

MP5

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.

```
[5]: # 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).

```
[6]: # 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

```
[7]: !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

```
[8]: !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)
Requirement already satisfied: ez_setup in /usr/local/lib/python3.10/dist-packages (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)
Requirement already satisfied: autorom[accept-rom-license]~=0.4.2 in /usr/local/lib/python3.10/dist-packages (from gym[accept-rom-license]) (0.4.2)
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)
Requirement already satisfied: AutoROM.accept-rom-license in /usr/local/lib/python3.10/dist-packages (from autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license]) (0.6.1)
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)

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**.

```
[9]: %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.

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

```
/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)
/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(
```

```
[11]: 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.

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

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

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

1.2.1 Main Training Loop

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

```
[13]: 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_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: 294    epsilon: 1.0    steps: 170    lr:
0.0001    evaluation reward: 0.5
episode: 2    score: 2.0    memory length: 493    epsilon: 1.0    steps: 199    lr:
0.0001    evaluation reward: 1.0
episode: 3    score: 2.0    memory length: 709    epsilon: 1.0    steps: 216    lr:
0.0001    evaluation reward: 1.25
episode: 4    score: 0.0    memory length: 833    epsilon: 1.0    steps: 124    lr:
0.0001    evaluation reward: 1.0
episode: 5    score: 0.0    memory length: 957    epsilon: 1.0    steps: 124    lr:
0.0001    evaluation reward: 0.8333333333333334
episode: 6    score: 4.0    memory length: 1238    epsilon: 1.0    steps: 281
lr: 0.0001    evaluation reward: 1.2857142857142858
episode: 7    score: 2.0    memory length: 1436    epsilon: 1.0    steps: 198
lr: 0.0001    evaluation reward: 1.375
episode: 8    score: 0.0    memory length: 1560    epsilon: 1.0    steps: 124
lr: 0.0001    evaluation reward: 1.2222222222222223
episode: 9    score: 2.0    memory length: 1759    epsilon: 1.0    steps: 199
lr: 0.0001    evaluation reward: 1.3
episode: 10   score: 5.0    memory length: 2109    epsilon: 1.0    steps: 350
lr: 0.0001    evaluation reward: 1.6363636363636365
episode: 11   score: 0.0    memory length: 2233    epsilon: 1.0    steps: 124
lr: 0.0001    evaluation reward: 1.5
episode: 12   score: 2.0    memory length: 2434    epsilon: 1.0    steps: 201
lr: 0.0001    evaluation reward: 1.5384615384615385
episode: 13   score: 2.0    memory length: 2650    epsilon: 1.0    steps: 216
lr: 0.0001    evaluation reward: 1.5714285714285714
episode: 14   score: 1.0    memory length: 2821    epsilon: 1.0    steps: 171
lr: 0.0001    evaluation reward: 1.5333333333333334
episode: 15   score: 3.0    memory length: 3068    epsilon: 1.0    steps: 247
lr: 0.0001    evaluation reward: 1.625
episode: 16   score: 2.0    memory length: 3284    epsilon: 1.0    steps: 216
lr: 0.0001    evaluation reward: 1.6470588235294117
episode: 17   score: 1.0    memory length: 3453    epsilon: 1.0    steps: 169
lr: 0.0001    evaluation reward: 1.6111111111111112
episode: 18   score: 1.0    memory length: 3605    epsilon: 1.0    steps: 152
lr: 0.0001    evaluation reward: 1.5789473684210527
episode: 19   score: 1.0    memory length: 3756    epsilon: 1.0    steps: 151
lr: 0.0001    evaluation reward: 1.55
episode: 20   score: 4.0    memory length: 4033    epsilon: 1.0    steps: 277
lr: 0.0001    evaluation reward: 1.6666666666666667
episode: 21   score: 0.0    memory length: 4156    epsilon: 1.0    steps: 123
lr: 0.0001    evaluation reward: 1.5909090909090908
episode: 22   score: 1.0    memory length: 4326    epsilon: 1.0    steps: 170

```

lr: 0.0001	evaluation reward: 1.565217391304348			
episode: 23	score: 0.0	memory length: 4450	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5			
episode: 24	score: 0.0	memory length: 4573	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 25	score: 2.0	memory length: 4771	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4615384615384615			
episode: 26	score: 2.0	memory length: 4989	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.4814814814814814			
episode: 27	score: 1.0	memory length: 5140	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.4642857142857142			
episode: 28	score: 0.0	memory length: 5263	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4137931034482758			
episode: 29	score: 0.0	memory length: 5387	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.3666666666666667			
episode: 30	score: 2.0	memory length: 5587	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.3870967741935485			
episode: 31	score: 4.0	memory length: 5885	epsilon: 1.0	steps: 298
lr: 0.0001	evaluation reward: 1.46875			
episode: 32	score: 3.0	memory length: 6131	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.5151515151515151			
episode: 33	score: 0.0	memory length: 6255	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.4705882352941178			
episode: 34	score: 3.0	memory length: 6505	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.5142857142857142			
episode: 35	score: 3.0	memory length: 6750	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.5555555555555556			
episode: 36	score: 2.0	memory length: 6968	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.5675675675675675			
episode: 37	score: 3.0	memory length: 7217	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.605263157894737			
episode: 38	score: 0.0	memory length: 7341	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.564102564102564			
episode: 39	score: 1.0	memory length: 7514	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.55			
episode: 40	score: 1.0	memory length: 7665	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.5365853658536586			
episode: 41	score: 3.0	memory length: 7931	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.5714285714285714			
episode: 42	score: 0.0	memory length: 8055	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5348837209302326			
episode: 43	score: 3.0	memory length: 8301	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.5681818181818181			
episode: 44	score: 2.0	memory length: 8500	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.5777777777777777			
episode: 45	score: 2.0	memory length: 8716	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.5869565217391304			
episode: 46	score: 0.0	memory length: 8839	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.553191489361702		
episode: 47	score: 3.0	memory length: 9087	epsilon: 1.0 steps: 248
lr: 0.0001	evaluation reward: 1.5833333333333333		
episode: 48	score: 5.0	memory length: 9382	epsilon: 1.0 steps: 295
lr: 0.0001	evaluation reward: 1.653061224489796		
episode: 49	score: 0.0	memory length: 9505	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.62		
episode: 50	score: 1.0	memory length: 9677	epsilon: 1.0 steps: 172
lr: 0.0001	evaluation reward: 1.607843137254902		
episode: 51	score: 0.0	memory length: 9801	epsilon: 1.0 steps: 124
lr: 0.0001	evaluation reward: 1.5769230769230769		
episode: 52	score: 1.0	memory length: 9953	epsilon: 1.0 steps: 152
lr: 0.0001	evaluation reward: 1.5660377358490567		
episode: 53	score: 1.0	memory length: 10125	epsilon: 1.0 steps: 172
lr: 0.0001	evaluation reward: 1.5555555555555556		
episode: 54	score: 2.0	memory length: 10345	epsilon: 1.0 steps: 220
lr: 0.0001	evaluation reward: 1.5636363636363637		
episode: 55	score: 1.0	memory length: 10515	epsilon: 1.0 steps: 170
lr: 0.0001	evaluation reward: 1.5535714285714286		
episode: 56	score: 1.0	memory length: 10667	epsilon: 1.0 steps: 152
lr: 0.0001	evaluation reward: 1.543859649122807		
episode: 57	score: 0.0	memory length: 10791	epsilon: 1.0 steps: 124
lr: 0.0001	evaluation reward: 1.5172413793103448		
episode: 58	score: 1.0	memory length: 10964	epsilon: 1.0 steps: 173
lr: 0.0001	evaluation reward: 1.5084745762711864		
episode: 59	score: 3.0	memory length: 11190	epsilon: 1.0 steps: 226
lr: 0.0001	evaluation reward: 1.5333333333333334		
episode: 60	score: 2.0	memory length: 11373	epsilon: 1.0 steps: 183
lr: 0.0001	evaluation reward: 1.540983606557377		
episode: 61	score: 2.0	memory length: 11572	epsilon: 1.0 steps: 199
lr: 0.0001	evaluation reward: 1.5483870967741935		
episode: 62	score: 2.0	memory length: 11789	epsilon: 1.0 steps: 217
lr: 0.0001	evaluation reward: 1.5555555555555556		
episode: 63	score: 2.0	memory length: 11989	epsilon: 1.0 steps: 200
lr: 0.0001	evaluation reward: 1.5625		
episode: 64	score: 1.0	memory length: 12159	epsilon: 1.0 steps: 170
lr: 0.0001	evaluation reward: 1.5538461538461539		
episode: 65	score: 0.0	memory length: 12282	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5303030303030303		
episode: 66	score: 2.0	memory length: 12502	epsilon: 1.0 steps: 220
lr: 0.0001	evaluation reward: 1.537313432835821		
episode: 67	score: 2.0	memory length: 12701	epsilon: 1.0 steps: 199
lr: 0.0001	evaluation reward: 1.5441176470588236		
episode: 68	score: 3.0	memory length: 12968	epsilon: 1.0 steps: 267
lr: 0.0001	evaluation reward: 1.565217391304348		
episode: 69	score: 2.0	memory length: 13166	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.5714285714285714		
episode: 70	score: 1.0	memory length: 13336	epsilon: 1.0 steps: 170

lr: 0.0001	evaluation reward: 1.5633802816901408		
episode: 71	score: 2.0 memory length: 13535	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.5694444444444444		
episode: 72	score: 2.0 memory length: 13733	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5753424657534247		
episode: 73	score: 0.0 memory length: 13857	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.554054054054054		
episode: 74	score: 3.0 memory length: 14124	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.5733333333333333		
episode: 75	score: 0.0 memory length: 14248	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5526315789473684		
episode: 76	score: 0.0 memory length: 14371	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5324675324675325		
episode: 77	score: 1.0 memory length: 14542	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.5256410256410255		
episode: 78	score: 2.0 memory length: 14761	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.5316455696202531		
episode: 79	score: 3.0 memory length: 15011	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.55		
episode: 80	score: 3.0 memory length: 15238	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.5679012345679013		
episode: 81	score: 1.0 memory length: 15407	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.5609756097560976		
episode: 82	score: 1.0 memory length: 15578	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.5542168674698795		
episode: 83	score: 1.0 memory length: 15748	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.5476190476190477		
episode: 84	score: 2.0 memory length: 15965	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.5529411764705883		
episode: 85	score: 2.0 memory length: 16166	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.558139534883721		
episode: 86	score: 4.0 memory length: 16484	epsilon: 1.0	steps: 318
lr: 0.0001	evaluation reward: 1.5862068965517242		
episode: 87	score: 5.0 memory length: 16815	epsilon: 1.0	steps: 331
lr: 0.0001	evaluation reward: 1.625		
episode: 88	score: 0.0 memory length: 16939	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.6067415730337078		
episode: 89	score: 1.0 memory length: 17111	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.6		
episode: 90	score: 1.0 memory length: 17281	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.5934065934065933		
episode: 91	score: 0.0 memory length: 17405	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.576086956521739		
episode: 92	score: 1.0 memory length: 17557	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.5698924731182795		
episode: 93	score: 1.0 memory length: 17709	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.5638297872340425		
episode: 94	score: 1.0 memory length: 17878	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.5578947368421052		
episode: 95	score: 0.0 memory length: 18002	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5416666666666667		
episode: 96	score: 0.0 memory length: 18126	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5257731958762886		
episode: 97	score: 1.0 memory length: 18299	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.5204081632653061		
episode: 98	score: 2.0 memory length: 18499	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.5252525252525253		
episode: 99	score: 2.0 memory length: 18680	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.53		
episode: 100	score: 2.0 memory length: 18878	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.55		
episode: 101	score: 2.0 memory length: 19079	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.56		
episode: 102	score: 2.0 memory length: 19298	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.56		
episode: 103	score: 1.0 memory length: 19468	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.55		
episode: 104	score: 2.0 memory length: 19686	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.57		
episode: 105	score: 3.0 memory length: 19958	epsilon: 1.0	steps: 272
lr: 0.0001	evaluation reward: 1.6		
episode: 106	score: 0.0 memory length: 20081	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.56		
episode: 107	score: 1.0 memory length: 20233	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.55		
episode: 108	score: 1.0 memory length: 20406	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.56		
episode: 109	score: 2.0 memory length: 20623	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.56		
episode: 110	score: 1.0 memory length: 20793	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.52		
episode: 111	score: 2.0 memory length: 20991	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.54		
episode: 112	score: 3.0 memory length: 21220	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.55		
episode: 113	score: 0.0 memory length: 21344	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.53		
episode: 114	score: 2.0 memory length: 21546	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.54		
episode: 115	score: 3.0 memory length: 21774	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.54		
episode: 116	score: 2.0 memory length: 21973	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.54		
episode: 117	score: 0.0 memory length: 22097	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.53		
episode: 118	score: 1.0 memory length: 22269	epsilon: 1.0	steps: 172

lr: 0.0001	evaluation reward: 1.53		
episode: 119	score: 2.0 memory length: 22468	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.54		
episode: 120	score: 2.0 memory length: 22687	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.52		
episode: 121	score: 1.0 memory length: 22839	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.53		
episode: 122	score: 1.0 memory length: 22991	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.53		
episode: 123	score: 1.0 memory length: 23144	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.54		
episode: 124	score: 3.0 memory length: 23393	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.57		
episode: 125	score: 3.0 memory length: 23643	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.58		
episode: 126	score: 2.0 memory length: 23842	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.58		
episode: 127	score: 0.0 memory length: 23965	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57		
episode: 128	score: 2.0 memory length: 24164	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.59		
episode: 129	score: 4.0 memory length: 24422	epsilon: 1.0	steps: 258
lr: 0.0001	evaluation reward: 1.63		
episode: 130	score: 2.0 memory length: 24623	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.63		
episode: 131	score: 0.0 memory length: 24747	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.59		
episode: 132	score: 0.0 memory length: 24871	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.56		
episode: 133	score: 4.0 memory length: 25148	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.6		
episode: 134	score: 3.0 memory length: 25375	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.6		
episode: 135	score: 1.0 memory length: 25546	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.58		
episode: 136	score: 2.0 memory length: 25767	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.58		
episode: 137	score: 1.0 memory length: 25937	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.56		
episode: 138	score: 4.0 memory length: 26212	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.6		
episode: 139	score: 0.0 memory length: 26336	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.59		
episode: 140	score: 0.0 memory length: 26460	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.58		
episode: 141	score: 1.0 memory length: 26631	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.56		
episode: 142	score: 2.0 memory length: 26829	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.58		
episode: 143	score: 3.0 memory length: 27095	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.58		
episode: 144	score: 0.0 memory length: 27219	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.56		
episode: 145	score: 0.0 memory length: 27343	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54		
episode: 146	score: 3.0 memory length: 27570	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.57		
episode: 147	score: 0.0 memory length: 27694	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54		
episode: 148	score: 2.0 memory length: 27892	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.51		
episode: 149	score: 2.0 memory length: 28108	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.53		
episode: 150	score: 2.0 memory length: 28329	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.54		
episode: 151	score: 2.0 memory length: 28548	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.56		
episode: 152	score: 1.0 memory length: 28718	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.56		
episode: 153	score: 2.0 memory length: 28900	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.57		
episode: 154	score: 3.0 memory length: 29131	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.58		
episode: 155	score: 1.0 memory length: 29283	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.58		
episode: 156	score: 0.0 memory length: 29407	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.57		
episode: 157	score: 2.0 memory length: 29606	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.59		
episode: 158	score: 0.0 memory length: 29730	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.58		
episode: 159	score: 1.0 memory length: 29900	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.56		
episode: 160	score: 0.0 memory length: 30024	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54		
episode: 161	score: 0.0 memory length: 30147	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 162	score: 0.0 memory length: 30271	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5		
episode: 163	score: 0.0 memory length: 30394	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48		
episode: 164	score: 0.0 memory length: 30518	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.47		
episode: 165	score: 1.0 memory length: 30689	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.48		
episode: 166	score: 0.0 memory length: 30813	epsilon: 1.0	steps: 124

lr: 0.0001	evaluation reward: 1.46		
episode: 167	score: 3.0 memory length: 31062	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.47		
episode: 168	score: 2.0 memory length: 31280	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.46		
episode: 169	score: 3.0 memory length: 31549	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.47		
episode: 170	score: 1.0 memory length: 31721	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.47		
episode: 171	score: 2.0 memory length: 31920	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.47		
episode: 172	score: 0.0 memory length: 32044	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.45		
episode: 173	score: 6.0 memory length: 32408	epsilon: 1.0	steps: 364
lr: 0.0001	evaluation reward: 1.51		
episode: 174	score: 5.0 memory length: 32699	epsilon: 1.0	steps: 291
lr: 0.0001	evaluation reward: 1.53		
episode: 175	score: 2.0 memory length: 32898	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.55		
episode: 176	score: 0.0 memory length: 33022	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.55		
episode: 177	score: 1.0 memory length: 33173	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.55		
episode: 178	score: 1.0 memory length: 33343	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.54		
episode: 179	score: 3.0 memory length: 33609	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.54		
episode: 180	score: 3.0 memory length: 33855	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.54		
episode: 181	score: 3.0 memory length: 34103	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.56		
episode: 182	score: 2.0 memory length: 34302	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.57		
episode: 183	score: 1.0 memory length: 34474	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.57		
episode: 184	score: 1.0 memory length: 34626	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.56		
episode: 185	score: 3.0 memory length: 34874	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.57		
episode: 186	score: 2.0 memory length: 35096	epsilon: 1.0	steps: 222
lr: 0.0001	evaluation reward: 1.55		
episode: 187	score: 3.0 memory length: 35343	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.53		
episode: 188	score: 2.0 memory length: 35545	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.55		
episode: 189	score: 0.0 memory length: 35668	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.54		
episode: 190	score: 3.0 memory length: 35935	epsilon: 1.0	steps: 267

lr: 0.0001	evaluation reward: 1.56		
episode: 191	score: 2.0 memory length: 36156	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.58		
episode: 192	score: 4.0 memory length: 36445	epsilon: 1.0	steps: 289
lr: 0.0001	evaluation reward: 1.61		
episode: 193	score: 1.0 memory length: 36614	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.61		
episode: 194	score: 1.0 memory length: 36765	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.61		
episode: 195	score: 2.0 memory length: 36983	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.63		
episode: 196	score: 1.0 memory length: 37153	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.64		
episode: 197	score: 4.0 memory length: 37410	epsilon: 1.0	steps: 257
lr: 0.0001	evaluation reward: 1.67		
episode: 198	score: 2.0 memory length: 37627	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.67		
episode: 199	score: 1.0 memory length: 37779	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.66		
episode: 200	score: 2.0 memory length: 38001	epsilon: 1.0	steps: 222
lr: 0.0001	evaluation reward: 1.66		
episode: 201	score: 2.0 memory length: 38200	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.66		
episode: 202	score: 2.0 memory length: 38398	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.66		
episode: 203	score: 0.0 memory length: 38522	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.65		
episode: 204	score: 2.0 memory length: 38720	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.65		
episode: 205	score: 4.0 memory length: 39016	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.66		
episode: 206	score: 3.0 memory length: 39283	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.69		
episode: 207	score: 2.0 memory length: 39502	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.7		
episode: 208	score: 3.0 memory length: 39747	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.72		
episode: 209	score: 0.0 memory length: 39870	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7		
episode: 210	score: 3.0 memory length: 40116	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.72		
episode: 211	score: 2.0 memory length: 40335	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.72		
episode: 212	score: 3.0 memory length: 40582	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.72		
episode: 213	score: 0.0 memory length: 40705	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.72		
episode: 214	score: 1.0 memory length: 40857	epsilon: 1.0	steps: 152

lr: 0.0001	evaluation reward: 1.71		
episode: 215	score: 0.0 memory length: 40981	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.68		
episode: 216	score: 2.0 memory length: 41198	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.68		
episode: 217	score: 2.0 memory length: 41397	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.7		
episode: 218	score: 0.0 memory length: 41521	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.69		
episode: 219	score: 1.0 memory length: 41673	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.68		
episode: 220	score: 0.0 memory length: 41796	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.66		
episode: 221	score: 1.0 memory length: 41966	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.66		
episode: 222	score: 1.0 memory length: 42118	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.66		
episode: 223	score: 2.0 memory length: 42321	epsilon: 1.0	steps: 203
lr: 0.0001	evaluation reward: 1.67		
episode: 224	score: 1.0 memory length: 42492	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.65		
episode: 225	score: 1.0 memory length: 42663	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.63		
episode: 226	score: 1.0 memory length: 42817	epsilon: 1.0	steps: 154
lr: 0.0001	evaluation reward: 1.62		
episode: 227	score: 0.0 memory length: 42940	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62		
episode: 228	score: 4.0 memory length: 43200	epsilon: 1.0	steps: 260
lr: 0.0001	evaluation reward: 1.64		
episode: 229	score: 2.0 memory length: 43399	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.62		
episode: 230	score: 2.0 memory length: 43597	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.62		
episode: 231	score: 1.0 memory length: 43769	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.63		
episode: 232	score: 1.0 memory length: 43939	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.64		
episode: 233	score: 1.0 memory length: 44092	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.61		
episode: 234	score: 4.0 memory length: 44348	epsilon: 1.0	steps: 256
lr: 0.0001	evaluation reward: 1.62		
episode: 235	score: 2.0 memory length: 44566	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.63		
episode: 236	score: 1.0 memory length: 44718	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.62		
episode: 237	score: 1.0 memory length: 44870	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.62		
episode: 238	score: 1.0 memory length: 45022	epsilon: 1.0	steps: 152

lr: 0.0001	evaluation reward: 1.59		
episode: 239	score: 0.0 memory length: 45146	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.59		
episode: 240	score: 3.0 memory length: 45375	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.62		
episode: 241	score: 2.0 memory length: 45574	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.63		
episode: 242	score: 2.0 memory length: 45792	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.63		
episode: 243	score: 3.0 memory length: 46059	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.63		
episode: 244	score: 0.0 memory length: 46183	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.63		
episode: 245	score: 3.0 memory length: 46431	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.66		
episode: 246	score: 0.0 memory length: 46555	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.63		
episode: 247	score: 2.0 memory length: 46754	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.65		
episode: 248	score: 1.0 memory length: 46923	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.64		
episode: 249	score: 0.0 memory length: 47047	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.62		
episode: 250	score: 3.0 memory length: 47275	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.63		
episode: 251	score: 1.0 memory length: 47445	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.62		
episode: 252	score: 1.0 memory length: 47615	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.62		
episode: 253	score: 1.0 memory length: 47766	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.61		
episode: 254	score: 4.0 memory length: 48082	epsilon: 1.0	steps: 316
lr: 0.0001	evaluation reward: 1.62		
episode: 255	score: 2.0 memory length: 48300	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.63		
episode: 256	score: 2.0 memory length: 48499	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.65		
episode: 257	score: 2.0 memory length: 48684	epsilon: 1.0	steps: 185
lr: 0.0001	evaluation reward: 1.65		
episode: 258	score: 2.0 memory length: 48903	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.67		
episode: 259	score: 3.0 memory length: 49150	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.69		
episode: 260	score: 2.0 memory length: 49367	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.71		
episode: 261	score: 1.0 memory length: 49538	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.72		
episode: 262	score: 2.0 memory length: 49736	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.74		
episode: 263	score: 2.0 memory length: 49935	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.76		
episode: 264	score: 2.0 memory length: 50121	epsilon: 1.0	steps: 186
lr: 0.0001	evaluation reward: 1.78		
episode: 265	score: 2.0 memory length: 50320	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.79		
episode: 266	score: 1.0 memory length: 50493	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.8		
episode: 267	score: 1.0 memory length: 50663	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.78		
episode: 268	score: 1.0 memory length: 50815	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.77		
episode: 269	score: 0.0 memory length: 50938	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.74		
episode: 270	score: 0.0 memory length: 51062	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.73		
episode: 271	score: 2.0 memory length: 51280	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.73		
episode: 272	score: 0.0 memory length: 51404	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.73		
episode: 273	score: 3.0 memory length: 51649	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.7		
episode: 274	score: 0.0 memory length: 51773	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.65		
episode: 275	score: 1.0 memory length: 51943	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.64		
episode: 276	score: 3.0 memory length: 52189	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.67		
episode: 277	score: 2.0 memory length: 52388	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.68		
episode: 278	score: 0.0 memory length: 52512	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.67		
episode: 279	score: 3.0 memory length: 52777	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.67		
episode: 280	score: 1.0 memory length: 52946	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.65		
episode: 281	score: 0.0 memory length: 53069	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62		
episode: 282	score: 6.0 memory length: 53445	epsilon: 1.0	steps: 376
lr: 0.0001	evaluation reward: 1.66		
episode: 283	score: 1.0 memory length: 53616	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.66		
episode: 284	score: 4.0 memory length: 53933	epsilon: 1.0	steps: 317
lr: 0.0001	evaluation reward: 1.69		
episode: 285	score: 0.0 memory length: 54057	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.66		
episode: 286	score: 1.0 memory length: 54208	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.65		
episode: 287	score: 0.0 memory length: 54332	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.62		
episode: 288	score: 2.0 memory length: 54531	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.62		
episode: 289	score: 1.0 memory length: 54701	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.63		
episode: 290	score: 0.0 memory length: 54825	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.6		
episode: 291	score: 1.0 memory length: 54996	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.59		
episode: 292	score: 2.0 memory length: 55195	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.57		
episode: 293	score: 0.0 memory length: 55319	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.56		
episode: 294	score: 0.0 memory length: 55442	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55		
episode: 295	score: 2.0 memory length: 55640	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.55		
episode: 296	score: 0.0 memory length: 55764	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.54		
episode: 297	score: 2.0 memory length: 55984	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.52		
episode: 298	score: 0.0 memory length: 56108	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.5		
episode: 299	score: 3.0 memory length: 56352	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.52		
episode: 300	score: 3.0 memory length: 56601	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.53		
episode: 301	score: 0.0 memory length: 56725	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.51		
episode: 302	score: 0.0 memory length: 56848	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 303	score: 2.0 memory length: 57067	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.51		
episode: 304	score: 0.0 memory length: 57191	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.49		
episode: 305	score: 2.0 memory length: 57410	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.47		
episode: 306	score: 1.0 memory length: 57562	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.45		
episode: 307	score: 0.0 memory length: 57686	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.43		
episode: 308	score: 1.0 memory length: 57856	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.41		
episode: 309	score: 2.0 memory length: 58077	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.43		
episode: 310	score: 2.0 memory length: 58260	epsilon: 1.0	steps: 183

lr: 0.0001	evaluation reward: 1.42		
episode: 311	score: 1.0 memory length: 58413	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.41		
episode: 312	score: 3.0 memory length: 58682	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.41		
episode: 313	score: 3.0 memory length: 58911	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.44		
episode: 314	score: 3.0 memory length: 59178	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.46		
episode: 315	score: 1.0 memory length: 59330	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.47		
episode: 316	score: 2.0 memory length: 59529	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.47		
episode: 317	score: 0.0 memory length: 59652	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45		
episode: 318	score: 2.0 memory length: 59850	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47		
episode: 319	score: 2.0 memory length: 60052	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.48		
episode: 320	score: 1.0 memory length: 60222	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.49		
episode: 321	score: 3.0 memory length: 60490	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.51		
episode: 322	score: 4.0 memory length: 60766	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.54		
episode: 323	score: 0.0 memory length: 60889	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 324	score: 1.0 memory length: 61040	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.52		
episode: 325	score: 0.0 memory length: 61164	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.51		
episode: 326	score: 2.0 memory length: 61345	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.52		
episode: 327	score: 2.0 memory length: 61544	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.54		
episode: 328	score: 3.0 memory length: 61796	epsilon: 1.0	steps: 252
lr: 0.0001	evaluation reward: 1.53		
episode: 329	score: 3.0 memory length: 62011	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.54		
episode: 330	score: 2.0 memory length: 62210	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.54		
episode: 331	score: 2.0 memory length: 62411	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.55		
episode: 332	score: 3.0 memory length: 62682	epsilon: 1.0	steps: 271
lr: 0.0001	evaluation reward: 1.57		
episode: 333	score: 1.0 memory length: 62852	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.57		
episode: 334	score: 3.0 memory length: 63101	epsilon: 1.0	steps: 249

lr: 0.0001	evaluation reward: 1.56			
episode: 335	score: 1.0	memory length: 63254	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.55			
episode: 336	score: 3.0	memory length: 63519	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.57			
episode: 337	score: 3.0	memory length: 63788	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.59			
episode: 338	score: 2.0	memory length: 64010	epsilon: 1.0	steps: 222
lr: 0.0001	evaluation reward: 1.6			
episode: 339	score: 3.0	memory length: 64275	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.63			
episode: 340	score: 0.0	memory length: 64398	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6			
episode: 341	score: 2.0	memory length: 64596	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6			
episode: 342	score: 4.0	memory length: 64876	epsilon: 1.0	steps: 280
lr: 0.0001	evaluation reward: 1.62			
episode: 343	score: 3.0	memory length: 65126	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.62			
episode: 344	score: 2.0	memory length: 65306	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.64			
episode: 345	score: 1.0	memory length: 65476	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.62			
episode: 346	score: 3.0	memory length: 65741	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.65			
episode: 347	score: 0.0	memory length: 65864	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.63			
episode: 348	score: 3.0	memory length: 66112	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.65			
episode: 349	score: 2.0	memory length: 66330	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.67			
episode: 350	score: 2.0	memory length: 66549	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.66			
episode: 351	score: 3.0	memory length: 66797	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.68			
episode: 352	score: 0.0	memory length: 66920	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67			
episode: 353	score: 2.0	memory length: 67119	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.68			
episode: 354	score: 3.0	memory length: 67364	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.67			
episode: 355	score: 3.0	memory length: 67610	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.68			
episode: 356	score: 1.0	memory length: 67781	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.67			
episode: 357	score: 1.0	memory length: 67934	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.66			
episode: 358	score: 4.0	memory length: 68226	epsilon: 1.0	steps: 292

lr: 0.0001	evaluation reward: 1.68		
episode: 359	score: 0.0 memory length: 68349	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65		
episode: 360	score: 3.0 memory length: 68582	epsilon: 1.0	steps: 233
lr: 0.0001	evaluation reward: 1.66		
episode: 361	score: 0.0 memory length: 68706	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.65		
episode: 362	score: 1.0 memory length: 68877	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.64		
episode: 363	score: 0.0 memory length: 69001	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.62		
episode: 364	score: 5.0 memory length: 69364	epsilon: 1.0	steps: 363
lr: 0.0001	evaluation reward: 1.65		
episode: 365	score: 0.0 memory length: 69488	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.63		
episode: 366	score: 2.0 memory length: 69669	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.64		
episode: 367	score: 2.0 memory length: 69869	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.65		
episode: 368	score: 0.0 memory length: 69992	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64		
episode: 369	score: 0.0 memory length: 70116	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.64		
episode: 370	score: 7.0 memory length: 70503	epsilon: 1.0	steps: 387
lr: 0.0001	evaluation reward: 1.71		
episode: 371	score: 3.0 memory length: 70748	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.72		
episode: 372	score: 0.0 memory length: 70872	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.72		
episode: 373	score: 0.0 memory length: 70996	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.69		
episode: 374	score: 0.0 memory length: 71120	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.69		
episode: 375	score: 2.0 memory length: 71319	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.7		
episode: 376	score: 0.0 memory length: 71442	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67		
episode: 377	score: 3.0 memory length: 71691	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.68		
episode: 378	score: 1.0 memory length: 71860	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.69		
episode: 379	score: 2.0 memory length: 72079	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.68		
episode: 380	score: 5.0 memory length: 72380	epsilon: 1.0	steps: 301
lr: 0.0001	evaluation reward: 1.72		
episode: 381	score: 0.0 memory length: 72503	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.72		
episode: 382	score: 3.0 memory length: 72750	epsilon: 1.0	steps: 247

lr: 0.0001	evaluation reward: 1.69		
episode: 383	score: 4.0 memory length: 73069	epsilon: 1.0	steps: 319
lr: 0.0001	evaluation reward: 1.72		
episode: 384	score: 0.0 memory length: 73193	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.68		
episode: 385	score: 2.0 memory length: 73411	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.7		
episode: 386	score: 2.0 memory length: 73610	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.71		
episode: 387	score: 2.0 memory length: 73812	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.73		
episode: 388	score: 2.0 memory length: 74013	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.73		
episode: 389	score: 2.0 memory length: 74212	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.74		
episode: 390	score: 1.0 memory length: 74381	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.75		
episode: 391	score: 0.0 memory length: 74505	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.74		
episode: 392	score: 3.0 memory length: 74752	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.75		
episode: 393	score: 1.0 memory length: 74924	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.76		
episode: 394	score: 2.0 memory length: 75140	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.78		
episode: 395	score: 5.0 memory length: 75482	epsilon: 1.0	steps: 342
lr: 0.0001	evaluation reward: 1.81		
episode: 396	score: 0.0 memory length: 75606	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.81		
episode: 397	score: 2.0 memory length: 75826	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.81		
episode: 398	score: 1.0 memory length: 75999	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.82		
episode: 399	score: 3.0 memory length: 76263	epsilon: 1.0	steps: 264
lr: 0.0001	evaluation reward: 1.82		
episode: 400	score: 1.0 memory length: 76433	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.8		
episode: 401	score: 3.0 memory length: 76660	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.83		
episode: 402	score: 0.0 memory length: 76783	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.83		
episode: 403	score: 2.0 memory length: 76981	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.83		
episode: 404	score: 4.0 memory length: 77274	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.87		
episode: 405	score: 3.0 memory length: 77540	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.88		
episode: 406	score: 0.0 memory length: 77664	epsilon: 1.0	steps: 124

lr: 0.0001	evaluation reward: 1.87			
episode: 407	score: 2.0	memory length: 77862	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.89			
episode: 408	score: 1.0	memory length: 78032	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.89			
episode: 409	score: 1.0	memory length: 78184	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.88			
episode: 410	score: 2.0	memory length: 78401	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.88			
episode: 411	score: 0.0	memory length: 78525	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.87			
episode: 412	score: 5.0	memory length: 78858	epsilon: 1.0	steps: 333
lr: 0.0001	evaluation reward: 1.89			
episode: 413	score: 0.0	memory length: 78982	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.86			
episode: 414	score: 1.0	memory length: 79153	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.84			
episode: 415	score: 1.0	memory length: 79322	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.84			
episode: 416	score: 1.0	memory length: 79474	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.83			
episode: 417	score: 2.0	memory length: 79693	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.85			
episode: 418	score: 0.0	memory length: 79817	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.83			
episode: 419	score: 1.0	memory length: 79969	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.82			
episode: 420	score: 4.0	memory length: 80265	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.85			
episode: 421	score: 1.0	memory length: 80435	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.83			
episode: 422	score: 2.0	memory length: 80656	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.81			
episode: 423	score: 1.0	memory length: 80827	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.82			
episode: 424	score: 2.0	memory length: 81008	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.83			
episode: 425	score: 2.0	memory length: 81190	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.85			
episode: 426	score: 2.0	memory length: 81409	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.85			
episode: 427	score: 2.0	memory length: 81628	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.85			
episode: 428	score: 0.0	memory length: 81751	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.82			
episode: 429	score: 0.0	memory length: 81875	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.79			
episode: 430	score: 2.0	memory length: 82056	epsilon: 1.0	steps: 181

