.366679180010011
steps: 152  lr: 6.4000000000000001e-06  evaluation reward: 6.41
episode: 1814  score: 7.0  memory length: 420253  epsilon:
0.36589708001000604  steps: 395  lr: 6.4000000000000001e-06  evaluation
reward: 6.33
episode: 1815  score: 4.0  memory length: 420513  epsilon: 0.3653822800100028
steps: 260  lr: 6.4000000000000001e-06  evaluation reward: 6.32
episode: 1816  score: 4.0  memory length: 420775  epsilon: 0.3648635200099995
steps: 262  lr: 6.4000000000000001e-06  evaluation reward: 6.29
episode: 1817  score: 5.0  memory length: 421056  epsilon: 0.364307140009996
steps: 281  lr: 6.4000000000000001e-06  evaluation reward: 6.3
episode: 1818  score: 5.0  memory length: 421344  epsilon:
0.36373690000999237  steps: 288  lr: 6.4000000000000001e-06  evaluation
reward: 6.31
episode: 1819  score: 4.0  memory length: 421604  epsilon: 0.3632221000099891
steps: 260  lr: 6.4000000000000001e-06  evaluation reward: 6.32
episode: 1820  score: 3.0  memory length: 421831  epsilon:
0.36277264000998627  steps: 227  lr: 6.4000000000000001e-06  evaluation
reward: 6.31
episode: 1821  score: 18.0  memory length: 422472  epsilon:
0.36150346000997824  steps: 641  lr: 6.4000000000000001e-06  evaluation
reward: 6.44
episode: 1822  score: 8.0  memory length: 422909  epsilon:
0.36063820000997276  steps: 437  lr: 6.4000000000000001e-06  evaluation
reward: 6.46
episode: 1823  score: 7.0  memory length: 423266  epsilon: 0.3599313400099683
steps: 357  lr: 6.4000000000000001e-06  evaluation reward: 6.49
episode: 1824  score: 11.0  memory length: 423807  epsilon:
0.3588601600099615  steps: 541  lr: 6.4000000000000001e-06  evaluation
reward: 6.53
episode: 1825  score: 10.0  memory length: 424282  epsilon:
0.35791966000995556  steps: 475  lr: 6.4000000000000001e-06  evaluation
reward: 6.59
episode: 1826  score: 8.0  memory length: 424724  epsilon: 0.35704450000995
steps: 442  lr: 6.4000000000000001e-06  evaluation reward: 6.64
episode: 1827  score: 10.0  memory length: 425210  epsilon:
0.35608222000994394  steps: 486  lr: 6.4000000000000001e-06  evaluation
reward: 6.7
episode: 1828  score: 8.0  memory length: 425638  epsilon: 0.3552347800099386
steps: 428  lr: 6.4000000000000001e-06  evaluation reward: 6.72
episode: 1829  score: 7.0  memory length: 426048  epsilon:
0.35442298000993344  steps: 410  lr: 6.4000000000000001e-06  evaluation
reward: 6.73

```

episode: 1830 score: 5.0 memory length: 426342 epsilon:
 0.35384086000992976 steps: 294 lr: 6.400000000000001e-06 evaluation
 reward: 6.73
 episode: 1831 score: 11.0 memory length: 426865 epsilon:
 0.3528053200099232 steps: 523 lr: 6.400000000000001e-06 evaluation
 reward: 6.79
 episode: 1832 score: 5.0 memory length: 427139 epsilon: 0.3522628000099198
 steps: 274 lr: 6.400000000000001e-06 evaluation reward: 6.81
 episode: 1833 score: 12.0 memory length: 427628 epsilon:
 0.35129458000991365 steps: 489 lr: 6.400000000000001e-06 evaluation
 reward: 6.86
 episode: 1834 score: 9.0 memory length: 428040 epsilon: 0.3504788200099085
 steps: 412 lr: 6.400000000000001e-06 evaluation reward: 6.91
 episode: 1835 score: 8.0 memory length: 428497 epsilon:
 0.34957396000990276 steps: 457 lr: 6.400000000000001e-06 evaluation
 reward: 6.91
 episode: 1836 score: 12.0 memory length: 429047 epsilon:
 0.34848496000989587 steps: 550 lr: 6.400000000000001e-06 evaluation
 reward: 6.99
 episode: 1837 score: 10.0 memory length: 429581 epsilon:
 0.3474276400098892 steps: 534 lr: 6.400000000000001e-06 evaluation
 reward: 7.03
 episode: 1838 score: 10.0 memory length: 430093 epsilon:
 0.34641388000988277 steps: 512 lr: 6.400000000000001e-06 evaluation
 reward: 7.06
 episode: 1839 score: 3.0 memory length: 430307 epsilon: 0.3459901600098801
 steps: 214 lr: 6.400000000000001e-06 evaluation reward: 7.02
 episode: 1840 score: 11.0 memory length: 430816 epsilon:
 0.3449823400098737 steps: 509 lr: 6.400000000000001e-06 evaluation
 reward: 7.06
 episode: 1841 score: 6.0 memory length: 431189 epsilon:
 0.34424380000986904 steps: 373 lr: 6.400000000000001e-06 evaluation
 reward: 7.05
 episode: 1842 score: 6.0 memory length: 431546 epsilon:
 0.34353694000986457 steps: 357 lr: 6.400000000000001e-06 evaluation
 reward: 7.02
 episode: 1843 score: 2.0 memory length: 431729 epsilon: 0.3431746000098623
 steps: 183 lr: 6.400000000000001e-06 evaluation reward: 6.95
 episode: 1844 score: 4.0 memory length: 431969 epsilon:
 0.34269940000985927 steps: 240 lr: 6.400000000000001e-06 evaluation
 reward: 6.95
 episode: 1845 score: 11.0 memory length: 432537 epsilon:
 0.34157476000985215 steps: 568 lr: 6.400000000000001e-06 evaluation
 reward: 7.01
 episode: 1846 score: 8.0 memory length: 432971 epsilon: 0.3407154400098467
 steps: 434 lr: 6.400000000000001e-06 evaluation reward: 7.04
 episode: 1847 score: 7.0 memory length: 433379 epsilon: 0.3399076000098416
 steps: 408 lr: 6.400000000000001e-06 evaluation reward: 7.03

episode: 1848 score: 7.0 memory length: 433749 epsilon:
 0.33917500000983697 steps: 370 lr: 6.400000000000001e-06 evaluation
 reward: 7.04
 episode: 1849 score: 6.0 memory length: 434057 epsilon: 0.3385651600098331
 steps: 308 lr: 6.400000000000001e-06 evaluation reward: 7.07
 episode: 1850 score: 7.0 memory length: 434425 epsilon: 0.3378365200098285
 steps: 368 lr: 6.400000000000001e-06 evaluation reward: 7.08
 episode: 1851 score: 7.0 memory length: 434804 epsilon:
 0.33708610000982375 steps: 379 lr: 6.400000000000001e-06 evaluation
 reward: 7.11
 episode: 1852 score: 6.0 memory length: 435159 epsilon: 0.3363832000098193
 steps: 355 lr: 6.400000000000001e-06 evaluation reward: 7.15
 episode: 1853 score: 13.0 memory length: 435707 epsilon:
 0.33529816000981244 steps: 548 lr: 6.400000000000001e-06 evaluation
 reward: 7.22
 episode: 1854 score: 6.0 memory length: 436066 epsilon:
 0.33458734000980794 steps: 359 lr: 6.400000000000001e-06 evaluation
 reward: 7.22
 episode: 1855 score: 6.0 memory length: 436404 epsilon: 0.3339181000098037
 steps: 338 lr: 6.400000000000001e-06 evaluation reward: 7.24
 episode: 1856 score: 8.0 memory length: 436849 epsilon:
 0.33303700000979813 steps: 445 lr: 6.400000000000001e-06 evaluation
 reward: 7.25
 episode: 1857 score: 9.0 memory length: 437341 epsilon:
 0.33206284000979197 steps: 492 lr: 6.400000000000001e-06 evaluation
 reward: 7.25
 episode: 1858 score: 6.0 memory length: 437694 epsilon:
 0.33136390000978755 steps: 353 lr: 6.400000000000001e-06 evaluation
 reward: 7.19
 episode: 1859 score: 13.0 memory length: 438276 epsilon:
 0.33021154000978026 steps: 582 lr: 6.400000000000001e-06 evaluation
 reward: 7.23
 episode: 1860 score: 5.0 memory length: 438583 epsilon: 0.3296036800097764
 steps: 307 lr: 6.400000000000001e-06 evaluation reward: 7.18
 episode: 1861 score: 7.0 memory length: 438976 epsilon: 0.3288255400097715
 steps: 393 lr: 6.400000000000001e-06 evaluation reward: 7.23
 episode: 1862 score: 7.0 memory length: 439354 epsilon:
 0.32807710000976675 steps: 378 lr: 6.400000000000001e-06 evaluation
 reward: 7.24
 episode: 1863 score: 7.0 memory length: 439745 epsilon:
 0.32730292000976186 steps: 391 lr: 6.400000000000001e-06 evaluation
 reward: 7.25
 episode: 1864 score: 8.0 memory length: 440176 epsilon:
 0.32644954000975646 steps: 431 lr: 6.400000000000001e-06 evaluation
 reward: 7.25
 episode: 1865 score: 5.0 memory length: 440504 epsilon:
 0.32580010000975235 steps: 328 lr: 6.400000000000001e-06 evaluation
 reward: 7.24

episode: 1866 score: 6.0 memory length: 440832 epsilon:
 0.32515066000974824 steps: 328 lr: 6.400000000000001e-06 evaluation
 reward: 7.1
 episode: 1867 score: 2.0 memory length: 441033 epsilon: 0.3247526800097457
 steps: 201 lr: 6.400000000000001e-06 evaluation reward: 7.07
 episode: 1868 score: 7.0 memory length: 441395 epsilon: 0.3240359200097412
 steps: 362 lr: 6.400000000000001e-06 evaluation reward: 7.09
 episode: 1869 score: 8.0 memory length: 441870 epsilon:
 0.32309542000973523 steps: 475 lr: 6.400000000000001e-06 evaluation
 reward: 7.08
 episode: 1870 score: 4.0 memory length: 442129 epsilon: 0.322582600009732
 steps: 259 lr: 6.400000000000001e-06 evaluation reward: 7.04
 episode: 1871 score: 12.0 memory length: 442677 epsilon:
 0.3214975600097251 steps: 548 lr: 6.400000000000001e-06 evaluation
 reward: 7.07
 episode: 1872 score: 5.0 memory length: 442986 epsilon:
 0.32088574000972125 steps: 309 lr: 6.400000000000001e-06 evaluation
 reward: 7.03
 episode: 1873 score: 10.0 memory length: 443494 epsilon:
 0.3198799000097149 steps: 508 lr: 6.400000000000001e-06 evaluation
 reward: 7.07
 episode: 1874 score: 8.0 memory length: 443919 epsilon:
 0.31903840000970957 steps: 425 lr: 6.400000000000001e-06 evaluation
 reward: 7.1
 episode: 1875 score: 4.0 memory length: 444216 epsilon:
 0.31845034000970585 steps: 297 lr: 6.400000000000001e-06 evaluation
 reward: 6.98
 episode: 1876 score: 6.0 memory length: 444561 epsilon: 0.3177672400097015
 steps: 345 lr: 6.400000000000001e-06 evaluation reward: 6.98
 episode: 1877 score: 8.0 memory length: 445019 epsilon: 0.3168604000096958
 steps: 458 lr: 6.400000000000001e-06 evaluation reward: 6.98
 episode: 1878 score: 10.0 memory length: 445490 epsilon:
 0.3159278200096899 steps: 471 lr: 6.400000000000001e-06 evaluation
 reward: 6.99
 episode: 1879 score: 12.0 memory length: 445962 epsilon: 0.314993260009684
 steps: 472 lr: 6.400000000000001e-06 evaluation reward: 7.07
 episode: 1880 score: 11.0 memory length: 446498 epsilon:
 0.31393198000967726 steps: 536 lr: 6.400000000000001e-06 evaluation
 reward: 7.09
 episode: 1881 score: 8.0 memory length: 446922 epsilon:
 0.31309246000967195 steps: 424 lr: 6.400000000000001e-06 evaluation
 reward: 7.11
 episode: 1882 score: 4.0 memory length: 447167 epsilon: 0.3126073600096689
 steps: 245 lr: 6.400000000000001e-06 evaluation reward: 7.12
 episode: 1883 score: 5.0 memory length: 447475 epsilon: 0.311997520009665
 steps: 308 lr: 6.400000000000001e-06 evaluation reward: 7.09
 episode: 1884 score: 10.0 memory length: 447999 epsilon:
 0.31096000000965845 steps: 524 lr: 6.400000000000001e-06 evaluation


