

```
In [1]: # This Python 3 environment comes with many helpful analytics libraries installed
# It is defined by the kaggle/python Docker image: https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

# Input data files are available in the read-only "../input/" directory
# For example, running this (by clicking run or pressing Shift+Enter) will list all files in the input directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 20GB to the current directory (/kaggle/working/) that gets preserved
# You can also write temporary files to /kaggle/temp/, but they won't be saved outside of the session

/kaggle/input/mp5dataset-1/assignment5_materials/agent.py
/kaggle/input/mp5dataset-1/assignment5_materials/MP5_DQN_LSTM_EC.ipynb
/kaggle/input/mp5dataset-1/assignment5_materials/memory.py
/kaggle/input/mp5dataset-1/assignment5_materials/model.py
/kaggle/input/mp5dataset-1/assignment5_materials/config.py
/kaggle/input/mp5dataset-1/assignment5_materials/agent_double.py
/kaggle/input/mp5dataset-1/assignment5_materials/Utils.py
/kaggle/input/mp5dataset-1/assignment5_materials/MP5.ipynb
/kaggle/input/mp5dataset-1/assignment5_materials/agent_lstm.py
/kaggle/input/mp5dataset-1/assignment5_materials/.ipynb_checkpoints/MP5-checkpoint.ipynb
```

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.

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

Requirement already satisfied: gym in /opt/conda/lib/python3.10/site-packages (0.26.2)
Collecting pyvirtualdisplay
 Downloading PyVirtualDisplay-3.0-py3-none-any.whl.metadata (943 bytes)
Requirement already satisfied: numpy>=1.18.0 in /opt/conda/lib/python3.10/site-packages (from gym) (1.26.4)
Requirement already satisfied: cloudpickle>=1.2.0 in /opt/conda/lib/python3.10/site-packages (from gym) (2.2.1)
Requirement already satisfied: gym-notices>=0.0.4 in /opt/conda/lib/python3.10/site-packages (from gym) (0.0.8)
Downloading PyVirtualDisplay-3.0-py3-none-any.whl (15 kB)
Installing collected packages: pyvirtualdisplay
Successfully installed pyvirtualdisplay-3.0
Reading package lists... Done
Building dependency tree
Reading state information... Done
ffmpeg is already the newest version (7:4.2.7-0ubuntu0.1).
xvfb is already the newest version (2:1.20.13-1ubuntu1~20.04.17).
The following additional packages will be installed:
 freeglut3 libglu1-mesa libpython2-stdlib libpython2.7-minimal
 libpython2.7-stdlib python2 python2-minimal python2.7 python2.7-minimal
Suggested packages:
 python-tk python-numpy libgle3 python2-doc python2.7-doc binfmt-support
The following NEW packages will be installed:

```
freeglut3 libglu1-mesa libpython2-stdlib libpython2.7-minimal
libpython2.7-stdlib python-opengl python2 python2-minimal python2.7
python2.7-minimal
0 upgraded, 10 newly installed, 0 to remove and 65 not upgraded.
Need to get 4540 kB of archives.
After this operation, 22.7 MB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu focal-updates/universe amd64 libpython2.7-minimal
amd64 2.7.18-1~20.04.4 [335 kB]
Get:2 http://archive.ubuntu.com/ubuntu focal-updates/universe amd64 python2.7-minimal am
d64 2.7.18-1~20.04.4 [1280 kB]
Get:3 http://archive.ubuntu.com/ubuntu focal/universe amd64 python2-minimal amd64 2.7.17
-2ubuntu4 [27.5 kB]
Get:4 http://archive.ubuntu.com/ubuntu focal-updates/universe amd64 libpython2.7-stdlib
amd64 2.7.18-1~20.04.4 [1887 kB]
Get:5 http://archive.ubuntu.com/ubuntu focal-updates/universe amd64 python2.7 amd64 2.7.
18-1~20.04.4 [248 kB]
Get:6 http://archive.ubuntu.com/ubuntu focal/universe amd64 libpython2-stdlib amd64 2.7.
17-2ubuntu4 [7072 B]
Get:7 http://archive.ubuntu.com/ubuntu focal/universe amd64 python2 amd64 2.7.17-2ubuntu
4 [26.5 kB]
Get:8 http://archive.ubuntu.com/ubuntu focal/universe amd64 freeglut3 amd64 2.8.1-3 [73.
6 kB]
Get:9 http://archive.ubuntu.com/ubuntu focal/main amd64 libglu1-mesa amd64 9.0.1-1build1
[168 kB]
Get:10 http://archive.ubuntu.com/ubuntu focal/universe amd64 python-opengl all 3.1.0+dfs
g-2build1 [486 kB]
Fetched 4540 kB in 0s (11.8 MB/s)
Selecting previously unselected package libpython2.7-minimal:amd64.
(Reading database ... 113807 files and directories currently installed.)
Preparing to unpack .../0-libpython2.7-minimal_2.7.18-1~20.04.4_amd64.deb ...
Unpacking libpython2.7-minimal:amd64 (2.7.18-1~20.04.4) ...
Selecting previously unselected package python2.7-minimal.
Preparing to unpack .../1-python2.7-minimal_2.7.18-1~20.04.4_amd64.deb ...
Unpacking python2.7-minimal (2.7.18-1~20.04.4) ...
Selecting previously unselected package python2-minimal.
Preparing to unpack .../2-python2-minimal_2.7.17-2ubuntu4_amd64.deb ...
Unpacking python2-minimal (2.7.17-2ubuntu4) ...
Selecting previously unselected package libpython2.7-stdlib:amd64.
Preparing to unpack .../3-libpython2.7-stdlib_2.7.18-1~20.04.4_amd64.deb ...
Unpacking libpython2.7-stdlib:amd64 (2.7.18-1~20.04.4) ...
Selecting previously unselected package python2.7.
Preparing to unpack .../4-python2.7_2.7.18-1~20.04.4_amd64.deb ...
Unpacking python2.7 (2.7.18-1~20.04.4) ...
Selecting previously unselected package libpython2-stdlib:amd64.
Preparing to unpack .../5-libpython2-stdlib_2.7.17-2ubuntu4_amd64.deb ...
Unpacking libpython2-stdlib:amd64 (2.7.17-2ubuntu4) ...
Setting up libpython2.7-minimal:amd64 (2.7.18-1~20.04.4) ...
Setting up python2.7-minimal (2.7.18-1~20.04.4) ...
Linking and byte-compiling packages for runtime python2.7...
Setting up python2-minimal (2.7.17-2ubuntu4) ...
Selecting previously unselected package python2.
(Reading database ... 114554 files and directories currently installed.)
Preparing to unpack .../python2_2.7.17-2ubuntu4_amd64.deb ...
Unpacking python2 (2.7.17-2ubuntu4) ...
Selecting previously unselected package freeglut3:amd64.
Preparing to unpack .../freeglut3_2.8.1-3_amd64.deb ...
Unpacking freeglut3:amd64 (2.8.1-3) ...
Selecting previously unselected package libglu1-mesa:amd64.
Preparing to unpack .../libglu1-mesa_9.0.1-1build1_amd64.deb ...
Unpacking libglu1-mesa:amd64 (9.0.1-1build1) ...
Selecting previously unselected package python-opengl.
Preparing to unpack .../python-opengl_3.1.0+dfsg-2build1_all.deb ...
Unpacking python-opengl (3.1.0+dfsg-2build1) ...
Setting up freeglut3:amd64 (2.8.1-3) ...
Setting up libpython2.7-stdlib:amd64 (2.7.18-1~20.04.4) ...
Setting up libglu1-mesa:amd64 (9.0.1-1build1) ...
```

```
Setting up python2.7 (2.7.18-1~20.04.4) ...
Setting up libpython2-stdlib:amd64 (2.7.17-2ubuntu4) ...
Setting up python2 (2.7.17-2ubuntu4) ...
Setting up python-opengl (3.1.0+dfsg-2build1) ...
Processing triggers for man-db (2.9.1-1) ...
Processing triggers for mime-support (3.64ubuntu1) ...
Processing triggers for libc-bin (2.31-0ubuntu9.14) ...
```

```
In [3]: !pip3 install --upgrade setuptools --user
!pip3 install ez_setup
!pip3 install gym[atari]
!pip3 install gym[accept-rom-license]
```

```
Requirement already satisfied: setuptools in /opt/conda/lib/python3.10/site-packages (69.0.3)
Collecting setuptools
  Downloading setuptools-69.5.1-py3-none-any.whl.metadata (6.2 kB)
Downloading setuptools-69.5.1-py3-none-any.whl (894 kB)
  894.6/894.6 kB 5.1 MB/s eta 0:00:00a 0:00:01m
Installing collected packages: setuptools
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.
tensorflow-decision-forests 1.8.1 requires wurlitizer, which is not installed.
tensorflow 2.15.0 requires keras<2.16,>=2.15.0, but you have keras 3.2.1 which is incompatible.
Successfully installed setuptools-69.5.1
Collecting ez_setup
  Downloading ez_setup-0.9.tar.gz (6.6 kB)
  Preparing metadata (setup.py) ... done
Building wheels for collected packages: ez_setup
  Building wheel for ez_setup (setup.py) ... done
  Created wheel for ez_setup: filename=ez_setup-0.9-py3-none-any.whl size=10995 sha256=253ee5a38b63b7766dc6d514be8b840cfc699e341f49d49eb61f8ce08abe8922
  Stored in directory: /root/.cache/pip/wheels/7a/d6/77/8f495e85fb7df23d41c328b9ea3cf0d9e83631b20bba479293
Successfully built ez_setup
Installing collected packages: ez_setup
Successfully installed ez_setup-0.9
Requirement already satisfied: gym[atari] in /opt/conda/lib/python3.10/site-packages (0.26.2)
Requirement already satisfied: numpy>=1.18.0 in /opt/conda/lib/python3.10/site-packages (from gym[atari]) (1.26.4)
Requirement already satisfied: cloudpickle>=1.2.0 in /opt/conda/lib/python3.10/site-packages (from gym[atari]) (2.2.1)
Requirement already satisfied: gym-notices>=0.0.4 in /opt/conda/lib/python3.10/site-packages (from gym[atari]) (0.0.8)
Collecting ale-py~=0.8.0 (from gym[atari])
  Downloading ale_py-0.8.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (8.1 kB)
Requirement already satisfied: importlib-resources in /opt/conda/lib/python3.10/site-packages (from ale-py~=0.8.0->gym[atari]) (6.1.1)
Requirement already satisfied: typing-extensions in /opt/conda/lib/python3.10/site-packages (from ale-py~=0.8.0->gym[atari]) (4.9.0)
Downloading ale_py-0.8.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.7 MB)
  1.7/1.7 MB 8.6 MB/s eta 0:00:0000:0100:01
Installing collected packages: ale-py
Successfully installed ale-py-0.8.1
Requirement already satisfied: gym[accept-rom-license] in /opt/conda/lib/python3.10/site-packages (0.26.2)
Requirement already satisfied: numpy>=1.18.0 in /opt/conda/lib/python3.10/site-packages (from gym[accept-rom-license]) (1.26.4)
Requirement already satisfied: cloudpickle>=1.2.0 in /opt/conda/lib/python3.10/site-packages (from gym[accept-rom-license]) (2.2.1)
Requirement already satisfied: gym-notices>=0.0.4 in /opt/conda/lib/python3.10/site-packages (from gym[accept-rom-license]) (0.0.8)
```

```

Collecting autorom~=0.4.2 (from autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-
-license"->gym[accept-rom-license])
  Downloading AutoROM-0.4.2-py3-none-any.whl.metadata (2.8 kB)
Requirement already satisfied: click in /opt/conda/lib/python3.10/site-packages (from au
torom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[acc
ept-rom-license]) (8.1.7)
Requirement already satisfied: requests in /opt/conda/lib/python3.10/site-packages (from
autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym
[accept-rom-license]) (2.31.0)
Requirement already satisfied: tqdm in /opt/conda/lib/python3.10/site-packages (from aut
orom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[acce
pt-rom-license]) (4.66.1)
Collecting AutoROM.accept-rom-license (from autorom[accept-rom-license]~=0.4.2; extra ==
"accept-rom-license"->gym[accept-rom-license])
  Downloading AutoROM.accept-rom-license-0.6.1.tar.gz (434 kB)
    434.7/434.7 kB 2.6 MB/s eta 0:00:00a 0:00:01
  Installing build dependencies ... done
  Getting requirements to build wheel ... done
  Preparing metadata (pyproject.toml) ... done
Requirement already satisfied: charset-normalizer<4,>=2 in /opt/conda/lib/python3.10/sit
e-packages (from requests->autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra ==
"accept-rom-license"->gym[accept-rom-license]) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.10/site-packages
(from requests->autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-ro
m-license"->gym[accept-rom-license]) (3.6)
Requirement already satisfied: urllib3<3,>=1.21.1 in /opt/conda/lib/python3.10/site-pack
ages (from requests->autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accep
t-rom-license"->gym[accept-rom-license]) (1.26.18)
Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.10/site-pack
ages (from requests->autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accep
t-rom-license"->gym[accept-rom-license]) (2024.2.2)
Downloading AutoROM-0.4.2-py3-none-any.whl (16 kB)
Building wheels for collected packages: AutoROM.accept-rom-license
  Building wheel for AutoROM.accept-rom-license (pyproject.toml) ... done
  Created wheel for AutoROM.accept-rom-license: filename=AutoROM.accept_rom_license-0.6.
1-py3-none-any.whl size=446659 sha256=30e8fef2a384c814d030ca2354cd7a3fa574df81d44130a9e4
23a00106b941b7
  Stored in directory: /root/.cache/pip/wheels/6b/1b/ef/a43ff1a2f1736d5711faa1ba4c1f61be
1131b8899e6a057811
Successfully built AutoROM.accept-rom-license
Installing collected packages: AutoROM.accept-rom-license, autorom
Successfully installed AutoROM.accept-rom-license-0.6.1 autorom-0.4.2

```

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

```

In [4]: %matplotlib inline

import sys
sys.path.insert(1, '/kaggle/input/mp5dataset-1/assignment5_materials')

import gym
import torch
import pylab
import random
import numpy as np
print(np.__version__)
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

```

1.26.4

In [5]: `## Understanding the environment`

In [6]: `env = gym.make('BreakoutDeterministic-v4')
state = env.reset()`

A.L.E: Arcade Learning Environment (version 0.8.1+53f58b7)
[Powered by Stella]

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

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

`env = gym.make('BreakoutDeterministic-v4') state = env.reset()`

In [7]: `number_lives = find_max_lives(env)
state_size = env.observation_space.shape
action_size = 3 #fire, left, and right`

/opt/conda/lib/python3.10/site-packages/gym/utils/passive_env_checker.py:233: Deprecatio
nWarning: `np.bool8` is a deprecated alias for `np.bool_`. (Deprecated NumPy 1.24)
if not isinstance(terminated, (bool, np.bool8)):

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.

In [8]: `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`

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"

```
In [9]: rewards, episodes = [], []
best_eval_reward = 0
EPISODES = 2500
print(EPISODES)
for e in range(EPISODES):
    print(e)
    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 lif
        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
        # print("reached")
        next_state, reward, terminated, truncated, info = env.step(action + 1)
        done = terminated or truncated

        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
        # print("Type of action:", type(action))
        # action = action.to(device)
        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 testi
            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("/kaggle/working/breakout_dqn.png") # save graph for training
#
pylab.savefig("./save_graph/breakout_dqn.png")

# 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:", 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_evaluation_reward:
#     torch.save(agent.policy_net, "/kaggle/working/save_model/breakout_dqn.pth")
#     best_evaluation_reward = np.mean(evaluation_reward)

```

```

2500
0
episode: 0  score: 0.0  memory length: 122  epsilon: 1.0  steps: 122  lr: 0.0001
evaluation reward: 0.0
1
episode: 1  score: 1.0  memory length: 294  epsilon: 1.0  steps: 172  lr: 0.0001
evaluation reward: 0.5
2
episode: 2  score: 0.0  memory length: 416  epsilon: 1.0  steps: 122  lr: 0.0001
evaluation reward: 0.3333333333333333
3
episode: 3  score: 7.0  memory length: 698  epsilon: 1.0  steps: 282  lr: 0.0001
evaluation reward: 2.0
4
episode: 4  score: 0.0  memory length: 820  epsilon: 1.0  steps: 122  lr: 0.0001
evaluation reward: 1.6
5
episode: 5  score: 3.0  memory length: 1071  epsilon: 1.0  steps: 251  lr: 0.0001
evaluation reward: 1.8333333333333333
6
episode: 6  score: 2.0  memory length: 1269  epsilon: 1.0  steps: 198  lr: 0.0001
evaluation reward: 1.8571428571428572
7
episode: 7  score: 2.0  memory length: 1468  epsilon: 1.0  steps: 199  lr: 0.0001
evaluation reward: 1.875
8
episode: 8  score: 0.0  memory length: 1590  epsilon: 1.0  steps: 122  lr: 0.0001
evaluation reward: 1.6666666666666667
9
episode: 9  score: 5.0  memory length: 1937  epsilon: 1.0  steps: 347  lr: 0.0001
evaluation reward: 2.0
10
episode: 10  score: 1.0  memory length: 2088  epsilon: 1.0  steps: 151  lr: 0.000
evaluation reward: 1.9090909090909092
11
episode: 11  score: 3.0  memory length: 2304  epsilon: 1.0  steps: 216  lr: 0.000
evaluation reward: 2.0
12
episode: 12  score: 2.0  memory length: 2486  epsilon: 1.0  steps: 182  lr: 0.000
evaluation reward: 2.0
13
episode: 13  score: 1.0  memory length: 2654  epsilon: 1.0  steps: 168  lr: 0.000
evaluation reward: 1.9285714285714286
14
episode: 14  score: 1.0  memory length: 2825  epsilon: 1.0  steps: 171  lr: 0.000
evaluation reward: 1.8666666666666667
15
episode: 15  score: 2.0  memory length: 3022  epsilon: 1.0  steps: 197  lr: 0.000
evaluation reward: 1.875
16

```


episode: 16	score: 0.0	memory length: 3145	epsilon: 1.0	steps: 123	lr: 0.000
1	evaluation reward: 1.7647058823529411				
17					
episode: 17	score: 2.0	memory length: 3327	epsilon: 1.0	steps: 182	lr: 0.000
1	evaluation reward: 1.7777777777777777				
18					
episode: 18	score: 2.0	memory length: 3545	epsilon: 1.0	steps: 218	lr: 0.000
1	evaluation reward: 1.7894736842105263				
19					
episode: 19	score: 0.0	memory length: 3667	epsilon: 1.0	steps: 122	lr: 0.000
1	evaluation reward: 1.7				
20					
episode: 20	score: 0.0	memory length: 3790	epsilon: 1.0	steps: 123	lr: 0.000
1	evaluation reward: 1.619047619047619				
21					
episode: 21	score: 1.0	memory length: 3941	epsilon: 1.0	steps: 151	lr: 0.000
1	evaluation reward: 1.5909090909090908				
22					
episode: 22	score: 0.0	memory length: 4063	epsilon: 1.0	steps: 122	lr: 0.000
1	evaluation reward: 1.5217391304347827				
23					
episode: 23	score: 2.0	memory length: 4260	epsilon: 1.0	steps: 197	lr: 0.000
1	evaluation reward: 1.5416666666666667				
24					
episode: 24	score: 1.0	memory length: 4411	epsilon: 1.0	steps: 151	lr: 0.000
1	evaluation reward: 1.52				
25					
episode: 25	score: 1.0	memory length: 4562	epsilon: 1.0	steps: 151	lr: 0.000
1	evaluation reward: 1.5				
26					
episode: 26	score: 2.0	memory length: 4762	epsilon: 1.0	steps: 200	lr: 0.000
1	evaluation reward: 1.5185185185185186				
27					
episode: 27	score: 0.0	memory length: 4885	epsilon: 1.0	steps: 123	lr: 0.000
1	evaluation reward: 1.4642857142857142				
28					
episode: 28	score: 1.0	memory length: 5057	epsilon: 1.0	steps: 172	lr: 0.000
1	evaluation reward: 1.4482758620689655				
29					
episode: 29	score: 2.0	memory length: 5275	epsilon: 1.0	steps: 218	lr: 0.000
1	evaluation reward: 1.4666666666666666				
30					
episode: 30	score: 2.0	memory length: 5455	epsilon: 1.0	steps: 180	lr: 0.000
1	evaluation reward: 1.4838709677419355				
31					
episode: 31	score: 1.0	memory length: 5605	epsilon: 1.0	steps: 150	lr: 0.000
1	evaluation reward: 1.46875				
32					
episode: 32	score: 2.0	memory length: 5824	epsilon: 1.0	steps: 219	lr: 0.000
1	evaluation reward: 1.4848484848484849				
33					
episode: 33	score: 2.0	memory length: 6023	epsilon: 1.0	steps: 199	lr: 0.000
1	evaluation reward: 1.5				
34					
episode: 34	score: 2.0	memory length: 6241	epsilon: 1.0	steps: 218	lr: 0.000
1	evaluation reward: 1.5142857142857142				
35					
episode: 35	score: 2.0	memory length: 6462	epsilon: 1.0	steps: 221	lr: 0.000
1	evaluation reward: 1.5277777777777777				
36					
episode: 36	score: 3.0	memory length: 6688	epsilon: 1.0	steps: 226	lr: 0.000
1	evaluation reward: 1.5675675675675675				
37					
episode: 37	score: 1.0	memory length: 6839	epsilon: 1.0	steps: 151	lr: 0.000
1	evaluation reward: 1.5526315789473684				
38					

episode: 38	score: 1.0	memory length: 7007	epsilon: 1.0	steps: 168	lr: 0.000
1	evaluation reward: 1.5384615384615385				
39					
episode: 39	score: 1.0	memory length: 7157	epsilon: 1.0	steps: 150	lr: 0.000
1	evaluation reward: 1.525				
40					
episode: 40	score: 1.0	memory length: 7328	epsilon: 1.0	steps: 171	lr: 0.000
1	evaluation reward: 1.5121951219512195				
41					
episode: 41	score: 2.0	memory length: 7526	epsilon: 1.0	steps: 198	lr: 0.000
1	evaluation reward: 1.5238095238095237				
42					
episode: 42	score: 2.0	memory length: 7723	epsilon: 1.0	steps: 197	lr: 0.000
1	evaluation reward: 1.5348837209302326				
43					
episode: 43	score: 0.0	memory length: 7846	epsilon: 1.0	steps: 123	lr: 0.000
1	evaluation reward: 1.5				
44					
episode: 44	score: 0.0	memory length: 7968	epsilon: 1.0	steps: 122	lr: 0.000
1	evaluation reward: 1.4666666666666666				
45					
episode: 45	score: 2.0	memory length: 8165	epsilon: 1.0	steps: 197	lr: 0.000
1	evaluation reward: 1.4782608695652173				
46					
episode: 46	score: 1.0	memory length: 8316	epsilon: 1.0	steps: 151	lr: 0.000
1	evaluation reward: 1.4680851063829787				
47					
episode: 47	score: 3.0	memory length: 8580	epsilon: 1.0	steps: 264	lr: 0.000
1	evaluation reward: 1.5				
48					
episode: 48	score: 0.0	memory length: 8703	epsilon: 1.0	steps: 123	lr: 0.000
1	evaluation reward: 1.469387755102041				
49					
episode: 49	score: 0.0	memory length: 8825	epsilon: 1.0	steps: 122	lr: 0.000
1	evaluation reward: 1.44				
50					
episode: 50	score: 4.0	memory length: 9105	epsilon: 1.0	steps: 280	lr: 0.000
1	evaluation reward: 1.4901960784313726				
51					
episode: 51	score: 2.0	memory length: 9303	epsilon: 1.0	steps: 198	lr: 0.000
1	evaluation reward: 1.5				
52					
episode: 52	score: 1.0	memory length: 9472	epsilon: 1.0	steps: 169	lr: 0.000
1	evaluation reward: 1.490566037735849				
53					
episode: 53	score: 1.0	memory length: 9623	epsilon: 1.0	steps: 151	lr: 0.000
1	evaluation reward: 1.4814814814814814				
54					
episode: 54	score: 1.0	memory length: 9773	epsilon: 1.0	steps: 150	lr: 0.000
1	evaluation reward: 1.4727272727272727				
55					
episode: 55	score: 2.0	memory length: 9971	epsilon: 1.0	steps: 198	lr: 0.000
1	evaluation reward: 1.4821428571428572				
56					
episode: 56	score: 1.0	memory length: 10142	epsilon: 1.0	steps: 171	lr: 0.00
01	evaluation reward: 1.4736842105263157				
57					
episode: 57	score: 0.0	memory length: 10264	epsilon: 1.0	steps: 122	lr: 0.00
01	evaluation reward: 1.4482758620689655				
58					
episode: 58	score: 2.0	memory length: 10482	epsilon: 1.0	steps: 218	lr: 0.00
01	evaluation reward: 1.4576271186440677				
59					
episode: 59	score: 0.0	memory length: 10604	epsilon: 1.0	steps: 122	lr: 0.00
01	evaluation reward: 1.4333333333333333				
60					

episode: 60	score: 0.0	memory length: 10727	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward:	1.4098360655737705			
61					
episode: 61	score: 2.0	memory length: 10943	epsilon: 1.0	steps: 216	lr: 0.00
01	evaluation reward:	1.4193548387096775			
62					
episode: 62	score: 1.0	memory length: 11095	epsilon: 1.0	steps: 152	lr: 0.00
01	evaluation reward:	1.4126984126984128			
63					
episode: 63	score: 0.0	memory length: 11218	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward:	1.390625			
64					
episode: 64	score: 0.0	memory length: 11341	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward:	1.3692307692307693			
65					
episode: 65	score: 0.0	memory length: 11464	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward:	1.3484848484848484			
66					
episode: 66	score: 0.0	memory length: 11587	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward:	1.328358208955224			
67					
episode: 67	score: 5.0	memory length: 11910	epsilon: 1.0	steps: 323	lr: 0.00
01	evaluation reward:	1.3823529411764706			
68					
episode: 68	score: 2.0	memory length: 12110	epsilon: 1.0	steps: 200	lr: 0.00
01	evaluation reward:	1.391304347826087			
69					
episode: 69	score: 2.0	memory length: 12328	epsilon: 1.0	steps: 218	lr: 0.00
01	evaluation reward:	1.4			
70					
episode: 70	score: 0.0	memory length: 12451	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward:	1.380281690140845			
71					
episode: 71	score: 1.0	memory length: 12601	epsilon: 1.0	steps: 150	lr: 0.00
01	evaluation reward:	1.375			
72					
episode: 72	score: 0.0	memory length: 12724	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward:	1.356164383561644			
73					
episode: 73	score: 1.0	memory length: 12892	epsilon: 1.0	steps: 168	lr: 0.00
01	evaluation reward:	1.3513513513513513			
74					
episode: 74	score: 2.0	memory length: 13091	epsilon: 1.0	steps: 199	lr: 0.00
01	evaluation reward:	1.36			
75					
episode: 75	score: 0.0	memory length: 13213	epsilon: 1.0	steps: 122	lr: 0.00
01	evaluation reward:	1.3421052631578947			
76					
episode: 76	score: 2.0	memory length: 13431	epsilon: 1.0	steps: 218	lr: 0.00
01	evaluation reward:	1.3506493506493507			
77					
episode: 77	score: 2.0	memory length: 13613	epsilon: 1.0	steps: 182	lr: 0.00
01	evaluation reward:	1.358974358974359			
78					
episode: 78	score: 2.0	memory length: 13811	epsilon: 1.0	steps: 198	lr: 0.00
01	evaluation reward:	1.3670886075949367			
79					
episode: 79	score: 1.0	memory length: 13980	epsilon: 1.0	steps: 169	lr: 0.00
01	evaluation reward:	1.3625			
80					
episode: 80	score: 0.0	memory length: 14102	epsilon: 1.0	steps: 122	lr: 0.00
01	evaluation reward:	1.345679012345679			
81					
episode: 81	score: 3.0	memory length: 14332	epsilon: 1.0	steps: 230	lr: 0.00
01	evaluation reward:	1.3658536585365855			
82					

episode: 82	score: 1.0	memory length: 14483	epsilon: 1.0	steps: 151	lr: 0.00
01	evaluation reward: 1.3614457831325302				
83					
episode: 83	score: 2.0	memory length: 14682	epsilon: 1.0	steps: 199	lr: 0.00
01	evaluation reward: 1.369047619047619				
84					
episode: 84	score: 0.0	memory length: 14805	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward: 1.3529411764705883				
85					
episode: 85	score: 0.0	memory length: 14928	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward: 1.3372093023255813				
86					
episode: 86	score: 2.0	memory length: 15146	epsilon: 1.0	steps: 218	lr: 0.00
01	evaluation reward: 1.3448275862068966				
87					
episode: 87	score: 0.0	memory length: 15269	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward: 1.3295454545454546				
88					
episode: 88	score: 3.0	memory length: 15495	epsilon: 1.0	steps: 226	lr: 0.00
01	evaluation reward: 1.348314606741573				
89					
episode: 89	score: 2.0	memory length: 15692	epsilon: 1.0	steps: 197	lr: 0.00
01	evaluation reward: 1.3555555555555556				
90					
episode: 90	score: 3.0	memory length: 15918	epsilon: 1.0	steps: 226	lr: 0.00
01	evaluation reward: 1.3736263736263736				
91					
episode: 91	score: 4.0	memory length: 16179	epsilon: 1.0	steps: 261	lr: 0.00
01	evaluation reward: 1.4021739130434783				
92					
episode: 92	score: 2.0	memory length: 16377	epsilon: 1.0	steps: 198	lr: 0.00
01	evaluation reward: 1.4086021505376345				
93					
episode: 93	score: 3.0	memory length: 16625	epsilon: 1.0	steps: 248	lr: 0.00
01	evaluation reward: 1.425531914893617				
94					
episode: 94	score: 2.0	memory length: 16809	epsilon: 1.0	steps: 184	lr: 0.00
01	evaluation reward: 1.431578947368421				
95					
episode: 95	score: 1.0	memory length: 16981	epsilon: 1.0	steps: 172	lr: 0.00
01	evaluation reward: 1.4270833333333333				
96					
episode: 96	score: 0.0	memory length: 17104	epsilon: 1.0	steps: 123	lr: 0.00
01	evaluation reward: 1.4123711340206186				
97					
episode: 97	score: 2.0	memory length: 17320	epsilon: 1.0	steps: 216	lr: 0.00
01	evaluation reward: 1.4183673469387754				
98					
episode: 98	score: 1.0	memory length: 17492	epsilon: 1.0	steps: 172	lr: 0.00
01	evaluation reward: 1.4141414141414141				
99					
episode: 99	score: 2.0	memory length: 17710	epsilon: 1.0	steps: 218	lr: 0.00
01	evaluation reward: 1.42				
100					
episode: 100	score: 2.0	memory length: 17908	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.44				
101					
episode: 101	score: 2.0	memory length: 18106	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.45				
102					
episode: 102	score: 1.0	memory length: 18278	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.46				
103					
episode: 103	score: 2.0	memory length: 18495	epsilon: 1.0	steps: 217	lr: 0.0
001	evaluation reward: 1.41				
104					

episode: 104	score: 2.0	memory length: 18715	epsilon: 1.0	steps: 220	lr: 0.0
001	evaluation reward: 1.43				
105					
episode: 105	score: 3.0	memory length: 18980	epsilon: 1.0	steps: 265	lr: 0.0
001	evaluation reward: 1.43				
106					
episode: 106	score: 2.0	memory length: 19164	epsilon: 1.0	steps: 184	lr: 0.0
001	evaluation reward: 1.43				
107					
episode: 107	score: 0.0	memory length: 19286	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.41				
108					
episode: 108	score: 1.0	memory length: 19456	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.42				
109					
episode: 109	score: 1.0	memory length: 19607	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.38				
110					
episode: 110	score: 4.0	memory length: 19905	epsilon: 1.0	steps: 298	lr: 0.0
001	evaluation reward: 1.41				
111					
episode: 111	score: 2.0	memory length: 20103	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.4				
112					
episode: 112	score: 2.0	memory length: 20301	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.4				
113					
episode: 113	score: 1.0	memory length: 20469	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.4				
114					
episode: 114	score: 2.0	memory length: 20690	epsilon: 1.0	steps: 221	lr: 0.0
001	evaluation reward: 1.41				
115					
episode: 115	score: 0.0	memory length: 20813	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.39				
116					
episode: 116	score: 3.0	memory length: 21061	epsilon: 1.0	steps: 248	lr: 0.0
001	evaluation reward: 1.42				
117					
episode: 117	score: 1.0	memory length: 21230	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.41				
118					
episode: 118	score: 0.0	memory length: 21353	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.39				
119					
episode: 119	score: 1.0	memory length: 21524	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.4				
120					
episode: 120	score: 2.0	memory length: 21740	epsilon: 1.0	steps: 216	lr: 0.0
001	evaluation reward: 1.42				
121					
episode: 121	score: 0.0	memory length: 21863	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
122					
episode: 122	score: 1.0	memory length: 22014	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.42				
123					
episode: 123	score: 2.0	memory length: 22235	epsilon: 1.0	steps: 221	lr: 0.0
001	evaluation reward: 1.42				
124					
episode: 124	score: 0.0	memory length: 22358	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
125					
episode: 125	score: 2.0	memory length: 22540	epsilon: 1.0	steps: 182	lr: 0.0
001	evaluation reward: 1.42				
126					

episode: 126	score: 1.0	memory length: 22691	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.41				
127					
episode: 127	score: 1.0	memory length: 22841	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.42				
128					
episode: 128	score: 0.0	memory length: 22964	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
129					
episode: 129	score: 1.0	memory length: 23114	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.4				
130					
episode: 130	score: 0.0	memory length: 23237	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.38				
131					
episode: 131	score: 1.0	memory length: 23388	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.38				
132					
episode: 132	score: 1.0	memory length: 23558	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.37				
133					
episode: 133	score: 1.0	memory length: 23730	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.36				
134					
episode: 134	score: 1.0	memory length: 23881	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.35				
135					
episode: 135	score: 1.0	memory length: 24032	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.34				
136					
episode: 136	score: 2.0	memory length: 24248	epsilon: 1.0	steps: 216	lr: 0.0
001	evaluation reward: 1.33				
137					
episode: 137	score: 2.0	memory length: 24446	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.34				
138					
episode: 138	score: 1.0	memory length: 24615	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.34				
139					
episode: 139	score: 3.0	memory length: 24844	epsilon: 1.0	steps: 229	lr: 0.0
001	evaluation reward: 1.36				
140					
episode: 140	score: 4.0	memory length: 25138	epsilon: 1.0	steps: 294	lr: 0.0
001	evaluation reward: 1.39				
141					
episode: 141	score: 0.0	memory length: 25261	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.37				
142					
episode: 142	score: 1.0	memory length: 25432	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.36				
143					
episode: 143	score: 1.0	memory length: 25603	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.37				
144					
episode: 144	score: 5.0	memory length: 25902	epsilon: 1.0	steps: 299	lr: 0.0
001	evaluation reward: 1.42				
145					
episode: 145	score: 1.0	memory length: 26072	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.41				
146					
episode: 146	score: 0.0	memory length: 26195	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.4				
147					
episode: 147	score: 1.0	memory length: 26364	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.38				
148					

episode: 148	score: 1.0	memory length: 26515	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.39				
149					
episode: 149	score: 3.0	memory length: 26761	epsilon: 1.0	steps: 246	lr: 0.0
001	evaluation reward: 1.42				
150					
episode: 150	score: 2.0	memory length: 26944	epsilon: 1.0	steps: 183	lr: 0.0
001	evaluation reward: 1.4				
151					
episode: 151	score: 1.0	memory length: 27113	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.39				
152					
episode: 152	score: 2.0	memory length: 27330	epsilon: 1.0	steps: 217	lr: 0.0
001	evaluation reward: 1.4				
153					
episode: 153	score: 0.0	memory length: 27453	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.39				
154					
episode: 154	score: 2.0	memory length: 27654	epsilon: 1.0	steps: 201	lr: 0.0
001	evaluation reward: 1.4				
155					
episode: 155	score: 2.0	memory length: 27852	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.4				
156					
episode: 156	score: 4.0	memory length: 28145	epsilon: 1.0	steps: 293	lr: 0.0
001	evaluation reward: 1.43				
157					
episode: 157	score: 0.0	memory length: 28268	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.43				
158					
episode: 158	score: 2.0	memory length: 28466	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.43				
159					
episode: 159	score: 0.0	memory length: 28589	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.43				
160					
episode: 160	score: 0.0	memory length: 28712	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.43				
161					
episode: 161	score: 0.0	memory length: 28835	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
162					
episode: 162	score: 1.0	memory length: 29007	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.41				
163					
episode: 163	score: 0.0	memory length: 29129	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.41				
164					
episode: 164	score: 0.0	memory length: 29251	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.41				
165					
episode: 165	score: 2.0	memory length: 29448	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.43				
166					
episode: 166	score: 3.0	memory length: 29693	epsilon: 1.0	steps: 245	lr: 0.0
001	evaluation reward: 1.46				
167					
episode: 167	score: 3.0	memory length: 29919	epsilon: 1.0	steps: 226	lr: 0.0
001	evaluation reward: 1.44				
168					
episode: 168	score: 2.0	memory length: 30117	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.44				
169					
episode: 169	score: 0.0	memory length: 30239	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.42				
170					

episode: 170	score: 0.0	memory length: 30362	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.42				
171					
episode: 171	score: 1.0	memory length: 30532	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.42				
172					
episode: 172	score: 2.0	memory length: 30729	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.44				
173					
episode: 173	score: 2.0	memory length: 30926	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.45				
174					
episode: 174	score: 0.0	memory length: 31048	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.43				
175					
episode: 175	score: 1.0	memory length: 31198	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.44				
176					
episode: 176	score: 2.0	memory length: 31400	epsilon: 1.0	steps: 202	lr: 0.0
001	evaluation reward: 1.44				
177					
episode: 177	score: 0.0	memory length: 31523	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.42				
178					
episode: 178	score: 1.0	memory length: 31674	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.41				
179					
episode: 179	score: 2.0	memory length: 31872	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.42				
180					
episode: 180	score: 2.0	memory length: 32070	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.44				
181					
episode: 181	score: 2.0	memory length: 32268	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.43				
182					
episode: 182	score: 1.0	memory length: 32439	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.43				
183					
episode: 183	score: 0.0	memory length: 32562	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
184					
episode: 184	score: 0.0	memory length: 32685	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
185					
episode: 185	score: 1.0	memory length: 32835	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.42				
186					
episode: 186	score: 0.0	memory length: 32958	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.4				
187					
episode: 187	score: 1.0	memory length: 33109	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.41				
188					
episode: 188	score: 1.0	memory length: 33261	epsilon: 1.0	steps: 152	lr: 0.0
001	evaluation reward: 1.39				
189					
episode: 189	score: 1.0	memory length: 33412	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.38				
190					
episode: 190	score: 4.0	memory length: 33689	epsilon: 1.0	steps: 277	lr: 0.0
001	evaluation reward: 1.39				
191					
episode: 191	score: 2.0	memory length: 33887	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.37				
192					

episode: 192	score: 1.0	memory length: 34057	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.36				
193					
episode: 193	score: 2.0	memory length: 34279	epsilon: 1.0	steps: 222	lr: 0.0
001	evaluation reward: 1.35				
194					
episode: 194	score: 3.0	memory length: 34524	epsilon: 1.0	steps: 245	lr: 0.0
001	evaluation reward: 1.36				
195					
episode: 195	score: 2.0	memory length: 34722	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.37				
196					
episode: 196	score: 3.0	memory length: 34967	epsilon: 1.0	steps: 245	lr: 0.0
001	evaluation reward: 1.4				
197					
episode: 197	score: 4.0	memory length: 35227	epsilon: 1.0	steps: 260	lr: 0.0
001	evaluation reward: 1.42				
198					
episode: 198	score: 2.0	memory length: 35411	epsilon: 1.0	steps: 184	lr: 0.0
001	evaluation reward: 1.43				
199					
episode: 199	score: 2.0	memory length: 35609	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.43				
200					
episode: 200	score: 3.0	memory length: 35855	epsilon: 1.0	steps: 246	lr: 0.0
001	evaluation reward: 1.44				
201					
episode: 201	score: 1.0	memory length: 36006	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.43				
202					
episode: 202	score: 0.0	memory length: 36129	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.42				
203					
episode: 203	score: 0.0	memory length: 36252	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.4				
204					
episode: 204	score: 0.0	memory length: 36375	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.38				
205					
episode: 205	score: 0.0	memory length: 36497	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.35				
206					
episode: 206	score: 1.0	memory length: 36665	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.34				
207					
episode: 207	score: 3.0	memory length: 36930	epsilon: 1.0	steps: 265	lr: 0.0
001	evaluation reward: 1.37				
208					
episode: 208	score: 1.0	memory length: 37080	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.37				
209					
episode: 209	score: 0.0	memory length: 37202	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.36				
210					
episode: 210	score: 3.0	memory length: 37430	epsilon: 1.0	steps: 228	lr: 0.0
001	evaluation reward: 1.35				
211					
episode: 211	score: 1.0	memory length: 37601	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.34				
212					
episode: 212	score: 0.0	memory length: 37724	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.32				
213					
episode: 213	score: 0.0	memory length: 37847	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.31				
214					

episode: 214	score: 0.0	memory length: 37970	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.29				
215					
episode: 215	score: 1.0	memory length: 38120	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.3				
216					
episode: 216	score: 2.0	memory length: 38319	epsilon: 1.0	steps: 199	lr: 0.0
001	evaluation reward: 1.29				
217					
episode: 217	score: 2.0	memory length: 38516	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.3				
218					
episode: 218	score: 4.0	memory length: 38816	epsilon: 1.0	steps: 300	lr: 0.0
001	evaluation reward: 1.34				
219					
episode: 219	score: 0.0	memory length: 38938	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.33				
220					
episode: 220	score: 2.0	memory length: 39136	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.33				
221					
episode: 221	score: 0.0	memory length: 39259	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.33				
222					
episode: 222	score: 2.0	memory length: 39476	epsilon: 1.0	steps: 217	lr: 0.0
001	evaluation reward: 1.34				
223					
episode: 223	score: 1.0	memory length: 39645	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.33				
224					
episode: 224	score: 1.0	memory length: 39795	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.34				
225					
episode: 225	score: 3.0	memory length: 40040	epsilon: 1.0	steps: 245	lr: 0.0
001	evaluation reward: 1.35				
226					
episode: 226	score: 0.0	memory length: 40163	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.34				
227					
episode: 227	score: 1.0	memory length: 40333	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.34				
228					
episode: 228	score: 0.0	memory length: 40455	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.34				
229					
episode: 229	score: 0.0	memory length: 40577	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.33				
230					
episode: 230	score: 0.0	memory length: 40700	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.33				
231					
episode: 231	score: 1.0	memory length: 40851	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.33				
232					
episode: 232	score: 2.0	memory length: 41052	epsilon: 1.0	steps: 201	lr: 0.0
001	evaluation reward: 1.34				
233					
episode: 233	score: 4.0	memory length: 41365	epsilon: 1.0	steps: 313	lr: 0.0
001	evaluation reward: 1.37				
234					
episode: 234	score: 3.0	memory length: 41632	epsilon: 1.0	steps: 267	lr: 0.0
001	evaluation reward: 1.39				
235					
episode: 235	score: 0.0	memory length: 41754	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.38				
236					

episode: 236	score: 3.0	memory length: 42002	epsilon: 1.0	steps: 248	lr: 0.0
001	evaluation reward: 1.39				
237					
episode: 237	score: 0.0	memory length: 42125	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.37				
238					
episode: 238	score: 3.0	memory length: 42390	epsilon: 1.0	steps: 265	lr: 0.0
001	evaluation reward: 1.39				
239					
episode: 239	score: 0.0	memory length: 42513	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.36				
240					
episode: 240	score: 2.0	memory length: 42732	epsilon: 1.0	steps: 219	lr: 0.0
001	evaluation reward: 1.34				
241					
episode: 241	score: 2.0	memory length: 42930	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.36				
242					
episode: 242	score: 0.0	memory length: 43052	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.35				
243					
episode: 243	score: 1.0	memory length: 43220	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.35				
244					
episode: 244	score: 2.0	memory length: 43417	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.32				
245					
episode: 245	score: 2.0	memory length: 43615	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.33				
246					
episode: 246	score: 3.0	memory length: 43843	epsilon: 1.0	steps: 228	lr: 0.0
001	evaluation reward: 1.36				
247					
episode: 247	score: 0.0	memory length: 43966	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.35				
248					
episode: 248	score: 3.0	memory length: 44229	epsilon: 1.0	steps: 263	lr: 0.0
001	evaluation reward: 1.37				
249					
episode: 249	score: 0.0	memory length: 44351	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.34				
250					
episode: 250	score: 1.0	memory length: 44523	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.33				
251					
episode: 251	score: 1.0	memory length: 44692	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.33				
252					
episode: 252	score: 0.0	memory length: 44815	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.31				
253					
episode: 253	score: 1.0	memory length: 44984	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.32				
254					
episode: 254	score: 2.0	memory length: 45202	epsilon: 1.0	steps: 218	lr: 0.0
001	evaluation reward: 1.32				
255					
episode: 255	score: 1.0	memory length: 45371	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.31				
256					
episode: 256	score: 1.0	memory length: 45522	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.28				
257					
episode: 257	score: 4.0	memory length: 45780	epsilon: 1.0	steps: 258	lr: 0.0
001	evaluation reward: 1.32				
258					

episode: 258	score: 4.0	memory length: 46055	epsilon: 1.0	steps: 275	lr: 0.0
001	evaluation reward: 1.34				
259					
episode: 259	score: 2.0	memory length: 46252	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.36				
260					
episode: 260	score: 0.0	memory length: 46375	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.36				
261					
episode: 261	score: 1.0	memory length: 46544	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.37				
262					
episode: 262	score: 2.0	memory length: 46745	epsilon: 1.0	steps: 201	lr: 0.0
001	evaluation reward: 1.38				
263					
episode: 263	score: 1.0	memory length: 46914	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.39				
264					
episode: 264	score: 2.0	memory length: 47135	epsilon: 1.0	steps: 221	lr: 0.0
001	evaluation reward: 1.41				
265					
episode: 265	score: 1.0	memory length: 47305	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.4				
266					
episode: 266	score: 0.0	memory length: 47428	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.37				
267					
episode: 267	score: 2.0	memory length: 47625	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.36				
268					
episode: 268	score: 0.0	memory length: 47748	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.34				
269					
episode: 269	score: 5.0	memory length: 48059	epsilon: 1.0	steps: 311	lr: 0.0
001	evaluation reward: 1.39				
270					
episode: 270	score: 2.0	memory length: 48256	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.41				
271					
episode: 271	score: 0.0	memory length: 48379	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.4				
272					
episode: 272	score: 1.0	memory length: 48530	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.39				
273					
episode: 273	score: 0.0	memory length: 48653	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.37				
274					
episode: 274	score: 0.0	memory length: 48775	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.37				
275					
episode: 275	score: 1.0	memory length: 48943	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.37				
276					
episode: 276	score: 1.0	memory length: 49095	epsilon: 1.0	steps: 152	lr: 0.0
001	evaluation reward: 1.36				
277					
episode: 277	score: 3.0	memory length: 49343	epsilon: 1.0	steps: 248	lr: 0.0
001	evaluation reward: 1.39				
278					
episode: 278	score: 1.0	memory length: 49512	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.39				
279					
episode: 279	score: 5.0	memory length: 49857	epsilon: 1.0	steps: 345	lr: 0.0
001	evaluation reward: 1.42				
280					