lr: 0.0001	evaluation reward: 1.79		
episode: 431	score: 3.0 memory length: 82323	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.8		
episode: 432	score: 0.0 memory length: 82447	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.77		
episode: 433	score: 3.0 memory length: 82696	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.79		
episode: 434	score: 0.0 memory length: 82820	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.76		
episode: 435	score: 0.0 memory length: 82944	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.75		
episode: 436	score: 1.0 memory length: 83095	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.73		
episode: 437	score: 1.0 memory length: 83265	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.71		
episode: 438	score: 1.0 memory length: 83418	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.7		
episode: 439	score: 0.0 memory length: 83541	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67		
episode: 440	score: 1.0 memory length: 83711	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.68		
episode: 441	score: 0.0 memory length: 83835	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.66		
episode: 442	score: 2.0 memory length: 84054	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.64		
episode: 443	score: 0.0 memory length: 84177	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61		
episode: 444	score: 0.0 memory length: 84301	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.59		
episode: 445	score: 0.0 memory length: 84424	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.58		
episode: 446	score: 3.0 memory length: 84691	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.58		
episode: 447	score: 0.0 memory length: 84815	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.58		
episode: 448	score: 0.0 memory length: 84939	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.55		
episode: 449	score: 0.0 memory length: 85063	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.53		
episode: 450	score: 3.0 memory length: 85307	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.54		
episode: 451	score: 2.0 memory length: 85506	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.53		
episode: 452	score: 0.0 memory length: 85630	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.53		
episode: 453	score: 0.0 memory length: 85754	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.51		
episode: 454	score: 3.0 memory length: 86000	epsilon: 1.0	steps: 246

lr: 0.0001	evaluation reward: 1.51		
episode: 455	score: 1.0 memory length: 86169	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.49		
episode: 456	score: 1.0 memory length: 86339	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.49		
episode: 457	score: 0.0 memory length: 86463	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.48		
episode: 458	score: 1.0 memory length: 86615	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.45		
episode: 459	score: 5.0 memory length: 86958	epsilon: 1.0	steps: 343
lr: 0.0001	evaluation reward: 1.5		
episode: 460	score: 2.0 memory length: 87157	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.49		
episode: 461	score: 3.0 memory length: 87404	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.52		
episode: 462	score: 1.0 memory length: 87574	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.52		
episode: 463	score: 0.0 memory length: 87698	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.52		
episode: 464	score: 1.0 memory length: 87868	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.48		
episode: 465	score: 2.0 memory length: 88087	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.5		
episode: 466	score: 2.0 memory length: 88306	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.5		
episode: 467	score: 1.0 memory length: 88458	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.49		
episode: 468	score: 0.0 memory length: 88582	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.49		
episode: 469	score: 0.0 memory length: 88705	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 470	score: 2.0 memory length: 88886	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.44		
episode: 471	score: 0.0 memory length: 89010	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.41		
episode: 472	score: 6.0 memory length: 89377	epsilon: 1.0	steps: 367
lr: 0.0001	evaluation reward: 1.47		
episode: 473	score: 0.0 memory length: 89501	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.47		
episode: 474	score: 0.0 memory length: 89625	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.47		
episode: 475	score: 0.0 memory length: 89749	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.45		
episode: 476	score: 5.0 memory length: 90087	epsilon: 1.0	steps: 338
lr: 0.0001	evaluation reward: 1.5		
episode: 477	score: 1.0 memory length: 90259	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.48		
episode: 478	score: 3.0 memory length: 90489	epsilon: 1.0	steps: 230

lr: 0.0001	evaluation reward: 1.5		
episode: 479	score: 3.0 memory length: 90735	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.51		
episode: 480	score: 2.0 memory length: 90936	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.48		
episode: 481	score: 2.0 memory length: 91135	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.5		
episode: 482	score: 0.0 memory length: 91259	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.47		
episode: 483	score: 2.0 memory length: 91478	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.45		
episode: 484	score: 3.0 memory length: 91725	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.48		
episode: 485	score: 2.0 memory length: 91925	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.48		
episode: 486	score: 1.0 memory length: 92077	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.47		
episode: 487	score: 1.0 memory length: 92250	epsilon: 1.0	steps: 173
lr: 0.0001	evaluation reward: 1.46		
episode: 488	score: 2.0 memory length: 92448	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.46		
episode: 489	score: 1.0 memory length: 92618	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.45		
episode: 490	score: 1.0 memory length: 92769	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.45		
episode: 491	score: 0.0 memory length: 92893	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.45		
episode: 492	score: 0.0 memory length: 93016	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.42		
episode: 493	score: 2.0 memory length: 93197	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.43		
episode: 494	score: 1.0 memory length: 93367	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.42		
episode: 495	score: 1.0 memory length: 93538	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.38		
episode: 496	score: 0.0 memory length: 93661	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.38		
episode: 497	score: 0.0 memory length: 93784	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36		
episode: 498	score: 0.0 memory length: 93907	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 499	score: 2.0 memory length: 94108	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.34		
episode: 500	score: 0.0 memory length: 94232	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.33		
episode: 501	score: 1.0 memory length: 94384	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.31		
episode: 502	score: 3.0 memory length: 94629	epsilon: 1.0	steps: 245

lr: 0.0001	evaluation reward: 1.34		
episode: 503	score: 2.0 memory length: 94831	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.34		
episode: 504	score: 0.0 memory length: 94955	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.3		
episode: 505	score: 3.0 memory length: 95224	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.3		
episode: 506	score: 0.0 memory length: 95348	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.3		
episode: 507	score: 1.0 memory length: 95517	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.29		
episode: 508	score: 3.0 memory length: 95785	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.31		
episode: 509	score: 1.0 memory length: 95957	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.31		
episode: 510	score: 3.0 memory length: 96187	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.32		
episode: 511	score: 2.0 memory length: 96385	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.34		
episode: 512	score: 4.0 memory length: 96679	epsilon: 1.0	steps: 294
lr: 0.0001	evaluation reward: 1.33		
episode: 513	score: 1.0 memory length: 96832	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.34		
episode: 514	score: 2.0 memory length: 97050	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.35		
episode: 515	score: 0.0 memory length: 97173	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.34		
episode: 516	score: 1.0 memory length: 97325	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.34		
episode: 517	score: 5.0 memory length: 97652	epsilon: 1.0	steps: 327
lr: 0.0001	evaluation reward: 1.37		
episode: 518	score: 3.0 memory length: 97881	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.4		
episode: 519	score: 0.0 memory length: 98005	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.39		
episode: 520	score: 1.0 memory length: 98177	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.36		
episode: 521	score: 0.0 memory length: 98301	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.35		
episode: 522	score: 2.0 memory length: 98500	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.35		
episode: 523	score: 0.0 memory length: 98624	epsilon: 1.0	steps: 124
lr: 0.0001	evaluation reward: 1.34		
episode: 524	score: 7.0 memory length: 98928	epsilon: 1.0	steps: 304
lr: 0.0001	evaluation reward: 1.39		
episode: 525	score: 0.0 memory length: 99051	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37		
episode: 526	score: 0.0 memory length: 99174	epsilon: 1.0	steps: 123

```

lr: 0.0001      evaluation reward: 1.35
episode: 527    score: 1.0    memory length: 99325    epsilon: 1.0    steps: 151
lr: 0.0001      evaluation reward: 1.34
episode: 528    score: 2.0    memory length: 99524    epsilon: 1.0    steps: 199
lr: 0.0001      evaluation reward: 1.36
episode: 529    score: 2.0    memory length: 99724    epsilon: 1.0    steps: 200
lr: 0.0001      evaluation reward: 1.38
episode: 530    score: 3.0    memory length: 99972    epsilon: 1.0    steps: 248
lr: 0.0001      evaluation reward: 1.39
episode: 531    score: 0.0    memory length: 100096    epsilon: 0.9998079400000042
steps: 124      lr: 0.0001      evaluation reward: 1.36
episode: 532    score: 0.0    memory length: 100220    epsilon: 0.9995624200000095
steps: 124      lr: 0.0001      evaluation reward: 1.36
episode: 533    score: 3.0    memory length: 100447    epsilon: 0.9991129600000193
steps: 227      lr: 0.0001      evaluation reward: 1.36
episode: 534    score: 0.0    memory length: 100570    epsilon: 0.9988694200000245
steps: 123      lr: 0.0001      evaluation reward: 1.36
episode: 535    score: 3.0    memory length: 100839    epsilon: 0.9983368000000361
steps: 269      lr: 0.0001      evaluation reward: 1.39
episode: 536    score: 1.0    memory length: 101010    epsilon: 0.9979982200000435
steps: 171      lr: 0.0001      evaluation reward: 1.39
episode: 537    score: 3.0    memory length: 101255    epsilon: 0.9975131200000054
steps: 245      lr: 0.0001      evaluation reward: 1.41
episode: 538    score: 8.0    memory length: 101650    epsilon: 0.9967310200000071
steps: 395      lr: 0.0001      evaluation reward: 1.48
episode: 539    score: 4.0    memory length: 101927    epsilon: 0.9961825600000829
steps: 277      lr: 0.0001      evaluation reward: 1.52
episode: 540    score: 1.0    memory length: 102079    epsilon: 0.9958816000000894
steps: 152      lr: 0.0001      evaluation reward: 1.52
episode: 541    score: 1.0    memory length: 102231    epsilon: 0.9955806400000959
steps: 152      lr: 0.0001      evaluation reward: 1.53
episode: 542    score: 0.0    memory length: 102355    epsilon: 0.99533512000001013
steps: 124      lr: 0.0001      evaluation reward: 1.51
episode: 543    score: 1.0    memory length: 102507    epsilon: 0.99503416000001078
steps: 152      lr: 0.0001      evaluation reward: 1.52
episode: 544    score: 7.0    memory length: 102935    epsilon: 0.99418672000001262
steps: 428      lr: 0.0001      evaluation reward: 1.59
episode: 545    score: 1.0    memory length: 103086    epsilon: 0.99388774000001327
steps: 151      lr: 0.0001      evaluation reward: 1.6
episode: 546    score: 0.0    memory length: 103210    epsilon: 0.9936422200000138
steps: 124      lr: 0.0001      evaluation reward: 1.57
episode: 547    score: 0.0    memory length: 103333    epsilon: 0.99339868000001433
steps: 123      lr: 0.0001      evaluation reward: 1.57
episode: 548    score: 0.0    memory length: 103457    epsilon: 0.99315316000001486
steps: 124      lr: 0.0001      evaluation reward: 1.57
episode: 549    score: 5.0    memory length: 103783    epsilon: 0.99250768000001627
steps: 326      lr: 0.0001      evaluation reward: 1.62
episode: 550    score: 0.0    memory length: 103907    epsilon: 0.9922621600000168

```

steps: 124	lr: 0.0001	evaluation reward: 1.59
episode: 551	score: 2.0	memory length: 104125 epsilon: 0.9918305200001774
steps: 218	lr: 0.0001	evaluation reward: 1.59
episode: 552	score: 5.0	memory length: 104442 epsilon: 0.991202860000191
steps: 317	lr: 0.0001	evaluation reward: 1.64
episode: 553	score: 1.0	memory length: 104612 epsilon: 0.9908662600001983
steps: 170	lr: 0.0001	evaluation reward: 1.65
episode: 554	score: 0.0	memory length: 104736 epsilon: 0.9906207400002036
steps: 124	lr: 0.0001	evaluation reward: 1.62
episode: 555	score: 3.0	memory length: 104963 epsilon: 0.9901712800002134
steps: 227	lr: 0.0001	evaluation reward: 1.64
episode: 556	score: 2.0	memory length: 105164 epsilon: 0.989773300000222
steps: 201	lr: 0.0001	evaluation reward: 1.65
episode: 557	score: 3.0	memory length: 105415 epsilon: 0.9892763200002328
steps: 251	lr: 0.0001	evaluation reward: 1.68
episode: 558	score: 1.0	memory length: 105566 epsilon: 0.9889773400002393
steps: 151	lr: 0.0001	evaluation reward: 1.68
episode: 559	score: 0.0	memory length: 105689 epsilon: 0.9887338000002446
steps: 123	lr: 0.0001	evaluation reward: 1.63
episode: 560	score: 1.0	memory length: 105859 epsilon: 0.9883972000002519
steps: 170	lr: 0.0001	evaluation reward: 1.62
episode: 561	score: 0.0	memory length: 105983 epsilon: 0.9881516800002572
steps: 124	lr: 0.0001	evaluation reward: 1.59
episode: 562	score: 3.0	memory length: 106234 epsilon: 0.987654700000268
steps: 251	lr: 0.0001	evaluation reward: 1.61
episode: 563	score: 0.0	memory length: 106358 epsilon: 0.9874091800002733
steps: 124	lr: 0.0001	evaluation reward: 1.61
episode: 564	score: 0.0	memory length: 106481 epsilon: 0.9871656400002786
steps: 123	lr: 0.0001	evaluation reward: 1.6
episode: 565	score: 0.0	memory length: 106604 epsilon: 0.9869221000002839
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 566	score: 2.0	memory length: 106820 epsilon: 0.9864944200002932
steps: 216	lr: 0.0001	evaluation reward: 1.58
episode: 567	score: 2.0	memory length: 107039 epsilon: 0.9860608000003026
steps: 219	lr: 0.0001	evaluation reward: 1.59
episode: 568	score: 0.0	memory length: 107162 epsilon: 0.9858172600003079
steps: 123	lr: 0.0001	evaluation reward: 1.59
episode: 569	score: 0.0	memory length: 107286 epsilon: 0.9855717400003132
steps: 124	lr: 0.0001	evaluation reward: 1.59
episode: 570	score: 2.0	memory length: 107484 epsilon: 0.9851797000003217
steps: 198	lr: 0.0001	evaluation reward: 1.59
episode: 571	score: 0.0	memory length: 107608 epsilon: 0.9849341800003271
steps: 124	lr: 0.0001	evaluation reward: 1.59
episode: 572	score: 0.0	memory length: 107732 epsilon: 0.9846886600003324
steps: 124	lr: 0.0001	evaluation reward: 1.53
episode: 573	score: 1.0	memory length: 107902 epsilon: 0.9843520600003397
steps: 170	lr: 0.0001	evaluation reward: 1.54
episode: 574	score: 0.0	memory length: 108026 epsilon: 0.984106540000345

steps: 124	lr: 0.0001	evaluation reward: 1.54
episode: 575	score: 2.0	memory length: 108226 epsilon: 0.9837105400003536
steps: 200	lr: 0.0001	evaluation reward: 1.56
episode: 576	score: 2.0	memory length: 108424 epsilon: 0.9833185000003621
steps: 198	lr: 0.0001	evaluation reward: 1.53
episode: 577	score: 4.0	memory length: 108709 epsilon: 0.9827542000003744
steps: 285	lr: 0.0001	evaluation reward: 1.56
episode: 578	score: 1.0	memory length: 108880 epsilon: 0.9824156200003817
steps: 171	lr: 0.0001	evaluation reward: 1.54
episode: 579	score: 0.0	memory length: 109004 epsilon: 0.9821701000003871
steps: 124	lr: 0.0001	evaluation reward: 1.51
episode: 580	score: 2.0	memory length: 109223 epsilon: 0.9817364800003965
steps: 219	lr: 0.0001	evaluation reward: 1.51
episode: 581	score: 1.0	memory length: 109393 epsilon: 0.9813998800004038
steps: 170	lr: 0.0001	evaluation reward: 1.5
episode: 582	score: 0.0	memory length: 109516 epsilon: 0.9811563400004091
steps: 123	lr: 0.0001	evaluation reward: 1.5
episode: 583	score: 0.0	memory length: 109640 epsilon: 0.9809108200004144
steps: 124	lr: 0.0001	evaluation reward: 1.48
episode: 584	score: 1.0	memory length: 109809 epsilon: 0.9805762000004217
steps: 169	lr: 0.0001	evaluation reward: 1.46
episode: 585	score: 0.0	memory length: 109933 epsilon: 0.980330680000427
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 586	score: 0.0	memory length: 110056 epsilon: 0.9800871400004323
steps: 123	lr: 0.0001	evaluation reward: 1.43
episode: 587	score: 0.0	memory length: 110180 epsilon: 0.9798416200004376
steps: 124	lr: 0.0001	evaluation reward: 1.42
episode: 588	score: 3.0	memory length: 110444 epsilon: 0.979318900000449
steps: 264	lr: 0.0001	evaluation reward: 1.43
episode: 589	score: 0.0	memory length: 110568 epsilon: 0.9790733800004543
steps: 124	lr: 0.0001	evaluation reward: 1.42
episode: 590	score: 1.0	memory length: 110741 epsilon: 0.9787308400004617
steps: 173	lr: 0.0001	evaluation reward: 1.42
episode: 591	score: 1.0	memory length: 110912 epsilon: 0.9783922600004691
steps: 171	lr: 0.0001	evaluation reward: 1.43
episode: 592	score: 1.0	memory length: 111082 epsilon: 0.9780556600004764
steps: 170	lr: 0.0001	evaluation reward: 1.44
episode: 593	score: 3.0	memory length: 111329 epsilon: 0.977566600000487
steps: 247	lr: 0.0001	evaluation reward: 1.45
episode: 594	score: 1.0	memory length: 111480 epsilon: 0.9772676200004935
steps: 151	lr: 0.0001	evaluation reward: 1.45
episode: 595	score: 0.0	memory length: 111603 epsilon: 0.9770240800004988
steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 596	score: 0.0	memory length: 111727 epsilon: 0.9767785600005041
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 597	score: 1.0	memory length: 111898 epsilon: 0.9764399800005115
steps: 171	lr: 0.0001	evaluation reward: 1.45
episode: 598	score: 1.0	memory length: 112068 epsilon: 0.9761033800005188

steps: 170	lr: 0.0001	evaluation reward: 1.46
episode: 599	score: 0.0	memory length: 112191 epsilon: 0.9758598400005241
steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 600	score: 3.0	memory length: 112440 epsilon: 0.9753668200005348
steps: 249	lr: 0.0001	evaluation reward: 1.47
episode: 601	score: 8.0	memory length: 112851 epsilon: 0.9745530400005524
steps: 411	lr: 0.0001	evaluation reward: 1.54
episode: 602	score: 2.0	memory length: 113055 epsilon: 0.9741491200005612
steps: 204	lr: 0.0001	evaluation reward: 1.53
episode: 603	score: 0.0	memory length: 113179 epsilon: 0.9739036000005665
steps: 124	lr: 0.0001	evaluation reward: 1.51
episode: 604	score: 0.0	memory length: 113303 epsilon: 0.9736580800005719
steps: 124	lr: 0.0001	evaluation reward: 1.51
episode: 605	score: 0.0	memory length: 113427 epsilon: 0.9734125600005772
steps: 124	lr: 0.0001	evaluation reward: 1.48
episode: 606	score: 1.0	memory length: 113578 epsilon: 0.9731135800005837
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 607	score: 4.0	memory length: 113872 epsilon: 0.9725314600005963
steps: 294	lr: 0.0001	evaluation reward: 1.52
episode: 608	score: 2.0	memory length: 114071 epsilon: 0.9721374400006049
steps: 199	lr: 0.0001	evaluation reward: 1.51
episode: 609	score: 1.0	memory length: 114241 epsilon: 0.9718008400006122
steps: 170	lr: 0.0001	evaluation reward: 1.51
episode: 610	score: 2.0	memory length: 114460 epsilon: 0.9713672200006216
steps: 219	lr: 0.0001	evaluation reward: 1.5
episode: 611	score: 1.0	memory length: 114611 epsilon: 0.9710682400006281
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 612	score: 3.0	memory length: 114882 epsilon: 0.9705316600006397
steps: 271	lr: 0.0001	evaluation reward: 1.48
episode: 613	score: 1.0	memory length: 115051 epsilon: 0.970197040000647
steps: 169	lr: 0.0001	evaluation reward: 1.48
episode: 614	score: 0.0	memory length: 115174 epsilon: 0.9699535000006523
steps: 123	lr: 0.0001	evaluation reward: 1.46
episode: 615	score: 1.0	memory length: 115344 epsilon: 0.9696169000006596
steps: 170	lr: 0.0001	evaluation reward: 1.47
episode: 616	score: 1.0	memory length: 115513 epsilon: 0.9692822800006669
steps: 169	lr: 0.0001	evaluation reward: 1.47
episode: 617	score: 2.0	memory length: 115715 epsilon: 0.9688823200006755
steps: 202	lr: 0.0001	evaluation reward: 1.44
episode: 618	score: 1.0	memory length: 115888 epsilon: 0.968539780000683
steps: 173	lr: 0.0001	evaluation reward: 1.42
episode: 619	score: 0.0	memory length: 116012 epsilon: 0.9682942600006883
steps: 124	lr: 0.0001	evaluation reward: 1.42
episode: 620	score: 3.0	memory length: 116256 epsilon: 0.9678111400006988
steps: 244	lr: 0.0001	evaluation reward: 1.44
episode: 621	score: 1.0	memory length: 116427 epsilon: 0.9674725600007061
steps: 171	lr: 0.0001	evaluation reward: 1.45
episode: 622	score: 1.0	memory length: 116596 epsilon: 0.9671379400007134

steps: 169	lr: 0.0001	evaluation reward: 1.44
episode: 623	score: 2.0	memory length: 116794 epsilon: 0.9667459000007219
steps: 198	lr: 0.0001	evaluation reward: 1.46
episode: 624	score: 1.0	memory length: 116963 epsilon: 0.9664112800007292
steps: 169	lr: 0.0001	evaluation reward: 1.4
episode: 625	score: 2.0	memory length: 117183 epsilon: 0.9659756800007386
steps: 220	lr: 0.0001	evaluation reward: 1.42
episode: 626	score: 1.0	memory length: 117353 epsilon: 0.9656390800007459
steps: 170	lr: 0.0001	evaluation reward: 1.43
episode: 627	score: 3.0	memory length: 117599 epsilon: 0.9651520000007565
steps: 246	lr: 0.0001	evaluation reward: 1.45
episode: 628	score: 1.0	memory length: 117769 epsilon: 0.9648154000007638
steps: 170	lr: 0.0001	evaluation reward: 1.44
episode: 629	score: 3.0	memory length: 118017 epsilon: 0.9643243600007745
steps: 248	lr: 0.0001	evaluation reward: 1.45
episode: 630	score: 1.0	memory length: 118187 epsilon: 0.9639877600007818
steps: 170	lr: 0.0001	evaluation reward: 1.43
episode: 631	score: 2.0	memory length: 118389 epsilon: 0.9635878000007905
steps: 202	lr: 0.0001	evaluation reward: 1.45
episode: 632	score: 2.0	memory length: 118588 epsilon: 0.963193780000799
steps: 199	lr: 0.0001	evaluation reward: 1.47
episode: 633	score: 2.0	memory length: 118807 epsilon: 0.9627601600008084
steps: 219	lr: 0.0001	evaluation reward: 1.46
episode: 634	score: 1.0	memory length: 118958 epsilon: 0.9624611800008149
steps: 151	lr: 0.0001	evaluation reward: 1.47
episode: 635	score: 2.0	memory length: 119177 epsilon: 0.9620275600008243
steps: 219	lr: 0.0001	evaluation reward: 1.46
episode: 636	score: 3.0	memory length: 119404 epsilon: 0.9615781000008341
steps: 227	lr: 0.0001	evaluation reward: 1.48
episode: 637	score: 3.0	memory length: 119653 epsilon: 0.9610850800008448
steps: 249	lr: 0.0001	evaluation reward: 1.48
episode: 638	score: 4.0	memory length: 119929 epsilon: 0.9605386000008567
steps: 276	lr: 0.0001	evaluation reward: 1.44
episode: 639	score: 0.0	memory length: 120053 epsilon: 0.960293080000862
steps: 124	lr: 0.0001	evaluation reward: 1.4
episode: 640	score: 1.0	memory length: 120224 epsilon: 0.9599545000008693
steps: 171	lr: 0.0001	evaluation reward: 1.4
episode: 641	score: 0.0	memory length: 120348 epsilon: 0.9597089800008747
steps: 124	lr: 0.0001	evaluation reward: 1.39
episode: 642	score: 1.0	memory length: 120518 epsilon: 0.959372380000882
steps: 170	lr: 0.0001	evaluation reward: 1.4
episode: 643	score: 1.0	memory length: 120669 epsilon: 0.9590734000008885
steps: 151	lr: 0.0001	evaluation reward: 1.4
episode: 644	score: 2.0	memory length: 120888 epsilon: 0.9586397800008979
steps: 219	lr: 0.0001	evaluation reward: 1.35
episode: 645	score: 3.0	memory length: 121153 epsilon: 0.9581150800009093
steps: 265	lr: 0.0001	evaluation reward: 1.37
episode: 646	score: 4.0	memory length: 121429 epsilon: 0.9575686000009211

steps: 276	lr: 0.0001	evaluation reward: 1.41
episode: 647	score: 2.0	memory length: 121647 epsilon: 0.9571369600009305
steps: 218	lr: 0.0001	evaluation reward: 1.43
episode: 648	score: 1.0	memory length: 121817 epsilon: 0.9568003600009378
steps: 170	lr: 0.0001	evaluation reward: 1.44
episode: 649	score: 2.0	memory length: 122016 epsilon: 0.9564063400009464
steps: 199	lr: 0.0001	evaluation reward: 1.41
episode: 650	score: 2.0	memory length: 122238 epsilon: 0.9559667800009559
steps: 222	lr: 0.0001	evaluation reward: 1.43
episode: 651	score: 5.0	memory length: 122599 epsilon: 0.9552520000009714
steps: 361	lr: 0.0001	evaluation reward: 1.46
episode: 652	score: 3.0	memory length: 122847 epsilon: 0.9547609600009821
steps: 248	lr: 0.0001	evaluation reward: 1.44
episode: 653	score: 1.0	memory length: 123018 epsilon: 0.9544223800009894
steps: 171	lr: 0.0001	evaluation reward: 1.44
episode: 654	score: 1.0	memory length: 123170 epsilon: 0.954121420000996
steps: 152	lr: 0.0001	evaluation reward: 1.45
episode: 655	score: 4.0	memory length: 123464 epsilon: 0.9535393000010086
steps: 294	lr: 0.0001	evaluation reward: 1.46
episode: 656	score: 2.0	memory length: 123683 epsilon: 0.953105680001018
steps: 219	lr: 0.0001	evaluation reward: 1.46
episode: 657	score: 4.0	memory length: 123982 epsilon: 0.9525136600010309
steps: 299	lr: 0.0001	evaluation reward: 1.47
episode: 658	score: 0.0	memory length: 124106 epsilon: 0.9522681400010362
steps: 124	lr: 0.0001	evaluation reward: 1.46
episode: 659	score: 1.0	memory length: 124279 epsilon: 0.9519256000010436
steps: 173	lr: 0.0001	evaluation reward: 1.47
episode: 660	score: 1.0	memory length: 124450 epsilon: 0.951587020001051
steps: 171	lr: 0.0001	evaluation reward: 1.47
episode: 661	score: 1.0	memory length: 124619 epsilon: 0.9512524000010583
steps: 169	lr: 0.0001	evaluation reward: 1.48
episode: 662	score: 2.0	memory length: 124838 epsilon: 0.9508187800010677
steps: 219	lr: 0.0001	evaluation reward: 1.47
episode: 663	score: 2.0	memory length: 125021 epsilon: 0.9504564400010755
steps: 183	lr: 0.0001	evaluation reward: 1.49
episode: 664	score: 1.0	memory length: 125172 epsilon: 0.950157460001082
steps: 151	lr: 0.0001	evaluation reward: 1.5
episode: 665	score: 0.0	memory length: 125296 epsilon: 0.9499119400010874
steps: 124	lr: 0.0001	evaluation reward: 1.5
episode: 666	score: 0.0	memory length: 125420 epsilon: 0.9496664200010927
steps: 124	lr: 0.0001	evaluation reward: 1.48
episode: 667	score: 0.0	memory length: 125544 epsilon: 0.949420900001098
steps: 124	lr: 0.0001	evaluation reward: 1.46
episode: 668	score: 4.0	memory length: 125839 epsilon: 0.9488368000011107
steps: 295	lr: 0.0001	evaluation reward: 1.5
episode: 669	score: 0.0	memory length: 125963 epsilon: 0.948591280001116
steps: 124	lr: 0.0001	evaluation reward: 1.5
episode: 670	score: 3.0	memory length: 126192 epsilon: 0.9481378600011259

steps: 229	lr: 0.0001	evaluation reward: 1.51
episode: 671	score: 1.0	memory length: 126343 epsilon: 0.9478388800011324
steps: 151	lr: 0.0001	evaluation reward: 1.52
episode: 672	score: 1.0	memory length: 126495 epsilon: 0.9475379200011389
steps: 152	lr: 0.0001	evaluation reward: 1.53
episode: 673	score: 3.0	memory length: 126727 epsilon: 0.9470785600011489
steps: 232	lr: 0.0001	evaluation reward: 1.55
episode: 674	score: 1.0	memory length: 126900 epsilon: 0.9467360200011563
steps: 173	lr: 0.0001	evaluation reward: 1.56
episode: 675	score: 2.0	memory length: 127099 epsilon: 0.9463420000011649
steps: 199	lr: 0.0001	evaluation reward: 1.56
episode: 676	score: 2.0	memory length: 127298 epsilon: 0.9459479800011734
steps: 199	lr: 0.0001	evaluation reward: 1.56
episode: 677	score: 2.0	memory length: 127479 epsilon: 0.9455896000011812
steps: 181	lr: 0.0001	evaluation reward: 1.54
episode: 678	score: 1.0	memory length: 127651 epsilon: 0.9452490400011886
steps: 172	lr: 0.0001	evaluation reward: 1.54
episode: 679	score: 0.0	memory length: 127775 epsilon: 0.9450035200011939
steps: 124	lr: 0.0001	evaluation reward: 1.54
episode: 680	score: 1.0	memory length: 127927 epsilon: 0.9447025600012005
steps: 152	lr: 0.0001	evaluation reward: 1.53
episode: 681	score: 0.0	memory length: 128050 epsilon: 0.9444590200012057
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 682	score: 4.0	memory length: 128312 epsilon: 0.943940260001217
steps: 262	lr: 0.0001	evaluation reward: 1.56
episode: 683	score: 0.0	memory length: 128436 epsilon: 0.9436947400012223
steps: 124	lr: 0.0001	evaluation reward: 1.56
episode: 684	score: 1.0	memory length: 128588 epsilon: 0.9433937800012289
steps: 152	lr: 0.0001	evaluation reward: 1.56
episode: 685	score: 0.0	memory length: 128712 epsilon: 0.9431482600012342
steps: 124	lr: 0.0001	evaluation reward: 1.56
episode: 686	score: 1.0	memory length: 128881 epsilon: 0.9428136400012415
steps: 169	lr: 0.0001	evaluation reward: 1.57
episode: 687	score: 2.0	memory length: 129098 epsilon: 0.9423839800012508
steps: 217	lr: 0.0001	evaluation reward: 1.59
episode: 688	score: 3.0	memory length: 129345 epsilon: 0.9418949200012614
steps: 247	lr: 0.0001	evaluation reward: 1.59
episode: 689	score: 0.0	memory length: 129468 epsilon: 0.9416513800012667
steps: 123	lr: 0.0001	evaluation reward: 1.59
episode: 690	score: 1.0	memory length: 129620 epsilon: 0.9413504200012732
steps: 152	lr: 0.0001	evaluation reward: 1.59
episode: 691	score: 1.0	memory length: 129772 epsilon: 0.9410494600012798
steps: 152	lr: 0.0001	evaluation reward: 1.59
episode: 692	score: 3.0	memory length: 129998 epsilon: 0.9406019800012895
steps: 226	lr: 0.0001	evaluation reward: 1.61
episode: 693	score: 1.0	memory length: 130168 epsilon: 0.9402653800012968
steps: 170	lr: 0.0001	evaluation reward: 1.59
episode: 694	score: 0.0	memory length: 130292 epsilon: 0.9400198600013021

steps: 124	lr: 0.0001	evaluation reward: 1.58
episode: 695	score: 2.0	memory length: 130472 epsilon: 0.9396634600013098
steps: 180	lr: 0.0001	evaluation reward: 1.6
episode: 696	score: 3.0	memory length: 130719 epsilon: 0.9391744000013205
steps: 247	lr: 0.0001	evaluation reward: 1.63
episode: 697	score: 2.0	memory length: 130900 epsilon: 0.9388160200013282
steps: 181	lr: 0.0001	evaluation reward: 1.64
episode: 698	score: 0.0	memory length: 131024 epsilon: 0.9385705000013336
steps: 124	lr: 0.0001	evaluation reward: 1.63
episode: 699	score: 0.0	memory length: 131148 epsilon: 0.9383249800013389
steps: 124	lr: 0.0001	evaluation reward: 1.63
episode: 700	score: 1.0	memory length: 131300 epsilon: 0.9380240200013454
steps: 152	lr: 0.0001	evaluation reward: 1.61
episode: 701	score: 0.0	memory length: 131424 epsilon: 0.9377785000013508
steps: 124	lr: 0.0001	evaluation reward: 1.53
episode: 702	score: 1.0	memory length: 131576 epsilon: 0.9374775400013573
steps: 152	lr: 0.0001	evaluation reward: 1.52
episode: 703	score: 0.0	memory length: 131700 epsilon: 0.9372320200013626
steps: 124	lr: 0.0001	evaluation reward: 1.52
episode: 704	score: 6.0	memory length: 132056 epsilon: 0.9365271400013779
steps: 356	lr: 0.0001	evaluation reward: 1.58
episode: 705	score: 2.0	memory length: 132257 epsilon: 0.9361291600013866
steps: 201	lr: 0.0001	evaluation reward: 1.6
episode: 706	score: 3.0	memory length: 132505 epsilon: 0.9356381200013972
steps: 248	lr: 0.0001	evaluation reward: 1.62
episode: 707	score: 0.0	memory length: 132628 epsilon: 0.9353945800014025
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 708	score: 2.0	memory length: 132827 epsilon: 0.9350005600014111
steps: 199	lr: 0.0001	evaluation reward: 1.58
episode: 709	score: 1.0	memory length: 132998 epsilon: 0.9346619800014184
steps: 171	lr: 0.0001	evaluation reward: 1.58
episode: 710	score: 1.0	memory length: 133150 epsilon: 0.934361020001425
steps: 152	lr: 0.0001	evaluation reward: 1.57
episode: 711	score: 2.0	memory length: 133369 epsilon: 0.9339274000014344
steps: 219	lr: 0.0001	evaluation reward: 1.58
episode: 712	score: 3.0	memory length: 133636 epsilon: 0.9333987400014458
steps: 267	lr: 0.0001	evaluation reward: 1.58
episode: 713	score: 2.0	memory length: 133854 epsilon: 0.9329671000014552
steps: 218	lr: 0.0001	evaluation reward: 1.59
episode: 714	score: 1.0	memory length: 134023 epsilon: 0.9326324800014625
steps: 169	lr: 0.0001	evaluation reward: 1.6
episode: 715	score: 0.0	memory length: 134147 epsilon: 0.9323869600014678
steps: 124	lr: 0.0001	evaluation reward: 1.59
episode: 716	score: 1.0	memory length: 134317 epsilon: 0.9320503600014751
steps: 170	lr: 0.0001	evaluation reward: 1.59
episode: 717	score: 1.0	memory length: 134487 epsilon: 0.9317137600014824
steps: 170	lr: 0.0001	evaluation reward: 1.58
episode: 718	score: 0.0	memory length: 134611 epsilon: 0.9314682400014878