```

reward: 7.15
episode: 1885  score: 9.0  memory length: 448470  epsilon:
0.31002742000965255  steps: 471  lr: 6.400000000000001e-06  evaluation
reward: 7.18
episode: 1886  score: 15.0  memory length: 449058  epsilon:
0.3088631800096452  steps: 588  lr: 6.400000000000001e-06  evaluation
reward: 7.27
episode: 1887  score: 10.0  memory length: 449527  epsilon:
0.3079345600096393  steps: 469  lr: 6.400000000000001e-06  evaluation
reward: 7.33
episode: 1888  score: 22.0  memory length: 450147  epsilon:
0.30670696000963155  steps: 620  lr: 6.400000000000001e-06  evaluation
reward: 7.49
episode: 1889  score: 9.0  memory length: 450594  epsilon:
0.30582190000962595  steps: 447  lr: 6.400000000000001e-06  evaluation
reward: 7.52
episode: 1890  score: 6.0  memory length: 450950  epsilon: 0.3051170200096215
steps: 356  lr: 6.400000000000001e-06  evaluation reward: 7.52
episode: 1891  score: 9.0  memory length: 451422  epsilon: 0.3041824600096156
steps: 472  lr: 6.400000000000001e-06  evaluation reward: 7.56
episode: 1892  score: 5.0  memory length: 451708  epsilon: 0.303616180009612
steps: 286  lr: 6.400000000000001e-06  evaluation reward: 7.57
episode: 1893  score: 6.0  memory length: 452035  epsilon: 0.3029687200096079
steps: 327  lr: 6.400000000000001e-06  evaluation reward: 7.59
episode: 1894  score: 9.0  memory length: 452505  epsilon: 0.302038120009602
steps: 470  lr: 6.400000000000001e-06  evaluation reward: 7.6
episode: 1895  score: 13.0  memory length: 453011  epsilon:
0.30103624000959567  steps: 506  lr: 6.400000000000001e-06  evaluation
reward: 7.63
episode: 1896  score: 11.0  memory length: 453561  epsilon:
0.2999472400095888  steps: 550  lr: 6.400000000000001e-06  evaluation
reward: 7.67
episode: 1897  score: 4.0  memory length: 453821  epsilon: 0.2994324400095855
steps: 260  lr: 6.400000000000001e-06  evaluation reward: 7.62
episode: 1898  score: 3.0  memory length: 454032  epsilon: 0.2990146600095829
steps: 211  lr: 6.400000000000001e-06  evaluation reward: 7.59
episode: 1899  score: 10.0  memory length: 454510  epsilon:
0.2980682200095769  steps: 478  lr: 6.400000000000001e-06  evaluation
reward: 7.5
episode: 1900  score: 6.0  memory length: 454840  epsilon:
0.29741482000957276  steps: 330  lr: 6.400000000000001e-06  evaluation
reward: 7.5
episode: 1901  score: 9.0  memory length: 455301  epsilon: 0.296502040009567
steps: 461  lr: 6.400000000000001e-06  evaluation reward: 7.53
episode: 1902  score: 8.0  memory length: 455757  epsilon:
0.29559916000956127  steps: 456  lr: 6.400000000000001e-06  evaluation
reward: 7.55
episode: 1903  score: 2.0  memory length: 455940  epsilon: 0.295236820009559

```