episode: 280	score: 1.0	memory length: 50025	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.41				
281					
episode: 281	score: 1.0	memory length: 50197	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.4				
282					
episode: 282	score: 0.0	memory length: 50320	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.39				
283					
episode: 283	score: 2.0	memory length: 50539	epsilon: 1.0	steps: 219	lr: 0.0
001	evaluation reward: 1.41				
284					
episode: 284	score: 2.0	memory length: 50737	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.43				
285					
episode: 285	score: 0.0	memory length: 50860	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.42				
286					
episode: 286	score: 0.0	memory length: 50983	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.42				
287					
episode: 287	score: 2.0	memory length: 51164	epsilon: 1.0	steps: 181	lr: 0.0
001	evaluation reward: 1.43				
288					
episode: 288	score: 3.0	memory length: 51397	epsilon: 1.0	steps: 233	lr: 0.0
001	evaluation reward: 1.45				
289					
episode: 289	score: 0.0	memory length: 51520	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.44				
290					
episode: 290	score: 3.0	memory length: 51749	epsilon: 1.0	steps: 229	lr: 0.0
001	evaluation reward: 1.43				
291					
episode: 291	score: 0.0	memory length: 51872	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
292					
episode: 292	score: 2.0	memory length: 52069	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.42				
293					
episode: 293	score: 2.0	memory length: 52267	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.42				
294					
episode: 294	score: 0.0	memory length: 52389	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.39				
295					
episode: 295	score: 3.0	memory length: 52619	epsilon: 1.0	steps: 230	lr: 0.0
001	evaluation reward: 1.4				
296					
episode: 296	score: 1.0	memory length: 52791	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.38				
297					
episode: 297	score: 4.0	memory length: 53088	epsilon: 1.0	steps: 297	lr: 0.0
001	evaluation reward: 1.38				
298					
episode: 298	score: 1.0	memory length: 53239	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.37				
299					
episode: 299	score: 1.0	memory length: 53407	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.36				
300					
episode: 300	score: 0.0	memory length: 53530	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.33				
301					
episode: 301	score: 1.0	memory length: 53681	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.33				
302					

episode: 302	score: 2.0	memory length: 53879	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.35				
303					
episode: 303	score: 2.0	memory length: 54080	epsilon: 1.0	steps: 201	lr: 0.0
001	evaluation reward: 1.37				
304					
episode: 304	score: 1.0	memory length: 54249	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.38				
305					
episode: 305	score: 0.0	memory length: 54371	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.38				
306					
episode: 306	score: 2.0	memory length: 54553	epsilon: 1.0	steps: 182	lr: 0.0
001	evaluation reward: 1.39				
307					
episode: 307	score: 1.0	memory length: 54723	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.37				
308					
episode: 308	score: 3.0	memory length: 54949	epsilon: 1.0	steps: 226	lr: 0.0
001	evaluation reward: 1.39				
309					
episode: 309	score: 2.0	memory length: 55147	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.41				
310					
episode: 310	score: 2.0	memory length: 55363	epsilon: 1.0	steps: 216	lr: 0.0
001	evaluation reward: 1.4				
311					
episode: 311	score: 0.0	memory length: 55486	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.39				
312					
episode: 312	score: 2.0	memory length: 55666	epsilon: 1.0	steps: 180	lr: 0.0
001	evaluation reward: 1.41				
313					
episode: 313	score: 0.0	memory length: 55789	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
314					
episode: 314	score: 0.0	memory length: 55912	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
315					
episode: 315	score: 1.0	memory length: 56063	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.41				
316					
episode: 316	score: 2.0	memory length: 56261	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.41				
317					
episode: 317	score: 2.0	memory length: 56458	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.41				
318					
episode: 318	score: 1.0	memory length: 56627	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.38				
319					
episode: 319	score: 0.0	memory length: 56750	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.38				
320					
episode: 320	score: 0.0	memory length: 56873	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.36				
321					
episode: 321	score: 0.0	memory length: 56996	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.36				
322					
episode: 322	score: 1.0	memory length: 57168	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.35				
323					
episode: 323	score: 3.0	memory length: 57396	epsilon: 1.0	steps: 228	lr: 0.0
001	evaluation reward: 1.37				
324					

episode: 324	score: 2.0	memory length: 57615	epsilon: 1.0	steps: 219	lr: 0.0
001	evaluation reward: 1.38				
325					
episode: 325	score: 1.0	memory length: 57786	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.36				
326					
episode: 326	score: 1.0	memory length: 57937	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.37				
327					
episode: 327	score: 1.0	memory length: 58087	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.37				
328					
episode: 328	score: 1.0	memory length: 58238	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.38				
329					
episode: 329	score: 0.0	memory length: 58361	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.38				
330					
episode: 330	score: 3.0	memory length: 58610	epsilon: 1.0	steps: 249	lr: 0.0
001	evaluation reward: 1.41				
331					
episode: 331	score: 2.0	memory length: 58808	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.42				
332					
episode: 332	score: 1.0	memory length: 58977	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.41				
333					
episode: 333	score: 2.0	memory length: 59193	epsilon: 1.0	steps: 216	lr: 0.0
001	evaluation reward: 1.39				
334					
episode: 334	score: 2.0	memory length: 59391	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.38				
335					
episode: 335	score: 0.0	memory length: 59514	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.38				
336					
episode: 336	score: 0.0	memory length: 59636	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.35				
337					
episode: 337	score: 6.0	memory length: 60011	epsilon: 1.0	steps: 375	lr: 0.0
001	evaluation reward: 1.41				
338					
episode: 338	score: 1.0	memory length: 60162	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.39				
339					
episode: 339	score: 4.0	memory length: 60437	epsilon: 1.0	steps: 275	lr: 0.0
001	evaluation reward: 1.43				
340					
episode: 340	score: 1.0	memory length: 60587	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.42				
341					
episode: 341	score: 0.0	memory length: 60709	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.4				
342					
episode: 342	score: 2.0	memory length: 60907	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.42				
343					
episode: 343	score: 0.0	memory length: 61030	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
344					
episode: 344	score: 2.0	memory length: 61228	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.41				
345					
episode: 345	score: 1.0	memory length: 61399	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.4				
346					

episode: 346	score: 1.0	memory length: 61570	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.38				
347					
episode: 347	score: 0.0	memory length: 61692	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.38				
348					
episode: 348	score: 3.0	memory length: 61918	epsilon: 1.0	steps: 226	lr: 0.0
001	evaluation reward: 1.38				
349					
episode: 349	score: 0.0	memory length: 62040	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.38				
350					
episode: 350	score: 2.0	memory length: 62221	epsilon: 1.0	steps: 181	lr: 0.0
001	evaluation reward: 1.39				
351					
episode: 351	score: 1.0	memory length: 62392	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.39				
352					
episode: 352	score: 2.0	memory length: 62609	epsilon: 1.0	steps: 217	lr: 0.0
001	evaluation reward: 1.41				
353					
episode: 353	score: 1.0	memory length: 62778	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.41				
354					
episode: 354	score: 2.0	memory length: 62976	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.41				
355					
episode: 355	score: 3.0	memory length: 63204	epsilon: 1.0	steps: 228	lr: 0.0
001	evaluation reward: 1.43				
356					
episode: 356	score: 2.0	memory length: 63421	epsilon: 1.0	steps: 217	lr: 0.0
001	evaluation reward: 1.44				
357					
episode: 357	score: 2.0	memory length: 63619	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.42				
358					
episode: 358	score: 0.0	memory length: 63742	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.38				
359					
episode: 359	score: 3.0	memory length: 64009	epsilon: 1.0	steps: 267	lr: 0.0
001	evaluation reward: 1.39				
360					
episode: 360	score: 1.0	memory length: 64160	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.4				
361					
episode: 361	score: 3.0	memory length: 64408	epsilon: 1.0	steps: 248	lr: 0.0
001	evaluation reward: 1.42				
362					
episode: 362	score: 2.0	memory length: 64609	epsilon: 1.0	steps: 201	lr: 0.0
001	evaluation reward: 1.42				
363					
episode: 363	score: 2.0	memory length: 64827	epsilon: 1.0	steps: 218	lr: 0.0
001	evaluation reward: 1.43				
364					
episode: 364	score: 2.0	memory length: 65028	epsilon: 1.0	steps: 201	lr: 0.0
001	evaluation reward: 1.43				
365					
episode: 365	score: 2.0	memory length: 65228	epsilon: 1.0	steps: 200	lr: 0.0
001	evaluation reward: 1.44				
366					
episode: 366	score: 0.0	memory length: 65350	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.44				
367					
episode: 367	score: 6.0	memory length: 65725	epsilon: 1.0	steps: 375	lr: 0.0
001	evaluation reward: 1.48				
368					

episode: 368	score: 4.0	memory length: 65988	epsilon: 1.0	steps: 263	lr: 0.0
001	evaluation reward: 1.52				
369					
episode: 369	score: 2.0	memory length: 66189	epsilon: 1.0	steps: 201	lr: 0.0
001	evaluation reward: 1.49				
370					
episode: 370	score: 1.0	memory length: 66357	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.48				
371					
episode: 371	score: 2.0	memory length: 66572	epsilon: 1.0	steps: 215	lr: 0.0
001	evaluation reward: 1.5				
372					
episode: 372	score: 2.0	memory length: 66790	epsilon: 1.0	steps: 218	lr: 0.0
001	evaluation reward: 1.51				
373					
episode: 373	score: 1.0	memory length: 66941	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.52				
374					
episode: 374	score: 1.0	memory length: 67092	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.53				
375					
episode: 375	score: 2.0	memory length: 67310	epsilon: 1.0	steps: 218	lr: 0.0
001	evaluation reward: 1.54				
376					
episode: 376	score: 0.0	memory length: 67433	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.53				
377					
episode: 377	score: 1.0	memory length: 67586	epsilon: 1.0	steps: 153	lr: 0.0
001	evaluation reward: 1.51				
378					
episode: 378	score: 2.0	memory length: 67767	epsilon: 1.0	steps: 181	lr: 0.0
001	evaluation reward: 1.52				
379					
episode: 379	score: 1.0	memory length: 67917	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.48				
380					
episode: 380	score: 2.0	memory length: 68115	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.49				
381					
episode: 381	score: 3.0	memory length: 68342	epsilon: 1.0	steps: 227	lr: 0.0
001	evaluation reward: 1.51				
382					
episode: 382	score: 0.0	memory length: 68465	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.51				
383					
episode: 383	score: 3.0	memory length: 68712	epsilon: 1.0	steps: 247	lr: 0.0
001	evaluation reward: 1.52				
384					
episode: 384	score: 1.0	memory length: 68864	epsilon: 1.0	steps: 152	lr: 0.0
001	evaluation reward: 1.51				
385					
episode: 385	score: 1.0	memory length: 69032	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.52				
386					
episode: 386	score: 0.0	memory length: 69155	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.52				
387					
episode: 387	score: 2.0	memory length: 69375	epsilon: 1.0	steps: 220	lr: 0.0
001	evaluation reward: 1.52				
388					
episode: 388	score: 0.0	memory length: 69498	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.49				
389					
episode: 389	score: 3.0	memory length: 69744	epsilon: 1.0	steps: 246	lr: 0.0
001	evaluation reward: 1.52				
390					

episode: 390	score: 0.0	memory length: 69867	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.49				
391					
episode: 391	score: 1.0	memory length: 70035	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.5				
392					
episode: 392	score: 1.0	memory length: 70204	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.49				
393					
episode: 393	score: 2.0	memory length: 70422	epsilon: 1.0	steps: 218	lr: 0.0
001	evaluation reward: 1.49				
394					
episode: 394	score: 3.0	memory length: 70651	epsilon: 1.0	steps: 229	lr: 0.0
001	evaluation reward: 1.52				
395					
episode: 395	score: 1.0	memory length: 70821	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.5				
396					
episode: 396	score: 1.0	memory length: 70971	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.5				
397					
episode: 397	score: 4.0	memory length: 71248	epsilon: 1.0	steps: 277	lr: 0.0
001	evaluation reward: 1.5				
398					
episode: 398	score: 2.0	memory length: 71446	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.51				
399					
episode: 399	score: 0.0	memory length: 71569	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.5				
400					
episode: 400	score: 0.0	memory length: 71691	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.5				
401					
episode: 401	score: 0.0	memory length: 71813	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.49				
402					
episode: 402	score: 3.0	memory length: 72056	epsilon: 1.0	steps: 243	lr: 0.0
001	evaluation reward: 1.5				
403					
episode: 403	score: 0.0	memory length: 72179	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.48				
404					
episode: 404	score: 6.0	memory length: 72530	epsilon: 1.0	steps: 351	lr: 0.0
001	evaluation reward: 1.53				
405					
episode: 405	score: 1.0	memory length: 72681	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.54				
406					
episode: 406	score: 0.0	memory length: 72803	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.52				
407					
episode: 407	score: 1.0	memory length: 72974	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.52				
408					
episode: 408	score: 3.0	memory length: 73221	epsilon: 1.0	steps: 247	lr: 0.0
001	evaluation reward: 1.52				
409					
episode: 409	score: 0.0	memory length: 73343	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.5				
410					
episode: 410	score: 1.0	memory length: 73493	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.49				
411					
episode: 411	score: 1.0	memory length: 73643	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.5				
412					

episode: 412	score: 0.0	memory length: 73765	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.48				
413					
episode: 413	score: 1.0	memory length: 73933	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.49				
414					
episode: 414	score: 1.0	memory length: 74084	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.5				
415					
episode: 415	score: 0.0	memory length: 74207	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.49				
416					
episode: 416	score: 2.0	memory length: 74405	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.49				
417					
episode: 417	score: 0.0	memory length: 74527	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.47				
418					
episode: 418	score: 0.0	memory length: 74649	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.46				
419					
episode: 419	score: 0.0	memory length: 74772	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.46				
420					
episode: 420	score: 2.0	memory length: 74970	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.48				
421					
episode: 421	score: 0.0	memory length: 75093	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.48				
422					
episode: 422	score: 1.0	memory length: 75244	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.48				
423					
episode: 423	score: 1.0	memory length: 75395	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.46				
424					
episode: 424	score: 0.0	memory length: 75517	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.44				
425					
episode: 425	score: 2.0	memory length: 75736	epsilon: 1.0	steps: 219	lr: 0.0
001	evaluation reward: 1.45				
426					
episode: 426	score: 3.0	memory length: 75982	epsilon: 1.0	steps: 246	lr: 0.0
001	evaluation reward: 1.47				
427					
episode: 427	score: 2.0	memory length: 76199	epsilon: 1.0	steps: 217	lr: 0.0
001	evaluation reward: 1.48				
428					
episode: 428	score: 1.0	memory length: 76371	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.48				
429					
episode: 429	score: 2.0	memory length: 76588	epsilon: 1.0	steps: 217	lr: 0.0
001	evaluation reward: 1.5				
430					
episode: 430	score: 2.0	memory length: 76786	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.49				
431					
episode: 431	score: 4.0	memory length: 77100	epsilon: 1.0	steps: 314	lr: 0.0
001	evaluation reward: 1.51				
432					
episode: 432	score: 2.0	memory length: 77298	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.52				
433					
episode: 433	score: 4.0	memory length: 77594	epsilon: 1.0	steps: 296	lr: 0.0
001	evaluation reward: 1.54				
434					

episode: 434	score: 2.0	memory length: 77793	epsilon: 1.0	steps: 199	lr: 0.0
001	evaluation reward: 1.54				
435					
episode: 435	score: 2.0	memory length: 78014	epsilon: 1.0	steps: 221	lr: 0.0
001	evaluation reward: 1.56				
436					
episode: 436	score: 0.0	memory length: 78136	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.56				
437					
episode: 437	score: 1.0	memory length: 78288	epsilon: 1.0	steps: 152	lr: 0.0
001	evaluation reward: 1.51				
438					
episode: 438	score: 1.0	memory length: 78456	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.51				
439					
episode: 439	score: 0.0	memory length: 78579	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.47				
440					
episode: 440	score: 2.0	memory length: 78779	epsilon: 1.0	steps: 200	lr: 0.0
001	evaluation reward: 1.48				
441					
episode: 441	score: 1.0	memory length: 78930	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.49				
442					
episode: 442	score: 3.0	memory length: 79198	epsilon: 1.0	steps: 268	lr: 0.0
001	evaluation reward: 1.5				
443					
episode: 443	score: 2.0	memory length: 79419	epsilon: 1.0	steps: 221	lr: 0.0
001	evaluation reward: 1.52				
444					
episode: 444	score: 0.0	memory length: 79542	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.5				
445					
episode: 445	score: 2.0	memory length: 79721	epsilon: 1.0	steps: 179	lr: 0.0
001	evaluation reward: 1.51				
446					
episode: 446	score: 2.0	memory length: 79919	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.52				
447					
episode: 447	score: 1.0	memory length: 80089	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.53				
448					
episode: 448	score: 1.0	memory length: 80257	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.51				
449					
episode: 449	score: 3.0	memory length: 80503	epsilon: 1.0	steps: 246	lr: 0.0
001	evaluation reward: 1.54				
450					
episode: 450	score: 2.0	memory length: 80719	epsilon: 1.0	steps: 216	lr: 0.0
001	evaluation reward: 1.54				
451					
episode: 451	score: 2.0	memory length: 80922	epsilon: 1.0	steps: 203	lr: 0.0
001	evaluation reward: 1.55				
452					
episode: 452	score: 1.0	memory length: 81092	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.54				
453					
episode: 453	score: 0.0	memory length: 81215	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.53				
454					
episode: 454	score: 0.0	memory length: 81337	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.51				
455					
episode: 455	score: 0.0	memory length: 81460	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.48				
456					

episode: 456	score: 2.0	memory length: 81658	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.48				
457					
episode: 457	score: 2.0	memory length: 81877	epsilon: 1.0	steps: 219	lr: 0.0
001	evaluation reward: 1.48				
458					
episode: 458	score: 0.0	memory length: 81999	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.48				
459					
episode: 459	score: 5.0	memory length: 82305	epsilon: 1.0	steps: 306	lr: 0.0
001	evaluation reward: 1.5				
460					
episode: 460	score: 1.0	memory length: 82457	epsilon: 1.0	steps: 152	lr: 0.0
001	evaluation reward: 1.5				
461					
episode: 461	score: 6.0	memory length: 82816	epsilon: 1.0	steps: 359	lr: 0.0
001	evaluation reward: 1.53				
462					
episode: 462	score: 1.0	memory length: 82967	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.52				
463					
episode: 463	score: 0.0	memory length: 83090	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.5				
464					
episode: 464	score: 0.0	memory length: 83213	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.48				
465					
episode: 465	score: 2.0	memory length: 83411	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.48				
466					
episode: 466	score: 1.0	memory length: 83581	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.49				
467					
episode: 467	score: 2.0	memory length: 83800	epsilon: 1.0	steps: 219	lr: 0.0
001	evaluation reward: 1.45				
468					
episode: 468	score: 2.0	memory length: 83997	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.43				
469					
episode: 469	score: 0.0	memory length: 84120	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
470					
episode: 470	score: 1.0	memory length: 84288	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.41				
471					
episode: 471	score: 4.0	memory length: 84546	epsilon: 1.0	steps: 258	lr: 0.0
001	evaluation reward: 1.43				
472					
episode: 472	score: 0.0	memory length: 84669	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
473					
episode: 473	score: 1.0	memory length: 84838	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.41				
474					
episode: 474	score: 2.0	memory length: 85038	epsilon: 1.0	steps: 200	lr: 0.0
001	evaluation reward: 1.42				
475					
episode: 475	score: 0.0	memory length: 85161	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.4				
476					
episode: 476	score: 3.0	memory length: 85406	epsilon: 1.0	steps: 245	lr: 0.0
001	evaluation reward: 1.43				
477					
episode: 477	score: 1.0	memory length: 85575	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.43				
478					

episode: 478	score: 2.0	memory length: 85772	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.43				
479					
episode: 479	score: 1.0	memory length: 85941	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.43				
480					
episode: 480	score: 2.0	memory length: 86139	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.43				
481					
episode: 481	score: 2.0	memory length: 86337	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.42				
482					
episode: 482	score: 0.0	memory length: 86460	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.42				
483					
episode: 483	score: 3.0	memory length: 86687	epsilon: 1.0	steps: 227	lr: 0.0
001	evaluation reward: 1.42				
484					
episode: 484	score: 0.0	memory length: 86810	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
485					
episode: 485	score: 2.0	memory length: 87027	epsilon: 1.0	steps: 217	lr: 0.0
001	evaluation reward: 1.42				
486					
episode: 486	score: 0.0	memory length: 87149	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.42				
487					
episode: 487	score: 4.0	memory length: 87400	epsilon: 1.0	steps: 251	lr: 0.0
001	evaluation reward: 1.44				
488					
episode: 488	score: 0.0	memory length: 87522	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.44				
489					
episode: 489	score: 2.0	memory length: 87720	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.43				
490					
episode: 490	score: 0.0	memory length: 87843	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.43				
491					
episode: 491	score: 2.0	memory length: 88043	epsilon: 1.0	steps: 200	lr: 0.0
001	evaluation reward: 1.44				
492					
episode: 492	score: 0.0	memory length: 88165	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.43				
493					
episode: 493	score: 3.0	memory length: 88410	epsilon: 1.0	steps: 245	lr: 0.0
001	evaluation reward: 1.44				
494					
episode: 494	score: 1.0	memory length: 88561	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.42				
495					
episode: 495	score: 0.0	memory length: 88684	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
496					
episode: 496	score: 1.0	memory length: 88852	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.41				
497					
episode: 497	score: 0.0	memory length: 88974	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.37				
498					
episode: 498	score: 2.0	memory length: 89172	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.37				
499					
episode: 499	score: 1.0	memory length: 89342	epsilon: 1.0	steps: 170	lr: 0.0
001	evaluation reward: 1.38				
500					

episode: 500	score: 0.0	memory length: 89465	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.38				
501					
episode: 501	score: 0.0	memory length: 89588	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.38				
502					
episode: 502	score: 2.0	memory length: 89786	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.37				
503					
episode: 503	score: 0.0	memory length: 89909	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.37				
504					
episode: 504	score: 0.0	memory length: 90032	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.31				
505					
episode: 505	score: 3.0	memory length: 90276	epsilon: 1.0	steps: 244	lr: 0.0
001	evaluation reward: 1.33				
506					
episode: 506	score: 0.0	memory length: 90399	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.33				
507					
episode: 507	score: 0.0	memory length: 90522	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.32				
508					
episode: 508	score: 2.0	memory length: 90742	epsilon: 1.0	steps: 220	lr: 0.0
001	evaluation reward: 1.31				
509					
episode: 509	score: 2.0	memory length: 90942	epsilon: 1.0	steps: 200	lr: 0.0
001	evaluation reward: 1.33				
510					
episode: 510	score: 4.0	memory length: 91198	epsilon: 1.0	steps: 256	lr: 0.0
001	evaluation reward: 1.36				
511					
episode: 511	score: 0.0	memory length: 91321	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.35				
512					
episode: 512	score: 1.0	memory length: 91493	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.36				
513					
episode: 513	score: 2.0	memory length: 91710	epsilon: 1.0	steps: 217	lr: 0.0
001	evaluation reward: 1.37				
514					
episode: 514	score: 5.0	memory length: 92040	epsilon: 1.0	steps: 330	lr: 0.0
001	evaluation reward: 1.41				
515					
episode: 515	score: 1.0	memory length: 92209	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.42				
516					
episode: 516	score: 1.0	memory length: 92359	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.41				
517					
episode: 517	score: 0.0	memory length: 92481	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.41				
518					
episode: 518	score: 0.0	memory length: 92603	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.41				
519					
episode: 519	score: 0.0	memory length: 92726	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.41				
520					
episode: 520	score: 2.0	memory length: 92923	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.41				
521					
episode: 521	score: 2.0	memory length: 93139	epsilon: 1.0	steps: 216	lr: 0.0
001	evaluation reward: 1.43				
522					

episode: 522	score: 4.0	memory length: 93454	epsilon: 1.0	steps: 315	lr: 0.0
001	evaluation reward: 1.46				
523					
episode: 523	score: 1.0	memory length: 93623	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.46				
524					
episode: 524	score: 1.0	memory length: 93774	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.47				
525					
episode: 525	score: 0.0	memory length: 93896	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.45				
526					
episode: 526	score: 1.0	memory length: 94068	epsilon: 1.0	steps: 172	lr: 0.0
001	evaluation reward: 1.43				
527					
episode: 527	score: 4.0	memory length: 94344	epsilon: 1.0	steps: 276	lr: 0.0
001	evaluation reward: 1.45				
528					
episode: 528	score: 2.0	memory length: 94542	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.46				
529					
episode: 529	score: 1.0	memory length: 94710	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.45				
530					
episode: 530	score: 4.0	memory length: 94987	epsilon: 1.0	steps: 277	lr: 0.0
001	evaluation reward: 1.47				
531					
episode: 531	score: 1.0	memory length: 95156	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.44				
532					
episode: 532	score: 2.0	memory length: 95374	epsilon: 1.0	steps: 218	lr: 0.0
001	evaluation reward: 1.44				
533					
episode: 533	score: 1.0	memory length: 95525	epsilon: 1.0	steps: 151	lr: 0.0
001	evaluation reward: 1.41				
534					
episode: 534	score: 1.0	memory length: 95696	epsilon: 1.0	steps: 171	lr: 0.0
001	evaluation reward: 1.4				
535					
episode: 535	score: 2.0	memory length: 95893	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.4				
536					
episode: 536	score: 2.0	memory length: 96090	epsilon: 1.0	steps: 197	lr: 0.0
001	evaluation reward: 1.42				
537					
episode: 537	score: 0.0	memory length: 96212	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.41				
538					
episode: 538	score: 2.0	memory length: 96410	epsilon: 1.0	steps: 198	lr: 0.0
001	evaluation reward: 1.42				
539					
episode: 539	score: 1.0	memory length: 96560	epsilon: 1.0	steps: 150	lr: 0.0
001	evaluation reward: 1.43				
540					
episode: 540	score: 1.0	memory length: 96728	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.42				
541					
episode: 541	score: 2.0	memory length: 96947	epsilon: 1.0	steps: 219	lr: 0.0
001	evaluation reward: 1.43				
542					
episode: 542	score: 2.0	memory length: 97165	epsilon: 1.0	steps: 218	lr: 0.0
001	evaluation reward: 1.42				
543					
episode: 543	score: 0.0	memory length: 97287	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.4				
544					

episode: 544	score: 0.0	memory length: 97409	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.4				
545					
episode: 545	score: 1.0	memory length: 97577	epsilon: 1.0	steps: 168	lr: 0.0
001	evaluation reward: 1.39				
546					
episode: 546	score: 2.0	memory length: 97777	epsilon: 1.0	steps: 200	lr: 0.0
001	evaluation reward: 1.39				
547					
episode: 547	score: 3.0	memory length: 98003	epsilon: 1.0	steps: 226	lr: 0.0
001	evaluation reward: 1.41				
548					
episode: 548	score: 1.0	memory length: 98172	epsilon: 1.0	steps: 169	lr: 0.0
001	evaluation reward: 1.41				
549					
episode: 549	score: 0.0	memory length: 98294	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.38				
550					
episode: 550	score: 0.0	memory length: 98416	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.36				
551					
episode: 551	score: 7.0	memory length: 98814	epsilon: 1.0	steps: 398	lr: 0.0
001	evaluation reward: 1.41				
552					
episode: 552	score: 3.0	memory length: 99061	epsilon: 1.0	steps: 247	lr: 0.0
001	evaluation reward: 1.43				
553					
episode: 553	score: 0.0	memory length: 99183	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.43				
554					
episode: 554	score: 5.0	memory length: 99527	epsilon: 1.0	steps: 344	lr: 0.0
001	evaluation reward: 1.48				
555					
episode: 555	score: 0.0	memory length: 99649	epsilon: 1.0	steps: 122	lr: 0.0
001	evaluation reward: 1.48				
556					
episode: 556	score: 2.0	memory length: 99829	epsilon: 1.0	steps: 180	lr: 0.0
001	evaluation reward: 1.48				
557					
episode: 557	score: 3.0	memory length: 100039	epsilon: 0.9999208000000017		step
s: 210	lr: 0.0001	evaluation reward: 1.49			
558					
episode: 558	score: 2.0	memory length: 100236	epsilon: 0.9995307400000102		step
s: 197	lr: 0.0001	evaluation reward: 1.51			
559					
episode: 559	score: 2.0	memory length: 100454	epsilon: 0.9990991000000196		step
s: 218	lr: 0.0001	evaluation reward: 1.48			
560					
episode: 560	score: 3.0	memory length: 100681	epsilon: 0.9986496400000293		step
s: 227	lr: 0.0001	evaluation reward: 1.5			
561					
episode: 561	score: 2.0	memory length: 100879	epsilon: 0.9982576000000378		step
s: 198	lr: 0.0001	evaluation reward: 1.46			
562					
episode: 562	score: 1.0	memory length: 101050	epsilon: 0.9979190200000452		step
s: 171	lr: 0.0001	evaluation reward: 1.46			
563					
episode: 563	score: 2.0	memory length: 101248	epsilon: 0.9975269800000537		step
s: 198	lr: 0.0001	evaluation reward: 1.48			
564					
episode: 564	score: 2.0	memory length: 101467	epsilon: 0.9970933600000631		step
s: 219	lr: 0.0001	evaluation reward: 1.5			
565					
episode: 565	score: 0.0	memory length: 101590	epsilon: 0.9968498200000684		step
s: 123	lr: 0.0001	evaluation reward: 1.48			
566					

episode: 566	score: 1.0	memory length: 101741	epsilon: 0.9965508400000749	step
s: 151	lr: 0.0001	evaluation reward: 1.48		
567				
episode: 567	score: 0.0	memory length: 101864	epsilon: 0.9963073000000802	step
s: 123	lr: 0.0001	evaluation reward: 1.46		
568				
episode: 568	score: 6.0	memory length: 102254	epsilon: 0.9955351000000969	step
s: 390	lr: 0.0001	evaluation reward: 1.5		
569				
episode: 569	score: 0.0	memory length: 102377	epsilon: 0.9952915600001022	step
s: 123	lr: 0.0001	evaluation reward: 1.5		
570				
episode: 570	score: 0.0	memory length: 102500	epsilon: 0.9950480200001075	step
s: 123	lr: 0.0001	evaluation reward: 1.49		
571				
episode: 571	score: 0.0	memory length: 102622	epsilon: 0.9948064600001127	step
s: 122	lr: 0.0001	evaluation reward: 1.45		
572				
episode: 572	score: 1.0	memory length: 102790	epsilon: 0.99447382000012	steps:
168	lr: 0.0001	evaluation reward: 1.46		
573				
episode: 573	score: 2.0	memory length: 102976	epsilon: 0.994105540000128	steps:
186	lr: 0.0001	evaluation reward: 1.47		
574				
episode: 574	score: 1.0	memory length: 103144	epsilon: 0.9937729000001352	step
s: 168	lr: 0.0001	evaluation reward: 1.46		
575				
episode: 575	score: 1.0	memory length: 103316	epsilon: 0.9934323400001426	step
s: 172	lr: 0.0001	evaluation reward: 1.47		
576				
episode: 576	score: 4.0	memory length: 103589	epsilon: 0.9928918000001543	step
s: 273	lr: 0.0001	evaluation reward: 1.48		
577				
episode: 577	score: 1.0	memory length: 103758	epsilon: 0.9925571800001616	step
s: 169	lr: 0.0001	evaluation reward: 1.48		
578				
episode: 578	score: 1.0	memory length: 103929	epsilon: 0.9922186000001689	step
s: 171	lr: 0.0001	evaluation reward: 1.47		
579				
episode: 579	score: 0.0	memory length: 104051	epsilon: 0.9919770400001742	step
s: 122	lr: 0.0001	evaluation reward: 1.46		
580				
episode: 580	score: 1.0	memory length: 104201	epsilon: 0.9916800400001806	step
s: 150	lr: 0.0001	evaluation reward: 1.45		
581				
episode: 581	score: 1.0	memory length: 104369	epsilon: 0.9913474000001878	step
s: 168	lr: 0.0001	evaluation reward: 1.44		
582				
episode: 582	score: 3.0	memory length: 104615	epsilon: 0.9908603200001984	step
s: 246	lr: 0.0001	evaluation reward: 1.47		
583				
episode: 583	score: 2.0	memory length: 104796	epsilon: 0.9905019400002062	step
s: 181	lr: 0.0001	evaluation reward: 1.46		
584				
episode: 584	score: 0.0	memory length: 104919	epsilon: 0.9902584000002115	step
s: 123	lr: 0.0001	evaluation reward: 1.46		
585				
episode: 585	score: 4.0	memory length: 105193	epsilon: 0.9897158800002233	step
s: 274	lr: 0.0001	evaluation reward: 1.48		
586				
episode: 586	score: 0.0	memory length: 105316	epsilon: 0.9894723400002285	step
s: 123	lr: 0.0001	evaluation reward: 1.48		
587				
episode: 587	score: 3.0	memory length: 105581	epsilon: 0.9889476400002399	step
s: 265	lr: 0.0001	evaluation reward: 1.47		
588				

episode: 588	score: 1.0	memory length: 105732	epsilon: 0.9886486600002464	step
s: 151	lr: 0.0001	evaluation reward: 1.48		
589				
episode: 589	score: 2.0	memory length: 105930	epsilon: 0.9882566200002549	step
s: 198	lr: 0.0001	evaluation reward: 1.48		
590				
episode: 590	score: 2.0	memory length: 106128	epsilon: 0.9878645800002634	step
s: 198	lr: 0.0001	evaluation reward: 1.5		
591				
episode: 591	score: 5.0	memory length: 106441	epsilon: 0.9872448400002769	step
s: 313	lr: 0.0001	evaluation reward: 1.53		
592				
episode: 592	score: 0.0	memory length: 106564	epsilon: 0.9870013000002822	step
s: 123	lr: 0.0001	evaluation reward: 1.53		
593				
episode: 593	score: 3.0	memory length: 106830	epsilon: 0.9864746200002936	step
s: 266	lr: 0.0001	evaluation reward: 1.53		
594				
episode: 594	score: 0.0	memory length: 106953	epsilon: 0.9862310800002989	step
s: 123	lr: 0.0001	evaluation reward: 1.52		
595				
episode: 595	score: 1.0	memory length: 107104	epsilon: 0.9859321000003054	step
s: 151	lr: 0.0001	evaluation reward: 1.53		
596				
episode: 596	score: 0.0	memory length: 107227	epsilon: 0.9856885600003107	step
s: 123	lr: 0.0001	evaluation reward: 1.52		
597				
episode: 597	score: 2.0	memory length: 107425	epsilon: 0.9852965200003192	step
s: 198	lr: 0.0001	evaluation reward: 1.54		
598				
episode: 598	score: 0.0	memory length: 107547	epsilon: 0.9850549600003244	step
s: 122	lr: 0.0001	evaluation reward: 1.52		
599				
episode: 599	score: 3.0	memory length: 107815	epsilon: 0.984524320000336	steps:
268	lr: 0.0001	evaluation reward: 1.54		
600				
episode: 600	score: 4.0	memory length: 108109	epsilon: 0.9839422000003486	step
s: 294	lr: 0.0001	evaluation reward: 1.58		
601				
episode: 601	score: 3.0	memory length: 108338	epsilon: 0.9834887800003584	step
s: 229	lr: 0.0001	evaluation reward: 1.61		
602				
episode: 602	score: 2.0	memory length: 108520	epsilon: 0.9831284200003663	step
s: 182	lr: 0.0001	evaluation reward: 1.61		
603				
episode: 603	score: 1.0	memory length: 108689	epsilon: 0.9827938000003735	step
s: 169	lr: 0.0001	evaluation reward: 1.62		
604				
episode: 604	score: 2.0	memory length: 108886	epsilon: 0.982403740000382	steps:
197	lr: 0.0001	evaluation reward: 1.64		
605				
episode: 605	score: 4.0	memory length: 109181	epsilon: 0.9818196400003947	step
s: 295	lr: 0.0001	evaluation reward: 1.65		
606				
episode: 606	score: 3.0	memory length: 109426	epsilon: 0.9813345400004052	step
s: 245	lr: 0.0001	evaluation reward: 1.68		
607				
episode: 607	score: 1.0	memory length: 109596	epsilon: 0.9809979400004125	step
s: 170	lr: 0.0001	evaluation reward: 1.69		
608				
episode: 608	score: 0.0	memory length: 109719	epsilon: 0.9807544000004178	step
s: 123	lr: 0.0001	evaluation reward: 1.67		
609				
episode: 609	score: 4.0	memory length: 110034	epsilon: 0.9801307000004313	step
s: 315	lr: 0.0001	evaluation reward: 1.69		
610				

episode: 610	score: 2.0	memory length: 110231	epsilon: 0.9797406400004398	step
s: 197	lr: 0.0001	evaluation reward: 1.67		
611				
episode: 611	score: 0.0	memory length: 110354	epsilon: 0.9794971000004451	step
s: 123	lr: 0.0001	evaluation reward: 1.67		
612				
episode: 612	score: 1.0	memory length: 110524	epsilon: 0.9791605000004524	step
s: 170	lr: 0.0001	evaluation reward: 1.67		
613				
episode: 613	score: 1.0	memory length: 110675	epsilon: 0.9788615200004589	step
s: 151	lr: 0.0001	evaluation reward: 1.66		
614				
episode: 614	score: 5.0	memory length: 110989	epsilon: 0.9782398000004724	step
s: 314	lr: 0.0001	evaluation reward: 1.66		
615				
episode: 615	score: 0.0	memory length: 111111	epsilon: 0.9779982400004776	step
s: 122	lr: 0.0001	evaluation reward: 1.65		
616				
episode: 616	score: 2.0	memory length: 111329	epsilon: 0.977566600000487	steps:
218	lr: 0.0001	evaluation reward: 1.66		
617				
episode: 617	score: 0.0	memory length: 111452	epsilon: 0.9773230600004923	step
s: 123	lr: 0.0001	evaluation reward: 1.66		
618				
episode: 618	score: 2.0	memory length: 111650	epsilon: 0.9769310200005008	step
s: 198	lr: 0.0001	evaluation reward: 1.68		
619				
episode: 619	score: 0.0	memory length: 111772	epsilon: 0.976689460000506	steps:
122	lr: 0.0001	evaluation reward: 1.68		
620				
episode: 620	score: 1.0	memory length: 111943	epsilon: 0.9763508800005134	step
s: 171	lr: 0.0001	evaluation reward: 1.67		
621				
episode: 621	score: 2.0	memory length: 112143	epsilon: 0.975954880000522	steps:
200	lr: 0.0001	evaluation reward: 1.67		
622				
episode: 622	score: 2.0	memory length: 112341	epsilon: 0.9755628400005305	step
s: 198	lr: 0.0001	evaluation reward: 1.65		
623				
episode: 623	score: 1.0	memory length: 112510	epsilon: 0.9752282200005378	step
s: 169	lr: 0.0001	evaluation reward: 1.65		
624				
episode: 624	score: 2.0	memory length: 112708	epsilon: 0.9748361800005463	step
s: 198	lr: 0.0001	evaluation reward: 1.66		
625				
episode: 625	score: 0.0	memory length: 112830	epsilon: 0.9745946200005515	step
s: 122	lr: 0.0001	evaluation reward: 1.66		
626				
episode: 626	score: 0.0	memory length: 112953	epsilon: 0.9743510800005568	step
s: 123	lr: 0.0001	evaluation reward: 1.65		
627				
episode: 627	score: 0.0	memory length: 113075	epsilon: 0.9741095200005621	step
s: 122	lr: 0.0001	evaluation reward: 1.61		
628				
episode: 628	score: 0.0	memory length: 113198	epsilon: 0.9738659800005673	step
s: 123	lr: 0.0001	evaluation reward: 1.59		
629				
episode: 629	score: 2.0	memory length: 113417	epsilon: 0.9734323600005768	step
s: 219	lr: 0.0001	evaluation reward: 1.6		
630				
episode: 630	score: 3.0	memory length: 113669	epsilon: 0.9729334000005876	step
s: 252	lr: 0.0001	evaluation reward: 1.59		
631				
episode: 631	score: 2.0	memory length: 113854	epsilon: 0.9725671000005955	step
s: 185	lr: 0.0001	evaluation reward: 1.6		
632				

episode: 632	score: 2.0	memory length: 114071	epsilon: 0.9721374400006049	step
s: 217	lr: 0.0001	evaluation reward: 1.6		
633				
episode: 633	score: 3.0	memory length: 114338	epsilon: 0.9716087800006163	step
s: 267	lr: 0.0001	evaluation reward: 1.62		
634				
episode: 634	score: 2.0	memory length: 114557	epsilon: 0.9711751600006258	step
s: 219	lr: 0.0001	evaluation reward: 1.63		
635				
episode: 635	score: 1.0	memory length: 114725	epsilon: 0.970842520000633	steps:
168	lr: 0.0001	evaluation reward: 1.62		
636				
episode: 636	score: 1.0	memory length: 114897	epsilon: 0.9705019600006404	step
s: 172	lr: 0.0001	evaluation reward: 1.61		
637				
episode: 637	score: 2.0	memory length: 115097	epsilon: 0.970105960000649	steps:
200	lr: 0.0001	evaluation reward: 1.63		
638				
episode: 638	score: 0.0	memory length: 115220	epsilon: 0.9698624200006543	step
s: 123	lr: 0.0001	evaluation reward: 1.61		
639				
episode: 639	score: 2.0	memory length: 115418	epsilon: 0.9694703800006628	step
s: 198	lr: 0.0001	evaluation reward: 1.62		
640				
episode: 640	score: 2.0	memory length: 115618	epsilon: 0.9690743800006714	step
s: 200	lr: 0.0001	evaluation reward: 1.63		
641				
episode: 641	score: 0.0	memory length: 115741	epsilon: 0.9688308400006767	step
s: 123	lr: 0.0001	evaluation reward: 1.61		
642				
episode: 642	score: 2.0	memory length: 115959	epsilon: 0.968399200000686	steps:
218	lr: 0.0001	evaluation reward: 1.61		
643				
episode: 643	score: 5.0	memory length: 116261	epsilon: 0.967801240000699	steps:
302	lr: 0.0001	evaluation reward: 1.66		
644				
episode: 644	score: 0.0	memory length: 116384	epsilon: 0.9675577000007043	step
s: 123	lr: 0.0001	evaluation reward: 1.66		
645				
episode: 645	score: 0.0	memory length: 116507	epsilon: 0.9673141600007096	step
s: 123	lr: 0.0001	evaluation reward: 1.65		
646				
episode: 646	score: 0.0	memory length: 116629	epsilon: 0.9670726000007148	step
s: 122	lr: 0.0001	evaluation reward: 1.63		
647				
episode: 647	score: 0.0	memory length: 116751	epsilon: 0.9668310400007201	step
s: 122	lr: 0.0001	evaluation reward: 1.6		
648				
episode: 648	score: 2.0	memory length: 116968	epsilon: 0.9664013800007294	step
s: 217	lr: 0.0001	evaluation reward: 1.61		
649				
episode: 649	score: 0.0	memory length: 117091	epsilon: 0.9661578400007347	step
s: 123	lr: 0.0001	evaluation reward: 1.61		
650				
episode: 650	score: 5.0	memory length: 117439	epsilon: 0.9654688000007496	step
s: 348	lr: 0.0001	evaluation reward: 1.66		
651				
episode: 651	score: 1.0	memory length: 117589	epsilon: 0.9651718000007561	step
s: 150	lr: 0.0001	evaluation reward: 1.6		
652				
episode: 652	score: 2.0	memory length: 117806	epsilon: 0.9647421400007654	step
s: 217	lr: 0.0001	evaluation reward: 1.59		
653				
episode: 653	score: 2.0	memory length: 118007	epsilon: 0.964344160000774	steps:
201	lr: 0.0001	evaluation reward: 1.61		
654				

episode: 654	score: 2.0	memory length: 118205	epsilon: 0.9639521200007826	step
s: 198	lr: 0.0001	evaluation reward: 1.58		
655				
episode: 655	score: 3.0	memory length: 118454	epsilon: 0.9634591000007933	step
s: 249	lr: 0.0001	evaluation reward: 1.61		
656				
episode: 656	score: 1.0	memory length: 118605	epsilon: 0.9631601200007998	step
s: 151	lr: 0.0001	evaluation reward: 1.6		
657				
episode: 657	score: 0.0	memory length: 118727	epsilon: 0.962918560000805	steps:
122	lr: 0.0001	evaluation reward: 1.57		
658				
episode: 658	score: 2.0	memory length: 118924	epsilon: 0.9625285000008135	step
s: 197	lr: 0.0001	evaluation reward: 1.57		
659				
episode: 659	score: 2.0	memory length: 119143	epsilon: 0.9620948800008229	step
s: 219	lr: 0.0001	evaluation reward: 1.57		
660				
episode: 660	score: 2.0	memory length: 119340	epsilon: 0.9617048200008314	step
s: 197	lr: 0.0001	evaluation reward: 1.56		
661				
episode: 661	score: 2.0	memory length: 119526	epsilon: 0.9613365400008393	step
s: 186	lr: 0.0001	evaluation reward: 1.56		
662				
episode: 662	score: 4.0	memory length: 119828	epsilon: 0.9607385800008523	step
s: 302	lr: 0.0001	evaluation reward: 1.59		
663				
episode: 663	score: 0.0	memory length: 119951	epsilon: 0.9604950400008576	step
s: 123	lr: 0.0001	evaluation reward: 1.57		
664				
episode: 664	score: 2.0	memory length: 120148	epsilon: 0.9601049800008661	step
s: 197	lr: 0.0001	evaluation reward: 1.57		
665				
episode: 665	score: 0.0	memory length: 120271	epsilon: 0.9598614400008714	step
s: 123	lr: 0.0001	evaluation reward: 1.57		
666				
episode: 666	score: 0.0	memory length: 120394	epsilon: 0.9596179000008767	step
s: 123	lr: 0.0001	evaluation reward: 1.56		
667				
episode: 667	score: 1.0	memory length: 120563	epsilon: 0.9592832800008839	step
s: 169	lr: 0.0001	evaluation reward: 1.57		
668				
episode: 668	score: 2.0	memory length: 120761	epsilon: 0.9588912400008924	step
s: 198	lr: 0.0001	evaluation reward: 1.53		
669				
episode: 669	score: 3.0	memory length: 121028	epsilon: 0.9583625800009039	step
s: 267	lr: 0.0001	evaluation reward: 1.56		
670				
episode: 670	score: 4.0	memory length: 121324	epsilon: 0.9577765000009166	step
s: 296	lr: 0.0001	evaluation reward: 1.6		
671				
episode: 671	score: 0.0	memory length: 121447	epsilon: 0.9575329600009219	step
s: 123	lr: 0.0001	evaluation reward: 1.6		
672				
episode: 672	score: 1.0	memory length: 121618	epsilon: 0.9571943800009293	step
s: 171	lr: 0.0001	evaluation reward: 1.6		
673				
episode: 673	score: 1.0	memory length: 121768	epsilon: 0.9568973800009357	step
s: 150	lr: 0.0001	evaluation reward: 1.59		
674				
episode: 674	score: 2.0	memory length: 121986	epsilon: 0.9564657400009451	step
s: 218	lr: 0.0001	evaluation reward: 1.6		
675				
episode: 675	score: 5.0	memory length: 122331	epsilon: 0.9557826400009599	step
s: 345	lr: 0.0001	evaluation reward: 1.64		
676				