steps: 124	lr: 0.0001	evaluation reward: 1.57
episode: 719	score: 0.0	memory length: 134735 epsilon: 0.9312227200014931
steps: 124	lr: 0.0001	evaluation reward: 1.57
episode: 720	score: 2.0	memory length: 134934 epsilon: 0.9308287000015016
steps: 199	lr: 0.0001	evaluation reward: 1.56
episode: 721	score: 2.0	memory length: 135134 epsilon: 0.9304327000015102
steps: 200	lr: 0.0001	evaluation reward: 1.57
episode: 722	score: 2.0	memory length: 135333 epsilon: 0.9300386800015188
steps: 199	lr: 0.0001	evaluation reward: 1.58
episode: 723	score: 0.0	memory length: 135456 epsilon: 0.9297951400015241
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 724	score: 1.0	memory length: 135625 epsilon: 0.9294605200015313
steps: 169	lr: 0.0001	evaluation reward: 1.56
episode: 725	score: 0.0	memory length: 135749 epsilon: 0.9292150000015367
steps: 124	lr: 0.0001	evaluation reward: 1.54
episode: 726	score: 0.0	memory length: 135872 epsilon: 0.928971460001542
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 727	score: 2.0	memory length: 136076 epsilon: 0.9285675400015507
steps: 204	lr: 0.0001	evaluation reward: 1.52
episode: 728	score: 1.0	memory length: 136246 epsilon: 0.9282309400015558
steps: 170	lr: 0.0001	evaluation reward: 1.52
episode: 729	score: 0.0	memory length: 136370 epsilon: 0.9279854200015634
steps: 124	lr: 0.0001	evaluation reward: 1.49
episode: 730	score: 1.0	memory length: 136521 epsilon: 0.9276864400015699
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 731	score: 2.0	memory length: 136740 epsilon: 0.9272528200015793
steps: 219	lr: 0.0001	evaluation reward: 1.49
episode: 732	score: 0.0	memory length: 136864 epsilon: 0.9270073000015846
steps: 124	lr: 0.0001	evaluation reward: 1.47
episode: 733	score: 0.0	memory length: 136988 epsilon: 0.9267617800015899
steps: 124	lr: 0.0001	evaluation reward: 1.45
episode: 734	score: 1.0	memory length: 137140 epsilon: 0.9264608200015965
steps: 152	lr: 0.0001	evaluation reward: 1.45
episode: 735	score: 1.0	memory length: 137310 epsilon: 0.9261242200016038
steps: 170	lr: 0.0001	evaluation reward: 1.44
episode: 736	score: 3.0	memory length: 137578 epsilon: 0.9255935800016153
steps: 268	lr: 0.0001	evaluation reward: 1.44
episode: 737	score: 1.0	memory length: 137730 epsilon: 0.9252926200016218
steps: 152	lr: 0.0001	evaluation reward: 1.42
episode: 738	score: 3.0	memory length: 137978 epsilon: 0.9248015800016325
steps: 248	lr: 0.0001	evaluation reward: 1.41
episode: 739	score: 2.0	memory length: 138177 epsilon: 0.924407560001641
steps: 199	lr: 0.0001	evaluation reward: 1.43
episode: 740	score: 2.0	memory length: 138376 epsilon: 0.9240135400016496
steps: 199	lr: 0.0001	evaluation reward: 1.44
episode: 741	score: 1.0	memory length: 138548 epsilon: 0.923672980001657
steps: 172	lr: 0.0001	evaluation reward: 1.45
episode: 742	score: 1.0	memory length: 138700 epsilon: 0.9233720200016635

steps: 152	lr: 0.0001	evaluation reward: 1.45
episode: 743	score: 0.0	memory length: 138823 epsilon: 0.9231284800016688
steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 744	score: 1.0	memory length: 138996 epsilon: 0.9227859400016762
steps: 173	lr: 0.0001	evaluation reward: 1.43
episode: 745	score: 1.0	memory length: 139168 epsilon: 0.9224453800016836
steps: 172	lr: 0.0001	evaluation reward: 1.41
episode: 746	score: 0.0	memory length: 139291 epsilon: 0.9222018400016889
steps: 123	lr: 0.0001	evaluation reward: 1.37
episode: 747	score: 2.0	memory length: 139489 epsilon: 0.9218098000016974
steps: 198	lr: 0.0001	evaluation reward: 1.37
episode: 748	score: 1.0	memory length: 139641 epsilon: 0.921508840001704
steps: 152	lr: 0.0001	evaluation reward: 1.37
episode: 749	score: 2.0	memory length: 139858 epsilon: 0.9210791800017133
steps: 217	lr: 0.0001	evaluation reward: 1.37
episode: 750	score: 1.0	memory length: 140027 epsilon: 0.9207445600017206
steps: 169	lr: 0.0001	evaluation reward: 1.36
episode: 751	score: 4.0	memory length: 140326 epsilon: 0.9201525400017334
steps: 299	lr: 0.0001	evaluation reward: 1.35
episode: 752	score: 0.0	memory length: 140450 epsilon: 0.9199070200017387
steps: 124	lr: 0.0001	evaluation reward: 1.32
episode: 753	score: 0.0	memory length: 140573 epsilon: 0.919663480001744
steps: 123	lr: 0.0001	evaluation reward: 1.31
episode: 754	score: 0.0	memory length: 140697 epsilon: 0.9194179600017494
steps: 124	lr: 0.0001	evaluation reward: 1.3
episode: 755	score: 2.0	memory length: 140896 epsilon: 0.9190239400017579
steps: 199	lr: 0.0001	evaluation reward: 1.28
episode: 756	score: 2.0	memory length: 141115 epsilon: 0.9185903200017673
steps: 219	lr: 0.0001	evaluation reward: 1.28
episode: 757	score: 0.0	memory length: 141239 epsilon: 0.9183448000017727
steps: 124	lr: 0.0001	evaluation reward: 1.24
episode: 758	score: 1.0	memory length: 141412 epsilon: 0.9180022600017801
steps: 173	lr: 0.0001	evaluation reward: 1.25
episode: 759	score: 0.0	memory length: 141536 epsilon: 0.9177567400017854
steps: 124	lr: 0.0001	evaluation reward: 1.24
episode: 760	score: 1.0	memory length: 141706 epsilon: 0.9174201400017927
steps: 170	lr: 0.0001	evaluation reward: 1.24
episode: 761	score: 2.0	memory length: 141905 epsilon: 0.9170261200018013
steps: 199	lr: 0.0001	evaluation reward: 1.25
episode: 762	score: 2.0	memory length: 142121 epsilon: 0.9165984400018106
steps: 216	lr: 0.0001	evaluation reward: 1.25
episode: 763	score: 3.0	memory length: 142367 epsilon: 0.9161113600018211
steps: 246	lr: 0.0001	evaluation reward: 1.26
episode: 764	score: 2.0	memory length: 142566 epsilon: 0.9157173400018297
steps: 199	lr: 0.0001	evaluation reward: 1.27
episode: 765	score: 3.0	memory length: 142793 epsilon: 0.9152678800018395
steps: 227	lr: 0.0001	evaluation reward: 1.3
episode: 766	score: 4.0	memory length: 143109 epsilon: 0.914642200001853

steps: 316	lr: 0.0001	evaluation reward: 1.34
episode: 767	score: 3.0	memory length: 143376 epsilon: 0.9141135400018645
steps: 267	lr: 0.0001	evaluation reward: 1.37
episode: 768	score: 0.0	memory length: 143499 epsilon: 0.9138700000018698
steps: 123	lr: 0.0001	evaluation reward: 1.33
episode: 769	score: 4.0	memory length: 143781 epsilon: 0.9133116400018819
steps: 282	lr: 0.0001	evaluation reward: 1.37
episode: 770	score: 0.0	memory length: 143905 epsilon: 0.9130661200018872
steps: 124	lr: 0.0001	evaluation reward: 1.34
episode: 771	score: 1.0	memory length: 144059 epsilon: 0.9127612000018939
steps: 154	lr: 0.0001	evaluation reward: 1.34
episode: 772	score: 1.0	memory length: 144231 epsilon: 0.9124206400019013
steps: 172	lr: 0.0001	evaluation reward: 1.34
episode: 773	score: 2.0	memory length: 144413 epsilon: 0.9120602800019091
steps: 182	lr: 0.0001	evaluation reward: 1.33
episode: 774	score: 3.0	memory length: 144680 epsilon: 0.9115316200019206
steps: 267	lr: 0.0001	evaluation reward: 1.35
episode: 775	score: 0.0	memory length: 144803 epsilon: 0.9112880800019258
steps: 123	lr: 0.0001	evaluation reward: 1.33
episode: 776	score: 1.0	memory length: 144955 epsilon: 0.9109871200019324
steps: 152	lr: 0.0001	evaluation reward: 1.32
episode: 777	score: 2.0	memory length: 145157 epsilon: 0.9105871600019411
steps: 202	lr: 0.0001	evaluation reward: 1.32
episode: 778	score: 2.0	memory length: 145374 epsilon: 0.9101575000019504
steps: 217	lr: 0.0001	evaluation reward: 1.33
episode: 779	score: 2.0	memory length: 145573 epsilon: 0.909763480001959
steps: 199	lr: 0.0001	evaluation reward: 1.35
episode: 780	score: 2.0	memory length: 145756 epsilon: 0.9094011400019668
steps: 183	lr: 0.0001	evaluation reward: 1.36
episode: 781	score: 3.0	memory length: 146022 epsilon: 0.9088744600019782
steps: 266	lr: 0.0001	evaluation reward: 1.39
episode: 782	score: 2.0	memory length: 146239 epsilon: 0.9084448000019876
steps: 217	lr: 0.0001	evaluation reward: 1.37
episode: 783	score: 4.0	memory length: 146533 epsilon: 0.9078626800020002
steps: 294	lr: 0.0001	evaluation reward: 1.41
episode: 784	score: 2.0	memory length: 146750 epsilon: 0.9074330200020095
steps: 217	lr: 0.0001	evaluation reward: 1.42
episode: 785	score: 4.0	memory length: 147011 epsilon: 0.9069162400020208
steps: 261	lr: 0.0001	evaluation reward: 1.46
episode: 786	score: 1.0	memory length: 147181 epsilon: 0.9065796400020281
steps: 170	lr: 0.0001	evaluation reward: 1.46
episode: 787	score: 0.0	memory length: 147305 epsilon: 0.9063341200020334
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 788	score: 1.0	memory length: 147474 epsilon: 0.9059995000020407
steps: 169	lr: 0.0001	evaluation reward: 1.42
episode: 789	score: 0.0	memory length: 147598 epsilon: 0.905753980002046
steps: 124	lr: 0.0001	evaluation reward: 1.42
episode: 790	score: 1.0	memory length: 147768 epsilon: 0.9054173800020533

steps: 170	lr: 0.0001	evaluation reward: 1.42
episode: 791	score: 1.0	memory length: 147938 epsilon: 0.9050807800020606
steps: 170	lr: 0.0001	evaluation reward: 1.42
episode: 792	score: 3.0	memory length: 148184 epsilon: 0.9045937000020712
steps: 246	lr: 0.0001	evaluation reward: 1.42
episode: 793	score: 0.0	memory length: 148308 epsilon: 0.9043481800020765
steps: 124	lr: 0.0001	evaluation reward: 1.41
episode: 794	score: 5.0	memory length: 148655 epsilon: 0.9036611200020914
steps: 347	lr: 0.0001	evaluation reward: 1.46
episode: 795	score: 3.0	memory length: 148881 epsilon: 0.9032136400021011
steps: 226	lr: 0.0001	evaluation reward: 1.47
episode: 796	score: 2.0	memory length: 149098 epsilon: 0.9027839800021105
steps: 217	lr: 0.0001	evaluation reward: 1.46
episode: 797	score: 2.0	memory length: 149317 epsilon: 0.9023503600021199
steps: 219	lr: 0.0001	evaluation reward: 1.46
episode: 798	score: 1.0	memory length: 149468 epsilon: 0.9020513800021264
steps: 151	lr: 0.0001	evaluation reward: 1.47
episode: 799	score: 1.0	memory length: 149620 epsilon: 0.9017504200021329
steps: 152	lr: 0.0001	evaluation reward: 1.48
episode: 800	score: 2.0	memory length: 149818 epsilon: 0.9013583800021414
steps: 198	lr: 0.0001	evaluation reward: 1.49
episode: 801	score: 2.0	memory length: 150017 epsilon: 0.90096436000215
steps: 199	lr: 0.0001	evaluation reward: 1.51
episode: 802	score: 1.0	memory length: 150186 epsilon: 0.9006297400021572
steps: 169	lr: 0.0001	evaluation reward: 1.51
episode: 803	score: 1.0	memory length: 150357 epsilon: 0.9002911600021646
steps: 171	lr: 0.0001	evaluation reward: 1.52
episode: 804	score: 1.0	memory length: 150509 epsilon: 0.8999902000021711
steps: 152	lr: 0.0001	evaluation reward: 1.47
episode: 805	score: 0.0	memory length: 150632 epsilon: 0.8997466600021764
steps: 123	lr: 0.0001	evaluation reward: 1.45
episode: 806	score: 1.0	memory length: 150802 epsilon: 0.8994100600021837
steps: 170	lr: 0.0001	evaluation reward: 1.43
episode: 807	score: 0.0	memory length: 150925 epsilon: 0.899166520002189
steps: 123	lr: 0.0001	evaluation reward: 1.43
episode: 808	score: 3.0	memory length: 151157 epsilon: 0.898707160002199
steps: 232	lr: 0.0001	evaluation reward: 1.44
episode: 809	score: 1.0	memory length: 151327 epsilon: 0.8983705600022063
steps: 170	lr: 0.0001	evaluation reward: 1.44
episode: 810	score: 0.0	memory length: 151451 epsilon: 0.8981250400022116
steps: 124	lr: 0.0001	evaluation reward: 1.43
episode: 811	score: 0.0	memory length: 151574 epsilon: 0.8978815000022169
steps: 123	lr: 0.0001	evaluation reward: 1.41
episode: 812	score: 3.0	memory length: 151823 epsilon: 0.8973884800022276
steps: 249	lr: 0.0001	evaluation reward: 1.41
episode: 813	score: 4.0	memory length: 152121 epsilon: 0.8967984400022404
steps: 298	lr: 0.0001	evaluation reward: 1.43
episode: 814	score: 0.0	memory length: 152245 epsilon: 0.8965529200022457

steps: 124	lr: 0.0001	evaluation reward: 1.42
episode: 815	score: 0.0	memory length: 152368 epsilon: 0.896309380002251
steps: 123	lr: 0.0001	evaluation reward: 1.42
episode: 816	score: 0.0	memory length: 152491 epsilon: 0.8960658400022563
steps: 123	lr: 0.0001	evaluation reward: 1.41
episode: 817	score: 0.0	memory length: 152615 epsilon: 0.8958203200022616
steps: 124	lr: 0.0001	evaluation reward: 1.4
episode: 818	score: 0.0	memory length: 152739 epsilon: 0.895574800002267
steps: 124	lr: 0.0001	evaluation reward: 1.4
episode: 819	score: 0.0	memory length: 152863 epsilon: 0.8953292800022723
steps: 124	lr: 0.0001	evaluation reward: 1.4
episode: 820	score: 1.0	memory length: 153014 epsilon: 0.8950303000022788
steps: 151	lr: 0.0001	evaluation reward: 1.39
episode: 821	score: 3.0	memory length: 153244 epsilon: 0.8945749000022887
steps: 230	lr: 0.0001	evaluation reward: 1.4
episode: 822	score: 0.0	memory length: 153368 epsilon: 0.894329380002294
steps: 124	lr: 0.0001	evaluation reward: 1.38
episode: 823	score: 1.0	memory length: 153520 epsilon: 0.8940284200023005
steps: 152	lr: 0.0001	evaluation reward: 1.39
episode: 824	score: 0.0	memory length: 153644 epsilon: 0.8937829000023059
steps: 124	lr: 0.0001	evaluation reward: 1.38
episode: 825	score: 2.0	memory length: 153865 epsilon: 0.8933453200023154
steps: 221	lr: 0.0001	evaluation reward: 1.4
episode: 826	score: 1.0	memory length: 154017 epsilon: 0.8930443600023219
steps: 152	lr: 0.0001	evaluation reward: 1.41
episode: 827	score: 0.0	memory length: 154140 epsilon: 0.8928008200023272
steps: 123	lr: 0.0001	evaluation reward: 1.39
episode: 828	score: 0.0	memory length: 154263 epsilon: 0.8925572800023325
steps: 123	lr: 0.0001	evaluation reward: 1.38
episode: 829	score: 3.0	memory length: 154511 epsilon: 0.8920662400023431
steps: 248	lr: 0.0001	evaluation reward: 1.41
episode: 830	score: 0.0	memory length: 154635 epsilon: 0.8918207200023485
steps: 124	lr: 0.0001	evaluation reward: 1.4
episode: 831	score: 2.0	memory length: 154834 epsilon: 0.891426700002357
steps: 199	lr: 0.0001	evaluation reward: 1.4
episode: 832	score: 4.0	memory length: 155133 epsilon: 0.8908346800023699
steps: 299	lr: 0.0001	evaluation reward: 1.44
episode: 833	score: 0.0	memory length: 155257 epsilon: 0.8905891600023752
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 834	score: 1.0	memory length: 155409 epsilon: 0.8902882000023817
steps: 152	lr: 0.0001	evaluation reward: 1.44
episode: 835	score: 4.0	memory length: 155669 epsilon: 0.8897734000023929
steps: 260	lr: 0.0001	evaluation reward: 1.47
episode: 836	score: 3.0	memory length: 155917 epsilon: 0.8892823600024036
steps: 248	lr: 0.0001	evaluation reward: 1.47
episode: 837	score: 1.0	memory length: 156087 epsilon: 0.8889457600024109
steps: 170	lr: 0.0001	evaluation reward: 1.47
episode: 838	score: 0.0	memory length: 156211 epsilon: 0.8887002400024162

steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 839	score: 0.0	memory length: 156335 epsilon: 0.8884547200024215
steps: 124	lr: 0.0001	evaluation reward: 1.42
episode: 840	score: 2.0	memory length: 156554 epsilon: 0.888021100002431
steps: 219	lr: 0.0001	evaluation reward: 1.42
episode: 841	score: 0.0	memory length: 156678 epsilon: 0.8877755800024363
steps: 124	lr: 0.0001	evaluation reward: 1.41
episode: 842	score: 3.0	memory length: 156927 epsilon: 0.887282560002447
steps: 249	lr: 0.0001	evaluation reward: 1.43
episode: 843	score: 0.0	memory length: 157051 epsilon: 0.8870370400024523
steps: 124	lr: 0.0001	evaluation reward: 1.43
episode: 844	score: 2.0	memory length: 157250 epsilon: 0.8866430200024609
steps: 199	lr: 0.0001	evaluation reward: 1.44
episode: 845	score: 2.0	memory length: 157433 epsilon: 0.8862806800024687
steps: 183	lr: 0.0001	evaluation reward: 1.45
episode: 846	score: 2.0	memory length: 157631 epsilon: 0.8858886400024772
steps: 198	lr: 0.0001	evaluation reward: 1.47
episode: 847	score: 1.0	memory length: 157801 epsilon: 0.8855520400024846
steps: 170	lr: 0.0001	evaluation reward: 1.46
episode: 848	score: 3.0	memory length: 158066 epsilon: 0.8850273400024959
steps: 265	lr: 0.0001	evaluation reward: 1.48
episode: 849	score: 3.0	memory length: 158313 epsilon: 0.8845382800025066
steps: 247	lr: 0.0001	evaluation reward: 1.49
episode: 850	score: 2.0	memory length: 158531 epsilon: 0.8841066400025159
steps: 218	lr: 0.0001	evaluation reward: 1.5
episode: 851	score: 0.0	memory length: 158655 epsilon: 0.8838611200025213
steps: 124	lr: 0.0001	evaluation reward: 1.46
episode: 852	score: 1.0	memory length: 158825 epsilon: 0.8835245200025286
steps: 170	lr: 0.0001	evaluation reward: 1.47
episode: 853	score: 2.0	memory length: 159023 epsilon: 0.8831324800025371
steps: 198	lr: 0.0001	evaluation reward: 1.49
episode: 854	score: 1.0	memory length: 159175 epsilon: 0.8828315200025436
steps: 152	lr: 0.0001	evaluation reward: 1.5
episode: 855	score: 3.0	memory length: 159422 epsilon: 0.8823424600025542
steps: 247	lr: 0.0001	evaluation reward: 1.51
episode: 856	score: 0.0	memory length: 159546 epsilon: 0.8820969400025596
steps: 124	lr: 0.0001	evaluation reward: 1.49
episode: 857	score: 1.0	memory length: 159716 epsilon: 0.8817603400025669
steps: 170	lr: 0.0001	evaluation reward: 1.5
episode: 858	score: 0.0	memory length: 159839 epsilon: 0.8815168000025722
steps: 123	lr: 0.0001	evaluation reward: 1.49
episode: 859	score: 1.0	memory length: 160010 epsilon: 0.8811782200025795
steps: 171	lr: 0.0001	evaluation reward: 1.5
episode: 860	score: 1.0	memory length: 160182 epsilon: 0.8808376600025869
steps: 172	lr: 0.0001	evaluation reward: 1.5
episode: 861	score: 0.0	memory length: 160306 epsilon: 0.8805921400025922
steps: 124	lr: 0.0001	evaluation reward: 1.48
episode: 862	score: 0.0	memory length: 160430 epsilon: 0.8803466200025976

steps: 124	lr: 0.0001	evaluation reward: 1.46
episode: 863	score: 0.0	memory length: 160554 epsilon: 0.8801011000026029
steps: 124	lr: 0.0001	evaluation reward: 1.43
episode: 864	score: 2.0	memory length: 160753 epsilon: 0.8797070800026114
steps: 199	lr: 0.0001	evaluation reward: 1.43
episode: 865	score: 1.0	memory length: 160905 epsilon: 0.879406120002618
steps: 152	lr: 0.0001	evaluation reward: 1.41
episode: 866	score: 0.0	memory length: 161028 epsilon: 0.8791625800026233
steps: 123	lr: 0.0001	evaluation reward: 1.37
episode: 867	score: 2.0	memory length: 161248 epsilon: 0.8787269800026327
steps: 220	lr: 0.0001	evaluation reward: 1.36
episode: 868	score: 0.0	memory length: 161372 epsilon: 0.878481460002638
steps: 124	lr: 0.0001	evaluation reward: 1.36
episode: 869	score: 0.0	memory length: 161496 epsilon: 0.8782359400026434
steps: 124	lr: 0.0001	evaluation reward: 1.32
episode: 870	score: 1.0	memory length: 161648 epsilon: 0.8779349800026499
steps: 152	lr: 0.0001	evaluation reward: 1.33
episode: 871	score: 4.0	memory length: 161945 epsilon: 0.8773469200026627
steps: 297	lr: 0.0001	evaluation reward: 1.36
episode: 872	score: 2.0	memory length: 162144 epsilon: 0.8769529000026712
steps: 199	lr: 0.0001	evaluation reward: 1.37
episode: 873	score: 0.0	memory length: 162267 epsilon: 0.8767093600026765
steps: 123	lr: 0.0001	evaluation reward: 1.35
episode: 874	score: 0.0	memory length: 162391 epsilon: 0.8764638400026818
steps: 124	lr: 0.0001	evaluation reward: 1.32
episode: 875	score: 2.0	memory length: 162590 epsilon: 0.8760698200026904
steps: 199	lr: 0.0001	evaluation reward: 1.34
episode: 876	score: 3.0	memory length: 162817 epsilon: 0.8756203600027002
steps: 227	lr: 0.0001	evaluation reward: 1.36
episode: 877	score: 0.0	memory length: 162941 epsilon: 0.8753748400027055
steps: 124	lr: 0.0001	evaluation reward: 1.34
episode: 878	score: 0.0	memory length: 163065 epsilon: 0.8751293200027108
steps: 124	lr: 0.0001	evaluation reward: 1.32
episode: 879	score: 2.0	memory length: 163264 epsilon: 0.8747353000027194
steps: 199	lr: 0.0001	evaluation reward: 1.32
episode: 880	score: 1.0	memory length: 163416 epsilon: 0.8744343400027259
steps: 152	lr: 0.0001	evaluation reward: 1.31
episode: 881	score: 5.0	memory length: 163760 epsilon: 0.8737532200027407
steps: 344	lr: 0.0001	evaluation reward: 1.33
episode: 882	score: 2.0	memory length: 163941 epsilon: 0.8733948400027485
steps: 181	lr: 0.0001	evaluation reward: 1.33
episode: 883	score: 2.0	memory length: 164159 epsilon: 0.8729632000027578
steps: 218	lr: 0.0001	evaluation reward: 1.31
episode: 884	score: 2.0	memory length: 164358 epsilon: 0.8725691800027664
steps: 199	lr: 0.0001	evaluation reward: 1.31
episode: 885	score: 1.0	memory length: 164527 epsilon: 0.8722345600027737
steps: 169	lr: 0.0001	evaluation reward: 1.28
episode: 886	score: 2.0	memory length: 164746 epsilon: 0.8718009400027831

steps: 219	lr: 0.0001	evaluation reward: 1.29	
episode: 887	score: 1.0	memory length: 164915	epsilon: 0.8714663200027903
steps: 169	lr: 0.0001	evaluation reward: 1.3	
episode: 888	score: 0.0	memory length: 165039	epsilon: 0.8712208000027957
steps: 124	lr: 0.0001	evaluation reward: 1.29	
episode: 889	score: 1.0	memory length: 165190	epsilon: 0.8709218200028022
steps: 151	lr: 0.0001	evaluation reward: 1.3	
episode: 890	score: 2.0	memory length: 165407	epsilon: 0.8704921600028115
steps: 217	lr: 0.0001	evaluation reward: 1.31	
episode: 891	score: 0.0	memory length: 165531	epsilon: 0.8702466400028168
steps: 124	lr: 0.0001	evaluation reward: 1.3	
episode: 892	score: 3.0	memory length: 165780	epsilon: 0.8697536200028275
steps: 249	lr: 0.0001	evaluation reward: 1.3	
episode: 893	score: 0.0	memory length: 165903	epsilon: 0.8695100800028328
steps: 123	lr: 0.0001	evaluation reward: 1.3	
episode: 894	score: 0.0	memory length: 166027	epsilon: 0.8692645600028381
steps: 124	lr: 0.0001	evaluation reward: 1.25	
episode: 895	score: 0.0	memory length: 166150	epsilon: 0.8690210200028434
steps: 123	lr: 0.0001	evaluation reward: 1.22	
episode: 896	score: 0.0	memory length: 166274	epsilon: 0.8687755000028488
steps: 124	lr: 0.0001	evaluation reward: 1.2	
episode: 897	score: 4.0	memory length: 166562	epsilon: 0.8682052600028611
steps: 288	lr: 0.0001	evaluation reward: 1.22	
episode: 898	score: 3.0	memory length: 166789	epsilon: 0.8677558000028709
steps: 227	lr: 0.0001	evaluation reward: 1.24	
episode: 899	score: 2.0	memory length: 167008	epsilon: 0.8673221800028803
steps: 219	lr: 0.0001	evaluation reward: 1.25	
episode: 900	score: 0.0	memory length: 167132	epsilon: 0.8670766600028856
steps: 124	lr: 0.0001	evaluation reward: 1.23	
episode: 901	score: 0.0	memory length: 167256	epsilon: 0.866831140002891
steps: 124	lr: 0.0001	evaluation reward: 1.21	
episode: 902	score: 1.0	memory length: 167426	epsilon: 0.8664945400028983
steps: 170	lr: 0.0001	evaluation reward: 1.21	
episode: 903	score: 1.0	memory length: 167599	epsilon: 0.8661520000029057
steps: 173	lr: 0.0001	evaluation reward: 1.21	
episode: 904	score: 1.0	memory length: 167768	epsilon: 0.865817380002913
steps: 169	lr: 0.0001	evaluation reward: 1.21	
episode: 905	score: 1.0	memory length: 167941	epsilon: 0.8654748400029204
steps: 173	lr: 0.0001	evaluation reward: 1.22	
episode: 906	score: 2.0	memory length: 168142	epsilon: 0.865076860002929
steps: 201	lr: 0.0001	evaluation reward: 1.23	
episode: 907	score: 2.0	memory length: 168341	epsilon: 0.8646828400029376
steps: 199	lr: 0.0001	evaluation reward: 1.25	
episode: 908	score: 0.0	memory length: 168465	epsilon: 0.8644373200029429
steps: 124	lr: 0.0001	evaluation reward: 1.22	
episode: 909	score: 1.0	memory length: 168636	epsilon: 0.8640987400029503
steps: 171	lr: 0.0001	evaluation reward: 1.22	
episode: 910	score: 1.0	memory length: 168788	epsilon: 0.8637977800029568

steps: 152	lr: 0.0001	evaluation reward: 1.23
episode: 911	score: 3.0	memory length: 169035 epsilon: 0.8633087200029674
steps: 247	lr: 0.0001	evaluation reward: 1.26
episode: 912	score: 2.0	memory length: 169218 epsilon: 0.8629463800029753
steps: 183	lr: 0.0001	evaluation reward: 1.25
episode: 913	score: 2.0	memory length: 169437 epsilon: 0.8625127600029847
steps: 219	lr: 0.0001	evaluation reward: 1.23
episode: 914	score: 3.0	memory length: 169654 epsilon: 0.862083100002994
steps: 217	lr: 0.0001	evaluation reward: 1.26
episode: 915	score: 4.0	memory length: 169949 epsilon: 0.8614990000030067
steps: 295	lr: 0.0001	evaluation reward: 1.3
episode: 916	score: 0.0	memory length: 170073 epsilon: 0.861253480003012
steps: 124	lr: 0.0001	evaluation reward: 1.3
episode: 917	score: 2.0	memory length: 170292 epsilon: 0.8608198600030215
steps: 219	lr: 0.0001	evaluation reward: 1.32
episode: 918	score: 3.0	memory length: 170541 epsilon: 0.8603268400030322
steps: 249	lr: 0.0001	evaluation reward: 1.35
episode: 919	score: 4.0	memory length: 170818 epsilon: 0.8597783800030441
steps: 277	lr: 0.0001	evaluation reward: 1.39
episode: 920	score: 0.0	memory length: 170942 epsilon: 0.8595328600030494
steps: 124	lr: 0.0001	evaluation reward: 1.38
episode: 921	score: 1.0	memory length: 171115 epsilon: 0.8591903200030568
steps: 173	lr: 0.0001	evaluation reward: 1.36
episode: 922	score: 2.0	memory length: 171314 epsilon: 0.8587963000030654
steps: 199	lr: 0.0001	evaluation reward: 1.38
episode: 923	score: 3.0	memory length: 171562 epsilon: 0.858305260003076
steps: 248	lr: 0.0001	evaluation reward: 1.4
episode: 924	score: 3.0	memory length: 171808 epsilon: 0.8578181800030866
steps: 246	lr: 0.0001	evaluation reward: 1.43
episode: 925	score: 4.0	memory length: 172101 epsilon: 0.8572380400030992
steps: 293	lr: 0.0001	evaluation reward: 1.45
episode: 926	score: 0.0	memory length: 172224 epsilon: 0.8569945000031045
steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 927	score: 0.0	memory length: 172348 epsilon: 0.8567489800031098
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 928	score: 3.0	memory length: 172595 epsilon: 0.8562599200031205
steps: 247	lr: 0.0001	evaluation reward: 1.47
episode: 929	score: 0.0	memory length: 172719 epsilon: 0.8560144000031258
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 930	score: 0.0	memory length: 172843 epsilon: 0.8557688800031311
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 931	score: 2.0	memory length: 173042 epsilon: 0.8553748600031397
steps: 199	lr: 0.0001	evaluation reward: 1.44
episode: 932	score: 2.0	memory length: 173261 epsilon: 0.8549412400031491
steps: 219	lr: 0.0001	evaluation reward: 1.42
episode: 933	score: 2.0	memory length: 173463 epsilon: 0.8545412800031578
steps: 202	lr: 0.0001	evaluation reward: 1.44
episode: 934	score: 4.0	memory length: 173742 epsilon: 0.8539888600031698

steps: 279	lr: 0.0001	evaluation reward: 1.47
episode: 935	score: 2.0	memory length: 173924 epsilon: 0.8536285000031776
steps: 182	lr: 0.0001	evaluation reward: 1.45
episode: 936	score: 2.0	memory length: 174123 epsilon: 0.8532344800031861
steps: 199	lr: 0.0001	evaluation reward: 1.44
episode: 937	score: 1.0	memory length: 174294 epsilon: 0.8528959000031935
steps: 171	lr: 0.0001	evaluation reward: 1.44
episode: 938	score: 0.0	memory length: 174418 epsilon: 0.8526503800031988
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 939	score: 0.0	memory length: 174542 epsilon: 0.8524048600032041
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 940	score: 5.0	memory length: 174866 epsilon: 0.8517633400032181
steps: 324	lr: 0.0001	evaluation reward: 1.47
episode: 941	score: 0.0	memory length: 174990 epsilon: 0.8515178200032234
steps: 124	lr: 0.0001	evaluation reward: 1.47
episode: 942	score: 3.0	memory length: 175235 epsilon: 0.8510327200032339
steps: 245	lr: 0.0001	evaluation reward: 1.47
episode: 943	score: 0.0	memory length: 175358 epsilon: 0.8507891800032392
steps: 123	lr: 0.0001	evaluation reward: 1.47
episode: 944	score: 3.0	memory length: 175605 epsilon: 0.8503001200032498
steps: 247	lr: 0.0001	evaluation reward: 1.48
episode: 945	score: 0.0	memory length: 175729 epsilon: 0.8500546000032552
steps: 124	lr: 0.0001	evaluation reward: 1.46
episode: 946	score: 0.0	memory length: 175852 epsilon: 0.8498110600032605
steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 947	score: 0.0	memory length: 175976 epsilon: 0.8495655400032658
steps: 124	lr: 0.0001	evaluation reward: 1.43
episode: 948	score: 2.0	memory length: 176195 epsilon: 0.8491319200032752
steps: 219	lr: 0.0001	evaluation reward: 1.42
episode: 949	score: 1.0	memory length: 176365 epsilon: 0.8487953200032825
steps: 170	lr: 0.0001	evaluation reward: 1.4
episode: 950	score: 5.0	memory length: 176711 epsilon: 0.8481102400032974
steps: 346	lr: 0.0001	evaluation reward: 1.43
episode: 951	score: 1.0	memory length: 176882 epsilon: 0.8477716600033047
steps: 171	lr: 0.0001	evaluation reward: 1.44
episode: 952	score: 1.0	memory length: 177051 epsilon: 0.847437040003312
steps: 169	lr: 0.0001	evaluation reward: 1.44
episode: 953	score: 1.0	memory length: 177223 epsilon: 0.8470964800033194
steps: 172	lr: 0.0001	evaluation reward: 1.43
episode: 954	score: 0.0	memory length: 177347 epsilon: 0.8468509600033247
steps: 124	lr: 0.0001	evaluation reward: 1.42
episode: 955	score: 0.0	memory length: 177470 epsilon: 0.84660742000333
steps: 123	lr: 0.0001	evaluation reward: 1.39
episode: 956	score: 0.0	memory length: 177594 epsilon: 0.8463619000033353
steps: 124	lr: 0.0001	evaluation reward: 1.39
episode: 957	score: 2.0	memory length: 177793 epsilon: 0.8459678800033439
steps: 199	lr: 0.0001	evaluation reward: 1.4
episode: 958	score: 1.0	memory length: 177966 epsilon: 0.8456253400033513