```

steps: 183    lr: 6.4000000000000001e-06    evaluation reward: 7.48
episode: 1904  score: 8.0    memory length: 456328    epsilon: 0.2944685800095541
steps: 388    lr: 6.4000000000000001e-06    evaluation reward: 7.51
episode: 1905  score: 8.0    memory length: 456735    epsilon: 0.293662720009549
steps: 407    lr: 6.4000000000000001e-06    evaluation reward: 7.54
episode: 1906  score: 11.0   memory length: 457214    epsilon: 0.292714300009543
steps: 479    lr: 6.4000000000000001e-06    evaluation reward: 7.6
episode: 1907  score: 6.0    memory length: 457570    epsilon:
0.29200942000953856    steps: 356    lr: 6.4000000000000001e-06    evaluation
reward: 7.59
episode: 1908  score: 5.0    memory length: 457901    epsilon: 0.2913540400095344
steps: 331    lr: 6.4000000000000001e-06    evaluation reward: 7.58
episode: 1909  score: 6.0    memory length: 458248    epsilon:
0.29066698000953006    steps: 347    lr: 6.4000000000000001e-06    evaluation
reward: 7.58
episode: 1910  score: 4.0    memory length: 458508    epsilon: 0.2901521800095268
steps: 260    lr: 6.4000000000000001e-06    evaluation reward: 7.58
episode: 1911  score: 5.0    memory length: 458821    epsilon: 0.2895324400095229
steps: 313    lr: 6.4000000000000001e-06    evaluation reward: 7.57
episode: 1912  score: 12.0   memory length: 459410    epsilon:
0.2883662200095155    steps: 589    lr: 6.4000000000000001e-06    evaluation
reward: 7.64
episode: 1913  score: 15.0   memory length: 459952    epsilon:
0.2872930600095087    steps: 542    lr: 6.4000000000000001e-06    evaluation
reward: 7.78
episode: 1914  score: 8.0    memory length: 460412    epsilon:
0.28638226000950295    steps: 460    lr: 6.4000000000000001e-06    evaluation
reward: 7.79
episode: 1915  score: 10.0   memory length: 460900    epsilon:
0.28541602000949684    steps: 488    lr: 6.4000000000000001e-06    evaluation
reward: 7.85
episode: 1916  score: 8.0    memory length: 461338    epsilon:
0.28454878000949135    steps: 438    lr: 6.4000000000000001e-06    evaluation
reward: 7.89
episode: 1917  score: 6.0    memory length: 461736    epsilon:
0.28376074000948637    steps: 398    lr: 6.4000000000000001e-06    evaluation
reward: 7.9
episode: 1918  score: 7.0    memory length: 462127    epsilon:
0.28298656000948147    steps: 391    lr: 6.4000000000000001e-06    evaluation
reward: 7.92
episode: 1919  score: 6.0    memory length: 462486    epsilon:
0.28227574000947697    steps: 359    lr: 6.4000000000000001e-06    evaluation
reward: 7.94
episode: 1920  score: 7.0    memory length: 462892    epsilon: 0.2814718600094719
steps: 406    lr: 6.4000000000000001e-06    evaluation reward: 7.98
episode: 1921  score: 6.0    memory length: 463235    epsilon: 0.2807927200094676
steps: 343    lr: 6.4000000000000001e-06    evaluation reward: 7.86
episode: 1922  score: 9.0    memory length: 463686    epsilon:

```

0.27989974000946194 steps: 451 lr: 6.4000000000000001e-06 evaluation
 reward: 7.87
 episode: 1923 score: 12.0 memory length: 464205 epsilon:
 0.27887212000945544 steps: 519 lr: 6.4000000000000001e-06 evaluation
 reward: 7.92
 episode: 1924 score: 12.0 memory length: 464719 epsilon: 0.277854400009449
 steps: 514 lr: 6.4000000000000001e-06 evaluation reward: 7.93
 episode: 1925 score: 6.0 memory length: 465076 epsilon: 0.2771475400094445
 steps: 357 lr: 6.4000000000000001e-06 evaluation reward: 7.89
 episode: 1926 score: 11.0 memory length: 465502 epsilon:
 0.2763040600094392 steps: 426 lr: 6.4000000000000001e-06 evaluation
 reward: 7.92
 episode: 1927 score: 16.0 memory length: 466114 epsilon:
 0.2750923000094315 steps: 612 lr: 6.4000000000000001e-06 evaluation
 reward: 7.98
 episode: 1928 score: 5.0 memory length: 466397 epsilon: 0.274531960009428
 steps: 283 lr: 6.4000000000000001e-06 evaluation reward: 7.95
 episode: 1929 score: 10.0 memory length: 466893 epsilon:
 0.27354988000942176 steps: 496 lr: 6.4000000000000001e-06 evaluation
 reward: 7.98
 episode: 1930 score: 6.0 memory length: 467249 epsilon: 0.2728450000094173
 steps: 356 lr: 6.4000000000000001e-06 evaluation reward: 7.99
 episode: 1931 score: 11.0 memory length: 467827 epsilon:
 0.27170056000941006 steps: 578 lr: 6.4000000000000001e-06 evaluation
 reward: 7.99
 episode: 1932 score: 14.0 memory length: 468423 epsilon:
 0.2705204800094026 steps: 596 lr: 6.4000000000000001e-06 evaluation
 reward: 8.08
 episode: 1933 score: 11.0 memory length: 468987 epsilon:
 0.26940376000939553 steps: 564 lr: 6.4000000000000001e-06 evaluation
 reward: 8.07
 episode: 1934 score: 6.0 memory length: 469326 epsilon: 0.2687325400093913
 steps: 339 lr: 6.4000000000000001e-06 evaluation reward: 8.04
 episode: 1935 score: 5.0 memory length: 469677 epsilon: 0.2680375600093869
 steps: 351 lr: 6.4000000000000001e-06 evaluation reward: 8.01
 episode: 1936 score: 6.0 memory length: 470018 epsilon: 0.2673623800093826
 steps: 341 lr: 6.4000000000000001e-06 evaluation reward: 7.95
 episode: 1937 score: 8.0 memory length: 470457 epsilon: 0.2664931600093771
 steps: 439 lr: 6.4000000000000001e-06 evaluation reward: 7.93
 episode: 1938 score: 8.0 memory length: 470864 epsilon: 0.265687300009372
 steps: 407 lr: 6.4000000000000001e-06 evaluation reward: 7.91
 episode: 1939 score: 6.0 memory length: 471227 epsilon:
 0.26496856000936747 steps: 363 lr: 6.4000000000000001e-06 evaluation
 reward: 7.94
 episode: 1940 score: 6.0 memory length: 471568 epsilon: 0.2642933800093632
 steps: 341 lr: 6.4000000000000001e-06 evaluation reward: 7.89
 episode: 1941 score: 6.0 memory length: 471908 epsilon:
 0.26362018000935894 steps: 340 lr: 6.4000000000000001e-06 evaluation