episode: 676	score: 4.0	memory length: 122609	epsilon: 0.9552322000009719	step
s: 278	lr: 0.0001	evaluation reward: 1.64		
677				
episode: 677	score: 3.0	memory length: 122857	epsilon: 0.9547411600009825	step
s: 248	lr: 0.0001	evaluation reward: 1.66		
678				
episode: 678	score: 2.0	memory length: 123057	epsilon: 0.9543451600009911	step
s: 200	lr: 0.0001	evaluation reward: 1.67		
679				
episode: 679	score: 4.0	memory length: 123358	epsilon: 0.9537491800010041	step
s: 301	lr: 0.0001	evaluation reward: 1.71		
680				
episode: 680	score: 3.0	memory length: 123586	epsilon: 0.9532977400010139	step
s: 228	lr: 0.0001	evaluation reward: 1.73		
681				
episode: 681	score: 2.0	memory length: 123801	epsilon: 0.9528720400010231	step
s: 215	lr: 0.0001	evaluation reward: 1.74		
682				
episode: 682	score: 0.0	memory length: 123923	epsilon: 0.9526304800010283	step
s: 122	lr: 0.0001	evaluation reward: 1.71		
683				
episode: 683	score: 0.0	memory length: 124046	epsilon: 0.9523869400010336	step
s: 123	lr: 0.0001	evaluation reward: 1.69		
684				
episode: 684	score: 0.0	memory length: 124169	epsilon: 0.9521434000010389	step
s: 123	lr: 0.0001	evaluation reward: 1.69		
685				
episode: 685	score: 1.0	memory length: 124320	epsilon: 0.9518444200010454	step
s: 151	lr: 0.0001	evaluation reward: 1.66		
686				
episode: 686	score: 2.0	memory length: 124535	epsilon: 0.9514187200010547	step
s: 215	lr: 0.0001	evaluation reward: 1.68		
687				
episode: 687	score: 2.0	memory length: 124753	epsilon: 0.950987080001064	steps:
218	lr: 0.0001	evaluation reward: 1.67		
688				
episode: 688	score: 3.0	memory length: 124982	epsilon: 0.9505336600010739	step
s: 229	lr: 0.0001	evaluation reward: 1.69		
689				
episode: 689	score: 3.0	memory length: 125229	epsilon: 0.9500446000010845	step
s: 247	lr: 0.0001	evaluation reward: 1.7		
690				
episode: 690	score: 2.0	memory length: 125447	epsilon: 0.9496129600010939	step
s: 218	lr: 0.0001	evaluation reward: 1.7		
691				
episode: 691	score: 1.0	memory length: 125598	epsilon: 0.9493139800011003	step
s: 151	lr: 0.0001	evaluation reward: 1.66		
692				
episode: 692	score: 2.0	memory length: 125796	epsilon: 0.9489219400011089	step
s: 198	lr: 0.0001	evaluation reward: 1.68		
693				
episode: 693	score: 1.0	memory length: 125967	epsilon: 0.9485833600011162	step
s: 171	lr: 0.0001	evaluation reward: 1.66		
694				
episode: 694	score: 2.0	memory length: 126164	epsilon: 0.9481933000011247	step
s: 197	lr: 0.0001	evaluation reward: 1.68		
695				
episode: 695	score: 0.0	memory length: 126287	epsilon: 0.947949760001113	steps:
123	lr: 0.0001	evaluation reward: 1.67		
696				
episode: 696	score: 3.0	memory length: 126512	epsilon: 0.9475042600011396	step
s: 225	lr: 0.0001	evaluation reward: 1.7		
697				
episode: 697	score: 2.0	memory length: 126710	epsilon: 0.9471122200011481	step
s: 198	lr: 0.0001	evaluation reward: 1.7		
698				

episode: 698	score: 0.0	memory length: 126833	epsilon: 0.9468686800011534	step
s: 123	lr: 0.0001	evaluation reward: 1.7		
699				
episode: 699	score: 2.0	memory length: 127033	epsilon: 0.946472680001162	steps:
200	lr: 0.0001	evaluation reward: 1.69		
700				
episode: 700	score: 5.0	memory length: 127359	epsilon: 0.945827200001176	steps:
326	lr: 0.0001	evaluation reward: 1.7		
701				
episode: 701	score: 3.0	memory length: 127586	epsilon: 0.9453777400011858	step
s: 227	lr: 0.0001	evaluation reward: 1.7		
702				
episode: 702	score: 1.0	memory length: 127757	epsilon: 0.9450391600011931	step
s: 171	lr: 0.0001	evaluation reward: 1.69		
703				
episode: 703	score: 1.0	memory length: 127908	epsilon: 0.9447401800011996	step
s: 151	lr: 0.0001	evaluation reward: 1.69		
704				
episode: 704	score: 2.0	memory length: 128128	epsilon: 0.9443045800012091	step
s: 220	lr: 0.0001	evaluation reward: 1.69		
705				
episode: 705	score: 0.0	memory length: 128251	epsilon: 0.9440610400012144	step
s: 123	lr: 0.0001	evaluation reward: 1.65		
706				
episode: 706	score: 0.0	memory length: 128373	epsilon: 0.9438194800012196	step
s: 122	lr: 0.0001	evaluation reward: 1.62		
707				
episode: 707	score: 1.0	memory length: 128544	epsilon: 0.943480900001227	steps:
171	lr: 0.0001	evaluation reward: 1.62		
708				
episode: 708	score: 1.0	memory length: 128713	epsilon: 0.9431462800012342	step
s: 169	lr: 0.0001	evaluation reward: 1.63		
709				
episode: 709	score: 4.0	memory length: 128989	epsilon: 0.9425998000012461	step
s: 276	lr: 0.0001	evaluation reward: 1.63		
710				
episode: 710	score: 2.0	memory length: 129207	epsilon: 0.9421681600012555	step
s: 218	lr: 0.0001	evaluation reward: 1.63		
711				
episode: 711	score: 2.0	memory length: 129406	epsilon: 0.941774140001264	steps:
199	lr: 0.0001	evaluation reward: 1.65		
712				
episode: 712	score: 0.0	memory length: 129528	epsilon: 0.9415325800012693	step
s: 122	lr: 0.0001	evaluation reward: 1.64		
713				
episode: 713	score: 3.0	memory length: 129773	epsilon: 0.9410474800012798	step
s: 245	lr: 0.0001	evaluation reward: 1.66		
714				
episode: 714	score: 0.0	memory length: 129895	epsilon: 0.940805920001285	steps:
122	lr: 0.0001	evaluation reward: 1.61		
715				
episode: 715	score: 1.0	memory length: 130064	epsilon: 0.9404713000012923	step
s: 169	lr: 0.0001	evaluation reward: 1.62		
716				
episode: 716	score: 1.0	memory length: 130234	epsilon: 0.9401347000012996	step
s: 170	lr: 0.0001	evaluation reward: 1.61		
717				
episode: 717	score: 2.0	memory length: 130431	epsilon: 0.9397446400013081	step
s: 197	lr: 0.0001	evaluation reward: 1.63		
718				
episode: 718	score: 2.0	memory length: 130612	epsilon: 0.9393862600013159	step
s: 181	lr: 0.0001	evaluation reward: 1.63		
719				
episode: 719	score: 0.0	memory length: 130734	epsilon: 0.9391447000013211	step
s: 122	lr: 0.0001	evaluation reward: 1.63		
720				

episode: 720	score: 2.0	memory length: 130953	epsilon: 0.9387110800013305	step
s: 219	lr: 0.0001	evaluation reward: 1.64		
721				
episode: 721	score: 2.0	memory length: 131151	epsilon: 0.938319040001339	steps:
198	lr: 0.0001	evaluation reward: 1.64		
722				
episode: 722	score: 3.0	memory length: 131397	epsilon: 0.9378319600013496	step
s: 246	lr: 0.0001	evaluation reward: 1.65		
723				
episode: 723	score: 1.0	memory length: 131567	epsilon: 0.9374953600013569	step
s: 170	lr: 0.0001	evaluation reward: 1.65		
724				
episode: 724	score: 3.0	memory length: 131813	epsilon: 0.9370082800013675	step
s: 246	lr: 0.0001	evaluation reward: 1.66		
725				
episode: 725	score: 2.0	memory length: 132012	epsilon: 0.936614260001376	steps:
199	lr: 0.0001	evaluation reward: 1.68		
726				
episode: 726	score: 1.0	memory length: 132181	epsilon: 0.9362796400013833	step
s: 169	lr: 0.0001	evaluation reward: 1.69		
727				
episode: 727	score: 4.0	memory length: 132480	epsilon: 0.9356876200013962	step
s: 299	lr: 0.0001	evaluation reward: 1.73		
728				
episode: 728	score: 2.0	memory length: 132678	epsilon: 0.9352955800014047	step
s: 198	lr: 0.0001	evaluation reward: 1.75		
729				
episode: 729	score: 0.0	memory length: 132800	epsilon: 0.9350540200014099	step
s: 122	lr: 0.0001	evaluation reward: 1.73		
730				
episode: 730	score: 0.0	memory length: 132923	epsilon: 0.9348104800014152	step
s: 123	lr: 0.0001	evaluation reward: 1.7		
731				
episode: 731	score: 1.0	memory length: 133091	epsilon: 0.9344778400014224	step
s: 168	lr: 0.0001	evaluation reward: 1.69		
732				
episode: 732	score: 2.0	memory length: 133309	epsilon: 0.9340462000014318	step
s: 218	lr: 0.0001	evaluation reward: 1.69		
733				
episode: 733	score: 3.0	memory length: 133557	epsilon: 0.9335551600014425	step
s: 248	lr: 0.0001	evaluation reward: 1.69		
734				
episode: 734	score: 2.0	memory length: 133757	epsilon: 0.933159160001451	steps:
200	lr: 0.0001	evaluation reward: 1.69		
735				
episode: 735	score: 3.0	memory length: 133983	epsilon: 0.9327116800014608	step
s: 226	lr: 0.0001	evaluation reward: 1.71		
736				
episode: 736	score: 0.0	memory length: 134106	epsilon: 0.932468140001466	steps:
123	lr: 0.0001	evaluation reward: 1.7		
737				
episode: 737	score: 3.0	memory length: 134352	epsilon: 0.9319810600014766	step
s: 246	lr: 0.0001	evaluation reward: 1.71		
738				
episode: 738	score: 0.0	memory length: 134475	epsilon: 0.9317375200014819	step
s: 123	lr: 0.0001	evaluation reward: 1.71		
739				
episode: 739	score: 4.0	memory length: 134754	epsilon: 0.9311851000014939	step
s: 279	lr: 0.0001	evaluation reward: 1.73		
740				
episode: 740	score: 4.0	memory length: 135067	epsilon: 0.9305653600015074	step
s: 313	lr: 0.0001	evaluation reward: 1.75		
741				
episode: 741	score: 3.0	memory length: 135313	epsilon: 0.9300782800015179	step
s: 246	lr: 0.0001	evaluation reward: 1.78		
742				

episode: 742	score: 1.0	memory length: 135463	epsilon: 0.9297812800015244	step
s: 150	lr: 0.0001	evaluation reward: 1.77		
743				
episode: 743	score: 2.0	memory length: 135678	epsilon: 0.9293555800015336	step
s: 215	lr: 0.0001	evaluation reward: 1.74		
744				
episode: 744	score: 0.0	memory length: 135800	epsilon: 0.9291140200015389	step
s: 122	lr: 0.0001	evaluation reward: 1.74		
745				
episode: 745	score: 0.0	memory length: 135923	epsilon: 0.9288704800015442	step
s: 123	lr: 0.0001	evaluation reward: 1.74		
746				
episode: 746	score: 4.0	memory length: 136236	epsilon: 0.9282507400015576	step
s: 313	lr: 0.0001	evaluation reward: 1.78		
747				
episode: 747	score: 2.0	memory length: 136434	epsilon: 0.9278587000015661	step
s: 198	lr: 0.0001	evaluation reward: 1.8		
748				
episode: 748	score: 0.0	memory length: 136557	epsilon: 0.9276151600015714	step
s: 123	lr: 0.0001	evaluation reward: 1.78		
749				
episode: 749	score: 0.0	memory length: 136680	epsilon: 0.9273716200015767	step
s: 123	lr: 0.0001	evaluation reward: 1.78		
750				
episode: 750	score: 1.0	memory length: 136830	epsilon: 0.9270746200015831	step
s: 150	lr: 0.0001	evaluation reward: 1.74		
751				
episode: 751	score: 2.0	memory length: 137048	epsilon: 0.9266429800015925	step
s: 218	lr: 0.0001	evaluation reward: 1.75		
752				
episode: 752	score: 3.0	memory length: 137295	epsilon: 0.9261539200016031	step
s: 247	lr: 0.0001	evaluation reward: 1.76		
753				
episode: 753	score: 2.0	memory length: 137493	epsilon: 0.9257618800016116	step
s: 198	lr: 0.0001	evaluation reward: 1.76		
754				
episode: 754	score: 5.0	memory length: 137828	epsilon: 0.925098580001626	steps:
335	lr: 0.0001	evaluation reward: 1.79		
755				
episode: 755	score: 2.0	memory length: 138046	epsilon: 0.9246669400016354	step
s: 218	lr: 0.0001	evaluation reward: 1.78		
756				
episode: 756	score: 2.0	memory length: 138262	epsilon: 0.9242392600016447	step
s: 216	lr: 0.0001	evaluation reward: 1.79		
757				
episode: 757	score: 3.0	memory length: 138509	epsilon: 0.9237502000016553	step
s: 247	lr: 0.0001	evaluation reward: 1.82		
758				
episode: 758	score: 1.0	memory length: 138678	epsilon: 0.9234155800016626	step
s: 169	lr: 0.0001	evaluation reward: 1.81		
759				
episode: 759	score: 0.0	memory length: 138801	epsilon: 0.9231720400016679	step
s: 123	lr: 0.0001	evaluation reward: 1.79		
760				
episode: 760	score: 1.0	memory length: 138952	epsilon: 0.9228730600016744	step
s: 151	lr: 0.0001	evaluation reward: 1.78		
761				
episode: 761	score: 2.0	memory length: 139150	epsilon: 0.9224810200016829	step
s: 198	lr: 0.0001	evaluation reward: 1.78		
762				
episode: 762	score: 1.0	memory length: 139301	epsilon: 0.9221820400016894	step
s: 151	lr: 0.0001	evaluation reward: 1.75		
763				
episode: 763	score: 1.0	memory length: 139452	epsilon: 0.9218830600016958	step
s: 151	lr: 0.0001	evaluation reward: 1.76		
764				

episode: 764	score: 1.0	memory length: 139621	epsilon: 0.9215484400017031	step
s: 169	lr: 0.0001	evaluation reward: 1.75		
765				
episode: 765	score: 0.0	memory length: 139744	epsilon: 0.9213049000017084	step
s: 123	lr: 0.0001	evaluation reward: 1.75		
766				
episode: 766	score: 0.0	memory length: 139867	epsilon: 0.9210613600017137	step
s: 123	lr: 0.0001	evaluation reward: 1.75		
767				
episode: 767	score: 3.0	memory length: 140134	epsilon: 0.9205327000017252	step
s: 267	lr: 0.0001	evaluation reward: 1.77		
768				
episode: 768	score: 2.0	memory length: 140332	epsilon: 0.9201406600017337	step
s: 198	lr: 0.0001	evaluation reward: 1.77		
769				
episode: 769	score: 1.0	memory length: 140483	epsilon: 0.9198416800017402	step
s: 151	lr: 0.0001	evaluation reward: 1.75		
770				
episode: 770	score: 3.0	memory length: 140714	epsilon: 0.9193843000017501	step
s: 231	lr: 0.0001	evaluation reward: 1.74		
771				
episode: 771	score: 1.0	memory length: 140883	epsilon: 0.9190496800017574	step
s: 169	lr: 0.0001	evaluation reward: 1.75		
772				
episode: 772	score: 1.0	memory length: 141055	epsilon: 0.9187091200017647	step
s: 172	lr: 0.0001	evaluation reward: 1.75		
773				
episode: 773	score: 1.0	memory length: 141225	epsilon: 0.918372520001772	steps:
170	lr: 0.0001	evaluation reward: 1.75		
774				
episode: 774	score: 4.0	memory length: 141518	epsilon: 0.9177923800017846	step
s: 293	lr: 0.0001	evaluation reward: 1.77		
775				
episode: 775	score: 6.0	memory length: 141889	epsilon: 0.9170578000018006	step
s: 371	lr: 0.0001	evaluation reward: 1.78		
776				
episode: 776	score: 1.0	memory length: 142041	epsilon: 0.9167568400018071	step
s: 152	lr: 0.0001	evaluation reward: 1.75		
777				
episode: 777	score: 3.0	memory length: 142286	epsilon: 0.9162717400018177	step
s: 245	lr: 0.0001	evaluation reward: 1.75		
778				
episode: 778	score: 1.0	memory length: 142436	epsilon: 0.9159747400018241	step
s: 150	lr: 0.0001	evaluation reward: 1.74		
779				
episode: 779	score: 0.0	memory length: 142559	epsilon: 0.9157312000018294	step
s: 123	lr: 0.0001	evaluation reward: 1.7		
780				
episode: 780	score: 0.0	memory length: 142682	epsilon: 0.9154876600018347	step
s: 123	lr: 0.0001	evaluation reward: 1.67		
781				
episode: 781	score: 2.0	memory length: 142902	epsilon: 0.9150520600018441	step
s: 220	lr: 0.0001	evaluation reward: 1.67		
782				
episode: 782	score: 3.0	memory length: 143128	epsilon: 0.9146045800018539	step
s: 226	lr: 0.0001	evaluation reward: 1.7		
783				
episode: 783	score: 1.0	memory length: 143279	epsilon: 0.9143056000018603	step
s: 151	lr: 0.0001	evaluation reward: 1.71		
784				
episode: 784	score: 1.0	memory length: 143448	epsilon: 0.9139709800018676	step
s: 169	lr: 0.0001	evaluation reward: 1.72		
785				
episode: 785	score: 1.0	memory length: 143616	epsilon: 0.9136383400018748	step
s: 168	lr: 0.0001	evaluation reward: 1.72		
786				

episode: 786	score: 3.0	memory length: 143843	epsilon: 0.9131888800018846	step
s: 227	lr: 0.0001	evaluation reward: 1.73		
787				
episode: 787	score: 1.0	memory length: 144011	epsilon: 0.9128562400018918	step
s: 168	lr: 0.0001	evaluation reward: 1.72		
788				
episode: 788	score: 3.0	memory length: 144260	epsilon: 0.9123632200019025	step
s: 249	lr: 0.0001	evaluation reward: 1.72		
789				
episode: 789	score: 2.0	memory length: 144458	epsilon: 0.911971180001911	steps:
198	lr: 0.0001	evaluation reward: 1.71		
790				
episode: 790	score: 1.0	memory length: 144609	epsilon: 0.9116722000019175	step
s: 151	lr: 0.0001	evaluation reward: 1.7		
791				
episode: 791	score: 2.0	memory length: 144807	epsilon: 0.911280160001926	steps:
198	lr: 0.0001	evaluation reward: 1.71		
792				
episode: 792	score: 1.0	memory length: 144957	epsilon: 0.9109831600019325	step
s: 150	lr: 0.0001	evaluation reward: 1.7		
793				
episode: 793	score: 4.0	memory length: 145252	epsilon: 0.9103990600019451	step
s: 295	lr: 0.0001	evaluation reward: 1.73		
794				
episode: 794	score: 0.0	memory length: 145374	epsilon: 0.9101575000019504	step
s: 122	lr: 0.0001	evaluation reward: 1.71		
795				
episode: 795	score: 1.0	memory length: 145543	epsilon: 0.9098228800019577	step
s: 169	lr: 0.0001	evaluation reward: 1.72		
796				
episode: 796	score: 3.0	memory length: 145769	epsilon: 0.9093754000019674	step
s: 226	lr: 0.0001	evaluation reward: 1.72		
797				
episode: 797	score: 0.0	memory length: 145891	epsilon: 0.9091338400019726	step
s: 122	lr: 0.0001	evaluation reward: 1.7		
798				
episode: 798	score: 2.0	memory length: 146089	epsilon: 0.9087418000019811	step
s: 198	lr: 0.0001	evaluation reward: 1.72		
799				
episode: 799	score: 0.0	memory length: 146212	epsilon: 0.9084982600019864	step
s: 123	lr: 0.0001	evaluation reward: 1.7		
800				
episode: 800	score: 0.0	memory length: 146334	epsilon: 0.9082567000019917	step
s: 122	lr: 0.0001	evaluation reward: 1.65		
801				
episode: 801	score: 1.0	memory length: 146485	epsilon: 0.9079577200019981	step
s: 151	lr: 0.0001	evaluation reward: 1.63		
802				
episode: 802	score: 1.0	memory length: 146636	epsilon: 0.9076587400020046	step
s: 151	lr: 0.0001	evaluation reward: 1.63		
803				
episode: 803	score: 2.0	memory length: 146833	epsilon: 0.9072686800020131	step
s: 197	lr: 0.0001	evaluation reward: 1.64		
804				
episode: 804	score: 3.0	memory length: 147080	epsilon: 0.9067796200020237	step
s: 247	lr: 0.0001	evaluation reward: 1.65		
805				
episode: 805	score: 0.0	memory length: 147203	epsilon: 0.906536080002029	steps:
123	lr: 0.0001	evaluation reward: 1.65		
806				
episode: 806	score: 3.0	memory length: 147468	epsilon: 0.9060113800020404	step
s: 265	lr: 0.0001	evaluation reward: 1.68		
807				
episode: 807	score: 2.0	memory length: 147667	epsilon: 0.905617360002049	steps:
199	lr: 0.0001	evaluation reward: 1.69		
808				

episode: 808	score: 1.0	memory length: 147817	epsilon: 0.9053203600020554	step
s: 150	lr: 0.0001	evaluation reward: 1.69		
809				
episode: 809	score: 1.0	memory length: 147968	epsilon: 0.9050213800020619	step
s: 151	lr: 0.0001	evaluation reward: 1.66		
810				
episode: 810	score: 3.0	memory length: 148235	epsilon: 0.9044927200020734	step
s: 267	lr: 0.0001	evaluation reward: 1.67		
811				
episode: 811	score: 1.0	memory length: 148406	epsilon: 0.9041541400020807	step
s: 171	lr: 0.0001	evaluation reward: 1.66		
812				
episode: 812	score: 0.0	memory length: 148529	epsilon: 0.903910600002086	steps:
123	lr: 0.0001	evaluation reward: 1.66		
813				
episode: 813	score: 0.0	memory length: 148651	epsilon: 0.9036690400020913	step
s: 122	lr: 0.0001	evaluation reward: 1.63		
814				
episode: 814	score: 3.0	memory length: 148898	epsilon: 0.9031799800021019	step
s: 247	lr: 0.0001	evaluation reward: 1.66		
815				
episode: 815	score: 3.0	memory length: 149162	epsilon: 0.9026572600021132	step
s: 264	lr: 0.0001	evaluation reward: 1.68		
816				
episode: 816	score: 2.0	memory length: 149362	epsilon: 0.9022612600021218	step
s: 200	lr: 0.0001	evaluation reward: 1.69		
817				
episode: 817	score: 3.0	memory length: 149588	epsilon: 0.9018137800021315	step
s: 226	lr: 0.0001	evaluation reward: 1.7		
818				
episode: 818	score: 0.0	memory length: 149710	epsilon: 0.9015722200021368	step
s: 122	lr: 0.0001	evaluation reward: 1.68		
819				
episode: 819	score: 0.0	memory length: 149832	epsilon: 0.901330660002142	steps:
122	lr: 0.0001	evaluation reward: 1.68		
820				
episode: 820	score: 3.0	memory length: 150077	epsilon: 0.9008455600021525	step
s: 245	lr: 0.0001	evaluation reward: 1.69		
821				
episode: 821	score: 1.0	memory length: 150227	epsilon: 0.900548560002159	steps:
150	lr: 0.0001	evaluation reward: 1.68		
822				
episode: 822	score: 1.0	memory length: 150395	epsilon: 0.9002159200021662	step
s: 168	lr: 0.0001	evaluation reward: 1.66		
823				
episode: 823	score: 3.0	memory length: 150621	epsilon: 0.8997684400021759	step
s: 226	lr: 0.0001	evaluation reward: 1.68		
824				
episode: 824	score: 0.0	memory length: 150744	epsilon: 0.8995249000021812	step
s: 123	lr: 0.0001	evaluation reward: 1.65		
825				
episode: 825	score: 2.0	memory length: 150963	epsilon: 0.8990912800021906	step
s: 219	lr: 0.0001	evaluation reward: 1.65		
826				
episode: 826	score: 2.0	memory length: 151179	epsilon: 0.8986636000021999	step
s: 216	lr: 0.0001	evaluation reward: 1.66		
827				
episode: 827	score: 3.0	memory length: 151424	epsilon: 0.8981785000022104	step
s: 245	lr: 0.0001	evaluation reward: 1.65		
828				
episode: 828	score: 5.0	memory length: 151767	epsilon: 0.8974993600022252	step
s: 343	lr: 0.0001	evaluation reward: 1.68		
829				
episode: 829	score: 0.0	memory length: 151889	epsilon: 0.8972578000022304	step
s: 122	lr: 0.0001	evaluation reward: 1.68		
830				

episode: 830	score: 1.0	memory length: 152040	epsilon: 0.8969588200022369	step
s: 151	lr: 0.0001	evaluation reward: 1.69		
831				
episode: 831	score: 2.0	memory length: 152258	epsilon: 0.8965271800022463	step
s: 218	lr: 0.0001	evaluation reward: 1.7		
832				
episode: 832	score: 2.0	memory length: 152477	epsilon: 0.8960935600022557	step
s: 219	lr: 0.0001	evaluation reward: 1.7		
833				
episode: 833	score: 0.0	memory length: 152600	epsilon: 0.895850020002261	steps:
123	lr: 0.0001	evaluation reward: 1.67		
834				
episode: 834	score: 1.0	memory length: 152769	epsilon: 0.8955154000022683	step
s: 169	lr: 0.0001	evaluation reward: 1.66		
835				
episode: 835	score: 4.0	memory length: 153059	epsilon: 0.8949412000022807	step
s: 290	lr: 0.0001	evaluation reward: 1.67		
836				
episode: 836	score: 3.0	memory length: 153304	epsilon: 0.8944561000022913	step
s: 245	lr: 0.0001	evaluation reward: 1.7		
837				
episode: 837	score: 4.0	memory length: 153582	epsilon: 0.8939056600023032	step
s: 278	lr: 0.0001	evaluation reward: 1.71		
838				
episode: 838	score: 2.0	memory length: 153803	epsilon: 0.8934680800023127	step
s: 221	lr: 0.0001	evaluation reward: 1.73		
839				
episode: 839	score: 1.0	memory length: 153954	epsilon: 0.8931691000023192	step
s: 151	lr: 0.0001	evaluation reward: 1.7		
840				
episode: 840	score: 1.0	memory length: 154105	epsilon: 0.8928701200023257	step
s: 151	lr: 0.0001	evaluation reward: 1.67		
841				
episode: 841	score: 1.0	memory length: 154256	epsilon: 0.8925711400023322	step
s: 151	lr: 0.0001	evaluation reward: 1.65		
842				
episode: 842	score: 2.0	memory length: 154454	epsilon: 0.8921791000023407	step
s: 198	lr: 0.0001	evaluation reward: 1.66		
843				
episode: 843	score: 1.0	memory length: 154605	epsilon: 0.8918801200023472	step
s: 151	lr: 0.0001	evaluation reward: 1.65		
844				
episode: 844	score: 2.0	memory length: 154803	epsilon: 0.8914880800023557	step
s: 198	lr: 0.0001	evaluation reward: 1.67		
845				
episode: 845	score: 3.0	memory length: 155049	epsilon: 0.8910010000023663	step
s: 246	lr: 0.0001	evaluation reward: 1.7		
846				
episode: 846	score: 0.0	memory length: 155172	epsilon: 0.8907574600023715	step
s: 123	lr: 0.0001	evaluation reward: 1.66		
847				
episode: 847	score: 0.0	memory length: 155294	epsilon: 0.8905159000023768	step
s: 122	lr: 0.0001	evaluation reward: 1.64		
848				
episode: 848	score: 2.0	memory length: 155492	epsilon: 0.8901238600023853	step
s: 198	lr: 0.0001	evaluation reward: 1.66		
849				
episode: 849	score: 2.0	memory length: 155690	epsilon: 0.8897318200023938	step
s: 198	lr: 0.0001	evaluation reward: 1.68		
850				
episode: 850	score: 0.0	memory length: 155813	epsilon: 0.8894882800023991	step
s: 123	lr: 0.0001	evaluation reward: 1.67		
851				
episode: 851	score: 0.0	memory length: 155936	epsilon: 0.8892447400024044	step
s: 123	lr: 0.0001	evaluation reward: 1.65		
852				

episode: 852	score: 3.0	memory length: 156181	epsilon: 0.8887596400024149	step
s: 245	lr: 0.0001	evaluation reward: 1.65		
853				
episode: 853	score: 5.0	memory length: 156496	epsilon: 0.8881359400024285	step
s: 315	lr: 0.0001	evaluation reward: 1.68		
854				
episode: 854	score: 0.0	memory length: 156619	epsilon: 0.8878924000024337	step
s: 123	lr: 0.0001	evaluation reward: 1.63		
855				
episode: 855	score: 2.0	memory length: 156821	epsilon: 0.8874924400024424	step
s: 202	lr: 0.0001	evaluation reward: 1.63		
856				
episode: 856	score: 3.0	memory length: 157046	epsilon: 0.8870469400024521	step
s: 225	lr: 0.0001	evaluation reward: 1.64		
857				
episode: 857	score: 2.0	memory length: 157229	epsilon: 0.88668460000246	steps:
183	lr: 0.0001	evaluation reward: 1.63		
858				
episode: 858	score: 3.0	memory length: 157493	epsilon: 0.8861618800024713	step
s: 264	lr: 0.0001	evaluation reward: 1.65		
859				
episode: 859	score: 2.0	memory length: 157709	epsilon: 0.8857342000024806	step
s: 216	lr: 0.0001	evaluation reward: 1.67		
860				
episode: 860	score: 2.0	memory length: 157890	epsilon: 0.8853758200024884	step
s: 181	lr: 0.0001	evaluation reward: 1.68		
861				
episode: 861	score: 3.0	memory length: 158117	epsilon: 0.8849263600024981	step
s: 227	lr: 0.0001	evaluation reward: 1.69		
862				
episode: 862	score: 3.0	memory length: 158361	epsilon: 0.8844432400025086	step
s: 244	lr: 0.0001	evaluation reward: 1.71		
863				
episode: 863	score: 2.0	memory length: 158558	epsilon: 0.8840531800025171	step
s: 197	lr: 0.0001	evaluation reward: 1.72		
864				
episode: 864	score: 1.0	memory length: 158709	epsilon: 0.8837542000025236	step
s: 151	lr: 0.0001	evaluation reward: 1.72		
865				
episode: 865	score: 1.0	memory length: 158860	epsilon: 0.8834552200025301	step
s: 151	lr: 0.0001	evaluation reward: 1.73		
866				
episode: 866	score: 2.0	memory length: 159061	epsilon: 0.8830572400025387	step
s: 201	lr: 0.0001	evaluation reward: 1.75		
867				
episode: 867	score: 0.0	memory length: 159183	epsilon: 0.882815680002544	steps:
122	lr: 0.0001	evaluation reward: 1.72		
868				
episode: 868	score: 2.0	memory length: 159401	epsilon: 0.8823840400025533	step
s: 218	lr: 0.0001	evaluation reward: 1.72		
869				
episode: 869	score: 0.0	memory length: 159523	epsilon: 0.8821424800025586	step
s: 122	lr: 0.0001	evaluation reward: 1.71		
870				
episode: 870	score: 3.0	memory length: 159734	epsilon: 0.8817247000025676	step
s: 211	lr: 0.0001	evaluation reward: 1.71		
871				
episode: 871	score: 1.0	memory length: 159884	epsilon: 0.8814277000025741	step
s: 150	lr: 0.0001	evaluation reward: 1.71		
872				
episode: 872	score: 1.0	memory length: 160034	epsilon: 0.8811307000025805	step
s: 150	lr: 0.0001	evaluation reward: 1.71		
873				
episode: 873	score: 2.0	memory length: 160232	epsilon: 0.880738660002589	steps:
198	lr: 0.0001	evaluation reward: 1.72		
874				

episode: 874	score: 0.0	memory length: 160354	epsilon: 0.8804971000025943	step
s: 122	lr: 0.0001	evaluation reward: 1.68		
875				
episode: 875	score: 4.0	memory length: 160671	epsilon: 0.8798694400026079	step
s: 317	lr: 0.0001	evaluation reward: 1.66		
876				
episode: 876	score: 2.0	memory length: 160869	epsilon: 0.8794774000026164	step
s: 198	lr: 0.0001	evaluation reward: 1.67		
877				
episode: 877	score: 4.0	memory length: 161127	epsilon: 0.8789665600026275	step
s: 258	lr: 0.0001	evaluation reward: 1.68		
878				
episode: 878	score: 1.0	memory length: 161295	epsilon: 0.8786339200026347	step
s: 168	lr: 0.0001	evaluation reward: 1.68		
879				
episode: 879	score: 2.0	memory length: 161493	epsilon: 0.8782418800026432	step
s: 198	lr: 0.0001	evaluation reward: 1.7		
880				
episode: 880	score: 2.0	memory length: 161709	epsilon: 0.8778142000026525	step
s: 216	lr: 0.0001	evaluation reward: 1.72		
881				
episode: 881	score: 2.0	memory length: 161907	epsilon: 0.877422160002661	steps:
s: 198	lr: 0.0001	evaluation reward: 1.72		
882				
episode: 882	score: 3.0	memory length: 162156	epsilon: 0.8769291400026717	step
s: 249	lr: 0.0001	evaluation reward: 1.72		
883				
episode: 883	score: 4.0	memory length: 162449	epsilon: 0.8763490000026843	step
s: 293	lr: 0.0001	evaluation reward: 1.75		
884				
episode: 884	score: 3.0	memory length: 162696	epsilon: 0.875859940002695	steps:
s: 247	lr: 0.0001	evaluation reward: 1.77		
885				
episode: 885	score: 1.0	memory length: 162847	epsilon: 0.8755609600027014	step
s: 151	lr: 0.0001	evaluation reward: 1.77		
886				
episode: 886	score: 5.0	memory length: 163193	epsilon: 0.8748758800027163	step
s: 346	lr: 0.0001	evaluation reward: 1.79		
887				
episode: 887	score: 0.0	memory length: 163316	epsilon: 0.8746323400027216	step
s: 123	lr: 0.0001	evaluation reward: 1.78		
888				
episode: 888	score: 3.0	memory length: 163541	epsilon: 0.8741868400027313	step
s: 225	lr: 0.0001	evaluation reward: 1.78		
889				
episode: 889	score: 1.0	memory length: 163691	epsilon: 0.8738898400027377	step
s: 150	lr: 0.0001	evaluation reward: 1.77		
890				
episode: 890	score: 3.0	memory length: 163917	epsilon: 0.8734423600027474	step
s: 226	lr: 0.0001	evaluation reward: 1.79		
891				
episode: 891	score: 1.0	memory length: 164067	epsilon: 0.8731453600027539	step
s: 150	lr: 0.0001	evaluation reward: 1.78		
892				
episode: 892	score: 2.0	memory length: 164284	epsilon: 0.8727157000027632	step
s: 217	lr: 0.0001	evaluation reward: 1.79		
893				
episode: 893	score: 3.0	memory length: 164510	epsilon: 0.8722682200027729	step
s: 226	lr: 0.0001	evaluation reward: 1.78		
894				
episode: 894	score: 2.0	memory length: 164708	epsilon: 0.8718761800027814	step
s: 198	lr: 0.0001	evaluation reward: 1.8		
895				
episode: 895	score: 1.0	memory length: 164859	epsilon: 0.8715772000027879	step
s: 151	lr: 0.0001	evaluation reward: 1.8		
896				

episode: 896	score: 3.0	memory length: 165085	epsilon: 0.8711297200027976	step
s: 226	lr: 0.0001	evaluation reward: 1.8		
897				
episode: 897	score: 2.0	memory length: 165264	epsilon: 0.8707753000028053	step
s: 179	lr: 0.0001	evaluation reward: 1.82		
898				
episode: 898	score: 0.0	memory length: 165387	epsilon: 0.8705317600028106	step
s: 123	lr: 0.0001	evaluation reward: 1.8		
899				
episode: 899	score: 0.0	memory length: 165509	epsilon: 0.8702902000028159	step
s: 122	lr: 0.0001	evaluation reward: 1.8		
900				
episode: 900	score: 4.0	memory length: 165790	epsilon: 0.869733820002828	steps:
281	lr: 0.0001	evaluation reward: 1.84		
901				
episode: 901	score: 3.0	memory length: 166003	epsilon: 0.8693120800028371	step
s: 213	lr: 0.0001	evaluation reward: 1.86		
902				
episode: 902	score: 2.0	memory length: 166183	epsilon: 0.8689556800028448	step
s: 180	lr: 0.0001	evaluation reward: 1.87		
903				
episode: 903	score: 2.0	memory length: 166364	epsilon: 0.8685973000028526	step
s: 181	lr: 0.0001	evaluation reward: 1.87		
904				
episode: 904	score: 3.0	memory length: 166591	epsilon: 0.8681478400028624	step
s: 227	lr: 0.0001	evaluation reward: 1.87		
905				
episode: 905	score: 3.0	memory length: 166838	epsilon: 0.867658780002873	steps:
247	lr: 0.0001	evaluation reward: 1.9		
906				
episode: 906	score: 2.0	memory length: 167018	epsilon: 0.8673023800028807	step
s: 180	lr: 0.0001	evaluation reward: 1.89		
907				
episode: 907	score: 0.0	memory length: 167140	epsilon: 0.867060820002886	steps:
122	lr: 0.0001	evaluation reward: 1.87		
908				
episode: 908	score: 4.0	memory length: 167419	epsilon: 0.866508400002898	steps:
279	lr: 0.0001	evaluation reward: 1.9		
909				
episode: 909	score: 1.0	memory length: 167570	epsilon: 0.8662094200029045	step
s: 151	lr: 0.0001	evaluation reward: 1.9		
910				
episode: 910	score: 0.0	memory length: 167693	epsilon: 0.8659658800029097	step
s: 123	lr: 0.0001	evaluation reward: 1.87		
911				
episode: 911	score: 0.0	memory length: 167816	epsilon: 0.865722340002915	steps:
123	lr: 0.0001	evaluation reward: 1.86		
912				
episode: 912	score: 2.0	memory length: 168034	epsilon: 0.8652907000029244	step
s: 218	lr: 0.0001	evaluation reward: 1.88		
913				
episode: 913	score: 2.0	memory length: 168216	epsilon: 0.8649303400029322	step
s: 182	lr: 0.0001	evaluation reward: 1.9		
914				
episode: 914	score: 0.0	memory length: 168339	epsilon: 0.8646868000029375	step
s: 123	lr: 0.0001	evaluation reward: 1.87		
915				
episode: 915	score: 3.0	memory length: 168583	epsilon: 0.864203680002948	steps:
244	lr: 0.0001	evaluation reward: 1.87		
916				
episode: 916	score: 5.0	memory length: 168908	epsilon: 0.863560180002962	steps:
325	lr: 0.0001	evaluation reward: 1.9		
917				
episode: 917	score: 2.0	memory length: 169105	epsilon: 0.8631701200029704	step
s: 197	lr: 0.0001	evaluation reward: 1.89		
918				

episode: 918	score: 2.0	memory length: 169325	epsilon: 0.8627345200029799	step
s: 220	lr: 0.0001	evaluation reward: 1.91		
919				
episode: 919	score: 2.0	memory length: 169523	epsilon: 0.8623424800029884	step
s: 198	lr: 0.0001	evaluation reward: 1.93		
920				
episode: 920	score: 1.0	memory length: 169695	epsilon: 0.8620019200029958	step
s: 172	lr: 0.0001	evaluation reward: 1.91		
921				
episode: 921	score: 2.0	memory length: 169913	epsilon: 0.8615702800030052	step
s: 218	lr: 0.0001	evaluation reward: 1.92		
922				
episode: 922	score: 2.0	memory length: 170113	epsilon: 0.8611742800030138	step
s: 200	lr: 0.0001	evaluation reward: 1.93		
923				
episode: 923	score: 5.0	memory length: 170457	epsilon: 0.8604931600030286	step
s: 344	lr: 0.0001	evaluation reward: 1.95		
924				
episode: 924	score: 0.0	memory length: 170579	epsilon: 0.8602516000030338	step
s: 122	lr: 0.0001	evaluation reward: 1.95		
925				
episode: 925	score: 2.0	memory length: 170777	epsilon: 0.8598595600030423	step
s: 198	lr: 0.0001	evaluation reward: 1.95		
926				
episode: 926	score: 0.0	memory length: 170899	epsilon: 0.8596180000030476	step
s: 122	lr: 0.0001	evaluation reward: 1.93		
927				
episode: 927	score: 2.0	memory length: 171098	epsilon: 0.8592239800030561	step
s: 199	lr: 0.0001	evaluation reward: 1.92		
928				
episode: 928	score: 0.0	memory length: 171221	epsilon: 0.8589804400030614	step
s: 123	lr: 0.0001	evaluation reward: 1.87		
929				
episode: 929	score: 4.0	memory length: 171494	epsilon: 0.8584399000030731	step
s: 273	lr: 0.0001	evaluation reward: 1.91		
930				
episode: 930	score: 0.0	memory length: 171617	epsilon: 0.8581963600030784	step
s: 123	lr: 0.0001	evaluation reward: 1.9		
931				
episode: 931	score: 4.0	memory length: 171876	epsilon: 0.8576835400030896	step
s: 259	lr: 0.0001	evaluation reward: 1.92		
932				
episode: 932	score: 2.0	memory length: 172074	epsilon: 0.8572915000030981	step
s: 198	lr: 0.0001	evaluation reward: 1.92		
933				
episode: 933	score: 3.0	memory length: 172322	epsilon: 0.8568004600031087	step
s: 248	lr: 0.0001	evaluation reward: 1.95		
934				
episode: 934	score: 0.0	memory length: 172444	epsilon: 0.8565589000031114	steps:
122	lr: 0.0001	evaluation reward: 1.94		
935				
episode: 935	score: 0.0	memory length: 172567	epsilon: 0.8563153600031193	step
s: 123	lr: 0.0001	evaluation reward: 1.9		
936				
episode: 936	score: 4.0	memory length: 172863	epsilon: 0.855729280003132	steps:
296	lr: 0.0001	evaluation reward: 1.91		
937				
episode: 937	score: 2.0	memory length: 173044	epsilon: 0.8553709000031398	step
s: 181	lr: 0.0001	evaluation reward: 1.89		
938				
episode: 938	score: 2.0	memory length: 173226	epsilon: 0.8550105400031476	step
s: 182	lr: 0.0001	evaluation reward: 1.89		
939				
episode: 939	score: 3.0	memory length: 173456	epsilon: 0.8545551400031575	step
s: 230	lr: 0.0001	evaluation reward: 1.91		
940				

episode: 940	score: 0.0	memory length: 173579	epsilon: 0.8543116000031628	step
s: 123	lr: 0.0001	evaluation reward: 1.9		
941				
episode: 941	score: 0.0	memory length: 173702	epsilon: 0.854068060003168	steps:
123	lr: 0.0001	evaluation reward: 1.89		
942				
episode: 942	score: 2.0	memory length: 173900	epsilon: 0.8536760200031765	step
s: 198	lr: 0.0001	evaluation reward: 1.89		
943				
episode: 943	score: 1.0	memory length: 174050	epsilon: 0.853379020003183	steps:
150	lr: 0.0001	evaluation reward: 1.89		
944				
episode: 944	score: 0.0	memory length: 174173	epsilon: 0.8531354800031883	step
s: 123	lr: 0.0001	evaluation reward: 1.87		
945				
episode: 945	score: 2.0	memory length: 174388	epsilon: 0.8527097800031975	step
s: 215	lr: 0.0001	evaluation reward: 1.86		
946				
episode: 946	score: 2.0	memory length: 174605	epsilon: 0.8522801200032069	step
s: 217	lr: 0.0001	evaluation reward: 1.88		
947				
episode: 947	score: 2.0	memory length: 174802	epsilon: 0.8518900600032153	step
s: 197	lr: 0.0001	evaluation reward: 1.9		
948				
episode: 948	score: 2.0	memory length: 175000	epsilon: 0.8514980200032238	step
s: 198	lr: 0.0001	evaluation reward: 1.9		
949				
episode: 949	score: 2.0	memory length: 175198	epsilon: 0.8511059800032323	step
s: 198	lr: 0.0001	evaluation reward: 1.9		
950				
episode: 950	score: 4.0	memory length: 175476	epsilon: 0.8505555400032443	step
s: 278	lr: 0.0001	evaluation reward: 1.94		
951				
episode: 951	score: 0.0	memory length: 175599	epsilon: 0.8503120000032496	step
s: 123	lr: 0.0001	evaluation reward: 1.94		
952				
episode: 952	score: 2.0	memory length: 175817	epsilon: 0.849880360003259	steps:
218	lr: 0.0001	evaluation reward: 1.93		
953				
episode: 953	score: 1.0	memory length: 175986	epsilon: 0.8495457400032662	step
s: 169	lr: 0.0001	evaluation reward: 1.89		
954				
episode: 954	score: 2.0	memory length: 176202	epsilon: 0.8491180600032755	step
s: 216	lr: 0.0001	evaluation reward: 1.91		
955				
episode: 955	score: 1.0	memory length: 176352	epsilon: 0.848821060003282	steps:
150	lr: 0.0001	evaluation reward: 1.9		
956				
episode: 956	score: 1.0	memory length: 176503	epsilon: 0.8485220800032884	step
s: 151	lr: 0.0001	evaluation reward: 1.88		
957				
episode: 957	score: 2.0	memory length: 176703	epsilon: 0.848126080003297	steps:
200	lr: 0.0001	evaluation reward: 1.88		
958				
episode: 958	score: 2.0	memory length: 176923	epsilon: 0.8476904800033065	step
s: 220	lr: 0.0001	evaluation reward: 1.87		
959				
episode: 959	score: 7.0	memory length: 177366	epsilon: 0.8468133400033255	step
s: 443	lr: 0.0001	evaluation reward: 1.92		
960				
episode: 960	score: 5.0	memory length: 177670	epsilon: 0.8462114200033386	step
s: 304	lr: 0.0001	evaluation reward: 1.95		
961				
episode: 961	score: 4.0	memory length: 177968	epsilon: 0.8456213800033514	step
s: 298	lr: 0.0001	evaluation reward: 1.96		
962				