steps: 173	lr: 0.0001	evaluation reward: 1.41
episode: 959	score: 1.0	memory length: 178135 epsilon: 0.8452907200033586
steps: 169	lr: 0.0001	evaluation reward: 1.41
episode: 960	score: 0.0	memory length: 178259 epsilon: 0.8450452000033639
steps: 124	lr: 0.0001	evaluation reward: 1.4
episode: 961	score: 2.0	memory length: 178441 epsilon: 0.8446848400033717
steps: 182	lr: 0.0001	evaluation reward: 1.42
episode: 962	score: 2.0	memory length: 178640 epsilon: 0.8442908200033803
steps: 199	lr: 0.0001	evaluation reward: 1.44
episode: 963	score: 3.0	memory length: 178887 epsilon: 0.8438017600033909
steps: 247	lr: 0.0001	evaluation reward: 1.47
episode: 964	score: 0.0	memory length: 179011 epsilon: 0.8435562400033962
steps: 124	lr: 0.0001	evaluation reward: 1.45
episode: 965	score: 0.0	memory length: 179135 epsilon: 0.8433107200034016
steps: 124	lr: 0.0001	evaluation reward: 1.44
episode: 966	score: 1.0	memory length: 179287 epsilon: 0.8430097600034081
steps: 152	lr: 0.0001	evaluation reward: 1.45
episode: 967	score: 5.0	memory length: 179626 epsilon: 0.8423385400034227
steps: 339	lr: 0.0001	evaluation reward: 1.48
episode: 968	score: 4.0	memory length: 179945 epsilon: 0.8417069200034364
steps: 319	lr: 0.0001	evaluation reward: 1.52
episode: 969	score: 0.0	memory length: 180068 epsilon: 0.8414633800034417
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 970	score: 1.0	memory length: 180240 epsilon: 0.8411228200034491
steps: 172	lr: 0.0001	evaluation reward: 1.52
episode: 971	score: 0.0	memory length: 180363 epsilon: 0.8408792800034544
steps: 123	lr: 0.0001	evaluation reward: 1.48
episode: 972	score: 2.0	memory length: 180566 epsilon: 0.8404773400034631
steps: 203	lr: 0.0001	evaluation reward: 1.48
episode: 973	score: 3.0	memory length: 180816 epsilon: 0.8399823400034738
steps: 250	lr: 0.0001	evaluation reward: 1.51
episode: 974	score: 7.0	memory length: 181191 epsilon: 0.83923984000349
steps: 375	lr: 0.0001	evaluation reward: 1.58
episode: 975	score: 2.0	memory length: 181389 epsilon: 0.8388478000034985
steps: 198	lr: 0.0001	evaluation reward: 1.58
episode: 976	score: 3.0	memory length: 181615 epsilon: 0.8384003200035082
steps: 226	lr: 0.0001	evaluation reward: 1.58
episode: 977	score: 2.0	memory length: 181814 epsilon: 0.8380063000035167
steps: 199	lr: 0.0001	evaluation reward: 1.6
episode: 978	score: 0.0	memory length: 181937 epsilon: 0.837762760003522
steps: 123	lr: 0.0001	evaluation reward: 1.6
episode: 979	score: 2.0	memory length: 182135 epsilon: 0.8373707200035305
steps: 198	lr: 0.0001	evaluation reward: 1.6
episode: 980	score: 2.0	memory length: 182334 epsilon: 0.8369767000035391
steps: 199	lr: 0.0001	evaluation reward: 1.61
episode: 981	score: 3.0	memory length: 182561 epsilon: 0.8365272400035488
steps: 227	lr: 0.0001	evaluation reward: 1.59
episode: 982	score: 2.0	memory length: 182760 epsilon: 0.8361332200035574

steps: 199	lr: 0.0001	evaluation reward: 1.59	
episode: 983	score: 3.0	memory length: 183007	epsilon: 0.835644160003568
steps: 247	lr: 0.0001	evaluation reward: 1.6	
episode: 984	score: 1.0	memory length: 183159	epsilon: 0.8353432000035745
steps: 152	lr: 0.0001	evaluation reward: 1.59	
episode: 985	score: 3.0	memory length: 183429	epsilon: 0.8348086000035861
steps: 270	lr: 0.0001	evaluation reward: 1.61	
episode: 986	score: 2.0	memory length: 183628	epsilon: 0.8344145800035947
steps: 199	lr: 0.0001	evaluation reward: 1.61	
episode: 987	score: 2.0	memory length: 183827	epsilon: 0.8340205600036033
steps: 199	lr: 0.0001	evaluation reward: 1.62	
episode: 988	score: 1.0	memory length: 183996	epsilon: 0.8336859400036105
steps: 169	lr: 0.0001	evaluation reward: 1.63	
episode: 989	score: 1.0	memory length: 184148	epsilon: 0.833384980003617
steps: 152	lr: 0.0001	evaluation reward: 1.63	
episode: 990	score: 2.0	memory length: 184347	epsilon: 0.8329909600036256
steps: 199	lr: 0.0001	evaluation reward: 1.63	
episode: 991	score: 0.0	memory length: 184471	epsilon: 0.8327454400036309
steps: 124	lr: 0.0001	evaluation reward: 1.63	
episode: 992	score: 0.0	memory length: 184595	epsilon: 0.8324999200036363
steps: 124	lr: 0.0001	evaluation reward: 1.6	
episode: 993	score: 4.0	memory length: 184891	epsilon: 0.831913840003649
steps: 296	lr: 0.0001	evaluation reward: 1.64	
episode: 994	score: 1.0	memory length: 185062	epsilon: 0.8315752600036563
steps: 171	lr: 0.0001	evaluation reward: 1.65	
episode: 995	score: 2.0	memory length: 185261	epsilon: 0.8311812400036649
steps: 199	lr: 0.0001	evaluation reward: 1.67	
episode: 996	score: 2.0	memory length: 185460	epsilon: 0.8307872200036734
steps: 199	lr: 0.0001	evaluation reward: 1.69	
episode: 997	score: 2.0	memory length: 185679	epsilon: 0.8303536000036829
steps: 219	lr: 0.0001	evaluation reward: 1.67	
episode: 998	score: 2.0	memory length: 185900	epsilon: 0.8299160200036924
steps: 221	lr: 0.0001	evaluation reward: 1.66	
episode: 999	score: 2.0	memory length: 186083	epsilon: 0.8295536800037002
steps: 183	lr: 0.0001	evaluation reward: 1.66	
episode: 1000	score: 5.0	memory length: 186429	epsilon: 0.8288686000037151
steps: 346	lr: 0.0001	evaluation reward: 1.71	
episode: 1001	score: 2.0	memory length: 186652	epsilon: 0.8284270600037247
steps: 223	lr: 0.0001	evaluation reward: 1.73	
episode: 1002	score: 1.0	memory length: 186804	epsilon: 0.8281261000037312
steps: 152	lr: 0.0001	evaluation reward: 1.73	
episode: 1003	score: 2.0	memory length: 187021	epsilon: 0.8276964400037405
steps: 217	lr: 0.0001	evaluation reward: 1.74	
episode: 1004	score: 6.0	memory length: 187396	epsilon: 0.8269539400037567
steps: 375	lr: 0.0001	evaluation reward: 1.79	
episode: 1005	score: 1.0	memory length: 187565	epsilon: 0.8266193200037639
steps: 169	lr: 0.0001	evaluation reward: 1.79	
episode: 1006	score: 2.0	memory length: 187784	epsilon: 0.8261857000037733

steps: 219	lr: 0.0001	evaluation reward: 1.79	
episode: 1007	score: 3.0	memory length: 188030	epsilon: 0.8256986200037839
steps: 246	lr: 0.0001	evaluation reward: 1.8	
episode: 1008	score: 2.0	memory length: 188212	epsilon: 0.8253382600037917
steps: 182	lr: 0.0001	evaluation reward: 1.82	
episode: 1009	score: 3.0	memory length: 188460	epsilon: 0.8248472200038024
steps: 248	lr: 0.0001	evaluation reward: 1.84	
episode: 1010	score: 1.0	memory length: 188612	epsilon: 0.8245462600038089
steps: 152	lr: 0.0001	evaluation reward: 1.84	
episode: 1011	score: 2.0	memory length: 188811	epsilon: 0.8241522400038175
steps: 199	lr: 0.0001	evaluation reward: 1.83	
episode: 1012	score: 3.0	memory length: 189079	epsilon: 0.823621600003829
steps: 268	lr: 0.0001	evaluation reward: 1.84	
episode: 1013	score: 3.0	memory length: 189347	epsilon: 0.8230909600038405
steps: 268	lr: 0.0001	evaluation reward: 1.85	
episode: 1014	score: 2.0	memory length: 189546	epsilon: 0.8226969400038491
steps: 199	lr: 0.0001	evaluation reward: 1.84	
episode: 1015	score: 2.0	memory length: 189745	epsilon: 0.8223029200038576
steps: 199	lr: 0.0001	evaluation reward: 1.82	
episode: 1016	score: 1.0	memory length: 189896	epsilon: 0.8220039400038641
steps: 151	lr: 0.0001	evaluation reward: 1.83	
episode: 1017	score: 1.0	memory length: 190065	epsilon: 0.8216693200038714
steps: 169	lr: 0.0001	evaluation reward: 1.82	
episode: 1018	score: 2.0	memory length: 190287	epsilon: 0.8212297600038809
steps: 222	lr: 0.0001	evaluation reward: 1.81	
episode: 1019	score: 3.0	memory length: 190517	epsilon: 0.8207743600038908
steps: 230	lr: 0.0001	evaluation reward: 1.8	
episode: 1020	score: 2.0	memory length: 190715	epsilon: 0.8203823200038993
steps: 198	lr: 0.0001	evaluation reward: 1.82	
episode: 1021	score: 5.0	memory length: 191081	epsilon: 0.8196576400039151
steps: 366	lr: 0.0001	evaluation reward: 1.86	
episode: 1022	score: 1.0	memory length: 191251	epsilon: 0.8193210400039224
steps: 170	lr: 0.0001	evaluation reward: 1.85	
episode: 1023	score: 5.0	memory length: 191574	epsilon: 0.8186815000039362
steps: 323	lr: 0.0001	evaluation reward: 1.87	
episode: 1024	score: 1.0	memory length: 191744	epsilon: 0.8183449000039436
steps: 170	lr: 0.0001	evaluation reward: 1.85	
episode: 1025	score: 2.0	memory length: 191927	epsilon: 0.8179825600039514
steps: 183	lr: 0.0001	evaluation reward: 1.83	
episode: 1026	score: 4.0	memory length: 192223	epsilon: 0.8173964800039641
steps: 296	lr: 0.0001	evaluation reward: 1.87	
episode: 1027	score: 5.0	memory length: 192569	epsilon: 0.816711400003979
steps: 346	lr: 0.0001	evaluation reward: 1.92	
episode: 1028	score: 5.0	memory length: 192878	epsilon: 0.8160995800039923
steps: 309	lr: 0.0001	evaluation reward: 1.94	
episode: 1029	score: 2.0	memory length: 193077	epsilon: 0.8157055600040009
steps: 199	lr: 0.0001	evaluation reward: 1.96	
episode: 1030	score: 4.0	memory length: 193373	epsilon: 0.8151194800040136

steps: 296	lr: 0.0001	evaluation reward: 2.0	
episode: 1031	score: 3.0	memory length: 193640	epsilon: 0.814590820004025
steps: 267	lr: 0.0001	evaluation reward: 2.01	
episode: 1032	score: 2.0	memory length: 193839	epsilon: 0.8141968000040336
steps: 199	lr: 0.0001	evaluation reward: 2.01	
episode: 1033	score: 0.0	memory length: 193962	epsilon: 0.8139532600040389
steps: 123	lr: 0.0001	evaluation reward: 1.99	
episode: 1034	score: 2.0	memory length: 194145	epsilon: 0.8135909200040468
steps: 183	lr: 0.0001	evaluation reward: 1.97	
episode: 1035	score: 2.0	memory length: 194344	epsilon: 0.8131969000040553
steps: 199	lr: 0.0001	evaluation reward: 1.97	
episode: 1036	score: 0.0	memory length: 194467	epsilon: 0.8129533600040606
steps: 123	lr: 0.0001	evaluation reward: 1.95	
episode: 1037	score: 2.0	memory length: 194666	epsilon: 0.8125593400040692
steps: 199	lr: 0.0001	evaluation reward: 1.96	
episode: 1038	score: 2.0	memory length: 194884	epsilon: 0.8121277000040785
steps: 218	lr: 0.0001	evaluation reward: 1.98	
episode: 1039	score: 3.0	memory length: 195136	epsilon: 0.8116287400040894
steps: 252	lr: 0.0001	evaluation reward: 2.01	
episode: 1040	score: 4.0	memory length: 195415	epsilon: 0.8110763200041013
steps: 279	lr: 0.0001	evaluation reward: 2.0	
episode: 1041	score: 1.0	memory length: 195584	epsilon: 0.8107417000041086
steps: 169	lr: 0.0001	evaluation reward: 2.01	
episode: 1042	score: 1.0	memory length: 195755	epsilon: 0.810403120004116
steps: 171	lr: 0.0001	evaluation reward: 1.99	
episode: 1043	score: 0.0	memory length: 195878	epsilon: 0.8101595800041212
steps: 123	lr: 0.0001	evaluation reward: 1.99	
episode: 1044	score: 4.0	memory length: 196138	epsilon: 0.8096447800041324
steps: 260	lr: 0.0001	evaluation reward: 2.0	
episode: 1045	score: 2.0	memory length: 196341	epsilon: 0.8092428400041412
steps: 203	lr: 0.0001	evaluation reward: 2.02	
episode: 1046	score: 2.0	memory length: 196540	epsilon: 0.8088488200041497
steps: 199	lr: 0.0001	evaluation reward: 2.04	
episode: 1047	score: 3.0	memory length: 196787	epsilon: 0.8083597600041603
steps: 247	lr: 0.0001	evaluation reward: 2.07	
episode: 1048	score: 1.0	memory length: 196958	epsilon: 0.8080211800041677
steps: 171	lr: 0.0001	evaluation reward: 2.06	
episode: 1049	score: 1.0	memory length: 197110	epsilon: 0.8077202200041742
steps: 152	lr: 0.0001	evaluation reward: 2.06	
episode: 1050	score: 3.0	memory length: 197357	epsilon: 0.8072311600041848
steps: 247	lr: 0.0001	evaluation reward: 2.04	
episode: 1051	score: 2.0	memory length: 197555	epsilon: 0.8068391200041933
steps: 198	lr: 0.0001	evaluation reward: 2.05	
episode: 1052	score: 1.0	memory length: 197727	epsilon: 0.8064985600042007
steps: 172	lr: 0.0001	evaluation reward: 2.05	
episode: 1053	score: 0.0	memory length: 197851	epsilon: 0.8062530400042061
steps: 124	lr: 0.0001	evaluation reward: 2.04	
episode: 1054	score: 3.0	memory length: 198103	epsilon: 0.8057540800042169

steps: 252	lr: 0.0001	evaluation reward: 2.07	
episode: 1055	score: 3.0	memory length: 198330	epsilon: 0.8053046200042266
steps: 227	lr: 0.0001	evaluation reward: 2.1	
episode: 1056	score: 1.0	memory length: 198482	epsilon: 0.8050036600042332
steps: 152	lr: 0.0001	evaluation reward: 2.11	
episode: 1057	score: 0.0	memory length: 198606	epsilon: 0.8047581400042385
steps: 124	lr: 0.0001	evaluation reward: 2.09	
episode: 1058	score: 0.0	memory length: 198729	epsilon: 0.8045146000042438
steps: 123	lr: 0.0001	evaluation reward: 2.08	
episode: 1059	score: 2.0	memory length: 198928	epsilon: 0.8041205800042523
steps: 199	lr: 0.0001	evaluation reward: 2.09	
episode: 1060	score: 0.0	memory length: 199052	epsilon: 0.8038750600042577
steps: 124	lr: 0.0001	evaluation reward: 2.09	
episode: 1061	score: 1.0	memory length: 199204	epsilon: 0.8035741000042642
steps: 152	lr: 0.0001	evaluation reward: 2.08	
episode: 1062	score: 3.0	memory length: 199474	epsilon: 0.8030395000042758
steps: 270	lr: 0.0001	evaluation reward: 2.09	
episode: 1063	score: 3.0	memory length: 199719	epsilon: 0.8025544000042864
steps: 245	lr: 0.0001	evaluation reward: 2.09	
episode: 1064	score: 2.0	memory length: 199920	epsilon: 0.802156420004295
steps: 201	lr: 0.0001	evaluation reward: 2.11	
episode: 1065	score: 2.0	memory length: 200119	epsilon: 0.8017624000043035
steps: 199	lr: 4e-05	evaluation reward: 2.13	
episode: 1066	score: 0.0	memory length: 200242	epsilon: 0.8015188600043088
steps: 123	lr: 4e-05	evaluation reward: 2.12	
episode: 1067	score: 1.0	memory length: 200412	epsilon: 0.8011822600043161
steps: 170	lr: 4e-05	evaluation reward: 2.08	
episode: 1068	score: 4.0	memory length: 200688	epsilon: 0.800635780004328
steps: 276	lr: 4e-05	evaluation reward: 2.08	
episode: 1069	score: 1.0	memory length: 200840	epsilon: 0.8003348200043345
steps: 152	lr: 4e-05	evaluation reward: 2.09	
episode: 1070	score: 5.0	memory length: 201189	epsilon: 0.7996438000043495
steps: 349	lr: 4e-05	evaluation reward: 2.13	
episode: 1071	score: 3.0	memory length: 201415	epsilon: 0.7991963200043593
steps: 226	lr: 4e-05	evaluation reward: 2.16	
episode: 1072	score: 5.0	memory length: 201739	epsilon: 0.7985548000043732
steps: 324	lr: 4e-05	evaluation reward: 2.19	
episode: 1073	score: 2.0	memory length: 201956	epsilon: 0.7981251400043825
steps: 217	lr: 4e-05	evaluation reward: 2.18	
episode: 1074	score: 2.0	memory length: 202138	epsilon: 0.7977647800043903
steps: 182	lr: 4e-05	evaluation reward: 2.13	
episode: 1075	score: 3.0	memory length: 202387	epsilon: 0.797271760004401
steps: 249	lr: 4e-05	evaluation reward: 2.14	
episode: 1076	score: 1.0	memory length: 202539	epsilon: 0.7969708000044076
steps: 152	lr: 4e-05	evaluation reward: 2.12	
episode: 1077	score: 6.0	memory length: 202917	epsilon: 0.7962223600044238
steps: 378	lr: 4e-05	evaluation reward: 2.16	
episode: 1078	score: 4.0	memory length: 203193	epsilon: 0.7956758800044357

steps: 276	lr: 4e-05	evaluation reward: 2.2	
episode: 1079	score: 3.0	memory length: 203425	epsilon: 0.7952165200044456
steps: 232	lr: 4e-05	evaluation reward: 2.21	
episode: 1080	score: 3.0	memory length: 203636	epsilon: 0.7947987400044547
steps: 211	lr: 4e-05	evaluation reward: 2.22	
episode: 1081	score: 3.0	memory length: 203866	epsilon: 0.7943433400044646
steps: 230	lr: 4e-05	evaluation reward: 2.22	
episode: 1082	score: 1.0	memory length: 204036	epsilon: 0.7940067400044719
steps: 170	lr: 4e-05	evaluation reward: 2.21	
episode: 1083	score: 2.0	memory length: 204216	epsilon: 0.7936503400044796
steps: 180	lr: 4e-05	evaluation reward: 2.2	
episode: 1084	score: 4.0	memory length: 204492	epsilon: 0.7931038600044915
steps: 276	lr: 4e-05	evaluation reward: 2.23	
episode: 1085	score: 2.0	memory length: 204694	epsilon: 0.7927039000045002
steps: 202	lr: 4e-05	evaluation reward: 2.22	
episode: 1086	score: 2.0	memory length: 204894	epsilon: 0.7923079000045088
steps: 200	lr: 4e-05	evaluation reward: 2.22	
episode: 1087	score: 1.0	memory length: 205064	epsilon: 0.7919713000045161
steps: 170	lr: 4e-05	evaluation reward: 2.21	
episode: 1088	score: 0.0	memory length: 205188	epsilon: 0.7917257800045214
steps: 124	lr: 4e-05	evaluation reward: 2.2	
episode: 1089	score: 3.0	memory length: 205418	epsilon: 0.7912703800045313
steps: 230	lr: 4e-05	evaluation reward: 2.22	
episode: 1090	score: 3.0	memory length: 205666	epsilon: 0.790779340004542
steps: 248	lr: 4e-05	evaluation reward: 2.23	
episode: 1091	score: 2.0	memory length: 205849	epsilon: 0.7904170000045498
steps: 183	lr: 4e-05	evaluation reward: 2.25	
episode: 1092	score: 4.0	memory length: 206128	epsilon: 0.7898645800045618
steps: 279	lr: 4e-05	evaluation reward: 2.29	
episode: 1093	score: 0.0	memory length: 206252	epsilon: 0.7896190600045672
steps: 124	lr: 4e-05	evaluation reward: 2.25	
episode: 1094	score: 3.0	memory length: 206520	epsilon: 0.7890884200045787
steps: 268	lr: 4e-05	evaluation reward: 2.27	
episode: 1095	score: 1.0	memory length: 206690	epsilon: 0.788751820004586
steps: 170	lr: 4e-05	evaluation reward: 2.26	
episode: 1096	score: 3.0	memory length: 206919	epsilon: 0.7882984000045958
steps: 229	lr: 4e-05	evaluation reward: 2.27	
episode: 1097	score: 4.0	memory length: 207221	epsilon: 0.7877004400046088
steps: 302	lr: 4e-05	evaluation reward: 2.29	
episode: 1098	score: 3.0	memory length: 207466	epsilon: 0.7872153400046193
steps: 245	lr: 4e-05	evaluation reward: 2.3	
episode: 1099	score: 0.0	memory length: 207590	epsilon: 0.7869698200046247
steps: 124	lr: 4e-05	evaluation reward: 2.28	
episode: 1100	score: 2.0	memory length: 207790	epsilon: 0.7865738200046333
steps: 200	lr: 4e-05	evaluation reward: 2.25	
episode: 1101	score: 3.0	memory length: 208017	epsilon: 0.786124360004643
steps: 227	lr: 4e-05	evaluation reward: 2.26	
episode: 1102	score: 0.0	memory length: 208141	epsilon: 0.7858788400046484

steps: 124	lr: 4e-05	evaluation reward: 2.25	
episode: 1103	score: 1.0	memory length: 208311	epsilon: 0.7855422400046557
steps: 170	lr: 4e-05	evaluation reward: 2.24	
episode: 1104	score: 3.0	memory length: 208558	epsilon: 0.7850531800046663
steps: 247	lr: 4e-05	evaluation reward: 2.21	
episode: 1105	score: 2.0	memory length: 208740	epsilon: 0.7846928200046741
steps: 182	lr: 4e-05	evaluation reward: 2.22	
episode: 1106	score: 2.0	memory length: 208958	epsilon: 0.7842611800046835
steps: 218	lr: 4e-05	evaluation reward: 2.22	
episode: 1107	score: 4.0	memory length: 209236	epsilon: 0.7837107400046954
steps: 278	lr: 4e-05	evaluation reward: 2.23	
episode: 1108	score: 1.0	memory length: 209409	epsilon: 0.7833682000047029
steps: 173	lr: 4e-05	evaluation reward: 2.22	
episode: 1109	score: 0.0	memory length: 209532	epsilon: 0.7831246600047082
steps: 123	lr: 4e-05	evaluation reward: 2.19	
episode: 1110	score: 2.0	memory length: 209715	epsilon: 0.782762320004716
steps: 183	lr: 4e-05	evaluation reward: 2.2	
episode: 1111	score: 3.0	memory length: 209944	epsilon: 0.7823089000047259
steps: 229	lr: 4e-05	evaluation reward: 2.21	
episode: 1112	score: 5.0	memory length: 210277	epsilon: 0.7816495600047402
steps: 333	lr: 4e-05	evaluation reward: 2.23	
episode: 1113	score: 1.0	memory length: 210429	epsilon: 0.7813486000047467
steps: 152	lr: 4e-05	evaluation reward: 2.21	
episode: 1114	score: 1.0	memory length: 210581	epsilon: 0.7810476400047532
steps: 152	lr: 4e-05	evaluation reward: 2.2	
episode: 1115	score: 2.0	memory length: 210780	epsilon: 0.7806536200047618
steps: 199	lr: 4e-05	evaluation reward: 2.2	
episode: 1116	score: 2.0	memory length: 210960	epsilon: 0.7802972200047695
steps: 180	lr: 4e-05	evaluation reward: 2.21	
episode: 1117	score: 3.0	memory length: 211186	epsilon: 0.7798497400047792
steps: 226	lr: 4e-05	evaluation reward: 2.23	
episode: 1118	score: 1.0	memory length: 211338	epsilon: 0.7795487800047858
steps: 152	lr: 4e-05	evaluation reward: 2.22	
episode: 1119	score: 2.0	memory length: 211537	epsilon: 0.7791547600047943
steps: 199	lr: 4e-05	evaluation reward: 2.21	
episode: 1120	score: 0.0	memory length: 211661	epsilon: 0.7789092400047997
steps: 124	lr: 4e-05	evaluation reward: 2.19	
episode: 1121	score: 3.0	memory length: 211889	epsilon: 0.7784578000048095
steps: 228	lr: 4e-05	evaluation reward: 2.17	
episode: 1122	score: 3.0	memory length: 212157	epsilon: 0.777927160004821
steps: 268	lr: 4e-05	evaluation reward: 2.19	
episode: 1123	score: 1.0	memory length: 212328	epsilon: 0.7775885800048283
steps: 171	lr: 4e-05	evaluation reward: 2.15	
episode: 1124	score: 2.0	memory length: 212527	epsilon: 0.7771945600048369
steps: 199	lr: 4e-05	evaluation reward: 2.16	
episode: 1125	score: 2.0	memory length: 212728	epsilon: 0.7767965800048455
steps: 201	lr: 4e-05	evaluation reward: 2.16	
episode: 1126	score: 2.0	memory length: 212927	epsilon: 0.7764025600048541

steps: 199	lr: 4e-05	evaluation reward: 2.14	
episode: 1127	score: 3.0	memory length: 213137	epsilon: 0.7759867600048631
steps: 210	lr: 4e-05	evaluation reward: 2.12	
episode: 1128	score: 2.0	memory length: 213360	epsilon: 0.7755452200048727
steps: 223	lr: 4e-05	evaluation reward: 2.09	
episode: 1129	score: 1.0	memory length: 213512	epsilon: 0.7752442600048792
steps: 152	lr: 4e-05	evaluation reward: 2.08	
episode: 1130	score: 5.0	memory length: 213837	epsilon: 0.7746007600048932
steps: 325	lr: 4e-05	evaluation reward: 2.09	
episode: 1131	score: 3.0	memory length: 214051	epsilon: 0.7741770400049024
steps: 214	lr: 4e-05	evaluation reward: 2.09	
episode: 1132	score: 4.0	memory length: 214332	epsilon: 0.7736206600049145
steps: 281	lr: 4e-05	evaluation reward: 2.11	
episode: 1133	score: 2.0	memory length: 214533	epsilon: 0.7732226800049231
steps: 201	lr: 4e-05	evaluation reward: 2.13	
episode: 1134	score: 2.0	memory length: 214716	epsilon: 0.772860340004931
steps: 183	lr: 4e-05	evaluation reward: 2.13	
episode: 1135	score: 4.0	memory length: 214983	epsilon: 0.7723316800049425
steps: 267	lr: 4e-05	evaluation reward: 2.15	
episode: 1136	score: 1.0	memory length: 215154	epsilon: 0.7719931000049498
steps: 171	lr: 4e-05	evaluation reward: 2.16	
episode: 1137	score: 3.0	memory length: 215381	epsilon: 0.7715436400049596
steps: 227	lr: 4e-05	evaluation reward: 2.17	
episode: 1138	score: 1.0	memory length: 215554	epsilon: 0.771201100004967
steps: 173	lr: 4e-05	evaluation reward: 2.16	
episode: 1139	score: 6.0	memory length: 215881	epsilon: 0.770553640004981
steps: 327	lr: 4e-05	evaluation reward: 2.19	
episode: 1140	score: 2.0	memory length: 216062	epsilon: 0.7701952600049888
steps: 181	lr: 4e-05	evaluation reward: 2.17	
episode: 1141	score: 1.0	memory length: 216232	epsilon: 0.7698586600049961
steps: 170	lr: 4e-05	evaluation reward: 2.17	
episode: 1142	score: 3.0	memory length: 216458	epsilon: 0.7694111800050059
steps: 226	lr: 4e-05	evaluation reward: 2.19	
episode: 1143	score: 3.0	memory length: 216685	epsilon: 0.7689617200050156
steps: 227	lr: 4e-05	evaluation reward: 2.22	
episode: 1144	score: 4.0	memory length: 216982	epsilon: 0.7683736600050284
steps: 297	lr: 4e-05	evaluation reward: 2.22	
episode: 1145	score: 3.0	memory length: 217230	epsilon: 0.767882620005039
steps: 248	lr: 4e-05	evaluation reward: 2.23	
episode: 1146	score: 3.0	memory length: 217475	epsilon: 0.7673975200050496
steps: 245	lr: 4e-05	evaluation reward: 2.24	
episode: 1147	score: 3.0	memory length: 217691	epsilon: 0.7669698400050589
steps: 216	lr: 4e-05	evaluation reward: 2.24	
episode: 1148	score: 4.0	memory length: 217967	epsilon: 0.7664233600050707
steps: 276	lr: 4e-05	evaluation reward: 2.27	
episode: 1149	score: 6.0	memory length: 218353	epsilon: 0.7656590800050873
steps: 386	lr: 4e-05	evaluation reward: 2.32	
episode: 1150	score: 2.0	memory length: 218535	epsilon: 0.7652987200050951

steps: 182	lr: 4e-05	evaluation reward: 2.31	
episode: 1151	score: 3.0	memory length: 218804	epsilon: 0.7647661000051067
steps: 269	lr: 4e-05	evaluation reward: 2.32	
episode: 1152	score: 3.0	memory length: 219018	epsilon: 0.7643423800051159
steps: 214	lr: 4e-05	evaluation reward: 2.34	
episode: 1153	score: 0.0	memory length: 219142	epsilon: 0.7640968600051212
steps: 124	lr: 4e-05	evaluation reward: 2.34	
episode: 1154	score: 3.0	memory length: 219368	epsilon: 0.7636493800051309
steps: 226	lr: 4e-05	evaluation reward: 2.34	
episode: 1155	score: 2.0	memory length: 219567	epsilon: 0.7632553600051395
steps: 199	lr: 4e-05	evaluation reward: 2.33	
episode: 1156	score: 4.0	memory length: 219882	epsilon: 0.762631660005153
steps: 315	lr: 4e-05	evaluation reward: 2.36	
episode: 1157	score: 4.0	memory length: 220160	epsilon: 0.762081220005165
steps: 278	lr: 4e-05	evaluation reward: 2.4	
episode: 1158	score: 3.0	memory length: 220386	epsilon: 0.7616337400051747
steps: 226	lr: 4e-05	evaluation reward: 2.43	
episode: 1159	score: 0.0	memory length: 220510	epsilon: 0.76138822000518
steps: 124	lr: 4e-05	evaluation reward: 2.41	
episode: 1160	score: 4.0	memory length: 220804	epsilon: 0.7608061000051927
steps: 294	lr: 4e-05	evaluation reward: 2.45	
episode: 1161	score: 2.0	memory length: 221007	epsilon: 0.7604041600052014
steps: 203	lr: 4e-05	evaluation reward: 2.46	
episode: 1162	score: 0.0	memory length: 221131	epsilon: 0.7601586400052067
steps: 124	lr: 4e-05	evaluation reward: 2.43	
episode: 1163	score: 1.0	memory length: 221282	epsilon: 0.7598596600052132
steps: 151	lr: 4e-05	evaluation reward: 2.41	
episode: 1164	score: 3.0	memory length: 221497	epsilon: 0.7594339600052225
steps: 215	lr: 4e-05	evaluation reward: 2.42	
episode: 1165	score: 2.0	memory length: 221678	epsilon: 0.7590755800052302
steps: 181	lr: 4e-05	evaluation reward: 2.42	
episode: 1166	score: 0.0	memory length: 221802	epsilon: 0.7588300600052356
steps: 124	lr: 4e-05	evaluation reward: 2.42	
episode: 1167	score: 4.0	memory length: 222078	epsilon: 0.7582835800052474
steps: 276	lr: 4e-05	evaluation reward: 2.45	
episode: 1168	score: 4.0	memory length: 222376	epsilon: 0.7576935400052602
steps: 298	lr: 4e-05	evaluation reward: 2.45	
episode: 1169	score: 4.0	memory length: 222638	epsilon: 0.7571747800052715
steps: 262	lr: 4e-05	evaluation reward: 2.48	
episode: 1170	score: 5.0	memory length: 222963	epsilon: 0.7565312800052855
steps: 325	lr: 4e-05	evaluation reward: 2.48	
episode: 1171	score: 2.0	memory length: 223182	epsilon: 0.7560976600052949
steps: 219	lr: 4e-05	evaluation reward: 2.47	
episode: 1172	score: 6.0	memory length: 223527	epsilon: 0.7554145600053097
steps: 345	lr: 4e-05	evaluation reward: 2.48	
episode: 1173	score: 2.0	memory length: 223749	epsilon: 0.7549750000053193
steps: 222	lr: 4e-05	evaluation reward: 2.48	
episode: 1174	score: 2.0	memory length: 223948	epsilon: 0.7545809800053278

steps: 199	lr: 4e-05	evaluation reward: 2.48	
episode: 1175	score: 1.0	memory length: 224100	epsilon: 0.7542800200053343
steps: 152	lr: 4e-05	evaluation reward: 2.46	
episode: 1176	score: 3.0	memory length: 224313	epsilon: 0.7538582800053435
steps: 213	lr: 4e-05	evaluation reward: 2.48	
episode: 1177	score: 1.0	memory length: 224465	epsilon: 0.75355732000535
steps: 152	lr: 4e-05	evaluation reward: 2.43	
episode: 1178	score: 2.0	memory length: 224664	epsilon: 0.7531633000053586
steps: 199	lr: 4e-05	evaluation reward: 2.41	
episode: 1179	score: 2.0	memory length: 224862	epsilon: 0.7527712600053671
steps: 198	lr: 4e-05	evaluation reward: 2.4	
episode: 1180	score: 5.0	memory length: 225132	epsilon: 0.7522366600053787
steps: 270	lr: 4e-05	evaluation reward: 2.42	
episode: 1181	score: 1.0	memory length: 225284	epsilon: 0.7519357000053852
steps: 152	lr: 4e-05	evaluation reward: 2.4	
episode: 1182	score: 3.0	memory length: 225498	epsilon: 0.7515119800053944
steps: 214	lr: 4e-05	evaluation reward: 2.42	
episode: 1183	score: 3.0	memory length: 225744	epsilon: 0.751024900005405
steps: 246	lr: 4e-05	evaluation reward: 2.43	
episode: 1184	score: 3.0	memory length: 225991	epsilon: 0.7505358400054156
steps: 247	lr: 4e-05	evaluation reward: 2.42	
episode: 1185	score: 3.0	memory length: 226237	epsilon: 0.7500487600054262
steps: 246	lr: 4e-05	evaluation reward: 2.43	
episode: 1186	score: 2.0	memory length: 226439	epsilon: 0.7496488000054349
steps: 202	lr: 4e-05	evaluation reward: 2.43	
episode: 1187	score: 4.0	memory length: 226719	epsilon: 0.7490944000054469
steps: 280	lr: 4e-05	evaluation reward: 2.46	
episode: 1188	score: 9.0	memory length: 227095	epsilon: 0.7483499200054631
steps: 376	lr: 4e-05	evaluation reward: 2.55	
episode: 1189	score: 4.0	memory length: 227370	epsilon: 0.7478054200054749
steps: 275	lr: 4e-05	evaluation reward: 2.56	
episode: 1190	score: 2.0	memory length: 227569	epsilon: 0.7474114000054835
steps: 199	lr: 4e-05	evaluation reward: 2.55	
episode: 1191	score: 4.0	memory length: 227845	epsilon: 0.7468649200054953
steps: 276	lr: 4e-05	evaluation reward: 2.57	
episode: 1192	score: 2.0	memory length: 228064	epsilon: 0.7464313000055047
steps: 219	lr: 4e-05	evaluation reward: 2.55	
episode: 1193	score: 2.0	memory length: 228263	epsilon: 0.7460372800055133
steps: 199	lr: 4e-05	evaluation reward: 2.57	
episode: 1194	score: 4.0	memory length: 228540	epsilon: 0.7454888200055252
steps: 277	lr: 4e-05	evaluation reward: 2.58	
episode: 1195	score: 2.0	memory length: 228741	epsilon: 0.7450908400055338
steps: 201	lr: 4e-05	evaluation reward: 2.59	
episode: 1196	score: 1.0	memory length: 228913	epsilon: 0.7447502800055412
steps: 172	lr: 4e-05	evaluation reward: 2.57	
episode: 1197	score: 0.0	memory length: 229037	epsilon: 0.7445047600055466
steps: 124	lr: 4e-05	evaluation reward: 2.53	
episode: 1198	score: 1.0	memory length: 229189	epsilon: 0.7442038000055531