```

reward: 7.89
episode: 1942  score: 5.0  memory length: 472217  epsilon:
0.26300836000935507  steps: 309  lr: 6.400000000000001e-06  evaluation
reward: 7.88
episode: 1943  score: 13.0  memory length: 472835  epsilon:
0.2617847200093473  steps: 618  lr: 6.400000000000001e-06  evaluation
reward: 7.99
episode: 1944  score: 14.0  memory length: 473464  epsilon:
0.26053930000933945  steps: 629  lr: 6.400000000000001e-06  evaluation
reward: 8.09
episode: 1945  score: 12.0  memory length: 474054  epsilon:
0.25937110000933206  steps: 590  lr: 6.400000000000001e-06  evaluation
reward: 8.1
episode: 1946  score: 9.0  memory length: 474530  epsilon: 0.2584286200093261
steps: 476  lr: 6.400000000000001e-06  evaluation reward: 8.11
episode: 1947  score: 6.0  memory length: 474836  epsilon:
0.25782274000932226  steps: 306  lr: 6.400000000000001e-06  evaluation
reward: 8.1
episode: 1948  score: 6.0  memory length: 475159  epsilon: 0.2571832000093182
steps: 323  lr: 6.400000000000001e-06  evaluation reward: 8.09
episode: 1949  score: 5.0  memory length: 475469  epsilon:
0.25656940000931433  steps: 310  lr: 6.400000000000001e-06  evaluation
reward: 8.08
episode: 1950  score: 10.0  memory length: 475992  epsilon:
0.2555338600093078  steps: 523  lr: 6.400000000000001e-06  evaluation
reward: 8.11
episode: 1951  score: 7.0  memory length: 476354  epsilon:
0.25481710000930324  steps: 362  lr: 6.400000000000001e-06  evaluation
reward: 8.11
episode: 1952  score: 8.0  memory length: 476810  epsilon:
0.25391422000929753  steps: 456  lr: 6.400000000000001e-06  evaluation
reward: 8.13
episode: 1953  score: 12.0  memory length: 477402  epsilon:
0.2527420600092901  steps: 592  lr: 6.400000000000001e-06  evaluation
reward: 8.12
episode: 1954  score: 6.0  memory length: 477781  epsilon:
0.25199164000928537  steps: 379  lr: 6.400000000000001e-06  evaluation
reward: 8.12
episode: 1955  score: 7.0  memory length: 478209  epsilon: 0.25114420000928
steps: 428  lr: 6.400000000000001e-06  evaluation reward: 8.13
episode: 1956  score: 6.0  memory length: 478547  epsilon:
0.25047496000927577  steps: 338  lr: 6.400000000000001e-06  evaluation
reward: 8.11
episode: 1957  score: 15.0  memory length: 479102  epsilon:
0.24937606000926882  steps: 555  lr: 6.400000000000001e-06  evaluation
reward: 8.17
episode: 1958  score: 18.0  memory length: 479773  epsilon:
0.2480474800092604  steps: 671  lr: 6.400000000000001e-06  evaluation

```