episode: 962	score: 1.0	memory length: 178119	epsilon: 0.8453224000033579	step
s: 151	lr: 0.0001	evaluation reward: 1.94		
963				
episode: 963	score: 2.0	memory length: 178339	epsilon: 0.8448868000033674	step
s: 220	lr: 0.0001	evaluation reward: 1.94		
964				
episode: 964	score: 1.0	memory length: 178508	epsilon: 0.8445521800033746	step
s: 169	lr: 0.0001	evaluation reward: 1.94		
965				
episode: 965	score: 1.0	memory length: 178677	epsilon: 0.8442175600033819	step
s: 169	lr: 0.0001	evaluation reward: 1.94		
966				
episode: 966	score: 2.0	memory length: 178857	epsilon: 0.8438611600033896	step
s: 180	lr: 0.0001	evaluation reward: 1.94		
967				
episode: 967	score: 0.0	memory length: 178979	epsilon: 0.8436196000033949	step
s: 122	lr: 0.0001	evaluation reward: 1.94		
968				
episode: 968	score: 1.0	memory length: 179148	epsilon: 0.8432849800034021	step
s: 169	lr: 0.0001	evaluation reward: 1.93		
969				
episode: 969	score: 3.0	memory length: 179373	epsilon: 0.8428394800034118	step
s: 225	lr: 0.0001	evaluation reward: 1.96		
970				
episode: 970	score: 2.0	memory length: 179573	epsilon: 0.8424434800034204	step
s: 200	lr: 0.0001	evaluation reward: 1.95		
971				
episode: 971	score: 4.0	memory length: 179887	epsilon: 0.8418217600034339	step
s: 314	lr: 0.0001	evaluation reward: 1.98		
972				
episode: 972	score: 3.0	memory length: 180096	epsilon: 0.8414079400034429	step
s: 209	lr: 0.0001	evaluation reward: 2.0		
973				
episode: 973	score: 2.0	memory length: 180278	epsilon: 0.8410475800034507	step
s: 182	lr: 0.0001	evaluation reward: 2.0		
974				
episode: 974	score: 2.0	memory length: 180476	epsilon: 0.8406555400034592	step
s: 198	lr: 0.0001	evaluation reward: 2.02		
975				
episode: 975	score: 1.0	memory length: 180644	epsilon: 0.8403229000034664	step
s: 168	lr: 0.0001	evaluation reward: 1.99		
976				
episode: 976	score: 0.0	memory length: 180766	epsilon: 0.8400813400034717	step
s: 122	lr: 0.0001	evaluation reward: 1.97		
977				
episode: 977	score: 4.0	memory length: 181062	epsilon: 0.8394952600034844	step
s: 296	lr: 0.0001	evaluation reward: 1.97		
978				
episode: 978	score: 0.0	memory length: 181184	epsilon: 0.8392537000034896	step
s: 122	lr: 0.0001	evaluation reward: 1.96		
979				
episode: 979	score: 3.0	memory length: 181429	epsilon: 0.8387686000035002	step
s: 245	lr: 0.0001	evaluation reward: 1.97		
980				
episode: 980	score: 2.0	memory length: 181644	epsilon: 0.8383429000035094	step
s: 215	lr: 0.0001	evaluation reward: 1.97		
981				
episode: 981	score: 2.0	memory length: 181842	epsilon: 0.8379508600035179	step
s: 198	lr: 0.0001	evaluation reward: 1.97		
982				
episode: 982	score: 0.0	memory length: 181964	epsilon: 0.8377093000035232	step
s: 122	lr: 0.0001	evaluation reward: 1.94		
983				
episode: 983	score: 1.0	memory length: 182133	epsilon: 0.8373746800035304	step
s: 169	lr: 0.0001	evaluation reward: 1.91		
984				

episode: 984	score: 3.0	memory length: 182363	epsilon: 0.8369192800035403	step
s: 230	lr: 0.0001	evaluation reward: 1.91		
985				
episode: 985	score: 2.0	memory length: 182562	epsilon: 0.8365252600035489	step
s: 199	lr: 0.0001	evaluation reward: 1.92		
986				
episode: 986	score: 3.0	memory length: 182790	epsilon: 0.8360738200035587	step
s: 228	lr: 0.0001	evaluation reward: 1.9		
987				
episode: 987	score: 1.0	memory length: 182958	epsilon: 0.8357411800035659	step
s: 168	lr: 0.0001	evaluation reward: 1.91		
988				
episode: 988	score: 1.0	memory length: 183109	epsilon: 0.8354422000035724	step
s: 151	lr: 0.0001	evaluation reward: 1.89		
989				
episode: 989	score: 1.0	memory length: 183280	epsilon: 0.8351036200035797	step
s: 171	lr: 0.0001	evaluation reward: 1.89		
990				
episode: 990	score: 1.0	memory length: 183431	epsilon: 0.8348046400035862	step
s: 151	lr: 0.0001	evaluation reward: 1.87		
991				
episode: 991	score: 2.0	memory length: 183628	epsilon: 0.8344145800035947	step
s: 197	lr: 0.0001	evaluation reward: 1.88		
992				
episode: 992	score: 2.0	memory length: 183828	epsilon: 0.8340185800036033	step
s: 200	lr: 0.0001	evaluation reward: 1.88		
993				
episode: 993	score: 1.0	memory length: 183999	epsilon: 0.8336800000036106	step
s: 171	lr: 0.0001	evaluation reward: 1.86		
994				
episode: 994	score: 3.0	memory length: 184226	epsilon: 0.8332305400036204	step
s: 227	lr: 0.0001	evaluation reward: 1.87		
995				
episode: 995	score: 2.0	memory length: 184424	epsilon: 0.8328385000036289	step
s: 198	lr: 0.0001	evaluation reward: 1.88		
996				
episode: 996	score: 5.0	memory length: 184742	epsilon: 0.8322088600036426	step
s: 318	lr: 0.0001	evaluation reward: 1.9		
997				
episode: 997	score: 0.0	memory length: 184865	epsilon: 0.8319653200036479	step
s: 123	lr: 0.0001	evaluation reward: 1.88		
998				
episode: 998	score: 2.0	memory length: 185062	epsilon: 0.8315752600036563	step
s: 197	lr: 0.0001	evaluation reward: 1.9		
999				
episode: 999	score: 3.0	memory length: 185290	epsilon: 0.8311238200036661	step
s: 228	lr: 0.0001	evaluation reward: 1.93		
1000				
episode: 1000	score: 2.0	memory length: 185507	epsilon: 0.8306941600036755	step
s: 217	lr: 0.0001	evaluation reward: 1.91		
1001				
episode: 1001	score: 6.0	memory length: 185883	epsilon: 0.8299496800036916	step
s: 376	lr: 0.0001	evaluation reward: 1.94		
1002				
episode: 1002	score: 1.0	memory length: 186034	epsilon: 0.8296507000036981	step
s: 151	lr: 0.0001	evaluation reward: 1.93		
1003				
episode: 1003	score: 3.0	memory length: 186283	epsilon: 0.8291576800037088	step
s: 249	lr: 0.0001	evaluation reward: 1.94		
1004				
episode: 1004	score: 6.0	memory length: 186611	epsilon: 0.8285082400037229	step
s: 328	lr: 0.0001	evaluation reward: 1.97		
1005				
episode: 1005	score: 2.0	memory length: 186809	epsilon: 0.8281162000037314	step
s: 198	lr: 0.0001	evaluation reward: 1.96		
1006				

episode: 1006	score: 2.0	memory length: 187011	epsilon: 0.8277162400037401	step
s: 202	lr: 0.0001	evaluation reward: 1.96		
1007				
episode: 1007	score: 1.0	memory length: 187162	epsilon: 0.8274172600037466	step
s: 151	lr: 0.0001	evaluation reward: 1.97		
1008				
episode: 1008	score: 4.0	memory length: 187424	epsilon: 0.8268985000037579	step
s: 262	lr: 0.0001	evaluation reward: 1.97		
1009				
episode: 1009	score: 2.0	memory length: 187642	epsilon: 0.8264668600037672	step
s: 218	lr: 0.0001	evaluation reward: 1.98		
1010				
episode: 1010	score: 2.0	memory length: 187842	epsilon: 0.8260708600037758	step
s: 200	lr: 0.0001	evaluation reward: 2.0		
1011				
episode: 1011	score: 0.0	memory length: 187965	epsilon: 0.8258273200037811	step
s: 123	lr: 0.0001	evaluation reward: 2.0		
1012				
episode: 1012	score: 2.0	memory length: 188162	epsilon: 0.8254372600037896	step
s: 197	lr: 0.0001	evaluation reward: 2.0		
1013				
episode: 1013	score: 2.0	memory length: 188359	epsilon: 0.825047200003798	step
s: 197	lr: 0.0001	evaluation reward: 2.0		
1014				
episode: 1014	score: 3.0	memory length: 188606	epsilon: 0.8245581400038087	step
s: 247	lr: 0.0001	evaluation reward: 2.03		
1015				
episode: 1015	score: 4.0	memory length: 188867	epsilon: 0.8240413600038199	step
s: 261	lr: 0.0001	evaluation reward: 2.04		
1016				
episode: 1016	score: 7.0	memory length: 189273	epsilon: 0.8232374800038373	step
s: 406	lr: 0.0001	evaluation reward: 2.06		
1017				
episode: 1017	score: 7.0	memory length: 189675	epsilon: 0.8224415200038546	step
s: 402	lr: 0.0001	evaluation reward: 2.11		
1018				
episode: 1018	score: 0.0	memory length: 189798	epsilon: 0.8221979800038599	step
s: 123	lr: 0.0001	evaluation reward: 2.09		
1019				
episode: 1019	score: 3.0	memory length: 190024	epsilon: 0.8217505000038696	step
s: 226	lr: 0.0001	evaluation reward: 2.1		
1020				
episode: 1020	score: 3.0	memory length: 190235	epsilon: 0.8213327200038787	step
s: 211	lr: 0.0001	evaluation reward: 2.12		
1021				
episode: 1021	score: 4.0	memory length: 190549	epsilon: 0.8207110000038922	step
s: 314	lr: 0.0001	evaluation reward: 2.14		
1022				
episode: 1022	score: 2.0	memory length: 190750	epsilon: 0.8203130200039008	step
s: 201	lr: 0.0001	evaluation reward: 2.14		
1023				
episode: 1023	score: 0.0	memory length: 190873	epsilon: 0.8200694800039061	step
s: 123	lr: 0.0001	evaluation reward: 2.09		
1024				
episode: 1024	score: 5.0	memory length: 191200	epsilon: 0.8194220200039202	step
s: 327	lr: 0.0001	evaluation reward: 2.14		
1025				
episode: 1025	score: 4.0	memory length: 191444	epsilon: 0.8189389000039307	step
s: 244	lr: 0.0001	evaluation reward: 2.16		
1026				
episode: 1026	score: 1.0	memory length: 191615	epsilon: 0.818600320003938	step
s: 171	lr: 0.0001	evaluation reward: 2.17		
1027				
episode: 1027	score: 0.0	memory length: 191738	epsilon: 0.8183567800039433	step
s: 123	lr: 0.0001	evaluation reward: 2.15		
1028				

episode: 1028	score: 3.0	memory length: 191970	epsilon: 0.8178974200039533	step
s: 232	lr: 0.0001	evaluation reward: 2.18		
1029				
episode: 1029	score: 2.0	memory length: 192152	epsilon: 0.8175370600039611	step
s: 182	lr: 0.0001	evaluation reward: 2.16		
1030				
episode: 1030	score: 5.0	memory length: 192518	epsilon: 0.8168123800039768	step
s: 366	lr: 0.0001	evaluation reward: 2.21		
1031				
episode: 1031	score: 1.0	memory length: 192669	epsilon: 0.8165134000039833	step
s: 151	lr: 0.0001	evaluation reward: 2.18		
1032				
episode: 1032	score: 3.0	memory length: 192894	epsilon: 0.816067900003993	step
s: 225	lr: 0.0001	evaluation reward: 2.19		
1033				
episode: 1033	score: 0.0	memory length: 193017	epsilon: 0.8158243600039983	step
s: 123	lr: 0.0001	evaluation reward: 2.16		
1034				
episode: 1034	score: 0.0	memory length: 193140	epsilon: 0.8155808200040036	step
s: 123	lr: 0.0001	evaluation reward: 2.16		
1035				
episode: 1035	score: 0.0	memory length: 193263	epsilon: 0.8153372800040088	step
s: 123	lr: 0.0001	evaluation reward: 2.16		
1036				
episode: 1036	score: 1.0	memory length: 193414	epsilon: 0.8150383000040153	step
s: 151	lr: 0.0001	evaluation reward: 2.13		
1037				
episode: 1037	score: 10.0	memory length: 193833	epsilon: 0.8142086800040333	step
ps: 419	lr: 0.0001	evaluation reward: 2.21		
1038				
episode: 1038	score: 6.0	memory length: 194199	epsilon: 0.8134840000040491	step
s: 366	lr: 0.0001	evaluation reward: 2.25		
1039				
episode: 1039	score: 4.0	memory length: 194438	epsilon: 0.8130107800040594	step
s: 239	lr: 0.0001	evaluation reward: 2.26		
1040				
episode: 1040	score: 3.0	memory length: 194667	epsilon: 0.8125573600040692	step
s: 229	lr: 0.0001	evaluation reward: 2.29		
1041				
episode: 1041	score: 1.0	memory length: 194819	epsilon: 0.8122564000040757	step
s: 152	lr: 0.0001	evaluation reward: 2.3		
1042				
episode: 1042	score: 2.0	memory length: 195038	epsilon: 0.8118227800040851	step
s: 219	lr: 0.0001	evaluation reward: 2.3		
1043				
episode: 1043	score: 4.0	memory length: 195333	epsilon: 0.8112386800040978	step
s: 295	lr: 0.0001	evaluation reward: 2.33		
1044				
episode: 1044	score: 4.0	memory length: 195593	epsilon: 0.810723880004109	step
s: 260	lr: 0.0001	evaluation reward: 2.37		
1045				
episode: 1045	score: 4.0	memory length: 195886	epsilon: 0.8101437400041216	step
s: 293	lr: 0.0001	evaluation reward: 2.39		
1046				
episode: 1046	score: 1.0	memory length: 196055	epsilon: 0.8098091200041289	step
s: 169	lr: 0.0001	evaluation reward: 2.38		
1047				
episode: 1047	score: 2.0	memory length: 196237	epsilon: 0.8094487600041367	step
s: 182	lr: 0.0001	evaluation reward: 2.38		
1048				
episode: 1048	score: 6.0	memory length: 196581	epsilon: 0.8087676400041515	step
s: 344	lr: 0.0001	evaluation reward: 2.42		
1049				
episode: 1049	score: 4.0	memory length: 196876	epsilon: 0.8081835400041641	step
s: 295	lr: 0.0001	evaluation reward: 2.44		
1050				

episode: 1050	score: 0.0	memory length: 196998	epsilon: 0.8079419800041694	step
s: 122	lr: 0.0001	evaluation reward: 2.4		
1051				
episode: 1051	score: 4.0	memory length: 197255	epsilon: 0.8074331200041804	step
s: 257	lr: 0.0001	evaluation reward: 2.44		
1052				
episode: 1052	score: 5.0	memory length: 197580	epsilon: 0.8067896200041944	step
s: 325	lr: 0.0001	evaluation reward: 2.47		
1053				
episode: 1053	score: 5.0	memory length: 197908	epsilon: 0.8061401800042085	step
s: 328	lr: 0.0001	evaluation reward: 2.51		
1054				
episode: 1054	score: 3.0	memory length: 198121	epsilon: 0.8057184400042177	step
s: 213	lr: 0.0001	evaluation reward: 2.52		
1055				
episode: 1055	score: 4.0	memory length: 198401	epsilon: 0.8051640400042297	step
s: 280	lr: 0.0001	evaluation reward: 2.55		
1056				
episode: 1056	score: 0.0	memory length: 198524	epsilon: 0.804920500004235	step
s: 123	lr: 0.0001	evaluation reward: 2.54		
1057				
episode: 1057	score: 1.0	memory length: 198675	epsilon: 0.8046215200042415	step
s: 151	lr: 0.0001	evaluation reward: 2.53		
1058				
episode: 1058	score: 0.0	memory length: 198798	epsilon: 0.8043779800042468	step
s: 123	lr: 0.0001	evaluation reward: 2.51		
1059				
episode: 1059	score: 3.0	memory length: 199008	epsilon: 0.8039621800042558	step
s: 210	lr: 0.0001	evaluation reward: 2.47		
1060				
episode: 1060	score: 3.0	memory length: 199233	epsilon: 0.8035166800042655	step
s: 225	lr: 0.0001	evaluation reward: 2.45		
1061				
episode: 1061	score: 0.0	memory length: 199355	epsilon: 0.8032751200042707	step
s: 122	lr: 0.0001	evaluation reward: 2.41		
1062				
episode: 1062	score: 3.0	memory length: 199601	epsilon: 0.8027880400042813	step
s: 246	lr: 0.0001	evaluation reward: 2.43		
1063				
episode: 1063	score: 5.0	memory length: 199925	epsilon: 0.8021465200042952	step
s: 324	lr: 0.0001	evaluation reward: 2.46		
1064				
episode: 1064	score: 3.0	memory length: 200153	epsilon: 0.801695080004305	step
s: 228	lr: 4e-05	evaluation reward: 2.48		
1065				
episode: 1065	score: 3.0	memory length: 200362	epsilon: 0.801281260004314	step
s: 209	lr: 4e-05	evaluation reward: 2.5		
1066				
episode: 1066	score: 3.0	memory length: 200613	epsilon: 0.8007842800043248	step
s: 251	lr: 4e-05	evaluation reward: 2.51		
1067				
episode: 1067	score: 2.0	memory length: 200811	epsilon: 0.8003922400043333	step
s: 198	lr: 4e-05	evaluation reward: 2.53		
1068				
episode: 1068	score: 2.0	memory length: 201030	epsilon: 0.7999586200043427	step
s: 219	lr: 4e-05	evaluation reward: 2.54		
1069				
episode: 1069	score: 3.0	memory length: 201258	epsilon: 0.7995071800043525	step
s: 228	lr: 4e-05	evaluation reward: 2.54		
1070				
episode: 1070	score: 1.0	memory length: 201409	epsilon: 0.799208200004359	step
s: 151	lr: 4e-05	evaluation reward: 2.53		
1071				
episode: 1071	score: 2.0	memory length: 201607	epsilon: 0.7988161600043675	step
s: 198	lr: 4e-05	evaluation reward: 2.51		
1072				

episode: 1072	score: 2.0	memory length: 201808	epsilon: 0.7984181800043761	step
s: 201	lr: 4e-05	evaluation reward: 2.5		
1073				
episode: 1073	score: 1.0	memory length: 201959	epsilon: 0.7981192000043826	step
s: 151	lr: 4e-05	evaluation reward: 2.49		
1074				
episode: 1074	score: 3.0	memory length: 202187	epsilon: 0.7976677600043924	step
s: 228	lr: 4e-05	evaluation reward: 2.5		
1075				
episode: 1075	score: 0.0	memory length: 202310	epsilon: 0.7974242200043977	step
s: 123	lr: 4e-05	evaluation reward: 2.49		
1076				
episode: 1076	score: 1.0	memory length: 202460	epsilon: 0.7971272200044042	step
s: 150	lr: 4e-05	evaluation reward: 2.5		
1077				
episode: 1077	score: 1.0	memory length: 202610	epsilon: 0.7968302200044106	step
s: 150	lr: 4e-05	evaluation reward: 2.47		
1078				
episode: 1078	score: 4.0	memory length: 202874	epsilon: 0.796307500004422	step
s: 264	lr: 4e-05	evaluation reward: 2.51		
1079				
episode: 1079	score: 2.0	memory length: 203053	epsilon: 0.7959530800044297	step
s: 179	lr: 4e-05	evaluation reward: 2.5		
1080				
episode: 1080	score: 2.0	memory length: 203268	epsilon: 0.7955273800044389	step
s: 215	lr: 4e-05	evaluation reward: 2.5		
1081				
episode: 1081	score: 1.0	memory length: 203437	epsilon: 0.7951927600044462	step
s: 169	lr: 4e-05	evaluation reward: 2.49		
1082				
episode: 1082	score: 2.0	memory length: 203618	epsilon: 0.7948343800044539	step
s: 181	lr: 4e-05	evaluation reward: 2.51		
1083				
episode: 1083	score: 7.0	memory length: 204015	epsilon: 0.794048320004471	step
s: 397	lr: 4e-05	evaluation reward: 2.57		
1084				
episode: 1084	score: 2.0	memory length: 204214	epsilon: 0.7936543000044796	step
s: 199	lr: 4e-05	evaluation reward: 2.56		
1085				
episode: 1085	score: 4.0	memory length: 204496	epsilon: 0.7930959400044917	step
s: 282	lr: 4e-05	evaluation reward: 2.58		
1086				
episode: 1086	score: 1.0	memory length: 204664	epsilon: 0.7927633000044989	step
s: 168	lr: 4e-05	evaluation reward: 2.56		
1087				
episode: 1087	score: 1.0	memory length: 204815	epsilon: 0.7924643200045054	step
s: 151	lr: 4e-05	evaluation reward: 2.56		
1088				
episode: 1088	score: 1.0	memory length: 204966	epsilon: 0.7921653400045119	step
s: 151	lr: 4e-05	evaluation reward: 2.56		
1089				
episode: 1089	score: 2.0	memory length: 205145	epsilon: 0.7918109200045196	step
s: 179	lr: 4e-05	evaluation reward: 2.57		
1090				
episode: 1090	score: 5.0	memory length: 205455	epsilon: 0.7911971200045329	step
s: 310	lr: 4e-05	evaluation reward: 2.61		
1091				
episode: 1091	score: 6.0	memory length: 205812	epsilon: 0.7904902600045483	step
s: 357	lr: 4e-05	evaluation reward: 2.65		
1092				
episode: 1092	score: 1.0	memory length: 205963	epsilon: 0.7901912800045547	step
s: 151	lr: 4e-05	evaluation reward: 2.64		
1093				
episode: 1093	score: 5.0	memory length: 206235	epsilon: 0.7896527200045664	step
s: 272	lr: 4e-05	evaluation reward: 2.68		
1094				

episode: 1094	score: 2.0	memory length: 206433	epsilon: 0.7892606800045749	step
s: 198	lr: 4e-05	evaluation reward: 2.67		
1095				
episode: 1095	score: 4.0	memory length: 206728	epsilon: 0.7886765800045876	step
s: 295	lr: 4e-05	evaluation reward: 2.69		
1096				
episode: 1096	score: 5.0	memory length: 207052	epsilon: 0.7880350600046016	step
s: 324	lr: 4e-05	evaluation reward: 2.69		
1097				
episode: 1097	score: 1.0	memory length: 207203	epsilon: 0.787736080004608	step
s: 151	lr: 4e-05	evaluation reward: 2.7		
1098				
episode: 1098	score: 3.0	memory length: 207428	epsilon: 0.7872905800046177	step
s: 225	lr: 4e-05	evaluation reward: 2.71		
1099				
episode: 1099	score: 2.0	memory length: 207646	epsilon: 0.7868589400046271	step
s: 218	lr: 4e-05	evaluation reward: 2.7		
1100				
episode: 1100	score: 6.0	memory length: 207961	epsilon: 0.7862352400046406	step
s: 315	lr: 4e-05	evaluation reward: 2.74		
1101				
episode: 1101	score: 4.0	memory length: 208237	epsilon: 0.7856887600046525	step
s: 276	lr: 4e-05	evaluation reward: 2.72		
1102				
episode: 1102	score: 2.0	memory length: 208418	epsilon: 0.7853303800046603	step
s: 181	lr: 4e-05	evaluation reward: 2.73		
1103				
episode: 1103	score: 5.0	memory length: 208745	epsilon: 0.7846829200046743	step
s: 327	lr: 4e-05	evaluation reward: 2.75		
1104				
episode: 1104	score: 2.0	memory length: 208942	epsilon: 0.7842928600046828	step
s: 197	lr: 4e-05	evaluation reward: 2.71		
1105				
episode: 1105	score: 6.0	memory length: 209313	epsilon: 0.7835582800046987	step
s: 371	lr: 4e-05	evaluation reward: 2.75		
1106				
episode: 1106	score: 2.0	memory length: 209493	epsilon: 0.7832018800047065	step
s: 180	lr: 4e-05	evaluation reward: 2.75		
1107				
episode: 1107	score: 0.0	memory length: 209616	epsilon: 0.7829583400047118	step
s: 123	lr: 4e-05	evaluation reward: 2.74		
1108				
episode: 1108	score: 4.0	memory length: 209937	epsilon: 0.7823227600047256	step
s: 321	lr: 4e-05	evaluation reward: 2.74		
1109				
episode: 1109	score: 1.0	memory length: 210088	epsilon: 0.782023780004732	step
s: 151	lr: 4e-05	evaluation reward: 2.73		
1110				
episode: 1110	score: 1.0	memory length: 210239	epsilon: 0.7817248000047385	step
s: 151	lr: 4e-05	evaluation reward: 2.72		
1111				
episode: 1111	score: 2.0	memory length: 210420	epsilon: 0.7813664200047463	step
s: 181	lr: 4e-05	evaluation reward: 2.74		
1112				
episode: 1112	score: 4.0	memory length: 210737	epsilon: 0.78073876000476	steps:
s: 317	lr: 4e-05	evaluation reward: 2.76		
1113				
episode: 1113	score: 5.0	memory length: 211083	epsilon: 0.7800536800047748	step
s: 346	lr: 4e-05	evaluation reward: 2.79		
1114				
episode: 1114	score: 5.0	memory length: 211408	epsilon: 0.7794101800047888	step
s: 325	lr: 4e-05	evaluation reward: 2.81		
1115				
episode: 1115	score: 4.0	memory length: 211703	epsilon: 0.7788260800048015	step
s: 295	lr: 4e-05	evaluation reward: 2.81		
1116				

episode: 1116	score: 3.0	memory length: 211952	epsilon: 0.7783330600048122	step
s: 249	lr: 4e-05	evaluation reward: 2.77		
1117				
episode: 1117	score: 3.0	memory length: 212178	epsilon: 0.7778855800048219	step
s: 226	lr: 4e-05	evaluation reward: 2.73		
1118				
episode: 1118	score: 1.0	memory length: 212328	epsilon: 0.7775885800048283	step
s: 150	lr: 4e-05	evaluation reward: 2.74		
1119				
episode: 1119	score: 4.0	memory length: 212588	epsilon: 0.7770737800048395	step
s: 260	lr: 4e-05	evaluation reward: 2.75		
1120				
episode: 1120	score: 4.0	memory length: 212847	epsilon: 0.7765609600048506	step
s: 259	lr: 4e-05	evaluation reward: 2.76		
1121				
episode: 1121	score: 4.0	memory length: 213145	epsilon: 0.7759709200048635	step
s: 298	lr: 4e-05	evaluation reward: 2.76		
1122				
episode: 1122	score: 1.0	memory length: 213296	epsilon: 0.7756719400048699	step
s: 151	lr: 4e-05	evaluation reward: 2.75		
1123				
episode: 1123	score: 3.0	memory length: 213542	epsilon: 0.7751848600048805	step
s: 246	lr: 4e-05	evaluation reward: 2.78		
1124				
episode: 1124	score: 8.0	memory length: 213983	epsilon: 0.7743116800048995	step
s: 441	lr: 4e-05	evaluation reward: 2.81		
1125				
episode: 1125	score: 1.0	memory length: 214152	epsilon: 0.7739770600049067	step
s: 169	lr: 4e-05	evaluation reward: 2.78		
1126				
episode: 1126	score: 3.0	memory length: 214381	epsilon: 0.7735236400049166	step
s: 229	lr: 4e-05	evaluation reward: 2.8		
1127				
episode: 1127	score: 4.0	memory length: 214676	epsilon: 0.7729395400049293	step
s: 295	lr: 4e-05	evaluation reward: 2.84		
1128				
episode: 1128	score: 3.0	memory length: 214922	epsilon: 0.7724524600049398	step
s: 246	lr: 4e-05	evaluation reward: 2.84		
1129				
episode: 1129	score: 4.0	memory length: 215218	epsilon: 0.7718663800049526	step
s: 296	lr: 4e-05	evaluation reward: 2.86		
1130				
episode: 1130	score: 2.0	memory length: 215416	epsilon: 0.7714743400049611	step
s: 198	lr: 4e-05	evaluation reward: 2.83		
1131				
episode: 1131	score: 2.0	memory length: 215597	epsilon: 0.7711159600049688	step
s: 181	lr: 4e-05	evaluation reward: 2.84		
1132				
episode: 1132	score: 3.0	memory length: 215847	epsilon: 0.7706209600049796	step
s: 250	lr: 4e-05	evaluation reward: 2.84		
1133				
episode: 1133	score: 5.0	memory length: 216190	epsilon: 0.7699418200049943	step
s: 343	lr: 4e-05	evaluation reward: 2.89		
1134				
episode: 1134	score: 2.0	memory length: 216370	epsilon: 0.7695854200050021	step
s: 180	lr: 4e-05	evaluation reward: 2.91		
1135				
episode: 1135	score: 6.0	memory length: 216713	epsilon: 0.7689062800050168	step
s: 343	lr: 4e-05	evaluation reward: 2.97		
1136				
episode: 1136	score: 2.0	memory length: 216912	epsilon: 0.7685122600050254	step
s: 199	lr: 4e-05	evaluation reward: 2.98		
1137				
episode: 1137	score: 2.0	memory length: 217110	epsilon: 0.7681202200050339	step
s: 198	lr: 4e-05	evaluation reward: 2.9		
1138				

episode: 1138	score: 5.0	memory length: 217455	epsilon: 0.7674371200050487	step
s: 345	lr: 4e-05	evaluation reward: 2.89		
1139				
episode: 1139	score: 3.0	memory length: 217702	epsilon: 0.7669480600050593	step
s: 247	lr: 4e-05	evaluation reward: 2.88		
1140				
episode: 1140	score: 3.0	memory length: 217969	epsilon: 0.7664194000050708	step
s: 267	lr: 4e-05	evaluation reward: 2.88		
1141				
episode: 1141	score: 2.0	memory length: 218169	epsilon: 0.7660234000050794	step
s: 200	lr: 4e-05	evaluation reward: 2.89		
1142				
episode: 1142	score: 2.0	memory length: 218387	epsilon: 0.7655917600050888	step
s: 218	lr: 4e-05	evaluation reward: 2.89		
1143				
episode: 1143	score: 2.0	memory length: 218585	epsilon: 0.7651997200050973	step
s: 198	lr: 4e-05	evaluation reward: 2.87		
1144				
episode: 1144	score: 2.0	memory length: 218786	epsilon: 0.7648017400051059	step
s: 201	lr: 4e-05	evaluation reward: 2.85		
1145				
episode: 1145	score: 4.0	memory length: 219069	epsilon: 0.7642414000051181	step
s: 283	lr: 4e-05	evaluation reward: 2.85		
1146				
episode: 1146	score: 2.0	memory length: 219267	epsilon: 0.7638493600051266	step
s: 198	lr: 4e-05	evaluation reward: 2.86		
1147				
episode: 1147	score: 4.0	memory length: 219526	epsilon: 0.7633365400051377	step
s: 259	lr: 4e-05	evaluation reward: 2.88		
1148				
episode: 1148	score: 2.0	memory length: 219726	epsilon: 0.7629405400051463	step
s: 200	lr: 4e-05	evaluation reward: 2.84		
1149				
episode: 1149	score: 4.0	memory length: 220043	epsilon: 0.76231288000516	steps:
317	lr: 4e-05	evaluation reward: 2.84		
1150				
episode: 1150	score: 2.0	memory length: 220241	epsilon: 0.7619208400051685	step
s: 198	lr: 4e-05	evaluation reward: 2.86		
1151				
episode: 1151	score: 8.0	memory length: 220677	epsilon: 0.7610575600051872	step
s: 436	lr: 4e-05	evaluation reward: 2.9		
1152				
episode: 1152	score: 5.0	memory length: 221022	epsilon: 0.760374460005202	step
s: 345	lr: 4e-05	evaluation reward: 2.9		
1153				
episode: 1153	score: 4.0	memory length: 221299	epsilon: 0.7598260000052139	step
s: 277	lr: 4e-05	evaluation reward: 2.89		
1154				
episode: 1154	score: 2.0	memory length: 221502	epsilon: 0.7594240600052227	step
s: 203	lr: 4e-05	evaluation reward: 2.88		
1155				
episode: 1155	score: 2.0	memory length: 221699	epsilon: 0.7590340000052311	step
s: 197	lr: 4e-05	evaluation reward: 2.86		
1156				
episode: 1156	score: 6.0	memory length: 222062	epsilon: 0.7583152600052467	step
s: 363	lr: 4e-05	evaluation reward: 2.92		
1157				
episode: 1157	score: 4.0	memory length: 222359	epsilon: 0.7577272000052595	step
s: 297	lr: 4e-05	evaluation reward: 2.95		
1158				
episode: 1158	score: 1.0	memory length: 222529	epsilon: 0.7573906000052668	step
s: 170	lr: 4e-05	evaluation reward: 2.96		
1159				
episode: 1159	score: 3.0	memory length: 222776	epsilon: 0.7569015400052774	step
s: 247	lr: 4e-05	evaluation reward: 2.96		
1160				

episode: 1160	score: 1.0	memory length: 222926	epsilon: 0.7566045400052839	step
s: 150	lr: 4e-05	evaluation reward: 2.94		
1161				
episode: 1161	score: 3.0	memory length: 223172	epsilon: 0.7561174600052945	step
s: 246	lr: 4e-05	evaluation reward: 2.97		
1162				
episode: 1162	score: 1.0	memory length: 223322	epsilon: 0.7558204600053009	step
s: 150	lr: 4e-05	evaluation reward: 2.95		
1163				
episode: 1163	score: 3.0	memory length: 223567	epsilon: 0.7553353600053114	step
s: 245	lr: 4e-05	evaluation reward: 2.93		
1164				
episode: 1164	score: 5.0	memory length: 223893	epsilon: 0.7546898800053254	step
s: 326	lr: 4e-05	evaluation reward: 2.95		
1165				
episode: 1165	score: 1.0	memory length: 224065	epsilon: 0.7543493200053328	step
s: 172	lr: 4e-05	evaluation reward: 2.93		
1166				
episode: 1166	score: 6.0	memory length: 224439	epsilon: 0.7536088000053489	step
s: 374	lr: 4e-05	evaluation reward: 2.96		
1167				
episode: 1167	score: 1.0	memory length: 224590	epsilon: 0.7533098200053554	step
s: 151	lr: 4e-05	evaluation reward: 2.95		
1168				
episode: 1168	score: 2.0	memory length: 224789	epsilon: 0.752915800005364	step
s: 199	lr: 4e-05	evaluation reward: 2.95		
1169				
episode: 1169	score: 6.0	memory length: 225127	epsilon: 0.7522465600053785	step
s: 338	lr: 4e-05	evaluation reward: 2.98		
1170				
episode: 1170	score: 3.0	memory length: 225391	epsilon: 0.7517238400053898	step
s: 264	lr: 4e-05	evaluation reward: 3.0		
1171				
episode: 1171	score: 2.0	memory length: 225588	epsilon: 0.7513337800053983	step
s: 197	lr: 4e-05	evaluation reward: 3.0		
1172				
episode: 1172	score: 3.0	memory length: 225818	epsilon: 0.7508783800054082	step
s: 230	lr: 4e-05	evaluation reward: 3.01		
1173				
episode: 1173	score: 0.0	memory length: 225941	epsilon: 0.7506348400054135	step
s: 123	lr: 4e-05	evaluation reward: 3.0		
1174				
episode: 1174	score: 6.0	memory length: 226293	epsilon: 0.7499378800054286	step
s: 352	lr: 4e-05	evaluation reward: 3.03		
1175				
episode: 1175	score: 4.0	memory length: 226587	epsilon: 0.7493557600054412	step
s: 294	lr: 4e-05	evaluation reward: 3.07		
1176				
episode: 1176	score: 9.0	memory length: 227087	epsilon: 0.7483657600054627	step
s: 500	lr: 4e-05	evaluation reward: 3.15		
1177				
episode: 1177	score: 3.0	memory length: 227334	epsilon: 0.7478767000054733	step
s: 247	lr: 4e-05	evaluation reward: 3.17		
1178				
episode: 1178	score: 6.0	memory length: 227671	epsilon: 0.7472094400054878	step
s: 337	lr: 4e-05	evaluation reward: 3.19		
1179				
episode: 1179	score: 3.0	memory length: 227882	epsilon: 0.7467916600054969	step
s: 211	lr: 4e-05	evaluation reward: 3.2		
1180				
episode: 1180	score: 1.0	memory length: 228033	epsilon: 0.7464926800055034	step
s: 151	lr: 4e-05	evaluation reward: 3.19		
1181				
episode: 1181	score: 3.0	memory length: 228282	epsilon: 0.7459996600055141	step
s: 249	lr: 4e-05	evaluation reward: 3.21		
1182				

episode: 1182	score: 3.0	memory length: 228510	epsilon: 0.7455482200055239	step
s: 228	lr: 4e-05	evaluation reward: 3.22		
1183				
episode: 1183	score: 4.0	memory length: 228770	epsilon: 0.7450334200055351	step
s: 260	lr: 4e-05	evaluation reward: 3.19		
1184				
episode: 1184	score: 4.0	memory length: 229045	epsilon: 0.7444889200055469	step
s: 275	lr: 4e-05	evaluation reward: 3.21		
1185				
episode: 1185	score: 4.0	memory length: 229345	epsilon: 0.7438949200055598	step
s: 300	lr: 4e-05	evaluation reward: 3.21		
1186				
episode: 1186	score: 2.0	memory length: 229543	epsilon: 0.7435028800055683	step
s: 198	lr: 4e-05	evaluation reward: 3.22		
1187				
episode: 1187	score: 2.0	memory length: 229743	epsilon: 0.7431068800055769	step
s: 200	lr: 4e-05	evaluation reward: 3.23		
1188				
episode: 1188	score: 5.0	memory length: 230087	epsilon: 0.7424257600055917	step
s: 344	lr: 4e-05	evaluation reward: 3.27		
1189				
episode: 1189	score: 2.0	memory length: 230305	epsilon: 0.741994120005601	step
s: 218	lr: 4e-05	evaluation reward: 3.27		
1190				
episode: 1190	score: 3.0	memory length: 230572	epsilon: 0.7414654600056125	step
s: 267	lr: 4e-05	evaluation reward: 3.25		
1191				
episode: 1191	score: 4.0	memory length: 230846	epsilon: 0.7409229400056243	step
s: 274	lr: 4e-05	evaluation reward: 3.23		
1192				
episode: 1192	score: 1.0	memory length: 230997	epsilon: 0.7406239600056308	step
s: 151	lr: 4e-05	evaluation reward: 3.23		
1193				
episode: 1193	score: 4.0	memory length: 231253	epsilon: 0.7401170800056418	step
s: 256	lr: 4e-05	evaluation reward: 3.22		
1194				
episode: 1194	score: 6.0	memory length: 231629	epsilon: 0.739372600005658	step
s: 376	lr: 4e-05	evaluation reward: 3.26		
1195				
episode: 1195	score: 3.0	memory length: 231877	epsilon: 0.7388815600056686	step
s: 248	lr: 4e-05	evaluation reward: 3.25		
1196				
episode: 1196	score: 5.0	memory length: 232169	epsilon: 0.7383034000056812	step
s: 292	lr: 4e-05	evaluation reward: 3.25		
1197				
episode: 1197	score: 6.0	memory length: 232509	epsilon: 0.7376302000056958	step
s: 340	lr: 4e-05	evaluation reward: 3.3		
1198				
episode: 1198	score: 6.0	memory length: 232850	epsilon: 0.7369550200057104	step
s: 341	lr: 4e-05	evaluation reward: 3.33		
1199				
episode: 1199	score: 1.0	memory length: 233021	epsilon: 0.7366164400057178	step
s: 171	lr: 4e-05	evaluation reward: 3.32		
1200				
episode: 1200	score: 4.0	memory length: 233316	epsilon: 0.7360323400057305	step
s: 295	lr: 4e-05	evaluation reward: 3.3		
1201				
episode: 1201	score: 5.0	memory length: 233681	epsilon: 0.7353096400057462	step
s: 365	lr: 4e-05	evaluation reward: 3.31		
1202				
episode: 1202	score: 2.0	memory length: 233861	epsilon: 0.7349532400057539	step
s: 180	lr: 4e-05	evaluation reward: 3.31		
1203				
episode: 1203	score: 3.0	memory length: 234107	epsilon: 0.7344661600057645	step
s: 246	lr: 4e-05	evaluation reward: 3.29		
1204				

episode: 1204	score: 2.0	memory length: 234289	epsilon: 0.7341058000057723	step
s: 182	lr: 4e-05	evaluation reward: 3.29		
1205				
episode: 1205	score: 4.0	memory length: 234563	epsilon: 0.7335632800057841	step
s: 274	lr: 4e-05	evaluation reward: 3.27		
1206				
episode: 1206	score: 4.0	memory length: 234861	epsilon: 0.7329732400057969	step
s: 298	lr: 4e-05	evaluation reward: 3.29		
1207				
episode: 1207	score: 2.0	memory length: 235061	epsilon: 0.7325772400058055	step
s: 200	lr: 4e-05	evaluation reward: 3.31		
1208				
episode: 1208	score: 1.0	memory length: 235211	epsilon: 0.7322802400058119	step
s: 150	lr: 4e-05	evaluation reward: 3.28		
1209				
episode: 1209	score: 5.0	memory length: 235517	epsilon: 0.7316743600058251	step
s: 306	lr: 4e-05	evaluation reward: 3.32		
1210				
episode: 1210	score: 4.0	memory length: 235794	epsilon: 0.731125900005837	step
s: 277	lr: 4e-05	evaluation reward: 3.35		
1211				
episode: 1211	score: 5.0	memory length: 236102	epsilon: 0.7305160600058502	step
s: 308	lr: 4e-05	evaluation reward: 3.38		
1212				
episode: 1212	score: 4.0	memory length: 236377	epsilon: 0.729971560005862	step
s: 275	lr: 4e-05	evaluation reward: 3.38		
1213				
episode: 1213	score: 3.0	memory length: 236622	epsilon: 0.7294864600058726	step
s: 245	lr: 4e-05	evaluation reward: 3.36		
1214				
episode: 1214	score: 4.0	memory length: 236899	epsilon: 0.7289380000058845	step
s: 277	lr: 4e-05	evaluation reward: 3.35		
1215				
episode: 1215	score: 4.0	memory length: 237143	epsilon: 0.728454880005895	step
s: 244	lr: 4e-05	evaluation reward: 3.35		
1216				
episode: 1216	score: 7.0	memory length: 237511	epsilon: 0.7277262400059108	step
s: 368	lr: 4e-05	evaluation reward: 3.39		
1217				
episode: 1217	score: 4.0	memory length: 237787	epsilon: 0.7271797600059227	step
s: 276	lr: 4e-05	evaluation reward: 3.4		
1218				
episode: 1218	score: 3.0	memory length: 238012	epsilon: 0.7267342600059323	step
s: 225	lr: 4e-05	evaluation reward: 3.42		
1219				
episode: 1219	score: 3.0	memory length: 238258	epsilon: 0.7262471800059429	step
s: 246	lr: 4e-05	evaluation reward: 3.41		
1220				
episode: 1220	score: 1.0	memory length: 238409	epsilon: 0.7259482000059494	step
s: 151	lr: 4e-05	evaluation reward: 3.38		
1221				
episode: 1221	score: 3.0	memory length: 238619	epsilon: 0.7255324000059584	step
s: 210	lr: 4e-05	evaluation reward: 3.37		
1222				
episode: 1222	score: 0.0	memory length: 238742	epsilon: 0.7252888600059637	step
s: 123	lr: 4e-05	evaluation reward: 3.36		
1223				
episode: 1223	score: 5.0	memory length: 239082	epsilon: 0.7246156600059783	step
s: 340	lr: 4e-05	evaluation reward: 3.38		
1224				
episode: 1224	score: 7.0	memory length: 239438	epsilon: 0.7239107800059936	step
s: 356	lr: 4e-05	evaluation reward: 3.37		
1225				
episode: 1225	score: 3.0	memory length: 239684	epsilon: 0.7234237000060042	step
s: 246	lr: 4e-05	evaluation reward: 3.39		
1226				