steps: 152	lr: 4e-05	evaluation reward: 2.51	
episode: 1199	score: 0.0	memory length: 229313	epsilon: 0.7439582800055584
steps: 124	lr: 4e-05	evaluation reward: 2.51	
episode: 1200	score: 7.0	memory length: 229681	epsilon: 0.7432296400055742
steps: 368	lr: 4e-05	evaluation reward: 2.56	
episode: 1201	score: 3.0	memory length: 229908	epsilon: 0.7427801800055584
steps: 227	lr: 4e-05	evaluation reward: 2.56	
episode: 1202	score: 3.0	memory length: 230175	epsilon: 0.7422515200055955
steps: 267	lr: 4e-05	evaluation reward: 2.59	
episode: 1203	score: 4.0	memory length: 230438	epsilon: 0.7417307800056068
steps: 263	lr: 4e-05	evaluation reward: 2.62	
episode: 1204	score: 0.0	memory length: 230562	epsilon: 0.7414852600056121
steps: 124	lr: 4e-05	evaluation reward: 2.59	
episode: 1205	score: 3.0	memory length: 230789	epsilon: 0.7410358000056219
steps: 227	lr: 4e-05	evaluation reward: 2.6	
episode: 1206	score: 1.0	memory length: 230941	epsilon: 0.7407348400056284
steps: 152	lr: 4e-05	evaluation reward: 2.59	
episode: 1207	score: 4.0	memory length: 231184	epsilon: 0.7402537000056388
steps: 243	lr: 4e-05	evaluation reward: 2.59	
episode: 1208	score: 2.0	memory length: 231383	epsilon: 0.7398596800056474
steps: 199	lr: 4e-05	evaluation reward: 2.6	
episode: 1209	score: 4.0	memory length: 231642	epsilon: 0.7393468600056585
steps: 259	lr: 4e-05	evaluation reward: 2.64	
episode: 1210	score: 1.0	memory length: 231793	epsilon: 0.739047880005665
steps: 151	lr: 4e-05	evaluation reward: 2.63	
episode: 1211	score: 6.0	memory length: 232143	epsilon: 0.7383548800056801
steps: 350	lr: 4e-05	evaluation reward: 2.66	
episode: 1212	score: 3.0	memory length: 232408	epsilon: 0.7378301800056914
steps: 265	lr: 4e-05	evaluation reward: 2.64	
episode: 1213	score: 3.0	memory length: 232637	epsilon: 0.7373767600057013
steps: 229	lr: 4e-05	evaluation reward: 2.66	
episode: 1214	score: 4.0	memory length: 232913	epsilon: 0.7368302800057132
steps: 276	lr: 4e-05	evaluation reward: 2.69	
episode: 1215	score: 7.0	memory length: 233337	epsilon: 0.7359907600057314
steps: 424	lr: 4e-05	evaluation reward: 2.74	
episode: 1216	score: 4.0	memory length: 233618	epsilon: 0.7354343800057435
steps: 281	lr: 4e-05	evaluation reward: 2.76	
episode: 1217	score: 1.0	memory length: 233770	epsilon: 0.73513342000575
steps: 152	lr: 4e-05	evaluation reward: 2.74	
episode: 1218	score: 1.0	memory length: 233940	epsilon: 0.7347968200057573
steps: 170	lr: 4e-05	evaluation reward: 2.74	
episode: 1219	score: 2.0	memory length: 234141	epsilon: 0.7343988400057659
steps: 201	lr: 4e-05	evaluation reward: 2.74	
episode: 1220	score: 4.0	memory length: 234408	epsilon: 0.7338701800057774
steps: 267	lr: 4e-05	evaluation reward: 2.78	
episode: 1221	score: 3.0	memory length: 234654	epsilon: 0.733383100005788
steps: 246	lr: 4e-05	evaluation reward: 2.78	
episode: 1222	score: 8.0	memory length: 235131	epsilon: 0.7324386400058085

steps: 477	lr: 4e-05	evaluation reward: 2.83	
episode: 1223	score: 6.0	memory length: 235501	epsilon: 0.7317060400058244
steps: 370	lr: 4e-05	evaluation reward: 2.88	
episode: 1224	score: 2.0	memory length: 235718	epsilon: 0.7312763800058337
steps: 217	lr: 4e-05	evaluation reward: 2.88	
episode: 1225	score: 3.0	memory length: 235931	epsilon: 0.7308546400058429
steps: 213	lr: 4e-05	evaluation reward: 2.89	
episode: 1226	score: 0.0	memory length: 236054	epsilon: 0.7306111000058482
steps: 123	lr: 4e-05	evaluation reward: 2.87	
episode: 1227	score: 2.0	memory length: 236235	epsilon: 0.730252720005856
steps: 181	lr: 4e-05	evaluation reward: 2.86	
episode: 1228	score: 2.0	memory length: 236436	epsilon: 0.7298547400058646
steps: 201	lr: 4e-05	evaluation reward: 2.86	
episode: 1229	score: 4.0	memory length: 236694	epsilon: 0.7293439000058757
steps: 258	lr: 4e-05	evaluation reward: 2.89	
episode: 1230	score: 1.0	memory length: 236846	epsilon: 0.7290429400058822
steps: 152	lr: 4e-05	evaluation reward: 2.85	
episode: 1231	score: 3.0	memory length: 237074	epsilon: 0.728591500005892
steps: 228	lr: 4e-05	evaluation reward: 2.85	
episode: 1232	score: 2.0	memory length: 237255	epsilon: 0.7282331200058998
steps: 181	lr: 4e-05	evaluation reward: 2.83	
episode: 1233	score: 3.0	memory length: 237485	epsilon: 0.7277777200059097
steps: 230	lr: 4e-05	evaluation reward: 2.84	
episode: 1234	score: 4.0	memory length: 237764	epsilon: 0.7272253000059217
steps: 279	lr: 4e-05	evaluation reward: 2.86	
episode: 1235	score: 3.0	memory length: 237995	epsilon: 0.7267679200059316
steps: 231	lr: 4e-05	evaluation reward: 2.85	
episode: 1236	score: 3.0	memory length: 238268	epsilon: 0.7262273800059433
steps: 273	lr: 4e-05	evaluation reward: 2.87	
episode: 1237	score: 3.0	memory length: 238515	epsilon: 0.725738320005954
steps: 247	lr: 4e-05	evaluation reward: 2.87	
episode: 1238	score: 7.0	memory length: 238962	epsilon: 0.7248532600059732
steps: 447	lr: 4e-05	evaluation reward: 2.93	
episode: 1239	score: 3.0	memory length: 239192	epsilon: 0.724397860005983
steps: 230	lr: 4e-05	evaluation reward: 2.9	
episode: 1240	score: 3.0	memory length: 239423	epsilon: 0.723940480005993
steps: 231	lr: 4e-05	evaluation reward: 2.91	
episode: 1241	score: 3.0	memory length: 239653	epsilon: 0.7234850800060029
steps: 230	lr: 4e-05	evaluation reward: 2.93	
episode: 1242	score: 6.0	memory length: 240004	epsilon: 0.722790100006018
steps: 351	lr: 4e-05	evaluation reward: 2.96	
episode: 1243	score: 4.0	memory length: 240267	epsilon: 0.7222693600060293
steps: 263	lr: 4e-05	evaluation reward: 2.97	
episode: 1244	score: 2.0	memory length: 240466	epsilon: 0.7218753400060378
steps: 199	lr: 4e-05	evaluation reward: 2.95	
episode: 1245	score: 3.0	memory length: 240714	epsilon: 0.7213843000060485
steps: 248	lr: 4e-05	evaluation reward: 2.95	
episode: 1246	score: 6.0	memory length: 241079	epsilon: 0.7206616000060642

steps: 365	lr: 4e-05	evaluation reward: 2.98	
episode: 1247	score: 3.0	memory length: 241327	epsilon: 0.7201705600060748
steps: 248	lr: 4e-05	evaluation reward: 2.98	
episode: 1248	score: 2.0	memory length: 241544	epsilon: 0.7197409000060841
steps: 217	lr: 4e-05	evaluation reward: 2.96	
episode: 1249	score: 2.0	memory length: 241724	epsilon: 0.7193845000060919
steps: 180	lr: 4e-05	evaluation reward: 2.92	
episode: 1250	score: 3.0	memory length: 241950	epsilon: 0.7189370200061016
steps: 226	lr: 4e-05	evaluation reward: 2.93	
episode: 1251	score: 3.0	memory length: 242180	epsilon: 0.7184816200061115
steps: 230	lr: 4e-05	evaluation reward: 2.93	
episode: 1252	score: 3.0	memory length: 242411	epsilon: 0.7180242400061214
steps: 231	lr: 4e-05	evaluation reward: 2.93	
episode: 1253	score: 3.0	memory length: 242640	epsilon: 0.7175708200061313
steps: 229	lr: 4e-05	evaluation reward: 2.96	
episode: 1254	score: 1.0	memory length: 242791	epsilon: 0.7172718400061378
steps: 151	lr: 4e-05	evaluation reward: 2.94	
episode: 1255	score: 2.0	memory length: 242991	epsilon: 0.7168758400061463
steps: 200	lr: 4e-05	evaluation reward: 2.94	
episode: 1256	score: 4.0	memory length: 243266	epsilon: 0.7163313400061582
steps: 275	lr: 4e-05	evaluation reward: 2.94	
episode: 1257	score: 4.0	memory length: 243544	epsilon: 0.7157809000061701
steps: 278	lr: 4e-05	evaluation reward: 2.94	
episode: 1258	score: 0.0	memory length: 243668	epsilon: 0.7155353800061754
steps: 124	lr: 4e-05	evaluation reward: 2.91	
episode: 1259	score: 3.0	memory length: 243902	epsilon: 0.7150720600061855
steps: 234	lr: 4e-05	evaluation reward: 2.94	
episode: 1260	score: 4.0	memory length: 244192	epsilon: 0.714497860006198
steps: 290	lr: 4e-05	evaluation reward: 2.94	
episode: 1261	score: 9.0	memory length: 244663	epsilon: 0.7135652800062182
steps: 471	lr: 4e-05	evaluation reward: 3.01	
episode: 1262	score: 5.0	memory length: 244969	epsilon: 0.7129594000062314
steps: 306	lr: 4e-05	evaluation reward: 3.06	
episode: 1263	score: 2.0	memory length: 245150	epsilon: 0.7126010200062391
steps: 181	lr: 4e-05	evaluation reward: 3.07	
episode: 1264	score: 2.0	memory length: 245348	epsilon: 0.7122089800062477
steps: 198	lr: 4e-05	evaluation reward: 3.06	
episode: 1265	score: 3.0	memory length: 245596	epsilon: 0.7117179400062583
steps: 248	lr: 4e-05	evaluation reward: 3.07	
episode: 1266	score: 3.0	memory length: 245823	epsilon: 0.7112684800062681
steps: 227	lr: 4e-05	evaluation reward: 3.1	
episode: 1267	score: 3.0	memory length: 246052	epsilon: 0.7108150600062779
steps: 229	lr: 4e-05	evaluation reward: 3.09	
episode: 1268	score: 3.0	memory length: 246279	epsilon: 0.7103656000062877
steps: 227	lr: 4e-05	evaluation reward: 3.08	
episode: 1269	score: 2.0	memory length: 246461	epsilon: 0.7100052400062955
steps: 182	lr: 4e-05	evaluation reward: 3.06	
episode: 1270	score: 3.0	memory length: 246672	epsilon: 0.7095874600063046

steps: 211	lr: 4e-05	evaluation reward: 3.04	
episode: 1271	score: 3.0	memory length: 246901	epsilon: 0.7091340400063144
steps: 229	lr: 4e-05	evaluation reward: 3.05	
episode: 1272	score: 2.0	memory length: 247100	epsilon: 0.708740020006323
steps: 199	lr: 4e-05	evaluation reward: 3.01	
episode: 1273	score: 3.0	memory length: 247313	epsilon: 0.7083182800063321
steps: 213	lr: 4e-05	evaluation reward: 3.02	
episode: 1274	score: 2.0	memory length: 247512	epsilon: 0.7079242600063407
steps: 199	lr: 4e-05	evaluation reward: 3.02	
episode: 1275	score: 9.0	memory length: 247972	epsilon: 0.7070134600063605
steps: 460	lr: 4e-05	evaluation reward: 3.1	
episode: 1276	score: 3.0	memory length: 248183	epsilon: 0.7065956800063695
steps: 211	lr: 4e-05	evaluation reward: 3.1	
episode: 1277	score: 3.0	memory length: 248428	epsilon: 0.70611058000638
steps: 245	lr: 4e-05	evaluation reward: 3.12	
episode: 1278	score: 2.0	memory length: 248627	epsilon: 0.7057165600063886
steps: 199	lr: 4e-05	evaluation reward: 3.12	
episode: 1279	score: 2.0	memory length: 248826	epsilon: 0.7053225400063972
steps: 199	lr: 4e-05	evaluation reward: 3.12	
episode: 1280	score: 3.0	memory length: 249075	epsilon: 0.7048295200064079
steps: 249	lr: 4e-05	evaluation reward: 3.1	
episode: 1281	score: 7.0	memory length: 249433	epsilon: 0.7041206800064232
steps: 358	lr: 4e-05	evaluation reward: 3.16	
episode: 1282	score: 4.0	memory length: 249710	epsilon: 0.7035722200064352
steps: 277	lr: 4e-05	evaluation reward: 3.17	
episode: 1283	score: 4.0	memory length: 249970	epsilon: 0.7030574200064463
steps: 260	lr: 4e-05	evaluation reward: 3.18	
episode: 1284	score: 4.0	memory length: 250251	epsilon: 0.7025010400064584
steps: 281	lr: 4e-05	evaluation reward: 3.19	
episode: 1285	score: 3.0	memory length: 250520	epsilon: 0.70196842000647
steps: 269	lr: 4e-05	evaluation reward: 3.19	
episode: 1286	score: 4.0	memory length: 250780	epsilon: 0.7014536200064811
steps: 260	lr: 4e-05	evaluation reward: 3.21	
episode: 1287	score: 2.0	memory length: 251001	epsilon: 0.7010160400064906
steps: 221	lr: 4e-05	evaluation reward: 3.19	
episode: 1288	score: 2.0	memory length: 251200	epsilon: 0.7006220200064992
steps: 199	lr: 4e-05	evaluation reward: 3.12	
episode: 1289	score: 2.0	memory length: 251399	epsilon: 0.7002280000065078
steps: 199	lr: 4e-05	evaluation reward: 3.1	
episode: 1290	score: 3.0	memory length: 251626	epsilon: 0.6997785400065175
steps: 227	lr: 4e-05	evaluation reward: 3.11	
episode: 1291	score: 2.0	memory length: 251827	epsilon: 0.6993805600065262
steps: 201	lr: 4e-05	evaluation reward: 3.09	
episode: 1292	score: 3.0	memory length: 252056	epsilon: 0.698927140006536
steps: 229	lr: 4e-05	evaluation reward: 3.1	
episode: 1293	score: 3.0	memory length: 252286	epsilon: 0.6984717400065459
steps: 230	lr: 4e-05	evaluation reward: 3.11	
episode: 1294	score: 2.0	memory length: 252506	epsilon: 0.6980361400065553

```

steps: 220    lr: 4e-05    evaluation reward: 3.09
episode: 1295  score: 2.0    memory length: 252705    epsilon: 0.6976421200065639
steps: 199    lr: 4e-05    evaluation reward: 3.09
episode: 1296  score: 1.0    memory length: 252857    epsilon: 0.6973411600065704
steps: 152    lr: 4e-05    evaluation reward: 3.09
episode: 1297  score: 11.0   memory length: 253260    epsilon:
0.6965432200065877    steps: 403    lr: 4e-05    evaluation reward: 3.2
episode: 1298  score: 2.0    memory length: 253479    epsilon: 0.6961096000065972
steps: 219    lr: 4e-05    evaluation reward: 3.21
episode: 1299  score: 3.0    memory length: 253706    epsilon: 0.6956601400066069
steps: 227    lr: 4e-05    evaluation reward: 3.24
episode: 1300  score: 3.0    memory length: 253956    epsilon: 0.6951651400066177
steps: 250    lr: 4e-05    evaluation reward: 3.2
episode: 1301  score: 5.0    memory length: 254300    epsilon: 0.6944840200066325
steps: 344    lr: 4e-05    evaluation reward: 3.22
episode: 1302  score: 3.0    memory length: 254526    epsilon: 0.6940365400066422
steps: 226    lr: 4e-05    evaluation reward: 3.22
episode: 1303  score: 3.0    memory length: 254753    epsilon: 0.6935870800066519
steps: 227    lr: 4e-05    evaluation reward: 3.21
episode: 1304  score: 3.0    memory length: 255003    epsilon: 0.6930920800066627
steps: 250    lr: 4e-05    evaluation reward: 3.24
episode: 1305  score: 1.0    memory length: 255175    epsilon: 0.6927515200066701
steps: 172    lr: 4e-05    evaluation reward: 3.22
episode: 1306  score: 5.0    memory length: 255473    epsilon: 0.6921614800066829
steps: 298    lr: 4e-05    evaluation reward: 3.26
episode: 1307  score: 3.0    memory length: 255723    epsilon: 0.6916664800066936
steps: 250    lr: 4e-05    evaluation reward: 3.25
episode: 1308  score: 3.0    memory length: 255950    epsilon: 0.6912170200067034
steps: 227    lr: 4e-05    evaluation reward: 3.26
episode: 1309  score: 7.0    memory length: 256345    epsilon: 0.6904349200067204
steps: 395    lr: 4e-05    evaluation reward: 3.29
episode: 1310  score: 1.0    memory length: 256497    epsilon: 0.6901339600067269
steps: 152    lr: 4e-05    evaluation reward: 3.29
episode: 1311  score: 2.0    memory length: 256713    epsilon: 0.6897062800067362
steps: 216    lr: 4e-05    evaluation reward: 3.25
episode: 1312  score: 2.0    memory length: 256912    epsilon: 0.6893122600067447
steps: 199    lr: 4e-05    evaluation reward: 3.24
episode: 1313  score: 2.0    memory length: 257111    epsilon: 0.6889182400067533
steps: 199    lr: 4e-05    evaluation reward: 3.23
episode: 1314  score: 4.0    memory length: 257386    epsilon: 0.6883737400067651
steps: 275    lr: 4e-05    evaluation reward: 3.23
episode: 1315  score: 1.0    memory length: 257538    epsilon: 0.6880727800067716
steps: 152    lr: 4e-05    evaluation reward: 3.17
episode: 1316  score: 1.0    memory length: 257690    epsilon: 0.6877718200067782
steps: 152    lr: 4e-05    evaluation reward: 3.14
episode: 1317  score: 0.0    memory length: 257813    epsilon: 0.6875282800067835
steps: 123    lr: 4e-05    evaluation reward: 3.13
episode: 1318  score: 3.0    memory length: 258044    epsilon: 0.6870709000067934

```

steps: 231	lr: 4e-05	evaluation reward: 3.15	
episode: 1319	score: 4.0	memory length: 258322	epsilon: 0.6865204600068053
steps: 278	lr: 4e-05	evaluation reward: 3.17	
episode: 1320	score: 4.0	memory length: 258620	epsilon: 0.6859304200068181
steps: 298	lr: 4e-05	evaluation reward: 3.17	
episode: 1321	score: 3.0	memory length: 258850	epsilon: 0.685475020006828
steps: 230	lr: 4e-05	evaluation reward: 3.17	
episode: 1322	score: 3.0	memory length: 259077	epsilon: 0.6850255600068378
steps: 227	lr: 4e-05	evaluation reward: 3.12	
episode: 1323	score: 5.0	memory length: 259392	epsilon: 0.6844018600068513
steps: 315	lr: 4e-05	evaluation reward: 3.11	
episode: 1324	score: 4.0	memory length: 259689	epsilon: 0.6838138000068641
steps: 297	lr: 4e-05	evaluation reward: 3.13	
episode: 1325	score: 6.0	memory length: 260053	epsilon: 0.6830930800068797
steps: 364	lr: 4e-05	evaluation reward: 3.16	
episode: 1326	score: 6.0	memory length: 260411	epsilon: 0.6823842400068951
steps: 358	lr: 4e-05	evaluation reward: 3.22	
episode: 1327	score: 8.0	memory length: 260815	epsilon: 0.6815843200069125
steps: 404	lr: 4e-05	evaluation reward: 3.28	
episode: 1328	score: 3.0	memory length: 261042	epsilon: 0.6811348600069222
steps: 227	lr: 4e-05	evaluation reward: 3.29	
episode: 1329	score: 0.0	memory length: 261166	epsilon: 0.6808893400069276
steps: 124	lr: 4e-05	evaluation reward: 3.25	
episode: 1330	score: 3.0	memory length: 261393	epsilon: 0.6804398800069373
steps: 227	lr: 4e-05	evaluation reward: 3.27	
episode: 1331	score: 3.0	memory length: 261607	epsilon: 0.6800161600069465
steps: 214	lr: 4e-05	evaluation reward: 3.27	
episode: 1332	score: 1.0	memory length: 261758	epsilon: 0.679717180006953
steps: 151	lr: 4e-05	evaluation reward: 3.26	
episode: 1333	score: 5.0	memory length: 262086	epsilon: 0.6790677400069671
steps: 328	lr: 4e-05	evaluation reward: 3.28	
episode: 1334	score: 4.0	memory length: 262384	epsilon: 0.6784777000069799
steps: 298	lr: 4e-05	evaluation reward: 3.28	
episode: 1335	score: 3.0	memory length: 262613	epsilon: 0.6780242800069898
steps: 229	lr: 4e-05	evaluation reward: 3.28	
episode: 1336	score: 3.0	memory length: 262845	epsilon: 0.6775649200069997
steps: 232	lr: 4e-05	evaluation reward: 3.28	
episode: 1337	score: 1.0	memory length: 262997	epsilon: 0.6772639600070063
steps: 152	lr: 4e-05	evaluation reward: 3.26	
episode: 1338	score: 4.0	memory length: 263255	epsilon: 0.6767531200070174
steps: 258	lr: 4e-05	evaluation reward: 3.23	
episode: 1339	score: 2.0	memory length: 263472	epsilon: 0.6763234600070267
steps: 217	lr: 4e-05	evaluation reward: 3.22	
episode: 1340	score: 6.0	memory length: 263829	epsilon: 0.675616600007042
steps: 357	lr: 4e-05	evaluation reward: 3.25	
episode: 1341	score: 2.0	memory length: 264049	epsilon: 0.6751810000070515
steps: 220	lr: 4e-05	evaluation reward: 3.24	
episode: 1342	score: 4.0	memory length: 264302	epsilon: 0.6746800600070624

steps: 253	lr: 4e-05	evaluation reward: 3.22	
episode: 1343	score: 3.0	memory length: 264529	epsilon: 0.6742306000070721
steps: 227	lr: 4e-05	evaluation reward: 3.21	
episode: 1344	score: 1.0	memory length: 264680	epsilon: 0.6739316200070786
steps: 151	lr: 4e-05	evaluation reward: 3.2	
episode: 1345	score: 5.0	memory length: 264991	epsilon: 0.673315840007092
steps: 311	lr: 4e-05	evaluation reward: 3.22	
episode: 1346	score: 1.0	memory length: 265143	epsilon: 0.6730148800070985
steps: 152	lr: 4e-05	evaluation reward: 3.17	
episode: 1347	score: 6.0	memory length: 265464	epsilon: 0.6723793000071123
steps: 321	lr: 4e-05	evaluation reward: 3.2	
episode: 1348	score: 2.0	memory length: 265662	epsilon: 0.6719872600071208
steps: 198	lr: 4e-05	evaluation reward: 3.2	
episode: 1349	score: 0.0	memory length: 265786	epsilon: 0.6717417400071262
steps: 124	lr: 4e-05	evaluation reward: 3.18	
episode: 1350	score: 2.0	memory length: 265985	epsilon: 0.6713477200071347
steps: 199	lr: 4e-05	evaluation reward: 3.17	
episode: 1351	score: 4.0	memory length: 266230	epsilon: 0.6708626200071452
steps: 245	lr: 4e-05	evaluation reward: 3.18	
episode: 1352	score: 4.0	memory length: 266510	epsilon: 0.6703082200071573
steps: 280	lr: 4e-05	evaluation reward: 3.19	
episode: 1353	score: 6.0	memory length: 266845	epsilon: 0.6696449200071717
steps: 335	lr: 4e-05	evaluation reward: 3.22	
episode: 1354	score: 3.0	memory length: 267075	epsilon: 0.6691895200071816
steps: 230	lr: 4e-05	evaluation reward: 3.24	
episode: 1355	score: 2.0	memory length: 267274	epsilon: 0.6687955000071901
steps: 199	lr: 4e-05	evaluation reward: 3.24	
episode: 1356	score: 3.0	memory length: 267505	epsilon: 0.6683381200072
steps: 231	lr: 4e-05	evaluation reward: 3.23	
episode: 1357	score: 3.0	memory length: 267719	epsilon: 0.6679144000072093
steps: 214	lr: 4e-05	evaluation reward: 3.22	
episode: 1358	score: 2.0	memory length: 267917	epsilon: 0.6675223600072178
steps: 198	lr: 4e-05	evaluation reward: 3.24	
episode: 1359	score: 1.0	memory length: 268088	epsilon: 0.6671837800072251
steps: 171	lr: 4e-05	evaluation reward: 3.22	
episode: 1360	score: 4.0	memory length: 268366	epsilon: 0.6666333400072371
steps: 278	lr: 4e-05	evaluation reward: 3.22	
episode: 1361	score: 7.0	memory length: 268767	epsilon: 0.6658393600072543
steps: 401	lr: 4e-05	evaluation reward: 3.2	
episode: 1362	score: 5.0	memory length: 269113	epsilon: 0.6651542800072692
steps: 346	lr: 4e-05	evaluation reward: 3.2	
episode: 1363	score: 4.0	memory length: 269410	epsilon: 0.6645662200072819
steps: 297	lr: 4e-05	evaluation reward: 3.22	
episode: 1364	score: 2.0	memory length: 269609	epsilon: 0.6641722000072905
steps: 199	lr: 4e-05	evaluation reward: 3.22	
episode: 1365	score: 2.0	memory length: 269808	epsilon: 0.663778180007299
steps: 199	lr: 4e-05	evaluation reward: 3.21	
episode: 1366	score: 6.0	memory length: 270147	epsilon: 0.6631069600073136

steps: 339	lr: 4e-05	evaluation reward: 3.24	
episode: 1367	score: 2.0	memory length: 270329	epsilon: 0.6627466000073214
steps: 182	lr: 4e-05	evaluation reward: 3.23	
episode: 1368	score: 6.0	memory length: 270685	epsilon: 0.6620417200073367
steps: 356	lr: 4e-05	evaluation reward: 3.26	
episode: 1369	score: 8.0	memory length: 271111	epsilon: 0.661198240007355
steps: 426	lr: 4e-05	evaluation reward: 3.32	
episode: 1370	score: 4.0	memory length: 271388	epsilon: 0.660649780007367
steps: 277	lr: 4e-05	evaluation reward: 3.33	
episode: 1371	score: 2.0	memory length: 271568	epsilon: 0.6602933800073747
steps: 180	lr: 4e-05	evaluation reward: 3.32	
episode: 1372	score: 4.0	memory length: 271824	epsilon: 0.6597865000073857
steps: 256	lr: 4e-05	evaluation reward: 3.34	
episode: 1373	score: 4.0	memory length: 272102	epsilon: 0.6592360600073977
steps: 278	lr: 4e-05	evaluation reward: 3.35	
episode: 1374	score: 1.0	memory length: 272254	epsilon: 0.6589351000074042
steps: 152	lr: 4e-05	evaluation reward: 3.34	
episode: 1375	score: 5.0	memory length: 272583	epsilon: 0.6582836800074183
steps: 329	lr: 4e-05	evaluation reward: 3.3	
episode: 1376	score: 6.0	memory length: 272941	epsilon: 0.6575748400074337
steps: 358	lr: 4e-05	evaluation reward: 3.33	
episode: 1377	score: 4.0	memory length: 273222	epsilon: 0.6570184600074458
steps: 281	lr: 4e-05	evaluation reward: 3.34	
episode: 1378	score: 1.0	memory length: 273373	epsilon: 0.6567194800074523
steps: 151	lr: 4e-05	evaluation reward: 3.33	
episode: 1379	score: 2.0	memory length: 273554	epsilon: 0.6563611000074601
steps: 181	lr: 4e-05	evaluation reward: 3.33	
episode: 1380	score: 6.0	memory length: 273931	epsilon: 0.6556146400074763
steps: 377	lr: 4e-05	evaluation reward: 3.36	
episode: 1381	score: 3.0	memory length: 274158	epsilon: 0.655165180007486
steps: 227	lr: 4e-05	evaluation reward: 3.32	
episode: 1382	score: 1.0	memory length: 274331	epsilon: 0.6548226400074935
steps: 173	lr: 4e-05	evaluation reward: 3.29	
episode: 1383	score: 7.0	memory length: 274685	epsilon: 0.6541217200075087
steps: 354	lr: 4e-05	evaluation reward: 3.32	
episode: 1384	score: 6.0	memory length: 275060	epsilon: 0.6533792200075248
steps: 375	lr: 4e-05	evaluation reward: 3.34	
episode: 1385	score: 6.0	memory length: 275418	epsilon: 0.6526703800075402
steps: 358	lr: 4e-05	evaluation reward: 3.37	
episode: 1386	score: 1.0	memory length: 275588	epsilon: 0.6523337800075475
steps: 170	lr: 4e-05	evaluation reward: 3.34	
episode: 1387	score: 4.0	memory length: 275867	epsilon: 0.6517813600075595
steps: 279	lr: 4e-05	evaluation reward: 3.36	
episode: 1388	score: 1.0	memory length: 276020	epsilon: 0.6514784200075661
steps: 153	lr: 4e-05	evaluation reward: 3.35	
episode: 1389	score: 2.0	memory length: 276219	epsilon: 0.6510844000075746
steps: 199	lr: 4e-05	evaluation reward: 3.35	
episode: 1390	score: 4.0	memory length: 276495	epsilon: 0.6505379200075865

steps: 276	lr: 4e-05	evaluation reward: 3.36	
episode: 1391	score: 4.0	memory length: 276756	epsilon: 0.6500211400075977
steps: 261	lr: 4e-05	evaluation reward: 3.38	
episode: 1392	score: 5.0	memory length: 277084	epsilon: 0.6493717000076118
steps: 328	lr: 4e-05	evaluation reward: 3.4	
episode: 1393	score: 5.0	memory length: 277431	epsilon: 0.6486846400076267
steps: 347	lr: 4e-05	evaluation reward: 3.42	
episode: 1394	score: 3.0	memory length: 277701	epsilon: 0.6481500400076383
steps: 270	lr: 4e-05	evaluation reward: 3.43	
episode: 1395	score: 5.0	memory length: 278029	epsilon: 0.6475006000076524
steps: 328	lr: 4e-05	evaluation reward: 3.46	
episode: 1396	score: 2.0	memory length: 278231	epsilon: 0.6471006400076611
steps: 202	lr: 4e-05	evaluation reward: 3.47	
episode: 1397	score: 2.0	memory length: 278412	epsilon: 0.6467422600076689
steps: 181	lr: 4e-05	evaluation reward: 3.38	
episode: 1398	score: 3.0	memory length: 278644	epsilon: 0.6462829000076789
steps: 232	lr: 4e-05	evaluation reward: 3.39	
episode: 1399	score: 6.0	memory length: 279019	epsilon: 0.645540400007695
steps: 375	lr: 4e-05	evaluation reward: 3.42	
episode: 1400	score: 1.0	memory length: 279171	epsilon: 0.6452394400077015
steps: 152	lr: 4e-05	evaluation reward: 3.4	
episode: 1401	score: 3.0	memory length: 279402	epsilon: 0.6447820600077114
steps: 231	lr: 4e-05	evaluation reward: 3.38	
episode: 1402	score: 3.0	memory length: 279651	epsilon: 0.6442890400077221
steps: 249	lr: 4e-05	evaluation reward: 3.38	
episode: 1403	score: 6.0	memory length: 280020	epsilon: 0.643558420007738
steps: 369	lr: 4e-05	evaluation reward: 3.41	
episode: 1404	score: 9.0	memory length: 280438	epsilon: 0.642730780007756
steps: 418	lr: 4e-05	evaluation reward: 3.47	
episode: 1405	score: 3.0	memory length: 280707	epsilon: 0.6421981600077675
steps: 269	lr: 4e-05	evaluation reward: 3.49	
episode: 1406	score: 4.0	memory length: 280988	epsilon: 0.6416417800077796
steps: 281	lr: 4e-05	evaluation reward: 3.48	
episode: 1407	score: 4.0	memory length: 281268	epsilon: 0.6410873800077916
steps: 280	lr: 4e-05	evaluation reward: 3.49	
episode: 1408	score: 3.0	memory length: 281515	epsilon: 0.6405983200078023
steps: 247	lr: 4e-05	evaluation reward: 3.49	
episode: 1409	score: 3.0	memory length: 281741	epsilon: 0.640150840007812
steps: 226	lr: 4e-05	evaluation reward: 3.45	
episode: 1410	score: 1.0	memory length: 281911	epsilon: 0.6398142400078193
steps: 170	lr: 4e-05	evaluation reward: 3.45	
episode: 1411	score: 3.0	memory length: 282140	epsilon: 0.6393608200078291
steps: 229	lr: 4e-05	evaluation reward: 3.46	
episode: 1412	score: 6.0	memory length: 282515	epsilon: 0.6386183200078452
steps: 375	lr: 4e-05	evaluation reward: 3.5	
episode: 1413	score: 2.0	memory length: 282714	epsilon: 0.6382243000078538
steps: 199	lr: 4e-05	evaluation reward: 3.5	
episode: 1414	score: 3.0	memory length: 282961	epsilon: 0.6377352400078644