```

reward: 8.29
episode: 1959  score: 12.0  memory length: 480354  epsilon:
0.24689710000925313  steps: 581  lr: 6.4000000000000001e-06  evaluation
reward: 8.28
episode: 1960  score: 8.0  memory length: 480746  epsilon:
0.24612094000924822  steps: 392  lr: 6.4000000000000001e-06  evaluation
reward: 8.31
episode: 1961  score: 21.0  memory length: 481391  epsilon:
0.24484384000924014  steps: 645  lr: 6.4000000000000001e-06  evaluation
reward: 8.45
episode: 1962  score: 9.0  memory length: 481865  epsilon: 0.2439053200092342
steps: 474  lr: 6.4000000000000001e-06  evaluation reward: 8.47
episode: 1963  score: 5.0  memory length: 482213  epsilon:
0.24321628000922985  steps: 348  lr: 6.4000000000000001e-06  evaluation
reward: 8.45
episode: 1964  score: 6.0  memory length: 482536  epsilon: 0.2425767400092258
steps: 323  lr: 6.4000000000000001e-06  evaluation reward: 8.43
episode: 1965  score: 10.0  memory length: 483056  epsilon:
0.24154714000921929  steps: 520  lr: 6.4000000000000001e-06  evaluation
reward: 8.48
episode: 1966  score: 9.0  memory length: 483400  epsilon:
0.24086602000921498  steps: 344  lr: 6.4000000000000001e-06  evaluation
reward: 8.51
episode: 1967  score: 12.0  memory length: 483834  epsilon:
0.24000670000920954  steps: 434  lr: 6.4000000000000001e-06  evaluation
reward: 8.61
episode: 1968  score: 8.0  memory length: 484235  epsilon:
0.23921272000920452  steps: 401  lr: 6.4000000000000001e-06  evaluation
reward: 8.62
episode: 1969  score: 9.0  memory length: 484690  epsilon:
0.23831182000919882  steps: 455  lr: 6.4000000000000001e-06  evaluation
reward: 8.63
episode: 1970  score: 9.0  memory length: 485128  epsilon:
0.23744458000919333  steps: 438  lr: 6.4000000000000001e-06  evaluation
reward: 8.68
episode: 1971  score: 4.0  memory length: 485386  epsilon: 0.2369337400091901
steps: 258  lr: 6.4000000000000001e-06  evaluation reward: 8.6
episode: 1972  score: 11.0  memory length: 485813  epsilon:
0.23608828000918475  steps: 427  lr: 6.4000000000000001e-06  evaluation
reward: 8.66
episode: 1973  score: 2.0  memory length: 485995  epsilon:
0.23572792000918247  steps: 182  lr: 6.4000000000000001e-06  evaluation
reward: 8.58
episode: 1974  score: 7.0  memory length: 486352  epsilon: 0.235021060009178
steps: 357  lr: 6.4000000000000001e-06  evaluation reward: 8.57
episode: 1975  score: 5.0  memory length: 486660  epsilon:
0.23441122000917414  steps: 308  lr: 6.4000000000000001e-06  evaluation
reward: 8.58

```

episode: 1976	score: 11.0	memory length: 487045	epsilon: 0.23364892000916931	steps: 385	lr: 6.400000000000001e-06	evaluation
reward: 8.63						
episode: 1977	score: 5.0	memory length: 487378	epsilon: 0.23298958000916514	steps: 333	lr: 6.400000000000001e-06	evaluation
reward: 8.6						
episode: 1978	score: 8.0	memory length: 487768	epsilon: 0.23221738000916026	steps: 390	lr: 6.400000000000001e-06	evaluation
reward: 8.58						
episode: 1979	score: 14.0	memory length: 488326	epsilon: 0.23111254000915327	steps: 558	lr: 6.400000000000001e-06	evaluation
reward: 8.6						
episode: 1980	score: 8.0	memory length: 488737	epsilon: 0.23029876000914812	steps: 411	lr: 6.400000000000001e-06	evaluation
reward: 8.57						
episode: 1981	score: 7.0	memory length: 489078	epsilon: 0.22962358000914385	steps: 341	lr: 6.400000000000001e-06	evaluation
reward: 8.56						
episode: 1982	score: 21.0	memory length: 489714	epsilon: 0.22836430000913588	steps: 636	lr: 6.400000000000001e-06	evaluation
reward: 8.73						
episode: 1983	score: 11.0	memory length: 490285	epsilon: 0.22723372000912873	steps: 571	lr: 6.400000000000001e-06	evaluation
reward: 8.79						
episode: 1984	score: 10.0	memory length: 490791	epsilon: 0.2262318400091224	steps: 506	lr: 6.400000000000001e-06	evaluation
reward: 8.79						
episode: 1985	score: 5.0	memory length: 491120	epsilon: 0.22558042000911827	steps: 329	lr: 6.400000000000001e-06	evaluation
reward: 8.75						
episode: 1986	score: 14.0	memory length: 491642	epsilon: 0.22454686000911173	steps: 522	lr: 6.400000000000001e-06	evaluation
reward: 8.74						
episode: 1987	score: 8.0	memory length: 492045	epsilon: 0.22374892000910668	steps: 403	lr: 6.400000000000001e-06	evaluation
reward: 8.72						
episode: 1988	score: 12.0	memory length: 492550	epsilon: 0.22274902000910035	steps: 505	lr: 6.400000000000001e-06	evaluation
reward: 8.62						
episode: 1989	score: 8.0	memory length: 492960	epsilon: 0.22193722000909522	steps: 410	lr: 6.400000000000001e-06	evaluation
reward: 8.61						
episode: 1990	score: 4.0	memory length: 493220	epsilon: 0.22142242000909196	steps: 260	lr: 6.400000000000001e-06	evaluation
reward: 8.59						
episode: 1991	score: 14.0	memory length: 493826	epsilon: 0.22022254000908437	steps: 606	lr: 6.400000000000001e-06	evaluation
reward: 8.64						

episode: 1992	score: 8.0	memory length: 494232	epsilon: 0.21941866000907928	steps: 406	lr: 6.400000000000001e-06	evaluation
reward: 8.67						
episode: 1993	score: 8.0	memory length: 494625	epsilon: 0.21864052000907436	steps: 393	lr: 6.400000000000001e-06	evaluation
reward: 8.69						
episode: 1994	score: 13.0	memory length: 495267	epsilon: 0.21736936000906631	steps: 642	lr: 6.400000000000001e-06	evaluation
reward: 8.73						
episode: 1995	score: 7.0	memory length: 495645	epsilon: 0.21662092000906158	steps: 378	lr: 6.400000000000001e-06	evaluation
reward: 8.67						
episode: 1996	score: 10.0	memory length: 496097	epsilon: 0.21572596000905592	steps: 452	lr: 6.400000000000001e-06	evaluation
reward: 8.66						
episode: 1997	score: 10.0	memory length: 496590	epsilon: 0.21474982000904974	steps: 493	lr: 6.400000000000001e-06	evaluation
reward: 8.72						
episode: 1998	score: 6.0	memory length: 496932	epsilon: 0.21407266000904546	steps: 342	lr: 6.400000000000001e-06	evaluation
reward: 8.75						
episode: 1999	score: 9.0	memory length: 497427	epsilon: 0.21309256000903926	steps: 495	lr: 6.400000000000001e-06	evaluation
reward: 8.74						
episode: 2000	score: 7.0	memory length: 497832	epsilon: 0.21229066000903418	steps: 405	lr: 6.400000000000001e-06	evaluation
reward: 8.75						
episode: 2001	score: 5.0	memory length: 498143	epsilon: 0.2116748800090303	steps: 311	lr: 6.400000000000001e-06	evaluation
reward: 8.71						
episode: 2002	score: 14.0	memory length: 498684	epsilon: 0.2106037000090235	steps: 541	lr: 6.400000000000001e-06	evaluation
reward: 8.77						
episode: 2003	score: 4.0	memory length: 498946	epsilon: 0.21008494000902023	steps: 262	lr: 6.400000000000001e-06	evaluation
reward: 8.79						
episode: 2004	score: 9.0	memory length: 499400	epsilon: 0.20918602000901454	steps: 454	lr: 6.400000000000001e-06	evaluation
reward: 8.8						
episode: 2005	score: 5.0	memory length: 499729	epsilon: 0.20853460000901042	steps: 329	lr: 6.400000000000001e-06	evaluation
reward: 8.77						
episode: 2006	score: 11.0	memory length: 500243	epsilon: 0.20751688000900398	steps: 514	lr: 2.560000000000001e-06	evaluation
reward: 8.77						
episode: 2007	score: 7.0	memory length: 500652	epsilon: 0.20670706000899886	steps: 409	lr: 2.560000000000001e-06	evaluation
reward: 8.78						
episode: 2008	score: 10.0	memory length: 501126	epsilon:			

0.20576854000899292 steps: 474 lr: 2.5600000000000001e-06 evaluation
 reward: 8.83
 episode: 2009 score: 9.0 memory length: 501599 epsilon: 0.204832000008987
 steps: 473 lr: 2.5600000000000001e-06 evaluation reward: 8.86
 episode: 2010 score: 7.0 memory length: 501969 epsilon:
 0.20409940000898236 steps: 370 lr: 2.5600000000000001e-06 evaluation
 reward: 8.89
 episode: 2011 score: 6.0 memory length: 502305 epsilon:
 0.20343412000897815 steps: 336 lr: 2.5600000000000001e-06 evaluation
 reward: 8.9
 episode: 2012 score: 12.0 memory length: 502896 epsilon:
 0.20226394000897074 steps: 591 lr: 2.5600000000000001e-06 evaluation
 reward: 8.9
 episode: 2013 score: 13.0 memory length: 503378 epsilon:
 0.2013095800089647 steps: 482 lr: 2.5600000000000001e-06 evaluation
 reward: 8.88
 episode: 2014 score: 14.0 memory length: 504033 epsilon:
 0.2000126800089565 steps: 655 lr: 2.5600000000000001e-06 evaluation
 reward: 8.94
 episode: 2015 score: 6.0 memory length: 504381 epsilon:
 0.19932364000895214 steps: 348 lr: 2.5600000000000001e-06 evaluation
 reward: 8.9
 episode: 2016 score: 8.0 memory length: 504857 epsilon:
 0.19838116000894618 steps: 476 lr: 2.5600000000000001e-06 evaluation
 reward: 8.9
 episode: 2017 score: 8.0 memory length: 505276 epsilon:
 0.19755154000894093 steps: 419 lr: 2.5600000000000001e-06 evaluation
 reward: 8.92
 episode: 2018 score: 10.0 memory length: 505745 epsilon:
 0.19662292000893505 steps: 469 lr: 2.5600000000000001e-06 evaluation
 reward: 8.95
 episode: 2019 score: 7.0 memory length: 506144 epsilon:
 0.19583290000893006 steps: 399 lr: 2.5600000000000001e-06 evaluation
 reward: 8.96
 episode: 2020 score: 7.0 memory length: 506484 epsilon: 0.1951597000089258
 steps: 340 lr: 2.5600000000000001e-06 evaluation reward: 8.96
 episode: 2021 score: 4.0 memory length: 506748 epsilon: 0.1946369800089225
 steps: 264 lr: 2.5600000000000001e-06 evaluation reward: 8.94
 episode: 2022 score: 14.0 memory length: 507320 epsilon:
 0.19350442000891532 steps: 572 lr: 2.5600000000000001e-06 evaluation
 reward: 8.99
 episode: 2023 score: 9.0 memory length: 507794 epsilon:
 0.19256590000890939 steps: 474 lr: 2.5600000000000001e-06 evaluation
 reward: 8.96
 episode: 2024 score: 9.0 memory length: 508251 epsilon:
 0.19166104000890366 steps: 457 lr: 2.5600000000000001e-06 evaluation
 reward: 8.93
 episode: 2025 score: 9.0 memory length: 508688 epsilon:

0.19079578000889819 steps: 437 lr: 2.5600000000000001e-06 evaluation
 reward: 8.96
 episode: 2026 score: 9.0 memory length: 509128 epsilon:
 0.18992458000889267 steps: 440 lr: 2.5600000000000001e-06 evaluation
 reward: 8.94
 episode: 2027 score: 11.0 memory length: 509652 epsilon:
 0.1888870600088861 steps: 524 lr: 2.5600000000000001e-06 evaluation
 reward: 8.89
 episode: 2028 score: 16.0 memory length: 510188 epsilon:
 0.1878257800088794 steps: 536 lr: 2.5600000000000001e-06 evaluation
 reward: 9.0
 episode: 2029 score: 7.0 memory length: 510562 epsilon: 0.1870852600088747
 steps: 374 lr: 2.5600000000000001e-06 evaluation reward: 8.97
 episode: 2030 score: 7.0 memory length: 510972 epsilon:
 0.18627346000886957 steps: 410 lr: 2.5600000000000001e-06 evaluation
 reward: 8.98
 episode: 2031 score: 8.0 memory length: 511358 epsilon:
 0.18550918000886474 steps: 386 lr: 2.5600000000000001e-06 evaluation
 reward: 8.95
 episode: 2032 score: 13.0 memory length: 511958 epsilon:
 0.18432118000885722 steps: 600 lr: 2.5600000000000001e-06 evaluation
 reward: 8.94
 episode: 2033 score: 9.0 memory length: 512449 epsilon:
 0.18334900000885107 steps: 491 lr: 2.5600000000000001e-06 evaluation
 reward: 8.92
 episode: 2034 score: 6.0 memory length: 512789 epsilon: 0.182675800088468
 steps: 340 lr: 2.5600000000000001e-06 evaluation reward: 8.92
 episode: 2035 score: 16.0 memory length: 513418 epsilon:
 0.18143038000883893 steps: 629 lr: 2.5600000000000001e-06 evaluation
 reward: 9.03
 episode: 2036 score: 15.0 memory length: 514006 epsilon:
 0.18026614000883157 steps: 588 lr: 2.5600000000000001e-06 evaluation
 reward: 9.12
 episode: 2037 score: 10.0 memory length: 514469 epsilon:
 0.17934940000882577 steps: 463 lr: 2.5600000000000001e-06 evaluation
 reward: 9.14
 episode: 2038 score: 9.0 memory length: 514953 epsilon: 0.1783910800088197
 steps: 484 lr: 2.5600000000000001e-06 evaluation reward: 9.15
 episode: 2039 score: 9.0 memory length: 515445 epsilon:
 0.17741692000881354 steps: 492 lr: 2.5600000000000001e-06 evaluation
 reward: 9.18
 episode: 2040 score: 9.0 memory length: 515883 epsilon:
 0.17654968000880805 steps: 438 lr: 2.5600000000000001e-06 evaluation
 reward: 9.21
 episode: 2041 score: 11.0 memory length: 516436 epsilon:
 0.17545474000880112 steps: 553 lr: 2.5600000000000001e-06 evaluation
 reward: 9.26
 episode: 2042 score: 9.0 memory length: 516911 epsilon:

0.17451424000879517 steps: 475 lr: 2.5600000000000001e-06 evaluation
 reward: 9.3
 episode: 2043 score: 15.0 memory length: 517616 epsilon:
 0.17311834000878634 steps: 705 lr: 2.5600000000000001e-06 evaluation
 reward: 9.32
 episode: 2044 score: 9.0 memory length: 518108 epsilon:
 0.17214418000878018 steps: 492 lr: 2.5600000000000001e-06 evaluation
 reward: 9.27
 episode: 2045 score: 10.0 memory length: 518663 epsilon:
 0.17104528000877323 steps: 555 lr: 2.5600000000000001e-06 evaluation
 reward: 9.25
 episode: 2046 score: 5.0 memory length: 518937 epsilon: 0.1705027600087698
 steps: 274 lr: 2.5600000000000001e-06 evaluation reward: 9.21
 episode: 2047 score: 14.0 memory length: 519526 epsilon:
 0.16933654000876242 steps: 589 lr: 2.5600000000000001e-06 evaluation
 reward: 9.29
 episode: 2048 score: 10.0 memory length: 520026 epsilon:
 0.16834654000875615 steps: 500 lr: 2.5600000000000001e-06 evaluation
 reward: 9.33
 episode: 2049 score: 15.0 memory length: 520690 epsilon:
 0.16703182000874783 steps: 664 lr: 2.5600000000000001e-06 evaluation
 reward: 9.43
 episode: 2050 score: 11.0 memory length: 521215 epsilon:
 0.16599232000874126 steps: 525 lr: 2.5600000000000001e-06 evaluation
 reward: 9.44
 episode: 2051 score: 9.0 memory length: 521631 epsilon:
 0.16516864000873605 steps: 416 lr: 2.5600000000000001e-06 evaluation
 reward: 9.46
 episode: 2052 score: 8.0 memory length: 522058 epsilon: 0.1643231800087307
 steps: 427 lr: 2.5600000000000001e-06 evaluation reward: 9.46
 episode: 2053 score: 9.0 memory length: 522552 epsilon: 0.1633450600087245
 steps: 494 lr: 2.5600000000000001e-06 evaluation reward: 9.43
 episode: 2054 score: 22.0 memory length: 523208 epsilon:
 0.1620461800087163 steps: 656 lr: 2.5600000000000001e-06 evaluation
 reward: 9.59
 episode: 2055 score: 9.0 memory length: 523712 epsilon:
 0.16104826000870998 steps: 504 lr: 2.5600000000000001e-06 evaluation
 reward: 9.61
 episode: 2056 score: 11.0 memory length: 524255 epsilon:
 0.15997312000870317 steps: 543 lr: 2.5600000000000001e-06 evaluation
 reward: 9.66
 episode: 2057 score: 12.0 memory length: 524697 epsilon:
 0.15909796000869764 steps: 442 lr: 2.5600000000000001e-06 evaluation
 reward: 9.63
 episode: 2058 score: 9.0 memory length: 525140 epsilon: 0.1582208200086921
 steps: 443 lr: 2.5600000000000001e-06 evaluation reward: 9.54
 episode: 2059 score: 7.0 memory length: 525517 epsilon:
 0.15747436000868736 steps: 377 lr: 2.5600000000000001e-06 evaluation

```

reward: 9.49
episode: 2060    score: 8.0    memory length: 525943    epsilon:
0.15663088000868203    steps: 426    lr: 2.5600000000000001e-06    evaluation
reward: 9.49
episode: 2061    score: 12.0   memory length: 526546    epsilon:
0.15543694000867447    steps: 603    lr: 2.5600000000000001e-06    evaluation
reward: 9.4
episode: 2062    score: 12.0   memory length: 527015    epsilon:
0.1545083200086686    steps: 469    lr: 2.5600000000000001e-06    evaluation
reward: 9.43
episode: 2063    score: 10.0   memory length: 527554    epsilon:
0.15344110000866185    steps: 539    lr: 2.5600000000000001e-06    evaluation
reward: 9.48
episode: 2064    score: 12.0   memory length: 528090    epsilon:
0.15237982000865513    steps: 536    lr: 2.5600000000000001e-06    evaluation
reward: 9.54
episode: 2065    score: 8.0    memory length: 528497    epsilon:
0.15157396000865003    steps: 407    lr: 2.5600000000000001e-06    evaluation
reward: 9.52
episode: 2066    score: 12.0   memory length: 529015    epsilon:
0.15054832000864354    steps: 518    lr: 2.5600000000000001e-06    evaluation
reward: 9.55
episode: 2067    score: 7.0    memory length: 529368    epsilon:
0.14984938000863912    steps: 353    lr: 2.5600000000000001e-06    evaluation
reward: 9.5
episode: 2068    score: 6.0    memory length: 529747    epsilon:
0.14909896000863437    steps: 379    lr: 2.5600000000000001e-06    evaluation
reward: 9.48
episode: 2069    score: 8.0    memory length: 530191    epsilon: 0.1482198400086288
steps: 444    lr: 2.5600000000000001e-06    evaluation reward: 9.47
episode: 2070    score: 7.0    memory length: 530570    epsilon:
0.14746942000862406    steps: 379    lr: 2.5600000000000001e-06    evaluation
reward: 9.45
episode: 2071    score: 8.0    memory length: 531027    epsilon:
0.14656456000861834    steps: 457    lr: 2.5600000000000001e-06    evaluation
reward: 9.49
episode: 2072    score: 22.0   memory length: 531752    epsilon:
0.14512906000860926    steps: 725    lr: 2.5600000000000001e-06    evaluation
reward: 9.6
episode: 2073    score: 9.0    memory length: 532208    epsilon:
0.14422618000860354    steps: 456    lr: 2.5600000000000001e-06    evaluation
reward: 9.67
episode: 2074    score: 9.0    memory length: 532719    epsilon:
0.14321440000859714    steps: 511    lr: 2.5600000000000001e-06    evaluation
reward: 9.69
episode: 2075    score: 9.0    memory length: 533155    epsilon:
0.14235112000859168    steps: 436    lr: 2.5600000000000001e-06    evaluation
reward: 9.73

```

episode: 2076 score: 9.0 memory length: 533630 epsilon:
 0.14141062000858573 steps: 475 lr: 2.5600000000000001e-06 evaluation
 reward: 9.71
 episode: 2077 score: 9.0 memory length: 534104 epsilon: 0.1404721000085798
 steps: 474 lr: 2.5600000000000001e-06 evaluation reward: 9.75
 episode: 2078 score: 8.0 memory length: 534504 epsilon:
 0.13968010000857478 steps: 400 lr: 2.5600000000000001e-06 evaluation
 reward: 9.75
 episode: 2079 score: 10.0 memory length: 534992 epsilon:
 0.13871386000856867 steps: 488 lr: 2.5600000000000001e-06 evaluation
 reward: 9.71
 episode: 2080 score: 16.0 memory length: 535693 epsilon:
 0.1373258800085599 steps: 701 lr: 2.5600000000000001e-06 evaluation
 reward: 9.79
 episode: 2081 score: 11.0 memory length: 536198 epsilon:
 0.13632598000855356 steps: 505 lr: 2.5600000000000001e-06 evaluation
 reward: 9.83
 episode: 2082 score: 10.0 memory length: 536706 epsilon:
 0.1353201400085472 steps: 508 lr: 2.5600000000000001e-06 evaluation
 reward: 9.72
 episode: 2083 score: 8.0 memory length: 537142 epsilon:
 0.13445686000854173 steps: 436 lr: 2.5600000000000001e-06 evaluation
 reward: 9.69
 episode: 2084 score: 10.0 memory length: 537633 epsilon:
 0.13348468000853558 steps: 491 lr: 2.5600000000000001e-06 evaluation
 reward: 9.69
 episode: 2085 score: 8.0 memory length: 538040 epsilon:
 0.13267882000853048 steps: 407 lr: 2.5600000000000001e-06 evaluation
 reward: 9.72
 episode: 2086 score: 10.0 memory length: 538569 epsilon:
 0.13163140000852386 steps: 529 lr: 2.5600000000000001e-06 evaluation
 reward: 9.68
 episode: 2087 score: 16.0 memory length: 539178 epsilon:
 0.13042558000851623 steps: 609 lr: 2.5600000000000001e-06 evaluation
 reward: 9.76
 episode: 2088 score: 20.0 memory length: 539792 epsilon:
 0.12920986000850854 steps: 614 lr: 2.5600000000000001e-06 evaluation
 reward: 9.84
 episode: 2089 score: 10.0 memory length: 540277 epsilon:
 0.12824956000850246 steps: 485 lr: 2.5600000000000001e-06 evaluation
 reward: 9.86
 episode: 2090 score: 11.0 memory length: 540816 epsilon:
 0.1271823400084957 steps: 539 lr: 2.5600000000000001e-06 evaluation
 reward: 9.93
 episode: 2091 score: 13.0 memory length: 541303 epsilon:
 0.1262180800084896 steps: 487 lr: 2.5600000000000001e-06 evaluation
 reward: 9.92
 episode: 2092 score: 21.0 memory length: 542013 epsilon:

0.12481228000848203	steps: 710	lr: 2.5600000000000001e-06	evaluation
reward: 10.05			
episode: 2093	score: 12.0	memory length: 542477	epsilon:
0.12389356000848266	steps: 464	lr: 2.5600000000000001e-06	evaluation
reward: 10.09			
episode: 2094	score: 11.0	memory length: 543022	epsilon:
0.1228144600084834	steps: 545	lr: 2.5600000000000001e-06	evaluation
reward: 10.07			
episode: 2095	score: 10.0	memory length: 543481	epsilon:
0.12190564000848401	steps: 459	lr: 2.5600000000000001e-06	evaluation
reward: 10.1			
episode: 2096	score: 13.0	memory length: 544066	epsilon:
0.1207473400084848	steps: 585	lr: 2.5600000000000001e-06	evaluation
reward: 10.13			
episode: 2097	score: 10.0	memory length: 544567	epsilon:
0.11975536000848548	steps: 501	lr: 2.5600000000000001e-06	evaluation
reward: 10.13			
episode: 2098	score: 9.0	memory length: 545042	epsilon:
0.11881486000848612	steps: 475	lr: 2.5600000000000001e-06	evaluation
reward: 10.16			
episode: 2099	score: 12.0	memory length: 545611	epsilon:
0.11768824000848689	steps: 569	lr: 2.5600000000000001e-06	evaluation
reward: 10.19			
episode: 2100	score: 23.0	memory length: 546380	epsilon:
0.11616562000848793	steps: 769	lr: 2.5600000000000001e-06	evaluation
reward: 10.35			
episode: 2101	score: 16.0	memory length: 546922	epsilon:
0.11509246000848866	steps: 542	lr: 2.5600000000000001e-06	evaluation
reward: 10.46			
episode: 2102	score: 7.0	memory length: 547278	epsilon:
0.11438758000848914	steps: 356	lr: 2.5600000000000001e-06	evaluation
reward: 10.39			
episode: 2103	score: 7.0	memory length: 547687	epsilon: 0.1135777600084897
steps: 409	lr: 2.5600000000000001e-06	evaluation reward: 10.42	
episode: 2104	score: 14.0	memory length: 548287	epsilon:
0.1123897600084905	steps: 600	lr: 2.5600000000000001e-06	evaluation
reward: 10.47			
episode: 2105	score: 9.0	memory length: 548763	epsilon:
0.11144728000849115	steps: 476	lr: 2.5600000000000001e-06	evaluation
reward: 10.51			
episode: 2106	score: 9.0	memory length: 549219	epsilon:
0.11054440000849176	steps: 456	lr: 2.5600000000000001e-06	evaluation
reward: 10.49			
episode: 2107	score: 9.0	memory length: 549677	epsilon:
0.10963756000849238	steps: 458	lr: 2.5600000000000001e-06	evaluation
reward: 10.51			
episode: 2108	score: 10.0	memory length: 550172	epsilon:
0.10865746000849305	steps: 495	lr: 2.5600000000000001e-06	evaluation

reward: 10.51			
episode: 2109	score: 11.0	memory length: 550707	epsilon: 0.10759816000849377
	steps: 535	lr: 2.5600000000000001e-06	evaluation
reward: 10.53			
episode: 2110	score: 8.0	memory length: 551164	epsilon: 0.10669330000849439
	steps: 457	lr: 2.5600000000000001e-06	evaluation
reward: 10.54			
episode: 2111	score: 14.0	memory length: 551743	epsilon: 0.10554688000849517
	steps: 579	lr: 2.5600000000000001e-06	evaluation
reward: 10.62			
episode: 2112	score: 11.0	memory length: 552282	epsilon: 0.1044796600084959
	steps: 539	lr: 2.5600000000000001e-06	evaluation
reward: 10.61			
episode: 2113	score: 6.0	memory length: 552621	epsilon: 0.10380844000849636
	steps: 339	lr: 2.5600000000000001e-06	evaluation
reward: 10.54			
episode: 2114	score: 14.0	memory length: 553172	epsilon: 0.1027174600084971
	steps: 551	lr: 2.5600000000000001e-06	evaluation
reward: 10.54			
episode: 2115	score: 8.0	memory length: 553579	epsilon: 0.10191160000849765
	steps: 407	lr: 2.5600000000000001e-06	evaluation
reward: 10.56			
episode: 2116	score: 10.0	memory length: 554103	epsilon: 0.10087408000849836
	steps: 524	lr: 2.5600000000000001e-06	evaluation
reward: 10.58			
episode: 2117	score: 10.0	memory length: 554557	epsilon: 0.09997516000849897
	steps: 454	lr: 2.5600000000000001e-06	evaluation
reward: 10.6			
episode: 2118	score: 8.0	memory length: 554981	epsilon: 0.09913564000849955
	steps: 424	lr: 2.5600000000000001e-06	evaluation
reward: 10.58			
episode: 2119	score: 12.0	memory length: 555558	epsilon: 0.09799318000850032
	steps: 577	lr: 2.5600000000000001e-06	evaluation
reward: 10.63			
episode: 2120	score: 13.0	memory length: 556165	epsilon: 0.09679132000850114
	steps: 607	lr: 2.5600000000000001e-06	evaluation
reward: 10.69			
episode: 2121	score: 11.0	memory length: 556752	epsilon: 0.09562906000850194
	steps: 587	lr: 2.5600000000000001e-06	evaluation
reward: 10.76			
episode: 2122	score: 6.0	memory length: 557081	epsilon: 0.09497764000850238
	steps: 329	lr: 2.5600000000000001e-06	evaluation
reward: 10.68			
episode: 2123	score: 11.0	memory length: 557637	epsilon: 0.09387676000850313
	steps: 556	lr: 2.5600000000000001e-06	evaluation
reward: 10.7			
episode: 2124	score: 9.0	memory length: 558113	epsilon: 0.09293428000850378
	steps: 476	lr: 2.5600000000000001e-06	evaluation

```

reward: 10.7
episode: 2125   score: 8.0   memory length: 558491   epsilon:
0.09218584000850429   steps: 378   lr: 2.5600000000000001e-06   evaluation
reward: 10.69
episode: 2126   score: 9.0   memory length: 558970   epsilon:
0.09123742000850493   steps: 479   lr: 2.5600000000000001e-06   evaluation
reward: 10.69
episode: 2127   score: 12.0  memory length: 559439   epsilon:
0.09030880000850557   steps: 469   lr: 2.5600000000000001e-06   evaluation
reward: 10.7
episode: 2128   score: 10.0  memory length: 559980   epsilon:
0.0892376200085063   steps: 541   lr: 2.5600000000000001e-06   evaluation
reward: 10.64
episode: 2129   score: 4.0   memory length: 560237   epsilon:
0.08872876000850664   steps: 257   lr: 2.5600000000000001e-06   evaluation
reward: 10.61
episode: 2130   score: 11.0  memory length: 560755   epsilon:
0.08770312000850734   steps: 518   lr: 2.5600000000000001e-06   evaluation
reward: 10.65
episode: 2131   score: 10.0  memory length: 561281   epsilon:
0.08666164000850805   steps: 526   lr: 2.5600000000000001e-06   evaluation
reward: 10.67
episode: 2132   score: 9.0   memory length: 561751   epsilon:
0.08573104000850869   steps: 470   lr: 2.5600000000000001e-06   evaluation
reward: 10.63
episode: 2133   score: 11.0  memory length: 562288   epsilon:
0.08466778000850941   steps: 537   lr: 2.5600000000000001e-06   evaluation
reward: 10.65
episode: 2134   score: 9.0   memory length: 562742   epsilon:
0.08376886000851003   steps: 454   lr: 2.5600000000000001e-06   evaluation
reward: 10.68
episode: 2135   score: 12.0  memory length: 563301   epsilon:
0.08266204000851078   steps: 559   lr: 2.5600000000000001e-06   evaluation
reward: 10.64
episode: 2136   score: 11.0  memory length: 563837   epsilon:
0.0816007600085115   steps: 536   lr: 2.5600000000000001e-06   evaluation
reward: 10.6
episode: 2137   score: 12.0  memory length: 564375   epsilon:
0.08053552000851223   steps: 538   lr: 2.5600000000000001e-06   evaluation
reward: 10.62
episode: 2138   score: 16.0  memory length: 565059   epsilon:
0.07918120000851316   steps: 684   lr: 2.5600000000000001e-06   evaluation
reward: 10.69
episode: 2139   score: 6.0   memory length: 565416   epsilon:
0.07847434000851364   steps: 357   lr: 2.5600000000000001e-06   evaluation
reward: 10.66
episode: 2140   score: 14.0  memory length: 566068   epsilon:
0.07718338000851452   steps: 652   lr: 2.5600000000000001e-06   evaluation

```

```

reward: 10.71
episode: 2141    score: 17.0    memory length: 566674    epsilon:
0.07598350000851534    steps: 606    lr: 2.5600000000000001e-06    evaluation
reward: 10.77
episode: 2142    score: 10.0    memory length: 567143    epsilon:
0.07505488000851597    steps: 469    lr: 2.5600000000000001e-06    evaluation
reward: 10.78
episode: 2143    score: 9.0    memory length: 567597    epsilon:
0.07415596000851658    steps: 454    lr: 2.5600000000000001e-06    evaluation
reward: 10.72
episode: 2144    score: 7.0    memory length: 567955    epsilon:
0.07344712000851707    steps: 358    lr: 2.5600000000000001e-06    evaluation
reward: 10.7
episode: 2145    score: 12.0    memory length: 568558    epsilon:
0.07225318000851788    steps: 603    lr: 2.5600000000000001e-06    evaluation
reward: 10.72

```

```

-----
KeyboardInterrupt                                Traceback (most recent call last)
<ipython-input-9-ccac64972853> in <cell line: 3>()
    36         # Start training after random sample generation
    37         if(frame >= train_frame): # You can set train_frame to a lower_
    ↪value while testing your starts training earlier
----> 38             agent.train_policy_net(frame)
    39             # Update the target network only for Double DQN only
    40             if double_dqn and (frame %_
    ↪update_target_network_frequency)== 0:

/content/drive/MyDrive/CS444/assignment5/assignment5_materials/agent_double.py_
    ↪in train_policy_net(self, frame)
    121         ### CODE ###
    122         self.optimizer.zero_grad()
--> 123         loss.backward()
    124         self.optimizer.step()
    125         self.scheduler.step()

/usr/local/lib/python3.10/dist-packages/torch/_tensor.py in backward(self,_
    ↪gradient, retain_graph, create_graph, inputs)
    520             inputs=inputs,
    521             )
--> 522         torch.autograd.backward(

    523             self, gradient, retain_graph, create_graph, inputs=inputs
    524             )

/usr/local/lib/python3.10/dist-packages/torch/autograd/__init__.py in_
    ↪backward(tensors, grad_tensors, retain_graph, create_graph, grad_variables,_
    ↪inputs)

```



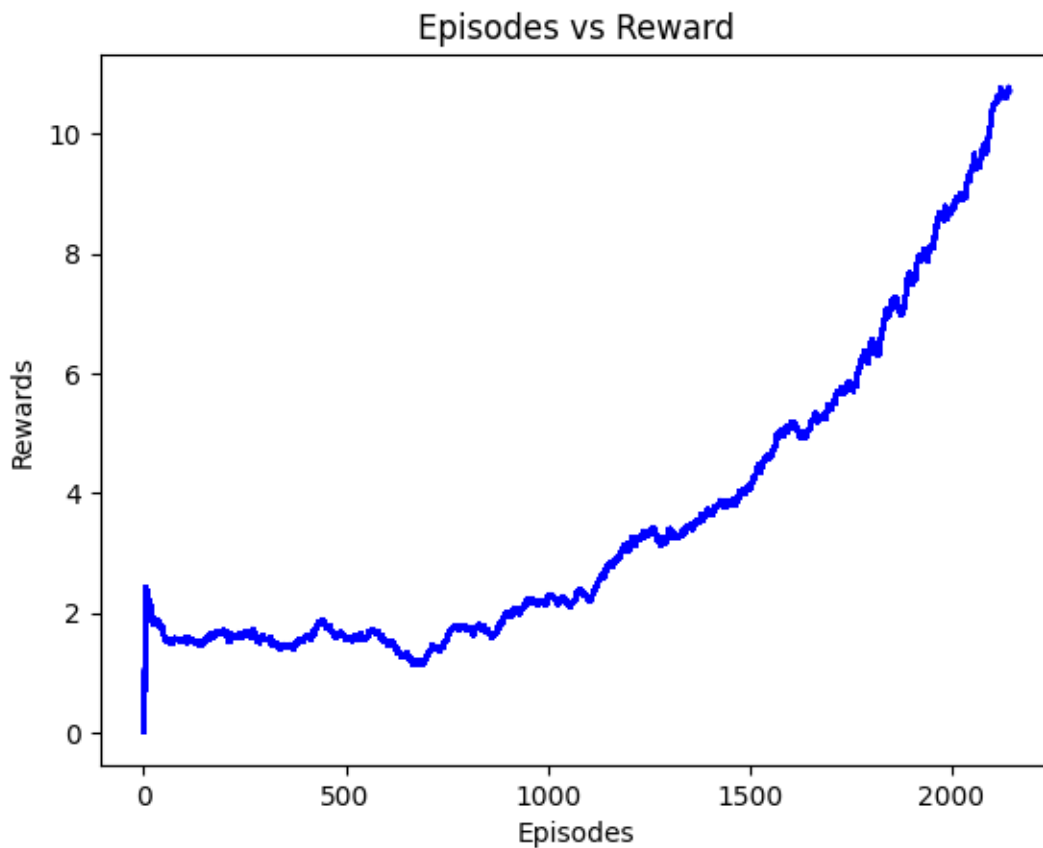
```

264     # some Python versions print out the first line of a multi-line_
↳function
265     # calls in the traceback and some print out the last line
--> 266     Variable._execution_engine.run_backward( # Calls into the C++_
↳engine to run the backward pass

267         tensors,
268         grad_tensors_,

```

KeyboardInterrupt:



2 Visualize Agent Performance

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

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

```
[ ]: !pip3 install xvfbwrapper
```

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

[ ]: # from gym.wrappers import Monitor # If importing monitor raises issues, try
      ↪ using `from gym.wrappers import RecordVideo`
from gym.wrappers import RecordVideo
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

[ ]: 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()
```