episode: 1226	score: 1.0	memory length: 239834	epsilon: 0.7231267000060106	step
s: 150	lr: 4e-05	evaluation reward: 3.37		
1227				
episode: 1227	score: 4.0	memory length: 240109	epsilon: 0.7225822000060225	step
s: 275	lr: 4e-05	evaluation reward: 3.37		
1228				
episode: 1228	score: 1.0	memory length: 240277	epsilon: 0.7222495600060297	step
s: 168	lr: 4e-05	evaluation reward: 3.35		
1229				
episode: 1229	score: 3.0	memory length: 240527	epsilon: 0.7217545600060404	step
s: 250	lr: 4e-05	evaluation reward: 3.34		
1230				
episode: 1230	score: 0.0	memory length: 240650	epsilon: 0.7215110200060457	step
s: 123	lr: 4e-05	evaluation reward: 3.32		
1231				
episode: 1231	score: 0.0	memory length: 240773	epsilon: 0.721267480006051	step
s: 123	lr: 4e-05	evaluation reward: 3.3		
1232				
episode: 1232	score: 3.0	memory length: 241019	epsilon: 0.7207804000060616	step
s: 246	lr: 4e-05	evaluation reward: 3.3		
1233				
episode: 1233	score: 1.0	memory length: 241190	epsilon: 0.7204418200060689	step
s: 171	lr: 4e-05	evaluation reward: 3.26		
1234				
episode: 1234	score: 6.0	memory length: 241547	epsilon: 0.7197349600060843	step
s: 357	lr: 4e-05	evaluation reward: 3.3		
1235				
episode: 1235	score: 4.0	memory length: 241862	epsilon: 0.7191112600060978	step
s: 315	lr: 4e-05	evaluation reward: 3.28		
1236				
episode: 1236	score: 7.0	memory length: 242223	epsilon: 0.7183964800061133	step
s: 361	lr: 4e-05	evaluation reward: 3.33		
1237				
episode: 1237	score: 5.0	memory length: 242566	epsilon: 0.7177173400061281	step
s: 343	lr: 4e-05	evaluation reward: 3.36		
1238				
episode: 1238	score: 2.0	memory length: 242746	epsilon: 0.7173609400061358	step
s: 180	lr: 4e-05	evaluation reward: 3.33		
1239				
episode: 1239	score: 8.0	memory length: 243189	epsilon: 0.7164838000061549	step
s: 443	lr: 4e-05	evaluation reward: 3.38		
1240				
episode: 1240	score: 4.0	memory length: 243483	epsilon: 0.7159016800061675	step
s: 294	lr: 4e-05	evaluation reward: 3.39		
1241				
episode: 1241	score: 2.0	memory length: 243663	epsilon: 0.7155452800061752	step
s: 180	lr: 4e-05	evaluation reward: 3.39		
1242				
episode: 1242	score: 3.0	memory length: 243891	epsilon: 0.715093840006185	step
s: 228	lr: 4e-05	evaluation reward: 3.4		
1243				
episode: 1243	score: 3.0	memory length: 244138	epsilon: 0.7146047800061957	step
s: 247	lr: 4e-05	evaluation reward: 3.41		
1244				
episode: 1244	score: 2.0	memory length: 244357	epsilon: 0.7141711600062051	step
s: 219	lr: 4e-05	evaluation reward: 3.41		
1245				
episode: 1245	score: 4.0	memory length: 244612	epsilon: 0.713666260006216	step
s: 255	lr: 4e-05	evaluation reward: 3.41		
1246				
episode: 1246	score: 4.0	memory length: 244904	epsilon: 0.7130881000062286	step
s: 292	lr: 4e-05	evaluation reward: 3.43		
1247				
episode: 1247	score: 6.0	memory length: 245243	epsilon: 0.7124168800062431	step
s: 339	lr: 4e-05	evaluation reward: 3.45		
1248				

episode: 1248	score: 7.0	memory length: 245647	epsilon: 0.7116169600062605	step
s: 404	lr: 4e-05	evaluation reward: 3.5		
1249				
episode: 1249	score: 3.0	memory length: 245857	epsilon: 0.7112011600062695	step
s: 210	lr: 4e-05	evaluation reward: 3.49		
1250				
episode: 1250	score: 5.0	memory length: 246166	epsilon: 0.7105893400062828	step
s: 309	lr: 4e-05	evaluation reward: 3.52		
1251				
episode: 1251	score: 4.0	memory length: 246442	epsilon: 0.7100428600062947	step
s: 276	lr: 4e-05	evaluation reward: 3.48		
1252				
episode: 1252	score: 2.0	memory length: 246622	epsilon: 0.7096864600063024	step
s: 180	lr: 4e-05	evaluation reward: 3.45		
1253				
episode: 1253	score: 1.0	memory length: 246773	epsilon: 0.7093874800063089	step
s: 151	lr: 4e-05	evaluation reward: 3.42		
1254				
episode: 1254	score: 3.0	memory length: 247020	epsilon: 0.7088984200063195	step
s: 247	lr: 4e-05	evaluation reward: 3.43		
1255				
episode: 1255	score: 6.0	memory length: 247363	epsilon: 0.7082192800063343	step
s: 343	lr: 4e-05	evaluation reward: 3.47		
1256				
episode: 1256	score: 3.0	memory length: 247591	epsilon: 0.7077678400063441	step
s: 228	lr: 4e-05	evaluation reward: 3.44		
1257				
episode: 1257	score: 4.0	memory length: 247866	epsilon: 0.7072233400063559	step
s: 275	lr: 4e-05	evaluation reward: 3.44		
1258				
episode: 1258	score: 4.0	memory length: 248142	epsilon: 0.7066768600063678	step
s: 276	lr: 4e-05	evaluation reward: 3.47		
1259				
episode: 1259	score: 1.0	memory length: 248293	epsilon: 0.7063778800063742	step
s: 151	lr: 4e-05	evaluation reward: 3.45		
1260				
episode: 1260	score: 6.0	memory length: 248632	epsilon: 0.7057066600063888	step
s: 339	lr: 4e-05	evaluation reward: 3.5		
1261				
episode: 1261	score: 5.0	memory length: 248978	epsilon: 0.7050215800064037	step
s: 346	lr: 4e-05	evaluation reward: 3.52		
1262				
episode: 1262	score: 13.0	memory length: 249445	epsilon: 0.7040969200064238	step
s: 467	lr: 4e-05	evaluation reward: 3.64		
1263				
episode: 1263	score: 1.0	memory length: 249596	epsilon: 0.7037979400064303	step
s: 151	lr: 4e-05	evaluation reward: 3.62		
1264				
episode: 1264	score: 5.0	memory length: 249919	epsilon: 0.7031584000064441	step
s: 323	lr: 4e-05	evaluation reward: 3.62		
1265				
episode: 1265	score: 5.0	memory length: 250246	epsilon: 0.7025109400064582	step
s: 327	lr: 4e-05	evaluation reward: 3.66		
1266				
episode: 1266	score: 4.0	memory length: 250522	epsilon: 0.7019644600064701	step
s: 276	lr: 4e-05	evaluation reward: 3.64		
1267				
episode: 1267	score: 4.0	memory length: 250797	epsilon: 0.7014199600064819	step
s: 275	lr: 4e-05	evaluation reward: 3.67		
1268				
episode: 1268	score: 6.0	memory length: 251157	epsilon: 0.7007071600064974	step
s: 360	lr: 4e-05	evaluation reward: 3.71		
1269				
episode: 1269	score: 2.0	memory length: 251337	epsilon: 0.7003507600065051	step
s: 180	lr: 4e-05	evaluation reward: 3.67		
1270				

episode: 1270	score: 4.0	memory length: 251633	epsilon: 0.6997646800065178	step
s: 296	lr: 4e-05	evaluation reward: 3.68		
1271				
episode: 1271	score: 1.0	memory length: 251785	epsilon: 0.6994637200065243	step
s: 152	lr: 4e-05	evaluation reward: 3.67		
1272				
episode: 1272	score: 2.0	memory length: 251965	epsilon: 0.6991073200065321	step
s: 180	lr: 4e-05	evaluation reward: 3.66		
1273				
episode: 1273	score: 6.0	memory length: 252324	epsilon: 0.6983965000065475	step
s: 359	lr: 4e-05	evaluation reward: 3.72		
1274				
episode: 1274	score: 3.0	memory length: 252555	epsilon: 0.6979391200065574	step
s: 231	lr: 4e-05	evaluation reward: 3.69		
1275				
episode: 1275	score: 2.0	memory length: 252736	epsilon: 0.6975807400065652	step
s: 181	lr: 4e-05	evaluation reward: 3.67		
1276				
episode: 1276	score: 4.0	memory length: 253015	epsilon: 0.6970283200065772	step
s: 279	lr: 4e-05	evaluation reward: 3.62		
1277				
episode: 1277	score: 6.0	memory length: 253352	epsilon: 0.6963610600065917	step
s: 337	lr: 4e-05	evaluation reward: 3.65		
1278				
episode: 1278	score: 6.0	memory length: 253726	epsilon: 0.6956205400066078	step
s: 374	lr: 4e-05	evaluation reward: 3.65		
1279				
episode: 1279	score: 5.0	memory length: 254063	epsilon: 0.6949532800066223	step
s: 337	lr: 4e-05	evaluation reward: 3.67		
1280				
episode: 1280	score: 5.0	memory length: 254411	epsilon: 0.6942642400066372	step
s: 348	lr: 4e-05	evaluation reward: 3.71		
1281				
episode: 1281	score: 3.0	memory length: 254637	epsilon: 0.6938167600066469	step
s: 226	lr: 4e-05	evaluation reward: 3.71		
1282				
episode: 1282	score: 6.0	memory length: 255012	epsilon: 0.6930742600066631	step
s: 375	lr: 4e-05	evaluation reward: 3.74		
1283				
episode: 1283	score: 2.0	memory length: 255209	epsilon: 0.6926842000066715	step
s: 197	lr: 4e-05	evaluation reward: 3.72		
1284				
episode: 1284	score: 7.0	memory length: 255639	epsilon: 0.69183280000669	steps:
430	lr: 4e-05	evaluation reward: 3.75		
1285				
episode: 1285	score: 4.0	memory length: 255916	epsilon: 0.6912843400067019	step
s: 277	lr: 4e-05	evaluation reward: 3.75		
1286				
episode: 1286	score: 2.0	memory length: 256113	epsilon: 0.6908942800067104	step
s: 197	lr: 4e-05	evaluation reward: 3.75		
1287				
episode: 1287	score: 5.0	memory length: 256439	epsilon: 0.6902488000067244	step
s: 326	lr: 4e-05	evaluation reward: 3.78		
1288				
episode: 1288	score: 3.0	memory length: 256668	epsilon: 0.6897953800067342	step
s: 229	lr: 4e-05	evaluation reward: 3.76		
1289				
episode: 1289	score: 3.0	memory length: 256894	epsilon: 0.689347900006744	step
s: 226	lr: 4e-05	evaluation reward: 3.77		
1290				
episode: 1290	score: 4.0	memory length: 257153	epsilon: 0.6888350800067551	step
s: 259	lr: 4e-05	evaluation reward: 3.78		
1291				
episode: 1291	score: 2.0	memory length: 257369	epsilon: 0.6884074000067644	step
s: 216	lr: 4e-05	evaluation reward: 3.76		
1292				

episode: 1292	score: 5.0	memory length: 257695	epsilon: 0.6877619200067784	step
s: 326	lr: 4e-05	evaluation reward: 3.8		
1293				
episode: 1293	score: 3.0	memory length: 257905	epsilon: 0.6873461200067874	step
s: 210	lr: 4e-05	evaluation reward: 3.79		
1294				
episode: 1294	score: 6.0	memory length: 258319	epsilon: 0.6865264000068052	step
s: 414	lr: 4e-05	evaluation reward: 3.79		
1295				
episode: 1295	score: 3.0	memory length: 258564	epsilon: 0.6860413000068157	step
s: 245	lr: 4e-05	evaluation reward: 3.79		
1296				
episode: 1296	score: 3.0	memory length: 258834	epsilon: 0.6855067000068273	step
s: 270	lr: 4e-05	evaluation reward: 3.77		
1297				
episode: 1297	score: 4.0	memory length: 259126	epsilon: 0.6849285400068399	step
s: 292	lr: 4e-05	evaluation reward: 3.75		
1298				
episode: 1298	score: 2.0	memory length: 259324	epsilon: 0.6845365000068484	step
s: 198	lr: 4e-05	evaluation reward: 3.71		
1299				
episode: 1299	score: 6.0	memory length: 259699	epsilon: 0.6837940000068645	step
s: 375	lr: 4e-05	evaluation reward: 3.76		
1300				
episode: 1300	score: 5.0	memory length: 260022	epsilon: 0.6831544600068784	step
s: 323	lr: 4e-05	evaluation reward: 3.77		
1301				
episode: 1301	score: 4.0	memory length: 260265	epsilon: 0.6826733200068889	step
s: 243	lr: 4e-05	evaluation reward: 3.76		
1302				
episode: 1302	score: 2.0	memory length: 260445	epsilon: 0.6823169200068966	step
s: 180	lr: 4e-05	evaluation reward: 3.76		
1303				
episode: 1303	score: 6.0	memory length: 260794	epsilon: 0.6816259000069116	step
s: 349	lr: 4e-05	evaluation reward: 3.79		
1304				
episode: 1304	score: 2.0	memory length: 260974	epsilon: 0.6812695000069193	step
s: 180	lr: 4e-05	evaluation reward: 3.79		
1305				
episode: 1305	score: 3.0	memory length: 261200	epsilon: 0.680822020006929	step
s: 226	lr: 4e-05	evaluation reward: 3.78		
1306				
episode: 1306	score: 5.0	memory length: 261524	epsilon: 0.680180500006943	step
s: 324	lr: 4e-05	evaluation reward: 3.79		
1307				
episode: 1307	score: 5.0	memory length: 261851	epsilon: 0.679533040006957	step
s: 327	lr: 4e-05	evaluation reward: 3.82		
1308				
episode: 1308	score: 5.0	memory length: 262158	epsilon: 0.6789251800069702	step
s: 307	lr: 4e-05	evaluation reward: 3.86		
1309				
episode: 1309	score: 3.0	memory length: 262387	epsilon: 0.6784717600069801	step
s: 229	lr: 4e-05	evaluation reward: 3.84		
1310				
episode: 1310	score: 13.0	memory length: 262876	epsilon: 0.6775035400070011	ste
ps: 489	lr: 4e-05	evaluation reward: 3.93		
1311				
episode: 1311	score: 3.0	memory length: 263086	epsilon: 0.6770877400070101	step
s: 210	lr: 4e-05	evaluation reward: 3.91		
1312				
episode: 1312	score: 3.0	memory length: 263334	epsilon: 0.6765967000070208	step
s: 248	lr: 4e-05	evaluation reward: 3.9		
1313				
episode: 1313	score: 3.0	memory length: 263564	epsilon: 0.6761413000070307	step
s: 230	lr: 4e-05	evaluation reward: 3.9		
1314				

episode: 1314	score: 3.0	memory length: 263828	epsilon: 0.675618580007042	step
s: 264	lr: 4e-05	evaluation reward: 3.89		
1315				
episode: 1315	score: 1.0	memory length: 263997	epsilon: 0.6752839600070493	step
s: 169	lr: 4e-05	evaluation reward: 3.86		
1316				
episode: 1316	score: 3.0	memory length: 264223	epsilon: 0.674836480007059	step
s: 226	lr: 4e-05	evaluation reward: 3.82		
1317				
episode: 1317	score: 1.0	memory length: 264375	epsilon: 0.6745355200070655	step
s: 152	lr: 4e-05	evaluation reward: 3.79		
1318				
episode: 1318	score: 0.0	memory length: 264497	epsilon: 0.6742939600070708	step
s: 122	lr: 4e-05	evaluation reward: 3.76		
1319				
episode: 1319	score: 3.0	memory length: 264742	epsilon: 0.6738088600070813	step
s: 245	lr: 4e-05	evaluation reward: 3.76		
1320				
episode: 1320	score: 7.0	memory length: 265170	epsilon: 0.6729614200070997	step
s: 428	lr: 4e-05	evaluation reward: 3.82		
1321				
episode: 1321	score: 5.0	memory length: 265497	epsilon: 0.6723139600071137	step
s: 327	lr: 4e-05	evaluation reward: 3.84		
1322				
episode: 1322	score: 2.0	memory length: 265695	epsilon: 0.6719219200071223	step
s: 198	lr: 4e-05	evaluation reward: 3.86		
1323				
episode: 1323	score: 2.0	memory length: 265893	epsilon: 0.6715298800071308	step
s: 198	lr: 4e-05	evaluation reward: 3.83		
1324				
episode: 1324	score: 6.0	memory length: 266228	epsilon: 0.6708665800071452	step
s: 335	lr: 4e-05	evaluation reward: 3.82		
1325				
episode: 1325	score: 3.0	memory length: 266494	epsilon: 0.6703399000071566	step
s: 266	lr: 4e-05	evaluation reward: 3.82		
1326				
episode: 1326	score: 2.0	memory length: 266692	epsilon: 0.6699478600071651	step
s: 198	lr: 4e-05	evaluation reward: 3.83		
1327				
episode: 1327	score: 2.0	memory length: 266889	epsilon: 0.6695578000071736	step
s: 197	lr: 4e-05	evaluation reward: 3.81		
1328				
episode: 1328	score: 2.0	memory length: 267069	epsilon: 0.6692014000071813	step
s: 180	lr: 4e-05	evaluation reward: 3.82		
1329				
episode: 1329	score: 5.0	memory length: 267388	epsilon: 0.668569780007195	step
s: 319	lr: 4e-05	evaluation reward: 3.84		
1330				
episode: 1330	score: 4.0	memory length: 267705	epsilon: 0.6679421200072087	step
s: 317	lr: 4e-05	evaluation reward: 3.88		
1331				
episode: 1331	score: 5.0	memory length: 268009	epsilon: 0.6673402000072217	step
s: 304	lr: 4e-05	evaluation reward: 3.93		
1332				
episode: 1332	score: 5.0	memory length: 268321	epsilon: 0.6667224400072351	step
s: 312	lr: 4e-05	evaluation reward: 3.95		
1333				
episode: 1333	score: 4.0	memory length: 268616	epsilon: 0.6661383400072478	step
s: 295	lr: 4e-05	evaluation reward: 3.98		
1334				
episode: 1334	score: 5.0	memory length: 268939	epsilon: 0.6654988000072617	step
s: 323	lr: 4e-05	evaluation reward: 3.97		
1335				
episode: 1335	score: 3.0	memory length: 269169	epsilon: 0.6650434000072716	step
s: 230	lr: 4e-05	evaluation reward: 3.96		
1336				

episode: 1336	score: 2.0	memory length: 269369	epsilon: 0.6646474000072802	step
s: 200	lr: 4e-05	evaluation reward: 3.91		
1337				
episode: 1337	score: 4.0	memory length: 269626	epsilon: 0.6641385400072912	step
s: 257	lr: 4e-05	evaluation reward: 3.9		
1338				
episode: 1338	score: 2.0	memory length: 269844	epsilon: 0.6637069000073006	step
s: 218	lr: 4e-05	evaluation reward: 3.9		
1339				
episode: 1339	score: 10.0	memory length: 270245	epsilon: 0.6629129200073178	ste
ps: 401	lr: 4e-05	evaluation reward: 3.92		
1340				
episode: 1340	score: 3.0	memory length: 270494	epsilon: 0.6624199000073285	step
s: 249	lr: 4e-05	evaluation reward: 3.91		
1341				
episode: 1341	score: 4.0	memory length: 270769	epsilon: 0.6618754000073404	step
s: 275	lr: 4e-05	evaluation reward: 3.93		
1342				
episode: 1342	score: 1.0	memory length: 270939	epsilon: 0.6615388000073477	step
s: 170	lr: 4e-05	evaluation reward: 3.91		
1343				
episode: 1343	score: 4.0	memory length: 271216	epsilon: 0.6609903400073596	step
s: 277	lr: 4e-05	evaluation reward: 3.92		
1344				
episode: 1344	score: 6.0	memory length: 271573	epsilon: 0.6602834800073749	step
s: 357	lr: 4e-05	evaluation reward: 3.96		
1345				
episode: 1345	score: 4.0	memory length: 271871	epsilon: 0.6596934400073877	step
s: 298	lr: 4e-05	evaluation reward: 3.96		
1346				
episode: 1346	score: 8.0	memory length: 272326	epsilon: 0.6587925400074073	step
s: 455	lr: 4e-05	evaluation reward: 4.0		
1347				
episode: 1347	score: 4.0	memory length: 272568	epsilon: 0.6583133800074177	step
s: 242	lr: 4e-05	evaluation reward: 3.98		
1348				
episode: 1348	score: 4.0	memory length: 272843	epsilon: 0.6577688800074295	step
s: 275	lr: 4e-05	evaluation reward: 3.95		
1349				
episode: 1349	score: 1.0	memory length: 272994	epsilon: 0.657469900007436	step
s: 151	lr: 4e-05	evaluation reward: 3.93		
1350				
episode: 1350	score: 7.0	memory length: 273383	epsilon: 0.6566996800074527	step
s: 389	lr: 4e-05	evaluation reward: 3.95		
1351				
episode: 1351	score: 0.0	memory length: 273506	epsilon: 0.656456140007458	step
s: 123	lr: 4e-05	evaluation reward: 3.91		
1352				
episode: 1352	score: 2.0	memory length: 273704	epsilon: 0.6560641000074665	step
s: 198	lr: 4e-05	evaluation reward: 3.91		
1353				
episode: 1353	score: 3.0	memory length: 273930	epsilon: 0.6556166200074762	step
s: 226	lr: 4e-05	evaluation reward: 3.93		
1354				
episode: 1354	score: 7.0	memory length: 274355	epsilon: 0.6547751200074945	step
s: 425	lr: 4e-05	evaluation reward: 3.97		
1355				
episode: 1355	score: 5.0	memory length: 274663	epsilon: 0.6541652800075077	step
s: 308	lr: 4e-05	evaluation reward: 3.96		
1356				
episode: 1356	score: 3.0	memory length: 274888	epsilon: 0.6537197800075174	step
s: 225	lr: 4e-05	evaluation reward: 3.96		
1357				
episode: 1357	score: 2.0	memory length: 275088	epsilon: 0.653323780007526	step
s: 200	lr: 4e-05	evaluation reward: 3.94		
1358				

episode: 1358	score: 5.0	memory length: 275393	epsilon: 0.6527198800075391	step
s: 305	lr: 4e-05	evaluation reward: 3.95		
1359				
episode: 1359	score: 2.0	memory length: 275590	epsilon: 0.6523298200075476	step
s: 197	lr: 4e-05	evaluation reward: 3.96		
1360				
episode: 1360	score: 8.0	memory length: 276005	epsilon: 0.6515081200075654	step
s: 415	lr: 4e-05	evaluation reward: 3.98		
1361				
episode: 1361	score: 5.0	memory length: 276348	epsilon: 0.6508289800075802	step
s: 343	lr: 4e-05	evaluation reward: 3.98		
1362				
episode: 1362	score: 2.0	memory length: 276530	epsilon: 0.650468620007588	step
s: 182	lr: 4e-05	evaluation reward: 3.87		
1363				
episode: 1363	score: 6.0	memory length: 276881	epsilon: 0.6497736400076031	step
s: 351	lr: 4e-05	evaluation reward: 3.92		
1364				
episode: 1364	score: 3.0	memory length: 277108	epsilon: 0.6493241800076128	step
s: 227	lr: 4e-05	evaluation reward: 3.9		
1365				
episode: 1365	score: 8.0	memory length: 277537	epsilon: 0.6484747600076313	step
s: 429	lr: 4e-05	evaluation reward: 3.93		
1366				
episode: 1366	score: 5.0	memory length: 277840	epsilon: 0.6478748200076443	step
s: 303	lr: 4e-05	evaluation reward: 3.94		
1367				
episode: 1367	score: 3.0	memory length: 278110	epsilon: 0.6473402200076559	step
s: 270	lr: 4e-05	evaluation reward: 3.93		
1368				
episode: 1368	score: 2.0	memory length: 278308	epsilon: 0.6469481800076644	step
s: 198	lr: 4e-05	evaluation reward: 3.89		
1369				
episode: 1369	score: 7.0	memory length: 278712	epsilon: 0.6461482600076818	step
s: 404	lr: 4e-05	evaluation reward: 3.94		
1370				
episode: 1370	score: 5.0	memory length: 279002	epsilon: 0.6455740600076942	step
s: 290	lr: 4e-05	evaluation reward: 3.95		
1371				
episode: 1371	score: 5.0	memory length: 279292	epsilon: 0.6449998600077067	step
s: 290	lr: 4e-05	evaluation reward: 3.99		
1372				
episode: 1372	score: 4.0	memory length: 279586	epsilon: 0.6444177400077193	step
s: 294	lr: 4e-05	evaluation reward: 4.01		
1373				
episode: 1373	score: 5.0	memory length: 279895	epsilon: 0.6438059200077326	step
s: 309	lr: 4e-05	evaluation reward: 4.0		
1374				
episode: 1374	score: 4.0	memory length: 280171	epsilon: 0.6432594400077445	step
s: 276	lr: 4e-05	evaluation reward: 4.01		
1375				
episode: 1375	score: 3.0	memory length: 280382	epsilon: 0.6428416600077536	step
s: 211	lr: 4e-05	evaluation reward: 4.02		
1376				
episode: 1376	score: 8.0	memory length: 280679	epsilon: 0.6422536000077663	step
s: 297	lr: 4e-05	evaluation reward: 4.06		
1377				
episode: 1377	score: 4.0	memory length: 280936	epsilon: 0.6417447400077774	step
s: 257	lr: 4e-05	evaluation reward: 4.04		
1378				
episode: 1378	score: 5.0	memory length: 281230	epsilon: 0.64116262000779	steps:
294	lr: 4e-05	evaluation reward: 4.03		
1379				
episode: 1379	score: 5.0	memory length: 281581	epsilon: 0.6404676400078051	step
s: 351	lr: 4e-05	evaluation reward: 4.03		
1380				

episode: 1380	score: 5.0	memory length: 281871	epsilon: 0.6398934400078176	step
s: 290	lr: 4e-05	evaluation reward: 4.03		
1381				
episode: 1381	score: 6.0	memory length: 282267	epsilon: 0.6391093600078346	step
s: 396	lr: 4e-05	evaluation reward: 4.06		
1382				
episode: 1382	score: 4.0	memory length: 282544	epsilon: 0.6385609000078465	step
s: 277	lr: 4e-05	evaluation reward: 4.04		
1383				
episode: 1383	score: 2.0	memory length: 282761	epsilon: 0.6381312400078558	step
s: 217	lr: 4e-05	evaluation reward: 4.04		
1384				
episode: 1384	score: 4.0	memory length: 283023	epsilon: 0.6376124800078671	step
s: 262	lr: 4e-05	evaluation reward: 4.01		
1385				
episode: 1385	score: 5.0	memory length: 283331	epsilon: 0.6370026400078803	step
s: 308	lr: 4e-05	evaluation reward: 4.02		
1386				
episode: 1386	score: 4.0	memory length: 283597	epsilon: 0.6364759600078918	step
s: 266	lr: 4e-05	evaluation reward: 4.04		
1387				
episode: 1387	score: 2.0	memory length: 283797	epsilon: 0.6360799600079003	step
s: 200	lr: 4e-05	evaluation reward: 4.01		
1388				
episode: 1388	score: 2.0	memory length: 283995	epsilon: 0.6356879200079089	step
s: 198	lr: 4e-05	evaluation reward: 4.0		
1389				
episode: 1389	score: 7.0	memory length: 284399	epsilon: 0.6348880000079262	step
s: 404	lr: 4e-05	evaluation reward: 4.04		
1390				
episode: 1390	score: 3.0	memory length: 284627	epsilon: 0.634436560007936	step
s: 228	lr: 4e-05	evaluation reward: 4.03		
1391				
episode: 1391	score: 6.0	memory length: 285003	epsilon: 0.6336920800079522	step
s: 376	lr: 4e-05	evaluation reward: 4.07		
1392				
episode: 1392	score: 3.0	memory length: 285214	epsilon: 0.6332743000079613	step
s: 211	lr: 4e-05	evaluation reward: 4.05		
1393				
episode: 1393	score: 6.0	memory length: 285611	epsilon: 0.6324882400079783	step
s: 397	lr: 4e-05	evaluation reward: 4.08		
1394				
episode: 1394	score: 3.0	memory length: 285860	epsilon: 0.631995220007989	step
s: 249	lr: 4e-05	evaluation reward: 4.05		
1395				
episode: 1395	score: 3.0	memory length: 286086	epsilon: 0.6315477400079987	step
s: 226	lr: 4e-05	evaluation reward: 4.05		
1396				
episode: 1396	score: 4.0	memory length: 286376	epsilon: 0.6309735400080112	step
s: 290	lr: 4e-05	evaluation reward: 4.06		
1397				
episode: 1397	score: 4.0	memory length: 286641	epsilon: 0.6304488400080226	step
s: 265	lr: 4e-05	evaluation reward: 4.06		
1398				
episode: 1398	score: 3.0	memory length: 286852	epsilon: 0.6300310600080317	step
s: 211	lr: 4e-05	evaluation reward: 4.07		
1399				
episode: 1399	score: 5.0	memory length: 287162	epsilon: 0.629417260008045	step
s: 310	lr: 4e-05	evaluation reward: 4.06		
1400				
episode: 1400	score: 4.0	memory length: 287439	epsilon: 0.6288688000080569	step
s: 277	lr: 4e-05	evaluation reward: 4.05		
1401				
episode: 1401	score: 5.0	memory length: 287762	epsilon: 0.6282292600080708	step
s: 323	lr: 4e-05	evaluation reward: 4.06		
1402				

episode: 1402	score: 6.0	memory length: 288157	epsilon: 0.6274471600080878	step
s: 395	lr: 4e-05	evaluation reward: 4.1		
1403				
episode: 1403	score: 1.0	memory length: 288308	epsilon: 0.6271481800080942	step
s: 151	lr: 4e-05	evaluation reward: 4.05		
1404				
episode: 1404	score: 6.0	memory length: 288686	epsilon: 0.6263997400081105	step
s: 378	lr: 4e-05	evaluation reward: 4.09		
1405				
episode: 1405	score: 5.0	memory length: 289023	epsilon: 0.625732480008125	step
s: 337	lr: 4e-05	evaluation reward: 4.11		
1406				
episode: 1406	score: 4.0	memory length: 289316	epsilon: 0.6251523400081376	step
s: 293	lr: 4e-05	evaluation reward: 4.1		
1407				
episode: 1407	score: 3.0	memory length: 289541	epsilon: 0.6247068400081472	step
s: 225	lr: 4e-05	evaluation reward: 4.08		
1408				
episode: 1408	score: 8.0	memory length: 289968	epsilon: 0.6238613800081656	step
s: 427	lr: 4e-05	evaluation reward: 4.11		
1409				
episode: 1409	score: 9.0	memory length: 290457	epsilon: 0.6228931600081866	step
s: 489	lr: 4e-05	evaluation reward: 4.17		
1410				
episode: 1410	score: 1.0	memory length: 290626	epsilon: 0.6225585400081939	step
s: 169	lr: 4e-05	evaluation reward: 4.05		
1411				
episode: 1411	score: 5.0	memory length: 290950	epsilon: 0.6219170200082078	step
s: 324	lr: 4e-05	evaluation reward: 4.07		
1412				
episode: 1412	score: 7.0	memory length: 291339	epsilon: 0.6211468000082245	step
s: 389	lr: 4e-05	evaluation reward: 4.11		
1413				
episode: 1413	score: 3.0	memory length: 291569	epsilon: 0.6206914000082344	step
s: 230	lr: 4e-05	evaluation reward: 4.11		
1414				
episode: 1414	score: 3.0	memory length: 291815	epsilon: 0.620204320008245	step
s: 246	lr: 4e-05	evaluation reward: 4.11		
1415				
episode: 1415	score: 5.0	memory length: 292121	epsilon: 0.6195984400082581	step
s: 306	lr: 4e-05	evaluation reward: 4.15		
1416				
episode: 1416	score: 6.0	memory length: 292456	epsilon: 0.6189351400082725	step
s: 335	lr: 4e-05	evaluation reward: 4.18		
1417				
episode: 1417	score: 5.0	memory length: 292782	epsilon: 0.6182896600082866	step
s: 326	lr: 4e-05	evaluation reward: 4.22		
1418				
episode: 1418	score: 5.0	memory length: 293128	epsilon: 0.6176045800083014	step
s: 346	lr: 4e-05	evaluation reward: 4.27		
1419				
episode: 1419	score: 9.0	memory length: 293554	epsilon: 0.6167611000083197	step
s: 426	lr: 4e-05	evaluation reward: 4.33		
1420				
episode: 1420	score: 7.0	memory length: 293929	epsilon: 0.6160186000083359	step
s: 375	lr: 4e-05	evaluation reward: 4.33		
1421				
episode: 1421	score: 3.0	memory length: 294196	epsilon: 0.6154899400083473	step
s: 267	lr: 4e-05	evaluation reward: 4.31		
1422				
episode: 1422	score: 4.0	memory length: 294492	epsilon: 0.6149038600083601	step
s: 296	lr: 4e-05	evaluation reward: 4.33		
1423				
episode: 1423	score: 6.0	memory length: 294889	epsilon: 0.6141178000083771	step
s: 397	lr: 4e-05	evaluation reward: 4.37		
1424				

episode: 1424	score: 2.0	memory length: 295071	epsilon: 0.613757440008385	step
s: 182	lr: 4e-05	evaluation reward: 4.33		
1425				
episode: 1425	score: 3.0	memory length: 295280	epsilon: 0.6133436200083939	step
s: 209	lr: 4e-05	evaluation reward: 4.33		
1426				
episode: 1426	score: 4.0	memory length: 295540	epsilon: 0.6128288200084051	step
s: 260	lr: 4e-05	evaluation reward: 4.35		
1427				
episode: 1427	score: 5.0	memory length: 295850	epsilon: 0.6122150200084184	step
s: 310	lr: 4e-05	evaluation reward: 4.38		
1428				
episode: 1428	score: 5.0	memory length: 296156	epsilon: 0.6116091400084316	step
s: 306	lr: 4e-05	evaluation reward: 4.41		
1429				
episode: 1429	score: 6.0	memory length: 296513	epsilon: 0.6109022800084469	step
s: 357	lr: 4e-05	evaluation reward: 4.42		
1430				
episode: 1430	score: 11.0	memory length: 297086	epsilon: 0.6097677400084716	step
s: 573	lr: 4e-05	evaluation reward: 4.49		
1431				
episode: 1431	score: 6.0	memory length: 297438	epsilon: 0.6090707800084867	step
s: 352	lr: 4e-05	evaluation reward: 4.5		
1432				
episode: 1432	score: 4.0	memory length: 297696	epsilon: 0.6085599400084978	step
s: 258	lr: 4e-05	evaluation reward: 4.49		
1433				
episode: 1433	score: 3.0	memory length: 297907	epsilon: 0.6081421600085068	step
s: 211	lr: 4e-05	evaluation reward: 4.48		
1434				
episode: 1434	score: 3.0	memory length: 298136	epsilon: 0.6076887400085167	step
s: 229	lr: 4e-05	evaluation reward: 4.46		
1435				
episode: 1435	score: 6.0	memory length: 298473	epsilon: 0.6070214800085312	step
s: 337	lr: 4e-05	evaluation reward: 4.49		
1436				
episode: 1436	score: 4.0	memory length: 298768	epsilon: 0.6064373800085439	step
s: 295	lr: 4e-05	evaluation reward: 4.51		
1437				
episode: 1437	score: 4.0	memory length: 299012	epsilon: 0.6059542600085543	step
s: 244	lr: 4e-05	evaluation reward: 4.51		
1438				
episode: 1438	score: 6.0	memory length: 299351	epsilon: 0.6052830400085689	step
s: 339	lr: 4e-05	evaluation reward: 4.55		
1439				
episode: 1439	score: 4.0	memory length: 299626	epsilon: 0.6047385400085807	step
s: 275	lr: 4e-05	evaluation reward: 4.49		
1440				
episode: 1440	score: 1.0	memory length: 299776	epsilon: 0.6044415400085872	step
s: 150	lr: 4e-05	evaluation reward: 4.47		
1441				
episode: 1441	score: 4.0	memory length: 300036	epsilon: 0.6039267400085984	step
s: 260	lr: 1.6000000000000003e-05	evaluation reward: 4.47		
1442				
episode: 1442	score: 2.0	memory length: 300233	epsilon: 0.6035366800086068	step
s: 197	lr: 1.6000000000000003e-05	evaluation reward: 4.48		
1443				
episode: 1443	score: 5.0	memory length: 300578	epsilon: 0.6028535800086217	step
s: 345	lr: 1.6000000000000003e-05	evaluation reward: 4.49		
1444				
episode: 1444	score: 4.0	memory length: 300853	epsilon: 0.6023090800086335	step
s: 275	lr: 1.6000000000000003e-05	evaluation reward: 4.47		
1445				
episode: 1445	score: 5.0	memory length: 301143	epsilon: 0.601734880008646	step
s: 290	lr: 1.6000000000000003e-05	evaluation reward: 4.48		
1446				

episode: 1446 score: 7.0 memory length: 301514 epsilon: 0.6010003000086619 step
s: 371 lr: 1.6000000000000003e-05 evaluation reward: 4.47
1447
episode: 1447 score: 4.0 memory length: 301755 epsilon: 0.6005231200086723 step
s: 241 lr: 1.6000000000000003e-05 evaluation reward: 4.47
1448
episode: 1448 score: 5.0 memory length: 302105 epsilon: 0.5998301200086873 step
s: 350 lr: 1.6000000000000003e-05 evaluation reward: 4.48
1449
episode: 1449 score: 3.0 memory length: 302371 epsilon: 0.5993034400086987 step
s: 266 lr: 1.6000000000000003e-05 evaluation reward: 4.5
1450
episode: 1450 score: 3.0 memory length: 302584 epsilon: 0.5988817000087079 step
s: 213 lr: 1.6000000000000003e-05 evaluation reward: 4.46
1451
episode: 1451 score: 4.0 memory length: 302879 epsilon: 0.5982976000087206 step
s: 295 lr: 1.6000000000000003e-05 evaluation reward: 4.5
1452
episode: 1452 score: 7.0 memory length: 303265 epsilon: 0.5975333200087372 step
s: 386 lr: 1.6000000000000003e-05 evaluation reward: 4.55
1453
episode: 1453 score: 4.0 memory length: 303525 epsilon: 0.5970185200087483 step
s: 260 lr: 1.6000000000000003e-05 evaluation reward: 4.56
1454
episode: 1454 score: 5.0 memory length: 303869 epsilon: 0.5963374000087631 step
s: 344 lr: 1.6000000000000003e-05 evaluation reward: 4.54
1455
episode: 1455 score: 1.0 memory length: 304019 epsilon: 0.5960404000087696 step
s: 150 lr: 1.6000000000000003e-05 evaluation reward: 4.5
1456
episode: 1456 score: 5.0 memory length: 304343 epsilon: 0.5953988800087835 step
s: 324 lr: 1.6000000000000003e-05 evaluation reward: 4.52
1457
episode: 1457 score: 3.0 memory length: 304556 epsilon: 0.5949771400087926 step
s: 213 lr: 1.6000000000000003e-05 evaluation reward: 4.53
1458
episode: 1458 score: 4.0 memory length: 304818 epsilon: 0.5944583800088039 step
s: 262 lr: 1.6000000000000003e-05 evaluation reward: 4.52
1459
episode: 1459 score: 4.0 memory length: 305092 epsilon: 0.5939158600088157 step
s: 274 lr: 1.6000000000000003e-05 evaluation reward: 4.54
1460
episode: 1460 score: 1.0 memory length: 305243 epsilon: 0.5936168800088222 step
s: 151 lr: 1.6000000000000003e-05 evaluation reward: 4.47
1461
episode: 1461 score: 3.0 memory length: 305455 epsilon: 0.5931971200088313 step
s: 212 lr: 1.6000000000000003e-05 evaluation reward: 4.45
1462
episode: 1462 score: 6.0 memory length: 305834 epsilon: 0.5924467000088476 step
s: 379 lr: 1.6000000000000003e-05 evaluation reward: 4.49
1463
episode: 1463 score: 5.0 memory length: 306148 epsilon: 0.5918249800088611 step
s: 314 lr: 1.6000000000000003e-05 evaluation reward: 4.48
1464
episode: 1464 score: 14.0 memory length: 306720 epsilon: 0.5906924200088857 step
s: 572 lr: 1.6000000000000003e-05 evaluation reward: 4.59
1465
episode: 1465 score: 7.0 memory length: 307120 epsilon: 0.5899004200089029 step
s: 400 lr: 1.6000000000000003e-05 evaluation reward: 4.58
1466
episode: 1466 score: 2.0 memory length: 307320 epsilon: 0.5895044200089115 step
s: 200 lr: 1.6000000000000003e-05 evaluation reward: 4.55
1467
episode: 1467 score: 1.0 memory length: 307470 epsilon: 0.5892074200089179 step
s: 150 lr: 1.6000000000000003e-05 evaluation reward: 4.53
1468

episode: 1468 score: 8.0 memory length: 307887 epsilon: 0.5883817600089358 step
s: 417 lr: 1.6000000000000003e-05 evaluation reward: 4.59
1469
episode: 1469 score: 6.0 memory length: 308236 epsilon: 0.5876907400089508 step
s: 349 lr: 1.6000000000000003e-05 evaluation reward: 4.58
1470
episode: 1470 score: 4.0 memory length: 308535 epsilon: 0.5870987200089637 step
s: 299 lr: 1.6000000000000003e-05 evaluation reward: 4.57
1471
episode: 1471 score: 4.0 memory length: 308831 epsilon: 0.5865126400089764 step
s: 296 lr: 1.6000000000000003e-05 evaluation reward: 4.56
1472
episode: 1472 score: 1.0 memory length: 308982 epsilon: 0.5862136600089829 step
s: 151 lr: 1.6000000000000003e-05 evaluation reward: 4.53
1473
episode: 1473 score: 5.0 memory length: 309306 epsilon: 0.5855721400089968 step
s: 324 lr: 1.6000000000000003e-05 evaluation reward: 4.53
1474
episode: 1474 score: 2.0 memory length: 309488 epsilon: 0.5852117800090046 step
s: 182 lr: 1.6000000000000003e-05 evaluation reward: 4.51
1475
episode: 1475 score: 5.0 memory length: 309796 epsilon: 0.5846019400090179 step
s: 308 lr: 1.6000000000000003e-05 evaluation reward: 4.53
1476
episode: 1476 score: 10.0 memory length: 310246 epsilon: 0.5837109400090372 ste
ps: 450 lr: 1.6000000000000003e-05 evaluation reward: 4.55
1477
episode: 1477 score: 2.0 memory length: 310462 epsilon: 0.5832832600090465 step
s: 216 lr: 1.6000000000000003e-05 evaluation reward: 4.53
1478
episode: 1478 score: 5.0 memory length: 310766 epsilon: 0.5826813400090596 step
s: 304 lr: 1.6000000000000003e-05 evaluation reward: 4.53
1479
episode: 1479 score: 6.0 memory length: 311128 epsilon: 0.5819645800090751 step
s: 362 lr: 1.6000000000000003e-05 evaluation reward: 4.54
1480
episode: 1480 score: 7.0 memory length: 311539 epsilon: 0.5811508000090928 step
s: 411 lr: 1.6000000000000003e-05 evaluation reward: 4.56
1481
episode: 1481 score: 2.0 memory length: 311737 epsilon: 0.5807587600091013 step
s: 198 lr: 1.6000000000000003e-05 evaluation reward: 4.52
1482
episode: 1482 score: 6.0 memory length: 312106 epsilon: 0.5800281400091172 step
s: 369 lr: 1.6000000000000003e-05 evaluation reward: 4.54
1483
episode: 1483 score: 1.0 memory length: 312257 epsilon: 0.5797291600091237 step
s: 151 lr: 1.6000000000000003e-05 evaluation reward: 4.53
1484
episode: 1484 score: 6.0 memory length: 312576 epsilon: 0.5790975400091374 step
s: 319 lr: 1.6000000000000003e-05 evaluation reward: 4.55
1485
episode: 1485 score: 1.0 memory length: 312745 epsilon: 0.5787629200091446 step
s: 169 lr: 1.6000000000000003e-05 evaluation reward: 4.51
1486
episode: 1486 score: 4.0 memory length: 313006 epsilon: 0.5782461400091559 step
s: 261 lr: 1.6000000000000003e-05 evaluation reward: 4.51
1487
episode: 1487 score: 4.0 memory length: 313286 epsilon: 0.5776917400091679 step
s: 280 lr: 1.6000000000000003e-05 evaluation reward: 4.53
1488
episode: 1488 score: 9.0 memory length: 313790 epsilon: 0.5766938200091896 step
s: 504 lr: 1.6000000000000003e-05 evaluation reward: 4.6
1489
episode: 1489 score: 4.0 memory length: 314048 epsilon: 0.5761829800092007 step
s: 258 lr: 1.6000000000000003e-05 evaluation reward: 4.57
1490

episode: 1490 score: 3.0 memory length: 314260 epsilon: 0.5757632200092098 step
s: 212 lr: 1.6000000000000003e-05 evaluation reward: 4.57
1491
episode: 1491 score: 6.0 memory length: 314597 epsilon: 0.5750959600092242 step
s: 337 lr: 1.6000000000000003e-05 evaluation reward: 4.57
1492
episode: 1492 score: 5.0 memory length: 314926 epsilon: 0.5744445400092384 step
s: 329 lr: 1.6000000000000003e-05 evaluation reward: 4.59
1493
episode: 1493 score: 8.0 memory length: 315372 epsilon: 0.5735614600092576 step
s: 446 lr: 1.6000000000000003e-05 evaluation reward: 4.61
1494
episode: 1494 score: 5.0 memory length: 315697 epsilon: 0.5729179600092715 step
s: 325 lr: 1.6000000000000003e-05 evaluation reward: 4.63
1495
episode: 1495 score: 5.0 memory length: 316004 epsilon: 0.5723101000092847 step
s: 307 lr: 1.6000000000000003e-05 evaluation reward: 4.65
1496
episode: 1496 score: 6.0 memory length: 316336 epsilon: 0.571652740009299 step
s: 332 lr: 1.6000000000000003e-05 evaluation reward: 4.67
1497
episode: 1497 score: 6.0 memory length: 316686 epsilon: 0.570959740009314 step
s: 350 lr: 1.6000000000000003e-05 evaluation reward: 4.69
1498
episode: 1498 score: 5.0 memory length: 316974 epsilon: 0.5703895000093264 step
s: 288 lr: 1.6000000000000003e-05 evaluation reward: 4.71
1499
episode: 1499 score: 3.0 memory length: 317201 epsilon: 0.5699400400093362 step
s: 227 lr: 1.6000000000000003e-05 evaluation reward: 4.69
1500
episode: 1500 score: 4.0 memory length: 317458 epsilon: 0.5694311800093472 step
s: 257 lr: 1.6000000000000003e-05 evaluation reward: 4.69
1501
episode: 1501 score: 11.0 memory length: 317837 epsilon: 0.5686807600093635 ste
ps: 379 lr: 1.6000000000000003e-05 evaluation reward: 4.75
1502
episode: 1502 score: 7.0 memory length: 318236 epsilon: 0.5678907400093807 step
s: 399 lr: 1.6000000000000003e-05 evaluation reward: 4.76
1503
episode: 1503 score: 5.0 memory length: 318561 epsilon: 0.5672472400093946 step
s: 325 lr: 1.6000000000000003e-05 evaluation reward: 4.8
1504
episode: 1504 score: 4.0 memory length: 318837 epsilon: 0.5667007600094065 step
s: 276 lr: 1.6000000000000003e-05 evaluation reward: 4.78
1505
episode: 1505 score: 6.0 memory length: 319214 epsilon: 0.5659543000094227 step
s: 377 lr: 1.6000000000000003e-05 evaluation reward: 4.79
1506
episode: 1506 score: 4.0 memory length: 319490 epsilon: 0.5654078200094346 step
s: 276 lr: 1.6000000000000003e-05 evaluation reward: 4.79
1507
episode: 1507 score: 5.0 memory length: 319797 epsilon: 0.5647999600094478 step
s: 307 lr: 1.6000000000000003e-05 evaluation reward: 4.81
1508
episode: 1508 score: 8.0 memory length: 320179 epsilon: 0.5640436000094642 step
s: 382 lr: 1.6000000000000003e-05 evaluation reward: 4.81
1509
episode: 1509 score: 3.0 memory length: 320407 epsilon: 0.563592160009474 step
s: 228 lr: 1.6000000000000003e-05 evaluation reward: 4.75
1510
episode: 1510 score: 3.0 memory length: 320652 epsilon: 0.5631070600094845 step
s: 245 lr: 1.6000000000000003e-05 evaluation reward: 4.77
1511
episode: 1511 score: 6.0 memory length: 320993 epsilon: 0.5624318800094992 step
s: 341 lr: 1.6000000000000003e-05 evaluation reward: 4.78
1512