steps: 247	lr: 4e-05	evaluation reward: 3.49	
episode: 1415	score: 6.0	memory length: 283319	epsilon: 0.6370264000078798
steps: 358	lr: 4e-05	evaluation reward: 3.54	
episode: 1416	score: 5.0	memory length: 283648	epsilon: 0.6363749800078939
steps: 329	lr: 4e-05	evaluation reward: 3.58	
episode: 1417	score: 3.0	memory length: 283879	epsilon: 0.6359176000079039
steps: 231	lr: 4e-05	evaluation reward: 3.61	
episode: 1418	score: 3.0	memory length: 284092	epsilon: 0.635495860007913
steps: 213	lr: 4e-05	evaluation reward: 3.61	
episode: 1419	score: 6.0	memory length: 284419	epsilon: 0.6348484000079271
steps: 327	lr: 4e-05	evaluation reward: 3.63	
episode: 1420	score: 6.0	memory length: 284724	epsilon: 0.6342445000079402
steps: 305	lr: 4e-05	evaluation reward: 3.65	
episode: 1421	score: 6.0	memory length: 285066	epsilon: 0.6335673400079549
steps: 342	lr: 4e-05	evaluation reward: 3.68	
episode: 1422	score: 3.0	memory length: 285312	epsilon: 0.6330802600079655
steps: 246	lr: 4e-05	evaluation reward: 3.68	
episode: 1423	score: 3.0	memory length: 285561	epsilon: 0.6325872400079762
steps: 249	lr: 4e-05	evaluation reward: 3.66	
episode: 1424	score: 6.0	memory length: 285960	epsilon: 0.6317972200079933
steps: 399	lr: 4e-05	evaluation reward: 3.68	
episode: 1425	score: 0.0	memory length: 286084	epsilon: 0.6315517000079987
steps: 124	lr: 4e-05	evaluation reward: 3.62	
episode: 1426	score: 4.0	memory length: 286358	epsilon: 0.6310091800080104
steps: 274	lr: 4e-05	evaluation reward: 3.6	
episode: 1427	score: 1.0	memory length: 286530	epsilon: 0.6306686200080178
steps: 172	lr: 4e-05	evaluation reward: 3.53	
episode: 1428	score: 4.0	memory length: 286808	epsilon: 0.6301181800080298
steps: 278	lr: 4e-05	evaluation reward: 3.54	
episode: 1429	score: 5.0	memory length: 287105	epsilon: 0.6295301200080425
steps: 297	lr: 4e-05	evaluation reward: 3.59	
episode: 1430	score: 6.0	memory length: 287466	epsilon: 0.628815340008058
steps: 361	lr: 4e-05	evaluation reward: 3.62	
episode: 1431	score: 5.0	memory length: 287775	epsilon: 0.6282035200080713
steps: 309	lr: 4e-05	evaluation reward: 3.64	
episode: 1432	score: 4.0	memory length: 288035	epsilon: 0.6276887200080825
steps: 260	lr: 4e-05	evaluation reward: 3.67	
episode: 1433	score: 2.0	memory length: 288234	epsilon: 0.6272947000080911
steps: 199	lr: 4e-05	evaluation reward: 3.64	
episode: 1434	score: 3.0	memory length: 288465	epsilon: 0.626837320008101
steps: 231	lr: 4e-05	evaluation reward: 3.63	
episode: 1435	score: 5.0	memory length: 288793	epsilon: 0.6261878800081151
steps: 328	lr: 4e-05	evaluation reward: 3.65	
episode: 1436	score: 1.0	memory length: 288945	epsilon: 0.6258869200081216
steps: 152	lr: 4e-05	evaluation reward: 3.63	
episode: 1437	score: 3.0	memory length: 289172	epsilon: 0.6254374600081314
steps: 227	lr: 4e-05	evaluation reward: 3.65	
episode: 1438	score: 2.0	memory length: 289355	epsilon: 0.6250751200081393

steps: 183	lr: 4e-05	evaluation reward: 3.63	
episode: 1439	score: 4.0	memory length: 289616	epsilon: 0.6245583400081505
steps: 261	lr: 4e-05	evaluation reward: 3.65	
episode: 1440	score: 5.0	memory length: 289924	epsilon: 0.6239485000081637
steps: 308	lr: 4e-05	evaluation reward: 3.64	
episode: 1441	score: 6.0	memory length: 290297	epsilon: 0.6232099600081797
steps: 373	lr: 4e-05	evaluation reward: 3.68	
episode: 1442	score: 4.0	memory length: 290591	epsilon: 0.6226278400081924
steps: 294	lr: 4e-05	evaluation reward: 3.68	
episode: 1443	score: 5.0	memory length: 290919	epsilon: 0.6219784000082065
steps: 328	lr: 4e-05	evaluation reward: 3.7	
episode: 1444	score: 7.0	memory length: 291302	epsilon: 0.6212200600082229
steps: 383	lr: 4e-05	evaluation reward: 3.76	
episode: 1445	score: 1.0	memory length: 291453	epsilon: 0.6209210800082294
steps: 151	lr: 4e-05	evaluation reward: 3.72	
episode: 1446	score: 4.0	memory length: 291729	epsilon: 0.6203746000082413
steps: 276	lr: 4e-05	evaluation reward: 3.75	
episode: 1447	score: 4.0	memory length: 291986	epsilon: 0.6198657400082523
steps: 257	lr: 4e-05	evaluation reward: 3.73	
episode: 1448	score: 4.0	memory length: 292261	epsilon: 0.6193212400082642
steps: 275	lr: 4e-05	evaluation reward: 3.75	
episode: 1449	score: 1.0	memory length: 292413	epsilon: 0.6190202800082707
steps: 152	lr: 4e-05	evaluation reward: 3.76	
episode: 1450	score: 10.0	memory length: 292951	epsilon: 0.6179550400082938
steps: 538	lr: 4e-05	evaluation reward: 3.84	
episode: 1451	score: 2.0	memory length: 293152	epsilon: 0.6175570600083025
steps: 201	lr: 4e-05	evaluation reward: 3.82	
episode: 1452	score: 2.0	memory length: 293371	epsilon: 0.6171234400083119
steps: 219	lr: 4e-05	evaluation reward: 3.8	
episode: 1453	score: 4.0	memory length: 293649	epsilon: 0.6165730000083238
steps: 278	lr: 4e-05	evaluation reward: 3.78	
episode: 1454	score: 3.0	memory length: 293879	epsilon: 0.6161176000083337
steps: 230	lr: 4e-05	evaluation reward: 3.78	
episode: 1455	score: 3.0	memory length: 294105	epsilon: 0.6156701200083434
steps: 226	lr: 4e-05	evaluation reward: 3.79	
episode: 1456	score: 5.0	memory length: 294396	epsilon: 0.6150939400083559
steps: 291	lr: 4e-05	evaluation reward: 3.81	
episode: 1457	score: 2.0	memory length: 294578	epsilon: 0.6147335800083638
steps: 182	lr: 4e-05	evaluation reward: 3.8	
episode: 1458	score: 4.0	memory length: 294859	epsilon: 0.6141772000083758
steps: 281	lr: 4e-05	evaluation reward: 3.82	
episode: 1459	score: 4.0	memory length: 295138	epsilon: 0.6136247800083878
steps: 279	lr: 4e-05	evaluation reward: 3.85	
episode: 1460	score: 5.0	memory length: 295450	epsilon: 0.6130070200084012
steps: 312	lr: 4e-05	evaluation reward: 3.86	
episode: 1461	score: 4.0	memory length: 295728	epsilon: 0.6124565800084132
steps: 278	lr: 4e-05	evaluation reward: 3.83	
episode: 1462	score: 6.0	memory length: 296122	epsilon: 0.6116764600084301

steps: 394	lr: 4e-05	evaluation reward: 3.84	
episode: 1463	score: 3.0	memory length: 296336	epsilon: 0.6112527400084393
steps: 214	lr: 4e-05	evaluation reward: 3.83	
episode: 1464	score: 2.0	memory length: 296518	epsilon: 0.6108923800084471
steps: 182	lr: 4e-05	evaluation reward: 3.83	
episode: 1465	score: 4.0	memory length: 296840	epsilon: 0.610254820008461
steps: 322	lr: 4e-05	evaluation reward: 3.85	
episode: 1466	score: 3.0	memory length: 297069	epsilon: 0.6098014000084708
steps: 229	lr: 4e-05	evaluation reward: 3.82	
episode: 1467	score: 1.0	memory length: 297221	epsilon: 0.6095004400084774
steps: 152	lr: 4e-05	evaluation reward: 3.81	
episode: 1468	score: 4.0	memory length: 297507	epsilon: 0.6089341600084897
steps: 286	lr: 4e-05	evaluation reward: 3.79	
episode: 1469	score: 1.0	memory length: 297679	epsilon: 0.608593600008497
steps: 172	lr: 4e-05	evaluation reward: 3.72	
episode: 1470	score: 4.0	memory length: 297955	epsilon: 0.6080471200085089
steps: 276	lr: 4e-05	evaluation reward: 3.72	
episode: 1471	score: 6.0	memory length: 298309	epsilon: 0.6073462000085241
steps: 354	lr: 4e-05	evaluation reward: 3.76	
episode: 1472	score: 5.0	memory length: 298655	epsilon: 0.606661120008539
steps: 346	lr: 4e-05	evaluation reward: 3.77	
episode: 1473	score: 3.0	memory length: 298902	epsilon: 0.6061720600085496
steps: 247	lr: 4e-05	evaluation reward: 3.76	
episode: 1474	score: 6.0	memory length: 299257	epsilon: 0.6054691600085649
steps: 355	lr: 4e-05	evaluation reward: 3.81	
episode: 1475	score: 7.0	memory length: 299630	epsilon: 0.6047306200085809
steps: 373	lr: 4e-05	evaluation reward: 3.83	
episode: 1476	score: 6.0	memory length: 299995	epsilon: 0.6040079200085966
steps: 365	lr: 4e-05	evaluation reward: 3.83	
episode: 1477	score: 1.0	memory length: 300166	epsilon: 0.603669340008604
steps: 171	lr: 1.6000000000000003e-05	evaluation reward: 3.8	
episode: 1478	score: 1.0	memory length: 300317	epsilon: 0.6033703600086104
steps: 151	lr: 1.6000000000000003e-05	evaluation reward: 3.8	
episode: 1479	score: 2.0	memory length: 300499	epsilon: 0.6030100000086183
steps: 182	lr: 1.6000000000000003e-05	evaluation reward: 3.8	
episode: 1480	score: 1.0	memory length: 300670	epsilon: 0.6026714200086256
steps: 171	lr: 1.6000000000000003e-05	evaluation reward: 3.75	
episode: 1481	score: 4.0	memory length: 300946	epsilon: 0.6021249400086375
steps: 276	lr: 1.6000000000000003e-05	evaluation reward: 3.76	
episode: 1482	score: 3.0	memory length: 301160	epsilon: 0.6017012200086467
steps: 214	lr: 1.6000000000000003e-05	evaluation reward: 3.78	
episode: 1483	score: 1.0	memory length: 301312	epsilon: 0.6014002600086532
steps: 152	lr: 1.6000000000000003e-05	evaluation reward: 3.72	
episode: 1484	score: 4.0	memory length: 301611	epsilon: 0.6008082400086661
steps: 299	lr: 1.6000000000000003e-05	evaluation reward: 3.7	
episode: 1485	score: 3.0	memory length: 301863	epsilon: 0.6003092800086769
steps: 252	lr: 1.6000000000000003e-05	evaluation reward: 3.67	
episode: 1486	score: 6.0	memory length: 302224	epsilon: 0.5995945000086924

```

steps: 361    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1487  score: 4.0    memory length: 302521    epsilon: 0.5990064400087052
steps: 297    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1488  score: 6.0    memory length: 302847    epsilon: 0.5983609600087192
steps: 326    lr: 1.6000000000000003e-05    evaluation reward: 3.77
episode: 1489  score: 2.0    memory length: 303045    epsilon: 0.5979689200087277
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.77
episode: 1490  score: 2.0    memory length: 303264    epsilon: 0.5975353000087371
steps: 219    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1491  score: 5.0    memory length: 303571    epsilon: 0.5969274400087503
steps: 307    lr: 1.6000000000000003e-05    evaluation reward: 3.76
episode: 1492  score: 2.0    memory length: 303772    epsilon: 0.596529460008759
steps: 201    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1493  score: 4.0    memory length: 304065    epsilon: 0.5959493200087715
steps: 293    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1494  score: 4.0    memory length: 304345    epsilon: 0.5953949200087836
steps: 280    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1495  score: 5.0    memory length: 304673    epsilon: 0.5947454800087977
steps: 328    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1496  score: 9.0    memory length: 305112    epsilon: 0.5938762600088165
steps: 439    lr: 1.6000000000000003e-05    evaluation reward: 3.8
episode: 1497  score: 3.0    memory length: 305383    epsilon: 0.5933396800088282
steps: 271    lr: 1.6000000000000003e-05    evaluation reward: 3.81
episode: 1498  score: 4.0    memory length: 305679    epsilon: 0.5927536000088409
steps: 296    lr: 1.6000000000000003e-05    evaluation reward: 3.82
episode: 1499  score: 4.0    memory length: 305924    epsilon: 0.5922685000088515
steps: 245    lr: 1.6000000000000003e-05    evaluation reward: 3.8
episode: 1500  score: 7.0    memory length: 306300    epsilon: 0.5915240200088676
steps: 376    lr: 1.6000000000000003e-05    evaluation reward: 3.86
episode: 1501  score: 4.0    memory length: 306580    epsilon: 0.5909696200088796
steps: 280    lr: 1.6000000000000003e-05    evaluation reward: 3.87
episode: 1502  score: 7.0    memory length: 307003    epsilon: 0.5901320800088978
steps: 423    lr: 1.6000000000000003e-05    evaluation reward: 3.91
episode: 1503  score: 6.0    memory length: 307357    epsilon: 0.589431160008913
steps: 354    lr: 1.6000000000000003e-05    evaluation reward: 3.91
episode: 1504  score: 4.0    memory length: 307639    epsilon: 0.5888728000089252
steps: 282    lr: 1.6000000000000003e-05    evaluation reward: 3.86
episode: 1505  score: 1.0    memory length: 307791    epsilon: 0.5885718400089317
steps: 152    lr: 1.6000000000000003e-05    evaluation reward: 3.84
episode: 1506  score: 7.0    memory length: 308222    epsilon: 0.5877184600089502
steps: 431    lr: 1.6000000000000003e-05    evaluation reward: 3.87
episode: 1507  score: 5.0    memory length: 308514    epsilon: 0.5871403000089628
steps: 292    lr: 1.6000000000000003e-05    evaluation reward: 3.88
episode: 1508  score: 3.0    memory length: 308729    epsilon: 0.586714600008972
steps: 215    lr: 1.6000000000000003e-05    evaluation reward: 3.88
episode: 1509  score: 7.0    memory length: 309106    epsilon: 0.5859681400089882
steps: 377    lr: 1.6000000000000003e-05    evaluation reward: 3.92
episode: 1510  score: 3.0    memory length: 309355    epsilon: 0.5854751200089989

```

```

steps: 249    lr: 1.6000000000000003e-05    evaluation reward: 3.94
episode: 1511  score: 9.0    memory length: 309817    epsilon: 0.5845603600090188
steps: 462    lr: 1.6000000000000003e-05    evaluation reward: 4.0
episode: 1512  score: 7.0    memory length: 310229    epsilon: 0.5837446000090365
steps: 412    lr: 1.6000000000000003e-05    evaluation reward: 4.01
episode: 1513  score: 6.0    memory length: 310624    epsilon: 0.5829625000090535
steps: 395    lr: 1.6000000000000003e-05    evaluation reward: 4.05
episode: 1514  score: 11.0   memory length: 311069    epsilon:
0.5820814000090726    steps: 445    lr: 1.6000000000000003e-05    evaluation
reward: 4.13
episode: 1515  score: 3.0    memory length: 311280    epsilon: 0.5816636200090817
steps: 211    lr: 1.6000000000000003e-05    evaluation reward: 4.1
episode: 1516  score: 7.0    memory length: 311636    epsilon: 0.580958740009097
steps: 356    lr: 1.6000000000000003e-05    evaluation reward: 4.12
episode: 1517  score: 5.0    memory length: 311964    epsilon: 0.5803093000091111
steps: 328    lr: 1.6000000000000003e-05    evaluation reward: 4.14
episode: 1518  score: 7.0    memory length: 312338    epsilon: 0.5795687800091271
steps: 374    lr: 1.6000000000000003e-05    evaluation reward: 4.18
episode: 1519  score: 12.0   memory length: 312806    epsilon:
0.5786421400091473    steps: 468    lr: 1.6000000000000003e-05    evaluation
reward: 4.24
episode: 1520  score: 7.0    memory length: 313234    epsilon: 0.5777947000091657
steps: 428    lr: 1.6000000000000003e-05    evaluation reward: 4.25
episode: 1521  score: 2.0    memory length: 313436    epsilon: 0.5773947400091743
steps: 202    lr: 1.6000000000000003e-05    evaluation reward: 4.21
episode: 1522  score: 5.0    memory length: 313742    epsilon: 0.5767888600091875
steps: 306    lr: 1.6000000000000003e-05    evaluation reward: 4.23
episode: 1523  score: 4.0    memory length: 314019    epsilon: 0.5762404000091994
steps: 277    lr: 1.6000000000000003e-05    evaluation reward: 4.24
episode: 1524  score: 10.0   memory length: 314541    epsilon:
0.5752068400092218    steps: 522    lr: 1.6000000000000003e-05    evaluation
reward: 4.28
episode: 1525  score: 3.0    memory length: 314774    epsilon: 0.5747455000092319
steps: 233    lr: 1.6000000000000003e-05    evaluation reward: 4.31
episode: 1526  score: 5.0    memory length: 315084    epsilon: 0.5741317000092452
steps: 310    lr: 1.6000000000000003e-05    evaluation reward: 4.32
episode: 1527  score: 4.0    memory length: 315346    epsilon: 0.5736129400092564
steps: 262    lr: 1.6000000000000003e-05    evaluation reward: 4.35
episode: 1528  score: 7.0    memory length: 315757    epsilon: 0.5727991600092741
steps: 411    lr: 1.6000000000000003e-05    evaluation reward: 4.38
episode: 1529  score: 5.0    memory length: 316052    epsilon: 0.5722150600092868
steps: 295    lr: 1.6000000000000003e-05    evaluation reward: 4.38
episode: 1530  score: 3.0    memory length: 316318    epsilon: 0.5716883800092982
steps: 266    lr: 1.6000000000000003e-05    evaluation reward: 4.35
episode: 1531  score: 8.0    memory length: 316709    epsilon: 0.570914200009315
steps: 391    lr: 1.6000000000000003e-05    evaluation reward: 4.38
episode: 1532  score: 9.0    memory length: 317192    epsilon: 0.5699578600093358
steps: 483    lr: 1.6000000000000003e-05    evaluation reward: 4.43

```

episode: 1533 score: 3.0 memory length: 317423 epsilon: 0.5695004800093457
 steps: 231 lr: 1.6000000000000003e-05 evaluation reward: 4.44
 episode: 1534 score: 5.0 memory length: 317752 epsilon: 0.5688490600093599
 steps: 329 lr: 1.6000000000000003e-05 evaluation reward: 4.46
 episode: 1535 score: 2.0 memory length: 317935 epsilon: 0.5684867200093677
 steps: 183 lr: 1.6000000000000003e-05 evaluation reward: 4.43
 episode: 1536 score: 11.0 memory length: 318486 epsilon:
 0.5673957400093914 steps: 551 lr: 1.6000000000000003e-05 evaluation
 reward: 4.53
 episode: 1537 score: 5.0 memory length: 318784 epsilon: 0.5668057000094042
 steps: 298 lr: 1.6000000000000003e-05 evaluation reward: 4.55
 episode: 1538 score: 5.0 memory length: 319093 epsilon: 0.5661938800094175
 steps: 309 lr: 1.6000000000000003e-05 evaluation reward: 4.58
 episode: 1539 score: 6.0 memory length: 319432 epsilon: 0.5655226600094321
 steps: 339 lr: 1.6000000000000003e-05 evaluation reward: 4.6
 episode: 1540 score: 6.0 memory length: 319786 epsilon: 0.5648217400094473
 steps: 354 lr: 1.6000000000000003e-05 evaluation reward: 4.61
 episode: 1541 score: 1.0 memory length: 319956 epsilon: 0.5644851400094546
 steps: 170 lr: 1.6000000000000003e-05 evaluation reward: 4.56
 episode: 1542 score: 4.0 memory length: 320254 epsilon: 0.5638951000094674
 steps: 298 lr: 1.6000000000000003e-05 evaluation reward: 4.56
 episode: 1543 score: 4.0 memory length: 320517 epsilon: 0.5633743600094787
 steps: 263 lr: 1.6000000000000003e-05 evaluation reward: 4.55
 episode: 1544 score: 2.0 memory length: 320718 epsilon: 0.5629763800094874
 steps: 201 lr: 1.6000000000000003e-05 evaluation reward: 4.5
 episode: 1545 score: 4.0 memory length: 320994 epsilon: 0.5624299000094992
 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 4.53
 episode: 1546 score: 5.0 memory length: 321304 epsilon: 0.5618161000095125
 steps: 310 lr: 1.6000000000000003e-05 evaluation reward: 4.54
 episode: 1547 score: 3.0 memory length: 321550 epsilon: 0.5613290200095231
 steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 4.53
 episode: 1548 score: 8.0 memory length: 321960 epsilon: 0.5605172200095407
 steps: 410 lr: 1.6000000000000003e-05 evaluation reward: 4.57
 episode: 1549 score: 7.0 memory length: 322379 epsilon: 0.5596876000095588
 steps: 419 lr: 1.6000000000000003e-05 evaluation reward: 4.63
 episode: 1550 score: 7.0 memory length: 322729 epsilon: 0.5589946000095738
 steps: 350 lr: 1.6000000000000003e-05 evaluation reward: 4.6
 episode: 1551 score: 10.0 memory length: 323240 epsilon:
 0.5579828200095958 steps: 511 lr: 1.6000000000000003e-05 evaluation
 reward: 4.68
 episode: 1552 score: 7.0 memory length: 323577 epsilon: 0.5573155600096102
 steps: 337 lr: 1.6000000000000003e-05 evaluation reward: 4.73
 episode: 1553 score: 4.0 memory length: 323819 epsilon: 0.5568364000096206
 steps: 242 lr: 1.6000000000000003e-05 evaluation reward: 4.73
 episode: 1554 score: 6.0 memory length: 324178 epsilon: 0.5561255800096361
 steps: 359 lr: 1.6000000000000003e-05 evaluation reward: 4.76
 episode: 1555 score: 7.0 memory length: 324564 epsilon: 0.5553613000096527
 steps: 386 lr: 1.6000000000000003e-05 evaluation reward: 4.8

episode: 1556 score: 3.0 memory length: 324775 epsilon: 0.5549435200096617
 steps: 211 lr: 1.6000000000000003e-05 evaluation reward: 4.78
 episode: 1557 score: 6.0 memory length: 325143 epsilon: 0.5542148800096776
 steps: 368 lr: 1.6000000000000003e-05 evaluation reward: 4.82
 episode: 1558 score: 4.0 memory length: 325422 epsilon: 0.5536624600096895
 steps: 279 lr: 1.6000000000000003e-05 evaluation reward: 4.82
 episode: 1559 score: 5.0 memory length: 325752 epsilon: 0.5530090600097037
 steps: 330 lr: 1.6000000000000003e-05 evaluation reward: 4.83
 episode: 1560 score: 3.0 memory length: 325980 epsilon: 0.5525576200097135
 steps: 228 lr: 1.6000000000000003e-05 evaluation reward: 4.81
 episode: 1561 score: 3.0 memory length: 326209 epsilon: 0.5521042000097234
 steps: 229 lr: 1.6000000000000003e-05 evaluation reward: 4.8
 episode: 1562 score: 2.0 memory length: 326390 epsilon: 0.5517458200097312
 steps: 181 lr: 1.6000000000000003e-05 evaluation reward: 4.76
 episode: 1563 score: 4.0 memory length: 326652 epsilon: 0.5512270600097424
 steps: 262 lr: 1.6000000000000003e-05 evaluation reward: 4.77
 episode: 1564 score: 5.0 memory length: 326960 epsilon: 0.5506172200097557
 steps: 308 lr: 1.6000000000000003e-05 evaluation reward: 4.8
 episode: 1565 score: 7.0 memory length: 327380 epsilon: 0.5497856200097737
 steps: 420 lr: 1.6000000000000003e-05 evaluation reward: 4.83
 episode: 1566 score: 4.0 memory length: 327660 epsilon: 0.5492312200097857
 steps: 280 lr: 1.6000000000000003e-05 evaluation reward: 4.84
 episode: 1567 score: 8.0 memory length: 328100 epsilon: 0.5483600200098047
 steps: 440 lr: 1.6000000000000003e-05 evaluation reward: 4.91
 episode: 1568 score: 9.0 memory length: 328425 epsilon: 0.5477165200098186
 steps: 325 lr: 1.6000000000000003e-05 evaluation reward: 4.96
 episode: 1569 score: 5.0 memory length: 328790 epsilon: 0.5469938200098343
 steps: 365 lr: 1.6000000000000003e-05 evaluation reward: 5.0
 episode: 1570 score: 8.0 memory length: 329225 epsilon: 0.546132520009853
 steps: 435 lr: 1.6000000000000003e-05 evaluation reward: 5.04
 episode: 1571 score: 8.0 memory length: 329699 epsilon: 0.545194000098734
 steps: 474 lr: 1.6000000000000003e-05 evaluation reward: 5.06
 episode: 1572 score: 8.0 memory length: 330104 epsilon: 0.5443921000098908
 steps: 405 lr: 1.6000000000000003e-05 evaluation reward: 5.09
 episode: 1573 score: 6.0 memory length: 330443 epsilon: 0.5437208800099054
 steps: 339 lr: 1.6000000000000003e-05 evaluation reward: 5.12
 episode: 1574 score: 5.0 memory length: 330774 epsilon: 0.5430655000099196
 steps: 331 lr: 1.6000000000000003e-05 evaluation reward: 5.11
 episode: 1575 score: 11.0 memory length: 331296 epsilon: 0.542031940009942
 steps: 522 lr: 1.6000000000000003e-05 evaluation reward: 5.15
 episode: 1576 score: 3.0 memory length: 331510 epsilon: 0.5416082200099512
 steps: 214 lr: 1.6000000000000003e-05 evaluation reward: 5.12
 episode: 1577 score: 7.0 memory length: 331870 epsilon: 0.5408954200099667
 steps: 360 lr: 1.6000000000000003e-05 evaluation reward: 5.18
 episode: 1578 score: 5.0 memory length: 332160 epsilon: 0.5403212200099792
 steps: 290 lr: 1.6000000000000003e-05 evaluation reward: 5.22
 episode: 1579 score: 3.0 memory length: 332370 epsilon: 0.5399054200099882
 steps: 210 lr: 1.6000000000000003e-05 evaluation reward: 5.23

episode: 1580 score: 3.0 memory length: 332618 epsilon: 0.5394143800099989
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 5.25
 episode: 1581 score: 3.0 memory length: 332848 epsilon: 0.5389589800100087
 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 5.24
 episode: 1582 score: 8.0 memory length: 333266 epsilon: 0.5381313400100267
 steps: 418 lr: 1.6000000000000003e-05 evaluation reward: 5.29
 episode: 1583 score: 3.0 memory length: 333513 epsilon: 0.5376422800100373
 steps: 247 lr: 1.6000000000000003e-05 evaluation reward: 5.31
 episode: 1584 score: 6.0 memory length: 333867 epsilon: 0.5369413600100525
 steps: 354 lr: 1.6000000000000003e-05 evaluation reward: 5.33
 episode: 1585 score: 7.0 memory length: 334280 epsilon: 0.5361236200100703
 steps: 413 lr: 1.6000000000000003e-05 evaluation reward: 5.37
 episode: 1586 score: 5.0 memory length: 334590 epsilon: 0.5355098200100836
 steps: 310 lr: 1.6000000000000003e-05 evaluation reward: 5.36
 episode: 1587 score: 6.0 memory length: 334929 epsilon: 0.5348386000100982
 steps: 339 lr: 1.6000000000000003e-05 evaluation reward: 5.38
 episode: 1588 score: 2.0 memory length: 335130 epsilon: 0.5344406200101068
 steps: 201 lr: 1.6000000000000003e-05 evaluation reward: 5.34
 episode: 1589 score: 5.0 memory length: 335437 epsilon: 0.53383276001012
 steps: 307 lr: 1.6000000000000003e-05 evaluation reward: 5.37
 episode: 1590 score: 5.0 memory length: 335733 epsilon: 0.5332466800101328
 steps: 296 lr: 1.6000000000000003e-05 evaluation reward: 5.4
 episode: 1591 score: 3.0 memory length: 335960 epsilon: 0.5327972200101425
 steps: 227 lr: 1.6000000000000003e-05 evaluation reward: 5.38
 episode: 1592 score: 3.0 memory length: 336171 epsilon: 0.5323794400101516
 steps: 211 lr: 1.6000000000000003e-05 evaluation reward: 5.39
 episode: 1593 score: 3.0 memory length: 336419 epsilon: 0.5318884000101622
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 5.38
 episode: 1594 score: 4.0 memory length: 336680 epsilon: 0.5313716200101735
 steps: 261 lr: 1.6000000000000003e-05 evaluation reward: 5.38
 episode: 1595 score: 3.0 memory length: 336928 epsilon: 0.5308805800101841
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 5.36
 episode: 1596 score: 2.0 memory length: 337127 epsilon: 0.5304865600101927
 steps: 199 lr: 1.6000000000000003e-05 evaluation reward: 5.29
 episode: 1597 score: 6.0 memory length: 337482 epsilon: 0.5297836600102079
 steps: 355 lr: 1.6000000000000003e-05 evaluation reward: 5.32
 episode: 1598 score: 5.0 memory length: 337808 epsilon: 0.529138180010222
 steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 5.33
 episode: 1599 score: 4.0 memory length: 338067 epsilon: 0.5286253600102331
 steps: 259 lr: 1.6000000000000003e-05 evaluation reward: 5.33
 episode: 1600 score: 7.0 memory length: 338515 epsilon: 0.5277383200102523
 steps: 448 lr: 1.6000000000000003e-05 evaluation reward: 5.33
 episode: 1601 score: 10.0 memory length: 338903 epsilon: 0.526970080010269
 steps: 388 lr: 1.6000000000000003e-05 evaluation reward: 5.39
 episode: 1602 score: 3.0 memory length: 339134 epsilon: 0.526512700010279
 steps: 231 lr: 1.6000000000000003e-05 evaluation reward: 5.35
 episode: 1603 score: 2.0 memory length: 339317 epsilon: 0.5261503600102868
 steps: 183 lr: 1.6000000000000003e-05 evaluation reward: 5.31

episode: 1604 score: 4.0 memory length: 339580 epsilon: 0.5256296200102981
 steps: 263 lr: 1.6000000000000003e-05 evaluation reward: 5.31
 episode: 1605 score: 2.0 memory length: 339779 epsilon: 0.5252356000103067
 steps: 199 lr: 1.6000000000000003e-05 evaluation reward: 5.32
 episode: 1606 score: 5.0 memory length: 340086 epsilon: 0.5246277400103199
 steps: 307 lr: 1.6000000000000003e-05 evaluation reward: 5.3
 episode: 1607 score: 3.0 memory length: 340333 epsilon: 0.5241386800103305
 steps: 247 lr: 1.6000000000000003e-05 evaluation reward: 5.28
 episode: 1608 score: 6.0 memory length: 340677 epsilon: 0.5234575600103453
 steps: 344 lr: 1.6000000000000003e-05 evaluation reward: 5.31
 episode: 1609 score: 5.0 memory length: 341011 epsilon: 0.5227962400103596
 steps: 334 lr: 1.6000000000000003e-05 evaluation reward: 5.29
 episode: 1610 score: 6.0 memory length: 341405 epsilon: 0.5220161200103766
 steps: 394 lr: 1.6000000000000003e-05 evaluation reward: 5.32
 episode: 1611 score: 5.0 memory length: 341747 epsilon: 0.5213389600103913
 steps: 342 lr: 1.6000000000000003e-05 evaluation reward: 5.28
 episode: 1612 score: 2.0 memory length: 341946 epsilon: 0.5209449400103998
 steps: 199 lr: 1.6000000000000003e-05 evaluation reward: 5.23
 episode: 1613 score: 3.0 memory length: 342177 epsilon: 0.5204875600104097
 steps: 231 lr: 1.6000000000000003e-05 evaluation reward: 5.2
 episode: 1614 score: 5.0 memory length: 342486 epsilon: 0.519875740010423
 steps: 309 lr: 1.6000000000000003e-05 evaluation reward: 5.14
 episode: 1615 score: 3.0 memory length: 342714 epsilon: 0.5194243000104328
 steps: 228 lr: 1.6000000000000003e-05 evaluation reward: 5.14
 episode: 1616 score: 3.0 memory length: 342943 epsilon: 0.5189708800104427
 steps: 229 lr: 1.6000000000000003e-05 evaluation reward: 5.1
 episode: 1617 score: 2.0 memory length: 343162 epsilon: 0.5185372600104521
 steps: 219 lr: 1.6000000000000003e-05 evaluation reward: 5.07
 episode: 1618 score: 9.0 memory length: 343649 epsilon: 0.517573000010473
 steps: 487 lr: 1.6000000000000003e-05 evaluation reward: 5.09
 episode: 1619 score: 7.0 memory length: 344022 epsilon: 0.516834460010489
 steps: 373 lr: 1.6000000000000003e-05 evaluation reward: 5.04
 episode: 1620 score: 6.0 memory length: 344379 epsilon: 0.5161276000105044
 steps: 357 lr: 1.6000000000000003e-05 evaluation reward: 5.03
 episode: 1621 score: 11.0 memory length: 344733 epsilon:
 0.5154266800105196 steps: 354 lr: 1.6000000000000003e-05 evaluation
 reward: 5.12
 episode: 1622 score: 4.0 memory length: 345010 epsilon: 0.5148782200105315
 steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 5.11
 episode: 1623 score: 5.0 memory length: 345319 epsilon: 0.5142664000105448
 steps: 309 lr: 1.6000000000000003e-05 evaluation reward: 5.12
 episode: 1624 score: 5.0 memory length: 345664 epsilon: 0.5135833000105596
 steps: 345 lr: 1.6000000000000003e-05 evaluation reward: 5.07
 episode: 1625 score: 6.0 memory length: 346056 epsilon: 0.5128071400105765
 steps: 392 lr: 1.6000000000000003e-05 evaluation reward: 5.1
 episode: 1626 score: 4.0 memory length: 346348 epsilon: 0.512228980010589
 steps: 292 lr: 1.6000000000000003e-05 evaluation reward: 5.09
 episode: 1627 score: 6.0 memory length: 346704 epsilon: 0.5115241000106043

```

steps: 356    lr: 1.6000000000000003e-05    evaluation reward: 5.11
episode: 1628  score: 8.0    memory length: 347096    epsilon: 0.5107479400106212
steps: 392    lr: 1.6000000000000003e-05    evaluation reward: 5.12
episode: 1629  score: 2.0    memory length: 347277    epsilon: 0.510389560010629
steps: 181    lr: 1.6000000000000003e-05    evaluation reward: 5.09
episode: 1630  score: 5.0    memory length: 347586    epsilon: 0.5097777400106422
steps: 309    lr: 1.6000000000000003e-05    evaluation reward: 5.11
episode: 1631  score: 5.0    memory length: 347912    epsilon: 0.5091322600106563
steps: 326    lr: 1.6000000000000003e-05    evaluation reward: 5.08
episode: 1632  score: 4.0    memory length: 348188    epsilon: 0.5085857800106681
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 5.03
episode: 1633  score: 5.0    memory length: 348513    epsilon: 0.5079422800106821
steps: 325    lr: 1.6000000000000003e-05    evaluation reward: 5.05
episode: 1634  score: 5.0    memory length: 348825    epsilon: 0.5073245200106955
steps: 312    lr: 1.6000000000000003e-05    evaluation reward: 5.05
episode: 1635  score: 5.0    memory length: 349139    epsilon: 0.506702800010709
steps: 314    lr: 1.6000000000000003e-05    evaluation reward: 5.08
episode: 1636  score: 9.0    memory length: 349606    epsilon: 0.5057781400107291
steps: 467    lr: 1.6000000000000003e-05    evaluation reward: 5.06
episode: 1637  score: 7.0    memory length: 350011    epsilon: 0.5049762400107465
steps: 405    lr: 1.6000000000000003e-05    evaluation reward: 5.08
episode: 1638  score: 14.0   memory length: 350400    epsilon:
0.5042060200107632    steps: 389    lr: 1.6000000000000003e-05    evaluation
reward: 5.17
episode: 1639  score: 6.0    memory length: 350739    epsilon: 0.5035348000107778
steps: 339    lr: 1.6000000000000003e-05    evaluation reward: 5.17
episode: 1640  score: 2.0    memory length: 350958    epsilon: 0.5031011800107872
steps: 219    lr: 1.6000000000000003e-05    evaluation reward: 5.13
episode: 1641  score: 6.0    memory length: 351298    epsilon: 0.5024279800108018
steps: 340    lr: 1.6000000000000003e-05    evaluation reward: 5.18
episode: 1642  score: 4.0    memory length: 351594    epsilon: 0.5018419000108145
steps: 296    lr: 1.6000000000000003e-05    evaluation reward: 5.18
episode: 1643  score: 2.0    memory length: 351793    epsilon: 0.5014478800108231
steps: 199    lr: 1.6000000000000003e-05    evaluation reward: 5.16
episode: 1644  score: 7.0    memory length: 352183    epsilon: 0.5006756800108398
steps: 390    lr: 1.6000000000000003e-05    evaluation reward: 5.21
episode: 1645  score: 5.0    memory length: 352514    epsilon: 0.5000203000108541
steps: 331    lr: 1.6000000000000003e-05    evaluation reward: 5.22
episode: 1646  score: 0.0    memory length: 352638    epsilon:
0.49977478001085307    steps: 124    lr: 1.6000000000000003e-05    evaluation
reward: 5.17
episode: 1647  score: 7.0    memory length: 353042    epsilon: 0.498974860010848
steps: 404    lr: 1.6000000000000003e-05    evaluation reward: 5.21
episode: 1648  score: 4.0    memory length: 353340    epsilon: 0.4983848200108443
steps: 298    lr: 1.6000000000000003e-05    evaluation reward: 5.17
episode: 1649  score: 6.0    memory length: 353700    epsilon:
0.49767202001083977    steps: 360    lr: 1.6000000000000003e-05    evaluation
reward: 5.16