episode: 1512 score: 9.0 memory length: 321454 epsilon: 0.561519100009519 step
s: 461 lr: 1.6000000000000003e-05 evaluation reward: 4.8
1513
episode: 1513 score: 6.0 memory length: 321869 epsilon: 0.5606974000095368 step
s: 415 lr: 1.6000000000000003e-05 evaluation reward: 4.83
1514
episode: 1514 score: 6.0 memory length: 322240 epsilon: 0.5599628200095528 step
s: 371 lr: 1.6000000000000003e-05 evaluation reward: 4.86
1515
episode: 1515 score: 4.0 memory length: 322499 epsilon: 0.5594500000095639 step
s: 259 lr: 1.6000000000000003e-05 evaluation reward: 4.85
1516
episode: 1516 score: 3.0 memory length: 322708 epsilon: 0.5590361800095729 step
s: 209 lr: 1.6000000000000003e-05 evaluation reward: 4.82
1517
episode: 1517 score: 6.0 memory length: 323046 epsilon: 0.5583669400095874 step
s: 338 lr: 1.6000000000000003e-05 evaluation reward: 4.83
1518
episode: 1518 score: 7.0 memory length: 323432 epsilon: 0.557602660009604 step
s: 386 lr: 1.6000000000000003e-05 evaluation reward: 4.85
1519
episode: 1519 score: 6.0 memory length: 323783 epsilon: 0.5569076800096191 step
s: 351 lr: 1.6000000000000003e-05 evaluation reward: 4.82
1520
episode: 1520 score: 3.0 memory length: 324029 epsilon: 0.5564206000096297 step
s: 246 lr: 1.6000000000000003e-05 evaluation reward: 4.78
1521
episode: 1521 score: 2.0 memory length: 324227 epsilon: 0.5560285600096382 step
s: 198 lr: 1.6000000000000003e-05 evaluation reward: 4.77
1522
episode: 1522 score: 2.0 memory length: 324425 epsilon: 0.5556365200096467 step
s: 198 lr: 1.6000000000000003e-05 evaluation reward: 4.75
1523
episode: 1523 score: 4.0 memory length: 324722 epsilon: 0.5550484600096595 step
s: 297 lr: 1.6000000000000003e-05 evaluation reward: 4.73
1524
episode: 1524 score: 6.0 memory length: 325077 epsilon: 0.5543455600096747 step
s: 355 lr: 1.6000000000000003e-05 evaluation reward: 4.77
1525
episode: 1525 score: 5.0 memory length: 325372 epsilon: 0.5537614600096874 step
s: 295 lr: 1.6000000000000003e-05 evaluation reward: 4.79
1526
episode: 1526 score: 4.0 memory length: 325647 epsilon: 0.5532169600096992 step
s: 275 lr: 1.6000000000000003e-05 evaluation reward: 4.79
1527
episode: 1527 score: 6.0 memory length: 326043 epsilon: 0.5524328800097162 step
s: 396 lr: 1.6000000000000003e-05 evaluation reward: 4.8
1528
episode: 1528 score: 3.0 memory length: 326314 epsilon: 0.5518963000097279 step
s: 271 lr: 1.6000000000000003e-05 evaluation reward: 4.78
1529
episode: 1529 score: 6.0 memory length: 326670 epsilon: 0.5511914200097432 step
s: 356 lr: 1.6000000000000003e-05 evaluation reward: 4.78
1530
episode: 1530 score: 6.0 memory length: 327009 epsilon: 0.5505202000097578 step
s: 339 lr: 1.6000000000000003e-05 evaluation reward: 4.73
1531
episode: 1531 score: 5.0 memory length: 327293 epsilon: 0.54995788000977 steps:
284 lr: 1.6000000000000003e-05 evaluation reward: 4.72
1532
episode: 1532 score: 4.0 memory length: 327559 epsilon: 0.5494312000097814 step
s: 266 lr: 1.6000000000000003e-05 evaluation reward: 4.72
1533
episode: 1533 score: 4.0 memory length: 327819 epsilon: 0.5489164000097926 step
s: 260 lr: 1.6000000000000003e-05 evaluation reward: 4.73
1534

episode: 1534 score: 6.0 memory length: 328167 epsilon: 0.5482273600098075 step
s: 348 lr: 1.6000000000000003e-05 evaluation reward: 4.76
1535
episode: 1535 score: 5.0 memory length: 328474 epsilon: 0.5476195000098207 step
s: 307 lr: 1.6000000000000003e-05 evaluation reward: 4.75
1536
episode: 1536 score: 4.0 memory length: 328749 epsilon: 0.5470750000098326 step
s: 275 lr: 1.6000000000000003e-05 evaluation reward: 4.75
1537
episode: 1537 score: 3.0 memory length: 328999 epsilon: 0.5465800000098433 step
s: 250 lr: 1.6000000000000003e-05 evaluation reward: 4.74
1538
episode: 1538 score: 4.0 memory length: 329260 epsilon: 0.5460632200098545 step
s: 261 lr: 1.6000000000000003e-05 evaluation reward: 4.72
1539
episode: 1539 score: 7.0 memory length: 329690 epsilon: 0.545211820009873 step
s: 430 lr: 1.6000000000000003e-05 evaluation reward: 4.75
1540
episode: 1540 score: 5.0 memory length: 330034 epsilon: 0.5445307000098878 step
s: 344 lr: 1.6000000000000003e-05 evaluation reward: 4.79
1541
episode: 1541 score: 8.0 memory length: 330499 epsilon: 0.5436100000099078 step
s: 465 lr: 1.6000000000000003e-05 evaluation reward: 4.83
1542
episode: 1542 score: 4.0 memory length: 330756 epsilon: 0.5431011400099188 step
s: 257 lr: 1.6000000000000003e-05 evaluation reward: 4.85
1543
episode: 1543 score: 4.0 memory length: 331053 epsilon: 0.5425130800099316 step
s: 297 lr: 1.6000000000000003e-05 evaluation reward: 4.84
1544
episode: 1544 score: 6.0 memory length: 331397 epsilon: 0.5418319600099464 step
s: 344 lr: 1.6000000000000003e-05 evaluation reward: 4.86
1545
episode: 1545 score: 5.0 memory length: 331729 epsilon: 0.5411746000099606 step
s: 332 lr: 1.6000000000000003e-05 evaluation reward: 4.86
1546
episode: 1546 score: 8.0 memory length: 332190 epsilon: 0.5402618200099805 step
s: 461 lr: 1.6000000000000003e-05 evaluation reward: 4.87
1547
episode: 1547 score: 4.0 memory length: 332446 epsilon: 0.5397549400099915 step
s: 256 lr: 1.6000000000000003e-05 evaluation reward: 4.87
1548
episode: 1548 score: 7.0 memory length: 332868 epsilon: 0.5389193800100096 step
s: 422 lr: 1.6000000000000003e-05 evaluation reward: 4.89
1549
episode: 1549 score: 5.0 memory length: 333217 epsilon: 0.5382283600100246 step
s: 349 lr: 1.6000000000000003e-05 evaluation reward: 4.91
1550
episode: 1550 score: 7.0 memory length: 333600 epsilon: 0.5374700200100411 step
s: 383 lr: 1.6000000000000003e-05 evaluation reward: 4.95
1551
episode: 1551 score: 5.0 memory length: 333927 epsilon: 0.5368225600100551 step
s: 327 lr: 1.6000000000000003e-05 evaluation reward: 4.96
1552
episode: 1552 score: 4.0 memory length: 334185 epsilon: 0.5363117200100662 step
s: 258 lr: 1.6000000000000003e-05 evaluation reward: 4.93
1553
episode: 1553 score: 9.0 memory length: 334550 epsilon: 0.5355890200100819 step
s: 365 lr: 1.6000000000000003e-05 evaluation reward: 4.98
1554
episode: 1554 score: 8.0 memory length: 334987 epsilon: 0.5347237600101007 step
s: 437 lr: 1.6000000000000003e-05 evaluation reward: 5.01
1555
episode: 1555 score: 3.0 memory length: 335198 epsilon: 0.5343059800101098 step
s: 211 lr: 1.6000000000000003e-05 evaluation reward: 5.03
1556

episode: 1556 score: 5.0 memory length: 335522 epsilon: 0.5336644600101237 step
s: 324 lr: 1.6000000000000003e-05 evaluation reward: 5.03
1557
episode: 1557 score: 5.0 memory length: 335845 epsilon: 0.5330249200101376 step
s: 323 lr: 1.6000000000000003e-05 evaluation reward: 5.05
1558
episode: 1558 score: 5.0 memory length: 336172 epsilon: 0.5323774600101516 step
s: 327 lr: 1.6000000000000003e-05 evaluation reward: 5.06
1559
episode: 1559 score: 3.0 memory length: 336382 epsilon: 0.5319616600101607 step
s: 210 lr: 1.6000000000000003e-05 evaluation reward: 5.05
1560
episode: 1560 score: 6.0 memory length: 336747 epsilon: 0.5312389600101763 step
s: 365 lr: 1.6000000000000003e-05 evaluation reward: 5.1
1561
episode: 1561 score: 4.0 memory length: 337043 epsilon: 0.5306528800101891 step
s: 296 lr: 1.6000000000000003e-05 evaluation reward: 5.11
1562
episode: 1562 score: 5.0 memory length: 337386 epsilon: 0.5299737400102038 step
s: 343 lr: 1.6000000000000003e-05 evaluation reward: 5.1
1563
episode: 1563 score: 5.0 memory length: 337672 epsilon: 0.5294074600102161 step
s: 286 lr: 1.6000000000000003e-05 evaluation reward: 5.1
1564
episode: 1564 score: 5.0 memory length: 337959 epsilon: 0.5288392000102284 step
s: 287 lr: 1.6000000000000003e-05 evaluation reward: 5.01
1565
episode: 1565 score: 9.0 memory length: 338437 epsilon: 0.527892760010249 step
s: 478 lr: 1.6000000000000003e-05 evaluation reward: 5.03
1566
episode: 1566 score: 5.0 memory length: 338747 epsilon: 0.5272789600102623 step
s: 310 lr: 1.6000000000000003e-05 evaluation reward: 5.06
1567
episode: 1567 score: 4.0 memory length: 338985 epsilon: 0.5268077200102725 step
s: 238 lr: 1.6000000000000003e-05 evaluation reward: 5.09
1568
episode: 1568 score: 5.0 memory length: 339311 epsilon: 0.5261622400102866 step
s: 326 lr: 1.6000000000000003e-05 evaluation reward: 5.06
1569
episode: 1569 score: 11.0 memory length: 339769 epsilon: 0.5252554000103062 ste
ps: 458 lr: 1.6000000000000003e-05 evaluation reward: 5.11
1570
episode: 1570 score: 5.0 memory length: 340074 epsilon: 0.5246515000103193 step
s: 305 lr: 1.6000000000000003e-05 evaluation reward: 5.12
1571
episode: 1571 score: 5.0 memory length: 340366 epsilon: 0.5240733400103319 step
s: 292 lr: 1.6000000000000003e-05 evaluation reward: 5.13
1572
episode: 1572 score: 3.0 memory length: 340598 epsilon: 0.5236139800103419 step
s: 232 lr: 1.6000000000000003e-05 evaluation reward: 5.15
1573
episode: 1573 score: 3.0 memory length: 340828 epsilon: 0.5231585800103518 step
s: 230 lr: 1.6000000000000003e-05 evaluation reward: 5.13
1574
episode: 1574 score: 5.0 memory length: 341116 epsilon: 0.5225883400103641 step
s: 288 lr: 1.6000000000000003e-05 evaluation reward: 5.16
1575
episode: 1575 score: 1.0 memory length: 341267 epsilon: 0.5222893600103706 step
s: 151 lr: 1.6000000000000003e-05 evaluation reward: 5.12
1576
episode: 1576 score: 7.0 memory length: 341663 epsilon: 0.5215052800103877 step
s: 396 lr: 1.6000000000000003e-05 evaluation reward: 5.09
1577
episode: 1577 score: 3.0 memory length: 341891 epsilon: 0.5210538400103975 step
s: 228 lr: 1.6000000000000003e-05 evaluation reward: 5.1
1578

episode: 1578	score: 5.0	memory length: 342177	epsilon: 0.5204875600104097	step
s: 286	lr: 1.6000000000000003e-05	evaluation reward: 5.1		
1579				
episode: 1579	score: 4.0	memory length: 342423	epsilon: 0.5200004800104203	step
s: 246	lr: 1.6000000000000003e-05	evaluation reward: 5.08		
1580				
episode: 1580	score: 2.0	memory length: 342624	epsilon: 0.519602500010429	step
s: 201	lr: 1.6000000000000003e-05	evaluation reward: 5.03		
1581				
episode: 1581	score: 6.0	memory length: 342994	epsilon: 0.5188699000104449	step
s: 370	lr: 1.6000000000000003e-05	evaluation reward: 5.07		
1582				
episode: 1582	score: 4.0	memory length: 343275	epsilon: 0.5183135200104569	step
s: 281	lr: 1.6000000000000003e-05	evaluation reward: 5.05		
1583				
episode: 1583	score: 4.0	memory length: 343554	epsilon: 0.5177611000104689	step
s: 279	lr: 1.6000000000000003e-05	evaluation reward: 5.08		
1584				
episode: 1584	score: 5.0	memory length: 343880	epsilon: 0.517115620010483	step
s: 326	lr: 1.6000000000000003e-05	evaluation reward: 5.07		
1585				
episode: 1585	score: 3.0	memory length: 344109	epsilon: 0.5166622000104928	step
s: 229	lr: 1.6000000000000003e-05	evaluation reward: 5.09		
1586				
episode: 1586	score: 4.0	memory length: 344383	epsilon: 0.5161196800105046	step
s: 274	lr: 1.6000000000000003e-05	evaluation reward: 5.09		
1587				
episode: 1587	score: 3.0	memory length: 344612	epsilon: 0.5156662600105144	step
s: 229	lr: 1.6000000000000003e-05	evaluation reward: 5.08		
1588				
episode: 1588	score: 3.0	memory length: 344842	epsilon: 0.5152108600105243	step
s: 230	lr: 1.6000000000000003e-05	evaluation reward: 5.02		
1589				
episode: 1589	score: 3.0	memory length: 345070	epsilon: 0.5147594200105341	step
s: 228	lr: 1.6000000000000003e-05	evaluation reward: 5.01		
1590				
episode: 1590	score: 3.0	memory length: 345280	epsilon: 0.5143436200105431	step
s: 210	lr: 1.6000000000000003e-05	evaluation reward: 5.01		
1591				
episode: 1591	score: 6.0	memory length: 345616	epsilon: 0.5136783400105576	step
s: 336	lr: 1.6000000000000003e-05	evaluation reward: 5.01		
1592				
episode: 1592	score: 5.0	memory length: 345946	epsilon: 0.5130249400105718	step
s: 330	lr: 1.6000000000000003e-05	evaluation reward: 5.01		
1593				
episode: 1593	score: 5.0	memory length: 346249	epsilon: 0.5124250000105848	step
s: 303	lr: 1.6000000000000003e-05	evaluation reward: 4.98		
1594				
episode: 1594	score: 5.0	memory length: 346581	epsilon: 0.511767640010599	step
s: 332	lr: 1.6000000000000003e-05	evaluation reward: 4.98		
1595				
episode: 1595	score: 7.0	memory length: 346979	epsilon: 0.5109796000106162	step
s: 398	lr: 1.6000000000000003e-05	evaluation reward: 5.0		
1596				
episode: 1596	score: 7.0	memory length: 347414	epsilon: 0.5101183000106349	step
s: 435	lr: 1.6000000000000003e-05	evaluation reward: 5.01		
1597				
episode: 1597	score: 3.0	memory length: 347660	epsilon: 0.5096312200106454	step
s: 246	lr: 1.6000000000000003e-05	evaluation reward: 4.98		
1598				
episode: 1598	score: 3.0	memory length: 347906	epsilon: 0.509144140010656	step
s: 246	lr: 1.6000000000000003e-05	evaluation reward: 4.96		
1599				
episode: 1599	score: 8.0	memory length: 348313	epsilon: 0.5083382800106735	step
s: 407	lr: 1.6000000000000003e-05	evaluation reward: 5.01		
1600				

```
episode: 1600    score: 2.0    memory length: 348510    epsilon: 0.507948220010682    step
s: 197    lr: 1.6000000000000003e-05    evaluation reward: 4.99
1601
episode: 1601    score: 6.0    memory length: 348881    epsilon: 0.5072136400106979    step
s: 371    lr: 1.6000000000000003e-05    evaluation reward: 4.94
1602
episode: 1602    score: 9.0    memory length: 349329    epsilon: 0.5063266000107172    step
s: 448    lr: 1.6000000000000003e-05    evaluation reward: 4.96
1603
episode: 1603    score: 12.0    memory length: 349930    epsilon: 0.505136620010743    step
s: 601    lr: 1.6000000000000003e-05    evaluation reward: 5.03
1604
episode: 1604    score: 5.0    memory length: 350257    epsilon: 0.504489160010757    step
s: 327    lr: 1.6000000000000003e-05    evaluation reward: 5.04
1605
episode: 1605    score: 5.0    memory length: 350579    epsilon: 0.5038516000107709    step
s: 322    lr: 1.6000000000000003e-05    evaluation reward: 5.03
1606
episode: 1606    score: 2.0    memory length: 350761    epsilon: 0.5034912400107787    step
s: 182    lr: 1.6000000000000003e-05    evaluation reward: 5.01
1607
episode: 1607    score: 9.0    memory length: 351086    epsilon: 0.5028477400107927    step
s: 325    lr: 1.6000000000000003e-05    evaluation reward: 5.05
1608
episode: 1608    score: 11.0    memory length: 351636    epsilon: 0.5017587400108163    ste
ps: 550    lr: 1.6000000000000003e-05    evaluation reward: 5.08
1609
episode: 1609    score: 11.0    memory length: 352202    epsilon: 0.5006380600108407    ste
ps: 566    lr: 1.6000000000000003e-05    evaluation reward: 5.16
1610
episode: 1610    score: 2.0    memory length: 352381    epsilon: 0.5002836400108484    step
s: 179    lr: 1.6000000000000003e-05    evaluation reward: 5.15
1611
episode: 1611    score: 8.0    memory length: 352866    epsilon: 0.4993233400108502    step
s: 485    lr: 1.6000000000000003e-05    evaluation reward: 5.17
1612
episode: 1612    score: 10.0    memory length: 353258    epsilon: 0.4985471800108453    ste
ps: 392    lr: 1.6000000000000003e-05    evaluation reward: 5.18
1613
episode: 1613    score: 4.0    memory length: 353556    epsilon: 0.49795714001084157    ste
ps: 298    lr: 1.6000000000000003e-05    evaluation reward: 5.16
1614
episode: 1614    score: 4.0    memory length: 353835    epsilon: 0.4974047200108381    step
s: 279    lr: 1.6000000000000003e-05    evaluation reward: 5.14
1615
episode: 1615    score: 4.0    memory length: 354114    epsilon: 0.4968523000108346    step
s: 279    lr: 1.6000000000000003e-05    evaluation reward: 5.14
1616
episode: 1616    score: 3.0    memory length: 354326    epsilon: 0.4964325400108319    step
s: 212    lr: 1.6000000000000003e-05    evaluation reward: 5.14
1617
episode: 1617    score: 6.0    memory length: 354681    epsilon: 0.4957296400108275    step
s: 355    lr: 1.6000000000000003e-05    evaluation reward: 5.14
1618
episode: 1618    score: 9.0    memory length: 355039    epsilon: 0.495020800010823    step
s: 358    lr: 1.6000000000000003e-05    evaluation reward: 5.16
1619
episode: 1619    score: 5.0    memory length: 355362    epsilon: 0.49438126001081895    ste
ps: 323    lr: 1.6000000000000003e-05    evaluation reward: 5.15
1620
episode: 1620    score: 8.0    memory length: 355767    epsilon: 0.4935793600108139    step
s: 405    lr: 1.6000000000000003e-05    evaluation reward: 5.2
1621
episode: 1621    score: 7.0    memory length: 356161    epsilon: 0.49279924001080894    ste
ps: 394    lr: 1.6000000000000003e-05    evaluation reward: 5.25
1622
```

episode: 1622	score: 1.0	memory length: 356312	epsilon: 0.49250026001080704	step
ps: 151	lr: 1.6000000000000003e-05	evaluation reward: 5.24		
1623				
episode: 1623	score: 7.0	memory length: 356723	epsilon: 0.4916864800108019	step
s: 411	lr: 1.6000000000000003e-05	evaluation reward: 5.27		
1624				
episode: 1624	score: 5.0	memory length: 357051	epsilon: 0.4910370400107978	step
s: 328	lr: 1.6000000000000003e-05	evaluation reward: 5.26		
1625				
episode: 1625	score: 7.0	memory length: 357446	epsilon: 0.49025494001079284	step
ps: 395	lr: 1.6000000000000003e-05	evaluation reward: 5.28		
1626				
episode: 1626	score: 3.0	memory length: 357691	epsilon: 0.48976984001078977	step
ps: 245	lr: 1.6000000000000003e-05	evaluation reward: 5.27		
1627				
episode: 1627	score: 4.0	memory length: 357951	epsilon: 0.4892550400107865	step
s: 260	lr: 1.6000000000000003e-05	evaluation reward: 5.25		
1628				
episode: 1628	score: 4.0	memory length: 358229	epsilon: 0.48870460001078303	step
ps: 278	lr: 1.6000000000000003e-05	evaluation reward: 5.26		
1629				
episode: 1629	score: 5.0	memory length: 358553	epsilon: 0.48806308001077897	step
ps: 324	lr: 1.6000000000000003e-05	evaluation reward: 5.25		
1630				
episode: 1630	score: 3.0	memory length: 358764	epsilon: 0.48764530001077633	step
ps: 211	lr: 1.6000000000000003e-05	evaluation reward: 5.22		
1631				
episode: 1631	score: 6.0	memory length: 359167	epsilon: 0.4868473600107713	step
s: 403	lr: 1.6000000000000003e-05	evaluation reward: 5.23		
1632				
episode: 1632	score: 6.0	memory length: 359507	epsilon: 0.486174160010767	step
s: 340	lr: 1.6000000000000003e-05	evaluation reward: 5.25		
1633				
episode: 1633	score: 3.0	memory length: 359735	epsilon: 0.48572272001076416	step
ps: 228	lr: 1.6000000000000003e-05	evaluation reward: 5.24		
1634				
episode: 1634	score: 3.0	memory length: 359964	epsilon: 0.4852693000107613	step
s: 229	lr: 1.6000000000000003e-05	evaluation reward: 5.21		
1635				
episode: 1635	score: 4.0	memory length: 360224	epsilon: 0.48475450001075804	step
ps: 260	lr: 1.6000000000000003e-05	evaluation reward: 5.2		
1636				
episode: 1636	score: 6.0	memory length: 360590	epsilon: 0.48402982001075345	step
ps: 366	lr: 1.6000000000000003e-05	evaluation reward: 5.22		
1637				
episode: 1637	score: 6.0	memory length: 360911	epsilon: 0.48339424001074943	step
ps: 321	lr: 1.6000000000000003e-05	evaluation reward: 5.25		
1638				
episode: 1638	score: 11.0	memory length: 361455	epsilon: 0.4823171200107426	step
ps: 544	lr: 1.6000000000000003e-05	evaluation reward: 5.32		
1639				
episode: 1639	score: 5.0	memory length: 361785	epsilon: 0.4816637200107385	step
s: 330	lr: 1.6000000000000003e-05	evaluation reward: 5.3		
1640				
episode: 1640	score: 7.0	memory length: 362151	epsilon: 0.4809390400107339	step
s: 366	lr: 1.6000000000000003e-05	evaluation reward: 5.32		
1641				
episode: 1641	score: 7.0	memory length: 362567	epsilon: 0.4801153600107287	step
s: 416	lr: 1.6000000000000003e-05	evaluation reward: 5.31		
1642				
episode: 1642	score: 6.0	memory length: 362943	epsilon: 0.479370880010724	step
s: 376	lr: 1.6000000000000003e-05	evaluation reward: 5.33		
1643				
episode: 1643	score: 4.0	memory length: 363241	epsilon: 0.47878084001072024	step
ps: 298	lr: 1.6000000000000003e-05	evaluation reward: 5.33		
1644				

episode: 1644	score: 6.0	memory length: 363599	epsilon: 0.47807200001071576	step
ps: 358	lr: 1.6000000000000003e-05	evaluation reward: 5.33		
1645				
episode: 1645	score: 6.0	memory length: 363978	epsilon: 0.477321580010711	step
s: 379	lr: 1.6000000000000003e-05	evaluation reward: 5.34		
1646				
episode: 1646	score: 5.0	memory length: 364305	epsilon: 0.4766741200107069	step
s: 327	lr: 1.6000000000000003e-05	evaluation reward: 5.31		
1647				
episode: 1647	score: 9.0	memory length: 364754	epsilon: 0.4757851000107013	step
s: 449	lr: 1.6000000000000003e-05	evaluation reward: 5.36		
1648				
episode: 1648	score: 11.0	memory length: 365192	epsilon: 0.4749178600106958	step
ps: 438	lr: 1.6000000000000003e-05	evaluation reward: 5.4		
1649				
episode: 1649	score: 5.0	memory length: 365508	epsilon: 0.47429218001069184	step
ps: 316	lr: 1.6000000000000003e-05	evaluation reward: 5.4		
1650				
episode: 1650	score: 6.0	memory length: 365872	epsilon: 0.4735714600106873	step
s: 364	lr: 1.6000000000000003e-05	evaluation reward: 5.39		
1651				
episode: 1651	score: 10.0	memory length: 366317	epsilon: 0.4726903600106817	step
ps: 445	lr: 1.6000000000000003e-05	evaluation reward: 5.44		
1652				
episode: 1652	score: 8.0	memory length: 366752	epsilon: 0.47182906001067626	step
ps: 435	lr: 1.6000000000000003e-05	evaluation reward: 5.48		
1653				
episode: 1653	score: 8.0	memory length: 367151	epsilon: 0.47103904001067126	step
ps: 399	lr: 1.6000000000000003e-05	evaluation reward: 5.47		
1654				
episode: 1654	score: 10.0	memory length: 367668	epsilon: 0.4700153800106648	step
ps: 517	lr: 1.6000000000000003e-05	evaluation reward: 5.49		
1655				
episode: 1655	score: 4.0	memory length: 367962	epsilon: 0.4694332600106611	step
s: 294	lr: 1.6000000000000003e-05	evaluation reward: 5.5		
1656				
episode: 1656	score: 5.0	memory length: 368285	epsilon: 0.46879372001065706	step
ps: 323	lr: 1.6000000000000003e-05	evaluation reward: 5.5		
1657				
episode: 1657	score: 6.0	memory length: 368663	epsilon: 0.4680452800106523	step
s: 378	lr: 1.6000000000000003e-05	evaluation reward: 5.51		
1658				
episode: 1658	score: 5.0	memory length: 368988	epsilon: 0.46740178001064825	step
ps: 325	lr: 1.6000000000000003e-05	evaluation reward: 5.51		
1659				
episode: 1659	score: 5.0	memory length: 369277	epsilon: 0.46682956001064463	step
ps: 289	lr: 1.6000000000000003e-05	evaluation reward: 5.53		
1660				
episode: 1660	score: 6.0	memory length: 369657	epsilon: 0.46607716001063987	step
ps: 380	lr: 1.6000000000000003e-05	evaluation reward: 5.53		
1661				
episode: 1661	score: 7.0	memory length: 370031	epsilon: 0.4653366400106352	step
s: 374	lr: 1.6000000000000003e-05	evaluation reward: 5.56		
1662				
episode: 1662	score: 9.0	memory length: 370486	epsilon: 0.4644357400106295	step
s: 455	lr: 1.6000000000000003e-05	evaluation reward: 5.6		
1663				
episode: 1663	score: 5.0	memory length: 370802	epsilon: 0.4638100600106255	step
s: 316	lr: 1.6000000000000003e-05	evaluation reward: 5.6		
1664				
episode: 1664	score: 6.0	memory length: 371158	epsilon: 0.46310518001062106	step
ps: 356	lr: 1.6000000000000003e-05	evaluation reward: 5.61		
1665				
episode: 1665	score: 3.0	memory length: 371387	epsilon: 0.4626517600106182	step
s: 229	lr: 1.6000000000000003e-05	evaluation reward: 5.55		
1666				

episode: 1666 score: 5.0 memory length: 371731 epsilon: 0.4619706400106139 step
s: 344 lr: 1.6000000000000003e-05 evaluation reward: 5.55
1667
episode: 1667 score: 5.0 memory length: 372053 epsilon: 0.46133308001060985 ste
ps: 322 lr: 1.6000000000000003e-05 evaluation reward: 5.56
1668
episode: 1668 score: 3.0 memory length: 372298 epsilon: 0.4608479800106068 step
s: 245 lr: 1.6000000000000003e-05 evaluation reward: 5.54
1669
episode: 1669 score: 7.0 memory length: 372707 epsilon: 0.46003816001060166 ste
ps: 409 lr: 1.6000000000000003e-05 evaluation reward: 5.5
1670
episode: 1670 score: 3.0 memory length: 372920 epsilon: 0.459616420010599 step
s: 213 lr: 1.6000000000000003e-05 evaluation reward: 5.48
1671
episode: 1671 score: 6.0 memory length: 373237 epsilon: 0.458988760010595 step
s: 317 lr: 1.6000000000000003e-05 evaluation reward: 5.49
1672
episode: 1672 score: 7.0 memory length: 373619 epsilon: 0.45823240001059024 ste
ps: 382 lr: 1.6000000000000003e-05 evaluation reward: 5.53
1673
episode: 1673 score: 6.0 memory length: 373991 epsilon: 0.4574958400105856 step
s: 372 lr: 1.6000000000000003e-05 evaluation reward: 5.56
1674
episode: 1674 score: 6.0 memory length: 374343 epsilon: 0.45679888001058117 ste
ps: 352 lr: 1.6000000000000003e-05 evaluation reward: 5.57
1675
episode: 1675 score: 3.0 memory length: 374571 epsilon: 0.4563474400105783 step
s: 228 lr: 1.6000000000000003e-05 evaluation reward: 5.59
1676
episode: 1676 score: 9.0 memory length: 375065 epsilon: 0.4553693200105721 step
s: 494 lr: 1.6000000000000003e-05 evaluation reward: 5.61
1677
episode: 1677 score: 3.0 memory length: 375277 epsilon: 0.45494956001056946 ste
ps: 212 lr: 1.6000000000000003e-05 evaluation reward: 5.61
1678
episode: 1678 score: 2.0 memory length: 375458 epsilon: 0.4545911800105672 step
s: 181 lr: 1.6000000000000003e-05 evaluation reward: 5.58
1679
episode: 1679 score: 4.0 memory length: 375716 epsilon: 0.45408034001056397 ste
ps: 258 lr: 1.6000000000000003e-05 evaluation reward: 5.58
1680
episode: 1680 score: 5.0 memory length: 376005 epsilon: 0.45350812001056034 ste
ps: 289 lr: 1.6000000000000003e-05 evaluation reward: 5.61
1681
episode: 1681 score: 9.0 memory length: 376452 epsilon: 0.45262306001055475 ste
ps: 447 lr: 1.6000000000000003e-05 evaluation reward: 5.64
1682
episode: 1682 score: 2.0 memory length: 376649 epsilon: 0.4522330000105523 step
s: 197 lr: 1.6000000000000003e-05 evaluation reward: 5.62
1683
episode: 1683 score: 11.0 memory length: 377074 epsilon: 0.45139150001054695 st
eps: 425 lr: 1.6000000000000003e-05 evaluation reward: 5.69
1684
episode: 1684 score: 8.0 memory length: 377544 epsilon: 0.45046090001054107 ste
ps: 470 lr: 1.6000000000000003e-05 evaluation reward: 5.72
1685
episode: 1685 score: 8.0 memory length: 377936 epsilon: 0.44968474001053615 ste
ps: 392 lr: 1.6000000000000003e-05 evaluation reward: 5.77
1686
episode: 1686 score: 5.0 memory length: 378275 epsilon: 0.4490135200105319 step
s: 339 lr: 1.6000000000000003e-05 evaluation reward: 5.78
1687
episode: 1687 score: 6.0 memory length: 378633 epsilon: 0.4483046800105274 step
s: 358 lr: 1.6000000000000003e-05 evaluation reward: 5.81
1688

episode: 1688 score: 5.0 memory length: 378961 epsilon: 0.4476552400105233 step
s: 328 lr: 1.6000000000000003e-05 evaluation reward: 5.83
1689
episode: 1689 score: 12.0 memory length: 379457 epsilon: 0.4466731600105171 ste
ps: 496 lr: 1.6000000000000003e-05 evaluation reward: 5.92
1690
episode: 1690 score: 5.0 memory length: 379732 epsilon: 0.44612866001051366 ste
ps: 275 lr: 1.6000000000000003e-05 evaluation reward: 5.94
1691
episode: 1691 score: 5.0 memory length: 380040 epsilon: 0.4455188200105098 step
s: 308 lr: 1.6000000000000003e-05 evaluation reward: 5.93
1692
episode: 1692 score: 4.0 memory length: 380306 epsilon: 0.44499214001050647 ste
ps: 266 lr: 1.6000000000000003e-05 evaluation reward: 5.92
1693
episode: 1693 score: 6.0 memory length: 380683 epsilon: 0.44424568001050174 ste
ps: 377 lr: 1.6000000000000003e-05 evaluation reward: 5.93
1694
episode: 1694 score: 3.0 memory length: 380893 epsilon: 0.4438298800104991 step
s: 210 lr: 1.6000000000000003e-05 evaluation reward: 5.91
1695
episode: 1695 score: 6.0 memory length: 381250 epsilon: 0.44312302001049464 ste
ps: 357 lr: 1.6000000000000003e-05 evaluation reward: 5.9
1696
episode: 1696 score: 5.0 memory length: 381552 epsilon: 0.44252506001049086 ste
ps: 302 lr: 1.6000000000000003e-05 evaluation reward: 5.88
1697
episode: 1697 score: 7.0 memory length: 381937 epsilon: 0.44176276001048603 ste
ps: 385 lr: 1.6000000000000003e-05 evaluation reward: 5.92
1698
episode: 1698 score: 3.0 memory length: 382147 epsilon: 0.4413469600104834 step
s: 210 lr: 1.6000000000000003e-05 evaluation reward: 5.92
1699
episode: 1699 score: 8.0 memory length: 382535 epsilon: 0.44057872001047854 ste
ps: 388 lr: 1.6000000000000003e-05 evaluation reward: 5.92
1700
episode: 1700 score: 13.0 memory length: 383039 epsilon: 0.43958080001047223 st
eps: 504 lr: 1.6000000000000003e-05 evaluation reward: 6.03
1701
episode: 1701 score: 7.0 memory length: 383446 epsilon: 0.43877494001046713 ste
ps: 407 lr: 1.6000000000000003e-05 evaluation reward: 6.04
1702
episode: 1702 score: 5.0 memory length: 383772 epsilon: 0.43812946001046305 ste
ps: 326 lr: 1.6000000000000003e-05 evaluation reward: 6.0
1703
episode: 1703 score: 5.0 memory length: 384066 epsilon: 0.43754734001045936 ste
ps: 294 lr: 1.6000000000000003e-05 evaluation reward: 5.93
1704
episode: 1704 score: 2.0 memory length: 384246 epsilon: 0.4371909400104571 step
s: 180 lr: 1.6000000000000003e-05 evaluation reward: 5.9
1705
episode: 1705 score: 6.0 memory length: 384588 epsilon: 0.4365137800104528 step
s: 342 lr: 1.6000000000000003e-05 evaluation reward: 5.91
1706
episode: 1706 score: 6.0 memory length: 384959 epsilon: 0.4357792000104482 step
s: 371 lr: 1.6000000000000003e-05 evaluation reward: 5.95
1707
episode: 1707 score: 7.0 memory length: 385315 epsilon: 0.4350743200104437 step
s: 356 lr: 1.6000000000000003e-05 evaluation reward: 5.93
1708
episode: 1708 score: 6.0 memory length: 385669 epsilon: 0.4343734000104393 step
s: 354 lr: 1.6000000000000003e-05 evaluation reward: 5.88
1709
episode: 1709 score: 7.0 memory length: 386020 epsilon: 0.4336784200104349 step
s: 351 lr: 1.6000000000000003e-05 evaluation reward: 5.84
1710

episode: 1710 score: 6.0 memory length: 386361 epsilon: 0.4330032400104306 step
s: 341 lr: 1.6000000000000003e-05 evaluation reward: 5.88
1711
episode: 1711 score: 5.0 memory length: 386686 epsilon: 0.43235974001042654 ste
ps: 325 lr: 1.6000000000000003e-05 evaluation reward: 5.85
1712
episode: 1712 score: 2.0 memory length: 386865 epsilon: 0.4320053200104243 step
s: 179 lr: 1.6000000000000003e-05 evaluation reward: 5.77
1713
episode: 1713 score: 6.0 memory length: 387187 epsilon: 0.43136776001042026 ste
ps: 322 lr: 1.6000000000000003e-05 evaluation reward: 5.79
1714
episode: 1714 score: 6.0 memory length: 387561 epsilon: 0.4306272400104156 step
s: 374 lr: 1.6000000000000003e-05 evaluation reward: 5.81
1715
episode: 1715 score: 2.0 memory length: 387743 epsilon: 0.4302668800104133 step
s: 182 lr: 1.6000000000000003e-05 evaluation reward: 5.79
1716
episode: 1716 score: 3.0 memory length: 387992 epsilon: 0.4297738600104102 step
s: 249 lr: 1.6000000000000003e-05 evaluation reward: 5.79
1717
episode: 1717 score: 5.0 memory length: 388317 epsilon: 0.4291303600104061 step
s: 325 lr: 1.6000000000000003e-05 evaluation reward: 5.78
1718
episode: 1718 score: 14.0 memory length: 388886 epsilon: 0.428003740010399 step
s: 569 lr: 1.6000000000000003e-05 evaluation reward: 5.83
1719
episode: 1719 score: 2.0 memory length: 389083 epsilon: 0.4276136800103965 step
s: 197 lr: 1.6000000000000003e-05 evaluation reward: 5.8
1720
episode: 1720 score: 5.0 memory length: 389395 epsilon: 0.4269959200103926 step
s: 312 lr: 1.6000000000000003e-05 evaluation reward: 5.77
1721
episode: 1721 score: 7.0 memory length: 389769 epsilon: 0.4262554000103879 step
s: 374 lr: 1.6000000000000003e-05 evaluation reward: 5.77
1722
episode: 1722 score: 10.0 memory length: 390124 epsilon: 0.42555250001038347 st
eps: 355 lr: 1.6000000000000003e-05 evaluation reward: 5.86
1723
episode: 1723 score: 4.0 memory length: 390404 epsilon: 0.42499810001037996 ste
ps: 280 lr: 1.6000000000000003e-05 evaluation reward: 5.83
1724
episode: 1724 score: 6.0 memory length: 390787 epsilon: 0.42423976001037517 ste
ps: 383 lr: 1.6000000000000003e-05 evaluation reward: 5.84
1725
episode: 1725 score: 5.0 memory length: 391097 epsilon: 0.4236259600103713 step
s: 310 lr: 1.6000000000000003e-05 evaluation reward: 5.82
1726
episode: 1726 score: 7.0 memory length: 391503 epsilon: 0.4228220800103662 step
s: 406 lr: 1.6000000000000003e-05 evaluation reward: 5.86
1727
episode: 1727 score: 8.0 memory length: 391960 epsilon: 0.42191722001036047 ste
ps: 457 lr: 1.6000000000000003e-05 evaluation reward: 5.9
1728
episode: 1728 score: 7.0 memory length: 392341 epsilon: 0.4211628400103557 step
s: 381 lr: 1.6000000000000003e-05 evaluation reward: 5.93
1729
episode: 1729 score: 8.0 memory length: 392766 epsilon: 0.4203213400103504 step
s: 425 lr: 1.6000000000000003e-05 evaluation reward: 5.96
1730
episode: 1730 score: 3.0 memory length: 392996 epsilon: 0.4198659400103475 step
s: 230 lr: 1.6000000000000003e-05 evaluation reward: 5.96
1731
episode: 1731 score: 6.0 memory length: 393335 epsilon: 0.41919472001034325 ste
ps: 339 lr: 1.6000000000000003e-05 evaluation reward: 5.96
1732

episode: 1732 score: 8.0 memory length: 393727 epsilon: 0.41841856001033834 step
ps: 392 lr: 1.6000000000000003e-05 evaluation reward: 5.98
1733
episode: 1733 score: 6.0 memory length: 394082 epsilon: 0.4177156600103339 step
s: 355 lr: 1.6000000000000003e-05 evaluation reward: 6.01
1734
episode: 1734 score: 4.0 memory length: 394345 epsilon: 0.4171949200103306 step
s: 263 lr: 1.6000000000000003e-05 evaluation reward: 6.02
1735
episode: 1735 score: 4.0 memory length: 394604 epsilon: 0.41668210001032735 step
ps: 259 lr: 1.6000000000000003e-05 evaluation reward: 6.02
1736
episode: 1736 score: 7.0 memory length: 395013 epsilon: 0.4158722800103222 step
s: 409 lr: 1.6000000000000003e-05 evaluation reward: 6.03
1737
episode: 1737 score: 3.0 memory length: 395279 epsilon: 0.4153456000103189 step
s: 266 lr: 1.6000000000000003e-05 evaluation reward: 6.0
1738
episode: 1738 score: 10.0 memory length: 395786 epsilon: 0.41434174001031254 step
eps: 507 lr: 1.6000000000000003e-05 evaluation reward: 5.99
1739
episode: 1739 score: 10.0 memory length: 396292 epsilon: 0.4133398600103062 step
ps: 506 lr: 1.6000000000000003e-05 evaluation reward: 6.04
1740
episode: 1740 score: 6.0 memory length: 396650 epsilon: 0.4126310200103017 step
s: 358 lr: 1.6000000000000003e-05 evaluation reward: 6.03
1741
episode: 1741 score: 7.0 memory length: 397043 epsilon: 0.4118528800102968 step
s: 393 lr: 1.6000000000000003e-05 evaluation reward: 6.03
1742
episode: 1742 score: 5.0 memory length: 397350 epsilon: 0.41124502001029295 step
ps: 307 lr: 1.6000000000000003e-05 evaluation reward: 6.02
1743
episode: 1743 score: 6.0 memory length: 397676 epsilon: 0.41059954001028887 step
ps: 326 lr: 1.6000000000000003e-05 evaluation reward: 6.04
1744
episode: 1744 score: 4.0 memory length: 397974 epsilon: 0.41000950001028513 step
ps: 298 lr: 1.6000000000000003e-05 evaluation reward: 6.02
1745
episode: 1745 score: 5.0 memory length: 398266 epsilon: 0.4094313400102815 step
s: 292 lr: 1.6000000000000003e-05 evaluation reward: 6.01
1746
episode: 1746 score: 4.0 memory length: 398540 epsilon: 0.40888882001027804 step
ps: 274 lr: 1.6000000000000003e-05 evaluation reward: 6.0
1747
episode: 1747 score: 7.0 memory length: 398965 epsilon: 0.4080473200102727 step
s: 425 lr: 1.6000000000000003e-05 evaluation reward: 5.98
1748
episode: 1748 score: 5.0 memory length: 399256 epsilon: 0.4074711400102691 step
s: 291 lr: 1.6000000000000003e-05 evaluation reward: 5.92
1749
episode: 1749 score: 9.0 memory length: 399709 epsilon: 0.4065742000102634 step
s: 453 lr: 1.6000000000000003e-05 evaluation reward: 5.96
1750
episode: 1750 score: 8.0 memory length: 400185 epsilon: 0.40563172001025743 step
ps: 476 lr: 6.400000000000001e-06 evaluation reward: 5.98
1751
episode: 1751 score: 7.0 memory length: 400612 epsilon: 0.4047862600102521 step
s: 427 lr: 6.400000000000001e-06 evaluation reward: 5.95
1752
episode: 1752 score: 8.0 memory length: 401018 epsilon: 0.403982380010247 step
s: 406 lr: 6.400000000000001e-06 evaluation reward: 5.95
1753
episode: 1753 score: 7.0 memory length: 401422 epsilon: 0.40318246001024194 step
ps: 404 lr: 6.400000000000001e-06 evaluation reward: 5.94
1754