```

episode: 1650 score: 7.0 memory length: 354064 epsilon: 0.4969513000108352
 steps: 364 lr: 1.6000000000000003e-05 evaluation reward: 5.16
 episode: 1651 score: 5.0 memory length: 354388 epsilon:
 0.49630978001083115 steps: 324 lr: 1.6000000000000003e-05 evaluation
 reward: 5.11
 episode: 1652 score: 9.0 memory length: 354880 epsilon: 0.495335620010825
 steps: 492 lr: 1.6000000000000003e-05 evaluation reward: 5.13
 episode: 1653 score: 3.0 memory length: 355092 epsilon:
 0.49491586001082233 steps: 212 lr: 1.6000000000000003e-05 evaluation
 reward: 5.12
 episode: 1654 score: 3.0 memory length: 355319 epsilon: 0.4944664000108195
 steps: 227 lr: 1.6000000000000003e-05 evaluation reward: 5.09
 episode: 1655 score: 4.0 memory length: 355615 epsilon: 0.4938803200108158
 steps: 296 lr: 1.6000000000000003e-05 evaluation reward: 5.06
 episode: 1656 score: 6.0 memory length: 355992 epsilon:
 0.49313386001081105 steps: 377 lr: 1.6000000000000003e-05 evaluation
 reward: 5.09
 episode: 1657 score: 5.0 memory length: 356316 epsilon: 0.492492340010807
 steps: 324 lr: 1.6000000000000003e-05 evaluation reward: 5.08
 episode: 1658 score: 3.0 memory length: 356530 epsilon: 0.4920686200108043
 steps: 214 lr: 1.6000000000000003e-05 evaluation reward: 5.07
 episode: 1659 score: 3.0 memory length: 356763 epsilon: 0.4916072800108014
 steps: 233 lr: 1.6000000000000003e-05 evaluation reward: 5.05
 episode: 1660 score: 6.0 memory length: 357137 epsilon: 0.4908667600107967
 steps: 374 lr: 1.6000000000000003e-05 evaluation reward: 5.08
 episode: 1661 score: 4.0 memory length: 357413 epsilon:
 0.49032028001079325 steps: 276 lr: 1.6000000000000003e-05 evaluation
 reward: 5.09
 episode: 1662 score: 7.0 memory length: 357789 epsilon:
 0.48957580001078854 steps: 376 lr: 1.6000000000000003e-05 evaluation
 reward: 5.14
 episode: 1663 score: 4.0 memory length: 358085 epsilon:
 0.48898972001078483 steps: 296 lr: 1.6000000000000003e-05 evaluation
 reward: 5.14
 episode: 1664 score: 7.0 memory length: 358511 epsilon: 0.4881462400107795
 steps: 426 lr: 1.6000000000000003e-05 evaluation reward: 5.16
 episode: 1665 score: 6.0 memory length: 358830 epsilon: 0.4875146200107755
 steps: 319 lr: 1.6000000000000003e-05 evaluation reward: 5.15
 episode: 1666 score: 3.0 memory length: 359061 epsilon: 0.4870572400107726
 steps: 231 lr: 1.6000000000000003e-05 evaluation reward: 5.14
 episode: 1667 score: 5.0 memory length: 359351 epsilon: 0.486483040010769
 steps: 290 lr: 1.6000000000000003e-05 evaluation reward: 5.11
 episode: 1668 score: 7.0 memory length: 359732 epsilon: 0.4857286600107642
 steps: 381 lr: 1.6000000000000003e-05 evaluation reward: 5.09
 episode: 1669 score: 5.0 memory length: 360042 epsilon: 0.4851148600107603
 steps: 310 lr: 1.6000000000000003e-05 evaluation reward: 5.09
 episode: 1670 score: 9.0 memory length: 360468 epsilon: 0.484271380010755
 steps: 426 lr: 1.6000000000000003e-05 evaluation reward: 5.1

episode: 1671 score: 7.0 memory length: 360854 epsilon:
 0.48350710001075015 steps: 386 lr: 1.6000000000000003e-05 evaluation
 reward: 5.09
 episode: 1672 score: 3.0 memory length: 361066 epsilon: 0.4830873400107475
 steps: 212 lr: 1.6000000000000003e-05 evaluation reward: 5.04
 episode: 1673 score: 3.0 memory length: 361280 epsilon: 0.4826636200107448
 steps: 214 lr: 1.6000000000000003e-05 evaluation reward: 5.01
 episode: 1674 score: 6.0 memory length: 361636 epsilon:
 0.48195874001074035 steps: 356 lr: 1.6000000000000003e-05 evaluation
 reward: 5.02
 episode: 1675 score: 2.0 memory length: 361834 epsilon:
 0.48156670001073787 steps: 198 lr: 1.6000000000000003e-05 evaluation
 reward: 4.93
 episode: 1676 score: 6.0 memory length: 362214 epsilon: 0.4808143000107331
 steps: 380 lr: 1.6000000000000003e-05 evaluation reward: 4.96
 episode: 1677 score: 5.0 memory length: 362541 epsilon: 0.480166840010729
 steps: 327 lr: 1.6000000000000003e-05 evaluation reward: 4.94
 episode: 1678 score: 8.0 memory length: 362939 epsilon: 0.479378800010724
 steps: 398 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1679 score: 3.0 memory length: 363166 epsilon: 0.4789293400107212
 steps: 227 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1680 score: 5.0 memory length: 363492 epsilon: 0.4782838600107171
 steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 4.99
 episode: 1681 score: 3.0 memory length: 363719 epsilon:
 0.47783440001071426 steps: 227 lr: 1.6000000000000003e-05 evaluation
 reward: 4.99
 episode: 1682 score: 6.0 memory length: 364075 epsilon: 0.4771295200107098
 steps: 356 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1683 score: 7.0 memory length: 364473 epsilon: 0.4763414800107048
 steps: 398 lr: 1.6000000000000003e-05 evaluation reward: 5.01
 episode: 1684 score: 4.0 memory length: 364755 epsilon: 0.4757831200107013
 steps: 282 lr: 1.6000000000000003e-05 evaluation reward: 4.99
 episode: 1685 score: 1.0 memory length: 364907 epsilon: 0.4754821600106994
 steps: 152 lr: 1.6000000000000003e-05 evaluation reward: 4.93
 episode: 1686 score: 3.0 memory length: 365134 epsilon:
 0.47503270001069653 steps: 227 lr: 1.6000000000000003e-05 evaluation
 reward: 4.91
 episode: 1687 score: 2.0 memory length: 365335 epsilon: 0.474634720010694
 steps: 201 lr: 1.6000000000000003e-05 evaluation reward: 4.87
 episode: 1688 score: 3.0 memory length: 365549 epsilon:
 0.47421100001069133 steps: 214 lr: 1.6000000000000003e-05 evaluation
 reward: 4.88
 episode: 1689 score: 2.0 memory length: 365748 epsilon:
 0.47381698001068884 steps: 199 lr: 1.6000000000000003e-05 evaluation
 reward: 4.85
 episode: 1690 score: 8.0 memory length: 366154 epsilon:
 0.47301310001068375 steps: 406 lr: 1.6000000000000003e-05 evaluation
 reward: 4.88

episode: 1691 score: 5.0 memory length: 366502 epsilon: 0.4723240600106794
 steps: 348 lr: 1.6000000000000003e-05 evaluation reward: 4.9
 episode: 1692 score: 4.0 memory length: 366742 epsilon: 0.4718488600106764
 steps: 240 lr: 1.6000000000000003e-05 evaluation reward: 4.91
 episode: 1693 score: 3.0 memory length: 366991 epsilon:
 0.47135584001067327 steps: 249 lr: 1.6000000000000003e-05 evaluation
 reward: 4.91
 episode: 1694 score: 6.0 memory length: 367346 epsilon: 0.4706529400106688
 steps: 355 lr: 1.6000000000000003e-05 evaluation reward: 4.93
 episode: 1695 score: 19.0 memory length: 367935 epsilon:
 0.46948672001066144 steps: 589 lr: 1.6000000000000003e-05 evaluation
 reward: 5.09
 episode: 1696 score: 6.0 memory length: 368254 epsilon:
 0.46885510001065744 steps: 319 lr: 1.6000000000000003e-05 evaluation
 reward: 5.13
 episode: 1697 score: 5.0 memory length: 368584 epsilon: 0.4682017000106533
 steps: 330 lr: 1.6000000000000003e-05 evaluation reward: 5.12
 episode: 1698 score: 5.0 memory length: 368930 epsilon: 0.467516620010649
 steps: 346 lr: 1.6000000000000003e-05 evaluation reward: 5.12
 episode: 1699 score: 7.0 memory length: 369265 epsilon: 0.4668533200106448
 steps: 335 lr: 1.6000000000000003e-05 evaluation reward: 5.15
 episode: 1700 score: 6.0 memory length: 369620 epsilon:
 0.46615042001064033 steps: 355 lr: 1.6000000000000003e-05 evaluation
 reward: 5.14
 episode: 1701 score: 5.0 memory length: 369945 epsilon:
 0.46550692001063626 steps: 325 lr: 1.6000000000000003e-05 evaluation
 reward: 5.09
 episode: 1702 score: 3.0 memory length: 370176 epsilon:
 0.46504954001063337 steps: 231 lr: 1.6000000000000003e-05 evaluation
 reward: 5.09
 episode: 1703 score: 5.0 memory length: 370503 epsilon:
 0.46440208001062927 steps: 327 lr: 1.6000000000000003e-05 evaluation
 reward: 5.12
 episode: 1704 score: 5.0 memory length: 370812 epsilon: 0.4637902600106254
 steps: 309 lr: 1.6000000000000003e-05 evaluation reward: 5.13
 episode: 1705 score: 7.0 memory length: 371201 epsilon: 0.4630200400106205
 steps: 389 lr: 1.6000000000000003e-05 evaluation reward: 5.18
 episode: 1706 score: 8.0 memory length: 371655 epsilon:
 0.46212112001061484 steps: 454 lr: 1.6000000000000003e-05 evaluation
 reward: 5.21
 episode: 1707 score: 6.0 memory length: 371982 epsilon:
 0.46147366001061074 steps: 327 lr: 1.6000000000000003e-05 evaluation
 reward: 5.24
 episode: 1708 score: 3.0 memory length: 372209 epsilon: 0.4610242000106079
 steps: 227 lr: 1.6000000000000003e-05 evaluation reward: 5.21
 episode: 1709 score: 6.0 memory length: 372584 epsilon: 0.4602817000106032
 steps: 375 lr: 1.6000000000000003e-05 evaluation reward: 5.22
 episode: 1710 score: 6.0 memory length: 372889 epsilon: 0.4596778000105994

```

steps: 305    lr: 1.6000000000000003e-05    evaluation reward: 5.22
episode: 1711  score: 4.0  memory length: 373150  epsilon: 0.4591610200105961
steps: 261    lr: 1.6000000000000003e-05    evaluation reward: 5.21
episode: 1712  score: 3.0  memory length: 373361  epsilon:
0.45874324001059347  steps: 211    lr: 1.6000000000000003e-05    evaluation
reward: 5.22
episode: 1713  score: 10.0  memory length: 373867  epsilon:
0.45774136001058713  steps: 506    lr: 1.6000000000000003e-05    evaluation
reward: 5.29
episode: 1714  score: 7.0  memory length: 374257  epsilon:
0.45696916001058224  steps: 390    lr: 1.6000000000000003e-05    evaluation
reward: 5.31
episode: 1715  score: 8.0  memory length: 374668  epsilon: 0.4561553800105771
steps: 411    lr: 1.6000000000000003e-05    evaluation reward: 5.36
episode: 1716  score: 6.0  memory length: 375022  epsilon:
0.45545446001057266  steps: 354    lr: 1.6000000000000003e-05    evaluation
reward: 5.39
episode: 1717  score: 5.0  memory length: 375330  epsilon: 0.4548446200105688
steps: 308    lr: 1.6000000000000003e-05    evaluation reward: 5.42
episode: 1718  score: 4.0  memory length: 375591  epsilon:
0.45432784001056553  steps: 261    lr: 1.6000000000000003e-05    evaluation
reward: 5.37
episode: 1719  score: 6.0  memory length: 375988  epsilon:
0.45354178001056056  steps: 397    lr: 1.6000000000000003e-05    evaluation
reward: 5.36
episode: 1720  score: 9.0  memory length: 376324  epsilon:
0.45287650001055635  steps: 336    lr: 1.6000000000000003e-05    evaluation
reward: 5.39
episode: 1721  score: 8.0  memory length: 376754  epsilon:
0.45202510001055096  steps: 430    lr: 1.6000000000000003e-05    evaluation
reward: 5.36
episode: 1722  score: 3.0  memory length: 376966  epsilon: 0.4516053400105483
steps: 212    lr: 1.6000000000000003e-05    evaluation reward: 5.35
episode: 1723  score: 5.0  memory length: 377293  epsilon: 0.4509578800105442
steps: 327    lr: 1.6000000000000003e-05    evaluation reward: 5.35
episode: 1724  score: 6.0  memory length: 377650  epsilon:
0.45025102001053974  steps: 357    lr: 1.6000000000000003e-05    evaluation
reward: 5.36
episode: 1725  score: 5.0  memory length: 377958  epsilon: 0.4496411800105359
steps: 308    lr: 1.6000000000000003e-05    evaluation reward: 5.35
episode: 1726  score: 7.0  memory length: 378346  epsilon: 0.448872940010531
steps: 388    lr: 1.6000000000000003e-05    evaluation reward: 5.38
episode: 1727  score: 8.0  memory length: 378762  epsilon: 0.4480492600105258
steps: 416    lr: 1.6000000000000003e-05    evaluation reward: 5.4
episode: 1728  score: 1.0  memory length: 378913  epsilon: 0.4477502800105239
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 5.33
episode: 1729  score: 5.0  memory length: 379243  epsilon: 0.4470968800105198
steps: 330    lr: 1.6000000000000003e-05    evaluation reward: 5.36

```

episode: 1730 score: 11.0 memory length: 379780 epsilon:
 0.44603362001051305 steps: 537 lr: 1.6000000000000003e-05 evaluation
 reward: 5.42
 episode: 1731 score: 5.0 memory length: 380096 epsilon: 0.4454079400105091
 steps: 316 lr: 1.6000000000000003e-05 evaluation reward: 5.42
 episode: 1732 score: 11.0 memory length: 380609 epsilon:
 0.44439220001050267 steps: 513 lr: 1.6000000000000003e-05 evaluation
 reward: 5.49
 episode: 1733 score: 8.0 memory length: 381033 epsilon:
 0.44355268001049736 steps: 424 lr: 1.6000000000000003e-05 evaluation
 reward: 5.52
 episode: 1734 score: 5.0 memory length: 381321 epsilon:
 0.44298244001049375 steps: 288 lr: 1.6000000000000003e-05 evaluation
 reward: 5.52
 episode: 1735 score: 8.0 memory length: 381764 epsilon: 0.4421053000104882
 steps: 443 lr: 1.6000000000000003e-05 evaluation reward: 5.55
 episode: 1736 score: 4.0 memory length: 382024 epsilon:
 0.44159050001048494 steps: 260 lr: 1.6000000000000003e-05 evaluation
 reward: 5.5
 episode: 1737 score: 3.0 memory length: 382271 epsilon:
 0.44110144001048185 steps: 247 lr: 1.6000000000000003e-05 evaluation
 reward: 5.46
 episode: 1738 score: 3.0 memory length: 382499 epsilon: 0.440650000010479
 steps: 228 lr: 1.6000000000000003e-05 evaluation reward: 5.35
 episode: 1739 score: 13.0 memory length: 383029 epsilon:
 0.43960060001047235 steps: 530 lr: 1.6000000000000003e-05 evaluation
 reward: 5.42
 episode: 1740 score: 7.0 memory length: 383436 epsilon:
 0.43879474001046725 steps: 407 lr: 1.6000000000000003e-05 evaluation
 reward: 5.47
 episode: 1741 score: 8.0 memory length: 383894 epsilon: 0.4378879000104615
 steps: 458 lr: 1.6000000000000003e-05 evaluation reward: 5.49
 episode: 1742 score: 6.0 memory length: 384238 epsilon: 0.4372067800104572
 steps: 344 lr: 1.6000000000000003e-05 evaluation reward: 5.51
 episode: 1743 score: 5.0 memory length: 384578 epsilon:
 0.43653358001045295 steps: 340 lr: 1.6000000000000003e-05 evaluation
 reward: 5.54
 episode: 1744 score: 6.0 memory length: 384920 epsilon:
 0.43585642001044866 steps: 342 lr: 1.6000000000000003e-05 evaluation
 reward: 5.53
 episode: 1745 score: 9.0 memory length: 385373 epsilon: 0.434959480010443
 steps: 453 lr: 1.6000000000000003e-05 evaluation reward: 5.57
 episode: 1746 score: 6.0 memory length: 385749 epsilon: 0.4342150000104383
 steps: 376 lr: 1.6000000000000003e-05 evaluation reward: 5.63
 episode: 1747 score: 5.0 memory length: 386056 epsilon:
 0.43360714001043443 steps: 307 lr: 1.6000000000000003e-05 evaluation
 reward: 5.61
 episode: 1748 score: 5.0 memory length: 386381 epsilon:

0.43296364001043036 steps: 325 lr: 1.6000000000000003e-05 evaluation
 reward: 5.62
 episode: 1749 score: 9.0 memory length: 386856 epsilon: 0.4320231400104244
 steps: 475 lr: 1.6000000000000003e-05 evaluation reward: 5.65
 episode: 1750 score: 9.0 memory length: 387360 epsilon: 0.4310252200104181
 steps: 504 lr: 1.6000000000000003e-05 evaluation reward: 5.67
 episode: 1751 score: 10.0 memory length: 387845 epsilon: 0.430064920010412
 steps: 485 lr: 1.6000000000000003e-05 evaluation reward: 5.72
 episode: 1752 score: 6.0 memory length: 388171 epsilon:
 0.42941944001040794 steps: 326 lr: 1.6000000000000003e-05 evaluation
 reward: 5.69
 episode: 1753 score: 8.0 memory length: 388614 epsilon: 0.4285423000104024
 steps: 443 lr: 1.6000000000000003e-05 evaluation reward: 5.74
 episode: 1754 score: 8.0 memory length: 389065 epsilon:
 0.42764932001039674 steps: 451 lr: 1.6000000000000003e-05 evaluation
 reward: 5.79
 episode: 1755 score: 11.0 memory length: 389602 epsilon: 0.42658606001039
 steps: 537 lr: 1.6000000000000003e-05 evaluation reward: 5.86
 episode: 1756 score: 7.0 memory length: 389950 epsilon:
 0.42589702001038565 steps: 348 lr: 1.6000000000000003e-05 evaluation
 reward: 5.87
 episode: 1757 score: 6.0 memory length: 390327 epsilon:
 0.42515056001038093 steps: 377 lr: 1.6000000000000003e-05 evaluation
 reward: 5.88
 episode: 1758 score: 6.0 memory length: 390703 epsilon: 0.4244060800103762
 steps: 376 lr: 1.6000000000000003e-05 evaluation reward: 5.91
 episode: 1759 score: 8.0 memory length: 391107 epsilon:
 0.42360616001037116 steps: 404 lr: 1.6000000000000003e-05 evaluation
 reward: 5.96
 episode: 1760 score: 4.0 memory length: 391384 epsilon: 0.4230577000103677
 steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 5.94
 episode: 1761 score: 6.0 memory length: 391741 epsilon: 0.4223508400103632
 steps: 357 lr: 1.6000000000000003e-05 evaluation reward: 5.96
 episode: 1762 score: 5.0 memory length: 392068 epsilon: 0.4217033800103591
 steps: 327 lr: 1.6000000000000003e-05 evaluation reward: 5.94
 episode: 1763 score: 7.0 memory length: 392435 epsilon: 0.4209767200103545
 steps: 367 lr: 1.6000000000000003e-05 evaluation reward: 5.97
 episode: 1764 score: 8.0 memory length: 392862 epsilon: 0.4201312600103492
 steps: 427 lr: 1.6000000000000003e-05 evaluation reward: 5.98
 episode: 1765 score: 9.0 memory length: 393315 epsilon: 0.4192343200103435
 steps: 453 lr: 1.6000000000000003e-05 evaluation reward: 6.01
 episode: 1766 score: 6.0 memory length: 393641 epsilon: 0.4185888400103394
 steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 6.04
 episode: 1767 score: 8.0 memory length: 394110 epsilon:
 0.41766022001033354 steps: 469 lr: 1.6000000000000003e-05 evaluation
 reward: 6.07
 episode: 1768 score: 7.0 memory length: 394502 epsilon: 0.4168840600103286
 steps: 392 lr: 1.6000000000000003e-05 evaluation reward: 6.07

episode: 1769 score: 7.0 memory length: 394892 epsilon:
 0.41611186001032374 steps: 390 lr: 1.6000000000000003e-05 evaluation
 reward: 6.09
 episode: 1770 score: 4.0 memory length: 395152 epsilon: 0.4155970600103205
 steps: 260 lr: 1.6000000000000003e-05 evaluation reward: 6.04
 episode: 1771 score: 4.0 memory length: 395451 epsilon:
 0.41500504001031674 steps: 299 lr: 1.6000000000000003e-05 evaluation
 reward: 6.01
 episode: 1772 score: 5.0 memory length: 395761 epsilon:
 0.41439124001031286 steps: 310 lr: 1.6000000000000003e-05 evaluation
 reward: 6.03
 episode: 1773 score: 5.0 memory length: 396073 epsilon:
 0.41377348001030895 steps: 312 lr: 1.6000000000000003e-05 evaluation
 reward: 6.05
 episode: 1774 score: 7.0 memory length: 396440 epsilon:
 0.41304682001030435 steps: 367 lr: 1.6000000000000003e-05 evaluation
 reward: 6.06
 episode: 1775 score: 8.0 memory length: 396878 epsilon:
 0.41217958001029886 steps: 438 lr: 1.6000000000000003e-05 evaluation
 reward: 6.12
 episode: 1776 score: 7.0 memory length: 397246 epsilon:
 0.41145094001029425 steps: 368 lr: 1.6000000000000003e-05 evaluation
 reward: 6.13
 episode: 1777 score: 3.0 memory length: 397480 epsilon: 0.4109876200102913
 steps: 234 lr: 1.6000000000000003e-05 evaluation reward: 6.11
 episode: 1778 score: 8.0 memory length: 397934 epsilon:
 0.41008870001028563 steps: 454 lr: 1.6000000000000003e-05 evaluation
 reward: 6.11
 episode: 1779 score: 6.0 memory length: 398309 epsilon:
 0.40934620001028094 steps: 375 lr: 1.6000000000000003e-05 evaluation
 reward: 6.14
 episode: 1780 score: 7.0 memory length: 398673 epsilon: 0.4086254800102764
 steps: 364 lr: 1.6000000000000003e-05 evaluation reward: 6.16
 episode: 1781 score: 7.0 memory length: 399077 epsilon: 0.4078255600102713
 steps: 404 lr: 1.6000000000000003e-05 evaluation reward: 6.2
 episode: 1782 score: 4.0 memory length: 399336 epsilon:
 0.40731274001026807 steps: 259 lr: 1.6000000000000003e-05 evaluation
 reward: 6.18
 episode: 1783 score: 5.0 memory length: 399663 epsilon: 0.406665280010264
 steps: 327 lr: 1.6000000000000003e-05 evaluation reward: 6.16
 episode: 1784 score: 9.0 memory length: 400186 epsilon: 0.4056297400102574
 steps: 523 lr: 6.4000000000000001e-06 evaluation reward: 6.21
 episode: 1785 score: 8.0 memory length: 400638 epsilon:
 0.40473478001025176 steps: 452 lr: 6.4000000000000001e-06 evaluation
 reward: 6.28
 episode: 1786 score: 10.0 memory length: 401117 epsilon:
 0.40378636001024576 steps: 479 lr: 6.4000000000000001e-06 evaluation
 reward: 6.35

episode: 1787 score: 5.0 memory length: 401407 epsilon: 0.4032121600102421
 steps: 290 lr: 6.4000000000000001e-06 evaluation reward: 6.38
 episode: 1788 score: 9.0 memory length: 401870 epsilon: 0.4022954200102363
 steps: 463 lr: 6.4000000000000001e-06 evaluation reward: 6.44
 episode: 1789 score: 7.0 memory length: 402295 epsilon: 0.401453920010231
 steps: 425 lr: 6.4000000000000001e-06 evaluation reward: 6.49
 episode: 1790 score: 5.0 memory length: 402640 epsilon: 0.4007708200102267
 steps: 345 lr: 6.4000000000000001e-06 evaluation reward: 6.46
 episode: 1791 score: 8.0 memory length: 403080 epsilon:
 0.39989962001022117 steps: 440 lr: 6.4000000000000001e-06 evaluation
 reward: 6.49
 episode: 1792 score: 8.0 memory length: 403500 epsilon: 0.3990680200102159
 steps: 420 lr: 6.4000000000000001e-06 evaluation reward: 6.53
 episode: 1793 score: 8.0 memory length: 403924 epsilon: 0.3982285000102106
 steps: 424 lr: 6.4000000000000001e-06 evaluation reward: 6.58
 episode: 1794 score: 6.0 memory length: 404288 epsilon:
 0.39750778001020604 steps: 364 lr: 6.4000000000000001e-06 evaluation
 reward: 6.58
 episode: 1795 score: 12.0 memory length: 404752 epsilon:
 0.3965890600102002 steps: 464 lr: 6.4000000000000001e-06 evaluation
 reward: 6.51
 episode: 1796 score: 4.0 memory length: 405016 epsilon: 0.3960663400101969
 steps: 264 lr: 6.4000000000000001e-06 evaluation reward: 6.49
 episode: 1797 score: 8.0 memory length: 405435 epsilon:
 0.39523672001019167 steps: 419 lr: 6.4000000000000001e-06 evaluation
 reward: 6.52
 episode: 1798 score: 8.0 memory length: 405874 epsilon:
 0.39436750001018617 steps: 439 lr: 6.4000000000000001e-06 evaluation
 reward: 6.55
 episode: 1799 score: 15.0 memory length: 406496 epsilon:
 0.3931359400101784 steps: 622 lr: 6.4000000000000001e-06 evaluation
 reward: 6.63
 episode: 1800 score: 4.0 memory length: 406774 epsilon: 0.3925855000101749
 steps: 278 lr: 6.4000000000000001e-06 evaluation reward: 6.61
 episode: 1801 score: 8.0 memory length: 407210 epsilon:
 0.39172222001016943 steps: 436 lr: 6.4000000000000001e-06 evaluation
 reward: 6.64
 episode: 1802 score: 7.0 memory length: 407559 epsilon:
 0.39103120001016506 steps: 349 lr: 6.4000000000000001e-06 evaluation
 reward: 6.68
 episode: 1803 score: 5.0 memory length: 407864 epsilon:
 0.39042730001016124 steps: 305 lr: 6.4000000000000001e-06 evaluation
 reward: 6.68
 episode: 1804 score: 8.0 memory length: 408276 epsilon: 0.3896115400101561
 steps: 412 lr: 6.4000000000000001e-06 evaluation reward: 6.71
 episode: 1805 score: 6.0 memory length: 408633 epsilon: 0.3889046800101516
 steps: 357 lr: 6.4000000000000001e-06 evaluation reward: 6.7
 episode: 1806 score: 8.0 memory length: 409106 epsilon: 0.3879681400101457

```

steps: 473    lr: 6.4000000000000001e-06    evaluation reward: 6.7
episode: 1807  score: 5.0  memory length: 409412  epsilon:
0.38736226001014185  steps: 306  lr: 6.4000000000000001e-06    evaluation
reward: 6.69
episode: 1808  score: 8.0  memory length: 409813  epsilon: 0.3865682800101368
steps: 401    lr: 6.4000000000000001e-06    evaluation reward: 6.74
episode: 1809  score: 11.0  memory length: 410367  epsilon:
0.3854713600101299  steps: 554  lr: 6.4000000000000001e-06    evaluation
reward: 6.79
episode: 1810  score: 9.0  memory length: 410825  epsilon:
0.38456452001012414  steps: 458  lr: 6.4000000000000001e-06    evaluation
reward: 6.82
episode: 1811  score: 11.0  memory length: 411360  epsilon:
0.38350522001011744  steps: 535  lr: 6.4000000000000001e-06    evaluation
reward: 6.89
episode: 1812  score: 8.0  memory length: 411730  epsilon: 0.3827726200101128
steps: 370    lr: 6.4000000000000001e-06    evaluation reward: 6.94
episode: 1813  score: 3.0  memory length: 411943  epsilon:
0.38235088001011014  steps: 213  lr: 6.4000000000000001e-06    evaluation
reward: 6.87
episode: 1814  score: 4.0  memory length: 412206  epsilon:
0.38183014001010684  steps: 263  lr: 6.4000000000000001e-06    evaluation
reward: 6.84
episode: 1815  score: 7.0  memory length: 412589  epsilon:
0.38107180001010205  steps: 383  lr: 6.4000000000000001e-06    evaluation
reward: 6.83
episode: 1816  score: 7.0  memory length: 412971  epsilon:
0.38031544001009726  steps: 382  lr: 6.4000000000000001e-06    evaluation
reward: 6.84
episode: 1817  score: 4.0  memory length: 413231  epsilon: 0.379800640010094
steps: 260    lr: 6.4000000000000001e-06    evaluation reward: 6.83
episode: 1818  score: 6.0  memory length: 413586  epsilon:
0.37909774001008956  steps: 355  lr: 6.4000000000000001e-06    evaluation
reward: 6.85
episode: 1819  score: 10.0  memory length: 414106  epsilon:
0.37806814001008304  steps: 520  lr: 6.4000000000000001e-06    evaluation
reward: 6.89
episode: 1820  score: 9.0  memory length: 414566  epsilon: 0.3771573400100773
steps: 460    lr: 6.4000000000000001e-06    evaluation reward: 6.89
episode: 1821  score: 11.0  memory length: 415083  epsilon:
0.3761336800100708  steps: 517  lr: 6.4000000000000001e-06    evaluation
reward: 6.92
episode: 1822  score: 8.0  memory length: 415537  epsilon: 0.3752347600100651
steps: 454    lr: 6.4000000000000001e-06    evaluation reward: 6.97
episode: 1823  score: 6.0  memory length: 415914  epsilon: 0.3744883000100604
steps: 377    lr: 6.4000000000000001e-06    evaluation reward: 6.98
episode: 1824  score: 9.0  memory length: 416409  epsilon: 0.3735082000100542
steps: 495    lr: 6.4000000000000001e-06    evaluation reward: 7.01

```

episode: 1825 score: 5.0 memory length: 416739 epsilon:
 0.37285480001005006 steps: 330 lr: 6.4000000000000001e-06 evaluation
 reward: 7.01
 episode: 1826 score: 10.0 memory length: 417220 epsilon:
 0.37190242001004403 steps: 481 lr: 6.4000000000000001e-06 evaluation
 reward: 7.04
 episode: 1827 score: 6.0 memory length: 417534 epsilon: 0.3712807000100401
 steps: 314 lr: 6.4000000000000001e-06 evaluation reward: 7.02
 episode: 1828 score: 7.0 memory length: 417901 epsilon: 0.3705540400100355
 steps: 367 lr: 6.4000000000000001e-06 evaluation reward: 7.08
 episode: 1829 score: 11.0 memory length: 418456 epsilon:
 0.36945514001002855 steps: 555 lr: 6.4000000000000001e-06 evaluation
 reward: 7.14
 episode: 1830 score: 9.0 memory length: 418930 epsilon: 0.3685166200100226
 steps: 474 lr: 6.4000000000000001e-06 evaluation reward: 7.12
 episode: 1831 score: 16.0 memory length: 419540 epsilon:
 0.36730882001001497 steps: 610 lr: 6.4000000000000001e-06 evaluation
 reward: 7.23
 episode: 1832 score: 12.0 memory length: 420103 epsilon:
 0.3661940800100079 steps: 563 lr: 6.4000000000000001e-06 evaluation
 reward: 7.24
 episode: 1833 score: 9.0 memory length: 420531 epsilon:
 0.36534664001000255 steps: 428 lr: 6.4000000000000001e-06 evaluation
 reward: 7.25
 episode: 1834 score: 12.0 memory length: 421157 epsilon:
 0.3641071600099947 steps: 626 lr: 6.4000000000000001e-06 evaluation
 reward: 7.32
 episode: 1835 score: 8.0 memory length: 421563 epsilon: 0.3633032800099896
 steps: 406 lr: 6.4000000000000001e-06 evaluation reward: 7.32
 episode: 1836 score: 7.0 memory length: 421973 epsilon: 0.3624914800099845
 steps: 410 lr: 6.4000000000000001e-06 evaluation reward: 7.35
 episode: 1837 score: 11.0 memory length: 422536 epsilon:
 0.36137674000997744 steps: 563 lr: 6.4000000000000001e-06 evaluation
 reward: 7.43
 episode: 1838 score: 7.0 memory length: 422986 epsilon: 0.3604857400099718
 steps: 450 lr: 6.4000000000000001e-06 evaluation reward: 7.47
 episode: 1839 score: 7.0 memory length: 423368 epsilon: 0.359729380009967
 steps: 382 lr: 6.4000000000000001e-06 evaluation reward: 7.41
 episode: 1840 score: 10.0 memory length: 423823 epsilon:
 0.3588284800099613 steps: 455 lr: 6.4000000000000001e-06 evaluation
 reward: 7.44
 episode: 1841 score: 2.0 memory length: 424006 epsilon: 0.358466140009959
 steps: 183 lr: 6.4000000000000001e-06 evaluation reward: 7.38
 episode: 1842 score: 9.0 memory length: 424462 epsilon: 0.3575632600099533
 steps: 456 lr: 6.4000000000000001e-06 evaluation reward: 7.41
 episode: 1843 score: 7.0 memory length: 424867 epsilon:
 0.35676136000994824 steps: 405 lr: 6.4000000000000001e-06 evaluation
 reward: 7.43

episode: 1844 score: 7.0 memory length: 425258 epsilon:
 0.35598718000994334 steps: 391 lr: 6.400000000000001e-06 evaluation
 reward: 7.44
 episode: 1845 score: 4.0 memory length: 425500 epsilon: 0.3555080200099403
 steps: 242 lr: 6.400000000000001e-06 evaluation reward: 7.39
 episode: 1846 score: 12.0 memory length: 426055 epsilon:
 0.35440912000993335 steps: 555 lr: 6.400000000000001e-06 evaluation
 reward: 7.45
 episode: 1847 score: 6.0 memory length: 426420 epsilon: 0.3536864200099288
 steps: 365 lr: 6.400000000000001e-06 evaluation reward: 7.46
 episode: 1848 score: 8.0 memory length: 426893 epsilon:
 0.35274988000992286 steps: 473 lr: 6.400000000000001e-06 evaluation
 reward: 7.49
 episode: 1849 score: 9.0 memory length: 427359 epsilon: 0.351827200009917
 steps: 466 lr: 6.400000000000001e-06 evaluation reward: 7.49
 episode: 1850 score: 7.0 memory length: 427767 epsilon: 0.3510193600099119
 steps: 408 lr: 6.400000000000001e-06 evaluation reward: 7.47
 episode: 1851 score: 10.0 memory length: 428289 epsilon:
 0.34998580000990537 steps: 522 lr: 6.400000000000001e-06 evaluation
 reward: 7.47
 episode: 1852 score: 5.0 memory length: 428599 epsilon: 0.3493720000099015
 steps: 310 lr: 6.400000000000001e-06 evaluation reward: 7.46
 episode: 1853 score: 2.0 memory length: 428797 epsilon: 0.348979960009899
 steps: 198 lr: 6.400000000000001e-06 evaluation reward: 7.4
 episode: 1854 score: 7.0 memory length: 429181 epsilon: 0.3482196400098942
 steps: 384 lr: 6.400000000000001e-06 evaluation reward: 7.39
 episode: 1855 score: 4.0 memory length: 429441 epsilon:
 0.34770484000989094 steps: 260 lr: 6.400000000000001e-06 evaluation
 reward: 7.32
 episode: 1856 score: 8.0 memory length: 429883 epsilon: 0.3468296800098854
 steps: 442 lr: 6.400000000000001e-06 evaluation reward: 7.33
 episode: 1857 score: 11.0 memory length: 430395 epsilon: 0.345815920009879
 steps: 512 lr: 6.400000000000001e-06 evaluation reward: 7.38
 episode: 1858 score: 4.0 memory length: 430656 epsilon: 0.3452991400098757
 steps: 261 lr: 6.400000000000001e-06 evaluation reward: 7.36
 episode: 1859 score: 5.0 memory length: 430965 epsilon:
 0.34468732000987184 steps: 309 lr: 6.400000000000001e-06 evaluation
 reward: 7.33
 episode: 1860 score: 9.0 memory length: 431441 epsilon: 0.3437448400098659
 steps: 476 lr: 6.400000000000001e-06 evaluation reward: 7.38
 episode: 1861 score: 9.0 memory length: 431766 epsilon: 0.3431013400098618
 steps: 325 lr: 6.400000000000001e-06 evaluation reward: 7.41
 episode: 1862 score: 6.0 memory length: 432126 epsilon: 0.3423885400098573
 steps: 360 lr: 6.400000000000001e-06 evaluation reward: 7.42
 episode: 1863 score: 10.0 memory length: 432628 epsilon: 0.341394580009851
 steps: 502 lr: 6.400000000000001e-06 evaluation reward: 7.45
 episode: 1864 score: 9.0 memory length: 433121 epsilon:
 0.34041844000984484 steps: 493 lr: 6.400000000000001e-06 evaluation

```

reward: 7.46
episode: 1865  score: 6.0  memory length: 433457  epsilon: 0.3397531600098406
steps: 336    lr: 6.4000000000000001e-06  evaluation reward: 7.43
episode: 1866  score: 6.0  memory length: 433786  epsilon: 0.3391017400098365
steps: 329    lr: 6.4000000000000001e-06  evaluation reward: 7.43
episode: 1867  score: 8.0  memory length: 434157  epsilon:
0.33836716000983186  steps: 371  lr: 6.4000000000000001e-06  evaluation
reward: 7.43
episode: 1868  score: 9.0  memory length: 434642  epsilon: 0.3374068600098258
steps: 485    lr: 6.4000000000000001e-06  evaluation reward: 7.45
episode: 1869  score: 5.0  memory length: 434967  epsilon: 0.3367633600098217
steps: 325    lr: 6.4000000000000001e-06  evaluation reward: 7.43
episode: 1870  score: 10.0  memory length: 435538  epsilon:
0.33563278000981456  steps: 571  lr: 6.4000000000000001e-06  evaluation
reward: 7.49
episode: 1871  score: 4.0  memory length: 435801  epsilon:
0.33511204000981126  steps: 263  lr: 6.4000000000000001e-06  evaluation
reward: 7.49
episode: 1872  score: 6.0  memory length: 436178  epsilon:
0.33436558000980654  steps: 377  lr: 6.4000000000000001e-06  evaluation
reward: 7.5
episode: 1873  score: 7.0  memory length: 436543  epsilon:
0.33364288000980197  steps: 365  lr: 6.4000000000000001e-06  evaluation
reward: 7.52
episode: 1874  score: 9.0  memory length: 437019  epsilon: 0.332700400009796
steps: 476    lr: 6.4000000000000001e-06  evaluation reward: 7.54
episode: 1875  score: 11.0  memory length: 437592  epsilon:
0.3315658600097888  steps: 573  lr: 6.4000000000000001e-06  evaluation
reward: 7.57
episode: 1876  score: 8.0  memory length: 437995  epsilon: 0.3307679200097838
steps: 403    lr: 6.4000000000000001e-06  evaluation reward: 7.58
episode: 1877  score: 4.0  memory length: 438290  epsilon: 0.3301838200097801
steps: 295    lr: 6.4000000000000001e-06  evaluation reward: 7.59
episode: 1878  score: 6.0  memory length: 438639  epsilon: 0.3294928000097757
steps: 349    lr: 6.4000000000000001e-06  evaluation reward: 7.57
episode: 1879  score: 3.0  memory length: 438852  epsilon:
0.32907106000977304  steps: 213  lr: 6.4000000000000001e-06  evaluation
reward: 7.54
episode: 1880  score: 7.0  memory length: 439258  epsilon:
0.32826718000976796  steps: 406  lr: 6.4000000000000001e-06  evaluation
reward: 7.54
episode: 1881  score: 4.0  memory length: 439538  epsilon:
0.32771278000976445  steps: 280  lr: 6.4000000000000001e-06  evaluation
reward: 7.51
episode: 1882  score: 9.0  memory length: 439969  epsilon:
0.32685940000975905  steps: 431  lr: 6.4000000000000001e-06  evaluation
reward: 7.56
episode: 1883  score: 8.0  memory length: 440360  epsilon:

```

0.32608522000975415 steps: 391 lr: 6.4000000000000001e-06 evaluation
 reward: 7.59
 episode: 1884 score: 2.0 memory length: 440559 epsilon:
 0.32569120000975166 steps: 199 lr: 6.4000000000000001e-06 evaluation
 reward: 7.52
 episode: 1885 score: 6.0 memory length: 440897 epsilon: 0.3250219600097474
 steps: 338 lr: 6.4000000000000001e-06 evaluation reward: 7.5
 episode: 1886 score: 9.0 memory length: 441371 epsilon: 0.3240834400097415
 steps: 474 lr: 6.4000000000000001e-06 evaluation reward: 7.49
 episode: 1887 score: 13.0 memory length: 441882 epsilon:
 0.3230716600097351 steps: 511 lr: 6.4000000000000001e-06 evaluation
 reward: 7.57
 episode: 1888 score: 11.0 memory length: 442422 epsilon:
 0.3220024600097283 steps: 540 lr: 6.4000000000000001e-06 evaluation
 reward: 7.59
 episode: 1889 score: 6.0 memory length: 442778 epsilon:
 0.32129758000972386 steps: 356 lr: 6.4000000000000001e-06 evaluation
 reward: 7.58
 episode: 1890 score: 4.0 memory length: 443059 epsilon:
 0.32074120000972034 steps: 281 lr: 6.4000000000000001e-06 evaluation
 reward: 7.57
 episode: 1891 score: 5.0 memory length: 443385 epsilon:
 0.32009572000971626 steps: 326 lr: 6.4000000000000001e-06 evaluation
 reward: 7.54
 episode: 1892 score: 4.0 memory length: 443630 epsilon: 0.3196106200097132
 steps: 245 lr: 6.4000000000000001e-06 evaluation reward: 7.5
 episode: 1893 score: 12.0 memory length: 444203 epsilon: 0.318476080009706
 steps: 573 lr: 6.4000000000000001e-06 evaluation reward: 7.54
 episode: 1894 score: 1.0 memory length: 444355 epsilon: 0.3181751200097041
 steps: 152 lr: 6.4000000000000001e-06 evaluation reward: 7.49
 episode: 1895 score: 10.0 memory length: 444869 epsilon:
 0.31715740000969767 steps: 514 lr: 6.4000000000000001e-06 evaluation
 reward: 7.47
 episode: 1896 score: 4.0 memory length: 445145 epsilon: 0.3166109200096942
 steps: 276 lr: 6.4000000000000001e-06 evaluation reward: 7.47
 episode: 1897 score: 6.0 memory length: 445500 epsilon:
 0.31590802000968976 steps: 355 lr: 6.4000000000000001e-06 evaluation
 reward: 7.45
 episode: 1898 score: 10.0 memory length: 446006 epsilon:
 0.3149061400096834 steps: 506 lr: 6.4000000000000001e-06 evaluation
 reward: 7.47
 episode: 1899 score: 13.0 memory length: 446556 epsilon:
 0.31381714000967653 steps: 550 lr: 6.4000000000000001e-06 evaluation
 reward: 7.45
 episode: 1900 score: 11.0 memory length: 447078 epsilon: 0.31278358000967
 steps: 522 lr: 6.4000000000000001e-06 evaluation reward: 7.52
 episode: 1901 score: 8.0 memory length: 447498 epsilon:
 0.31195198000966473 steps: 420 lr: 6.4000000000000001e-06 evaluation


```

reward: 7.52
episode: 1902  score: 10.0  memory length: 448002  epsilon:
0.3109540600096584  steps: 504  lr: 6.4000000000000001e-06  evaluation
reward: 7.55
episode: 1903  score: 9.0  memory length: 448497  epsilon: 0.3099739600096522
steps: 495  lr: 6.4000000000000001e-06  evaluation reward: 7.59
episode: 1904  score: 6.0  memory length: 448833  epsilon: 0.309308680009648
steps: 336  lr: 6.4000000000000001e-06  evaluation reward: 7.57
episode: 1905  score: 14.0  memory length: 449444  epsilon:
0.30809890000964035  steps: 611  lr: 6.4000000000000001e-06  evaluation
reward: 7.65
episode: 1906  score: 5.0  memory length: 449749  epsilon:
0.30749500000963653  steps: 305  lr: 6.4000000000000001e-06  evaluation
reward: 7.62
episode: 1907  score: 9.0  memory length: 450233  epsilon:
0.30653668000963047  steps: 484  lr: 6.4000000000000001e-06  evaluation
reward: 7.66
episode: 1908  score: 11.0  memory length: 450725  epsilon:
0.3055625200096243  steps: 492  lr: 6.4000000000000001e-06  evaluation
reward: 7.69
episode: 1909  score: 6.0  memory length: 451083  epsilon: 0.3048536800096198
steps: 358  lr: 6.4000000000000001e-06  evaluation reward: 7.64
episode: 1910  score: 8.0  memory length: 451521  epsilon:
0.30398644000961433  steps: 438  lr: 6.4000000000000001e-06  evaluation
reward: 7.63
episode: 1911  score: 4.0  memory length: 451782  epsilon:
0.30346966000961106  steps: 261  lr: 6.4000000000000001e-06  evaluation
reward: 7.56
episode: 1912  score: 4.0  memory length: 452062  epsilon:
0.30291526000960756  steps: 280  lr: 6.4000000000000001e-06  evaluation
reward: 7.52
episode: 1913  score: 6.0  memory length: 452416  epsilon: 0.3022143400096031
steps: 354  lr: 6.4000000000000001e-06  evaluation reward: 7.55
episode: 1914  score: 4.0  memory length: 452676  epsilon:
0.30169954000959986  steps: 260  lr: 6.4000000000000001e-06  evaluation
reward: 7.55
episode: 1915  score: 11.0  memory length: 453191  epsilon:
0.3006798400095934  steps: 515  lr: 6.4000000000000001e-06  evaluation
reward: 7.59
episode: 1916  score: 8.0  memory length: 453585  epsilon: 0.2998997200095885
steps: 394  lr: 6.4000000000000001e-06  evaluation reward: 7.6
episode: 1917  score: 9.0  memory length: 454090  epsilon:
0.29889982000958215  steps: 505  lr: 6.4000000000000001e-06  evaluation
reward: 7.65
episode: 1918  score: 9.0  memory length: 454511  epsilon: 0.2980662400095769
steps: 421  lr: 6.4000000000000001e-06  evaluation reward: 7.68
episode: 1919  score: 9.0  memory length: 454969  epsilon:
0.29715940000957114  steps: 458  lr: 6.4000000000000001e-06  evaluation

```

```

reward: 7.67
episode: 1920  score: 8.0  memory length: 455370  epsilon: 0.2963654200095661
steps: 401  lr: 6.4000000000000001e-06  evaluation reward: 7.66
episode: 1921  score: 7.0  memory length: 455759  epsilon:
0.29559520000956124  steps: 389  lr: 6.4000000000000001e-06  evaluation
reward: 7.62
episode: 1922  score: 8.0  memory length: 456189  epsilon:
0.29474380000955586  steps: 430  lr: 6.4000000000000001e-06  evaluation
reward: 7.62
episode: 1923  score: 8.0  memory length: 456614  epsilon:
0.29390230000955053  steps: 425  lr: 6.4000000000000001e-06  evaluation
reward: 7.64
episode: 1924  score: 9.0  memory length: 457071  epsilon: 0.2929974400095448
steps: 457  lr: 6.4000000000000001e-06  evaluation reward: 7.64
episode: 1925  score: 10.0  memory length: 457601  epsilon:
0.29194804000953817  steps: 530  lr: 6.4000000000000001e-06  evaluation
reward: 7.69
episode: 1926  score: 5.0  memory length: 457911  epsilon: 0.2913342400095343
steps: 310  lr: 6.4000000000000001e-06  evaluation reward: 7.64
episode: 1927  score: 7.0  memory length: 458310  epsilon: 0.2905442200095293
steps: 399  lr: 6.4000000000000001e-06  evaluation reward: 7.65
episode: 1928  score: 6.0  memory length: 458654  epsilon: 0.289863100009525
steps: 344  lr: 6.4000000000000001e-06  evaluation reward: 7.64
episode: 1929  score: 7.0  memory length: 459029  epsilon: 0.2891206000095203
steps: 375  lr: 6.4000000000000001e-06  evaluation reward: 7.6
episode: 1930  score: 6.0  memory length: 459405  epsilon:
0.28837612000951557  steps: 376  lr: 6.4000000000000001e-06  evaluation
reward: 7.57
episode: 1931  score: 15.0  memory length: 460014  epsilon:
0.28717030000950794  steps: 609  lr: 6.4000000000000001e-06  evaluation
reward: 7.56
episode: 1932  score: 11.0  memory length: 460506  epsilon:
0.2861961400095018  steps: 492  lr: 6.4000000000000001e-06  evaluation
reward: 7.55
episode: 1933  score: 8.0  memory length: 460939  epsilon:
0.28533880000949635  steps: 433  lr: 6.4000000000000001e-06  evaluation
reward: 7.54
episode: 1934  score: 4.0  memory length: 461259  epsilon:
0.28470520000949234  steps: 320  lr: 6.4000000000000001e-06  evaluation
reward: 7.46
episode: 1935  score: 13.0  memory length: 461855  epsilon:
0.2835251200094849  steps: 596  lr: 6.4000000000000001e-06  evaluation
reward: 7.51
episode: 1936  score: 8.0  memory length: 462284  epsilon: 0.2826757000094795
steps: 429  lr: 6.4000000000000001e-06  evaluation reward: 7.52
episode: 1937  score: 9.0  memory length: 462759  epsilon:
0.28173520000947355  steps: 475  lr: 6.4000000000000001e-06  evaluation
reward: 7.5

```

episode: 1938 score: 5.0 memory length: 463069 epsilon:
 0.28112140000946967 steps: 310 lr: 6.400000000000001e-06 evaluation
 reward: 7.48
 episode: 1939 score: 6.0 memory length: 463413 epsilon:
 0.28044028000946536 steps: 344 lr: 6.400000000000001e-06 evaluation
 reward: 7.47
 episode: 1940 score: 5.0 memory length: 463688 epsilon: 0.2798957800094619
 steps: 275 lr: 6.400000000000001e-06 evaluation reward: 7.42
 episode: 1941 score: 4.0 memory length: 463947 epsilon:
 0.27938296000945867 steps: 259 lr: 6.400000000000001e-06 evaluation
 reward: 7.44
 episode: 1942 score: 17.0 memory length: 464607 epsilon:
 0.2780761600094504 steps: 660 lr: 6.400000000000001e-06 evaluation
 reward: 7.52
 episode: 1943 score: 6.0 memory length: 464984 epsilon: 0.2773297000094457
 steps: 377 lr: 6.400000000000001e-06 evaluation reward: 7.51
 episode: 1944 score: 11.0 memory length: 465473 epsilon:
 0.27636148000943955 steps: 489 lr: 6.400000000000001e-06 evaluation
 reward: 7.55
 episode: 1945 score: 9.0 memory length: 465952 epsilon:
 0.27541306000943355 steps: 479 lr: 6.400000000000001e-06 evaluation
 reward: 7.6
 episode: 1946 score: 5.0 memory length: 466259 epsilon: 0.2748052000094297
 steps: 307 lr: 6.400000000000001e-06 evaluation reward: 7.53
 episode: 1947 score: 11.0 memory length: 466814 epsilon:
 0.27370630000942275 steps: 555 lr: 6.400000000000001e-06 evaluation
 reward: 7.58
 episode: 1948 score: 6.0 memory length: 467136 epsilon: 0.2730687400094187
 steps: 322 lr: 6.400000000000001e-06 evaluation reward: 7.56
 episode: 1949 score: 9.0 memory length: 467639 epsilon: 0.2720728000094124
 steps: 503 lr: 6.400000000000001e-06 evaluation reward: 7.56
 episode: 1950 score: 14.0 memory length: 468277 epsilon:
 0.2708095600094044 steps: 638 lr: 6.400000000000001e-06 evaluation
 reward: 7.63
 episode: 1951 score: 4.0 memory length: 468519 epsilon: 0.2703304000094014
 steps: 242 lr: 6.400000000000001e-06 evaluation reward: 7.57
 episode: 1952 score: 7.0 memory length: 468874 epsilon:
 0.26962750000939695 steps: 355 lr: 6.400000000000001e-06 evaluation
 reward: 7.59
 episode: 1953 score: 10.0 memory length: 469362 epsilon:
 0.26866126000939083 steps: 488 lr: 6.400000000000001e-06 evaluation
 reward: 7.67
 episode: 1954 score: 8.0 memory length: 469813 epsilon: 0.2677682800093852
 steps: 451 lr: 6.400000000000001e-06 evaluation reward: 7.68
 episode: 1955 score: 6.0 memory length: 470136 epsilon:
 0.26712874000938114 steps: 323 lr: 6.400000000000001e-06 evaluation
 reward: 7.7
 episode: 1956 score: 9.0 memory length: 470618 epsilon: 0.2661743800093751

```

steps: 482    lr: 6.4000000000000001e-06    evaluation reward: 7.71
episode: 1957  score: 4.0    memory length: 470882    epsilon: 0.2656516600093718
steps: 264    lr: 6.4000000000000001e-06    evaluation reward: 7.64
episode: 1958  score: 6.0    memory length: 471243    epsilon:
0.26493688000936727    steps: 361    lr: 6.4000000000000001e-06    evaluation
reward: 7.66
episode: 1959  score: 15.0    memory length: 471824    epsilon: 0.26378650000936
steps: 581    lr: 6.4000000000000001e-06    evaluation reward: 7.76
episode: 1960  score: 13.0    memory length: 472453    epsilon:
0.2625410800093521    steps: 629    lr: 6.4000000000000001e-06    evaluation
reward: 7.8
episode: 1961  score: 7.0    memory length: 472835    epsilon: 0.2617847200093473
steps: 382    lr: 6.4000000000000001e-06    evaluation reward: 7.78
episode: 1962  score: 7.0    memory length: 473225    epsilon:
0.26101252000934244    steps: 390    lr: 6.4000000000000001e-06    evaluation
reward: 7.79
episode: 1963  score: 7.0    memory length: 473614    epsilon:
0.26024230000933757    steps: 389    lr: 6.4000000000000001e-06    evaluation
reward: 7.76
episode: 1964  score: 2.0    memory length: 473797    epsilon: 0.2598799600093353
steps: 183    lr: 6.4000000000000001e-06    evaluation reward: 7.69
episode: 1965  score: 16.0    memory length: 474448    epsilon:
0.2585909800093271    steps: 651    lr: 6.4000000000000001e-06    evaluation
reward: 7.79
episode: 1966  score: 7.0    memory length: 474833    epsilon: 0.2578286800093223
steps: 385    lr: 6.4000000000000001e-06    evaluation reward: 7.8
episode: 1967  score: 8.0    memory length: 475260    epsilon:
0.25698322000931695    steps: 427    lr: 6.4000000000000001e-06    evaluation
reward: 7.8
episode: 1968  score: 10.0    memory length: 475749    epsilon:
0.2560150000093108    steps: 489    lr: 6.4000000000000001e-06    evaluation
reward: 7.81
episode: 1969  score: 5.0    memory length: 476023    epsilon: 0.2554724800093074
steps: 274    lr: 6.4000000000000001e-06    evaluation reward: 7.81
episode: 1970  score: 14.0    memory length: 476662    epsilon:
0.2542072600092994    steps: 639    lr: 6.4000000000000001e-06    evaluation
reward: 7.85
episode: 1971  score: 5.0    memory length: 476955    epsilon: 0.2536271200092957
steps: 293    lr: 6.4000000000000001e-06    evaluation reward: 7.86
episode: 1972  score: 7.0    memory length: 477350    epsilon:
0.25284502000929077    steps: 395    lr: 6.4000000000000001e-06    evaluation
reward: 7.87
episode: 1973  score: 9.0    memory length: 477855    epsilon:
0.25184512000928444    steps: 505    lr: 6.4000000000000001e-06    evaluation
reward: 7.89
episode: 1974  score: 6.0    memory length: 478213    epsilon:
0.25113628000927996    steps: 358    lr: 6.4000000000000001e-06    evaluation
reward: 7.86

```

episode: 1975 score: 9.0 memory length: 478689 epsilon: 0.250193800009274
 steps: 476 lr: 6.400000000000001e-06 evaluation reward: 7.84
 episode: 1976 score: 5.0 memory length: 479015 epsilon: 0.2495483200092699
 steps: 326 lr: 6.400000000000001e-06 evaluation reward: 7.81
 episode: 1977 score: 15.0 memory length: 479583 epsilon:
 0.2484236800092628 steps: 568 lr: 6.400000000000001e-06 evaluation
 reward: 7.92
 episode: 1978 score: 10.0 memory length: 480072 epsilon:
 0.24745546000925667 steps: 489 lr: 6.400000000000001e-06 evaluation
 reward: 7.96
 episode: 1979 score: 10.0 memory length: 480427 epsilon:
 0.24675256000925222 steps: 355 lr: 6.400000000000001e-06 evaluation
 reward: 8.03
 episode: 1980 score: 6.0 memory length: 480763 epsilon: 0.246087280009248
 steps: 336 lr: 6.400000000000001e-06 evaluation reward: 8.02
 episode: 1981 score: 9.0 memory length: 481245 epsilon:
 0.24513292000924197 steps: 482 lr: 6.400000000000001e-06 evaluation
 reward: 8.07
 episode: 1982 score: 11.0 memory length: 481743 epsilon:
 0.24414688000923573 steps: 498 lr: 6.400000000000001e-06 evaluation
 reward: 8.09
 episode: 1983 score: 9.0 memory length: 482200 epsilon: 0.24324202000923
 steps: 457 lr: 6.400000000000001e-06 evaluation reward: 8.1
 episode: 1984 score: 5.0 memory length: 482491 epsilon:
 0.24266584000922636 steps: 291 lr: 6.400000000000001e-06 evaluation
 reward: 8.13
 episode: 1985 score: 7.0 memory length: 482897 epsilon:
 0.24186196000922128 steps: 406 lr: 6.400000000000001e-06 evaluation
 reward: 8.14
 episode: 1986 score: 9.0 memory length: 483224 epsilon:
 0.24121450000921718 steps: 327 lr: 6.400000000000001e-06 evaluation
 reward: 8.14
 episode: 1987 score: 7.0 memory length: 483604 epsilon:
 0.24046210000921242 steps: 380 lr: 6.400000000000001e-06 evaluation
 reward: 8.08
 episode: 1988 score: 6.0 memory length: 483963 epsilon:
 0.23975128000920792 steps: 359 lr: 6.400000000000001e-06 evaluation
 reward: 8.03
 episode: 1989 score: 12.0 memory length: 484467 epsilon:
 0.2387533600092016 steps: 504 lr: 6.400000000000001e-06 evaluation
 reward: 8.09
 episode: 1990 score: 6.0 memory length: 484804 epsilon: 0.2380861000091974
 steps: 337 lr: 6.400000000000001e-06 evaluation reward: 8.11
 episode: 1991 score: 9.0 memory length: 485258 epsilon: 0.2371871800091917
 steps: 454 lr: 6.400000000000001e-06 evaluation reward: 8.15
 episode: 1992 score: 17.0 memory length: 485897 epsilon:
 0.2359219600091837 steps: 639 lr: 6.400000000000001e-06 evaluation
 reward: 8.28

episode: 1993 score: 5.0 memory length: 486226 epsilon:
 0.23527054000917957 steps: 329 lr: 6.4000000000000001e-06 evaluation
 reward: 8.21
 episode: 1994 score: 9.0 memory length: 486554 epsilon:
 0.23462110000917547 steps: 328 lr: 6.4000000000000001e-06 evaluation
 reward: 8.29
 episode: 1995 score: 6.0 memory length: 486872 epsilon:
 0.23399146000917148 steps: 318 lr: 6.4000000000000001e-06 evaluation
 reward: 8.25
 episode: 1996 score: 9.0 memory length: 487339 epsilon:
 0.23306680000916563 steps: 467 lr: 6.4000000000000001e-06 evaluation
 reward: 8.3
 episode: 1997 score: 5.0 memory length: 487613 epsilon: 0.2325242800091622
 steps: 274 lr: 6.4000000000000001e-06 evaluation reward: 8.29
 episode: 1998 score: 15.0 memory length: 488274 epsilon:
 0.23121550000915392 steps: 661 lr: 6.4000000000000001e-06 evaluation
 reward: 8.34
 episode: 1999 score: 10.0 memory length: 488799 epsilon:
 0.23017600000914734 steps: 525 lr: 6.4000000000000001e-06 evaluation
 reward: 8.31
 episode: 2000 score: 10.0 memory length: 489323 epsilon:
 0.22913848000914078 steps: 524 lr: 6.4000000000000001e-06 evaluation
 reward: 8.3
 episode: 2001 score: 9.0 memory length: 489842 epsilon:
 0.22811086000913428 steps: 519 lr: 6.4000000000000001e-06 evaluation
 reward: 8.31
 episode: 2002 score: 11.0 memory length: 490340 epsilon:
 0.22712482000912804 steps: 498 lr: 6.4000000000000001e-06 evaluation
 reward: 8.32
 episode: 2003 score: 8.0 memory length: 490748 epsilon:
 0.22631698000912293 steps: 408 lr: 6.4000000000000001e-06 evaluation
 reward: 8.31
 episode: 2004 score: 10.0 memory length: 491253 epsilon:
 0.2253170800091166 steps: 505 lr: 6.4000000000000001e-06 evaluation
 reward: 8.35
 episode: 2005 score: 8.0 memory length: 491643 epsilon: 0.2245448800091117
 steps: 390 lr: 6.4000000000000001e-06 evaluation reward: 8.29
 episode: 2006 score: 5.0 memory length: 491956 epsilon: 0.2239251400091078
 steps: 313 lr: 6.4000000000000001e-06 evaluation reward: 8.29
 episode: 2007 score: 9.0 memory length: 492431 epsilon:
 0.22298464000910184 steps: 475 lr: 6.4000000000000001e-06 evaluation
 reward: 8.29
 episode: 2008 score: 11.0 memory length: 492954 epsilon:
 0.2219491000090953 steps: 523 lr: 6.4000000000000001e-06 evaluation
 reward: 8.29
 episode: 2009 score: 6.0 memory length: 493302 epsilon:
 0.22126006000909093 steps: 348 lr: 6.4000000000000001e-06 evaluation
 reward: 8.29

episode: 2010 score: 10.0 memory length: 493760 epsilon:
 0.2203532200090852 steps: 458 lr: 6.4000000000000001e-06 evaluation
 reward: 8.31
 episode: 2011 score: 11.0 memory length: 494292 epsilon:
 0.21929986000907853 steps: 532 lr: 6.4000000000000001e-06 evaluation
 reward: 8.38
 episode: 2012 score: 12.0 memory length: 494719 epsilon:
 0.21845440000907318 steps: 427 lr: 6.4000000000000001e-06 evaluation
 reward: 8.46
 episode: 2013 score: 9.0 memory length: 495028 epsilon: 0.2178425800090693
 steps: 309 lr: 6.4000000000000001e-06 evaluation reward: 8.49
 episode: 2014 score: 7.0 memory length: 495439 epsilon:
 0.21702880000906416 steps: 411 lr: 6.4000000000000001e-06 evaluation
 reward: 8.52
 episode: 2015 score: 8.0 memory length: 495816 epsilon:
 0.21628234000905944 steps: 377 lr: 6.4000000000000001e-06 evaluation
 reward: 8.49
 episode: 2016 score: 5.0 memory length: 496108 epsilon:
 0.21570418000905578 steps: 292 lr: 6.4000000000000001e-06 evaluation
 reward: 8.46
 episode: 2017 score: 10.0 memory length: 496592 epsilon:
 0.21474586000904972 steps: 484 lr: 6.4000000000000001e-06 evaluation
 reward: 8.47
 episode: 2018 score: 10.0 memory length: 497118 epsilon:
 0.21370438000904313 steps: 526 lr: 6.4000000000000001e-06 evaluation
 reward: 8.48
 episode: 2019 score: 18.0 memory length: 497789 epsilon:
 0.21237580000903472 steps: 671 lr: 6.4000000000000001e-06 evaluation
 reward: 8.57
 episode: 2020 score: 11.0 memory length: 498372 epsilon:
 0.21122146000902742 steps: 583 lr: 6.4000000000000001e-06 evaluation
 reward: 8.6
 episode: 2021 score: 15.0 memory length: 498942 epsilon:
 0.21009286000902028 steps: 570 lr: 6.4000000000000001e-06 evaluation
 reward: 8.68
 episode: 2022 score: 11.0 memory length: 499390 epsilon:
 0.20920582000901466 steps: 448 lr: 6.4000000000000001e-06 evaluation
 reward: 8.71
 episode: 2023 score: 9.0 memory length: 499877 epsilon:
 0.20824156000900856 steps: 487 lr: 6.4000000000000001e-06 evaluation
 reward: 8.72
 episode: 2024 score: 16.0 memory length: 500449 epsilon:
 0.2071090000090014 steps: 572 lr: 2.5600000000000001e-06 evaluation
 reward: 8.79
 episode: 2025 score: 5.0 memory length: 500741 epsilon:
 0.20653084000899774 steps: 292 lr: 2.5600000000000001e-06 evaluation
 reward: 8.74
 episode: 2026 score: 6.0 memory length: 501079 epsilon: 0.2058616000089935

```

steps: 338    lr: 2.5600000000000001e-06    evaluation reward: 8.75
episode: 2027    score: 5.0    memory length: 501406    epsilon: 0.2052141400089894
steps: 327    lr: 2.5600000000000001e-06    evaluation reward: 8.73
episode: 2028    score: 12.0    memory length: 501989    epsilon:
0.2040598000089821    steps: 583    lr: 2.5600000000000001e-06    evaluation
reward: 8.79
episode: 2029    score: 8.0    memory length: 502432    epsilon:
0.20318266000897656    steps: 443    lr: 2.5600000000000001e-06    evaluation
reward: 8.8
episode: 2030    score: 11.0    memory length: 503009    epsilon:
0.20204020000896933    steps: 577    lr: 2.5600000000000001e-06    evaluation
reward: 8.85

```

```

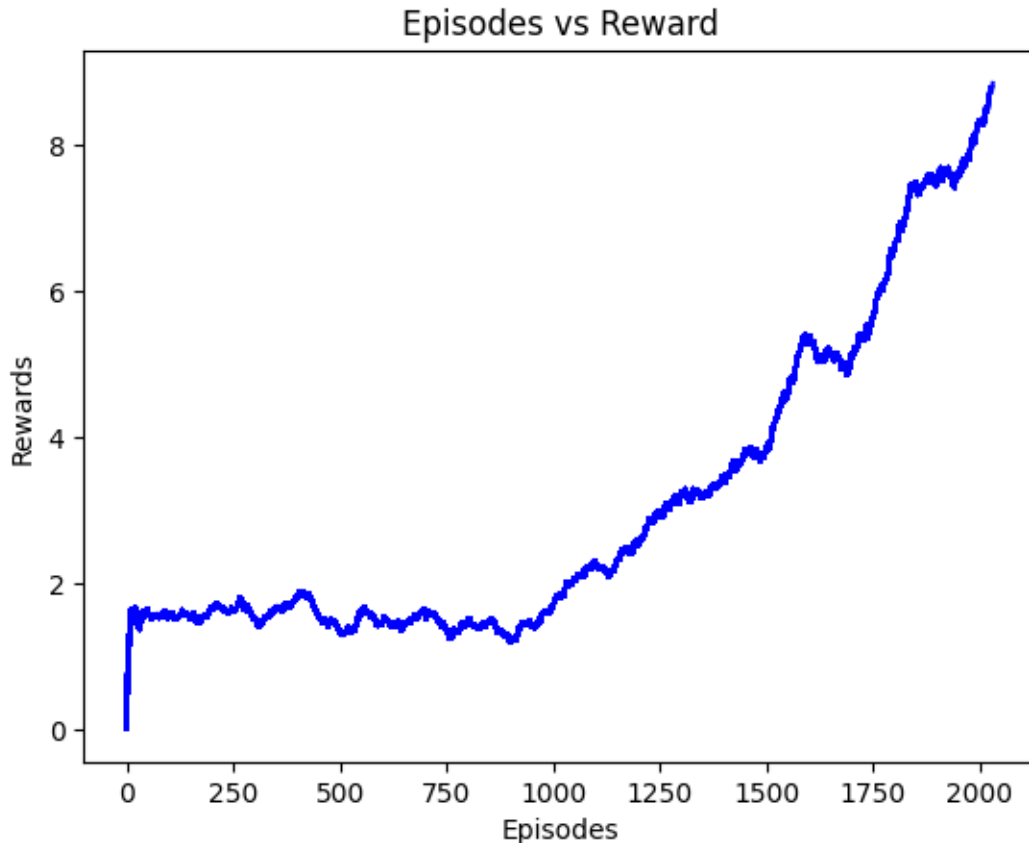
-----
KeyboardInterrupt                                Traceback (most recent call last)
<ipython-input-13-3b074d992556> in <cell line: 3>()
    36         # Start training after random sample generation
    37         if(frame >= train_frame): # You can set train_frame to a lower_
↳value while testing 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.py in_
↳train_policy_net(self, frame)
    61             self.epsilon -= self.epsilon_decay
    62
----> 63             mini_batch = self.memory.sample_mini_batch(frame)
    64             mini_batch = np.array(mini_batch, dtype=object).transpose()
    65

/content/drive/MyDrive/CS444/assignment5/assignment5_materials/memory.py in_
↳sample_mini_batch(self, frame)
    26             sample = []
    27             for j in range(HISTORY_SIZE + 1):
----> 28                 sample.append(self.memory[i + j])
    29
    30             sample = np.array(sample, dtype=object)

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
```

Requirement already satisfied: xvfbwrapper in /usr/local/lib/python3.10/dist-packages (0.2.9)

```
[ ]: 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()

```