episode: 1754 score: 4.0 memory length: 401682 epsilon: 0.4026676600102387 step
s: 260 lr: 6.400000000000001e-06 evaluation reward: 5.88
1755
episode: 1755 score: 5.0 memory length: 401986 epsilon: 0.4020657400102349 step
s: 304 lr: 6.400000000000001e-06 evaluation reward: 5.89
1756
episode: 1756 score: 10.0 memory length: 402556 epsilon: 0.40093714001022773 st
eps: 570 lr: 6.400000000000001e-06 evaluation reward: 5.94
1757
episode: 1757 score: 3.0 memory length: 402769 epsilon: 0.40051540001022506 ste
ps: 213 lr: 6.400000000000001e-06 evaluation reward: 5.91
1758
episode: 1758 score: 5.0 memory length: 403094 epsilon: 0.399871900010221 step
s: 325 lr: 6.400000000000001e-06 evaluation reward: 5.91
1759
episode: 1759 score: 5.0 memory length: 403441 epsilon: 0.39918484001021665 ste
ps: 347 lr: 6.400000000000001e-06 evaluation reward: 5.91
1760
episode: 1760 score: 17.0 memory length: 403971 epsilon: 0.39813544001021 step
s: 530 lr: 6.400000000000001e-06 evaluation reward: 6.02
1761
episode: 1761 score: 7.0 memory length: 404396 epsilon: 0.3972939400102047 step
s: 425 lr: 6.400000000000001e-06 evaluation reward: 6.02
1762
episode: 1762 score: 11.0 memory length: 404935 epsilon: 0.39622672001019793 st
eps: 539 lr: 6.400000000000001e-06 evaluation reward: 6.04
1763
episode: 1763 score: 8.0 memory length: 405345 epsilon: 0.3954149200101928 step
s: 410 lr: 6.400000000000001e-06 evaluation reward: 6.07
1764
episode: 1764 score: 9.0 memory length: 405859 epsilon: 0.39439720001018636 ste
ps: 514 lr: 6.400000000000001e-06 evaluation reward: 6.1
1765
episode: 1765 score: 6.0 memory length: 406188 epsilon: 0.39374578001018223 ste
ps: 329 lr: 6.400000000000001e-06 evaluation reward: 6.13
1766
episode: 1766 score: 4.0 memory length: 406485 epsilon: 0.3931577200101785 step
s: 297 lr: 6.400000000000001e-06 evaluation reward: 6.12
1767
episode: 1767 score: 7.0 memory length: 406874 epsilon: 0.39238750001017364 ste
ps: 389 lr: 6.400000000000001e-06 evaluation reward: 6.14
1768
episode: 1768 score: 5.0 memory length: 407200 epsilon: 0.39174202001016956 ste
ps: 326 lr: 6.400000000000001e-06 evaluation reward: 6.16
1769
episode: 1769 score: 7.0 memory length: 407598 epsilon: 0.39095398001016457 ste
ps: 398 lr: 6.400000000000001e-06 evaluation reward: 6.16
1770
episode: 1770 score: 5.0 memory length: 407907 epsilon: 0.3903421600101607 step
s: 309 lr: 6.400000000000001e-06 evaluation reward: 6.18
1771
episode: 1771 score: 9.0 memory length: 408379 epsilon: 0.3894076000101548 step
s: 472 lr: 6.400000000000001e-06 evaluation reward: 6.21
1772
episode: 1772 score: 5.0 memory length: 408706 epsilon: 0.3887601400101507 step
s: 327 lr: 6.400000000000001e-06 evaluation reward: 6.19
1773
episode: 1773 score: 6.0 memory length: 409047 epsilon: 0.3880849600101464 step
s: 341 lr: 6.400000000000001e-06 evaluation reward: 6.19
1774
episode: 1774 score: 6.0 memory length: 409401 epsilon: 0.387384040010142 step
s: 354 lr: 6.400000000000001e-06 evaluation reward: 6.19
1775
episode: 1775 score: 9.0 memory length: 409858 epsilon: 0.38647918001013626 ste
ps: 457 lr: 6.400000000000001e-06 evaluation reward: 6.25
1776

episode: 1776 score: 7.0 memory length: 410273 epsilon: 0.38565748001013106 step
ps: 415 lr: 6.400000000000001e-06 evaluation reward: 6.23
1777
episode: 1777 score: 2.0 memory length: 410454 epsilon: 0.3852991000101288 step
s: 181 lr: 6.400000000000001e-06 evaluation reward: 6.22
1778
episode: 1778 score: 9.0 memory length: 410884 epsilon: 0.3844477000101234 step
s: 430 lr: 6.400000000000001e-06 evaluation reward: 6.29
1779
episode: 1779 score: 5.0 memory length: 411180 epsilon: 0.3838616200101197 step
s: 296 lr: 6.400000000000001e-06 evaluation reward: 6.3
1780
episode: 1780 score: 5.0 memory length: 411488 epsilon: 0.38325178001011584 step
ps: 308 lr: 6.400000000000001e-06 evaluation reward: 6.3
1781
episode: 1781 score: 8.0 memory length: 411924 epsilon: 0.3823885000101104 step
s: 436 lr: 6.400000000000001e-06 evaluation reward: 6.29
1782
episode: 1782 score: 4.0 memory length: 412190 epsilon: 0.38186182001010704 step
ps: 266 lr: 6.400000000000001e-06 evaluation reward: 6.31
1783
episode: 1783 score: 9.0 memory length: 412682 epsilon: 0.3808876600101009 step
s: 492 lr: 6.400000000000001e-06 evaluation reward: 6.29
1784
episode: 1784 score: 4.0 memory length: 412957 epsilon: 0.38034316001009744 step
ps: 275 lr: 6.400000000000001e-06 evaluation reward: 6.25
1785
episode: 1785 score: 8.0 memory length: 413403 epsilon: 0.37946008001009185 step
ps: 446 lr: 6.400000000000001e-06 evaluation reward: 6.25
1786
episode: 1786 score: 8.0 memory length: 413854 epsilon: 0.3785671000100862 step
s: 451 lr: 6.400000000000001e-06 evaluation reward: 6.28
1787
episode: 1787 score: 4.0 memory length: 414111 epsilon: 0.378058240010083 step
s: 257 lr: 6.400000000000001e-06 evaluation reward: 6.26
1788
episode: 1788 score: 5.0 memory length: 414424 epsilon: 0.37743850001007906 step
ps: 313 lr: 6.400000000000001e-06 evaluation reward: 6.26
1789
episode: 1789 score: 5.0 memory length: 414732 epsilon: 0.3768286600100752 step
s: 308 lr: 6.400000000000001e-06 evaluation reward: 6.19
1790
episode: 1790 score: 5.0 memory length: 415024 epsilon: 0.37625050001007154 step
ps: 292 lr: 6.400000000000001e-06 evaluation reward: 6.19
1791
episode: 1791 score: 8.0 memory length: 415459 epsilon: 0.3753892000100661 step
s: 435 lr: 6.400000000000001e-06 evaluation reward: 6.22
1792
episode: 1792 score: 6.0 memory length: 415818 epsilon: 0.3746783800100616 step
s: 359 lr: 6.400000000000001e-06 evaluation reward: 6.24
1793
episode: 1793 score: 4.0 memory length: 416075 epsilon: 0.3741695200100584 step
s: 257 lr: 6.400000000000001e-06 evaluation reward: 6.22
1794
episode: 1794 score: 7.0 memory length: 416425 epsilon: 0.373476520010054 step
s: 350 lr: 6.400000000000001e-06 evaluation reward: 6.26
1795
episode: 1795 score: 3.0 memory length: 416637 epsilon: 0.37305676001005134 step
ps: 212 lr: 6.400000000000001e-06 evaluation reward: 6.23
1796
episode: 1796 score: 5.0 memory length: 416941 epsilon: 0.3724548400100475 step
s: 304 lr: 6.400000000000001e-06 evaluation reward: 6.23
1797
episode: 1797 score: 6.0 memory length: 417261 epsilon: 0.3718212400100435 step
s: 320 lr: 6.400000000000001e-06 evaluation reward: 6.22
1798

episode: 1798 score: 6.0 memory length: 417565 epsilon: 0.3712193200100397 step
s: 304 lr: 6.400000000000001e-06 evaluation reward: 6.25
1799
episode: 1799 score: 6.0 memory length: 417903 epsilon: 0.3705500800100355 step
s: 338 lr: 6.400000000000001e-06 evaluation reward: 6.23
1800
episode: 1800 score: 6.0 memory length: 418261 epsilon: 0.369841240010031 step
s: 358 lr: 6.400000000000001e-06 evaluation reward: 6.16
1801
episode: 1801 score: 11.0 memory length: 418794 epsilon: 0.3687859000100243 ste
ps: 533 lr: 6.400000000000001e-06 evaluation reward: 6.2
1802
episode: 1802 score: 10.0 memory length: 419290 epsilon: 0.3678038200100181 ste
ps: 496 lr: 6.400000000000001e-06 evaluation reward: 6.25
1803
episode: 1803 score: 6.0 memory length: 419648 epsilon: 0.3670949800100136 step
s: 358 lr: 6.400000000000001e-06 evaluation reward: 6.26
1804
episode: 1804 score: 7.0 memory length: 420037 epsilon: 0.36632476001000874 ste
ps: 389 lr: 6.400000000000001e-06 evaluation reward: 6.31
1805
episode: 1805 score: 5.0 memory length: 420342 epsilon: 0.3657208600100049 step
s: 305 lr: 6.400000000000001e-06 evaluation reward: 6.3
1806
episode: 1806 score: 5.0 memory length: 420649 epsilon: 0.3651130000100011 step
s: 307 lr: 6.400000000000001e-06 evaluation reward: 6.29
1807
episode: 1807 score: 4.0 memory length: 420906 epsilon: 0.36460414000999786 ste
ps: 257 lr: 6.400000000000001e-06 evaluation reward: 6.26
1808
episode: 1808 score: 11.0 memory length: 421461 epsilon: 0.3635052400099909 ste
ps: 555 lr: 6.400000000000001e-06 evaluation reward: 6.31
1809
episode: 1809 score: 8.0 memory length: 421896 epsilon: 0.36264394000998545 ste
ps: 435 lr: 6.400000000000001e-06 evaluation reward: 6.32
1810
episode: 1810 score: 6.0 memory length: 422231 epsilon: 0.36198064000998126 ste
ps: 335 lr: 6.400000000000001e-06 evaluation reward: 6.32
1811
episode: 1811 score: 8.0 memory length: 422674 epsilon: 0.3611035000099757 step
s: 443 lr: 6.400000000000001e-06 evaluation reward: 6.35
1812
episode: 1812 score: 9.0 memory length: 423133 epsilon: 0.36019468000996996 ste
ps: 459 lr: 6.400000000000001e-06 evaluation reward: 6.42
1813
episode: 1813 score: 9.0 memory length: 423584 epsilon: 0.3593017000099643 step
s: 451 lr: 6.400000000000001e-06 evaluation reward: 6.45
1814
episode: 1814 score: 7.0 memory length: 423973 epsilon: 0.35853148000995944 ste
ps: 389 lr: 6.400000000000001e-06 evaluation reward: 6.46
1815
episode: 1815 score: 8.0 memory length: 424455 epsilon: 0.3575771200099534 step
s: 482 lr: 6.400000000000001e-06 evaluation reward: 6.52
1816
episode: 1816 score: 8.0 memory length: 424888 epsilon: 0.356719780009948 step
s: 433 lr: 6.400000000000001e-06 evaluation reward: 6.57
1817
episode: 1817 score: 10.0 memory length: 425265 epsilon: 0.35597332000994325 st
eps: 377 lr: 6.400000000000001e-06 evaluation reward: 6.62
1818
episode: 1818 score: 5.0 memory length: 425590 epsilon: 0.3553298200099392 step
s: 325 lr: 6.400000000000001e-06 evaluation reward: 6.53
1819
episode: 1819 score: 10.0 memory length: 426142 epsilon: 0.35423686000993226 st
eps: 552 lr: 6.400000000000001e-06 evaluation reward: 6.61
1820

episode: 1820 score: 6.0 memory length: 426464 epsilon: 0.35359930000992823 step
ps: 322 lr: 6.400000000000001e-06 evaluation reward: 6.62
1821
episode: 1821 score: 7.0 memory length: 426850 epsilon: 0.3528350200099234 step
s: 386 lr: 6.400000000000001e-06 evaluation reward: 6.62
1822
episode: 1822 score: 10.0 memory length: 427354 epsilon: 0.3518371000099171 step
ps: 504 lr: 6.400000000000001e-06 evaluation reward: 6.62
1823
episode: 1823 score: 12.0 memory length: 427930 epsilon: 0.35069662000990987 step
eps: 576 lr: 6.400000000000001e-06 evaluation reward: 6.7
1824
episode: 1824 score: 3.0 memory length: 428143 epsilon: 0.3502748800099072 step
s: 213 lr: 6.400000000000001e-06 evaluation reward: 6.67
1825
episode: 1825 score: 4.0 memory length: 428419 epsilon: 0.34972840000990374 step
ps: 276 lr: 6.400000000000001e-06 evaluation reward: 6.66
1826
episode: 1826 score: 7.0 memory length: 428790 epsilon: 0.3489938200098991 step
s: 371 lr: 6.400000000000001e-06 evaluation reward: 6.66
1827
episode: 1827 score: 9.0 memory length: 429236 epsilon: 0.3481107400098935 step
s: 446 lr: 6.400000000000001e-06 evaluation reward: 6.67
1828
episode: 1828 score: 11.0 memory length: 429770 epsilon: 0.3470534200098868 step
ps: 534 lr: 6.400000000000001e-06 evaluation reward: 6.71
1829
episode: 1829 score: 9.0 memory length: 430291 epsilon: 0.3460218400098803 step
s: 521 lr: 6.400000000000001e-06 evaluation reward: 6.72
1830
episode: 1830 score: 3.0 memory length: 430504 epsilon: 0.3456001000098776 step
s: 213 lr: 6.400000000000001e-06 evaluation reward: 6.72
1831
episode: 1831 score: 8.0 memory length: 430907 epsilon: 0.34480216000987257 step
ps: 403 lr: 6.400000000000001e-06 evaluation reward: 6.74
1832
episode: 1832 score: 4.0 memory length: 431169 epsilon: 0.3442834000098693 step
s: 262 lr: 6.400000000000001e-06 evaluation reward: 6.7
1833
episode: 1833 score: 6.0 memory length: 431500 epsilon: 0.34362802000986514 step
ps: 331 lr: 6.400000000000001e-06 evaluation reward: 6.7
1834
episode: 1834 score: 10.0 memory length: 432030 epsilon: 0.3425786200098585 step
ps: 530 lr: 6.400000000000001e-06 evaluation reward: 6.76
1835
episode: 1835 score: 4.0 memory length: 432271 epsilon: 0.3421014400098555 step
s: 241 lr: 6.400000000000001e-06 evaluation reward: 6.76
1836
episode: 1836 score: 13.0 memory length: 432906 epsilon: 0.34084414000984753 step
eps: 635 lr: 6.400000000000001e-06 evaluation reward: 6.82
1837
episode: 1837 score: 4.0 memory length: 433205 epsilon: 0.3402521200098438 step
s: 299 lr: 6.400000000000001e-06 evaluation reward: 6.83
1838
episode: 1838 score: 11.0 memory length: 433769 epsilon: 0.3391354000098367 step
ps: 564 lr: 6.400000000000001e-06 evaluation reward: 6.84
1839
episode: 1839 score: 4.0 memory length: 434027 epsilon: 0.3386245600098335 step
s: 258 lr: 6.400000000000001e-06 evaluation reward: 6.78
1840
episode: 1840 score: 5.0 memory length: 434320 epsilon: 0.3380444200098298 step
s: 293 lr: 6.400000000000001e-06 evaluation reward: 6.77
1841
episode: 1841 score: 7.0 memory length: 434711 epsilon: 0.3372702400098249 step
s: 391 lr: 6.400000000000001e-06 evaluation reward: 6.77
1842

episode: 1842 score: 8.0 memory length: 435169 epsilon: 0.3363634000098192 step
s: 458 lr: 6.400000000000001e-06 evaluation reward: 6.8
1843
episode: 1843 score: 6.0 memory length: 435560 epsilon: 0.3355892200098143 step
s: 391 lr: 6.400000000000001e-06 evaluation reward: 6.8
1844
episode: 1844 score: 10.0 memory length: 436051 epsilon: 0.33461704000980813 st
eps: 491 lr: 6.400000000000001e-06 evaluation reward: 6.86
1845
episode: 1845 score: 12.0 memory length: 436641 epsilon: 0.33344884000980074 st
eps: 590 lr: 6.400000000000001e-06 evaluation reward: 6.93
1846
episode: 1846 score: 12.0 memory length: 437089 epsilon: 0.3325618000097951 ste
ps: 448 lr: 6.400000000000001e-06 evaluation reward: 7.01
1847
episode: 1847 score: 4.0 memory length: 437330 epsilon: 0.3320846200097921 step
s: 241 lr: 6.400000000000001e-06 evaluation reward: 6.98
1848
episode: 1848 score: 7.0 memory length: 437701 epsilon: 0.33135004000978746 ste
ps: 371 lr: 6.400000000000001e-06 evaluation reward: 7.0
1849
episode: 1849 score: 7.0 memory length: 438121 epsilon: 0.3305184400097822 step
s: 420 lr: 6.400000000000001e-06 evaluation reward: 6.98
1850
episode: 1850 score: 5.0 memory length: 438415 epsilon: 0.3299363200097785 step
s: 294 lr: 6.400000000000001e-06 evaluation reward: 6.95
1851
episode: 1851 score: 5.0 memory length: 438721 epsilon: 0.3293304400097747 step
s: 306 lr: 6.400000000000001e-06 evaluation reward: 6.93
1852
episode: 1852 score: 6.0 memory length: 439074 epsilon: 0.32863150000977026 ste
ps: 353 lr: 6.400000000000001e-06 evaluation reward: 6.91
1853
episode: 1853 score: 5.0 memory length: 439364 epsilon: 0.32805730000976663 ste
ps: 290 lr: 6.400000000000001e-06 evaluation reward: 6.89
1854
episode: 1854 score: 9.0 memory length: 439812 epsilon: 0.327170260009761 step
s: 448 lr: 6.400000000000001e-06 evaluation reward: 6.94
1855
episode: 1855 score: 7.0 memory length: 440207 epsilon: 0.32638816000975607 ste
ps: 395 lr: 6.400000000000001e-06 evaluation reward: 6.96
1856
episode: 1856 score: 6.0 memory length: 440535 epsilon: 0.32573872000975196 ste
ps: 328 lr: 6.400000000000001e-06 evaluation reward: 6.92
1857
episode: 1857 score: 6.0 memory length: 440891 epsilon: 0.3250338400097475 step
s: 356 lr: 6.400000000000001e-06 evaluation reward: 6.95
1858
episode: 1858 score: 11.0 memory length: 441347 epsilon: 0.3241309600097418 ste
ps: 456 lr: 6.400000000000001e-06 evaluation reward: 7.01
1859
episode: 1859 score: 5.0 memory length: 441638 epsilon: 0.32355478000973814 ste
ps: 291 lr: 6.400000000000001e-06 evaluation reward: 7.01
1860
episode: 1860 score: 3.0 memory length: 441867 epsilon: 0.3231013600097353 step
s: 229 lr: 6.400000000000001e-06 evaluation reward: 6.87
1861
episode: 1861 score: 6.0 memory length: 442228 epsilon: 0.32238658000973075 ste
ps: 361 lr: 6.400000000000001e-06 evaluation reward: 6.86
1862
episode: 1862 score: 9.0 memory length: 442659 epsilon: 0.32153320000972535 ste
ps: 431 lr: 6.400000000000001e-06 evaluation reward: 6.84
1863
episode: 1863 score: 6.0 memory length: 442977 epsilon: 0.32090356000972137 ste
ps: 318 lr: 6.400000000000001e-06 evaluation reward: 6.82
1864

episode: 1864 score: 6.0 memory length: 443356 epsilon: 0.3201531400097166 step
s: 379 lr: 6.400000000000001e-06 evaluation reward: 6.79
1865
episode: 1865 score: 8.0 memory length: 443764 epsilon: 0.3193453000097115 step
s: 408 lr: 6.400000000000001e-06 evaluation reward: 6.81
1866
episode: 1866 score: 10.0 memory length: 444221 epsilon: 0.3184404400097058 ste
ps: 457 lr: 6.400000000000001e-06 evaluation reward: 6.87
1867
episode: 1867 score: 5.0 memory length: 444531 epsilon: 0.3178266400097019 step
s: 310 lr: 6.400000000000001e-06 evaluation reward: 6.85
1868
episode: 1868 score: 8.0 memory length: 444954 epsilon: 0.3169891000096966 step
s: 423 lr: 6.400000000000001e-06 evaluation reward: 6.88
1869
episode: 1869 score: 5.0 memory length: 445263 epsilon: 0.31637728000969273 ste
ps: 309 lr: 6.400000000000001e-06 evaluation reward: 6.86
1870
episode: 1870 score: 11.0 memory length: 445772 epsilon: 0.31536946000968635 st
eps: 509 lr: 6.400000000000001e-06 evaluation reward: 6.92
1871
episode: 1871 score: 9.0 memory length: 446249 epsilon: 0.3144250000096804 step
s: 477 lr: 6.400000000000001e-06 evaluation reward: 6.92
1872
episode: 1872 score: 6.0 memory length: 446568 epsilon: 0.3137933800096764 step
s: 319 lr: 6.400000000000001e-06 evaluation reward: 6.93
1873
episode: 1873 score: 9.0 memory length: 447031 epsilon: 0.3128766400096706 step
s: 463 lr: 6.400000000000001e-06 evaluation reward: 6.96
1874
episode: 1874 score: 8.0 memory length: 447472 epsilon: 0.31200346000966506 ste
ps: 441 lr: 6.400000000000001e-06 evaluation reward: 6.98
1875
episode: 1875 score: 10.0 memory length: 447967 epsilon: 0.31102336000965886 st
eps: 495 lr: 6.400000000000001e-06 evaluation reward: 6.99
1876
episode: 1876 score: 4.0 memory length: 448242 epsilon: 0.3104788600096554 step
s: 275 lr: 6.400000000000001e-06 evaluation reward: 6.96
1877
episode: 1877 score: 5.0 memory length: 448531 epsilon: 0.3099066400096518 step
s: 289 lr: 6.400000000000001e-06 evaluation reward: 6.99
1878
episode: 1878 score: 8.0 memory length: 448973 epsilon: 0.30903148000964625 ste
ps: 442 lr: 6.400000000000001e-06 evaluation reward: 6.98
1879
episode: 1879 score: 10.0 memory length: 449471 epsilon: 0.30804544000964 step
s: 498 lr: 6.400000000000001e-06 evaluation reward: 7.03
1880
episode: 1880 score: 13.0 memory length: 449962 epsilon: 0.30707326000963386 st
eps: 491 lr: 6.400000000000001e-06 evaluation reward: 7.11
1881
episode: 1881 score: 6.0 memory length: 450285 epsilon: 0.3064337200096298 step
s: 323 lr: 6.400000000000001e-06 evaluation reward: 7.09
1882
episode: 1882 score: 16.0 memory length: 450892 epsilon: 0.3052318600096222 ste
ps: 607 lr: 6.400000000000001e-06 evaluation reward: 7.21
1883
episode: 1883 score: 5.0 memory length: 451201 epsilon: 0.30462004000961834 ste
ps: 309 lr: 6.400000000000001e-06 evaluation reward: 7.17
1884
episode: 1884 score: 11.0 memory length: 451575 epsilon: 0.30387952000961366 st
eps: 374 lr: 6.400000000000001e-06 evaluation reward: 7.24
1885
episode: 1885 score: 8.0 memory length: 451975 epsilon: 0.30308752000960865 ste
ps: 400 lr: 6.400000000000001e-06 evaluation reward: 7.24
1886

episode: 1886 score: 5.0 memory length: 452298 epsilon: 0.3024479800096046 step
s: 323 lr: 6.400000000000001e-06 evaluation reward: 7.21
1887
episode: 1887 score: 6.0 memory length: 452642 epsilon: 0.3017668600096003 step
s: 344 lr: 6.400000000000001e-06 evaluation reward: 7.23
1888
episode: 1888 score: 10.0 memory length: 453120 epsilon: 0.3008204200095943 ste
ps: 478 lr: 6.400000000000001e-06 evaluation reward: 7.28
1889
episode: 1889 score: 6.0 memory length: 453493 epsilon: 0.30008188000958963 ste
ps: 373 lr: 6.400000000000001e-06 evaluation reward: 7.29
1890
episode: 1890 score: 14.0 memory length: 453998 epsilon: 0.2990819800095833 ste
ps: 505 lr: 6.400000000000001e-06 evaluation reward: 7.38
1891
episode: 1891 score: 10.0 memory length: 454386 epsilon: 0.29831374000957844 st
eps: 388 lr: 6.400000000000001e-06 evaluation reward: 7.4
1892
episode: 1892 score: 10.0 memory length: 454919 epsilon: 0.29725840000957177 st
eps: 533 lr: 6.400000000000001e-06 evaluation reward: 7.44
1893
episode: 1893 score: 5.0 memory length: 455228 epsilon: 0.2966465800095679 step
s: 309 lr: 6.400000000000001e-06 evaluation reward: 7.45
1894
episode: 1894 score: 6.0 memory length: 455604 epsilon: 0.2959021000095632 step
s: 376 lr: 6.400000000000001e-06 evaluation reward: 7.44
1895
episode: 1895 score: 6.0 memory length: 455956 epsilon: 0.2952051400095588 step
s: 352 lr: 6.400000000000001e-06 evaluation reward: 7.47
1896
episode: 1896 score: 4.0 memory length: 456214 epsilon: 0.29469430000955554 ste
ps: 258 lr: 6.400000000000001e-06 evaluation reward: 7.46
1897
episode: 1897 score: 8.0 memory length: 456615 epsilon: 0.2939003200095505 step
s: 401 lr: 6.400000000000001e-06 evaluation reward: 7.48
1898
episode: 1898 score: 4.0 memory length: 456875 epsilon: 0.29338552000954726 ste
ps: 260 lr: 6.400000000000001e-06 evaluation reward: 7.46
1899
episode: 1899 score: 7.0 memory length: 457274 epsilon: 0.29259550000954226 ste
ps: 399 lr: 6.400000000000001e-06 evaluation reward: 7.47
1900
episode: 1900 score: 11.0 memory length: 457654 epsilon: 0.2918431000095375 ste
ps: 380 lr: 6.400000000000001e-06 evaluation reward: 7.52
1901
episode: 1901 score: 8.0 memory length: 458067 epsilon: 0.29102536000953233 ste
ps: 413 lr: 6.400000000000001e-06 evaluation reward: 7.49
1902
episode: 1902 score: 7.0 memory length: 458472 epsilon: 0.29022346000952726 ste
ps: 405 lr: 6.400000000000001e-06 evaluation reward: 7.46
1903
episode: 1903 score: 3.0 memory length: 458685 epsilon: 0.2898017200095246 step
s: 213 lr: 6.400000000000001e-06 evaluation reward: 7.43
1904
episode: 1904 score: 8.0 memory length: 459094 epsilon: 0.28899190000951946 ste
ps: 409 lr: 6.400000000000001e-06 evaluation reward: 7.44
1905
episode: 1905 score: 18.0 memory length: 459725 epsilon: 0.28774252000951156 st
eps: 631 lr: 6.400000000000001e-06 evaluation reward: 7.57
1906
episode: 1906 score: 11.0 memory length: 460271 epsilon: 0.2866614400095047 ste
ps: 546 lr: 6.400000000000001e-06 evaluation reward: 7.63
1907
episode: 1907 score: 9.0 memory length: 460744 epsilon: 0.2857249000094988 step
s: 473 lr: 6.400000000000001e-06 evaluation reward: 7.68
1908

episode: 1908 score: 13.0 memory length: 461291 epsilon: 0.28464184000949194 step
eps: 547 lr: 6.400000000000001e-06 evaluation reward: 7.7
1909
episode: 1909 score: 7.0 memory length: 461700 epsilon: 0.2838320200094868 step
s: 409 lr: 6.400000000000001e-06 evaluation reward: 7.69
1910
episode: 1910 score: 7.0 memory length: 462067 epsilon: 0.2831053600094822 step
s: 367 lr: 6.400000000000001e-06 evaluation reward: 7.7
1911
episode: 1911 score: 10.0 memory length: 462564 epsilon: 0.282121300009476 step
s: 497 lr: 6.400000000000001e-06 evaluation reward: 7.72
1912
episode: 1912 score: 5.0 memory length: 462854 epsilon: 0.28154710000947236 step
ps: 290 lr: 6.400000000000001e-06 evaluation reward: 7.68
1913
episode: 1913 score: 9.0 memory length: 463177 epsilon: 0.2809075600094683 step
s: 323 lr: 6.400000000000001e-06 evaluation reward: 7.68
1914
episode: 1914 score: 8.0 memory length: 463543 epsilon: 0.28018288000946373 step
ps: 366 lr: 6.400000000000001e-06 evaluation reward: 7.69
1915
episode: 1915 score: 8.0 memory length: 463950 epsilon: 0.27937702000945863 step
ps: 407 lr: 6.400000000000001e-06 evaluation reward: 7.69
1916
episode: 1917 score: 7.0 memory length: 464686 epsilon: 0.2779197400094494 step
s: 376 lr: 6.400000000000001e-06 evaluation reward: 7.64
1918
episode: 1918 score: 6.0 memory length: 465008 epsilon: 0.2772821800094454 step
s: 322 lr: 6.400000000000001e-06 evaluation reward: 7.65
1919
episode: 1919 score: 7.0 memory length: 465357 epsilon: 0.276591160009441 step
s: 349 lr: 6.400000000000001e-06 evaluation reward: 7.62
1920
episode: 1920 score: 10.0 memory length: 465685 epsilon: 0.2759417200094369 step
ps: 328 lr: 6.400000000000001e-06 evaluation reward: 7.66
1921
episode: 1921 score: 4.0 memory length: 465927 epsilon: 0.27546256000943387 step
ps: 242 lr: 6.400000000000001e-06 evaluation reward: 7.63
1922
episode: 1922 score: 9.0 memory length: 466408 epsilon: 0.27451018000942784 step
ps: 481 lr: 6.400000000000001e-06 evaluation reward: 7.62
1923
episode: 1923 score: 10.0 memory length: 466899 epsilon: 0.2735380000094217 step
ps: 491 lr: 6.400000000000001e-06 evaluation reward: 7.6
1924
episode: 1924 score: 8.0 memory length: 467345 epsilon: 0.2726549200094161 step
s: 446 lr: 6.400000000000001e-06 evaluation reward: 7.65
1925
episode: 1925 score: 9.0 memory length: 467833 epsilon: 0.27168868000941 steps:
488 lr: 6.400000000000001e-06 evaluation reward: 7.7
1926
episode: 1926 score: 8.0 memory length: 468266 epsilon: 0.27083134000940456 step
ps: 433 lr: 6.400000000000001e-06 evaluation reward: 7.71
1927
episode: 1927 score: 6.0 memory length: 468632 epsilon: 0.2701066600094 steps:
366 lr: 6.400000000000001e-06 evaluation reward: 7.68
1928
episode: 1928 score: 12.0 memory length: 469092 epsilon: 0.2691958600093942 step
ps: 460 lr: 6.400000000000001e-06 evaluation reward: 7.69
1929
episode: 1929 score: 4.0 memory length: 469352 epsilon: 0.26868106000939096 step
ps: 260 lr: 6.400000000000001e-06 evaluation reward: 7.64
1930
episode: 1930 score: 4.0 memory length: 469591 epsilon: 0.26820784000938797 step
ps: 239 lr: 6.400000000000001e-06 evaluation reward: 7.65
1931

episode: 1931 score: 10.0 memory length: 470095 epsilon: 0.26720992000938165 step
eps: 504 lr: 6.400000000000001e-06 evaluation reward: 7.67
1932
episode: 1932 score: 5.0 memory length: 470402 epsilon: 0.2666020600093778 step
s: 307 lr: 6.400000000000001e-06 evaluation reward: 7.68
1933
episode: 1933 score: 13.0 memory length: 471025 epsilon: 0.26536852000937 step
s: 623 lr: 6.400000000000001e-06 evaluation reward: 7.75
1934
episode: 1934 score: 8.0 memory length: 471430 epsilon: 0.26456662000936493 step
ps: 405 lr: 6.400000000000001e-06 evaluation reward: 7.73
1935
episode: 1935 score: 8.0 memory length: 471859 epsilon: 0.26371720000935955 step
ps: 429 lr: 6.400000000000001e-06 evaluation reward: 7.77
1936
episode: 1936 score: 11.0 memory length: 472234 epsilon: 0.26297470000935486 step
eps: 375 lr: 6.400000000000001e-06 evaluation reward: 7.75
1937
episode: 1937 score: 10.0 memory length: 472738 epsilon: 0.26197678000934854 step
eps: 504 lr: 6.400000000000001e-06 evaluation reward: 7.81
1938
episode: 1938 score: 7.0 memory length: 473103 epsilon: 0.26125408000934397 step
ps: 365 lr: 6.400000000000001e-06 evaluation reward: 7.77
1939
episode: 1939 score: 8.0 memory length: 473538 epsilon: 0.2603927800093385 step
s: 435 lr: 6.400000000000001e-06 evaluation reward: 7.81
1940
episode: 1940 score: 4.0 memory length: 473797 epsilon: 0.2598799600093353 step
s: 259 lr: 6.400000000000001e-06 evaluation reward: 7.8
1941
episode: 1941 score: 11.0 memory length: 474317 epsilon: 0.25885036000932876 step
eps: 520 lr: 6.400000000000001e-06 evaluation reward: 7.84
1942
episode: 1942 score: 6.0 memory length: 474638 epsilon: 0.25821478000932474 step
ps: 321 lr: 6.400000000000001e-06 evaluation reward: 7.82
1943
episode: 1943 score: 5.0 memory length: 474930 epsilon: 0.2576366200093211 step
s: 292 lr: 6.400000000000001e-06 evaluation reward: 7.81
1944
episode: 1944 score: 5.0 memory length: 475222 epsilon: 0.2570584600093174 step
s: 292 lr: 6.400000000000001e-06 evaluation reward: 7.76
1945
episode: 1945 score: 8.0 memory length: 475639 epsilon: 0.2562328000093122 step
s: 417 lr: 6.400000000000001e-06 evaluation reward: 7.72
1946
episode: 1946 score: 10.0 memory length: 476050 epsilon: 0.25541902000930705 step
eps: 411 lr: 6.400000000000001e-06 evaluation reward: 7.7
1947
episode: 1947 score: 8.0 memory length: 476431 epsilon: 0.2546646400093023 step
s: 381 lr: 6.400000000000001e-06 evaluation reward: 7.74
1948
episode: 1948 score: 5.0 memory length: 476757 epsilon: 0.2540191600092982 step
s: 326 lr: 6.400000000000001e-06 evaluation reward: 7.72
1949
episode: 1949 score: 8.0 memory length: 477212 epsilon: 0.2531182600092925 step
s: 455 lr: 6.400000000000001e-06 evaluation reward: 7.73
1950
episode: 1950 score: 8.0 memory length: 477627 epsilon: 0.2522965600092873 step
s: 415 lr: 6.400000000000001e-06 evaluation reward: 7.76
1951
episode: 1951 score: 3.0 memory length: 477838 epsilon: 0.25187878000928465 step
ps: 211 lr: 6.400000000000001e-06 evaluation reward: 7.74
1952
episode: 1952 score: 7.0 memory length: 478203 epsilon: 0.2511560800092801 step
s: 365 lr: 6.400000000000001e-06 evaluation reward: 7.75
1953

episode: 1953 score: 13.0 memory length: 478788 epsilon: 0.24999778000927275 st
eps: 585 lr: 6.400000000000001e-06 evaluation reward: 7.83
1954
episode: 1954 score: 8.0 memory length: 479182 epsilon: 0.24921766000926782 ste
ps: 394 lr: 6.400000000000001e-06 evaluation reward: 7.82
1955
episode: 1955 score: 8.0 memory length: 479637 epsilon: 0.24831676000926212 ste
ps: 455 lr: 6.400000000000001e-06 evaluation reward: 7.83
1956
episode: 1956 score: 10.0 memory length: 480152 epsilon: 0.24729706000925566 st
eps: 515 lr: 6.400000000000001e-06 evaluation reward: 7.87
1957
episode: 1957 score: 16.0 memory length: 480789 epsilon: 0.24603580000924768 st
eps: 637 lr: 6.400000000000001e-06 evaluation reward: 7.97
1958
episode: 1958 score: 14.0 memory length: 481338 epsilon: 0.2449487800092408 ste
ps: 549 lr: 6.400000000000001e-06 evaluation reward: 8.0
1959
episode: 1959 score: 7.0 memory length: 481752 epsilon: 0.24412906000923562 ste
ps: 414 lr: 6.400000000000001e-06 evaluation reward: 8.02
1960
episode: 1960 score: 6.0 memory length: 482103 epsilon: 0.24343408000923122 ste
ps: 351 lr: 6.400000000000001e-06 evaluation reward: 8.05
1961
episode: 1961 score: 17.0 memory length: 482746 epsilon: 0.24216094000922317 st
eps: 643 lr: 6.400000000000001e-06 evaluation reward: 8.16
1962
episode: 1962 score: 5.0 memory length: 483036 epsilon: 0.24158674000921954 ste
ps: 290 lr: 6.400000000000001e-06 evaluation reward: 8.12
1963
episode: 1963 score: 9.0 memory length: 483484 epsilon: 0.24069970000921392 ste
ps: 448 lr: 6.400000000000001e-06 evaluation reward: 8.15
1964
episode: 1964 score: 5.0 memory length: 483793 epsilon: 0.24008788000921005 ste
ps: 309 lr: 6.400000000000001e-06 evaluation reward: 8.14
1965
episode: 1965 score: 9.0 memory length: 484244 epsilon: 0.2391949000092044 step
s: 451 lr: 6.400000000000001e-06 evaluation reward: 8.15
1966
episode: 1966 score: 5.0 memory length: 484535 epsilon: 0.23861872000920076 ste
ps: 291 lr: 6.400000000000001e-06 evaluation reward: 8.1
1967
episode: 1967 score: 5.0 memory length: 484843 epsilon: 0.2380088800091969 step
s: 308 lr: 6.400000000000001e-06 evaluation reward: 8.1
1968
episode: 1968 score: 7.0 memory length: 485193 epsilon: 0.23731588000919251 ste
ps: 350 lr: 6.400000000000001e-06 evaluation reward: 8.09
1969
episode: 1969 score: 10.0 memory length: 485578 epsilon: 0.2365535800091877 ste
ps: 385 lr: 6.400000000000001e-06 evaluation reward: 8.14
1970
episode: 1970 score: 8.0 memory length: 485962 epsilon: 0.23579326000918288 ste
ps: 384 lr: 6.400000000000001e-06 evaluation reward: 8.11
1971
episode: 1971 score: 8.0 memory length: 486386 epsilon: 0.23495374000917757 ste
ps: 424 lr: 6.400000000000001e-06 evaluation reward: 8.1
1972
episode: 1972 score: 6.0 memory length: 486702 epsilon: 0.2343280600091736 step
s: 316 lr: 6.400000000000001e-06 evaluation reward: 8.1
1973
episode: 1973 score: 16.0 memory length: 487372 epsilon: 0.23300146000916522 st
eps: 670 lr: 6.400000000000001e-06 evaluation reward: 8.17
1974
episode: 1974 score: 9.0 memory length: 487838 epsilon: 0.23207878000915938 ste
ps: 466 lr: 6.400000000000001e-06 evaluation reward: 8.18
1975

episode: 1975 score: 6.0 memory length: 488174 epsilon: 0.23141350000915517 step
ps: 336 lr: 6.400000000000001e-06 evaluation reward: 8.14
1976
episode: 1976 score: 10.0 memory length: 488660 epsilon: 0.23045122000914908 step
eps: 486 lr: 6.400000000000001e-06 evaluation reward: 8.2
1977
episode: 1977 score: 13.0 memory length: 489151 epsilon: 0.22947904000914293 step
eps: 491 lr: 6.400000000000001e-06 evaluation reward: 8.28
1978
episode: 1978 score: 10.0 memory length: 489611 epsilon: 0.22856824000913717 step
eps: 460 lr: 6.400000000000001e-06 evaluation reward: 8.3
1979
episode: 1979 score: 8.0 memory length: 490007 epsilon: 0.2277841600091322 step
s: 396 lr: 6.400000000000001e-06 evaluation reward: 8.28
1980
episode: 1980 score: 5.0 memory length: 490354 epsilon: 0.22709710000912786 step
ps: 347 lr: 6.400000000000001e-06 evaluation reward: 8.2
1981
episode: 1981 score: 9.0 memory length: 490840 epsilon: 0.22613482000912177 step
ps: 486 lr: 6.400000000000001e-06 evaluation reward: 8.23
1982
episode: 1982 score: 4.0 memory length: 491083 epsilon: 0.22565368000911873 step
ps: 243 lr: 6.400000000000001e-06 evaluation reward: 8.11
1983
episode: 1983 score: 13.0 memory length: 491663 epsilon: 0.22450528000911146 step
eps: 580 lr: 6.400000000000001e-06 evaluation reward: 8.19
1984
episode: 1984 score: 3.0 memory length: 491875 epsilon: 0.2240855200091088 step
s: 212 lr: 6.400000000000001e-06 evaluation reward: 8.11
1985
episode: 1985 score: 10.0 memory length: 492398 epsilon: 0.22304998000910226 step
eps: 523 lr: 6.400000000000001e-06 evaluation reward: 8.13
1986
episode: 1986 score: 9.0 memory length: 492858 epsilon: 0.2221391800090965 step
s: 460 lr: 6.400000000000001e-06 evaluation reward: 8.17
1987
episode: 1987 score: 11.0 memory length: 493430 epsilon: 0.22100662000908933 step
eps: 572 lr: 6.400000000000001e-06 evaluation reward: 8.22
1988
episode: 1988 score: 10.0 memory length: 493957 epsilon: 0.21996316000908273 step
eps: 527 lr: 6.400000000000001e-06 evaluation reward: 8.22
1989
episode: 1989 score: 9.0 memory length: 494423 epsilon: 0.2190404800090769 step
s: 466 lr: 6.400000000000001e-06 evaluation reward: 8.25
1990
episode: 1990 score: 6.0 memory length: 494759 epsilon: 0.21837520000907268 step
ps: 336 lr: 6.400000000000001e-06 evaluation reward: 8.17
1991
episode: 1991 score: 14.0 memory length: 495300 epsilon: 0.2173040200090659 step
ps: 541 lr: 6.400000000000001e-06 evaluation reward: 8.21
1992
episode: 1992 score: 10.0 memory length: 495814 epsilon: 0.21628630000905946 step
eps: 514 lr: 6.400000000000001e-06 evaluation reward: 8.21
1993
episode: 1993 score: 9.0 memory length: 496284 epsilon: 0.21535570000905357 step
ps: 470 lr: 6.400000000000001e-06 evaluation reward: 8.25
1994
episode: 1994 score: 15.0 memory length: 496706 epsilon: 0.2145201400090483 step
ps: 422 lr: 6.400000000000001e-06 evaluation reward: 8.34
1995
episode: 1995 score: 7.0 memory length: 497108 epsilon: 0.21372418000904325 step
ps: 402 lr: 6.400000000000001e-06 evaluation reward: 8.35
1996
episode: 1996 score: 6.0 memory length: 497467 epsilon: 0.21301336000903875 step
ps: 359 lr: 6.400000000000001e-06 evaluation reward: 8.37
1997

episode: 1997 score: 8.0 memory length: 497868 epsilon: 0.21221938000903373 step
ps: 401 lr: 6.400000000000001e-06 evaluation reward: 8.37
1998
episode: 1998 score: 7.0 memory length: 498294 epsilon: 0.2113759000090284 step
s: 426 lr: 6.400000000000001e-06 evaluation reward: 8.4
1999
episode: 1999 score: 8.0 memory length: 498698 epsilon: 0.21057598000902333 step
ps: 404 lr: 6.400000000000001e-06 evaluation reward: 8.41
2000
episode: 2000 score: 7.0 memory length: 499107 epsilon: 0.2097661600090182 step
s: 409 lr: 6.400000000000001e-06 evaluation reward: 8.37
2001
episode: 2001 score: 5.0 memory length: 499432 epsilon: 0.20912266000901414 step
ps: 325 lr: 6.400000000000001e-06 evaluation reward: 8.34
2002
episode: 2002 score: 6.0 memory length: 499756 epsilon: 0.20848114000901008 step
ps: 324 lr: 6.400000000000001e-06 evaluation reward: 8.33
2003
episode: 2003 score: 8.0 memory length: 500158 epsilon: 0.20768518000900504 step
ps: 402 lr: 2.5600000000000001e-06 evaluation reward: 8.38
2004
episode: 2004 score: 9.0 memory length: 500625 epsilon: 0.2067605200089992 step
s: 467 lr: 2.5600000000000001e-06 evaluation reward: 8.39
2005
episode: 2005 score: 5.0 memory length: 500933 epsilon: 0.20615068000899534 step
ps: 308 lr: 2.5600000000000001e-06 evaluation reward: 8.26
2006
episode: 2006 score: 4.0 memory length: 501195 epsilon: 0.20563192000899205 step
ps: 262 lr: 2.5600000000000001e-06 evaluation reward: 8.19
2007
episode: 2007 score: 5.0 memory length: 501499 epsilon: 0.20503000000898824 step
ps: 304 lr: 2.5600000000000001e-06 evaluation reward: 8.15
2008
episode: 2008 score: 21.0 memory length: 502140 epsilon: 0.20376082000898021 step
eps: 641 lr: 2.5600000000000001e-06 evaluation reward: 8.23
2009
episode: 2009 score: 6.0 memory length: 502513 epsilon: 0.20302228000897554 step
ps: 373 lr: 2.5600000000000001e-06 evaluation reward: 8.22
2010
episode: 2010 score: 8.0 memory length: 503002 epsilon: 0.20205406000896942 step
ps: 489 lr: 2.5600000000000001e-06 evaluation reward: 8.23
2011
episode: 2011 score: 9.0 memory length: 503490 epsilon: 0.2010878200089633 step
s: 488 lr: 2.5600000000000001e-06 evaluation reward: 8.22
2012
episode: 2012 score: 5.0 memory length: 503814 epsilon: 0.20044630000895924 step
ps: 324 lr: 2.5600000000000001e-06 evaluation reward: 8.22
2013
episode: 2013 score: 6.0 memory length: 504190 epsilon: 0.19970182000895453 step
ps: 376 lr: 2.5600000000000001e-06 evaluation reward: 8.19
2014
episode: 2014 score: 6.0 memory length: 504529 epsilon: 0.1990306000089503 step
s: 339 lr: 2.5600000000000001e-06 evaluation reward: 8.17
2015
episode: 2015 score: 8.0 memory length: 504941 epsilon: 0.19821484000894513 step
ps: 412 lr: 2.5600000000000001e-06 evaluation reward: 8.17
2016
episode: 2016 score: 10.0 memory length: 505419 epsilon: 0.19726840000893914 step
eps: 478 lr: 2.5600000000000001e-06 evaluation reward: 8.21
2017
episode: 2017 score: 11.0 memory length: 505796 epsilon: 0.19652194000893441 step
eps: 377 lr: 2.5600000000000001e-06 evaluation reward: 8.25
2018
episode: 2018 score: 12.0 memory length: 506398 epsilon: 0.19532998000892687 step
eps: 602 lr: 2.5600000000000001e-06 evaluation reward: 8.31
2019

episode: 2019	score: 6.0	memory length: 506771	epsilon: 0.1945914400089222	step
s: 373	lr: 2.5600000000000001e-06	evaluation reward: 8.3		
2020				
episode: 2020	score: 7.0	memory length: 507145	epsilon: 0.19385092000891752	ste
ps: 374	lr: 2.5600000000000001e-06	evaluation reward: 8.27		
2021				
episode: 2021	score: 10.0	memory length: 507587	epsilon: 0.19297576000891198	st
eps: 442	lr: 2.5600000000000001e-06	evaluation reward: 8.33		
2022				
episode: 2022	score: 11.0	memory length: 508130	epsilon: 0.19190062000890518	st
eps: 543	lr: 2.5600000000000001e-06	evaluation reward: 8.35		
2023				
episode: 2023	score: 5.0	memory length: 508416	epsilon: 0.1913343400089016	step
s: 286	lr: 2.5600000000000001e-06	evaluation reward: 8.3		
2024				
episode: 2024	score: 5.0	memory length: 508723	epsilon: 0.19072648000889775	ste
ps: 307	lr: 2.5600000000000001e-06	evaluation reward: 8.27		
2025				
episode: 2025	score: 5.0	memory length: 509029	epsilon: 0.19012060000889391	ste
ps: 306	lr: 2.5600000000000001e-06	evaluation reward: 8.23		
2026				
episode: 2026	score: 10.0	memory length: 509543	epsilon: 0.18910288000888748	st
eps: 514	lr: 2.5600000000000001e-06	evaluation reward: 8.25		
2027				
episode: 2027	score: 12.0	memory length: 509949	epsilon: 0.1882990000088824	ste
ps: 406	lr: 2.5600000000000001e-06	evaluation reward: 8.31		
2028				
episode: 2028	score: 9.0	memory length: 510429	epsilon: 0.18734860000887638	ste
ps: 480	lr: 2.5600000000000001e-06	evaluation reward: 8.28		
2029				
episode: 2029	score: 3.0	memory length: 510641	epsilon: 0.18692884000887372	ste
ps: 212	lr: 2.5600000000000001e-06	evaluation reward: 8.27		
2030				
episode: 2030	score: 9.0	memory length: 511069	epsilon: 0.18608140000886836	ste
ps: 428	lr: 2.5600000000000001e-06	evaluation reward: 8.32		
2031				
episode: 2031	score: 12.0	memory length: 511667	epsilon: 0.18489736000886087	st
eps: 598	lr: 2.5600000000000001e-06	evaluation reward: 8.34		
2032				
episode: 2032	score: 6.0	memory length: 512006	epsilon: 0.18422614000885662	ste
ps: 339	lr: 2.5600000000000001e-06	evaluation reward: 8.35		
2033				
episode: 2033	score: 10.0	memory length: 512507	epsilon: 0.18323416000885034	st
eps: 501	lr: 2.5600000000000001e-06	evaluation reward: 8.32		
2034				
episode: 2034	score: 5.0	memory length: 512814	epsilon: 0.1826263000088465	step
s: 307	lr: 2.5600000000000001e-06	evaluation reward: 8.29		
2035				
episode: 2035	score: 10.0	memory length: 513356	epsilon: 0.1815531400088397	ste
ps: 542	lr: 2.5600000000000001e-06	evaluation reward: 8.31		
2036				
episode: 2036	score: 13.0	memory length: 513846	epsilon: 0.18058294000883357	st
eps: 490	lr: 2.5600000000000001e-06	evaluation reward: 8.33		
2037				
episode: 2037	score: 14.0	memory length: 514359	epsilon: 0.17956720000882714	st
eps: 513	lr: 2.5600000000000001e-06	evaluation reward: 8.37		
2038				
episode: 2038	score: 6.0	memory length: 514690	epsilon: 0.178911820008823	step
s: 331	lr: 2.5600000000000001e-06	evaluation reward: 8.36		
2039				
episode: 2039	score: 18.0	memory length: 515303	epsilon: 0.17769808000881532	st
eps: 613	lr: 2.5600000000000001e-06	evaluation reward: 8.46		
2040				
episode: 2040	score: 13.0	memory length: 515861	epsilon: 0.17659324000880833	st
eps: 558	lr: 2.5600000000000001e-06	evaluation reward: 8.55		
2041				

episode: 2041 score: 8.0 memory length: 516256 epsilon: 0.17581114000880338 ste
ps: 395 lr: 2.5600000000000001e-06 evaluation reward: 8.52
2042
episode: 2042 score: 5.0 memory length: 516546 epsilon: 0.17523694000879975 ste
ps: 290 lr: 2.5600000000000001e-06 evaluation reward: 8.51
2043
episode: 2043 score: 5.0 memory length: 516835 epsilon: 0.17466472000879613 ste
ps: 289 lr: 2.5600000000000001e-06 evaluation reward: 8.51
2044
episode: 2044 score: 4.0 memory length: 517093 epsilon: 0.1741538800087929 step
s: 258 lr: 2.5600000000000001e-06 evaluation reward: 8.5
2045
episode: 2045 score: 8.0 memory length: 517546 epsilon: 0.17325694000878722 ste
ps: 453 lr: 2.5600000000000001e-06 evaluation reward: 8.5
2046
episode: 2046 score: 11.0 memory length: 518062 epsilon: 0.17223526000878076 st
eps: 516 lr: 2.5600000000000001e-06 evaluation reward: 8.51
2047
episode: 2047 score: 12.0 memory length: 518469 epsilon: 0.17142940000877566 st
eps: 407 lr: 2.5600000000000001e-06 evaluation reward: 8.55
2048
episode: 2048 score: 6.0 memory length: 518849 epsilon: 0.1706770000087709 step
s: 380 lr: 2.5600000000000001e-06 evaluation reward: 8.56
2049
episode: 2049 score: 6.0 memory length: 519170 epsilon: 0.17004142000876687 ste
ps: 321 lr: 2.5600000000000001e-06 evaluation reward: 8.54
2050
episode: 2050 score: 9.0 memory length: 519606 epsilon: 0.1691781400087614 step
s: 436 lr: 2.5600000000000001e-06 evaluation reward: 8.55
2051
episode: 2051 score: 10.0 memory length: 520113 epsilon: 0.16817428000875506 st
eps: 507 lr: 2.5600000000000001e-06 evaluation reward: 8.62
2052
episode: 2052 score: 9.0 memory length: 520611 epsilon: 0.16718824000874882 ste
ps: 498 lr: 2.5600000000000001e-06 evaluation reward: 8.64
2053
episode: 2053 score: 9.0 memory length: 521071 epsilon: 0.16627744000874306 ste
ps: 460 lr: 2.5600000000000001e-06 evaluation reward: 8.6
2054
episode: 2054 score: 6.0 memory length: 521400 epsilon: 0.16562602000873894 ste
ps: 329 lr: 2.5600000000000001e-06 evaluation reward: 8.58
2055
episode: 2055 score: 10.0 memory length: 521894 epsilon: 0.16464790000873275 st
eps: 494 lr: 2.5600000000000001e-06 evaluation reward: 8.6
2056
episode: 2056 score: 8.0 memory length: 522372 epsilon: 0.16370146000872676 ste
ps: 478 lr: 2.5600000000000001e-06 evaluation reward: 8.58
2057
episode: 2057 score: 6.0 memory length: 522746 epsilon: 0.16296094000872208 ste
ps: 374 lr: 2.5600000000000001e-06 evaluation reward: 8.48
2058
episode: 2058 score: 7.0 memory length: 523155 epsilon: 0.16215112000871695 ste
ps: 409 lr: 2.5600000000000001e-06 evaluation reward: 8.41
2059
episode: 2059 score: 9.0 memory length: 523618 epsilon: 0.16123438000871115 ste
ps: 463 lr: 2.5600000000000001e-06 evaluation reward: 8.43
2060
episode: 2060 score: 13.0 memory length: 524058 epsilon: 0.16036318000870564 st
eps: 440 lr: 2.5600000000000001e-06 evaluation reward: 8.5
2061
episode: 2061 score: 8.0 memory length: 524487 epsilon: 0.15951376000870027 ste
ps: 429 lr: 2.5600000000000001e-06 evaluation reward: 8.41
2062
episode: 2062 score: 7.0 memory length: 524904 epsilon: 0.15868810000869504 ste
ps: 417 lr: 2.5600000000000001e-06 evaluation reward: 8.43
2063

episode: 2063 score: 6.0 memory length: 525274 epsilon: 0.1579555000086904 step
s: 370 lr: 2.5600000000000001e-06 evaluation reward: 8.4
2064
episode: 2064 score: 9.0 memory length: 525727 epsilon: 0.15705856000868473 ste
ps: 453 lr: 2.5600000000000001e-06 evaluation reward: 8.44
2065
episode: 2065 score: 10.0 memory length: 526194 epsilon: 0.15613390000867888 st
eps: 467 lr: 2.5600000000000001e-06 evaluation reward: 8.45
2066
episode: 2066 score: 12.0 memory length: 526653 epsilon: 0.15522508000867313 st
eps: 459 lr: 2.5600000000000001e-06 evaluation reward: 8.52
2067
episode: 2067 score: 6.0 memory length: 526996 epsilon: 0.15454594000866884 ste
ps: 343 lr: 2.5600000000000001e-06 evaluation reward: 8.53
2068
episode: 2068 score: 3.0 memory length: 527206 epsilon: 0.1541301400086662 step
s: 210 lr: 2.5600000000000001e-06 evaluation reward: 8.49
2069
episode: 2069 score: 8.0 memory length: 527609 epsilon: 0.15333220000866116 ste
ps: 403 lr: 2.5600000000000001e-06 evaluation reward: 8.47
2070
episode: 2071 score: 18.0 memory length: 528723 epsilon: 0.1511264800086472 ste
ps: 534 lr: 2.5600000000000001e-06 evaluation reward: 8.6
2072
episode: 2072 score: 10.0 memory length: 529214 epsilon: 0.15015430000864105 st
eps: 491 lr: 2.5600000000000001e-06 evaluation reward: 8.64
2073
episode: 2073 score: 9.0 memory length: 529643 epsilon: 0.14930488000863568 ste
ps: 429 lr: 2.5600000000000001e-06 evaluation reward: 8.57
2074
episode: 2074 score: 8.0 memory length: 530078 epsilon: 0.14844358000863023 ste
ps: 435 lr: 2.5600000000000001e-06 evaluation reward: 8.56
2075
episode: 2075 score: 12.0 memory length: 530607 epsilon: 0.1473961600086236 ste
ps: 529 lr: 2.5600000000000001e-06 evaluation reward: 8.62
2076
episode: 2076 score: 14.0 memory length: 531140 epsilon: 0.14634082000861692 st
eps: 533 lr: 2.5600000000000001e-06 evaluation reward: 8.66
2077
episode: 2077 score: 16.0 memory length: 531767 epsilon: 0.14509936000860907 st
eps: 627 lr: 2.5600000000000001e-06 evaluation reward: 8.69
2078
episode: 2078 score: 6.0 memory length: 532091 epsilon: 0.144457840008605 step
s: 324 lr: 2.5600000000000001e-06 evaluation reward: 8.65
2079
episode: 2079 score: 10.0 memory length: 532478 epsilon: 0.14369158000860016 st
eps: 387 lr: 2.5600000000000001e-06 evaluation reward: 8.67
2080
episode: 2080 score: 11.0 memory length: 532982 epsilon: 0.14269366000859385 st
eps: 504 lr: 2.5600000000000001e-06 evaluation reward: 8.73
2081
episode: 2081 score: 12.0 memory length: 533533 epsilon: 0.14160268000858695 st
eps: 551 lr: 2.5600000000000001e-06 evaluation reward: 8.76
2082
episode: 2082 score: 9.0 memory length: 533976 epsilon: 0.1407255400085814 step
s: 443 lr: 2.5600000000000001e-06 evaluation reward: 8.81
2083
episode: 2083 score: 12.0 memory length: 534512 epsilon: 0.13966426000857468 st
eps: 536 lr: 2.5600000000000001e-06 evaluation reward: 8.8
2084
episode: 2084 score: 7.0 memory length: 534899 epsilon: 0.13889800000856983 ste
ps: 387 lr: 2.5600000000000001e-06 evaluation reward: 8.84
2085
episode: 2085 score: 9.0 memory length: 535330 epsilon: 0.13804462000856443 ste
ps: 431 lr: 2.5600000000000001e-06 evaluation reward: 8.83
2086

episode: 2086 score: 12.0 memory length: 535830 epsilon: 0.13705462000855817 st
eps: 500 lr: 2.5600000000000001e-06 evaluation reward: 8.86
2087
episode: 2087 score: 6.0 memory length: 536184 epsilon: 0.13635370000855374 ste
ps: 354 lr: 2.5600000000000001e-06 evaluation reward: 8.81
2088
episode: 2088 score: 12.0 memory length: 536685 epsilon: 0.13536172000854746 st
eps: 501 lr: 2.5600000000000001e-06 evaluation reward: 8.83
2089
episode: 2089 score: 5.0 memory length: 537008 epsilon: 0.1347221800085434 step
s: 323 lr: 2.5600000000000001e-06 evaluation reward: 8.79
2090
episode: 2090 score: 8.0 memory length: 537480 epsilon: 0.1337876200085375 step
s: 472 lr: 2.5600000000000001e-06 evaluation reward: 8.81
2091
episode: 2091 score: 8.0 memory length: 537901 epsilon: 0.13295404000853223 ste
ps: 421 lr: 2.5600000000000001e-06 evaluation reward: 8.75
2092
episode: 2092 score: 10.0 memory length: 538447 epsilon: 0.1318729600085254 ste
ps: 546 lr: 2.5600000000000001e-06 evaluation reward: 8.75
2093
episode: 2093 score: 12.0 memory length: 539015 epsilon: 0.13074832000851827 st
eps: 568 lr: 2.5600000000000001e-06 evaluation reward: 8.78
2094
episode: 2094 score: 5.0 memory length: 539338 epsilon: 0.13010878000851422 ste
ps: 323 lr: 2.5600000000000001e-06 evaluation reward: 8.68
2095
episode: 2095 score: 9.0 memory length: 539771 epsilon: 0.1292514400085088 step
s: 433 lr: 2.5600000000000001e-06 evaluation reward: 8.7
2096
episode: 2096 score: 12.0 memory length: 540408 epsilon: 0.12799018000850082 st
eps: 637 lr: 2.5600000000000001e-06 evaluation reward: 8.76
2097
episode: 2097 score: 13.0 memory length: 540863 epsilon: 0.12708928000849512 st
eps: 455 lr: 2.5600000000000001e-06 evaluation reward: 8.81
2098
episode: 2098 score: 8.0 memory length: 541271 epsilon: 0.12628144000849 steps:
408 lr: 2.5600000000000001e-06 evaluation reward: 8.82
2099
episode: 2099 score: 9.0 memory length: 541709 epsilon: 0.12541420000848452 ste
ps: 438 lr: 2.5600000000000001e-06 evaluation reward: 8.83
2100
episode: 2100 score: 12.0 memory length: 542160 epsilon: 0.12452122000848223 st
eps: 451 lr: 2.5600000000000001e-06 evaluation reward: 8.88
2101
episode: 2101 score: 10.0 memory length: 542639 epsilon: 0.12357280000848288 st
eps: 479 lr: 2.5600000000000001e-06 evaluation reward: 8.93
2102
episode: 2102 score: 9.0 memory length: 543101 epsilon: 0.1226580400084835 step
s: 462 lr: 2.5600000000000001e-06 evaluation reward: 8.96
2103
episode: 2103 score: 13.0 memory length: 543691 epsilon: 0.1214898400084843 ste
ps: 590 lr: 2.5600000000000001e-06 evaluation reward: 9.01
2104
episode: 2104 score: 17.0 memory length: 544262 epsilon: 0.12035926000848507 st
eps: 571 lr: 2.5600000000000001e-06 evaluation reward: 9.09
2105
episode: 2105 score: 6.0 memory length: 544600 epsilon: 0.11969002000848553 ste
ps: 338 lr: 2.5600000000000001e-06 evaluation reward: 9.1
2106
episode: 2106 score: 10.0 memory length: 545078 epsilon: 0.11874358000848617 st
eps: 478 lr: 2.5600000000000001e-06 evaluation reward: 9.16
2107
episode: 2107 score: 15.0 memory length: 545489 epsilon: 0.11792980000848673 st
eps: 411 lr: 2.5600000000000001e-06 evaluation reward: 9.26
2108

episode: 2108 score: 15.0 memory length: 546169 epsilon: 0.11658340000848764 st
eps: 680 lr: 2.5600000000000001e-06 evaluation reward: 9.2
2109
episode: 2109 score: 11.0 memory length: 546643 epsilon: 0.11564488000848828 st
eps: 474 lr: 2.5600000000000001e-06 evaluation reward: 9.25
2110
episode: 2110 score: 12.0 memory length: 547198 epsilon: 0.11454598000848903 st
eps: 555 lr: 2.5600000000000001e-06 evaluation reward: 9.29
2111
episode: 2111 score: 9.0 memory length: 547625 epsilon: 0.11370052000848961 ste
ps: 427 lr: 2.5600000000000001e-06 evaluation reward: 9.29
2112
episode: 2112 score: 9.0 memory length: 548137 epsilon: 0.1126867600084903 step
s: 512 lr: 2.5600000000000001e-06 evaluation reward: 9.33
2113
episode: 2113 score: 7.0 memory length: 548521 epsilon: 0.11192644000849082 ste
ps: 384 lr: 2.5600000000000001e-06 evaluation reward: 9.34
2114
episode: 2114 score: 7.0 memory length: 548864 epsilon: 0.11124730000849128 ste
ps: 343 lr: 2.5600000000000001e-06 evaluation reward: 9.35
2115
episode: 2115 score: 15.0 memory length: 549422 epsilon: 0.11014246000849204 st
eps: 558 lr: 2.5600000000000001e-06 evaluation reward: 9.42
2116
episode: 2116 score: 8.0 memory length: 549845 epsilon: 0.10930492000849261 ste
ps: 423 lr: 2.5600000000000001e-06 evaluation reward: 9.4
2117
episode: 2117 score: 10.0 memory length: 550325 epsilon: 0.10835452000849326 st
eps: 480 lr: 2.5600000000000001e-06 evaluation reward: 9.39
2118
episode: 2118 score: 7.0 memory length: 550745 epsilon: 0.10752292000849382 ste
ps: 420 lr: 2.5600000000000001e-06 evaluation reward: 9.34
2119
episode: 2119 score: 10.0 memory length: 551270 epsilon: 0.10648342000849453 st
eps: 525 lr: 2.5600000000000001e-06 evaluation reward: 9.38
2120
episode: 2120 score: 8.0 memory length: 551667 epsilon: 0.10569736000849507 ste
ps: 397 lr: 2.5600000000000001e-06 evaluation reward: 9.39
2121
episode: 2121 score: 4.0 memory length: 551930 epsilon: 0.10517662000849543 ste
ps: 263 lr: 2.5600000000000001e-06 evaluation reward: 9.33
2122
episode: 2122 score: 8.0 memory length: 552355 epsilon: 0.104335120008496 step
s: 425 lr: 2.5600000000000001e-06 evaluation reward: 9.3
2123
episode: 2123 score: 8.0 memory length: 552776 epsilon: 0.10350154000849657 ste
ps: 421 lr: 2.5600000000000001e-06 evaluation reward: 9.33
2124
episode: 2124 score: 10.0 memory length: 553237 epsilon: 0.10258876000849719 st
eps: 461 lr: 2.5600000000000001e-06 evaluation reward: 9.38
2125
episode: 2125 score: 16.0 memory length: 553787 epsilon: 0.10149976000849793 st
eps: 550 lr: 2.5600000000000001e-06 evaluation reward: 9.49
2126
episode: 2126 score: 8.0 memory length: 554244 epsilon: 0.10059490000849855 ste
ps: 457 lr: 2.5600000000000001e-06 evaluation reward: 9.47
2127
episode: 2127 score: 10.0 memory length: 554740 epsilon: 0.09961282000849922 st
eps: 496 lr: 2.5600000000000001e-06 evaluation reward: 9.45
2128
episode: 2128 score: 12.0 memory length: 555179 epsilon: 0.09874360000849981 st
eps: 439 lr: 2.5600000000000001e-06 evaluation reward: 9.48
2129
episode: 2129 score: 7.0 memory length: 555537 epsilon: 0.0980347600085003 step
s: 358 lr: 2.5600000000000001e-06 evaluation reward: 9.52
2130

```
episode: 2130    score: 8.0    memory length: 555942    epsilon: 0.09723286000850084    ste
ps: 405    lr: 2.5600000000000001e-06    evaluation reward: 9.51
2131
episode: 2131    score: 8.0    memory length: 556362    epsilon: 0.09640126000850141    ste
ps: 420    lr: 2.5600000000000001e-06    evaluation reward: 9.47
2132
episode: 2132    score: 17.0    memory length: 556996    epsilon: 0.09514594000850227    st
eps: 634    lr: 2.5600000000000001e-06    evaluation reward: 9.58
2133
episode: 2133    score: 11.0    memory length: 557505    epsilon: 0.09413812000850295    st
eps: 509    lr: 2.5600000000000001e-06    evaluation reward: 9.59
2134
episode: 2134    score: 12.0    memory length: 558103    epsilon: 0.09295408000850376    st
eps: 598    lr: 2.5600000000000001e-06    evaluation reward: 9.66
2135
episode: 2135    score: 6.0    memory length: 558425    epsilon: 0.0923165200085042    step
s: 322    lr: 2.5600000000000001e-06    evaluation reward: 9.62
2136
episode: 2136    score: 11.0    memory length: 558939    epsilon: 0.09129880000850489    st
eps: 514    lr: 2.5600000000000001e-06    evaluation reward: 9.6
2137
episode: 2137    score: 5.0    memory length: 559262    epsilon: 0.09065926000850533    ste
ps: 323    lr: 2.5600000000000001e-06    evaluation reward: 9.51
2138
episode: 2138    score: 12.0    memory length: 559771    epsilon: 0.08965144000850601    st
eps: 509    lr: 2.5600000000000001e-06    evaluation reward: 9.57
2139
episode: 2139    score: 6.0    memory length: 560110    epsilon: 0.08898022000850647    ste
ps: 339    lr: 2.5600000000000001e-06    evaluation reward: 9.45
2140
episode: 2140    score: 5.0    memory length: 560383    epsilon: 0.08843968000850684    ste
ps: 273    lr: 2.5600000000000001e-06    evaluation reward: 9.37
2141
episode: 2141    score: 7.0    memory length: 560765    epsilon: 0.08768332000850736    ste
ps: 382    lr: 2.5600000000000001e-06    evaluation reward: 9.36
2142
episode: 2142    score: 14.0    memory length: 561295    epsilon: 0.08663392000850807    st
eps: 530    lr: 2.5600000000000001e-06    evaluation reward: 9.45
2143
episode: 2143    score: 10.0    memory length: 561632    epsilon: 0.08596666000850853    st
eps: 337    lr: 2.5600000000000001e-06    evaluation reward: 9.5
2144
episode: 2144    score: 9.0    memory length: 562106    epsilon: 0.08502814000850917    ste
ps: 474    lr: 2.5600000000000001e-06    evaluation reward: 9.55
2145
episode: 2145    score: 9.0    memory length: 562536    epsilon: 0.08417674000850975    ste
ps: 430    lr: 2.5600000000000001e-06    evaluation reward: 9.56
2146
episode: 2146    score: 7.0    memory length: 562920    epsilon: 0.08341642000851027    ste
ps: 384    lr: 2.5600000000000001e-06    evaluation reward: 9.52
2147
episode: 2147    score: 9.0    memory length: 563390    epsilon: 0.0824858200085109    step
s: 470    lr: 2.5600000000000001e-06    evaluation reward: 9.49
2148
episode: 2148    score: 10.0    memory length: 563849    epsilon: 0.08157700000851152    st
eps: 459    lr: 2.5600000000000001e-06    evaluation reward: 9.53
2149
episode: 2149    score: 7.0    memory length: 564238    epsilon: 0.08080678000851205    ste
ps: 389    lr: 2.5600000000000001e-06    evaluation reward: 9.54
2150
episode: 2150    score: 4.0    memory length: 564481    epsilon: 0.08032564000851238    ste
ps: 243    lr: 2.5600000000000001e-06    evaluation reward: 9.49
2151
```

KeyboardInterrupt

Traceback (most recent call last)

```

Cell In[9], line 47
    45 # Start training after random sample generation
    46 if(frame >= train_frame): # You can set train_frame to a lower value while testing
    your starts training earlier
--> 47     agent.train_policy_net(frame)
    48     # Update the target network only for Double DQN only
    49     if double_dqn and (frame % update_target_network_frequency)== 0:

```

File /kaggle/input/mp5dataset-1/assignment5_materials/agent.py:62, in Agent.train_policy_net(self, frame)

```

    59 if self.epsilon > self.epsilon_min:
    60     self.epsilon -= self.epsilon_decay
--> 62 mini_batch = self.memory.sample_mini_batch(frame)
    63 mini_batch = np.array(mini_batch, dtype=object).transpose()
    65 history = np.stack(mini_batch[0], axis=0)

```

File /kaggle/input/mp5dataset-1/assignment5_materials/memory.py:30, in ReplayMemory.sample_mini_batch(self, frame)

```

    27 for j in range(HISTORY_SIZE + 1):
    28     sample.append(self.memory[i + j])
--> 30 sample = np.array(sample, dtype=object)
    31 mini_batch.append((np.stack(sample[:, 0], axis=0), sample[3, 1], sample[3, 2],
    sample[3, 3]))
    33 return mini_batch

```

KeyboardInterrupt:

Error in callback <function flush_figures at 0x7f5a1c276440> (for post_execute), with arguments args (),kwargs {}:

```

-----
KeyboardInterrupt                                Traceback (most recent call last)
File /opt/conda/lib/python3.10/site-packages/matplotlib_inline/backend_inline.py:126, in
flush_figures()

```

```

    123 if InlineBackend.instance().close_figures:
    124     # ignore the tracking, just draw and close all figures
    125     try:
--> 126         return show(True)
    127     except Exception as e:
    128         # safely show traceback if in IPython, else raise
    129         ip = get_ipython()

```

File /opt/conda/lib/python3.10/site-packages/matplotlib_inline/backend_inline.py:90, in show(close, block)

```

    88 try:
    89     for figure_manager in Gcf.get_all_fig_managers():
--> 90         display(
    91             figure_manager.canvas.figure,
    92             metadata=_fetch_figure_metadata(figure_manager.canvas.figure)
    93         )
    94 finally:
    95     show._to_draw = []

```

File /opt/conda/lib/python3.10/site-packages/IPython/core/display_functions.py:298, in display(include, exclude, metadata, transient, display_id, raw, clear, *objs, **kwargs)

```

    296     publish_display_data(data=obj, metadata=metadata, **kwargs)
    297 else:
--> 298     format_dict, md_dict = format(obj, include=include, exclude=exclude)
    299     if not format_dict:
    300         # nothing to display (e.g. _ipython_display_ took over)
    301         continue

```

File /opt/conda/lib/python3.10/site-packages/IPython/core/formatters.py:179, in DisplayFormatter.format(self, obj, include, exclude)

```

    177 md = None
    178 try:
--> 179     data = formatter(obj)

```

```

180 except:
181     # FIXME: log the exception
182     raise

```

File /opt/conda/lib/python3.10/site-packages/decorator.py:232, in decorate.<locals>.fun(*args, **kw)

```

230 if not kwsyntax:
231     args, kw = fix(args, kw, sig)
--> 232 return caller(func, *(extras + args), **kw)

```

File /opt/conda/lib/python3.10/site-packages/IPython/core/formatters.py:223, in catch_format_error(method, self, *args, **kwargs)

```

221 """show traceback on failed format call"""
222 try:
--> 223     r = method(self, *args, **kwargs)
224 except NotImplementedError:
225     # don't warn on NotImplementedError
226     return self._check_return(None, args[0])

```

File /opt/conda/lib/python3.10/site-packages/IPython/core/formatters.py:340, in BaseFormatter.__call__(self, obj)

```

338     pass
339 else:
--> 340     return printer(obj)
341 # Finally look for special method names
342 method = get_real_method(obj, self.print_method)

```

File /opt/conda/lib/python3.10/site-packages/IPython/core/pylabtools.py:152, in print_figure(fig, fmt, bbox_inches, base64, **kwargs)

```

149     from matplotlib.backend_bases import FigureCanvasBase
150     FigureCanvasBase(fig)
--> 152 fig.canvas.print_figure(bytes_io, **kw)
153 data = bytes_io.getvalue()
154 if fmt == 'svg':

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/backend_bases.py:2366, in FigureCanvasBase.print_figure(self, filename, dpi, facecolor, edgecolor, orientation, format, bbox_inches, pad_inches, bbox_extra_artists, backend, **kwargs)

```

2362 try:
2363     # _get_renderer may change the figure dpi (as vector formats
2364     # force the figure dpi to 72), so we need to set it again here.
2365     with cbook._setattr_cm(self.figure, dpi=dpi):
--> 2366         result = print_method(
2367             filename,
2368             facecolor=facecolor,
2369             edgecolor=edgecolor,
2370             orientation=orientation,
2371             bbox_inches_restore=_bbox_inches_restore,
2372             **kwargs)
2373 finally:
2374     if bbox_inches and restore_bbox:

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/backend_bases.py:2232, in FigureCanvasBase._switch_canvas_and_return_print_method.<locals>.<lambda>(*args, **kwargs)

```

2228     optional_kws = { # Passed by print_figure for other renderers.
2229         "dpi", "facecolor", "edgecolor", "orientation",
2230         "bbox_inches_restore"}
2231     skip = optional_kws - {*inspect.signature(meth).parameters}
--> 2232     print_method = functools.wraps(meth)(lambda *args, **kwargs: meth(
2233         *args, **{k: v for k, v in kwargs.items() if k not in skip}))
2234 else: # Let third-parties do as they see fit.
2235     print_method = meth

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/backends/backend_agg.py:509, in FigureCanvasAgg.print_png(self, filename_or_obj, metadata, pil_kwargs)

```

462 def print_png(self, filename_or_obj, *, metadata=None, pil_kwargs=None):

```

```

463         """
464         Write the figure to a PNG file.
465
466     (...)
467     """
468     *metadata*, including the default 'Software' key.
469     """
--> 509     self._print_pil(filename_or_obj, "png", pil_kwargs, metadata)

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/backends/backend_agg.py:457, in FigureCanvasAgg._print_pil(self, filename_or_obj, fmt, pil_kwargs, metadata)

```

452 def _print_pil(self, filename_or_obj, fmt, pil_kwargs, metadata=None):
453     """
454     Draw the canvas, then save it using `.image.imsave` (to which
455     *pil_kwargs* and *metadata* are forwarded).
456     """
--> 457     FigureCanvasAgg.draw(self)
458     mpl.image.imsave(
459         filename_or_obj, self.buffer_rgba(), format=fmt, origin="upper",
460         dpi=self.figure.dpi, metadata=metadata, pil_kwargs=pil_kwargs)

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/backends/backend_agg.py:400, in FigureCanvasAgg.draw(self)

```

396 # Acquire a lock on the shared font cache.
397 with RendererAgg.lock, \
398     (self.toolbar._wait_cursor_for_draw_cm() if self.toolbar
399     else nullcontext()):
--> 400     self.figure.draw(self.renderer)
401     # A GUI class may be need to update a window using this draw, so
402     # don't forget to call the superclass.
403     super().draw()

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/artist.py:95, in _finalize_rasterization.<locals>.draw_wrapper(artist, renderer, *args, **kwargs)

```

93 @wraps(draw)
94 def draw_wrapper(artist, renderer, *args, **kwargs):
---> 95     result = draw(artist, renderer, *args, **kwargs)
96     if renderer._rasterizing:
97         renderer.stop_rasterizing()

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/artist.py:72, in allow_rasterization.<locals>.draw_wrapper(artist, renderer)

```

69 if artist.get_agg_filter() is not None:
70     renderer.start_filter()
---> 72     return draw(artist, renderer)
73 finally:
74     if artist.get_agg_filter() is not None:

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/figure.py:3175, in Figure.draw(self, renderer)

```

3172 # ValueError can occur when resizing a window.
3173 self.patch.draw(renderer)
-> 3175 mimage._draw_list_compositing_images(
3176     renderer, self, artists, self.suppressComposite)
3177 for sfig in self.subfigs:
3178     sfig.draw(renderer)

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/image.py:131, in _draw_list_compositing_images(renderer, parent, artists, suppress_composite)

```

129 if not_composite or not has_images:
130     for a in artists:
--> 131         a.draw(renderer)
132 else:
133     # Composite any adjacent images together
134     image_group = []

```

File /opt/conda/lib/python3.10/site-packages/matplotlib/artist.py:72, in allow_rasterization


```

tion.<locals>.draw_wrapper(artist, renderer)
    69     if artist.get_agg_filter() is not None:
    70         renderer.start_filter()
---> 72     return draw(artist, renderer)
    73 finally:
    74     if artist.get_agg_filter() is not None:

File /opt/conda/lib/python3.10/site-packages/matplotlib/axes/_base.py:3064, in _AxesBase.draw(self, renderer)
    3061 if artists_rasterized:
    3062     _draw_rasterized(self.figure, artists_rasterized, renderer)
-> 3064 mimage._draw_list_compositing_images(
    3065     renderer, self, artists, self.figure.suppressComposite)
    3067 renderer.close_group('axes')
    3068 self.stale = False

File /opt/conda/lib/python3.10/site-packages/matplotlib/image.py:131, in _draw_list_compositing_images(renderer, parent, artists, suppress_composite)
    129 if not_composite or not has_images:
    130     for a in artists:
--> 131         a.draw(renderer)
    132 else:
    133     # Composite any adjacent images together
    134     image_group = []

File /opt/conda/lib/python3.10/site-packages/matplotlib/artist.py:72, in allow_rasterization.<locals>.draw_wrapper(artist, renderer)
    69     if artist.get_agg_filter() is not None:
    70         renderer.start_filter()
---> 72     return draw(artist, renderer)
    73 finally:
    74     if artist.get_agg_filter() is not None:

File /opt/conda/lib/python3.10/site-packages/matplotlib/lines.py:797, in Line2D.draw(self, renderer)
    794         gc.set_foreground(lc_rgba, isRGBA=True)
    796         gc.set_dashes(*self._dash_pattern)
--> 797         renderer.draw_path(gc, tpath, affine.frozen())
    798         gc.restore()
    800 if self._marker and self._markersize > 0:

File /opt/conda/lib/python3.10/site-packages/matplotlib/backends/backend_agg.py:146, in RendererAgg.draw_path(self, gc, path, transform, rgbFace)
    144 else:
    145     try:
--> 146         self._renderer.draw_path(gc, path, transform, rgbFace)
    147     except OverflowError:
    148         cant_chunk = ''

KeyboardInterrupt:

```

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.

```
In [11]: torch.save(agent.policy_net, "/kaggle/working/breakout_dqn_latest.pth")
```

```
In [13]: # from gym.wrappers import Monitor # If importing monitor raises issues, try using `from
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

```

In [14]: `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 gam

    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()

```

```

-----
NameError                                Traceback (most recent call last)
Cell In[14], line 9
      6 agent.epsilon = 0.0 # Set agent to only exploit the best action
      8 env = gym.make('BreakoutDeterministic-v4')
----> 9 env = wrap_env(env)
      11 done = False
      12 score = 0

Cell In[13], line 39, in wrap_env(env)
      38 def wrap_env(env):
----> 39     env = Monitor(env, './video', force=True)
      40     return env

NameError: name 'Monitor' is not defined

```

In [17]: !pip install moviepy

```

Collecting moviepy
  Downloading moviepy-1.0.3.tar.gz (388 kB)
    ooooooooooooooooooooooooooooooooooooooooooooooooooooo 388.3/388.3 kB 2.6 MB/s eta 0:00:0000:0100:01
  Preparing metadata (setup.py) ... done
Collecting decorator<5.0,>=4.0.2 (from moviepy)
  Downloading decorator-4.4.2-py2.py3-none-any.whl.metadata (4.2 kB)
Requirement already satisfied: tqdm<5.0,>=4.11.2 in /opt/conda/lib/python3.10/site-packa
ges (from moviepy) (4.66.1)
Requirement already satisfied: requests<3.0,>=2.8.1 in /opt/conda/lib/python3.10/site-pa
ckages (from moviepy) (2.31.0)
Collecting proglog<=1.0.0 (from moviepy)
  Downloading proglog-0.1.10-py3-none-any.whl.metadata (639 bytes)
Requirement already satisfied: numpy>=1.17.3 in /opt/conda/lib/python3.10/site-packages
 (from moviepy) (1.26.4)
Requirement already satisfied: imageio<3.0,>=2.5 in /opt/conda/lib/python3.10/site-packa
ges (from moviepy) (2.33.1)
Collecting imageio_ffmpeg>=0.2.0 (from moviepy)
  Downloading imageio_ffmpeg-0.4.9-py3-none-manylinux2010_x86_64.whl.metadata (1.7 kB)
Requirement already satisfied: pillow>=8.3.2 in /opt/conda/lib/python3.10/site-packages
 (from imageio<3.0,>=2.5->moviepy) (9.5.0)

```

```

Requirement already satisfied: setuptools in /root/.local/lib/python3.10/site-packages
(from imageio_ffmpeg>=0.2.0->moviepy) (69.5.1)
Requirement already satisfied: charset-normalizer<4,>=2 in /opt/conda/lib/python3.10/site-packages (from requests<3.0,>=2.8.1->moviepy) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.10/site-packages (from requests<3.0,>=2.8.1->moviepy) (3.6)
Requirement already satisfied: urllib3<3,>=1.21.1 in /opt/conda/lib/python3.10/site-packages (from requests<3.0,>=2.8.1->moviepy) (1.26.18)
Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.10/site-packages (from requests<3.0,>=2.8.1->moviepy) (2024.2.2)
Downloading decorator-4.4.2-py2.py3-none-any.whl (9.2 kB)
Downloading imageio_ffmpeg-0.4.9-py3-none-manylinux2010_x86_64.whl (26.9 MB)
  26.9/26.9 MB 50.1 MB/s eta 0:00:00:00:0100:01
Downloading proglog-0.1.10-py3-none-any.whl (6.1 kB)
Building wheels for collected packages: moviepy
  Building wheel for moviepy (setup.py) ... done
  Created wheel for moviepy: filename=moviepy-1.0.3-py3-none-any.whl size=110720 sha256=d79f96ba9e4ec5e57c1cd2476f0e33a25ab46bb4088cae66bc8dee83e9d3b465
  Stored in directory: /root/.cache/pip/wheels/96/32/2d/e10123bd88fbfc02fed53cc18c80a171d3c87479ed845fa7c1
Successfully built moviepy
Installing collected packages: proglog, imageio_ffmpeg, decorator, moviepy
  Attempting uninstall: decorator
    Found existing installation: decorator 5.1.1
    Uninstalling decorator-5.1.1:
      Successfully uninstalled decorator-5.1.1
Successfully installed decorator-4.4.2 imageio_ffmpeg-0.4.9 moviepy-1.0.3 proglog-0.1.10

```

In [18]:

```

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 torch.utils.tensorboard import SummaryWriter
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

from gym.wrappers import RecordVideo # If importing monitor raises issues, try using `fr
import glob
import io
import base64

from IPython.display import HTML
from IPython import display as ipythondisplay

from pyvirtualdisplay import Display

env = gym.make('BreakoutDeterministic-v4', render_mode='rgb_array')
state = env.reset()
number_lives = find_max_lives(env)
state_size = env.observation_space.shape
action_size = 3 #fire, left, and right

```

```

# 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 = RecordVideo(env, './video')
    return env

from agent import Agent
action_size = 3

display = Display(visible=0, size=(300, 200))
display.start()

# Load agent
agent = Agent(action_size)
agent.load_policy_net("/kaggle/working/breakout_dqn_latest.pth")
agent.epsilon = 0.0 # Set agent to only exploit the best action

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, HISTORY_SIZE)
frame = 0
while not done:
    # show_state(env, step) # uncommenting this provides another way to visualize the gam
    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['lives'])

life = info['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()

```

```

/opt/conda/lib/python3.10/site-packages/gym/utils/passive_env_checker.py:233: Deprecatio
nWarning: `np.bool8` is a deprecated alias for `np.bool_`. (Deprecated NumPy 1.24)
  if not isinstance(terminated, (bool, np.bool8)):
/opt/conda/lib/python3.10/site-packages/gym/wrappers/record_video.py:75: UserWarning: WA
RN: Overwriting existing videos at /kaggle/working/video folder (try specifying a differ
ent `video_folder` for the `RecordVideo` wrapper if this is not desired)
  logger.warn(
/opt/conda/lib/python3.10/site-packages/gym/utils/passive_env_checker.py:289: UserWarnin
g: WARN: No render fps was declared in the environment (env.metadata['render_fps'] is No
ne or not defined), rendering may occur at inconsistent fps.
  logger.warn(
Moviepy - Building video /kaggle/working/video/rl-video-episode-0.mp4.
Moviepy - Writing video /kaggle/working/video/rl-video-episode-0.mp4

```

```

-----
TypeError                                Traceback (most recent call last)

```

```

Cell In[18], line 103

```

```

    100 action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
    101 state = next_state
--> 103 next_state, reward, done, _, info = env.step(action + 1)
    105 frame_next_state = get_frame(next_state)
    106 history[4, :, :] = frame_next_state

```

```

File /opt/conda/lib/python3.10/site-packages/gym/wrappers/record_video.py:167, in Record
Video.step(self, action)

```

```

    165 if not self.is_vector_env:
    166     if terminateds or truncateds:
--> 167         self.close_video_recorder()
    168 elif terminateds[0] or truncateds[0]:
    169     self.close_video_recorder()

```

```

File /opt/conda/lib/python3.10/site-packages/gym/wrappers/record_video.py:180, in Record
Video.close_video_recorder(self)

```

```

    178 if self.recording:
    179     assert self.video_recorder is not None
--> 180     self.video_recorder.close()
    181 self.recording = False
    182 self.recorded_frames = 1

```

```

File /opt/conda/lib/python3.10/site-packages/gym/wrappers/monitoring/video_recorder.py:1
58, in VideoRecorder.close(self)

```

```

    156 logger.debug(f"Closing video encoder: path={self.path}")
    157 clip = ImageSequenceClip(self.recorded_frames, fps=self.frames_per_sec)
--> 158 clip.write_videofile(self.path)
    159 else:
    160     # No frames captured. Set metadata.

```

```

161         if self.metadata is None:

File /opt/conda/lib/python3.10/site-packages/decorator.py:232, in decorate.<locals>.fun
(*args, **kw)
    230 if not kwsyntax:
    231     args, kw = fix(args, kw, sig)
--> 232 return caller(func, *(extras + args), **kw)

File /opt/conda/lib/python3.10/site-packages/moviepy/decorators.py:54, in requires_durat
ion(f, clip, *a, **k)
    52     raise ValueError("Attribute 'duration' not set")
    53 else:
---> 54     return f(clip, *a, **k)

File /opt/conda/lib/python3.10/site-packages/decorator.py:232, in decorate.<locals>.fun
(*args, **kw)
    230 if not kwsyntax:
    231     args, kw = fix(args, kw, sig)
--> 232 return caller(func, *(extras + args), **kw)

File /opt/conda/lib/python3.10/site-packages/moviepy/decorators.py:135, in use_clip_fps_
by_default(f, clip, *a, **k)
    130 new_a = [fun(arg) if (name=='fps') else arg
    131             for (arg, name) in zip(a, names)]
    132 new_kw = {k: fun(v) if k=='fps' else v
    133             for (k,v) in k.items()}
--> 135 return f(clip, *new_a, **new_kw)

File /opt/conda/lib/python3.10/site-packages/decorator.py:232, in decorate.<locals>.fun
(*args, **kw)
    230 if not kwsyntax:
    231     args, kw = fix(args, kw, sig)
--> 232 return caller(func, *(extras + args), **kw)

File /opt/conda/lib/python3.10/site-packages/moviepy/decorators.py:22, in convert_masks_
to_RGB(f, clip, *a, **k)
    20 if clip.ismask:
    21     clip = clip.to_RGB()
---> 22 return f(clip, *a, **k)

File /opt/conda/lib/python3.10/site-packages/moviepy/video/VideoClip.py:300, in VideoCli
p.write_videofile(self, filename, fps, codec, bitrate, audio, audio_fps, preset, audio_n
bytes, audio_codec, audio_bitrate, audio_bufsize, temp_audiofile, rewrite_audio, remove_
temp, write_logfile, verbose, threads, ffmpeg_params, logger)
    292 if make_audio:
    293     self.audio.write_audiofile(audiofile, audio_fps,
    294                               audio_nbytes, audio_bufsize,
    295                               audio_codec, bitrate=audio_bitrate,
    296                               write_logfile=write_logfile,
    297                               verbose=verbose,
    298                               logger=logger)
--> 300 ffmpeg_write_video(self, filename, fps, codec,
    301                      bitrate=bitrate,
    302                      preset=preset,
    303                      write_logfile=write_logfile,
    304                      audiofile=audiofile,
    305                      verbose=verbose, threads=threads,
    306                      ffmpeg_params=ffmpeg_params,
    307                      logger=logger)
    309 if remove_temp and make_audio:
    310     if os.path.exists(audiofile):

File /opt/conda/lib/python3.10/site-packages/moviepy/video/io/ffmpeg_writer.py:213, in f
fmpeg_write_video(clip, filename, fps, codec, bitrate, preset, withmask, write_logfile,
audiofile, verbose, threads, ffmpeg_params, logger)
    211     logfile = None

```

```

212 logger(message='Moviepy - Writing video %s\n' % filename)
--> 213 with FFMPEG_VideoWriter(filename, clip.size, fps, codec = codec,
214                             preset=preset, bitrate=bitrate, logfile=logfile,
215                             audiofile=audiofile, threads=threads,
216                             ffmpeg_params=ffmpeg_params) as writer:
218     nframes = int(clip.duration*fps)
220     for t,frame in clip.iter_frames(logger=logger, with_times=True,
221                                     fps=fps, dtype="uint8"):

```

File /opt/conda/lib/python3.10/site-packages/moviepy/video/io/ffmpeg_writer.py:88, in FFMPEG_VideoWriter.__init__(self, filename, size, fps, codec, audiofile, preset, bitrate, withmask, logfile, threads, ffmpeg_params)

```

77 self.ext = self.filename.split(".")[1]
79 # order is important
80 cmd = [
81     get_setting("FFMPEG_BINARY"),
82     '-y',
83     '-loglevel', 'error' if logfile == sp.PIPE else 'info',
84     '-f', 'rawvideo',
85     '-vcodec', 'rawvideo',
86     '-s', '%dx%d' % (size[0], size[1]),
87     '-pix_fmt', 'rgba' if withmask else 'rgb24',
--> 88     '-r', '%.02f' % fps,
89     '-an', '-i', '-'
90 ]
91 if audiofile is not None:
92     cmd.extend([
93         '-i', audiofile,
94         '-acodec', 'copy'
95     ])

```

TypeError: must be real number, not NoneType

In []: