

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 [32]: !pip3 install gym gym==0.15.3  
!sudo apt-get install -y xvfb python-opengl ffmpeg
```

```
Requirement already satisfied: gym in /opt/conda/lib/python3.7/site-packages (0.23.1)
Collecting gym
  Downloading gym-0.15.3.tar.gz (1.6 MB)
  ━━━━━━━━━━━━━━━━ 1.6/1.6 MB 14.5 MB/s eta 0:00:000:010
0:01
  Preparing metadata (setup.py) ... done
Requirement already satisfied: scipy in /opt/conda/lib/python3.7/site-packages (from gym) (1.7.3)
Requirement already satisfied: numpy>=1.10.4 in /opt/conda/lib/python3.7/site-packages (from gym) (1.19.5)
Requirement already satisfied: six in /opt/conda/lib/python3.7/site-packages (from gym) (1.16.0)
Collecting pyglet<=1.3.2,>=1.2.0
  Downloading pyglet-1.3.2-py2.py3-none-any.whl (1.0 MB)
  ━━━━━━━━━━━━━━ 1.0/1.0 MB 53.5 MB/s eta 0:00:00
Collecting cloudpickle~=1.2.0
  Downloading cloudpickle-1.2.2-py2.py3-none-any.whl (25 kB)
Collecting future
  Downloading future-0.18.2.tar.gz (829 kB)
  ━━━━━━━━━━━━━━ 829.2/829.2 KB 64.3 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Building wheels for collected packages: gym, future
  Building wheel for gym (setup.py) ... done
    Created wheel for gym: filename=gym-0.15.3-py3-none-any.whl size=1644970 sha256=49c2244485d5a0fc89beafc3f6cbd71cb670f0ae466a171ebf1f43a5f3b16ea3
    Stored in directory: /home/zacharytgeorge/.cache/pip/wheels/55/16/6b/2250ca4f9f050a4d27d8bed287e57bbb3c33fc4066f557cc75
  Building wheel for future (setup.py) ... done
    Created wheel for future: filename=future-0.18.2-py3-none-any.whl size=491070 sha256=8e3eb12581370abe9fc33fc5819f4f49ae5ccda5abad1a010c19d7441a70d3
    Stored in directory: /home/zacharytgeorge/.cache/pip/wheels/56/b0/fe/4410d17b32f1fc3cf54cdfb2bc04d7b4b8f4ae377e2229ba0
Successfully built gym future
Installing collected packages: cloudpickle, future, pyglet, gym
Attempting uninstall: cloudpickle
  Found existing installation: cloudpickle 2.0.0
  Uninstalling cloudpickle-2.0.0:
    Successfully uninstalled cloudpickle-2.0.0
Attempting uninstall: gym
  Found existing installation: gym 0.23.1
  Uninstalling gym-0.23.1:
    Successfully uninstalled gym-0.23.1
Successfully installed cloudpickle-1.2.2 future-0.18.2 gym-0.15.3 pyglet-1.3.2
Reading package lists... Done
Building dependency tree
Reading state information... Done
python-opengl is already the newest version (3.1.0+dfsg-2).
xvfb is already the newest version (2:1.20.4-1+deb10u4).
ffmpeg is already the newest version (7:4.1.9-0+deb10u1).
0 upgraded, 0 newly installed, 0 to remove and 5 not upgraded.
```

```
In [2]: !pip3 install --upgrade setuptools --user
!pip3 install ez_setup
!pip3 install gym[atari]
!pip3 install gym[accept-rom-license]
```

```
Requirement already satisfied: setuptools in /home/zacharytgeorge/.local/lib/python3.7/site-packages (62.1.0)
Requirement already satisfied: ez_setup in /opt/conda/lib/python3.7/site-packages (0.9)
Requirement already satisfied: gym[atari] in /opt/conda/lib/python3.7/site-packages (0.23.1)
Requirement already satisfied: gym-notices>=0.0.4 in /opt/conda/lib/python3.7/site-packages (from gym[atari]) (0.0.6)
Requirement already satisfied: importlib-metadata>=4.10.0 in /opt/conda/lib/python3.7/site-packages (from gym[atari]) (4.11.3)
Requirement already satisfied: numpy>=1.18.0 in /opt/conda/lib/python3.7/site-packages (from gym[atari]) (1.19.5)
Requirement already satisfied:云pickle>=1.2.0 in /opt/conda/lib/python3.7/site-packages (from gym[atari]) (2.0.0)
Requirement already satisfied: ale-py~0.7.4 in /opt/conda/lib/python3.7/site-packages (from gym[atari]) (0.7.5)
Requirement already satisfied: importlib-resources in /opt/conda/lib/python3.7/site-packages (from ale-py~0.7.4->gym[atari]) (5.4.0)
Requirement already satisfied: typing-extensions>=3.6.4 in /opt/conda/lib/python3.7/site-packages (from importlib-metadata>=4.10.0->gym[atari]) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /opt/conda/lib/python3.7/site-packages (from importlib-metadata>=4.10.0->gym[atari]) (3.7.0)
Requirement already satisfied: gym[accept-rom-license] in /opt/conda/lib/python3.7/site-packages (0.23.1)
Requirement already satisfied: numpy>=1.18.0 in /opt/conda/lib/python3.7/site-packages (from gym[accept-rom-license]) (1.19.5)
Requirement already satisfied:云pickle>=1.2.0 in /opt/conda/lib/python3.7/site-packages (from gym[accept-rom-license]) (2.0.0)
Requirement already satisfied: importlib-metadata>=4.10.0 in /opt/conda/lib/python3.7/site-packages (from gym[accept-rom-license]) (4.11.3)
Requirement already satisfied: gym-notices>=0.0.4 in /opt/conda/lib/python3.7/site-packages (from gym[accept-rom-license]) (0.0.6)
Requirement already satisfied: autorom[accept-rom-license]~0.4.2 in /opt/conda/lib/python3.7/site-packages (from gym[accept-rom-license]) (0.4.2)
Requirement already satisfied: importlib-resources in /opt/conda/lib/python3.7/site-packages (from autorom[accept-rom-license]~0.4.2->gym[accept-rom-license]) (5.4.0)
Requirement already satisfied: requests in /opt/conda/lib/python3.7/site-packages (from autorom[accept-rom-license]~0.4.2->gym[accept-rom-license]) (2.27.1)
Requirement already satisfied: tqdm in /opt/conda/lib/python3.7/site-packages (from a
utorom[accept-rom-license]~0.4.2->gym[accept-rom-license]) (4.63.0)
Requirement already satisfied: click in /opt/conda/lib/python3.7/site-packages (from autorom[accept-rom-license]~0.4.2->gym[accept-rom-license]) (8.0.4)
Requirement already satisfied: AutorOM.accept-rom-license in /opt/conda/lib/python3.7/site-packages (from autorom[accept-rom-license]~0.4.2->gym[accept-rom-license]) (0.4.2)
Requirement already satisfied: typing-extensions>=3.6.4 in /opt/conda/lib/python3.7/site-packages (from importlib-metadata>=4.10.0->gym[accept-rom-license]) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /opt/conda/lib/python3.7/site-packages (from importlib-metadata>=4.10.0->gym[accept-rom-license]) (3.7.0)
Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.7/site-packages (from requests->autorom[accept-rom-license]~0.4.2->gym[accept-rom-license]) (2021.10.8)
Requirement already satisfied: charset-normalizer~2.0.0 in /opt/conda/lib/python3.7/site-packages (from requests->autorom[accept-rom-license]~0.4.2->gym[accept-rom-license]) (2.0.12)
Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.7/site-packages (from requests->autorom[accept-rom-license]~0.4.2->gym[accept-rom-license]) (3.3)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /opt/conda/lib/python3.7/site-packages (from requests->autorom[accept-rom-license]~0.4.2->gym[accept-rom-license]) (1.26.8)
```

In [3]: `pip install scikit-image`

```
Requirement already satisfied: scikit-image in /opt/conda/lib/python3.7/site-packages (0.19.2)
Requirement already satisfied: imageio>=2.4.1 in /opt/conda/lib/python3.7/site-packages (from scikit-image) (2.19.0)
Requirement already satisfied: packaging>=20.0 in /opt/conda/lib/python3.7/site-packages (from scikit-image) (21.3)
Requirement already satisfied: pillow!=7.1.0,!=7.1.1,!=8.3.0,>=6.1.0 in /opt/conda/lib/python3.7/site-packages (from scikit-image) (9.0.1)
Requirement already satisfied: PyWavelets>=1.1.1 in /opt/conda/lib/python3.7/site-packages (from scikit-image) (1.3.0)
Requirement already satisfied: numpy>=1.17.0 in /opt/conda/lib/python3.7/site-packages (from scikit-image) (1.19.5)
Requirement already satisfied: networkx>=2.2 in /opt/conda/lib/python3.7/site-packages (from scikit-image) (2.7.1)
Requirement already satisfied: scipy>=1.4.1 in /opt/conda/lib/python3.7/site-packages (from scikit-image) (1.7.3)
Requirement already satisfied: tifffile>=2019.7.26 in /opt/conda/lib/python3.7/site-packages (from scikit-image) (2021.11.2)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /opt/conda/lib/python3.7/site-packages (from packaging>=20.0->scikit-image) (3.0.7)
Note: you may need to restart the kernel to use updated packages.
```

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 [1]: `%matplotlib inline`

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

import matplotlib.pyplot as plt
# %Load_ext autoreload
# %autoreload 2
```

Understanding the environment

In the following cell, we initialize our game of **Breakout** and you can see how the environment looks like. For further documentation of the 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.

```
In [2]: env = gym.make('BreakoutDeterministic-v4')
state = env.reset()
```

```
A.L.E: Arcade Learning Environment (version 0.7.5+db37282)
[Powered by Stella]
```

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

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

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

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

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

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

```

next_state = state
life = number_lives

get_init_state(history, state, HISTORY_SIZE)

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

    # Perform a fire action if ball is no longer on screen to continue onto next level
    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 = reward

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

    if done:
        evaluation_reward.append(score)
        rewards.append(np.mean(evaluation_reward))
        episodes.append(e)
        pylab.plot(episodes, rewards, 'b')
        pylab.xlabel('Episodes')
        pylab.ylabel('Rewards')
        pylab.title('Episodes vs Reward')
        pylab.savefig("./save_graph/breakout_dqn.png") # save graph for training visualization

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

# if the mean of scores of last 100 episode is bigger than 5 save model
### Change this save condition to whatever you prefer ####
if np.mean(evaluation_reward) > 5 and np.mean(evaluation_reward) > best_eval_reward:
    torch.save(agent.policy_net, "./save_model/breakout_dqn.pth")
    best_eval_reward = np.mean(evaluation_reward)

```

episode: 0	score: 4.0	memory length: 278	epsilon: 1.0	steps: 278	lr: 0.00
001	evaluation reward: 4.0				
episode: 1	score: 0.0	memory length: 402	epsilon: 1.0	steps: 124	lr: 0.00
001	evaluation reward: 2.0				
episode: 2	score: 2.0	memory length: 605	epsilon: 1.0	steps: 203	lr: 0.00
001	evaluation reward: 2.0				
episode: 3	score: 1.0	memory length: 757	epsilon: 1.0	steps: 152	lr: 0.00
001	evaluation reward: 1.75				
episode: 4	score: 0.0	memory length: 881	epsilon: 1.0	steps: 124	lr: 0.00
001	evaluation reward: 1.4				
episode: 5	score: 3.0	memory length: 1133	epsilon: 1.0	steps: 252	lr: 0.0
001	evaluation reward: 1.6666666666666667				
episode: 6	score: 0.0	memory length: 1257	epsilon: 1.0	steps: 124	lr: 0.0
001	evaluation reward: 1.4285714285714286				
episode: 7	score: 0.0	memory length: 1380	epsilon: 1.0	steps: 123	lr: 0.0
001	evaluation reward: 1.25				
episode: 8	score: 0.0	memory length: 1504	epsilon: 1.0	steps: 124	lr: 0.0
001	evaluation reward: 1.1111111111111112				
episode: 9	score: 0.0	memory length: 1628	epsilon: 1.0	steps: 124	lr: 0.0
001	evaluation reward: 1.0				
episode: 10	score: 0.0	memory length: 1752	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 0.9090909090909091				
episode: 11	score: 3.0	memory length: 2018	epsilon: 1.0	steps: 266	lr: 0.
0001	evaluation reward: 1.0833333333333333				
episode: 12	score: 1.0	memory length: 2169	epsilon: 1.0	steps: 151	lr: 0.
0001	evaluation reward: 1.0769230769230769				
episode: 13	score: 0.0	memory length: 2292	epsilon: 1.0	steps: 123	lr: 0.
0001	evaluation reward: 1.0				
episode: 14	score: 2.0	memory length: 2492	epsilon: 1.0	steps: 200	lr: 0.
0001	evaluation reward: 1.0666666666666667				
episode: 15	score: 0.0	memory length: 2616	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.0				
episode: 16	score: 3.0	memory length: 2864	epsilon: 1.0	steps: 248	lr: 0.
0001	evaluation reward: 1.1176470588235294				
episode: 17	score: 3.0	memory length: 3090	epsilon: 1.0	steps: 226	lr: 0.
0001	evaluation reward: 1.2222222222222223				
episode: 18	score: 2.0	memory length: 3307	epsilon: 1.0	steps: 217	lr: 0.
0001	evaluation reward: 1.263157894736842				
episode: 19	score: 2.0	memory length: 3506	epsilon: 1.0	steps: 199	lr: 0.
0001	evaluation reward: 1.3				
episode: 20	score: 1.0	memory length: 3675	epsilon: 1.0	steps: 169	lr: 0.
0001	evaluation reward: 1.2857142857142858				
episode: 21	score: 3.0	memory length: 3905	epsilon: 1.0	steps: 230	lr: 0.
0001	evaluation reward: 1.3636363636363635				
episode: 22	score: 1.0	memory length: 4077	epsilon: 1.0	steps: 172	lr: 0.
0001	evaluation reward: 1.3478260869565217				
episode: 23	score: 0.0	memory length: 4201	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.2916666666666667				
episode: 24	score: 1.0	memory length: 4371	epsilon: 1.0	steps: 170	lr: 0.
0001	evaluation reward: 1.28				
episode: 25	score: 0.0	memory length: 4495	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.2307692307692308				
episode: 26	score: 2.0	memory length: 4679	epsilon: 1.0	steps: 184	lr: 0.
0001	evaluation reward: 1.2592592592592593				
episode: 27	score: 0.0	memory length: 4803	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.2142857142857142				
episode: 28	score: 1.0	memory length: 4973	epsilon: 1.0	steps: 170	lr: 0.
0001	evaluation reward: 1.206896551724138				
episode: 29	score: 2.0	memory length: 5189	epsilon: 1.0	steps: 216	lr: 0.
0001	evaluation reward: 1.2333333333333334				

episode: 30	score: 2.0	memory length: 5408	epsilon: 1.0	steps: 219	lr: 0.
0001	evaluation reward: 1.2580645161290323				
episode: 31	score: 2.0	memory length: 5610	epsilon: 1.0	steps: 202	lr: 0.
0001	evaluation reward: 1.28125				
episode: 32	score: 2.0	memory length: 5831	epsilon: 1.0	steps: 221	lr: 0.
0001	evaluation reward: 1.303030303030303				
episode: 33	score: 0.0	memory length: 5955	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.2647058823529411				
episode: 34	score: 2.0	memory length: 6173	epsilon: 1.0	steps: 218	lr: 0.
0001	evaluation reward: 1.2857142857142858				
episode: 35	score: 0.0	memory length: 6296	epsilon: 1.0	steps: 123	lr: 0.
0001	evaluation reward: 1.25				
episode: 36	score: 0.0	memory length: 6420	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.2162162162162162				
episode: 37	score: 1.0	memory length: 6571	epsilon: 1.0	steps: 151	lr: 0.
0001	evaluation reward: 1.2105263157894737				
episode: 38	score: 0.0	memory length: 6695	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.1794871794871795				
episode: 39	score: 2.0	memory length: 6914	epsilon: 1.0	steps: 219	lr: 0.
0001	evaluation reward: 1.2				
episode: 40	score: 2.0	memory length: 7113	epsilon: 1.0	steps: 199	lr: 0.
0001	evaluation reward: 1.2195121951219512				
episode: 41	score: 3.0	memory length: 7329	epsilon: 1.0	steps: 216	lr: 0.
0001	evaluation reward: 1.2619047619047619				
episode: 42	score: 2.0	memory length: 7530	epsilon: 1.0	steps: 201	lr: 0.
0001	evaluation reward: 1.2790697674418605				
episode: 43	score: 2.0	memory length: 7731	epsilon: 1.0	steps: 201	lr: 0.
0001	evaluation reward: 1.2954545454545454				
episode: 44	score: 3.0	memory length: 7958	epsilon: 1.0	steps: 227	lr: 0.
0001	evaluation reward: 1.3333333333333333				
episode: 45	score: 2.0	memory length: 8178	epsilon: 1.0	steps: 220	lr: 0.
0001	evaluation reward: 1.3478260869565217				
episode: 46	score: 2.0	memory length: 8359	epsilon: 1.0	steps: 181	lr: 0.
0001	evaluation reward: 1.3617021276595744				
episode: 47	score: 5.0	memory length: 8698	epsilon: 1.0	steps: 339	lr: 0.
0001	evaluation reward: 1.4375				
episode: 48	score: 0.0	memory length: 8822	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.4081632653061225				
episode: 49	score: 0.0	memory length: 8946	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.38				
episode: 50	score: 1.0	memory length: 9097	epsilon: 1.0	steps: 151	lr: 0.
0001	evaluation reward: 1.3725490196078431				
episode: 51	score: 1.0	memory length: 9267	epsilon: 1.0	steps: 170	lr: 0.
0001	evaluation reward: 1.3653846153846154				
episode: 52	score: 1.0	memory length: 9437	epsilon: 1.0	steps: 170	lr: 0.
0001	evaluation reward: 1.3584905660377358				
episode: 53	score: 0.0	memory length: 9561	epsilon: 1.0	steps: 124	lr: 0.
0001	evaluation reward: 1.3333333333333333				
episode: 54	score: 1.0	memory length: 9731	epsilon: 1.0	steps: 170	lr: 0.
0001	evaluation reward: 1.3272727272727274				
episode: 55	score: 2.0	memory length: 9931	epsilon: 1.0	steps: 200	lr: 0.
0001	evaluation reward: 1.3392857142857142				
episode: 56	score: 2.0	memory length: 10150	epsilon: 1.0	steps: 219	lr:
0.0001	evaluation reward: 1.3508771929824561				
episode: 57	score: 0.0	memory length: 10274	epsilon: 1.0	steps: 124	lr:
0.0001	evaluation reward: 1.3275862068965518				
episode: 58	score: 0.0	memory length: 10398	epsilon: 1.0	steps: 124	lr:
0.0001	evaluation reward: 1.305084745762712				
episode: 59	score: 3.0	memory length: 10646	epsilon: 1.0	steps: 248	lr:
0.0001	evaluation reward: 1.3333333333333333				

episode: 60 score: 1.0 memory length: 10798 0.0001 evaluation reward: 1.3278688524590163	epsilon: 1.0 steps: 152 lr:
episode: 61 score: 3.0 memory length: 11046 0.0001 evaluation reward: 1.3548387096774193	epsilon: 1.0 steps: 248 lr:
episode: 62 score: 0.0 memory length: 11170 0.0001 evaluation reward: 1.3333333333333333	epsilon: 1.0 steps: 124 lr:
episode: 63 score: 3.0 memory length: 11416 0.0001 evaluation reward: 1.359375	epsilon: 1.0 steps: 246 lr:
episode: 64 score: 0.0 memory length: 11540 0.0001 evaluation reward: 1.3384615384615384	epsilon: 1.0 steps: 124 lr:
episode: 65 score: 4.0 memory length: 11833 0.0001 evaluation reward: 1.378787878787879	epsilon: 1.0 steps: 293 lr:
episode: 66 score: 0.0 memory length: 11957 0.0001 evaluation reward: 1.3582089552238805	epsilon: 1.0 steps: 124 lr:
episode: 67 score: 2.0 memory length: 12178 0.0001 evaluation reward: 1.3676470588235294	epsilon: 1.0 steps: 221 lr:
episode: 68 score: 2.0 memory length: 12396 0.0001 evaluation reward: 1.3768115942028984	epsilon: 1.0 steps: 218 lr:
episode: 69 score: 1.0 memory length: 12567 0.0001 evaluation reward: 1.3714285714285714	epsilon: 1.0 steps: 171 lr:
episode: 70 score: 2.0 memory length: 12770 0.0001 evaluation reward: 1.380281690140845	epsilon: 1.0 steps: 203 lr:
episode: 71 score: 2.0 memory length: 12951 0.0001 evaluation reward: 1.3888888888888888	epsilon: 1.0 steps: 181 lr:
episode: 72 score: 0.0 memory length: 13074 0.0001 evaluation reward: 1.36986301369863	epsilon: 1.0 steps: 123 lr:
episode: 73 score: 4.0 memory length: 13342 0.0001 evaluation reward: 1.4054054054054055	epsilon: 1.0 steps: 268 lr:
episode: 74 score: 1.0 memory length: 13513 0.0001 evaluation reward: 1.4	epsilon: 1.0 steps: 171 lr:
episode: 75 score: 0.0 memory length: 13636 0.0001 evaluation reward: 1.381578947368421	epsilon: 1.0 steps: 123 lr:
episode: 76 score: 1.0 memory length: 13807 0.0001 evaluation reward: 1.3766233766233766	epsilon: 1.0 steps: 171 lr:
episode: 77 score: 0.0 memory length: 13931 0.0001 evaluation reward: 1.358974358974359	epsilon: 1.0 steps: 124 lr:
episode: 78 score: 1.0 memory length: 14082 0.0001 evaluation reward: 1.3544303797468353	epsilon: 1.0 steps: 151 lr:
episode: 79 score: 4.0 memory length: 14376 0.0001 evaluation reward: 1.3875	epsilon: 1.0 steps: 294 lr:
episode: 80 score: 3.0 memory length: 14644 0.0001 evaluation reward: 1.4074074074074074	epsilon: 1.0 steps: 268 lr:
episode: 81 score: 0.0 memory length: 14768 0.0001 evaluation reward: 1.3902439024390243	epsilon: 1.0 steps: 124 lr:
episode: 82 score: 2.0 memory length: 14987 0.0001 evaluation reward: 1.3975903614457832	epsilon: 1.0 steps: 219 lr:
episode: 83 score: 3.0 memory length: 15213 0.0001 evaluation reward: 1.4166666666666667	epsilon: 1.0 steps: 226 lr:
episode: 84 score: 0.0 memory length: 15337 0.0001 evaluation reward: 1.4	epsilon: 1.0 steps: 124 lr:
episode: 85 score: 2.0 memory length: 15560 0.0001 evaluation reward: 1.4069767441860466	epsilon: 1.0 steps: 223 lr:
episode: 86 score: 0.0 memory length: 15684 0.0001 evaluation reward: 1.3908045977011494	epsilon: 1.0 steps: 124 lr:
episode: 87 score: 3.0 memory length: 15913 0.0001 evaluation reward: 1.4090909090909092	epsilon: 1.0 steps: 229 lr:
episode: 88 score: 0.0 memory length: 16037 0.0001 evaluation reward: 1.3932584269662922	epsilon: 1.0 steps: 124 lr:
episode: 89 score: 2.0 memory length: 16218 0.0001 evaluation reward: 1.4	epsilon: 1.0 steps: 181 lr:

episode: 90	score: 0.0	memory length: 16342	epsilon: 1.0	steps: 124	lr:
0.0001	evaluation reward: 1.3846153846153846				
episode: 91	score: 1.0	memory length: 16512	epsilon: 1.0	steps: 170	lr:
0.0001	evaluation reward: 1.3804347826086956				
episode: 92	score: 3.0	memory length: 16760	epsilon: 1.0	steps: 248	lr:
0.0001	evaluation reward: 1.3978494623655915				
episode: 93	score: 4.0	memory length: 17037	epsilon: 1.0	steps: 277	lr:
0.0001	evaluation reward: 1.425531914893617				
episode: 94	score: 0.0	memory length: 17161	epsilon: 1.0	steps: 124	lr:
0.0001	evaluation reward: 1.4105263157894736				
episode: 95	score: 2.0	memory length: 17359	epsilon: 1.0	steps: 198	lr:
0.0001	evaluation reward: 1.4166666666666667				
episode: 96	score: 0.0	memory length: 17483	epsilon: 1.0	steps: 124	lr:
0.0001	evaluation reward: 1.402061855670103				
episode: 97	score: 4.0	memory length: 17759	epsilon: 1.0	steps: 276	lr:
0.0001	evaluation reward: 1.4285714285714286				
episode: 98	score: 0.0	memory length: 17883	epsilon: 1.0	steps: 124	lr:
0.0001	evaluation reward: 1.4141414141414141				
episode: 99	score: 1.0	memory length: 18054	epsilon: 1.0	steps: 171	lr:
0.0001	evaluation reward: 1.41				
episode: 100	score: 0.0	memory length: 18177	epsilon: 1.0	steps: 123	lr:
0.0001	evaluation reward: 1.37				
episode: 101	score: 0.0	memory length: 18301	epsilon: 1.0	steps: 124	lr:
0.0001	evaluation reward: 1.37				
episode: 102	score: 2.0	memory length: 18520	epsilon: 1.0	steps: 219	lr:
0.0001	evaluation reward: 1.37				
episode: 103	score: 2.0	memory length: 18739	epsilon: 1.0	steps: 219	lr:
0.0001	evaluation reward: 1.38				
episode: 104	score: 3.0	memory length: 18988	epsilon: 1.0	steps: 249	lr:
0.0001	evaluation reward: 1.41				
episode: 105	score: 1.0	memory length: 19158	epsilon: 1.0	steps: 170	lr:
0.0001	evaluation reward: 1.39				
episode: 106	score: 2.0	memory length: 19338	epsilon: 1.0	steps: 180	lr:
0.0001	evaluation reward: 1.41				
episode: 107	score: 1.0	memory length: 19511	epsilon: 1.0	steps: 173	lr:
0.0001	evaluation reward: 1.42				
episode: 108	score: 0.0	memory length: 19635	epsilon: 1.0	steps: 124	lr:
0.0001	evaluation reward: 1.42				
episode: 109	score: 2.0	memory length: 19833	epsilon: 1.0	steps: 198	lr:
0.0001	evaluation reward: 1.44				
episode: 110	score: 1.0	memory length: 19985	epsilon: 1.0	steps: 152	lr:
0.0001	evaluation reward: 1.45				
episode: 111	score: 3.0	memory length: 20233	epsilon: 1.0	steps: 248	lr:
0.0001	evaluation reward: 1.45				
episode: 112	score: 2.0	memory length: 20431	epsilon: 1.0	steps: 198	lr:
0.0001	evaluation reward: 1.46				
episode: 113	score: 2.0	memory length: 20630	epsilon: 1.0	steps: 199	lr:
0.0001	evaluation reward: 1.48				
episode: 114	score: 1.0	memory length: 20800	epsilon: 1.0	steps: 170	lr:
0.0001	evaluation reward: 1.47				
episode: 115	score: 1.0	memory length: 20970	epsilon: 1.0	steps: 170	lr:
0.0001	evaluation reward: 1.48				
episode: 116	score: 0.0	memory length: 21093	epsilon: 1.0	steps: 123	lr:
0.0001	evaluation reward: 1.45				
episode: 117	score: 3.0	memory length: 21339	epsilon: 1.0	steps: 246	lr:
0.0001	evaluation reward: 1.45				
episode: 118	score: 2.0	memory length: 21537	epsilon: 1.0	steps: 198	lr:
0.0001	evaluation reward: 1.45				
episode: 119	score: 0.0	memory length: 21661	epsilon: 1.0	steps: 124	lr:
0.0001	evaluation reward: 1.43				

episode: 120 score: 1.0 memory length: 21832 epsilon: 1.0 steps: 171 lr: 0.0001 evaluation reward: 1.43
episode: 121 score: 1.0 memory length: 21984 epsilon: 1.0 steps: 152 lr: 0.0001 evaluation reward: 1.41
episode: 122 score: 2.0 memory length: 22183 epsilon: 1.0 steps: 199 lr: 0.0001 evaluation reward: 1.42
episode: 123 score: 0.0 memory length: 22306 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.42
episode: 124 score: 2.0 memory length: 22527 epsilon: 1.0 steps: 221 lr: 0.0001 evaluation reward: 1.43
episode: 125 score: 2.0 memory length: 22725 epsilon: 1.0 steps: 198 lr: 0.0001 evaluation reward: 1.45
episode: 126 score: 2.0 memory length: 22923 epsilon: 1.0 steps: 198 lr: 0.0001 evaluation reward: 1.45
episode: 127 score: 0.0 memory length: 23046 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.45
episode: 128 score: 0.0 memory length: 23170 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.44
episode: 129 score: 0.0 memory length: 23294 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.42
episode: 130 score: 2.0 memory length: 23512 epsilon: 1.0 steps: 218 lr: 0.0001 evaluation reward: 1.42
episode: 131 score: 0.0 memory length: 23635 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.4
episode: 132 score: 1.0 memory length: 23805 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.39
episode: 133 score: 0.0 memory length: 23929 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.39
episode: 134 score: 0.0 memory length: 24052 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.37
episode: 135 score: 4.0 memory length: 24332 epsilon: 1.0 steps: 280 lr: 0.0001 evaluation reward: 1.41
episode: 136 score: 0.0 memory length: 24456 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.41
episode: 137 score: 1.0 memory length: 24626 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.41
episode: 138 score: 2.0 memory length: 24824 epsilon: 1.0 steps: 198 lr: 0.0001 evaluation reward: 1.43
episode: 139 score: 1.0 memory length: 24976 epsilon: 1.0 steps: 152 lr: 0.0001 evaluation reward: 1.42
episode: 140 score: 3.0 memory length: 25245 epsilon: 1.0 steps: 269 lr: 0.0001 evaluation reward: 1.43
episode: 141 score: 1.0 memory length: 25397 epsilon: 1.0 steps: 152 lr: 0.0001 evaluation reward: 1.41
episode: 142 score: 1.0 memory length: 25567 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.4
episode: 143 score: 2.0 memory length: 25749 epsilon: 1.0 steps: 182 lr: 0.0001 evaluation reward: 1.4
episode: 144 score: 2.0 memory length: 25966 epsilon: 1.0 steps: 217 lr: 0.0001 evaluation reward: 1.39
episode: 145 score: 2.0 memory length: 26189 epsilon: 1.0 steps: 223 lr: 0.0001 evaluation reward: 1.39
episode: 146 score: 1.0 memory length: 26359 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.38
episode: 147 score: 1.0 memory length: 26511 epsilon: 1.0 steps: 152 lr: 0.0001 evaluation reward: 1.34
episode: 148 score: 1.0 memory length: 26684 epsilon: 1.0 steps: 173 lr: 0.0001 evaluation reward: 1.35
episode: 149 score: 2.0 memory length: 26901 epsilon: 1.0 steps: 217 lr: 0.0001 evaluation reward: 1.37

episode: 150	score: 0.0	memory length: 27025	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.36					
episode: 151	score: 3.0	memory length: 27253	epsilon: 1.0	steps: 228	lr: 0.0001
evaluation reward: 1.38					
episode: 152	score: 2.0	memory length: 27472	epsilon: 1.0	steps: 219	lr: 0.0001
evaluation reward: 1.39					
episode: 153	score: 3.0	memory length: 27699	epsilon: 1.0	steps: 227	lr: 0.0001
evaluation reward: 1.42					
episode: 154	score: 1.0	memory length: 27870	epsilon: 1.0	steps: 171	lr: 0.0001
evaluation reward: 1.42					
episode: 155	score: 1.0	memory length: 28022	epsilon: 1.0	steps: 152	lr: 0.0001
evaluation reward: 1.41					
episode: 156	score: 0.0	memory length: 28146	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.39					
episode: 157	score: 1.0	memory length: 28315	epsilon: 1.0	steps: 169	lr: 0.0001
evaluation reward: 1.4					
episode: 158	score: 4.0	memory length: 28631	epsilon: 1.0	steps: 316	lr: 0.0001
evaluation reward: 1.44					
episode: 159	score: 0.0	memory length: 28755	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.41					
episode: 160	score: 1.0	memory length: 28924	epsilon: 1.0	steps: 169	lr: 0.0001
evaluation reward: 1.41					
episode: 161	score: 2.0	memory length: 29140	epsilon: 1.0	steps: 216	lr: 0.0001
evaluation reward: 1.4					
episode: 162	score: 1.0	memory length: 29310	epsilon: 1.0	steps: 170	lr: 0.0001
evaluation reward: 1.41					
episode: 163	score: 2.0	memory length: 29532	epsilon: 1.0	steps: 222	lr: 0.0001
evaluation reward: 1.4					
episode: 164	score: 4.0	memory length: 29819	epsilon: 1.0	steps: 287	lr: 0.0001
evaluation reward: 1.44					
episode: 165	score: 2.0	memory length: 30018	epsilon: 1.0	steps: 199	lr: 0.0001
evaluation reward: 1.42					
episode: 166	score: 0.0	memory length: 30141	epsilon: 1.0	steps: 123	lr: 0.0001
evaluation reward: 1.42					
episode: 167	score: 0.0	memory length: 30265	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.4					
episode: 168	score: 2.0	memory length: 30464	epsilon: 1.0	steps: 199	lr: 0.0001
evaluation reward: 1.4					
episode: 169	score: 5.0	memory length: 30792	epsilon: 1.0	steps: 328	lr: 0.0001
evaluation reward: 1.44					
episode: 170	score: 0.0	memory length: 30915	epsilon: 1.0	steps: 123	lr: 0.0001
evaluation reward: 1.42					
episode: 171	score: 2.0	memory length: 31113	epsilon: 1.0	steps: 198	lr: 0.0001
evaluation reward: 1.42					
episode: 172	score: 0.0	memory length: 31237	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.42					
episode: 173	score: 2.0	memory length: 31455	epsilon: 1.0	steps: 218	lr: 0.0001
evaluation reward: 1.4					
episode: 174	score: 0.0	memory length: 31578	epsilon: 1.0	steps: 123	lr: 0.0001
evaluation reward: 1.39					
episode: 175	score: 3.0	memory length: 31829	epsilon: 1.0	steps: 251	lr: 0.0001
evaluation reward: 1.42					
episode: 176	score: 1.0	memory length: 31999	epsilon: 1.0	steps: 170	lr: 0.0001
evaluation reward: 1.42					
episode: 177	score: 4.0	memory length: 32313	epsilon: 1.0	steps: 314	lr: 0.0001
evaluation reward: 1.46					
episode: 178	score: 3.0	memory length: 32544	epsilon: 1.0	steps: 231	lr: 0.0001
evaluation reward: 1.48					
episode: 179	score: 3.0	memory length: 32776	epsilon: 1.0	steps: 232	lr: 0.0001
evaluation reward: 1.47					

episode: 180	score: 1.0	memory length: 32946	epsilon: 1.0	steps: 170	lr: 0.0001
					evaluation reward: 1.45
episode: 181	score: 3.0	memory length: 33194	epsilon: 1.0	steps: 248	lr: 0.0001
					evaluation reward: 1.48
episode: 182	score: 0.0	memory length: 33318	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.46
episode: 183	score: 5.0	memory length: 33647	epsilon: 1.0	steps: 329	lr: 0.0001
					evaluation reward: 1.48
episode: 184	score: 1.0	memory length: 33799	epsilon: 1.0	steps: 152	lr: 0.0001
					evaluation reward: 1.49
episode: 185	score: 2.0	memory length: 34018	epsilon: 1.0	steps: 219	lr: 0.0001
					evaluation reward: 1.49
episode: 186	score: 2.0	memory length: 34217	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.51
episode: 187	score: 0.0	memory length: 34340	epsilon: 1.0	steps: 123	lr: 0.0001
					evaluation reward: 1.48
episode: 188	score: 2.0	memory length: 34539	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.5
episode: 189	score: 3.0	memory length: 34807	epsilon: 1.0	steps: 268	lr: 0.0001
					evaluation reward: 1.51
episode: 190	score: 2.0	memory length: 35006	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.53
episode: 191	score: 2.0	memory length: 35224	epsilon: 1.0	steps: 218	lr: 0.0001
					evaluation reward: 1.54
episode: 192	score: 1.0	memory length: 35377	epsilon: 1.0	steps: 153	lr: 0.0001
					evaluation reward: 1.52
episode: 193	score: 0.0	memory length: 35500	epsilon: 1.0	steps: 123	lr: 0.0001
					evaluation reward: 1.48
episode: 194	score: 3.0	memory length: 35745	epsilon: 1.0	steps: 245	lr: 0.0001
					evaluation reward: 1.51
episode: 195	score: 0.0	memory length: 35869	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.49
episode: 196	score: 1.0	memory length: 36039	epsilon: 1.0	steps: 170	lr: 0.0001
					evaluation reward: 1.5
episode: 197	score: 2.0	memory length: 36237	epsilon: 1.0	steps: 198	lr: 0.0001
					evaluation reward: 1.48
episode: 198	score: 0.0	memory length: 36360	epsilon: 1.0	steps: 123	lr: 0.0001
					evaluation reward: 1.48
episode: 199	score: 2.0	memory length: 36559	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.49
episode: 200	score: 3.0	memory length: 36804	epsilon: 1.0	steps: 245	lr: 0.0001
					evaluation reward: 1.52
episode: 201	score: 0.0	memory length: 36928	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.52
episode: 202	score: 0.0	memory length: 37052	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.5
episode: 203	score: 2.0	memory length: 37233	epsilon: 1.0	steps: 181	lr: 0.0001
					evaluation reward: 1.5
episode: 204	score: 1.0	memory length: 37402	epsilon: 1.0	steps: 169	lr: 0.0001
					evaluation reward: 1.48
episode: 205	score: 1.0	memory length: 37571	epsilon: 1.0	steps: 169	lr: 0.0001
					evaluation reward: 1.48
episode: 206	score: 4.0	memory length: 37841	epsilon: 1.0	steps: 270	lr: 0.0001
					evaluation reward: 1.5
episode: 207	score: 6.0	memory length: 38234	epsilon: 1.0	steps: 393	lr: 0.0001
					evaluation reward: 1.55
episode: 208	score: 1.0	memory length: 38385	epsilon: 1.0	steps: 151	lr: 0.0001
					evaluation reward: 1.56
episode: 209	score: 2.0	memory length: 38584	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.56

episode: 210	score: 2.0	memory length: 38803	epsilon: 1.0	steps: 219	lr: 0.0001
					evaluation reward: 1.57
episode: 211	score: 3.0	memory length: 39051	epsilon: 1.0	steps: 248	lr: 0.0001
					evaluation reward: 1.57
episode: 212	score: 0.0	memory length: 39175	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.55
episode: 213	score: 0.0	memory length: 39298	epsilon: 1.0	steps: 123	lr: 0.0001
					evaluation reward: 1.53
episode: 214	score: 1.0	memory length: 39468	epsilon: 1.0	steps: 170	lr: 0.0001
					evaluation reward: 1.53
episode: 215	score: 2.0	memory length: 39649	epsilon: 1.0	steps: 181	lr: 0.0001
					evaluation reward: 1.54
episode: 216	score: 0.0	memory length: 39773	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.54
episode: 217	score: 0.0	memory length: 39896	epsilon: 1.0	steps: 123	lr: 0.0001
					evaluation reward: 1.51
episode: 218	score: 2.0	memory length: 40116	epsilon: 1.0	steps: 220	lr: 0.0001
					evaluation reward: 1.51
episode: 219	score: 1.0	memory length: 40269	epsilon: 1.0	steps: 153	lr: 0.0001
					evaluation reward: 1.52
episode: 220	score: 2.0	memory length: 40467	epsilon: 1.0	steps: 198	lr: 0.0001
					evaluation reward: 1.53
episode: 221	score: 1.0	memory length: 40618	epsilon: 1.0	steps: 151	lr: 0.0001
					evaluation reward: 1.53
episode: 222	score: 8.0	memory length: 41049	epsilon: 1.0	steps: 431	lr: 0.0001
					evaluation reward: 1.59
episode: 223	score: 2.0	memory length: 41268	epsilon: 1.0	steps: 219	lr: 0.0001
					evaluation reward: 1.61
episode: 224	score: 3.0	memory length: 41516	epsilon: 1.0	steps: 248	lr: 0.0001
					evaluation reward: 1.62
episode: 225	score: 1.0	memory length: 41668	epsilon: 1.0	steps: 152	lr: 0.0001
					evaluation reward: 1.61
episode: 226	score: 2.0	memory length: 41867	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.61
episode: 227	score: 4.0	memory length: 42165	epsilon: 1.0	steps: 298	lr: 0.0001
					evaluation reward: 1.65
episode: 228	score: 1.0	memory length: 42335	epsilon: 1.0	steps: 170	lr: 0.0001
					evaluation reward: 1.66
episode: 229	score: 2.0	memory length: 42534	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.68
episode: 230	score: 1.0	memory length: 42703	epsilon: 1.0	steps: 169	lr: 0.0001
					evaluation reward: 1.67
episode: 231	score: 0.0	memory length: 42827	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.67
episode: 232	score: 2.0	memory length: 43030	epsilon: 1.0	steps: 203	lr: 0.0001
					evaluation reward: 1.68
episode: 233	score: 2.0	memory length: 43247	epsilon: 1.0	steps: 217	lr: 0.0001
					evaluation reward: 1.7
episode: 234	score: 1.0	memory length: 43417	epsilon: 1.0	steps: 170	lr: 0.0001
					evaluation reward: 1.71
episode: 235	score: 1.0	memory length: 43586	epsilon: 1.0	steps: 169	lr: 0.0001
					evaluation reward: 1.68
episode: 236	score: 2.0	memory length: 43804	epsilon: 1.0	steps: 218	lr: 0.0001
					evaluation reward: 1.7
episode: 237	score: 1.0	memory length: 43956	epsilon: 1.0	steps: 152	lr: 0.0001
					evaluation reward: 1.7
episode: 238	score: 0.0	memory length: 44079	epsilon: 1.0	steps: 123	lr: 0.0001
					evaluation reward: 1.68
episode: 239	score: 1.0	memory length: 44249	epsilon: 1.0	steps: 170	lr: 0.0001
					evaluation reward: 1.68

episode: 240	score: 2.0	memory length: 44448	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.67
episode: 241	score: 1.0	memory length: 44621	epsilon: 1.0	steps: 173	lr: 0.0001
					evaluation reward: 1.67
episode: 242	score: 0.0	memory length: 44745	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.66
episode: 243	score: 0.0	memory length: 44869	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.64
episode: 244	score: 3.0	memory length: 45120	epsilon: 1.0	steps: 251	lr: 0.0001
					evaluation reward: 1.65
episode: 245	score: 2.0	memory length: 45319	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.65
episode: 246	score: 3.0	memory length: 45567	epsilon: 1.0	steps: 248	lr: 0.0001
					evaluation reward: 1.67
episode: 247	score: 2.0	memory length: 45766	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.68
episode: 248	score: 6.0	memory length: 46165	epsilon: 1.0	steps: 399	lr: 0.0001
					evaluation reward: 1.73
episode: 249	score: 3.0	memory length: 46393	epsilon: 1.0	steps: 228	lr: 0.0001
					evaluation reward: 1.74
episode: 250	score: 3.0	memory length: 46641	epsilon: 1.0	steps: 248	lr: 0.0001
					evaluation reward: 1.77
episode: 251	score: 2.0	memory length: 46839	epsilon: 1.0	steps: 198	lr: 0.0001
					evaluation reward: 1.76
episode: 252	score: 1.0	memory length: 46991	epsilon: 1.0	steps: 152	lr: 0.0001
					evaluation reward: 1.75
episode: 253	score: 0.0	memory length: 47115	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.72
episode: 254	score: 2.0	memory length: 47335	epsilon: 1.0	steps: 220	lr: 0.0001
					evaluation reward: 1.73
episode: 255	score: 0.0	memory length: 47459	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.72
episode: 256	score: 1.0	memory length: 47628	epsilon: 1.0	steps: 169	lr: 0.0001
					evaluation reward: 1.73
episode: 257	score: 2.0	memory length: 47844	epsilon: 1.0	steps: 216	lr: 0.0001
					evaluation reward: 1.74
episode: 258	score: 1.0	memory length: 48015	epsilon: 1.0	steps: 171	lr: 0.0001
					evaluation reward: 1.71
episode: 259	score: 0.0	memory length: 48139	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.71
episode: 260	score: 3.0	memory length: 48368	epsilon: 1.0	steps: 229	lr: 0.0001
					evaluation reward: 1.73
episode: 261	score: 2.0	memory length: 48567	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.73
episode: 262	score: 1.0	memory length: 48718	epsilon: 1.0	steps: 151	lr: 0.0001
					evaluation reward: 1.73
episode: 263	score: 0.0	memory length: 48842	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.71
episode: 264	score: 1.0	memory length: 49012	epsilon: 1.0	steps: 170	lr: 0.0001
					evaluation reward: 1.68
episode: 265	score: 0.0	memory length: 49136	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.66
episode: 266	score: 1.0	memory length: 49307	epsilon: 1.0	steps: 171	lr: 0.0001
					evaluation reward: 1.67
episode: 267	score: 2.0	memory length: 49506	epsilon: 1.0	steps: 199	lr: 0.0001
					evaluation reward: 1.69
episode: 268	score: 1.0	memory length: 49678	epsilon: 1.0	steps: 172	lr: 0.0001
					evaluation reward: 1.68
episode: 269	score: 0.0	memory length: 49802	epsilon: 1.0	steps: 124	lr: 0.0001
					evaluation reward: 1.63

episode: 270	score: 1.0	memory length: 49972	epsilon: 1.0	steps: 170	lr: 0.0001
evaluation reward: 1.64					
episode: 271	score: 2.0	memory length: 50171	epsilon: 1.0	steps: 199	lr: 0.0001
evaluation reward: 1.64					
episode: 272	score: 1.0	memory length: 50340	epsilon: 1.0	steps: 169	lr: 0.0001
evaluation reward: 1.65					
episode: 273	score: 1.0	memory length: 50492	epsilon: 1.0	steps: 152	lr: 0.0001
evaluation reward: 1.64					
episode: 274	score: 3.0	memory length: 50740	epsilon: 1.0	steps: 248	lr: 0.0001
evaluation reward: 1.67					
episode: 275	score: 3.0	memory length: 50989	epsilon: 1.0	steps: 249	lr: 0.0001
evaluation reward: 1.67					
episode: 276	score: 1.0	memory length: 51140	epsilon: 1.0	steps: 151	lr: 0.0001
evaluation reward: 1.67					
episode: 277	score: 1.0	memory length: 51309	epsilon: 1.0	steps: 169	lr: 0.0001
evaluation reward: 1.64					
episode: 278	score: 1.0	memory length: 51461	epsilon: 1.0	steps: 152	lr: 0.0001
evaluation reward: 1.62					
episode: 279	score: 0.0	memory length: 51585	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.59					
episode: 280	score: 1.0	memory length: 51755	epsilon: 1.0	steps: 170	lr: 0.0001
evaluation reward: 1.59					
episode: 281	score: 3.0	memory length: 51983	epsilon: 1.0	steps: 228	lr: 0.0001
evaluation reward: 1.59					
episode: 282	score: 0.0	memory length: 52107	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.59					
episode: 283	score: 0.0	memory length: 52230	epsilon: 1.0	steps: 123	lr: 0.0001
evaluation reward: 1.54					
episode: 284	score: 2.0	memory length: 52432	epsilon: 1.0	steps: 202	lr: 0.0001
evaluation reward: 1.55					
episode: 285	score: 3.0	memory length: 52679	epsilon: 1.0	steps: 247	lr: 0.0001
evaluation reward: 1.56					
episode: 286	score: 2.0	memory length: 52897	epsilon: 1.0	steps: 218	lr: 0.0001
evaluation reward: 1.56					
episode: 287	score: 3.0	memory length: 53146	epsilon: 1.0	steps: 249	lr: 0.0001
evaluation reward: 1.59					
episode: 288	score: 0.0	memory length: 53269	epsilon: 1.0	steps: 123	lr: 0.0001
evaluation reward: 1.57					
episode: 289	score: 1.0	memory length: 53439	epsilon: 1.0	steps: 170	lr: 0.0001
evaluation reward: 1.55					
episode: 290	score: 0.0	memory length: 53563	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.53					
episode: 291	score: 1.0	memory length: 53714	epsilon: 1.0	steps: 151	lr: 0.0001
evaluation reward: 1.52					
episode: 292	score: 0.0	memory length: 53838	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.51					
episode: 293	score: 1.0	memory length: 54007	epsilon: 1.0	steps: 169	lr: 0.0001
evaluation reward: 1.52					
episode: 294	score: 1.0	memory length: 54160	epsilon: 1.0	steps: 153	lr: 0.0001
evaluation reward: 1.5					
episode: 295	score: 2.0	memory length: 54341	epsilon: 1.0	steps: 181	lr: 0.0001
evaluation reward: 1.52					
episode: 296	score: 2.0	memory length: 54539	epsilon: 1.0	steps: 198	lr: 0.0001
evaluation reward: 1.53					
episode: 297	score: 0.0	memory length: 54663	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.51					
episode: 298	score: 2.0	memory length: 54885	epsilon: 1.0	steps: 222	lr: 0.0001
evaluation reward: 1.53					
episode: 299	score: 0.0	memory length: 55009	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.51					

episode: 300 score: 2.0 memory length: 55229 epsilon: 1.0 steps: 220 lr: 0.0001 evaluation reward: 1.5
episode: 301 score: 5.0 memory length: 55533 epsilon: 1.0 steps: 304 lr: 0.0001 evaluation reward: 1.55
episode: 302 score: 0.0 memory length: 55657 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.55
episode: 303 score: 1.0 memory length: 55809 epsilon: 1.0 steps: 152 lr: 0.0001 evaluation reward: 1.54
episode: 304 score: 0.0 memory length: 55933 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.53
episode: 305 score: 0.0 memory length: 56057 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.52
episode: 306 score: 0.0 memory length: 56180 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.48
episode: 307 score: 0.0 memory length: 56304 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.42
episode: 308 score: 0.0 memory length: 56427 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.41
episode: 309 score: 0.0 memory length: 56551 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.39
episode: 310 score: 5.0 memory length: 56869 epsilon: 1.0 steps: 318 lr: 0.0001 evaluation reward: 1.42
episode: 311 score: 5.0 memory length: 57171 epsilon: 1.0 steps: 302 lr: 0.0001 evaluation reward: 1.44
episode: 312 score: 0.0 memory length: 57294 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.44
episode: 313 score: 0.0 memory length: 57418 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.44
episode: 314 score: 2.0 memory length: 57616 epsilon: 1.0 steps: 198 lr: 0.0001 evaluation reward: 1.45
episode: 315 score: 0.0 memory length: 57739 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.43
episode: 316 score: 3.0 memory length: 57966 epsilon: 1.0 steps: 227 lr: 0.0001 evaluation reward: 1.46
episode: 317 score: 0.0 memory length: 58089 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.46
episode: 318 score: 5.0 memory length: 58362 epsilon: 1.0 steps: 273 lr: 0.0001 evaluation reward: 1.49
episode: 319 score: 1.0 memory length: 58513 epsilon: 1.0 steps: 151 lr: 0.0001 evaluation reward: 1.49
episode: 320 score: 2.0 memory length: 58732 epsilon: 1.0 steps: 219 lr: 0.0001 evaluation reward: 1.49
episode: 321 score: 1.0 memory length: 58902 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.49
episode: 322 score: 1.0 memory length: 59073 epsilon: 1.0 steps: 171 lr: 0.0001 evaluation reward: 1.42
episode: 323 score: 3.0 memory length: 59341 epsilon: 1.0 steps: 268 lr: 0.0001 evaluation reward: 1.43
episode: 324 score: 4.0 memory length: 59662 epsilon: 1.0 steps: 321 lr: 0.0001 evaluation reward: 1.44
episode: 325 score: 3.0 memory length: 59910 epsilon: 1.0 steps: 248 lr: 0.0001 evaluation reward: 1.46
episode: 326 score: 1.0 memory length: 60061 epsilon: 1.0 steps: 151 lr: 0.0001 evaluation reward: 1.45
episode: 327 score: 2.0 memory length: 60260 epsilon: 1.0 steps: 199 lr: 0.0001 evaluation reward: 1.43
episode: 328 score: 1.0 memory length: 60411 epsilon: 1.0 steps: 151 lr: 0.0001 evaluation reward: 1.43
episode: 329 score: 0.0 memory length: 60534 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.41

episode: 330	score: 2.0	memory length: 60753	epsilon: 1.0	steps: 219	lr: 0.0001
	evaluation reward: 1.42				
episode: 331	score: 4.0	memory length: 61010	epsilon: 1.0	steps: 257	lr: 0.0001
	evaluation reward: 1.46				
episode: 332	score: 0.0	memory length: 61134	epsilon: 1.0	steps: 124	lr: 0.0001
	evaluation reward: 1.44				
episode: 333	score: 2.0	memory length: 61332	epsilon: 1.0	steps: 198	lr: 0.0001
	evaluation reward: 1.44				
episode: 334	score: 7.0	memory length: 61760	epsilon: 1.0	steps: 428	lr: 0.0001
	evaluation reward: 1.5				
episode: 335	score: 2.0	memory length: 61959	epsilon: 1.0	steps: 199	lr: 0.0001
	evaluation reward: 1.51				
episode: 336	score: 2.0	memory length: 62177	epsilon: 1.0	steps: 218	lr: 0.0001
	evaluation reward: 1.51				
episode: 337	score: 0.0	memory length: 62301	epsilon: 1.0	steps: 124	lr: 0.0001
	evaluation reward: 1.5				
episode: 338	score: 0.0	memory length: 62425	epsilon: 1.0	steps: 124	lr: 0.0001
	evaluation reward: 1.5				
episode: 339	score: 1.0	memory length: 62596	epsilon: 1.0	steps: 171	lr: 0.0001
	evaluation reward: 1.5				
episode: 340	score: 1.0	memory length: 62766	epsilon: 1.0	steps: 170	lr: 0.0001
	evaluation reward: 1.49				
episode: 341	score: 2.0	memory length: 62985	epsilon: 1.0	steps: 219	lr: 0.0001
	evaluation reward: 1.5				
episode: 342	score: 2.0	memory length: 63205	epsilon: 1.0	steps: 220	lr: 0.0001
	evaluation reward: 1.52				
episode: 343	score: 0.0	memory length: 63329	epsilon: 1.0	steps: 124	lr: 0.0001
	evaluation reward: 1.52				
episode: 344	score: 1.0	memory length: 63501	epsilon: 1.0	steps: 172	lr: 0.0001
	evaluation reward: 1.5				
episode: 345	score: 4.0	memory length: 63782	epsilon: 1.0	steps: 281	lr: 0.0001
	evaluation reward: 1.52				
episode: 346	score: 1.0	memory length: 63933	epsilon: 1.0	steps: 151	lr: 0.0001
	evaluation reward: 1.5				
episode: 347	score: 0.0	memory length: 64057	epsilon: 1.0	steps: 124	lr: 0.0001
	evaluation reward: 1.48				
episode: 348	score: 1.0	memory length: 64208	epsilon: 1.0	steps: 151	lr: 0.0001
	evaluation reward: 1.43				
episode: 349	score: 0.0	memory length: 64332	epsilon: 1.0	steps: 124	lr: 0.0001
	evaluation reward: 1.4				
episode: 350	score: 0.0	memory length: 64456	epsilon: 1.0	steps: 124	lr: 0.0001
	evaluation reward: 1.37				
episode: 351	score: 1.0	memory length: 64608	epsilon: 1.0	steps: 152	lr: 0.0001
	evaluation reward: 1.36				
episode: 352	score: 2.0	memory length: 64828	epsilon: 1.0	steps: 220	lr: 0.0001
	evaluation reward: 1.37				
episode: 353	score: 2.0	memory length: 65026	epsilon: 1.0	steps: 198	lr: 0.0001
	evaluation reward: 1.39				
episode: 354	score: 0.0	memory length: 65150	epsilon: 1.0	steps: 124	lr: 0.0001
	evaluation reward: 1.37				
episode: 355	score: 0.0	memory length: 65274	epsilon: 1.0	steps: 124	lr: 0.0001
	evaluation reward: 1.37				
episode: 356	score: 2.0	memory length: 65494	epsilon: 1.0	steps: 220	lr: 0.0001
	evaluation reward: 1.38				
episode: 357	score: 2.0	memory length: 65694	epsilon: 1.0	steps: 200	lr: 0.0001
	evaluation reward: 1.38				
episode: 358	score: 4.0	memory length: 65977	epsilon: 1.0	steps: 283	lr: 0.0001
	evaluation reward: 1.41				
episode: 359	score: 0.0	memory length: 66100	epsilon: 1.0	steps: 123	lr: 0.0001
	evaluation reward: 1.41				

episode: 360 score: 0.0 memory length: 66223 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.38
episode: 361 score: 0.0 memory length: 66347 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.36
episode: 362 score: 3.0 memory length: 66577 epsilon: 1.0 steps: 230 lr: 0.0001 evaluation reward: 1.38
episode: 363 score: 1.0 memory length: 66750 epsilon: 1.0 steps: 173 lr: 0.0001 evaluation reward: 1.39
episode: 364 score: 2.0 memory length: 66968 epsilon: 1.0 steps: 218 lr: 0.0001 evaluation reward: 1.4
episode: 365 score: 2.0 memory length: 67189 epsilon: 1.0 steps: 221 lr: 0.0001 evaluation reward: 1.42
episode: 366 score: 1.0 memory length: 67358 epsilon: 1.0 steps: 169 lr: 0.0001 evaluation reward: 1.42
episode: 367 score: 2.0 memory length: 67557 epsilon: 1.0 steps: 199 lr: 0.0001 evaluation reward: 1.42
episode: 368 score: 3.0 memory length: 67786 epsilon: 1.0 steps: 229 lr: 0.0001 evaluation reward: 1.44
episode: 369 score: 0.0 memory length: 67910 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.44
episode: 370 score: 0.0 memory length: 68034 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.43
episode: 371 score: 2.0 memory length: 68252 epsilon: 1.0 steps: 218 lr: 0.0001 evaluation reward: 1.43
episode: 372 score: 1.0 memory length: 68404 epsilon: 1.0 steps: 152 lr: 0.0001 evaluation reward: 1.43
episode: 373 score: 3.0 memory length: 68649 epsilon: 1.0 steps: 245 lr: 0.0001 evaluation reward: 1.45
episode: 374 score: 0.0 memory length: 68772 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.42
episode: 375 score: 3.0 memory length: 69039 epsilon: 1.0 steps: 267 lr: 0.0001 evaluation reward: 1.42
episode: 376 score: 1.0 memory length: 69212 epsilon: 1.0 steps: 173 lr: 0.0001 evaluation reward: 1.42
episode: 377 score: 3.0 memory length: 69442 epsilon: 1.0 steps: 230 lr: 0.0001 evaluation reward: 1.44
episode: 378 score: 0.0 memory length: 69566 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.43
episode: 379 score: 2.0 memory length: 69765 epsilon: 1.0 steps: 199 lr: 0.0001 evaluation reward: 1.45
episode: 380 score: 1.0 memory length: 69934 epsilon: 1.0 steps: 169 lr: 0.0001 evaluation reward: 1.45
episode: 381 score: 2.0 memory length: 70136 epsilon: 1.0 steps: 202 lr: 0.0001 evaluation reward: 1.44
episode: 382 score: 0.0 memory length: 70260 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.44
episode: 383 score: 4.0 memory length: 70554 epsilon: 1.0 steps: 294 lr: 0.0001 evaluation reward: 1.48
episode: 384 score: 0.0 memory length: 70678 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.46
episode: 385 score: 1.0 memory length: 70848 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.44
episode: 386 score: 3.0 memory length: 71076 epsilon: 1.0 steps: 228 lr: 0.0001 evaluation reward: 1.45
episode: 387 score: 2.0 memory length: 71296 epsilon: 1.0 steps: 220 lr: 0.0001 evaluation reward: 1.44
episode: 388 score: 1.0 memory length: 71468 epsilon: 1.0 steps: 172 lr: 0.0001 evaluation reward: 1.45
episode: 389 score: 1.0 memory length: 71638 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.45

episode: 390 score: 3.0 memory length: 71906 epsilon: 1.0 steps: 268 lr: 0.0001 evaluation reward: 1.48
episode: 391 score: 1.0 memory length: 72076 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.48
episode: 392 score: 2.0 memory length: 72295 epsilon: 1.0 steps: 219 lr: 0.0001 evaluation reward: 1.5
episode: 393 score: 0.0 memory length: 72418 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.49
episode: 394 score: 4.0 memory length: 72736 epsilon: 1.0 steps: 318 lr: 0.0001 evaluation reward: 1.52
episode: 395 score: 0.0 memory length: 72860 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.5
episode: 396 score: 3.0 memory length: 73131 epsilon: 1.0 steps: 271 lr: 0.0001 evaluation reward: 1.51
episode: 397 score: 2.0 memory length: 73311 epsilon: 1.0 steps: 180 lr: 0.0001 evaluation reward: 1.53
episode: 398 score: 0.0 memory length: 73435 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.51
episode: 399 score: 2.0 memory length: 73655 epsilon: 1.0 steps: 220 lr: 0.0001 evaluation reward: 1.53
episode: 400 score: 2.0 memory length: 73873 epsilon: 1.0 steps: 218 lr: 0.0001 evaluation reward: 1.53
episode: 401 score: 3.0 memory length: 74144 epsilon: 1.0 steps: 271 lr: 0.0001 evaluation reward: 1.51
episode: 402 score: 0.0 memory length: 74268 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.51
episode: 403 score: 3.0 memory length: 74533 epsilon: 1.0 steps: 265 lr: 0.0001 evaluation reward: 1.53
episode: 404 score: 0.0 memory length: 74656 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.53
episode: 405 score: 2.0 memory length: 74857 epsilon: 1.0 steps: 201 lr: 0.0001 evaluation reward: 1.55
episode: 406 score: 1.0 memory length: 75028 epsilon: 1.0 steps: 171 lr: 0.0001 evaluation reward: 1.56
episode: 407 score: 0.0 memory length: 75152 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.56
episode: 408 score: 1.0 memory length: 75321 epsilon: 1.0 steps: 169 lr: 0.0001 evaluation reward: 1.57
episode: 409 score: 2.0 memory length: 75502 epsilon: 1.0 steps: 181 lr: 0.0001 evaluation reward: 1.59
episode: 410 score: 1.0 memory length: 75671 epsilon: 1.0 steps: 169 lr: 0.0001 evaluation reward: 1.55
episode: 411 score: 0.0 memory length: 75795 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.5
episode: 412 score: 2.0 memory length: 75975 epsilon: 1.0 steps: 180 lr: 0.0001 evaluation reward: 1.52
episode: 413 score: 0.0 memory length: 76099 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.52
episode: 414 score: 0.0 memory length: 76223 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.5
episode: 415 score: 4.0 memory length: 76481 epsilon: 1.0 steps: 258 lr: 0.0001 evaluation reward: 1.54
episode: 416 score: 5.0 memory length: 76806 epsilon: 1.0 steps: 325 lr: 0.0001 evaluation reward: 1.56
episode: 417 score: 1.0 memory length: 76958 epsilon: 1.0 steps: 152 lr: 0.0001 evaluation reward: 1.57
episode: 418 score: 2.0 memory length: 77156 epsilon: 1.0 steps: 198 lr: 0.0001 evaluation reward: 1.54
episode: 419 score: 4.0 memory length: 77432 epsilon: 1.0 steps: 276 lr: 0.0001 evaluation reward: 1.57

episode	score	memory length	epsilon	steps	lr:
420	0.0	77556	1.0	124	0.0001
421	2.0	77755	1.0	199	0.0001
422	7.0	78158	1.0	403	0.0001
423	0.0	78282	1.0	124	0.0001
424	3.0	78529	1.0	247	0.0001
425	0.0	78653	1.0	124	0.0001
426	0.0	78776	1.0	123	0.0001
427	3.0	79022	1.0	246	0.0001
428	2.0	79221	1.0	199	0.0001
429	0.0	79345	1.0	124	0.0001
430	3.0	79590	1.0	245	0.0001
431	0.0	79714	1.0	124	0.0001
432	3.0	79963	1.0	249	0.0001
433	2.0	80163	1.0	200	0.0001
434	2.0	80361	1.0	198	0.0001
435	1.0	80512	1.0	151	0.0001
436	1.0	80664	1.0	152	0.0001
437	0.0	80788	1.0	124	0.0001
438	1.0	80960	1.0	172	0.0001
439	1.0	81111	1.0	151	0.0001
440	0.0	81235	1.0	124	0.0001
441	2.0	81457	1.0	222	0.0001
442	1.0	81627	1.0	170	0.0001
443	0.0	81750	1.0	123	0.0001
444	0.0	81873	1.0	123	0.0001
445	2.0	82090	1.0	217	0.0001
446	1.0	82260	1.0	170	0.0001
447	0.0	82383	1.0	123	0.0001
448	2.0	82584	1.0	201	0.0001
449	4.0	82862	1.0	278	0.0001

episode	score	memory length	epsilon	steps	lr:
450	1.0	83014	1.0	152	0.0001
451	1.0	83183	1.0	169	0.0001
452	0.0	83306	1.0	123	0.0001
453	1.0	83475	1.0	169	0.0001
454	1.0	83645	1.0	170	0.0001
455	3.0	83872	1.0	227	0.0001
456	1.0	84024	1.0	152	0.0001
457	1.0	84194	1.0	170	0.0001
458	3.0	84420	1.0	226	0.0001
459	2.0	84619	1.0	199	0.0001
460	4.0	84938	1.0	319	0.0001
461	0.0	85062	1.0	124	0.0001
462	1.0	85214	1.0	152	0.0001
463	2.0	85413	1.0	199	0.0001
464	2.0	85612	1.0	199	0.0001
465	6.0	85985	1.0	373	0.0001
466	2.0	86202	1.0	217	0.0001
467	3.0	86431	1.0	229	0.0001
468	1.0	86582	1.0	151	0.0001
469	5.0	86919	1.0	337	0.0001
470	2.0	87118	1.0	199	0.0001
471	0.0	87242	1.0	124	0.0001
472	0.0	87366	1.0	124	0.0001
473	2.0	87582	1.0	216	0.0001
474	0.0	87705	1.0	123	0.0001
475	1.0	87875	1.0	170	0.0001
476	0.0	87998	1.0	123	0.0001
477	1.0	88150	1.0	152	0.0001
478	0.0	88274	1.0	124	0.0001
479	2.0	88457	1.0	183	0.0001

episode: 480	score: 0.0	memory length: 88581	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.55					
episode: 481	score: 0.0	memory length: 88705	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.53					
episode: 482	score: 1.0	memory length: 88875	epsilon: 1.0	steps: 170	lr: 0.0001
evaluation reward: 1.54					
episode: 483	score: 0.0	memory length: 88999	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.5					
episode: 484	score: 2.0	memory length: 89200	epsilon: 1.0	steps: 201	lr: 0.0001
evaluation reward: 1.52					
episode: 485	score: 0.0	memory length: 89324	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.51					
episode: 486	score: 2.0	memory length: 89523	epsilon: 1.0	steps: 199	lr: 0.0001
evaluation reward: 1.5					
episode: 487	score: 1.0	memory length: 89693	epsilon: 1.0	steps: 170	lr: 0.0001
evaluation reward: 1.49					
episode: 488	score: 4.0	memory length: 90011	epsilon: 1.0	steps: 318	lr: 0.0001
evaluation reward: 1.52					
episode: 489	score: 3.0	memory length: 90257	epsilon: 1.0	steps: 246	lr: 0.0001
evaluation reward: 1.54					
episode: 490	score: 1.0	memory length: 90427	epsilon: 1.0	steps: 170	lr: 0.0001
evaluation reward: 1.52					
episode: 491	score: 0.0	memory length: 90551	epsilon: 1.0	steps: 124	lr: 0.0001
evaluation reward: 1.51					
episode: 492	score: 2.0	memory length: 90750	epsilon: 1.0	steps: 199	lr: 0.0001
evaluation reward: 1.51					
episode: 493	score: 1.0	memory length: 90903	epsilon: 1.0	steps: 153	lr: 0.0001
evaluation reward: 1.52					
episode: 494	score: 3.0	memory length: 91150	epsilon: 1.0	steps: 247	lr: 0.0001
evaluation reward: 1.51					
episode: 495	score: 2.0	memory length: 91369	epsilon: 1.0	steps: 219	lr: 0.0001
evaluation reward: 1.53					
episode: 496	score: 2.0	memory length: 91567	epsilon: 1.0	steps: 198	lr: 0.0001
evaluation reward: 1.52					
episode: 497	score: 1.0	memory length: 91736	epsilon: 1.0	steps: 169	lr: 0.0001
evaluation reward: 1.51					
episode: 498	score: 1.0	memory length: 91887	epsilon: 1.0	steps: 151	lr: 0.0001
evaluation reward: 1.52					
episode: 499	score: 3.0	memory length: 92133	epsilon: 1.0	steps: 246	lr: 0.0001
evaluation reward: 1.53					
episode: 500	score: 2.0	memory length: 92331	epsilon: 1.0	steps: 198	lr: 0.0001
evaluation reward: 1.53					
episode: 501	score: 0.0	memory length: 92454	epsilon: 1.0	steps: 123	lr: 0.0001
evaluation reward: 1.5					
episode: 502	score: 2.0	memory length: 92653	epsilon: 1.0	steps: 199	lr: 0.0001
evaluation reward: 1.52					
episode: 503	score: 2.0	memory length: 92872	epsilon: 1.0	steps: 219	lr: 0.0001
evaluation reward: 1.51					
episode: 504	score: 1.0	memory length: 93023	epsilon: 1.0	steps: 151	lr: 0.0001
evaluation reward: 1.52					
episode: 505	score: 2.0	memory length: 93221	epsilon: 1.0	steps: 198	lr: 0.0001
evaluation reward: 1.52					
episode: 506	score: 3.0	memory length: 93448	epsilon: 1.0	steps: 227	lr: 0.0001
evaluation reward: 1.54					
episode: 507	score: 1.0	memory length: 93618	epsilon: 1.0	steps: 170	lr: 0.0001
evaluation reward: 1.55					
episode: 508	score: 1.0	memory length: 93789	epsilon: 1.0	steps: 171	lr: 0.0001
evaluation reward: 1.55					
episode: 509	score: 4.0	memory length: 94084	epsilon: 1.0	steps: 295	lr: 0.0001
evaluation reward: 1.57					

episode: 510 score: 0.0 memory length: 94207 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.56
episode: 511 score: 1.0 memory length: 94376 epsilon: 1.0 steps: 169 lr: 0.0001 evaluation reward: 1.57
episode: 512 score: 1.0 memory length: 94529 epsilon: 1.0 steps: 153 lr: 0.0001 evaluation reward: 1.56
episode: 513 score: 3.0 memory length: 94773 epsilon: 1.0 steps: 244 lr: 0.0001 evaluation reward: 1.59
episode: 514 score: 3.0 memory length: 95041 epsilon: 1.0 steps: 268 lr: 0.0001 evaluation reward: 1.62
episode: 515 score: 0.0 memory length: 95165 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.58
episode: 516 score: 3.0 memory length: 95430 epsilon: 1.0 steps: 265 lr: 0.0001 evaluation reward: 1.56
episode: 517 score: 3.0 memory length: 95642 epsilon: 1.0 steps: 212 lr: 0.0001 evaluation reward: 1.58
episode: 518 score: 1.0 memory length: 95794 epsilon: 1.0 steps: 152 lr: 0.0001 evaluation reward: 1.57
episode: 519 score: 0.0 memory length: 95918 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.53
episode: 520 score: 1.0 memory length: 96069 epsilon: 1.0 steps: 151 lr: 0.0001 evaluation reward: 1.54
episode: 521 score: 2.0 memory length: 96268 epsilon: 1.0 steps: 199 lr: 0.0001 evaluation reward: 1.54
episode: 522 score: 1.0 memory length: 96438 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.48
episode: 523 score: 3.0 memory length: 96652 epsilon: 1.0 steps: 214 lr: 0.0001 evaluation reward: 1.51
episode: 524 score: 1.0 memory length: 96823 epsilon: 1.0 steps: 171 lr: 0.0001 evaluation reward: 1.49
episode: 525 score: 5.0 memory length: 97140 epsilon: 1.0 steps: 317 lr: 0.0001 evaluation reward: 1.54
episode: 526 score: 3.0 memory length: 97367 epsilon: 1.0 steps: 227 lr: 0.0001 evaluation reward: 1.57
episode: 527 score: 1.0 memory length: 97537 epsilon: 1.0 steps: 170 lr: 0.0001 evaluation reward: 1.55
episode: 528 score: 0.0 memory length: 97661 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.53
episode: 529 score: 1.0 memory length: 97812 epsilon: 1.0 steps: 151 lr: 0.0001 evaluation reward: 1.54
episode: 530 score: 0.0 memory length: 97935 epsilon: 1.0 steps: 123 lr: 0.0001 evaluation reward: 1.51
episode: 531 score: 1.0 memory length: 98087 epsilon: 1.0 steps: 152 lr: 0.0001 evaluation reward: 1.52
episode: 532 score: 0.0 memory length: 98211 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.49
episode: 533 score: 0.0 memory length: 98335 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.47
episode: 534 score: 0.0 memory length: 98459 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.45
episode: 535 score: 2.0 memory length: 98677 epsilon: 1.0 steps: 218 lr: 0.0001 evaluation reward: 1.46
episode: 536 score: 1.0 memory length: 98849 epsilon: 1.0 steps: 172 lr: 0.0001 evaluation reward: 1.46
episode: 537 score: 0.0 memory length: 98973 epsilon: 1.0 steps: 124 lr: 0.0001 evaluation reward: 1.46
episode: 538 score: 2.0 memory length: 99171 epsilon: 1.0 steps: 198 lr: 0.0001 evaluation reward: 1.47
episode: 539 score: 1.0 memory length: 99322 epsilon: 1.0 steps: 151 lr: 0.0001 evaluation reward: 1.47

```
episode: 540    score: 0.0    memory length: 99446    epsilon: 1.0    steps: 124    lr: 0.0001    evaluation reward: 1.47
episode: 541    score: 1.0    memory length: 99616    epsilon: 1.0    steps: 170    lr: 0.0001    evaluation reward: 1.46
episode: 542    score: 2.0    memory length: 99818    epsilon: 1.0    steps: 202    lr: 0.0001    evaluation reward: 1.47
episode: 543    score: 0.0    memory length: 99942    epsilon: 1.0    steps: 124    lr: 0.0001    evaluation reward: 1.47
/home/zacharytgeorge/assignment5_materials/assignment5_materials/memory.py:29: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray
    sample = np.array(sample)
/home/zacharytgeorge/assignment5_materials/assignment5_materials/agent_double.py:73: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray
    mini_batch = np.array(mini_batch).transpose()
```

episode: 544	score: 2.0	memory length: 100164	epsilon: 0.9996733000000071	st
eps: 222	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.9994277800000124	st
episode: 545	score: 0.0	memory length: 100288	epsilon: 0.998895160000024	ste
eps: 124	lr: 0.0001	evaluation reward: 1.47	epsilon: 0.9985605400000312	st
episode: 546	score: 4.0	memory length: 100557	epsilon: 0.9981071200000411	st
ps: 269	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9978061600000476	st
episode: 547	score: 1.0	memory length: 100726	epsilon: 0.997560640000053	ste
eps: 169	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.9971171200000626	st
episode: 548	score: 3.0	memory length: 100955	epsilon: 0.9965290600000754	st
eps: 229	lr: 0.0001	evaluation reward: 1.52	epsilon: 0.9962835400000807	st
episode: 549	score: 1.0	memory length: 101107	epsilon: 0.9958499200000901	st
eps: 152	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.9956044000000954	st
episode: 550	score: 0.0	memory length: 101231	epsilon: 0.995208400000104	ste
ps: 124	lr: 0.0001	evaluation reward: 1.48	epsilon: 0.994196620000126	ste
episode: 551	score: 2.0	memory length: 101455	epsilon: 0.9938956600001325	st
eps: 224	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.993367000000144	ste
episode: 552	score: 4.0	memory length: 101752	epsilon: 0.9928799200001546	st
eps: 297	lr: 0.0001	evaluation reward: 1.53	epsilon: 0.9926344000001599	st
episode: 553	score: 0.0	memory length: 101876	epsilon: 0.9923908600001652	st
eps: 124	lr: 0.0001	evaluation reward: 1.52	epsilon: 0.9921453400001705	st
episode: 554	score: 2.0	memory length: 102095	epsilon: 0.9917156800001798	st
eps: 219	lr: 0.0001	evaluation reward: 1.53	epsilon: 0.9913790800001872	st
episode: 555	score: 0.0	memory length: 102219	epsilon: 0.9911335600001925	st
eps: 124	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9907058800002018	st
episode: 556	score: 2.0	memory length: 102419	epsilon: 0.9900425800002162	st
ps: 200	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.9896069800002256	st
episode: 557	score: 8.0	memory length: 102930	epsilon: 0.989361460000231	ste
ps: 511	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.9890209000002383	st
episode: 558	score: 1.0	memory length: 103082	epsilon: 0.9886862800002456	st
eps: 152	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.9882843400002543	st
episode: 559	score: 4.0	memory length: 103349		
ps: 267	lr: 0.0001	evaluation reward: 1.58		
episode: 560	score: 3.0	memory length: 103595		
eps: 246	lr: 0.0001	evaluation reward: 1.57		
episode: 561	score: 0.0	memory length: 103719		
eps: 124	lr: 0.0001	evaluation reward: 1.57		
episode: 562	score: 0.0	memory length: 103842		
eps: 123	lr: 0.0001	evaluation reward: 1.56		
episode: 563	score: 0.0	memory length: 103966		
eps: 124	lr: 0.0001	evaluation reward: 1.54		
episode: 564	score: 2.0	memory length: 104183		
eps: 217	lr: 0.0001	evaluation reward: 1.54		
episode: 565	score: 1.0	memory length: 104353		
eps: 170	lr: 0.0001	evaluation reward: 1.49		
episode: 566	score: 0.0	memory length: 104477		
eps: 124	lr: 0.0001	evaluation reward: 1.47		
episode: 567	score: 2.0	memory length: 104693		
eps: 216	lr: 0.0001	evaluation reward: 1.46		
episode: 568	score: 5.0	memory length: 105028		
eps: 335	lr: 0.0001	evaluation reward: 1.5		
episode: 569	score: 2.0	memory length: 105248		
eps: 220	lr: 0.0001	evaluation reward: 1.47		
episode: 570	score: 0.0	memory length: 105372		
ps: 124	lr: 0.0001	evaluation reward: 1.45		
episode: 571	score: 1.0	memory length: 105544		
eps: 172	lr: 0.0001	evaluation reward: 1.46		
episode: 572	score: 1.0	memory length: 105713		
eps: 169	lr: 0.0001	evaluation reward: 1.47		
episode: 573	score: 2.0	memory length: 105916		
eps: 203	lr: 0.0001	evaluation reward: 1.47		

episode: 574	score: 0.0	memory length: 106040	epsilon: 0.9880388200002597	st
eps: 124	lr: 0.0001	evaluation reward: 1.47	epsilon: 0.9876804400002674	st
episode: 575	score: 2.0	memory length: 106221	epsilon: 0.9873458200002747	st
eps: 181	lr: 0.0001	evaluation reward: 1.48	epsilon: 0.98710030000028	step
episode: 576	score: 1.0	memory length: 106390	epsilon: 0.9866666800002895	st
eps: 169	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.986274640000298	ste
episode: 577	score: 0.0	memory length: 106514	epsilon: 0.9860291200003033	st
s: 124	lr: 0.0001	evaluation reward: 1.48	epsilon: 0.985579660000313	ste
episode: 578	score: 2.0	memory length: 106733	epsilon: 0.9852806800003195	st
eps: 219	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.984847060000329	ste
episode: 579	score: 2.0	memory length: 106931	epsilon: 0.9844174000003383	st
ps: 198	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9838729000003501	st
episode: 580	score: 0.0	memory length: 107055	epsilon: 0.9836293600003554	st
eps: 124	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9832927600003627	st
episode: 581	score: 3.0	memory length: 107282	epsilon: 0.98295616000037	step
ps: 227	lr: 0.0001	evaluation reward: 1.53	epsilon: 0.9825284800003793	st
episode: 582	score: 1.0	memory length: 107433	epsilon: 0.9821918800003866	st
eps: 151	lr: 0.0001	evaluation reward: 1.53	epsilon: 0.9817978600003952	st
episode: 583	score: 2.0	memory length: 107652	epsilon: 0.9813642400004046	st
ps: 219	lr: 0.0001	evaluation reward: 1.55	epsilon: 0.9810632800004111	st
episode: 584	score: 2.0	memory length: 107869	epsilon: 0.9808177600004164	st
eps: 217	lr: 0.0001	evaluation reward: 1.55	epsilon: 0.9804257200004249	st
episode: 585	score: 1.0	memory length: 108020	epsilon: 0.9801247600004315	st
eps: 151	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.9798792400004368	st
episode: 586	score: 0.0	memory length: 108144	epsilon: 0.9795446200004441	st
eps: 124	lr: 0.0001	evaluation reward: 1.54	epsilon: 0.9790615000004546	st
episode: 587	score: 0.0	memory length: 108267	epsilon: 0.9786674800004631	st
eps: 123	lr: 0.0001	evaluation reward: 1.53	epsilon: 0.9782734600004717	st
episode: 588	score: 1.0	memory length: 108437	epsilon: 0.978027940000477	ste
eps: 170	lr: 0.0001	evaluation reward: 1.5		
episode: 589	score: 1.0	memory length: 108607		
s: 170	lr: 0.0001	evaluation reward: 1.48		
episode: 590	score: 2.0	memory length: 108823		
eps: 216	lr: 0.0001	evaluation reward: 1.49		
episode: 591	score: 1.0	memory length: 108993		
eps: 170	lr: 0.0001	evaluation reward: 1.5		
episode: 592	score: 2.0	memory length: 109192		
eps: 199	lr: 0.0001	evaluation reward: 1.5		
episode: 593	score: 2.0	memory length: 109411		
eps: 219	lr: 0.0001	evaluation reward: 1.51		
episode: 594	score: 1.0	memory length: 109563		
eps: 152	lr: 0.0001	evaluation reward: 1.49		
episode: 595	score: 0.0	memory length: 109687		
eps: 124	lr: 0.0001	evaluation reward: 1.47		
episode: 596	score: 2.0	memory length: 109885		
eps: 198	lr: 0.0001	evaluation reward: 1.47		
episode: 597	score: 1.0	memory length: 110037		
eps: 152	lr: 0.0001	evaluation reward: 1.47		
episode: 598	score: 0.0	memory length: 110161		
eps: 124	lr: 0.0001	evaluation reward: 1.46		
episode: 599	score: 1.0	memory length: 110330		
eps: 169	lr: 0.0001	evaluation reward: 1.44		
episode: 600	score: 3.0	memory length: 110574		
eps: 244	lr: 0.0001	evaluation reward: 1.45		
episode: 601	score: 2.0	memory length: 110773		
eps: 199	lr: 0.0001	evaluation reward: 1.47		
episode: 602	score: 2.0	memory length: 110972		
eps: 199	lr: 0.0001	evaluation reward: 1.47		
episode: 603	score: 0.0	memory length: 111096		
ps: 124	lr: 0.0001	evaluation reward: 1.45		

episode: 604	score: 1.0	memory length: 111248	epsilon: 0.9777269800004835	st
eps: 152	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9773329600004921	st
episode: 605	score: 2.0	memory length: 111447	epsilon: 0.9768102400005034	st
eps: 199	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9764716600005108	st
episode: 606	score: 3.0	memory length: 111711	epsilon: 0.9761707000005173	st
eps: 264	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9758360800005246	st
episode: 607	score: 1.0	memory length: 111882	epsilon: 0.9754420600005331	st
eps: 171	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9751965400005385	st
episode: 608	score: 1.0	memory length: 112034	epsilon: 0.9749530000005437	st
eps: 152	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9747074800005491	st
episode: 609	score: 1.0	memory length: 112203	epsilon: 0.9744619600005544	st
eps: 169	lr: 0.0001	evaluation reward: 1.42	epsilon: 0.974067940000563	ste
episode: 610	score: 2.0	memory length: 112402	epsilon: 0.9738224200005683	st
eps: 199	lr: 0.0001	evaluation reward: 1.44	epsilon: 0.9732719800005802	st
episode: 611	score: 0.0	memory length: 112526	epsilon: 0.9730264600005856	st
eps: 124	lr: 0.0001	evaluation reward: 1.43	epsilon: 0.9726918400005928	st
episode: 612	score: 0.0	memory length: 112649	epsilon: 0.9722027800006035	st
eps: 123	lr: 0.0001	evaluation reward: 1.42	epsilon: 0.9718602400006109	st
episode: 613	score: 0.0	memory length: 112773	epsilon: 0.9716167000006162	st
eps: 124	lr: 0.0001	evaluation reward: 1.39	epsilon: 0.9711771400006257	st
episode: 614	score: 0.0	memory length: 112897	epsilon: 0.9707237200006356	st
eps: 124	lr: 0.0001	evaluation reward: 1.36	epsilon: 0.9703891000006428	st
episode: 615	score: 2.0	memory length: 113096	epsilon: 0.9701455600006481	st
ps: 199	lr: 0.0001	evaluation reward: 1.38	epsilon: 0.9699000400006534	st
episode: 616	score: 0.0	memory length: 113220	epsilon: 0.969506020000662	ste
eps: 124	lr: 0.0001	evaluation reward: 1.35	epsilon: 0.9692605000006673	st
episode: 617	score: 4.0	memory length: 113498	epsilon: 0.9688070800006772	st
eps: 278	lr: 0.0001	evaluation reward: 1.36	epsilon: 0.9684150400006857	st
episode: 618	score: 0.0	memory length: 113622	epsilon: 0.9680170600006943	st
eps: 124	lr: 0.0001	evaluation reward: 1.35	epsilon: 0.9675834400007037	st
episode: 619	score: 1.0	memory length: 113791		
eps: 169	lr: 0.0001	evaluation reward: 1.36		
episode: 620	score: 3.0	memory length: 114038		
eps: 247	lr: 0.0001	evaluation reward: 1.38		
episode: 621	score: 1.0	memory length: 114211		
eps: 173	lr: 0.0001	evaluation reward: 1.37		
episode: 622	score: 0.0	memory length: 114334		
eps: 123	lr: 0.0001	evaluation reward: 1.36		
episode: 623	score: 2.0	memory length: 114556		
eps: 222	lr: 0.0001	evaluation reward: 1.35		
episode: 624	score: 3.0	memory length: 114785		
eps: 229	lr: 0.0001	evaluation reward: 1.37		
episode: 625	score: 1.0	memory length: 114954		
eps: 169	lr: 0.0001	evaluation reward: 1.33		
episode: 626	score: 0.0	memory length: 115077		
eps: 123	lr: 0.0001	evaluation reward: 1.3		
episode: 627	score: 0.0	memory length: 115201		
eps: 124	lr: 0.0001	evaluation reward: 1.29		
episode: 628	score: 2.0	memory length: 115400		
ps: 199	lr: 0.0001	evaluation reward: 1.31		
episode: 629	score: 0.0	memory length: 115524		
eps: 124	lr: 0.0001	evaluation reward: 1.3		
episode: 630	score: 3.0	memory length: 115753		
eps: 229	lr: 0.0001	evaluation reward: 1.33		
episode: 631	score: 2.0	memory length: 115951		
eps: 198	lr: 0.0001	evaluation reward: 1.34		
episode: 632	score: 2.0	memory length: 116152		
eps: 201	lr: 0.0001	evaluation reward: 1.36		
episode: 633	score: 2.0	memory length: 116371		
eps: 219	lr: 0.0001	evaluation reward: 1.38		

episode: 634	score: 3.0	memory length: 116599	epsilon: 0.9671320000007135	st
eps: 228	lr: 0.0001	evaluation reward: 1.41	epsilon: 0.9667973800007208	st
episode: 635	score: 1.0	memory length: 116768	epsilon: 0.9663023800007315	st
eps: 169	lr: 0.0001	evaluation reward: 1.4	epsilon: 0.965864800000741	ste
episode: 636	score: 3.0	memory length: 117018	epsilon: 0.9655222600007485	st
eps: 250	lr: 0.0001	evaluation reward: 1.42	epsilon: 0.965221300000755	ste
episode: 637	score: 2.0	memory length: 117239	epsilon: 0.9649223200007615	st
ps: 221	lr: 0.0001	evaluation reward: 1.44	epsilon: 0.9644312800007722	st
episode: 638	score: 1.0	memory length: 117412	epsilon: 0.9639026200007836	st
eps: 173	lr: 0.0001	evaluation reward: 1.43	epsilon: 0.9635105800007921	st
episode: 639	score: 1.0	memory length: 117564	epsilon: 0.9632650600007975	st
ps: 152	lr: 0.0001	evaluation reward: 1.43	epsilon: 0.9630195400008028	st
episode: 640	score: 1.0	memory length: 117715	epsilon: 0.9625898800008121	st
eps: 151	lr: 0.0001	evaluation reward: 1.44	epsilon: 0.9619127200008268	st
episode: 641	score: 3.0	memory length: 117963	epsilon: 0.96130882000084	step
eps: 248	lr: 0.0001	evaluation reward: 1.46	epsilon: 0.9609148000008485	st
episode: 642	score: 3.0	memory length: 118230	epsilon: 0.9605782000008558	st
eps: 267	lr: 0.0001	evaluation reward: 1.47	epsilon: 0.9603346600008611	st
episode: 643	score: 2.0	memory length: 118428	epsilon: 0.9597466000008739	st
eps: 198	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.959188240000886	ste
episode: 644	score: 0.0	memory length: 118552	epsilon: 0.9589427200008913	st
eps: 124	lr: 0.0001	evaluation reward: 1.47	epsilon: 0.9583922800009033	st
episode: 645	score: 0.0	memory length: 118676	epsilon: 0.9578438200009152	st
eps: 124	lr: 0.0001	evaluation reward: 1.47	epsilon: 0.9574062400009247	st
episode: 646	score: 2.0	memory length: 118893	epsilon: 0.9571052800009312	st
eps: 217	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9568597600009365	st
episode: 647	score: 5.0	memory length: 119235	epsilon: 0.9564202000009461	st
eps: 342	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.9561746800009514	st
episode: 648	score: 5.0	memory length: 119540	epsilon: 0.9558737200009579	st
s: 305	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.9554203000009678	st
episode: 649	score: 2.0	memory length: 119739		
eps: 199	lr: 0.0001	evaluation reward: 1.52		
episode: 650	score: 1.0	memory length: 119909		
eps: 170	lr: 0.0001	evaluation reward: 1.53		
episode: 651	score: 0.0	memory length: 120032		
eps: 123	lr: 0.0001	evaluation reward: 1.51		
episode: 652	score: 4.0	memory length: 120329		
eps: 297	lr: 0.0001	evaluation reward: 1.51		
episode: 653	score: 4.0	memory length: 120611		
ps: 282	lr: 0.0001	evaluation reward: 1.55		
episode: 654	score: 0.0	memory length: 120735		
eps: 124	lr: 0.0001	evaluation reward: 1.53		
episode: 655	score: 4.0	memory length: 121013		
eps: 278	lr: 0.0001	evaluation reward: 1.57		
episode: 656	score: 4.0	memory length: 121290		
eps: 277	lr: 0.0001	evaluation reward: 1.59		
episode: 657	score: 2.0	memory length: 121511		
eps: 221	lr: 0.0001	evaluation reward: 1.53		
episode: 658	score: 1.0	memory length: 121663		
eps: 152	lr: 0.0001	evaluation reward: 1.53		
episode: 659	score: 0.0	memory length: 121787		
eps: 124	lr: 0.0001	evaluation reward: 1.49		
episode: 660	score: 2.0	memory length: 122009		
eps: 222	lr: 0.0001	evaluation reward: 1.48		
episode: 661	score: 0.0	memory length: 122133		
eps: 124	lr: 0.0001	evaluation reward: 1.48		
episode: 662	score: 1.0	memory length: 122285		
eps: 152	lr: 0.0001	evaluation reward: 1.49		
episode: 663	score: 3.0	memory length: 122514		
eps: 229	lr: 0.0001	evaluation reward: 1.52		

episode: 664	score: 6.0	memory length: 122765	epsilon: 0.9549233200009786	st
eps: 251	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.9543847600009903	st
episode: 665	score: 3.0	memory length: 123037	epsilon: 0.9541392400009956	st
eps: 272	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.9538937200010009	st
episode: 666	score: 0.0	memory length: 123161	epsilon: 0.9534640600010102	st
eps: 124	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.952877980001023	ste
episode: 667	score: 0.0	memory length: 123285	epsilon: 0.9525354400010304	st
eps: 124	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.9521988400010377	st
episode: 668	score: 2.0	memory length: 123502	epsilon: 0.95195530001043	ste
eps: 217	lr: 0.0001	evaluation reward: 1.53	epsilon: 0.9515216800010524	st
episode: 669	score: 4.0	memory length: 123798	epsilon: 0.9512761600010577	st
ps: 296	lr: 0.0001	evaluation reward: 1.55	epsilon: 0.9507494800010692	st
episode: 670	score: 1.0	memory length: 123971	epsilon: 0.9504485200010757	st
eps: 173	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.950204980001081	ste
episode: 671	score: 1.0	memory length: 124141	epsilon: 0.9499594600010863	st
eps: 170	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.9496208800010937	st
episode: 672	score: 0.0	memory length: 124264	epsilon: 0.9490803400011054	st
ps: 123	lr: 0.0001	evaluation reward: 1.55	epsilon: 0.9486289000011152	st
episode: 673	score: 2.0	memory length: 124483	epsilon: 0.948272500001123	ste
eps: 219	lr: 0.0001	evaluation reward: 1.55	epsilon: 0.9480269800011283	st
episode: 674	score: 0.0	memory length: 124607	epsilon: 0.9477834400011336	st
eps: 124	lr: 0.0001	evaluation reward: 1.55	epsilon: 0.9474211000011414	st
episode: 675	score: 3.0	memory length: 124873	epsilon: 0.94702906000115	step
eps: 266	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.94656574000116	step
episode: 676	score: 1.0	memory length: 125025	epsilon: 0.9460727200011707	st
eps: 152	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.945736120001178	ste
episode: 677	score: 0.0	memory length: 125148	epsilon: 0.9454906000011833	st
ps: 123	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.9451916200011898	st
episode: 678	score: 0.0	memory length: 125272	epsilon: 0.9449461000011952	st
eps: 124	lr: 0.0001	evaluation reward: 1.54	epsilon: 0.9443164600012088	st
episode: 679	score: 1.0	memory length: 125443		
eps: 171	lr: 0.0001	evaluation reward: 1.53		
episode: 680	score: 3.0	memory length: 125716		
eps: 273	lr: 0.0001	evaluation reward: 1.56		
episode: 681	score: 3.0	memory length: 125944		
eps: 228	lr: 0.0001	evaluation reward: 1.56		
episode: 682	score: 2.0	memory length: 126124		
ps: 180	lr: 0.0001	evaluation reward: 1.57		
episode: 683	score: 0.0	memory length: 126248		
eps: 124	lr: 0.0001	evaluation reward: 1.55		
episode: 684	score: 0.0	memory length: 126371		
eps: 123	lr: 0.0001	evaluation reward: 1.53		
episode: 685	score: 2.0	memory length: 126554		
eps: 183	lr: 0.0001	evaluation reward: 1.54		
episode: 686	score: 2.0	memory length: 126752		
s: 198	lr: 0.0001	evaluation reward: 1.56		
episode: 687	score: 3.0	memory length: 126986		
s: 234	lr: 0.0001	evaluation reward: 1.59		
episode: 688	score: 3.0	memory length: 127235		
eps: 249	lr: 0.0001	evaluation reward: 1.61		
episode: 689	score: 1.0	memory length: 127405		
ps: 170	lr: 0.0001	evaluation reward: 1.61		
episode: 690	score: 0.0	memory length: 127529		
eps: 124	lr: 0.0001	evaluation reward: 1.59		
episode: 691	score: 1.0	memory length: 127680		
eps: 151	lr: 0.0001	evaluation reward: 1.59		
episode: 692	score: 0.0	memory length: 127804		
eps: 124	lr: 0.0001	evaluation reward: 1.57		
episode: 693	score: 4.0	memory length: 128122		
eps: 318	lr: 0.0001	evaluation reward: 1.59		

episode: 694	score: 2.0	memory length: 128302	epsilon: 0.9439600600012166	st
eps: 180	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.9436591000012231	st
episode: 695	score: 1.0	memory length: 128454	epsilon: 0.9434135800012284	st
eps: 152	lr: 0.0001	evaluation reward: 1.61	epsilon: 0.94287898000124	step
episode: 696	score: 0.0	memory length: 128578	epsilon: 0.9426334600012454	st
eps: 124	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.9423899200012507	st
episode: 697	score: 3.0	memory length: 128848	epsilon: 0.9419068000012611	st
s: 270	lr: 0.0001	evaluation reward: 1.61	epsilon: 0.9416632600012664	st
episode: 698	score: 0.0	memory length: 128972	epsilon: 0.9413068600012742	st
eps: 124	lr: 0.0001	evaluation reward: 1.61	epsilon: 0.9409663000012816	st
episode: 699	score: 0.0	memory length: 129095	epsilon: 0.940532680001291	ste
eps: 123	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.9401050000013003	st
episode: 700	score: 3.0	memory length: 129339	epsilon: 0.9396496000013101	st
eps: 244	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.9394040800013155	st
episode: 701	score: 0.0	memory length: 129462	epsilon: 0.939010060001324	ste
eps: 123	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.9384754600013356	st
episode: 702	score: 2.0	memory length: 129642	epsilon: 0.938229940001341	ste
eps: 180	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.9378715600013487	st
episode: 703	score: 1.0	memory length: 129814	epsilon: 0.9374795200013573	st
eps: 172	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.9371429200013646	st
episode: 704	score: 2.0	memory length: 130033	epsilon: 0.9366103000013761	st
ps: 219	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.9362182600013846	st
episode: 705	score: 2.0	memory length: 130249	epsilon: 0.9356856400013962	st
eps: 216	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.9354401200014015	st
episode: 706	score: 3.0	memory length: 130479	epsilon: 0.9351035200014088	st
eps: 230	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.9347114800014173	st
episode: 707	score: 0.0	memory length: 130603	epsilon: 0.934311520001426	ste
eps: 124	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.9338224600014366	st
episode: 708	score: 2.0	memory length: 130802	epsilon: 0.9334244800014453	st
ps: 199	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.9329908600014547	st
episode: 709	score: 3.0	memory length: 131072		
eps: 270	lr: 0.0001	evaluation reward: 1.62		
episode: 710	score: 0.0	memory length: 131196		
ps: 124	lr: 0.0001	evaluation reward: 1.6		
episode: 711	score: 2.0	memory length: 131377		
eps: 181	lr: 0.0001	evaluation reward: 1.62		
episode: 712	score: 2.0	memory length: 131575		
eps: 198	lr: 0.0001	evaluation reward: 1.64		
episode: 713	score: 1.0	memory length: 131745		
eps: 170	lr: 0.0001	evaluation reward: 1.65		
episode: 714	score: 3.0	memory length: 132014		
eps: 269	lr: 0.0001	evaluation reward: 1.68		
episode: 715	score: 2.0	memory length: 132212		
eps: 198	lr: 0.0001	evaluation reward: 1.68		
episode: 716	score: 3.0	memory length: 132481		
eps: 269	lr: 0.0001	evaluation reward: 1.71		
episode: 717	score: 0.0	memory length: 132605		
eps: 124	lr: 0.0001	evaluation reward: 1.67		
episode: 718	score: 1.0	memory length: 132775		
eps: 170	lr: 0.0001	evaluation reward: 1.68		
episode: 719	score: 2.0	memory length: 132973		
eps: 198	lr: 0.0001	evaluation reward: 1.69		
episode: 720	score: 2.0	memory length: 133175		
ps: 202	lr: 0.0001	evaluation reward: 1.68		
episode: 721	score: 3.0	memory length: 133422		
eps: 247	lr: 0.0001	evaluation reward: 1.7		
episode: 722	score: 2.0	memory length: 133623		
eps: 201	lr: 0.0001	evaluation reward: 1.72		
episode: 723	score: 2.0	memory length: 133842		
eps: 219	lr: 0.0001	evaluation reward: 1.72		

episode: 724	score: 3.0	memory length: 134093	epsilon: 0.9324938800014655	st
eps: 251	lr: 0.0001	evaluation reward: 1.72	epsilon: 0.9322483600014708	st
episode: 725	score: 0.0	memory length: 134217	epsilon: 0.9317216800014823	st
eps: 124	lr: 0.0001	evaluation reward: 1.71	epsilon: 0.9313831000014896	st
episode: 726	score: 3.0	memory length: 134483	epsilon: 0.9310821400014961	st
eps: 266	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.9307415800015035	st
episode: 727	score: 1.0	memory length: 134654	epsilon: 0.9303475600015121	st
eps: 171	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.9300070000015195	st
episode: 728	score: 1.0	memory length: 134806	epsilon: 0.9295139800015302	st
eps: 152	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.9292704400015355	st
episode: 729	score: 1.0	memory length: 134978	epsilon: 0.9289318600015428	st
eps: 172	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.9285378400015514	st
episode: 730	score: 2.0	memory length: 135177	epsilon: 0.9281438200015599	st
eps: 199	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.9277498000015685	st
episode: 731	score: 1.0	memory length: 135349	epsilon: 0.9272983600015783	st
eps: 172	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.9269637400015855	st
episode: 732	score: 3.0	memory length: 135598	epsilon: 0.9264746800015962	st
eps: 249	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.9260410600016056	st
episode: 733	score: 0.0	memory length: 135721	epsilon: 0.9255480400016163	st
eps: 123	lr: 0.0001	evaluation reward: 1.72	epsilon: 0.9252470800016228	st
episode: 734	score: 1.0	memory length: 135892	epsilon: 0.9247501000016336	st
eps: 171	lr: 0.0001	evaluation reward: 1.7	epsilon: 0.9243560800016422	st
episode: 735	score: 2.0	memory length: 136091	epsilon: 0.9239620600016507	st
eps: 199	lr: 0.0001	evaluation reward: 1.71	epsilon: 0.9235284400016601	st
episode: 736	score: 2.0	memory length: 136290	epsilon: 0.9229443400016728	st
eps: 199	lr: 0.0001	evaluation reward: 1.7	epsilon: 0.9226453600016793	st
episode: 737	score: 2.0	memory length: 136489	epsilon: 0.9224018200016846	st
eps: 199	lr: 0.0001	evaluation reward: 1.7	epsilon: 0.9221028400016911	st
episode: 738	score: 3.0	memory length: 136717	epsilon: 0.9218573200016964	st
eps: 228	lr: 0.0001	evaluation reward: 1.72		
episode: 739	score: 1.0	memory length: 136886		
eps: 169	lr: 0.0001	evaluation reward: 1.72		
episode: 740	score: 3.0	memory length: 137133		
eps: 247	lr: 0.0001	evaluation reward: 1.74		
episode: 741	score: 2.0	memory length: 137352		
eps: 219	lr: 0.0001	evaluation reward: 1.73		
episode: 742	score: 3.0	memory length: 137601		
eps: 249	lr: 0.0001	evaluation reward: 1.73		
episode: 743	score: 1.0	memory length: 137753		
eps: 152	lr: 0.0001	evaluation reward: 1.72		
episode: 744	score: 3.0	memory length: 138004		
eps: 251	lr: 0.0001	evaluation reward: 1.75		
episode: 745	score: 2.0	memory length: 138203		
eps: 199	lr: 0.0001	evaluation reward: 1.77		
episode: 746	score: 2.0	memory length: 138402		
eps: 199	lr: 0.0001	evaluation reward: 1.77		
episode: 747	score: 2.0	memory length: 138621		
eps: 219	lr: 0.0001	evaluation reward: 1.74		
episode: 748	score: 0.0	memory length: 138744		
eps: 123	lr: 0.0001	evaluation reward: 1.69		
episode: 749	score: 1.0	memory length: 138916		
eps: 172	lr: 0.0001	evaluation reward: 1.68		
episode: 750	score: 1.0	memory length: 139067		
eps: 151	lr: 0.0001	evaluation reward: 1.68		
episode: 751	score: 0.0	memory length: 139190		
eps: 123	lr: 0.0001	evaluation reward: 1.68		
episode: 752	score: 1.0	memory length: 139341		
eps: 151	lr: 0.0001	evaluation reward: 1.65		
episode: 753	score: 0.0	memory length: 139465		
eps: 124	lr: 0.0001	evaluation reward: 1.61		

episode: 754	score: 3.0	memory length: 139735	epsilon: 0.921322720001708	ste
ps: 270	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.9207762400017199	st
episode: 755	score: 4.0	memory length: 140011	epsilon: 0.9203842000017284	st
eps: 276	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.9200456200017357	st
episode: 756	score: 2.0	memory length: 140209	epsilon: 0.919709020001743	ste
eps: 198	lr: 0.0001	evaluation reward: 1.62	epsilon: 0.9193169800017515	st
episode: 757	score: 1.0	memory length: 140380	epsilon: 0.919018000001758	ste
eps: 171	lr: 0.0001	evaluation reward: 1.61	epsilon: 0.9186200200017667	st
episode: 758	score: 1.0	memory length: 140550	epsilon: 0.9182794600017741	st
ps: 170	lr: 0.0001	evaluation reward: 1.61	epsilon: 0.9179210800017819	st
episode: 759	score: 2.0	memory length: 140748	epsilon: 0.9176221000017883	st
eps: 198	lr: 0.0001	evaluation reward: 1.63	epsilon: 0.9172855000017957	st
episode: 760	score: 1.0	memory length: 140899	epsilon: 0.9169291000018034	st
ps: 151	lr: 0.0001	evaluation reward: 1.62	epsilon: 0.9165687400018112	st
episode: 761	score: 2.0	memory length: 141100	epsilon: 0.9162658000018178	st
eps: 201	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.9160222600018231	st
episode: 762	score: 1.0	memory length: 141272	epsilon: 0.915564880001833	ste
eps: 172	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.9150778000018436	st
episode: 763	score: 2.0	memory length: 141453	epsilon: 0.914737240001851	ste
eps: 181	lr: 0.0001	evaluation reward: 1.63	epsilon: 0.9143768800018588	st
episode: 764	score: 1.0	memory length: 141604	epsilon: 0.9138502000018702	st
eps: 151	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.9134146000018797	st
episode: 765	score: 1.0	memory length: 141774	epsilon: 0.9126107200018971	st
eps: 170	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.9122127400019058	st
episode: 766	score: 2.0	memory length: 141954	epsilon: 0.9118187200019143	st
eps: 180	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.9113316400019249	st
episode: 767	score: 2.0	memory length: 142136	epsilon: 0.9110306800019314	st
eps: 182	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.9106901200019388	st
episode: 768	score: 1.0	memory length: 142289	epsilon: 0.9101971000019495	st
eps: 153	lr: 0.0001	evaluation reward: 1.59		
episode: 769	score: 0.0	memory length: 142412		
eps: 123	lr: 0.0001	evaluation reward: 1.55		
episode: 770	score: 3.0	memory length: 142643		
ps: 231	lr: 0.0001	evaluation reward: 1.57		
episode: 771	score: 3.0	memory length: 142889		
eps: 246	lr: 0.0001	evaluation reward: 1.59		
episode: 772	score: 1.0	memory length: 143061		
ps: 172	lr: 0.0001	evaluation reward: 1.6		
episode: 773	score: 2.0	memory length: 143243		
eps: 182	lr: 0.0001	evaluation reward: 1.6		
episode: 774	score: 3.0	memory length: 143509		
eps: 266	lr: 0.0001	evaluation reward: 1.63		
episode: 775	score: 2.0	memory length: 143729		
eps: 220	lr: 0.0001	evaluation reward: 1.62		
episode: 776	score: 0.0	memory length: 143853		
ps: 124	lr: 0.0001	evaluation reward: 1.61		
episode: 777	score: 5.0	memory length: 144135		
eps: 282	lr: 0.0001	evaluation reward: 1.66		
episode: 778	score: 2.0	memory length: 144336		
eps: 201	lr: 0.0001	evaluation reward: 1.68		
episode: 779	score: 2.0	memory length: 144535		
eps: 199	lr: 0.0001	evaluation reward: 1.69		
episode: 780	score: 3.0	memory length: 144781		
eps: 246	lr: 0.0001	evaluation reward: 1.69		
episode: 781	score: 1.0	memory length: 144933		
eps: 152	lr: 0.0001	evaluation reward: 1.67		
episode: 782	score: 1.0	memory length: 145105		
eps: 172	lr: 0.0001	evaluation reward: 1.66		
episode: 783	score: 3.0	memory length: 145354		
eps: 249	lr: 0.0001	evaluation reward: 1.69		

episode: 784	score: 4.0	memory length: 145614	epsilon: 0.909682300019607	st
eps: 260	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.9093813400019672	st
episode: 785	score: 1.0	memory length: 145766	epsilon: 0.9090803800019738	st
eps: 152	lr: 0.0001	evaluation reward: 1.72	epsilon: 0.9088348600019791	st
episode: 786	score: 1.0	memory length: 145918	epsilon: 0.9085358800019856	st
eps: 152	lr: 0.0001	evaluation reward: 1.71	epsilon: 0.9082369000019921	st
episode: 787	score: 0.0	memory length: 146042	epsilon: 0.9079359400019986	st
eps: 124	lr: 0.0001	evaluation reward: 1.68	epsilon: 0.9074449000020093	st
episode: 788	score: 1.0	memory length: 146193	epsilon: 0.9070865200020171	st
eps: 151	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.9066034000020275	st
episode: 789	score: 1.0	memory length: 146344	epsilon: 0.9062093800020361	st
eps: 151	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.9057757600020455	st
episode: 790	score: 1.0	memory length: 146496	epsilon: 0.9054154000020533	st
eps: 152	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.9050570200020611	st
episode: 791	score: 3.0	memory length: 146744	epsilon: 0.9047580400020676	st
eps: 248	lr: 0.0001	evaluation reward: 1.69	epsilon: 0.9045145000020729	st
episode: 792	score: 2.0	memory length: 146925	epsilon: 0.9040808800020823	st
eps: 181	lr: 0.0001	evaluation reward: 1.71	epsilon: 0.9037185400020902	st
episode: 793	score: 3.0	memory length: 147169	epsilon: 0.9030017800021057	st
eps: 244	lr: 0.0001	evaluation reward: 1.7	epsilon: 0.902665180002113	ste
episode: 794	score: 2.0	memory length: 147368	epsilon: 0.9023642200021196	st
eps: 199	lr: 0.0001	evaluation reward: 1.7	epsilon: 0.9021187000021249	st
episode: 795	score: 2.0	memory length: 147587	epsilon: 0.9018177400021314	st
eps: 219	lr: 0.0001	evaluation reward: 1.71	epsilon: 0.9013999600021405	st
episode: 796	score: 2.0	memory length: 147769	epsilon: 0.9009722800021498	st
eps: 182	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.9004871800021603	st
episode: 797	score: 2.0	memory length: 147950	epsilon: 0.9002416600021657	st
eps: 181	lr: 0.0001	evaluation reward: 1.72	epsilon: 0.8999981200021709	st
episode: 798	score: 1.0	memory length: 148101	epsilon: 0.8994853000021821	st
eps: 151	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.8989309000021941	st
episode: 799	score: 0.0	memory length: 148224		
eps: 123	lr: 0.0001	evaluation reward: 1.73		
episode: 800	score: 2.0	memory length: 148443		
eps: 219	lr: 0.0001	evaluation reward: 1.72		
episode: 801	score: 2.0	memory length: 148626		
eps: 183	lr: 0.0001	evaluation reward: 1.74		
episode: 802	score: 5.0	memory length: 148988		
eps: 362	lr: 0.0001	evaluation reward: 1.77		
episode: 803	score: 1.0	memory length: 149158		
ps: 170	lr: 0.0001	evaluation reward: 1.77		
episode: 804	score: 1.0	memory length: 149310		
eps: 152	lr: 0.0001	evaluation reward: 1.76		
episode: 805	score: 0.0	memory length: 149434		
eps: 124	lr: 0.0001	evaluation reward: 1.74		
episode: 806	score: 1.0	memory length: 149586		
eps: 152	lr: 0.0001	evaluation reward: 1.72		
episode: 807	score: 3.0	memory length: 149797		
eps: 211	lr: 0.0001	evaluation reward: 1.75		
episode: 808	score: 2.0	memory length: 150013		
eps: 216	lr: 0.0001	evaluation reward: 1.75		
episode: 809	score: 3.0	memory length: 150258		
eps: 245	lr: 0.0001	evaluation reward: 1.75		
episode: 810	score: 0.0	memory length: 150382		
eps: 124	lr: 0.0001	evaluation reward: 1.75		
episode: 811	score: 0.0	memory length: 150505		
eps: 123	lr: 0.0001	evaluation reward: 1.73		
episode: 812	score: 4.0	memory length: 150764		
eps: 259	lr: 0.0001	evaluation reward: 1.75		
episode: 813	score: 4.0	memory length: 151044		
eps: 280	lr: 0.0001	evaluation reward: 1.78		

episode: 814	score: 1.0	memory length: 151196	epsilon: 0.8986299400022006	st
eps: 152	lr: 0.0001	evaluation reward: 1.76	epsilon: 0.898384420002206	ste
episode: 815	score: 0.0	memory length: 151320	epsilon: 0.8980418800022134	st
ps: 124	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8974973800022252	st
episode: 816	score: 1.0	memory length: 151493	epsilon: 0.8972538400022305	st
eps: 173	lr: 0.0001	evaluation reward: 1.72	epsilon: 0.896952880002237	ste
episode: 817	score: 4.0	memory length: 151768	epsilon: 0.8964618400022477	st
eps: 275	lr: 0.0001	evaluation reward: 1.76	epsilon: 0.8960678200022563	st
episode: 818	score: 0.0	memory length: 151891	epsilon: 0.8956302400022658	st
eps: 123	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.8951728600022757	st
episode: 819	score: 1.0	memory length: 152043	epsilon: 0.8948699200022823	st
ps: 152	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8945709400022888	st
episode: 820	score: 3.0	memory length: 152291	epsilon: 0.8941769200022973	st
eps: 248	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.8937829000023059	st
episode: 821	score: 2.0	memory length: 152490	epsilon: 0.8934819400023124	st
eps: 199	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8927731000023278	st
episode: 822	score: 2.0	memory length: 152711	epsilon: 0.8919613000023454	st
eps: 221	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8916266800023527	st
episode: 823	score: 3.0	memory length: 152942	epsilon: 0.891199000002362	ste
eps: 231	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.8908069600023705	st
episode: 824	score: 1.0	memory length: 153095	epsilon: 0.8902347400023829	st
eps: 153	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.8899337800023894	st
episode: 825	score: 1.0	memory length: 153246	epsilon: 0.889537780002398	ste
eps: 151	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8892922600024034	st
episode: 826	score: 2.0	memory length: 153445	epsilon: 0.888892300002412	ste
eps: 199	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.8885913400024186	st
episode: 827	score: 2.0	memory length: 153644	epsilon: 0.8881418800024283	st
eps: 199	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8876924200024381	st
episode: 828	score: 1.0	memory length: 153796	epsilon: 0.8874488800024434	st
eps: 152	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8871479200024499	st
episode: 829	score: 6.0	memory length: 154154		
eps: 358	lr: 0.0001	evaluation reward: 1.79		
episode: 830	score: 7.0	memory length: 154564		
eps: 410	lr: 0.0001	evaluation reward: 1.84		
episode: 831	score: 1.0	memory length: 154733		
eps: 169	lr: 0.0001	evaluation reward: 1.84		
episode: 832	score: 2.0	memory length: 154949		
ps: 216	lr: 0.0001	evaluation reward: 1.83		
episode: 833	score: 2.0	memory length: 155147		
eps: 198	lr: 0.0001	evaluation reward: 1.85		
episode: 834	score: 5.0	memory length: 155436		
eps: 289	lr: 0.0001	evaluation reward: 1.89		
episode: 835	score: 1.0	memory length: 155588		
eps: 152	lr: 0.0001	evaluation reward: 1.88		
episode: 836	score: 2.0	memory length: 155788		
ps: 200	lr: 0.0001	evaluation reward: 1.88		
episode: 837	score: 0.0	memory length: 155912		
eps: 124	lr: 0.0001	evaluation reward: 1.86		
episode: 838	score: 2.0	memory length: 156114		
ps: 202	lr: 0.0001	evaluation reward: 1.85		
episode: 839	score: 1.0	memory length: 156266		
eps: 152	lr: 0.0001	evaluation reward: 1.85		
episode: 840	score: 3.0	memory length: 156493		
eps: 227	lr: 0.0001	evaluation reward: 1.85		
episode: 841	score: 3.0	memory length: 156720		
eps: 227	lr: 0.0001	evaluation reward: 1.86		
episode: 842	score: 0.0	memory length: 156843		
eps: 123	lr: 0.0001	evaluation reward: 1.83		
episode: 843	score: 1.0	memory length: 156995		
eps: 152	lr: 0.0001	evaluation reward: 1.83		

episode: 844	score: 1.0	memory length: 157148	epsilon: 0.8868449800024565	st
eps: 153	lr: 0.0001	evaluation reward: 1.81		
episode: 845	score: 0.0	memory length: 157272	epsilon: 0.8865994600024618	st
eps: 124	lr: 0.0001	evaluation reward: 1.79		
episode: 846	score: 6.0	memory length: 157667	epsilon: 0.8858173600024788	st
eps: 395	lr: 0.0001	evaluation reward: 1.83		
episode: 847	score: 1.0	memory length: 157837	epsilon: 0.8854807600024861	st
eps: 170	lr: 0.0001	evaluation reward: 1.82		
episode: 848	score: 2.0	memory length: 158036	epsilon: 0.8850867400024947	st
eps: 199	lr: 0.0001	evaluation reward: 1.84		
episode: 849	score: 5.0	memory length: 158382	epsilon: 0.8844016600025095	st
eps: 346	lr: 0.0001	evaluation reward: 1.88		
episode: 850	score: 1.0	memory length: 158552	epsilon: 0.8840650600025168	st
eps: 170	lr: 0.0001	evaluation reward: 1.88		
episode: 851	score: 1.0	memory length: 158723	epsilon: 0.8837264800025242	st
eps: 171	lr: 0.0001	evaluation reward: 1.89		
episode: 852	score: 3.0	memory length: 158970	epsilon: 0.8832374200025348	st
eps: 247	lr: 0.0001	evaluation reward: 1.91		
episode: 853	score: 0.0	memory length: 159093	epsilon: 0.8829938800025401	st
eps: 123	lr: 0.0001	evaluation reward: 1.91		
episode: 854	score: 2.0	memory length: 159291	epsilon: 0.8826018400025486	st
eps: 198	lr: 0.0001	evaluation reward: 1.9		
episode: 855	score: 2.0	memory length: 159490	epsilon: 0.8822078200025572	st
eps: 199	lr: 0.0001	evaluation reward: 1.88		
episode: 856	score: 4.0	memory length: 159784	epsilon: 0.8816257000025698	st
eps: 294	lr: 0.0001	evaluation reward: 1.9		
episode: 857	score: 0.0	memory length: 159907	epsilon: 0.8813821600025751	st
eps: 123	lr: 0.0001	evaluation reward: 1.89		
episode: 858	score: 0.0	memory length: 160031	epsilon: 0.8811366400025804	st
eps: 124	lr: 0.0001	evaluation reward: 1.88		
episode: 859	score: 1.0	memory length: 160203	epsilon: 0.8807960800025878	st
eps: 172	lr: 0.0001	evaluation reward: 1.87		
episode: 860	score: 4.0	memory length: 160496	epsilon: 0.8802159400026004	st
eps: 293	lr: 0.0001	evaluation reward: 1.9		
episode: 861	score: 5.0	memory length: 160847	epsilon: 0.8795209600026155	st
eps: 351	lr: 0.0001	evaluation reward: 1.93		
episode: 862	score: 1.0	memory length: 161017	epsilon: 0.8791843600026228	st
eps: 170	lr: 0.0001	evaluation reward: 1.93		
episode: 863	score: 1.0	memory length: 161168	epsilon: 0.8788853800026293	st
eps: 151	lr: 0.0001	evaluation reward: 1.92		
episode: 864	score: 3.0	memory length: 161421	epsilon: 0.8783844400026402	st
eps: 253	lr: 0.0001	evaluation reward: 1.94		
episode: 865	score: 3.0	memory length: 161650	epsilon: 0.87793102000265	step
s: 229	lr: 0.0001	evaluation reward: 1.96		
episode: 866	score: 0.0	memory length: 161774	epsilon: 0.8776855000026553	st
eps: 124	lr: 0.0001	evaluation reward: 1.94		
episode: 867	score: 1.0	memory length: 161945	epsilon: 0.8773469200026627	st
eps: 171	lr: 0.0001	evaluation reward: 1.93		
episode: 868	score: 0.0	memory length: 162068	epsilon: 0.877103380002668	ste
ps: 123	lr: 0.0001	evaluation reward: 1.92		
episode: 869	score: 2.0	memory length: 162288	epsilon: 0.8766677800026774	st
eps: 220	lr: 0.0001	evaluation reward: 1.94		
episode: 870	score: 3.0	memory length: 162554	epsilon: 0.8761411000026889	st
eps: 266	lr: 0.0001	evaluation reward: 1.94		
episode: 871	score: 1.0	memory length: 162706	epsilon: 0.8758401400026954	st
eps: 152	lr: 0.0001	evaluation reward: 1.92		
episode: 872	score: 1.0	memory length: 162877	epsilon: 0.8755015600027027	st
eps: 171	lr: 0.0001	evaluation reward: 1.92		
episode: 873	score: 2.0	memory length: 163077	epsilon: 0.8751055600027113	st
eps: 200	lr: 0.0001	evaluation reward: 1.92		

episode: 874	score: 2.0	memory length: 163258	epsilon: 0.8747471800027191	st
eps: 181	lr: 0.0001	evaluation reward: 1.91		
episode: 875	score: 3.0	memory length: 163505	epsilon: 0.8742581200027297	st
eps: 247	lr: 0.0001	evaluation reward: 1.92		
episode: 876	score: 2.0	memory length: 163686	epsilon: 0.8738997400027375	st
eps: 181	lr: 0.0001	evaluation reward: 1.94		
episode: 877	score: 1.0	memory length: 163838	epsilon: 0.873598780002744	ste
ps: 152	lr: 0.0001	evaluation reward: 1.9		
episode: 878	score: 4.0	memory length: 164135	epsilon: 0.8730107200027568	st
eps: 297	lr: 0.0001	evaluation reward: 1.92		
episode: 879	score: 5.0	memory length: 164465	epsilon: 0.872357320002771	ste
ps: 330	lr: 0.0001	evaluation reward: 1.95		
episode: 880	score: 2.0	memory length: 164685	epsilon: 0.8719217200027805	st
eps: 220	lr: 0.0001	evaluation reward: 1.94		
episode: 881	score: 1.0	memory length: 164855	epsilon: 0.8715851200027878	st
eps: 170	lr: 0.0001	evaluation reward: 1.94		
episode: 882	score: 1.0	memory length: 165026	epsilon: 0.8712465400027951	st
eps: 171	lr: 0.0001	evaluation reward: 1.94		
episode: 883	score: 0.0	memory length: 165150	epsilon: 0.8710010200028004	st
eps: 124	lr: 0.0001	evaluation reward: 1.91		
episode: 884	score: 3.0	memory length: 165400	epsilon: 0.8705060200028112	st
eps: 250	lr: 0.0001	evaluation reward: 1.9		
episode: 885	score: 2.0	memory length: 165598	epsilon: 0.8701139800028197	st
eps: 198	lr: 0.0001	evaluation reward: 1.91		
episode: 886	score: 4.0	memory length: 165861	epsilon: 0.869593240002831	ste
ps: 263	lr: 0.0001	evaluation reward: 1.94		
episode: 887	score: 3.0	memory length: 166110	epsilon: 0.8691002200028417	st
eps: 249	lr: 0.0001	evaluation reward: 1.97		
episode: 888	score: 0.0	memory length: 166234	epsilon: 0.868854700002847	ste
ps: 124	lr: 0.0001	evaluation reward: 1.96		
episode: 889	score: 2.0	memory length: 166433	epsilon: 0.8684606800028556	st
eps: 199	lr: 0.0001	evaluation reward: 1.97		
episode: 890	score: 1.0	memory length: 166586	epsilon: 0.8681577400028622	st
eps: 153	lr: 0.0001	evaluation reward: 1.97		
episode: 891	score: 0.0	memory length: 166709	epsilon: 0.8679142000028675	st
eps: 123	lr: 0.0001	evaluation reward: 1.94		
episode: 892	score: 3.0	memory length: 166936	epsilon: 0.8674647400028772	st
eps: 227	lr: 0.0001	evaluation reward: 1.95		
episode: 893	score: 3.0	memory length: 167166	epsilon: 0.8670093400028871	st
eps: 230	lr: 0.0001	evaluation reward: 1.95		
episode: 894	score: 4.0	memory length: 167482	epsilon: 0.8663836600029007	st
eps: 316	lr: 0.0001	evaluation reward: 1.97		
episode: 895	score: 0.0	memory length: 167606	epsilon: 0.866138140002906	ste
ps: 124	lr: 0.0001	evaluation reward: 1.95		
episode: 896	score: 4.0	memory length: 167902	epsilon: 0.8655520600029187	st
eps: 296	lr: 0.0001	evaluation reward: 1.97		
episode: 897	score: 1.0	memory length: 168074	epsilon: 0.8652115000029261	st
eps: 172	lr: 0.0001	evaluation reward: 1.96		
episode: 898	score: 2.0	memory length: 168274	epsilon: 0.8648155000029347	st
eps: 200	lr: 0.0001	evaluation reward: 1.97		
episode: 899	score: 2.0	memory length: 168476	epsilon: 0.8644155400029434	st
eps: 202	lr: 0.0001	evaluation reward: 1.99		
episode: 900	score: 0.0	memory length: 168600	epsilon: 0.8641700200029487	st
eps: 124	lr: 0.0001	evaluation reward: 1.97		
episode: 901	score: 3.0	memory length: 168867	epsilon: 0.8636413600029602	st
eps: 267	lr: 0.0001	evaluation reward: 1.98		
episode: 902	score: 1.0	memory length: 169038	epsilon: 0.8633027800029676	st
eps: 171	lr: 0.0001	evaluation reward: 1.94		
episode: 903	score: 4.0	memory length: 169315	epsilon: 0.8627543200029795	st
eps: 277	lr: 0.0001	evaluation reward: 1.97		

episode: 904	score: 2.0	memory length: 169533	epsilon: 0.8623226800029888	st
eps: 218	lr: 0.0001	evaluation reward: 1.98		
episode: 905	score: 2.0	memory length: 169752	epsilon: 0.8618890600029983	st
eps: 219	lr: 0.0001	evaluation reward: 2.0		
episode: 906	score: 1.0	memory length: 169903	epsilon: 0.8615900800030047	st
eps: 151	lr: 0.0001	evaluation reward: 2.0		
episode: 907	score: 2.0	memory length: 170084	epsilon: 0.8612317000030125	st
eps: 181	lr: 0.0001	evaluation reward: 1.99		
episode: 908	score: 1.0	memory length: 170254	epsilon: 0.8608951000030198	st
eps: 170	lr: 0.0001	evaluation reward: 1.98		
episode: 909	score: 4.0	memory length: 170511	epsilon: 0.8603862400030309	st
eps: 257	lr: 0.0001	evaluation reward: 1.99		
episode: 910	score: 2.0	memory length: 170729	epsilon: 0.8599546000030402	st
eps: 218	lr: 0.0001	evaluation reward: 2.01		
episode: 911	score: 0.0	memory length: 170852	epsilon: 0.8597110600030455	st
eps: 123	lr: 0.0001	evaluation reward: 2.01		
episode: 912	score: 2.0	memory length: 171051	epsilon: 0.8593170400030541	st
eps: 199	lr: 0.0001	evaluation reward: 1.99		
episode: 913	score: 3.0	memory length: 171299	epsilon: 0.8588260000030647	st
eps: 248	lr: 0.0001	evaluation reward: 1.98		
episode: 914	score: 2.0	memory length: 171498	epsilon: 0.8584319800030733	st
eps: 199	lr: 0.0001	evaluation reward: 1.99		
episode: 915	score: 0.0	memory length: 171622	epsilon: 0.8581864600030786	st
eps: 124	lr: 0.0001	evaluation reward: 1.99		
episode: 916	score: 1.0	memory length: 171774	epsilon: 0.8578855000030852	st
eps: 152	lr: 0.0001	evaluation reward: 1.99		
episode: 917	score: 2.0	memory length: 171973	epsilon: 0.8574914800030937	st
eps: 199	lr: 0.0001	evaluation reward: 1.97		
episode: 918	score: 4.0	memory length: 172271	epsilon: 0.8569014400031065	st
eps: 298	lr: 0.0001	evaluation reward: 2.01		
episode: 919	score: 2.0	memory length: 172491	epsilon: 0.856465840003116	ste
ps: 220	lr: 0.0001	evaluation reward: 2.02		
episode: 920	score: 0.0	memory length: 172615	epsilon: 0.8562203200031213	st
eps: 124	lr: 0.0001	evaluation reward: 1.99		
episode: 921	score: 1.0	memory length: 172785	epsilon: 0.8558837200031286	st
eps: 170	lr: 0.0001	evaluation reward: 1.98		
episode: 922	score: 2.0	memory length: 172985	epsilon: 0.8554877200031372	st
eps: 200	lr: 0.0001	evaluation reward: 1.98		
episode: 923	score: 4.0	memory length: 173261	epsilon: 0.8549412400031491	st
eps: 276	lr: 0.0001	evaluation reward: 1.99		
episode: 924	score: 2.0	memory length: 173479	epsilon: 0.8545096000031585	st
eps: 218	lr: 0.0001	evaluation reward: 2.0		
episode: 925	score: 2.0	memory length: 173698	epsilon: 0.8540759800031679	st
eps: 219	lr: 0.0001	evaluation reward: 2.01		
episode: 926	score: 3.0	memory length: 173942	epsilon: 0.8535928600031784	st
eps: 244	lr: 0.0001	evaluation reward: 2.02		
episode: 927	score: 2.0	memory length: 174125	epsilon: 0.8532305200031862	st
eps: 183	lr: 0.0001	evaluation reward: 2.02		
episode: 928	score: 1.0	memory length: 174276	epsilon: 0.8529315400031927	st
eps: 151	lr: 0.0001	evaluation reward: 2.02		
episode: 929	score: 0.0	memory length: 174400	epsilon: 0.852686020003198	ste
ps: 124	lr: 0.0001	evaluation reward: 1.96		
episode: 930	score: 1.0	memory length: 174552	epsilon: 0.8523850600032046	st
eps: 152	lr: 0.0001	evaluation reward: 1.9		
episode: 931	score: 1.0	memory length: 174722	epsilon: 0.8520484600032119	st
eps: 170	lr: 0.0001	evaluation reward: 1.9		
episode: 932	score: 3.0	memory length: 174971	epsilon: 0.8515554400032226	st
eps: 249	lr: 0.0001	evaluation reward: 1.91		
episode: 933	score: 4.0	memory length: 175250	epsilon: 0.8510030200032346	st
eps: 279	lr: 0.0001	evaluation reward: 1.93		

episode: 934	score: 1.0	memory length: 175402	epsilon: 0.8507020600032411	st
eps: 152	lr: 0.0001	evaluation reward: 1.89	epsilon: 0.8504030800032476	st
episode: 935	score: 1.0	memory length: 175553	epsilon: 0.8500090600032562	st
eps: 151	lr: 0.0001	evaluation reward: 1.89	epsilon: 0.8495774200032655	st
episode: 936	score: 2.0	memory length: 175752	epsilon: 0.8493338800032708	st
eps: 199	lr: 0.0001	evaluation reward: 1.89	epsilon: 0.8490329200032773	st
episode: 937	score: 2.0	memory length: 175970	epsilon: 0.8484785200032894	st
eps: 218	lr: 0.0001	evaluation reward: 1.91	epsilon: 0.8479855000033001	st
episode: 938	score: 0.0	memory length: 176093	epsilon: 0.8475914800033086	st
eps: 123	lr: 0.0001	evaluation reward: 1.89	epsilon: 0.847345960003314	ste
episode: 939	score: 1.0	memory length: 176245	epsilon: 0.8469123400033234	st
eps: 152	lr: 0.0001	evaluation reward: 1.89	epsilon: 0.8463599200033354	st
episode: 940	score: 4.0	memory length: 176525	epsilon: 0.8459104600033451	st
eps: 280	lr: 0.0001	evaluation reward: 1.9	epsilon: 0.8455184200033536	st
episode: 941	score: 3.0	memory length: 176774	epsilon: 0.845088760003363	ste
eps: 249	lr: 0.0001	evaluation reward: 1.9	epsilon: 0.8443739800033785	st
episode: 942	score: 2.0	memory length: 176973	epsilon: 0.8439423400033879	st
eps: 199	lr: 0.0001	evaluation reward: 1.92	epsilon: 0.8436077200033951	st
episode: 943	score: 0.0	memory length: 177097	epsilon: 0.8431840000034043	st
ps: 124	lr: 0.0001	evaluation reward: 1.91	epsilon: 0.8427325600034141	st
episode: 944	score: 2.0	memory length: 177316	epsilon: 0.8424335800034206	st
eps: 219	lr: 0.0001	evaluation reward: 1.92	epsilon: 0.8419821400034304	st
episode: 945	score: 4.0	memory length: 177595	epsilon: 0.8416435600034378	st
eps: 279	lr: 0.0001	evaluation reward: 1.96	epsilon: 0.8413980400034431	st
episode: 946	score: 3.0	memory length: 177822	epsilon: 0.8409644200034525	st
eps: 227	lr: 0.0001	evaluation reward: 1.93	epsilon: 0.8403229000034664	st
episode: 947	score: 2.0	memory length: 178020	epsilon: 0.840021940003473	ste
eps: 198	lr: 0.0001	evaluation reward: 1.94	epsilon: 0.8396853400034803	st
episode: 948	score: 2.0	memory length: 178237	epsilon: 0.8392873600034889	st
ps: 217	lr: 0.0001	evaluation reward: 1.94	epsilon: 0.8388577000034982	st
episode: 949	score: 6.0	memory length: 178598		
eps: 361	lr: 0.0001	evaluation reward: 1.95		
episode: 950	score: 2.0	memory length: 178816		
eps: 218	lr: 0.0001	evaluation reward: 1.96		
episode: 951	score: 1.0	memory length: 178985		
eps: 169	lr: 0.0001	evaluation reward: 1.96		
episode: 952	score: 3.0	memory length: 179199		
eps: 214	lr: 0.0001	evaluation reward: 1.96		
episode: 953	score: 3.0	memory length: 179427		
eps: 228	lr: 0.0001	evaluation reward: 1.99		
episode: 954	score: 1.0	memory length: 179578		
eps: 151	lr: 0.0001	evaluation reward: 1.98		
episode: 955	score: 3.0	memory length: 179806		
eps: 228	lr: 0.0001	evaluation reward: 1.99		
episode: 956	score: 1.0	memory length: 179977		
eps: 171	lr: 0.0001	evaluation reward: 1.96		
episode: 957	score: 0.0	memory length: 180101		
eps: 124	lr: 0.0001	evaluation reward: 1.96		
episode: 958	score: 2.0	memory length: 180320		
eps: 219	lr: 0.0001	evaluation reward: 1.98		
episode: 959	score: 5.0	memory length: 180644		
eps: 324	lr: 0.0001	evaluation reward: 2.02		
episode: 960	score: 1.0	memory length: 180796		
ps: 152	lr: 0.0001	evaluation reward: 1.99		
episode: 961	score: 1.0	memory length: 180966		
eps: 170	lr: 0.0001	evaluation reward: 1.95		
episode: 962	score: 2.0	memory length: 181167		
eps: 201	lr: 0.0001	evaluation reward: 1.96		
episode: 963	score: 2.0	memory length: 181384		
eps: 217	lr: 0.0001	evaluation reward: 1.97		

episode: 964	score: 1.0	memory length: 181555	epsilon: 0.8385191200035056	st
eps: 171	lr: 0.0001	evaluation reward: 1.95	epsilon: 0.8377845400035215	st
episode: 965	score: 6.0	memory length: 181926	epsilon: 0.8373608200035307	st
eps: 371	lr: 0.0001	evaluation reward: 1.98	epsilon: 0.8367094000035449	st
episode: 966	score: 3.0	memory length: 182140	epsilon: 0.8363728000035522	st
eps: 214	lr: 0.0001	evaluation reward: 2.01	epsilon: 0.8359787800035607	st
episode: 967	score: 5.0	memory length: 182469	epsilon: 0.8354560600035721	st
eps: 329	lr: 0.0001	evaluation reward: 2.05	epsilon: 0.834216580003599	ste
episode: 968	score: 1.0	memory length: 182639	epsilon: 0.8346640600035893	st
eps: 170	lr: 0.0001	evaluation reward: 2.06	epsilon: 0.8337592000036089	st
episode: 969	score: 2.0	memory length: 182838	epsilon: 0.8333651800036175	st
eps: 199	lr: 0.0001	evaluation reward: 2.06	epsilon: 0.833064220003624	ste
episode: 970	score: 3.0	memory length: 183102	epsilon: 0.8327632600036305	st
eps: 264	lr: 0.0001	evaluation reward: 2.06	epsilon: 0.8324029000036384	st
episode: 971	score: 2.0	memory length: 183300	epsilon: 0.8319554200036481	st
eps: 198	lr: 0.0001	evaluation reward: 2.07	epsilon: 0.8313277600036617	st
episode: 972	score: 2.0	memory length: 183502	epsilon: 0.8309674000036695	st
eps: 202	lr: 0.0001	evaluation reward: 2.08	epsilon: 0.8303793400036823	st
episode: 973	score: 3.0	memory length: 183728	epsilon: 0.8300427400036896	st
ps: 226	lr: 0.0001	evaluation reward: 2.09	epsilon: 0.8294329000037028	st
episode: 974	score: 3.0	memory length: 183959	epsilon: 0.8288785000037149	st
eps: 231	lr: 0.0001	evaluation reward: 2.1	epsilon: 0.8285775400037214	st
episode: 975	score: 2.0	memory length: 184158	epsilon: 0.8281478800037307	st
eps: 199	lr: 0.0001	evaluation reward: 2.09	epsilon: 0.8278469200037373	st
episode: 976	score: 1.0	memory length: 184310	epsilon: 0.827446960003746	ste
ps: 152	lr: 0.0001	evaluation reward: 2.08	epsilon: 0.8271479800037524	st
episode: 977	score: 1.0	memory length: 184462	epsilon: 0.826847020003759	ste
eps: 152	lr: 0.0001	evaluation reward: 2.08	epsilon: 0.8264549800037675	st
episode: 978	score: 2.0	memory length: 184644	epsilon: 0.8260253200037768	st
eps: 182	lr: 0.0001	evaluation reward: 2.06		
episode: 979	score: 3.0	memory length: 184870		
eps: 226	lr: 0.0001	evaluation reward: 2.04		
episode: 980	score: 6.0	memory length: 185187		
eps: 317	lr: 0.0001	evaluation reward: 2.08		
episode: 981	score: 2.0	memory length: 185369		
eps: 182	lr: 0.0001	evaluation reward: 2.09		
episode: 982	score: 4.0	memory length: 185666		
eps: 297	lr: 0.0001	evaluation reward: 2.12		
episode: 983	score: 1.0	memory length: 185836		
eps: 170	lr: 0.0001	evaluation reward: 2.13		
episode: 984	score: 5.0	memory length: 186144		
eps: 308	lr: 0.0001	evaluation reward: 2.15		
episode: 985	score: 4.0	memory length: 186424		
eps: 280	lr: 0.0001	evaluation reward: 2.17		
episode: 986	score: 1.0	memory length: 186576		
eps: 152	lr: 0.0001	evaluation reward: 2.14		
episode: 987	score: 2.0	memory length: 186793		
eps: 217	lr: 0.0001	evaluation reward: 2.13		
episode: 988	score: 1.0	memory length: 186945		
eps: 152	lr: 0.0001	evaluation reward: 2.14		
episode: 989	score: 2.0	memory length: 187147		
ps: 202	lr: 0.0001	evaluation reward: 2.14		
episode: 990	score: 1.0	memory length: 187298		
eps: 151	lr: 0.0001	evaluation reward: 2.14		
episode: 991	score: 1.0	memory length: 187450		
ps: 152	lr: 0.0001	evaluation reward: 2.15		
episode: 992	score: 2.0	memory length: 187648		
eps: 198	lr: 0.0001	evaluation reward: 2.14		
episode: 993	score: 2.0	memory length: 187865		
eps: 217	lr: 0.0001	evaluation reward: 2.13		

episode: 994	score: 5.0	memory length: 188193	epsilon: 0.8253758800037909	st
eps: 328	lr: 0.0001	evaluation reward: 2.14	epsilon: 0.8248848400038016	st
episode: 995	score: 3.0	memory length: 188441	epsilon: 0.8242967800038143	st
eps: 248	lr: 0.0001	evaluation reward: 2.17	epsilon: 0.8238631600038238	st
episode: 996	score: 4.0	memory length: 188738	epsilon: 0.8233661800038345	st
eps: 297	lr: 0.0001	evaluation reward: 2.17	epsilon: 0.8229721600038431	st
episode: 997	score: 2.0	memory length: 188957	epsilon: 0.8225187400038529	s
eps: 219	lr: 0.0001	evaluation reward: 2.18	epsilon: 0.8221267000038615	s
episode: 998	score: 3.0	memory length: 189208	epsilon: 0.82173268000387	ste
eps: 251	lr: 0.0001	evaluation reward: 2.19	epsilon: 0.8213050000038793	s
episode: 999	score: 2.0	memory length: 189407	epsilon: 0.820948600003887	st
eps: 199	lr: 0.0001	evaluation reward: 2.19	epsilon: 0.8203110400039009	s
episode: 1000	score: 3.0	memory length: 189636	epsilon: 0.8199764200039081	s
teps: 229	lr: 0.0001	evaluation reward: 2.22	epsilon: 0.8194576600039194	s
episode: 1001	score: 2.0	memory length: 189834	epsilon: 0.8191012600039271	s
teps: 198	lr: 0.0001	evaluation reward: 2.21	epsilon: 0.8186834800039362	s
episode: 1002	score: 2.0	memory length: 190033	epsilon: 0.8183211400039441	s
ps: 199	lr: 0.0001	evaluation reward: 2.22	epsilon: 0.8179825600039514	s
episode: 1003	score: 2.0	memory length: 190249	epsilon: 0.8177390200039567	s
teps: 216	lr: 0.0001	evaluation reward: 2.2	epsilon: 0.8173786600039645	s
episode: 1004	score: 2.0	memory length: 190429	epsilon: 0.8168717800039755	s
eps: 180	lr: 0.0001	evaluation reward: 2.2	epsilon: 0.8165708200039821	s
episode: 1005	score: 5.0	memory length: 190751	epsilon: 0.8160599800039932	s
teps: 322	lr: 0.0001	evaluation reward: 2.23	epsilon: 0.8152323400040111	s
episode: 1006	score: 1.0	memory length: 190920	epsilon: 0.8148343600040198	s
teps: 169	lr: 0.0001	evaluation reward: 2.23	epsilon: 0.8144007400040292	s
episode: 1007	score: 4.0	memory length: 191182	epsilon: 0.8140621600040365	s
teps: 262	lr: 0.0001	evaluation reward: 2.25	epsilon: 0.8136681400040451	s
episode: 1008	score: 2.0	memory length: 191362	epsilon: 0.8132701600040537	s
teps: 180	lr: 0.0001	evaluation reward: 2.26	epsilon: 0.8127791200040644	s
episode: 1009	score: 3.0	memory length: 191573	epsilon: 0.8123234000407321	s
teps: 211	lr: 0.0001	evaluation reward: 2.25	epsilon: 0.8119764200040811	s
episode: 1010	score: 2.0	memory length: 191756	epsilon: 0.81168717800040982	s
teps: 183	lr: 0.0001	evaluation reward: 2.25	epsilon: 0.8113050000041111	s
episode: 1011	score: 1.0	memory length: 191927	epsilon: 0.8110621600041291	s
teps: 171	lr: 0.0001	evaluation reward: 2.26	epsilon: 0.8107739020004244	s
episode: 1012	score: 0.0	memory length: 192050	epsilon: 0.8105099800042611	s
teps: 123	lr: 0.0001	evaluation reward: 2.24	epsilon: 0.8102323400042791	s
episode: 1013	score: 2.0	memory length: 192232	epsilon: 0.8100000000043071	s
teps: 182	lr: 0.0001	evaluation reward: 2.23	epsilon: 0.8097642000043251	s
episode: 1014	score: 4.0	memory length: 192488	epsilon: 0.8095323400043431	s
teps: 256	lr: 0.0001	evaluation reward: 2.25	epsilon: 0.8093099800043611	s
episode: 1015	score: 1.0	memory length: 192640	epsilon: 0.8090876600043791	s
teps: 152	lr: 0.0001	evaluation reward: 2.26	epsilon: 0.8088644200043971	s
episode: 1016	score: 4.0	memory length: 192898	epsilon: 0.8086411800044151	s
teps: 258	lr: 0.0001	evaluation reward: 2.29	epsilon: 0.8084179800044331	s
episode: 1017	score: 7.0	memory length: 193316	epsilon: 0.8081946400044511	s
teps: 418	lr: 0.0001	evaluation reward: 2.34	epsilon: 0.8079714000044691	s
episode: 1018	score: 2.0	memory length: 193517	epsilon: 0.8077481600044871	s
teps: 201	lr: 0.0001	evaluation reward: 2.32	epsilon: 0.8075249200045051	s
episode: 1019	score: 2.0	memory length: 193736	epsilon: 0.8073016800045231	s
teps: 219	lr: 0.0001	evaluation reward: 2.32	epsilon: 0.8070784400045411	s
episode: 1020	score: 1.0	memory length: 193907	epsilon: 0.8068552000045591	s
teps: 171	lr: 0.0001	evaluation reward: 2.33	epsilon: 0.8066319600045771	s
episode: 1021	score: 2.0	memory length: 194106	epsilon: 0.8064087200045951	s
teps: 199	lr: 0.0001	evaluation reward: 2.34	epsilon: 0.8061854800046131	s
episode: 1022	score: 2.0	memory length: 194307	epsilon: 0.8059622400046311	s
teps: 201	lr: 0.0001	evaluation reward: 2.34	epsilon: 0.8057389000046491	s
episode: 1023	score: 3.0	memory length: 194555	epsilon: 0.8055156600046671	s
teps: 248	lr: 0.0001	evaluation reward: 2.33	epsilon: 0.8052923200046851	s

episode: 1024	score: 2.0	memory length: 194754	epsilon: 0.8123851000040729	s
teps: 199	lr: 0.0001	evaluation reward: 2.33		
episode: 1025	score: 4.0	memory length: 195050	epsilon: 0.8117990200040857	s
teps: 296	lr: 0.0001	evaluation reward: 2.35		
episode: 1026	score: 2.0	memory length: 195249	epsilon: 0.8114050000040942	s
teps: 199	lr: 0.0001	evaluation reward: 2.34		
episode: 1027	score: 0.0	memory length: 195373	epsilon: 0.8111594800040995	s
teps: 124	lr: 0.0001	evaluation reward: 2.32		
episode: 1028	score: 2.0	memory length: 195556	epsilon: 0.8107971400041074	s
teps: 183	lr: 0.0001	evaluation reward: 2.33		
episode: 1029	score: 4.0	memory length: 195817	epsilon: 0.8102803600041186	s
teps: 261	lr: 0.0001	evaluation reward: 2.37		
episode: 1030	score: 4.0	memory length: 196115	epsilon: 0.8096903200041314	s
teps: 298	lr: 0.0001	evaluation reward: 2.4		
episode: 1031	score: 2.0	memory length: 196314	epsilon: 0.80929630000414	ste
ps: 199	lr: 0.0001	evaluation reward: 2.41		
episode: 1032	score: 2.0	memory length: 196532	epsilon: 0.8088646600041494	s
teps: 218	lr: 0.0001	evaluation reward: 2.4		
episode: 1033	score: 3.0	memory length: 196780	epsilon: 0.80837362000416	ste
ps: 248	lr: 0.0001	evaluation reward: 2.39		
episode: 1034	score: 3.0	memory length: 197008	epsilon: 0.8079221800041698	s
teps: 228	lr: 0.0001	evaluation reward: 2.41		
episode: 1035	score: 2.0	memory length: 197190	epsilon: 0.8075618200041776	s
teps: 182	lr: 0.0001	evaluation reward: 2.42		
episode: 1036	score: 5.0	memory length: 197521	epsilon: 0.8069064400041919	s
teps: 331	lr: 0.0001	evaluation reward: 2.45		
episode: 1037	score: 3.0	memory length: 197772	epsilon: 0.8064094600042027	s
teps: 251	lr: 0.0001	evaluation reward: 2.46		
episode: 1038	score: 3.0	memory length: 198017	epsilon: 0.8059243600042132	s
teps: 245	lr: 0.0001	evaluation reward: 2.49		
episode: 1039	score: 1.0	memory length: 198190	epsilon: 0.8055818200042206	s
teps: 173	lr: 0.0001	evaluation reward: 2.49		
episode: 1040	score: 0.0	memory length: 198314	epsilon: 0.805336300004226	st
eps: 124	lr: 0.0001	evaluation reward: 2.45		
episode: 1041	score: 8.0	memory length: 198648	epsilon: 0.8046749800042403	s
teps: 334	lr: 0.0001	evaluation reward: 2.5		
episode: 1042	score: 3.0	memory length: 198876	epsilon: 0.8042235400042501	s
teps: 228	lr: 0.0001	evaluation reward: 2.51		
episode: 1043	score: 0.0	memory length: 198999	epsilon: 0.803980000042554	s
teps: 123	lr: 0.0001	evaluation reward: 2.51		
episode: 1044	score: 3.0	memory length: 199229	epsilon: 0.8035246000042653	s
teps: 230	lr: 0.0001	evaluation reward: 2.52		
episode: 1045	score: 1.0	memory length: 199399	epsilon: 0.8031880000042726	s
teps: 170	lr: 0.0001	evaluation reward: 2.49		
episode: 1046	score: 0.0	memory length: 199523	epsilon: 0.8029424800042779	s
teps: 124	lr: 0.0001	evaluation reward: 2.46		
episode: 1047	score: 3.0	memory length: 199794	epsilon: 0.8024059000042896	s
teps: 271	lr: 0.0001	evaluation reward: 2.47		
episode: 1048	score: 4.0	memory length: 200109	epsilon: 0.8017822000043031	s
teps: 315	lr: 4e-05	evaluation reward: 2.49		
episode: 1049	score: 0.0	memory length: 200232	epsilon: 0.8015386600043084	s
teps: 123	lr: 4e-05	evaluation reward: 2.43		
episode: 1050	score: 8.0	memory length: 200695	epsilon: 0.8006219200043283	s
teps: 463	lr: 4e-05	evaluation reward: 2.49		
episode: 1051	score: 1.0	memory length: 200847	epsilon: 0.8003209600043348	s
teps: 152	lr: 4e-05	evaluation reward: 2.49		
episode: 1052	score: 1.0	memory length: 200999	epsilon: 0.8000200000043414	s
teps: 152	lr: 4e-05	evaluation reward: 2.47		
episode: 1053	score: 3.0	memory length: 201209	epsilon: 0.7996042000043504	s
teps: 210	lr: 4e-05	evaluation reward: 2.47		

episode: 1054 score: 2.0 memory length: 201408	epsilon: 0.799210180004359	st
eps: 199 lr: 4e-05 evaluation reward: 2.48	epsilon: 0.7985429200043734	s
episode: 1055 score: 6.0 memory length: 201745	epsilon: 0.79824196000438	ste
teps: 337 lr: 4e-05 evaluation reward: 2.51	epsilon: 0.7977568600043905	s
episode: 1056 score: 1.0 memory length: 201897	epsilon: 0.7971707800044032	s
ps: 152 lr: 4e-05 evaluation reward: 2.51	epsilon: 0.7966361800044148	s
episode: 1057 score: 3.0 memory length: 202142	epsilon: 0.7962401800044234	s
teps: 245 lr: 4e-05 evaluation reward: 2.54	epsilon: 0.7958996200044308	s
episode: 1058 score: 4.0 memory length: 202438	epsilon: 0.7952699800044445	s
teps: 296 lr: 4e-05 evaluation reward: 2.56	epsilon: 0.794969020004451	st
episode: 1059 score: 3.0 memory length: 202708	epsilon: 0.7945750000044596	s
teps: 270 lr: 4e-05 evaluation reward: 2.54	epsilon: 0.7940839600044702	s
episode: 1060 score: 2.0 memory length: 202908	epsilon: 0.7938404200044755	s
teps: 200 lr: 4e-05 evaluation reward: 2.55	epsilon: 0.7932998800044873	s
episode: 1061 score: 1.0 memory length: 203080	epsilon: 0.7928246800044976	s
teps: 172 lr: 4e-05 evaluation reward: 2.55	epsilon: 0.7919178400045173	s
episode: 1062 score: 4.0 memory length: 203398	epsilon: 0.7913674000045292	s
teps: 318 lr: 4e-05 evaluation reward: 2.57	epsilon: 0.7908763600045399	s
episode: 1063 score: 1.0 memory length: 203550	epsilon: 0.7905358000045473	s
eps: 152 lr: 4e-05 evaluation reward: 2.56	epsilon: 0.7901101000045565	s
episode: 1064 score: 2.0 memory length: 203749	epsilon: 0.7896467800045666	s
teps: 199 lr: 4e-05 evaluation reward: 2.57	epsilon: 0.7891953400045764	s
episode: 1065 score: 3.0 memory length: 203997	epsilon: 0.7887102400045869	s
teps: 248 lr: 4e-05 evaluation reward: 2.54	epsilon: 0.7882211800045975	s
episode: 1066 score: 0.0 memory length: 204120	epsilon: 0.7873440400046166	s
teps: 123 lr: 4e-05 evaluation reward: 2.51	epsilon: 0.7869460600046252	s
episode: 1067 score: 4.0 memory length: 204393	epsilon: 0.7867005400046305	s
teps: 273 lr: 4e-05 evaluation reward: 2.5	epsilon: 0.7863639400046378	s
episode: 1068 score: 4.0 memory length: 204633	epsilon: 0.7858333000046493	s
teps: 240 lr: 4e-05 evaluation reward: 2.53	epsilon: 0.78534424000466	ste
episode: 1069 score: 8.0 memory length: 205091		
teps: 458 lr: 4e-05 evaluation reward: 2.59		
episode: 1070 score: 4.0 memory length: 205369		
teps: 278 lr: 4e-05 evaluation reward: 2.6		
episode: 1071 score: 3.0 memory length: 205617		
teps: 248 lr: 4e-05 evaluation reward: 2.61		
episode: 1072 score: 1.0 memory length: 205789		
teps: 172 lr: 4e-05 evaluation reward: 2.6		
episode: 1073 score: 3.0 memory length: 206004		
teps: 215 lr: 4e-05 evaluation reward: 2.6		
episode: 1074 score: 3.0 memory length: 206238		
teps: 234 lr: 4e-05 evaluation reward: 2.6		
episode: 1075 score: 3.0 memory length: 206466		
teps: 228 lr: 4e-05 evaluation reward: 2.61		
episode: 1076 score: 3.0 memory length: 206711		
teps: 245 lr: 4e-05 evaluation reward: 2.63		
episode: 1077 score: 3.0 memory length: 206958		
teps: 247 lr: 4e-05 evaluation reward: 2.65		
episode: 1078 score: 8.0 memory length: 207401		
teps: 443 lr: 4e-05 evaluation reward: 2.71		
episode: 1079 score: 2.0 memory length: 207602		
teps: 201 lr: 4e-05 evaluation reward: 2.7		
episode: 1080 score: 0.0 memory length: 207726		
teps: 124 lr: 4e-05 evaluation reward: 2.64		
episode: 1081 score: 1.0 memory length: 207896		
teps: 170 lr: 4e-05 evaluation reward: 2.63		
episode: 1082 score: 3.0 memory length: 208164		
teps: 268 lr: 4e-05 evaluation reward: 2.62		
episode: 1083 score: 3.0 memory length: 208411		
ps: 247 lr: 4e-05 evaluation reward: 2.64		

episode: 1084 score: 3.0 memory length: 208639	epsilon: 0.7848928000046698	s
teps: 228 lr: 4e-05 evaluation reward: 2.62	epsilon: 0.7846472800046751	s
episode: 1085 score: 0.0 memory length: 208763	epsilon: 0.7838097400046933	s
teps: 124 lr: 4e-05 evaluation reward: 2.58	epsilon: 0.7833187000047039	s
episode: 1086 score: 7.0 memory length: 209186	epsilon: 0.7827742000047158	s
teps: 423 lr: 4e-05 evaluation reward: 2.64	epsilon: 0.7824158200047235	s
episode: 1087 score: 3.0 memory length: 209434	epsilon: 0.7820812000047308	s
teps: 248 lr: 4e-05 evaluation reward: 2.65	epsilon: 0.7813446400047468	s
episode: 1088 score: 4.0 memory length: 209709	epsilon: 0.7808555800047574	s
teps: 275 lr: 4e-05 evaluation reward: 2.68	epsilon: 0.7801130800047735	s
episode: 1089 score: 2.0 memory length: 209890	epsilon: 0.7797190600047821	s
teps: 181 lr: 4e-05 evaluation reward: 2.68	epsilon: 0.77935474000479	ste
episode: 1090 score: 1.0 memory length: 210059	epsilon: 0.7789627000047985	s
teps: 169 lr: 4e-05 evaluation reward: 2.68	epsilon: 0.7784281000048101	s
episode: 1091 score: 6.0 memory length: 210431	epsilon: 0.7779430000048206	s
teps: 372 lr: 4e-05 evaluation reward: 2.73	epsilon: 0.7775074000048301	s
episode: 1092 score: 3.0 memory length: 210678	epsilon: 0.7769193400048429	s
teps: 247 lr: 4e-05 evaluation reward: 2.74	epsilon: 0.7765253200048514	s
episode: 1093 score: 6.0 memory length: 211053	epsilon: 0.7759333000048643	s
teps: 375 lr: 4e-05 evaluation reward: 2.78	epsilon: 0.7755194800048733	s
episode: 1094 score: 2.0 memory length: 211252	epsilon: 0.7751828800048806	s
teps: 199 lr: 4e-05 evaluation reward: 2.75	epsilon: 0.7745948200048933	s
episode: 1095 score: 2.0 memory length: 211436	epsilon: 0.7741612000049027	s
ps: 184 lr: 4e-05 evaluation reward: 2.74	epsilon: 0.7737256000049122	s
episode: 1096 score: 2.0 memory length: 211634	epsilon: 0.7733632600049201	s
teps: 198 lr: 4e-05 evaluation reward: 2.72	epsilon: 0.772810840004932	st
episode: 1097 score: 3.0 memory length: 211904	epsilon: 0.7723178200049428	s
teps: 270 lr: 4e-05 evaluation reward: 2.73	epsilon: 0.7716307600049577	s
episode: 1098 score: 3.0 memory length: 212149	epsilon: 0.7709793400049718	s
teps: 245 lr: 4e-05 evaluation reward: 2.73	epsilon: 0.770555620004981	st
episode: 1099 score: 2.0 memory length: 212369		
teps: 220 lr: 4e-05 evaluation reward: 2.73		
episode: 1100 score: 4.0 memory length: 212666		
teps: 297 lr: 4e-05 evaluation reward: 2.74		
episode: 1101 score: 2.0 memory length: 212865		
teps: 199 lr: 4e-05 evaluation reward: 2.74		
episode: 1102 score: 4.0 memory length: 213164		
teps: 299 lr: 4e-05 evaluation reward: 2.76		
episode: 1103 score: 3.0 memory length: 213373		
teps: 209 lr: 4e-05 evaluation reward: 2.77		
episode: 1104 score: 1.0 memory length: 213543		
teps: 170 lr: 4e-05 evaluation reward: 2.76		
episode: 1105 score: 4.0 memory length: 213840		
teps: 297 lr: 4e-05 evaluation reward: 2.75		
episode: 1106 score: 2.0 memory length: 214059		
teps: 219 lr: 4e-05 evaluation reward: 2.76		
episode: 1107 score: 2.0 memory length: 214279		
teps: 220 lr: 4e-05 evaluation reward: 2.74		
episode: 1108 score: 2.0 memory length: 214462		
teps: 183 lr: 4e-05 evaluation reward: 2.74		
episode: 1109 score: 4.0 memory length: 214741		
eps: 279 lr: 4e-05 evaluation reward: 2.75		
episode: 1110 score: 3.0 memory length: 214990		
teps: 249 lr: 4e-05 evaluation reward: 2.76		
episode: 1111 score: 6.0 memory length: 215337		
teps: 347 lr: 4e-05 evaluation reward: 2.81		
episode: 1112 score: 5.0 memory length: 215666		
teps: 329 lr: 4e-05 evaluation reward: 2.86		
episode: 1113 score: 3.0 memory length: 215880		
eps: 214 lr: 4e-05 evaluation reward: 2.87		

episode: 1114	score: 7.0	memory length: 216284	epsilon: 0.7697557000049984	s
teps: 404	lr: 4e-05	evaluation reward: 2.9	epsilon: 0.769359700005007	st
episode: 1115	score: 2.0	memory length: 216484	epsilon: 0.7690607200050135	s
eps: 200	lr: 4e-05	evaluation reward: 2.91	epsilon: 0.7685142400050253	s
episode: 1116	score: 1.0	memory length: 216635	epsilon: 0.7677281800050424	s
teps: 151	lr: 4e-05	evaluation reward: 2.88	epsilon: 0.7672648600050525	s
episode: 1117	score: 4.0	memory length: 216911	epsilon: 0.7668134200050623	s
teps: 276	lr: 4e-05	evaluation reward: 2.85	epsilon: 0.7665124600050688	s
episode: 1118	score: 7.0	memory length: 217308	epsilon: 0.7659263800050815	s
teps: 397	lr: 4e-05	evaluation reward: 2.9	epsilon: 0.765441280005092	st
episode: 1119	score: 3.0	memory length: 217542	epsilon: 0.7650472600051006	s
teps: 234	lr: 4e-05	evaluation reward: 2.91	epsilon: 0.7647126400051079	s
episode: 1120	score: 3.0	memory length: 217770	epsilon: 0.7641641800051198	s
teps: 228	lr: 4e-05	evaluation reward: 2.93	epsilon: 0.7637147200051295	s
episode: 1121	score: 1.0	memory length: 217922	epsilon: 0.7630890400051431	s
teps: 152	lr: 4e-05	evaluation reward: 2.92	epsilon: 0.7627900600051496	s
episode: 1122	score: 4.0	memory length: 218218	epsilon: 0.7623960400051581	s
teps: 296	lr: 4e-05	evaluation reward: 2.94	epsilon: 0.7620950800051647	s
episode: 1123	score: 3.0	memory length: 218463	epsilon: 0.7616971000051733	s
eps: 245	lr: 4e-05	evaluation reward: 2.94	epsilon: 0.7614515800051787	s
episode: 1124	score: 2.0	memory length: 218662	epsilon: 0.761206060005184	st
teps: 199	lr: 4e-05	evaluation reward: 2.94	epsilon: 0.7607071000051948	s
episode: 1125	score: 1.0	memory length: 218831	epsilon: 0.7601170600052076	s
teps: 169	lr: 4e-05	evaluation reward: 2.91	epsilon: 0.7594359400052224	s
episode: 1126	score: 4.0	memory length: 219108	epsilon: 0.7589092600052338	s
teps: 277	lr: 4e-05	evaluation reward: 2.93	epsilon: 0.7581410200052505	s
episode: 1127	score: 3.0	memory length: 219335	epsilon: 0.7574876200052647	s
teps: 227	lr: 4e-05	evaluation reward: 2.96	epsilon: 0.7566045400052839	s
episode: 1128	score: 5.0	memory length: 219651	epsilon: 0.7561511200052937	s
teps: 316	lr: 4e-05	evaluation reward: 2.99	epsilon: 0.755814520005301	st
episode: 1129	score: 1.0	memory length: 219802		
teps: 151	lr: 4e-05	evaluation reward: 2.96		
episode: 1130	score: 2.0	memory length: 220001		
teps: 199	lr: 4e-05	evaluation reward: 2.94		
episode: 1131	score: 1.0	memory length: 220153		
teps: 152	lr: 4e-05	evaluation reward: 2.93		
episode: 1132	score: 2.0	memory length: 220354		
teps: 201	lr: 4e-05	evaluation reward: 2.93		
episode: 1133	score: 0.0	memory length: 220478		
teps: 124	lr: 4e-05	evaluation reward: 2.9		
episode: 1134	score: 0.0	memory length: 220602		
eps: 124	lr: 4e-05	evaluation reward: 2.87		
episode: 1135	score: 3.0	memory length: 220854		
teps: 252	lr: 4e-05	evaluation reward: 2.88		
episode: 1136	score: 4.0	memory length: 221152		
teps: 298	lr: 4e-05	evaluation reward: 2.87		
episode: 1137	score: 5.0	memory length: 221496		
teps: 344	lr: 4e-05	evaluation reward: 2.89		
episode: 1138	score: 3.0	memory length: 221762		
teps: 266	lr: 4e-05	evaluation reward: 2.89		
episode: 1139	score: 7.0	memory length: 222150		
teps: 388	lr: 4e-05	evaluation reward: 2.95		
episode: 1140	score: 5.0	memory length: 222480		
teps: 330	lr: 4e-05	evaluation reward: 3.0		
episode: 1141	score: 7.0	memory length: 222926		
teps: 446	lr: 4e-05	evaluation reward: 2.99		
episode: 1142	score: 3.0	memory length: 223155		
teps: 229	lr: 4e-05	evaluation reward: 2.99		
episode: 1143	score: 1.0	memory length: 223325		
eps: 170	lr: 4e-05	evaluation reward: 3.0		

episode: 1144	score: 3.0	memory length: 223577	epsilon: 0.7553155600053119	s
teps: 252	lr: 4e-05	evaluation reward: 3.0	epsilon: 0.7547651200053238	s
episode: 1145	score: 4.0	memory length: 223855	epsilon: 0.7543077400053337	s
teps: 278	lr: 4e-05	evaluation reward: 3.03	epsilon: 0.7538780800053431	s
episode: 1146	score: 3.0	memory length: 224086	epsilon: 0.7533652600053542	s
teps: 231	lr: 4e-05	evaluation reward: 3.06	epsilon: 0.7527791800053669	s
episode: 1147	score: 2.0	memory length: 224303	epsilon: 0.7523317000053766	s
teps: 217	lr: 4e-05	evaluation reward: 3.05	epsilon: 0.7518406600053873	s
episode: 1148	score: 4.0	memory length: 224562	epsilon: 0.7513872400053971	s
teps: 259	lr: 4e-05	evaluation reward: 3.05	epsilon: 0.750933820005407	st
episode: 1149	score: 4.0	memory length: 224858	epsilon: 0.7501220200054246	s
teps: 296	lr: 4e-05	evaluation reward: 3.09	epsilon: 0.7498765000054299	s
episode: 1150	score: 3.0	memory length: 225084	epsilon: 0.7495399000054372	s
teps: 226	lr: 4e-05	evaluation reward: 3.04	epsilon: 0.7492389400054438	s
episode: 1151	score: 3.0	memory length: 225332	epsilon: 0.7489023400054511	s
teps: 248	lr: 4e-05	evaluation reward: 3.06	epsilon: 0.7484132800054617	s
episode: 1152	score: 3.0	memory length: 225561	epsilon: 0.7479202600054724	s
teps: 229	lr: 4e-05	evaluation reward: 3.08	epsilon: 0.7473935800054838	s
episode: 1153	score: 3.0	memory length: 225790	epsilon: 0.7470312400054917	s
eps: 229	lr: 4e-05	evaluation reward: 3.08	epsilon: 0.746785720005497	st
episode: 1154	score: 7.0	memory length: 226200	epsilon: 0.7461976600055098	s
teps: 410	lr: 4e-05	evaluation reward: 3.13	epsilon: 0.7456848400055209	s
episode: 1155	score: 0.0	memory length: 226324	epsilon: 0.7453838800055275	s
teps: 124	lr: 4e-05	evaluation reward: 3.07	epsilon: 0.744989860005536	st
episode: 1156	score: 1.0	memory length: 226494	epsilon: 0.7442711200055516	s
teps: 170	lr: 4e-05	evaluation reward: 3.07	epsilon: 0.7437444400055631	s
episode: 1157	score: 1.0	memory length: 226646	epsilon: 0.7431583600055758	s
teps: 152	lr: 4e-05	evaluation reward: 3.05	epsilon: 0.7424752600055906	s
episode: 1158	score: 1.0	memory length: 226816	epsilon: 0.7420198600056005	s
teps: 170	lr: 4e-05	evaluation reward: 3.02	epsilon: 0.7417169200056071	s
episode: 1159	score: 7.0	memory length: 227063		
teps: 247	lr: 4e-05	evaluation reward: 3.06		
episode: 1160	score: 3.0	memory length: 227312		
teps: 249	lr: 4e-05	evaluation reward: 3.07		
episode: 1161	score: 3.0	memory length: 227578		
teps: 266	lr: 4e-05	evaluation reward: 3.09		
episode: 1162	score: 2.0	memory length: 227761		
teps: 183	lr: 4e-05	evaluation reward: 3.07		
episode: 1163	score: 0.0	memory length: 227885		
eps: 124	lr: 4e-05	evaluation reward: 3.06		
episode: 1164	score: 4.0	memory length: 228182		
teps: 297	lr: 4e-05	evaluation reward: 3.08		
episode: 1165	score: 4.0	memory length: 228441		
teps: 259	lr: 4e-05	evaluation reward: 3.09		
episode: 1166	score: 1.0	memory length: 228593		
teps: 152	lr: 4e-05	evaluation reward: 3.1		
episode: 1167	score: 2.0	memory length: 228792		
eps: 199	lr: 4e-05	evaluation reward: 3.08		
episode: 1168	score: 5.0	memory length: 229155		
teps: 363	lr: 4e-05	evaluation reward: 3.09		
episode: 1169	score: 3.0	memory length: 229421		
teps: 266	lr: 4e-05	evaluation reward: 3.04		
episode: 1170	score: 4.0	memory length: 229717		
teps: 296	lr: 4e-05	evaluation reward: 3.04		
episode: 1171	score: 5.0	memory length: 230062		
teps: 345	lr: 4e-05	evaluation reward: 3.06		
episode: 1172	score: 3.0	memory length: 230292		
teps: 230	lr: 4e-05	evaluation reward: 3.08		
episode: 1173	score: 1.0	memory length: 230445		
teps: 153	lr: 4e-05	evaluation reward: 3.06		

episode: 1174	score: 3.0	memory length: 230672	epsilon: 0.7412674600056168	s
teps: 227	lr: 4e-05	evaluation reward: 3.06	epsilon: 0.7407388000056283	s
episode: 1175	score: 3.0	memory length: 230939	epsilon: 0.7403764600056362	s
teps: 267	lr: 4e-05	evaluation reward: 3.06	epsilon: 0.7399270000056459	s
episode: 1176	score: 2.0	memory length: 231122	epsilon: 0.7393330000056588	s
teps: 183	lr: 4e-05	evaluation reward: 3.05	epsilon: 0.7385390200056761	s
episode: 1177	score: 3.0	memory length: 231349	epsilon: 0.7380242200056872	s
teps: 227	lr: 4e-05	evaluation reward: 3.05	epsilon: 0.7375351600056979	s
episode: 1178	score: 4.0	memory length: 231649	epsilon: 0.7368520600057127	s
teps: 300	lr: 4e-05	evaluation reward: 3.01	epsilon: 0.7364580400057212	s
episode: 1179	score: 7.0	memory length: 232050	epsilon: 0.7359412600057325	s
teps: 401	lr: 4e-05	evaluation reward: 3.06	epsilon: 0.7354918000057422	s
episode: 1180	score: 4.0	memory length: 232310	epsilon: 0.7350997600057507	s
teps: 260	lr: 4e-05	evaluation reward: 3.1	epsilon: 0.7346661400057601	s
episode: 1181	score: 3.0	memory length: 232557	epsilon: 0.7342087600057701	s
teps: 247	lr: 4e-05	evaluation reward: 3.12	epsilon: 0.7336979200057812	s
episode: 1182	score: 5.0	memory length: 232902	epsilon: 0.7330999600057941	s
teps: 345	lr: 4e-05	evaluation reward: 3.14	epsilon: 0.7327059400058027	s
episode: 1183	score: 2.0	memory length: 233101	epsilon: 0.7322010400058137	s
teps: 199	lr: 4e-05	evaluation reward: 3.13	epsilon: 0.7318090000058222	s
episode: 1184	score: 4.0	memory length: 233362	epsilon: 0.731262520005834	st
teps: 261	lr: 4e-05	evaluation reward: 3.14	epsilon: 0.7305794200058489	s
episode: 1185	score: 3.0	memory length: 233589	epsilon: 0.7301458000058583	s
teps: 227	lr: 4e-05	evaluation reward: 3.17	epsilon: 0.7296904000058682	s
episode: 1186	score: 2.0	memory length: 233787	epsilon: 0.7293894400058747	s
teps: 198	lr: 4e-05	evaluation reward: 3.12	epsilon: 0.7289261200058847	s
episode: 1187	score: 2.0	memory length: 234006	epsilon: 0.7282766800058988	s
teps: 219	lr: 4e-05	evaluation reward: 3.11	epsilon: 0.727668820005912	st
episode: 1188	score: 3.0	memory length: 234237	epsilon: 0.7270154200059262	s
teps: 231	lr: 4e-05	evaluation reward: 3.1	epsilon: 0.7264313200059389	s
episode: 1189	score: 4.0	memory length: 234495		
teps: 258	lr: 4e-05	evaluation reward: 3.12		
episode: 1190	score: 4.0	memory length: 234797		
teps: 302	lr: 4e-05	evaluation reward: 3.15		
episode: 1191	score: 2.0	memory length: 234996		
teps: 199	lr: 4e-05	evaluation reward: 3.11		
episode: 1192	score: 4.0	memory length: 235251		
teps: 255	lr: 4e-05	evaluation reward: 3.12		
episode: 1193	score: 2.0	memory length: 235449		
teps: 198	lr: 4e-05	evaluation reward: 3.08		
episode: 1194	score: 4.0	memory length: 235725		
eps: 276	lr: 4e-05	evaluation reward: 3.1		
episode: 1195	score: 5.0	memory length: 236070		
teps: 345	lr: 4e-05	evaluation reward: 3.13		
episode: 1196	score: 2.0	memory length: 236289		
teps: 219	lr: 4e-05	evaluation reward: 3.13		
episode: 1197	score: 3.0	memory length: 236519		
teps: 230	lr: 4e-05	evaluation reward: 3.13		
episode: 1198	score: 1.0	memory length: 236671		
teps: 152	lr: 4e-05	evaluation reward: 3.11		
episode: 1199	score: 3.0	memory length: 236905		
teps: 234	lr: 4e-05	evaluation reward: 3.12		
episode: 1200	score: 5.0	memory length: 237233		
teps: 328	lr: 4e-05	evaluation reward: 3.13		
episode: 1201	score: 5.0	memory length: 237540		
eps: 307	lr: 4e-05	evaluation reward: 3.16		
episode: 1202	score: 8.0	memory length: 237870		
teps: 330	lr: 4e-05	evaluation reward: 3.2		
episode: 1203	score: 4.0	memory length: 238165		
teps: 295	lr: 4e-05	evaluation reward: 3.21		

episode: 1204	score: 1.0	memory length: 238337	epsilon: 0.7260907600059463	s
teps: 172	lr: 4e-05	evaluation reward: 3.21	epsilon: 0.7255383400059583	s
episode: 1205	score: 4.0	memory length: 238616	epsilon: 0.7249780000059705	s
teps: 279	lr: 4e-05	evaluation reward: 3.21	epsilon: 0.7243384600059843	s
episode: 1206	score: 4.0	memory length: 238899	epsilon: 0.7237860400059963	s
teps: 283	lr: 4e-05	evaluation reward: 3.23	epsilon: 0.7232890600060071	s
episode: 1207	score: 5.0	memory length: 239222	epsilon: 0.7226317000060214	s
teps: 323	lr: 4e-05	evaluation reward: 3.26	epsilon: 0.7220872000060332	s
episode: 1208	score: 4.0	memory length: 239501	epsilon: 0.7215743800060443	s
teps: 279	lr: 4e-05	evaluation reward: 3.28	epsilon: 0.7210536400060557	s
episode: 1209	score: 3.0	memory length: 239752	epsilon: 0.7205447800060667	s
teps: 251	lr: 4e-05	evaluation reward: 3.27	epsilon: 0.7202438200060732	s
episode: 1210	score: 5.0	memory length: 240084	epsilon: 0.7199092000060805	s
teps: 332	lr: 4e-05	evaluation reward: 3.29	epsilon: 0.7194597400060903	s
episode: 1211	score: 4.0	memory length: 240359	epsilon: 0.7190063200061001	s
teps: 275	lr: 4e-05	evaluation reward: 3.27	epsilon: 0.7183073800061153	s
episode: 1212	score: 4.0	memory length: 240618	epsilon: 0.717719320006128	st
teps: 259	lr: 4e-05	evaluation reward: 3.26	epsilon: 0.717075820006142	st
episode: 1213	score: 4.0	memory length: 240881	epsilon: 0.7164877600061548	s
teps: 263	lr: 4e-05	evaluation reward: 3.27	epsilon: 0.7158997000061675	s
episode: 1214	score: 4.0	memory length: 241138	epsilon: 0.7153809400061788	s
teps: 257	lr: 4e-05	evaluation reward: 3.24	epsilon: 0.7148245600061909	s
episode: 1215	score: 1.0	memory length: 241290	epsilon: 0.7144661800061987	s
teps: 152	lr: 4e-05	evaluation reward: 3.23	epsilon: 0.7140701800062073	s
episode: 1216	score: 1.0	memory length: 241459	epsilon: 0.7136761600062158	s
teps: 169	lr: 4e-05	evaluation reward: 3.23	epsilon: 0.7132841200062243	s
episode: 1217	score: 3.0	memory length: 241686	epsilon: 0.7128247600062343	s
teps: 227	lr: 4e-05	evaluation reward: 3.22	epsilon: 0.7120822600062504	s
episode: 1218	score: 3.0	memory length: 241915	epsilon: 0.7115357800062623	s
teps: 229	lr: 4e-05	evaluation reward: 3.18	epsilon: 0.7107457600062794	s
episode: 1219	score: 6.0	memory length: 242268		
teps: 353	lr: 4e-05	evaluation reward: 3.21		
episode: 1220	score: 4.0	memory length: 242565		
eps: 297	lr: 4e-05	evaluation reward: 3.22		
episode: 1221	score: 5.0	memory length: 242890		
eps: 325	lr: 4e-05	evaluation reward: 3.26		
episode: 1222	score: 4.0	memory length: 243187		
teps: 297	lr: 4e-05	evaluation reward: 3.26		
episode: 1223	score: 4.0	memory length: 243484		
teps: 297	lr: 4e-05	evaluation reward: 3.27		
episode: 1224	score: 4.0	memory length: 243746		
teps: 262	lr: 4e-05	evaluation reward: 3.29		
episode: 1225	score: 4.0	memory length: 244027		
teps: 281	lr: 4e-05	evaluation reward: 3.32		
episode: 1226	score: 2.0	memory length: 244208		
teps: 181	lr: 4e-05	evaluation reward: 3.3		
episode: 1227	score: 2.0	memory length: 244408		
teps: 200	lr: 4e-05	evaluation reward: 3.29		
episode: 1228	score: 2.0	memory length: 244607		
teps: 199	lr: 4e-05	evaluation reward: 3.26		
episode: 1229	score: 2.0	memory length: 244805		
teps: 198	lr: 4e-05	evaluation reward: 3.27		
episode: 1230	score: 3.0	memory length: 245037		
teps: 232	lr: 4e-05	evaluation reward: 3.28		
episode: 1231	score: 6.0	memory length: 245412		
teps: 375	lr: 4e-05	evaluation reward: 3.33		
episode: 1232	score: 4.0	memory length: 245688		
teps: 276	lr: 4e-05	evaluation reward: 3.35		
episode: 1233	score: 7.0	memory length: 246087		
teps: 399	lr: 4e-05	evaluation reward: 3.42		

episode: 1234	score: 3.0	memory length: 246336	epsilon: 0.7102527400062901	s
teps: 249	lr: 4e-05	evaluation reward: 3.45	epsilon: 0.7097181400063017	s
episode: 1235	score: 3.0	memory length: 246606	epsilon: 0.7094171800063083	s
teps: 270	lr: 4e-05	evaluation reward: 3.45	epsilon: 0.708740020006323	st
episode: 1236	score: 1.0	memory length: 246758	epsilon: 0.708185620006335	st
teps: 152	lr: 4e-05	evaluation reward: 3.42	epsilon: 0.7076074600063476	s
episode: 1237	score: 6.0	memory length: 247100	epsilon: 0.706942180006362	st
eps: 342	lr: 4e-05	evaluation reward: 3.43	epsilon: 0.706298680006376	st
episode: 1238	score: 4.0	memory length: 247380	epsilon: 0.7059403000063837	s
eps: 280	lr: 4e-05	evaluation reward: 3.44	epsilon: 0.7055066800063932	s
episode: 1239	score: 5.0	memory length: 247672	epsilon: 0.7048988200064064	s
teps: 292	lr: 4e-05	evaluation reward: 3.42	epsilon: 0.7043127400064191	s
episode: 1240	score: 5.0	memory length: 248008	epsilon: 0.7039781200064263	s
eps: 336	lr: 4e-05	evaluation reward: 3.42	epsilon: 0.7034316400064382	s
episode: 1241	score: 5.0	memory length: 248333	epsilon: 0.7030376200064468	s
eps: 325	lr: 4e-05	evaluation reward: 3.4	epsilon: 0.7025842000064566	s
episode: 1242	score: 2.0	memory length: 248514	epsilon: 0.7018912000064716	s
teps: 181	lr: 4e-05	evaluation reward: 3.39	epsilon: 0.7012833400064848	s
episode: 1243	score: 2.0	memory length: 248733	epsilon: 0.7005943000064998	s
teps: 219	lr: 4e-05	evaluation reward: 3.4	epsilon: 0.6999488200065138	s
episode: 1244	score: 5.0	memory length: 249040	epsilon: 0.699525100006523	st
teps: 307	lr: 4e-05	evaluation reward: 3.42	epsilon: 0.6990419800065335	s
episode: 1245	score: 4.0	memory length: 249336	epsilon: 0.6986479600065421	s
teps: 296	lr: 4e-05	evaluation reward: 3.42	epsilon: 0.6980579200065549	s
episode: 1246	score: 1.0	memory length: 249505	epsilon: 0.6972817600065717	
teps: 169	lr: 4e-05	evaluation reward: 3.4	epsilon: 0.6965075800065885	s
episode: 1247	score: 4.0	memory length: 249781	epsilon: 0.6959195200066013	s
teps: 276	lr: 4e-05	evaluation reward: 3.42	epsilon: 0.6954601600066113	s
episode: 1248	score: 2.0	memory length: 249980	epsilon: 0.6950305000066206	s
teps: 199	lr: 4e-05	evaluation reward: 3.4	epsilon: 0.6946345000066292	s
episode: 1249	score: 3.0	memory length: 250209		
teps: 229	lr: 4e-05	evaluation reward: 3.39		
episode: 1250	score: 5.0	memory length: 250559		
teps: 350	lr: 4e-05	evaluation reward: 3.41		
episode: 1251	score: 5.0	memory length: 250866		
teps: 307	lr: 4e-05	evaluation reward: 3.43		
episode: 1252	score: 5.0	memory length: 251214		
teps: 348	lr: 4e-05	evaluation reward: 3.45		
episode: 1253	score: 5.0	memory length: 251540		
teps: 326	lr: 4e-05	evaluation reward: 3.47		
episode: 1254	score: 3.0	memory length: 251754		
eps: 214	lr: 4e-05	evaluation reward: 3.43		
episode: 1255	score: 4.0	memory length: 251998		
teps: 244	lr: 4e-05	evaluation reward: 3.47		
episode: 1256	score: 2.0	memory length: 252197		
teps: 199	lr: 4e-05	evaluation reward: 3.48		
episode: 1257	score: 4.0	memory length: 252495		
teps: 298	lr: 4e-05	evaluation reward: 3.51		
episode: 1258	score: 10.0	memory length: 252887		
steps: 392	lr: 4e-05	evaluation reward: 3.6		
episode: 1259	score: 7.0	memory length: 253278		
teps: 391	lr: 4e-05	evaluation reward: 3.6		
episode: 1260	score: 4.0	memory length: 253575		
teps: 297	lr: 4e-05	evaluation reward: 3.61		
episode: 1261	score: 3.0	memory length: 253807		
teps: 232	lr: 4e-05	evaluation reward: 3.61		
episode: 1262	score: 2.0	memory length: 254024		
teps: 217	lr: 4e-05	evaluation reward: 3.61		
episode: 1263	score: 2.0	memory length: 254224		
teps: 200	lr: 4e-05	evaluation reward: 3.63		

episode: 1264	score: 5.0	memory length: 254535	epsilon: 0.6940187200066426	s
teps: 311	lr: 4e-05	evaluation reward: 3.64	epsilon: 0.6935613400066525	s
episode: 1265	score: 3.0	memory length: 254766	epsilon: 0.6928168600066686	s
teps: 231	lr: 4e-05	evaluation reward: 3.63	epsilon: 0.6922684000066806	s
episode: 1266	score: 6.0	memory length: 255142	epsilon: 0.6917773600066912	s
teps: 376	lr: 4e-05	evaluation reward: 3.68	epsilon: 0.6912289000067031	s
episode: 1267	score: 4.0	memory length: 255419	epsilon: 0.6904408600067202	s
teps: 277	lr: 4e-05	evaluation reward: 3.7	epsilon: 0.6896191600067381	s
episode: 1268	score: 3.0	memory length: 255667	epsilon: 0.689163760006748	st
teps: 248	lr: 4e-05	evaluation reward: 3.68	epsilon: 0.6885836200067605	s
episode: 1269	score: 4.0	memory length: 255944	epsilon: 0.6880609000067719	s
teps: 277	lr: 4e-05	evaluation reward: 3.69	epsilon: 0.6875124400067838	s
episode: 1270	score: 7.0	memory length: 256342	epsilon: 0.6870233800067944	s
teps: 398	lr: 4e-05	evaluation reward: 3.72	epsilon: 0.6863442400068092	s
episode: 1271	score: 6.0	memory length: 256757	epsilon: 0.6857007400068231	s
teps: 415	lr: 4e-05	evaluation reward: 3.73	epsilon: 0.6851483200068351	s
episode: 1272	score: 3.0	memory length: 256987	epsilon: 0.6845602600068479	s
eps: 230	lr: 4e-05	evaluation reward: 3.73	epsilon: 0.6842593000068544	s
episode: 1273	score: 5.0	memory length: 257280	epsilon: 0.6834791800068714	s
teps: 293	lr: 4e-05	evaluation reward: 3.77	epsilon: 0.6830871400068799	s
episode: 1274	score: 3.0	memory length: 257544	epsilon: 0.6824396800068939	s
teps: 264	lr: 4e-05	evaluation reward: 3.77	epsilon: 0.6820060600069033	s
episode: 1275	score: 4.0	memory length: 257821	epsilon: 0.681513040006914	st
teps: 277	lr: 4e-05	evaluation reward: 3.78	epsilon: 0.6810200200069247	s
episode: 1276	score: 3.0	memory length: 258068	epsilon: 0.6806616400069325	s
teps: 247	lr: 4e-05	evaluation reward: 3.79	epsilon: 0.6802003000069425	s
episode: 1277	score: 6.0	memory length: 258411	epsilon: 0.6797429200069525	s
teps: 343	lr: 4e-05	evaluation reward: 3.82	epsilon: 0.6789964600069687	s
episode: 1278	score: 5.0	memory length: 258736	epsilon: 0.678152980006987	st
teps: 325	lr: 4e-05	evaluation reward: 3.83	epsilon: 0.6775966000069991	s
episode: 1279	score: 4.0	memory length: 259015		
teps: 279	lr: 4e-05	evaluation reward: 3.8		
episode: 1280	score: 4.0	memory length: 259312		
teps: 297	lr: 4e-05	evaluation reward: 3.8		
episode: 1281	score: 1.0	memory length: 259464		
teps: 152	lr: 4e-05	evaluation reward: 3.78		
episode: 1282	score: 6.0	memory length: 259858		
teps: 394	lr: 4e-05	evaluation reward: 3.79		
episode: 1283	score: 2.0	memory length: 260056		
teps: 198	lr: 4e-05	evaluation reward: 3.79		
episode: 1284	score: 5.0	memory length: 260383		
teps: 327	lr: 4e-05	evaluation reward: 3.8		
episode: 1285	score: 2.0	memory length: 260602		
teps: 219	lr: 4e-05	evaluation reward: 3.79		
episode: 1286	score: 3.0	memory length: 260851		
eps: 249	lr: 4e-05	evaluation reward: 3.8		
episode: 1287	score: 3.0	memory length: 261100		
teps: 249	lr: 4e-05	evaluation reward: 3.81		
episode: 1288	score: 2.0	memory length: 261281		
teps: 181	lr: 4e-05	evaluation reward: 3.8		
episode: 1289	score: 3.0	memory length: 261514		
teps: 233	lr: 4e-05	evaluation reward: 3.79		
episode: 1290	score: 3.0	memory length: 261745		
teps: 231	lr: 4e-05	evaluation reward: 3.78		
episode: 1291	score: 7.0	memory length: 262122		
teps: 377	lr: 4e-05	evaluation reward: 3.83		
episode: 1292	score: 7.0	memory length: 262548		
eps: 426	lr: 4e-05	evaluation reward: 3.86		
episode: 1293	score: 4.0	memory length: 262829		
teps: 281	lr: 4e-05	evaluation reward: 3.88		

episode: 1294	score: 1.0	memory length: 262980	epsilon: 0.6772976200070056	s
teps: 151	lr: 4e-05	evaluation reward: 3.85	epsilon: 0.6767471800070175	s
episode: 1295	score: 4.0	memory length: 263258	epsilon: 0.676353160007026	st
teps: 278	lr: 4e-05	evaluation reward: 3.84	epsilon: 0.6757235200070397	s
episode: 1296	score: 2.0	memory length: 263457	epsilon: 0.6750780400070537	s
eps: 199	lr: 4e-05	evaluation reward: 3.84	epsilon: 0.6745315600070656	s
episode: 1297	score: 4.0	memory length: 263775	epsilon: 0.6739415200070784	s
teps: 318	lr: 4e-05	evaluation reward: 3.85	epsilon: 0.6735257200070874	s
episode: 1298	score: 9.0	memory length: 264101	epsilon: 0.6726941200071055	s
teps: 326	lr: 4e-05	evaluation reward: 3.93	epsilon: 0.6722723800071146	s
episode: 1299	score: 4.0	memory length: 264377	epsilon: 0.6716229400071287	s
teps: 276	lr: 4e-05	evaluation reward: 3.94	epsilon: 0.6709814200071427	s
episode: 1300	score: 4.0	memory length: 264675	epsilon: 0.6706468000071499	s
teps: 298	lr: 4e-05	evaluation reward: 3.93	epsilon: 0.6699716200071646	s
episode: 1301	score: 3.0	memory length: 264885	epsilon: 0.6693855400071773	s
teps: 210	lr: 4e-05	evaluation reward: 3.91	epsilon: 0.6689261800071873	s
episode: 1302	score: 7.0	memory length: 265305	epsilon: 0.6679599400072083	s
teps: 420	lr: 4e-05	evaluation reward: 3.9	epsilon: 0.6674708800072189	s
episode: 1303	score: 3.0	memory length: 265518	epsilon: 0.6669857800072294	s
teps: 213	lr: 4e-05	evaluation reward: 3.89	epsilon: 0.6665640400072386	s
episode: 1304	score: 5.0	memory length: 265846	epsilon: 0.6661700200072471	s
teps: 328	lr: 4e-05	evaluation reward: 3.93	epsilon: 0.6657205600072569	s
episode: 1305	score: 5.0	memory length: 266170	epsilon: 0.6651661600072689	s
teps: 324	lr: 4e-05	evaluation reward: 3.94	epsilon: 0.6645602800072821	s
episode: 1306	score: 1.0	memory length: 266339	epsilon: 0.6637148200073004	s
teps: 169	lr: 4e-05	evaluation reward: 3.91	epsilon: 0.6631327000073131	s
episode: 1307	score: 6.0	memory length: 266680	epsilon: 0.6627347200073217	s
teps: 341	lr: 4e-05	evaluation reward: 3.92	epsilon: 0.6619902400073379	s
episode: 1308	score: 5.0	memory length: 266976	epsilon: 0.6614120800073504	s
teps: 296	lr: 4e-05	evaluation reward: 3.93		
episode: 1309	score: 3.0	memory length: 267208		
teps: 232	lr: 4e-05	evaluation reward: 3.93		
episode: 1310	score: 9.0	memory length: 267696		
teps: 488	lr: 4e-05	evaluation reward: 3.97		
episode: 1311	score: 3.0	memory length: 267943		
teps: 247	lr: 4e-05	evaluation reward: 3.96		
episode: 1312	score: 3.0	memory length: 268188		
teps: 245	lr: 4e-05	evaluation reward: 3.95		
episode: 1313	score: 3.0	memory length: 268401		
teps: 213	lr: 4e-05	evaluation reward: 3.94		
episode: 1314	score: 2.0	memory length: 268600		
teps: 199	lr: 4e-05	evaluation reward: 3.92		
episode: 1315	score: 3.0	memory length: 268827		
teps: 227	lr: 4e-05	evaluation reward: 3.94		
episode: 1316	score: 4.0	memory length: 269107		
teps: 280	lr: 4e-05	evaluation reward: 3.97		
episode: 1317	score: 6.0	memory length: 269413		
teps: 306	lr: 4e-05	evaluation reward: 4.0		
episode: 1318	score: 3.0	memory length: 269628		
teps: 215	lr: 4e-05	evaluation reward: 4.0		
episode: 1319	score: 3.0	memory length: 269840		
teps: 212	lr: 4e-05	evaluation reward: 3.97		
episode: 1320	score: 5.0	memory length: 270134		
teps: 294	lr: 4e-05	evaluation reward: 3.98		
episode: 1321	score: 2.0	memory length: 270335		
teps: 201	lr: 4e-05	evaluation reward: 3.95		
episode: 1322	score: 6.0	memory length: 270711		
teps: 376	lr: 4e-05	evaluation reward: 3.97		
episode: 1323	score: 5.0	memory length: 271003		
teps: 292	lr: 4e-05	evaluation reward: 3.98		

episode: 1324 score: 5.0 memory length: 271351	epsilon: 0.6607230400073654	s
teps: 348 lr: 4e-05 evaluation reward: 3.99		
episode: 1325 score: 1.0 memory length: 271502	epsilon: 0.6604240600073719	s
teps: 151 lr: 4e-05 evaluation reward: 3.96		
episode: 1326 score: 5.0 memory length: 271827	epsilon: 0.6597805600073858	s
teps: 325 lr: 4e-05 evaluation reward: 3.99		
episode: 1327 score: 6.0 memory length: 272224	epsilon: 0.6589945000074029	s
teps: 397 lr: 4e-05 evaluation reward: 4.03		
episode: 1328 score: 2.0 memory length: 272424	epsilon: 0.6585985000074115	s
teps: 200 lr: 4e-05 evaluation reward: 4.03		
episode: 1329 score: 10.0 memory length: 272812	epsilon: 0.6578302600074282	
steps: 388 lr: 4e-05 evaluation reward: 4.11		
episode: 1330 score: 3.0 memory length: 273040	epsilon: 0.657378820007438	st
eps: 228 lr: 4e-05 evaluation reward: 4.11		
episode: 1331 score: 8.0 memory length: 273443	epsilon: 0.6565808800074553	s
teps: 403 lr: 4e-05 evaluation reward: 4.13		
episode: 1332 score: 4.0 memory length: 273719	epsilon: 0.6560344000074672	s
teps: 276 lr: 4e-05 evaluation reward: 4.13		
episode: 1333 score: 5.0 memory length: 274026	epsilon: 0.6554265400074804	s
teps: 307 lr: 4e-05 evaluation reward: 4.11		
episode: 1334 score: 2.0 memory length: 274244	epsilon: 0.6549949000074897	s
teps: 218 lr: 4e-05 evaluation reward: 4.1		
episode: 1335 score: 6.0 memory length: 274638	epsilon: 0.6542147800075067	s
teps: 394 lr: 4e-05 evaluation reward: 4.13		
episode: 1336 score: 5.0 memory length: 274943	epsilon: 0.6536108800075198	s
teps: 305 lr: 4e-05 evaluation reward: 4.17		
episode: 1337 score: 3.0 memory length: 275207	epsilon: 0.6530881600075311	s
teps: 264 lr: 4e-05 evaluation reward: 4.14		
episode: 1338 score: 4.0 memory length: 275505	epsilon: 0.6524981200075439	s
teps: 298 lr: 4e-05 evaluation reward: 4.14		
episode: 1339 score: 6.0 memory length: 275860	epsilon: 0.6517952200075592	s
teps: 355 lr: 4e-05 evaluation reward: 4.15		
episode: 1340 score: 4.0 memory length: 276142	epsilon: 0.6512368600075713	s
teps: 282 lr: 4e-05 evaluation reward: 4.14		
episode: 1341 score: 6.0 memory length: 276513	epsilon: 0.6505022800075873	s
teps: 371 lr: 4e-05 evaluation reward: 4.15		
episode: 1342 score: 2.0 memory length: 276733	epsilon: 0.6500666800075967	s
teps: 220 lr: 4e-05 evaluation reward: 4.15		
episode: 1343 score: 3.0 memory length: 276984	epsilon: 0.6495697000076075	s
teps: 251 lr: 4e-05 evaluation reward: 4.16		
episode: 1344 score: 6.0 memory length: 277360	epsilon: 0.6488252200076237	s
teps: 376 lr: 4e-05 evaluation reward: 4.17		
episode: 1345 score: 3.0 memory length: 277611	epsilon: 0.6483282400076344	s
teps: 251 lr: 4e-05 evaluation reward: 4.16		
episode: 1346 score: 3.0 memory length: 277844	epsilon: 0.6478669000076445	s
teps: 233 lr: 4e-05 evaluation reward: 4.18		
episode: 1347 score: 5.0 memory length: 278167	epsilon: 0.6472273600076583	s
teps: 323 lr: 4e-05 evaluation reward: 4.19		
episode: 1348 score: 7.0 memory length: 278607	epsilon: 0.6463561600076773	s
teps: 440 lr: 4e-05 evaluation reward: 4.24		
episode: 1349 score: 4.0 memory length: 278896	epsilon: 0.6457839400076897	s
teps: 289 lr: 4e-05 evaluation reward: 4.25		
episode: 1350 score: 6.0 memory length: 279292	epsilon: 0.6449998600077067	s
teps: 396 lr: 4e-05 evaluation reward: 4.26		
episode: 1351 score: 5.0 memory length: 279586	epsilon: 0.6444177400077193	s
teps: 294 lr: 4e-05 evaluation reward: 4.26		
episode: 1352 score: 3.0 memory length: 279834	epsilon: 0.64392670000773	ste
ps: 248 lr: 4e-05 evaluation reward: 4.24		
episode: 1353 score: 4.0 memory length: 280116	epsilon: 0.6433683400077421	s
teps: 282 lr: 4e-05 evaluation reward: 4.23		

episode: 1354	score: 6.0	memory length: 280472	epsilon: 0.6426634600077574	s
teps: 356	lr: 4e-05	evaluation reward: 4.26	epsilon: 0.6422417200077666	s
episode: 1355	score: 3.0	memory length: 280685	epsilon: 0.6415348600077819	s
teps: 213	lr: 4e-05	evaluation reward: 4.25	epsilon: 0.6410854000077917	s
episode: 1356	score: 6.0	memory length: 281042	epsilon: 0.6403706200078072	s
teps: 357	lr: 4e-05	evaluation reward: 4.29	epsilon: 0.639919180007817	st
episode: 1357	score: 3.0	memory length: 281269	epsilon: 0.6392954800078305	s
teps: 227	lr: 4e-05	evaluation reward: 4.28	epsilon: 0.6388400800078404	s
episode: 1358	score: 6.0	memory length: 281630	epsilon: 0.6384797200078483	s
teps: 361	lr: 4e-05	evaluation reward: 4.24	epsilon: 0.6377966200078631	s
episode: 1359	score: 3.0	memory length: 281858	epsilon: 0.6372105400078758	s
eps: 228	lr: 4e-05	evaluation reward: 4.2	epsilon: 0.6368482000078837	s
episode: 1360	score: 4.0	memory length: 282173	epsilon: 0.635911660007904	st
teps: 315	lr: 4e-05	evaluation reward: 4.2	epsilon: 0.6354780400079134	s
episode: 1361	score: 3.0	memory length: 282403	epsilon: 0.6347751400079287	s
teps: 230	lr: 4e-05	evaluation reward: 4.2	epsilon: 0.634163320007942	st
episode: 1362	score: 2.0	memory length: 282585	epsilon: 0.6335673400079549	s
teps: 182	lr: 4e-05	evaluation reward: 4.2	epsilon: 0.6332050000079628	s
episode: 1363	score: 5.0	memory length: 282930	epsilon: 0.6326486200079748	s
teps: 345	lr: 4e-05	evaluation reward: 4.23	epsilon: 0.6323476600079814	s
episode: 1364	score: 4.0	memory length: 283226	epsilon: 0.6318388000079924	s
teps: 296	lr: 4e-05	evaluation reward: 4.22	epsilon: 0.6310388800080098	s
episode: 1365	score: 2.0	memory length: 283409	epsilon: 0.630522100008021	st
teps: 183	lr: 4e-05	evaluation reward: 4.21	epsilon: 0.6301280800080296	s
episode: 1366	score: 8.0	memory length: 283882	epsilon: 0.6292351000080489	s
eps: 473	lr: 4e-05	evaluation reward: 4.23	epsilon: 0.6289895800080543	s
episode: 1367	score: 2.0	memory length: 284101	epsilon: 0.6284708200080655	s
teps: 219	lr: 4e-05	evaluation reward: 4.21	epsilon: 0.6276075400080843	s
episode: 1368	score: 6.0	memory length: 284456	epsilon: 0.6268690000081003	s
teps: 355	lr: 4e-05	evaluation reward: 4.24	epsilon: 0.6264532000081093	s
episode: 1369	score: 5.0	memory length: 284765		
eps: 309	lr: 4e-05	evaluation reward: 4.25		
episode: 1370	score: 4.0	memory length: 285066		
teps: 301	lr: 4e-05	evaluation reward: 4.22		
episode: 1371	score: 2.0	memory length: 285249		
teps: 183	lr: 4e-05	evaluation reward: 4.18		
episode: 1372	score: 4.0	memory length: 285530		
teps: 281	lr: 4e-05	evaluation reward: 4.19		
episode: 1373	score: 1.0	memory length: 285682		
teps: 152	lr: 4e-05	evaluation reward: 4.15		
episode: 1374	score: 4.0	memory length: 285939		
teps: 257	lr: 4e-05	evaluation reward: 4.16		
episode: 1375	score: 7.0	memory length: 286343		
teps: 404	lr: 4e-05	evaluation reward: 4.19		
episode: 1376	score: 4.0	memory length: 286604		
eps: 261	lr: 4e-05	evaluation reward: 4.2		
episode: 1377	score: 2.0	memory length: 286803		
teps: 199	lr: 4e-05	evaluation reward: 4.16		
episode: 1378	score: 9.0	memory length: 287254		
teps: 451	lr: 4e-05	evaluation reward: 4.2		
episode: 1379	score: 0.0	memory length: 287378		
teps: 124	lr: 4e-05	evaluation reward: 4.16		
episode: 1380	score: 4.0	memory length: 287640		
teps: 262	lr: 4e-05	evaluation reward: 4.16		
episode: 1381	score: 7.0	memory length: 288076		
teps: 436	lr: 4e-05	evaluation reward: 4.22		
episode: 1382	score: 6.0	memory length: 288449		
teps: 373	lr: 4e-05	evaluation reward: 4.22		
episode: 1383	score: 3.0	memory length: 288659		
teps: 210	lr: 4e-05	evaluation reward: 4.23		

episode: 1384 score: 4.0 memory length: 288919	epsilon: 0.6259384000081205	s
teps: 260 lr: 4e-05 evaluation reward: 4.22		
episode: 1385 score: 4.0 memory length: 289195	epsilon: 0.6253919200081324	s
teps: 276 lr: 4e-05 evaluation reward: 4.24		
episode: 1386 score: 3.0 memory length: 289424	epsilon: 0.6249385000081422	s
teps: 229 lr: 4e-05 evaluation reward: 4.24		
episode: 1387 score: 5.0 memory length: 289771	epsilon: 0.6242514400081571	s
teps: 347 lr: 4e-05 evaluation reward: 4.26		
episode: 1388 score: 5.0 memory length: 290077	epsilon: 0.6236455600081703	s
teps: 306 lr: 4e-05 evaluation reward: 4.29		
episode: 1389 score: 5.0 memory length: 290403	epsilon: 0.6230000800081843	s
teps: 326 lr: 4e-05 evaluation reward: 4.31		
episode: 1390 score: 5.0 memory length: 290710	epsilon: 0.6223922200081975	s
teps: 307 lr: 4e-05 evaluation reward: 4.33		
episode: 1391 score: 3.0 memory length: 290938	epsilon: 0.6219407800082073	s
teps: 228 lr: 4e-05 evaluation reward: 4.29		
episode: 1392 score: 7.0 memory length: 291329	epsilon: 0.6211666000082241	s
teps: 391 lr: 4e-05 evaluation reward: 4.29		
episode: 1393 score: 2.0 memory length: 291528	epsilon: 0.6207725800082327	s
teps: 199 lr: 4e-05 evaluation reward: 4.27		
episode: 1394 score: 6.0 memory length: 291913	epsilon: 0.6200102800082492	s
teps: 385 lr: 4e-05 evaluation reward: 4.32		
episode: 1395 score: 6.0 memory length: 292310	epsilon: 0.6192242200082663	s
teps: 397 lr: 4e-05 evaluation reward: 4.34		
episode: 1396 score: 8.0 memory length: 292759	epsilon: 0.6183352000082856	s
teps: 449 lr: 4e-05 evaluation reward: 4.4		
episode: 1397 score: 16.0 memory length: 293398	epsilon: 0.617069980008313	s
teps: 639 lr: 4e-05 evaluation reward: 4.52		
episode: 1398 score: 6.0 memory length: 293757	epsilon: 0.6163591600083285	s
teps: 359 lr: 4e-05 evaluation reward: 4.49		
episode: 1399 score: 7.0 memory length: 294163	epsilon: 0.6155552800083459	s
teps: 406 lr: 4e-05 evaluation reward: 4.52		
episode: 1400 score: 4.0 memory length: 294462	epsilon: 0.6149632600083588	s
teps: 299 lr: 4e-05 evaluation reward: 4.52		
episode: 1401 score: 6.0 memory length: 294842	epsilon: 0.6142108600083751	s
teps: 380 lr: 4e-05 evaluation reward: 4.55		
episode: 1402 score: 6.0 memory length: 295207	epsilon: 0.6134881600083908	s
teps: 365 lr: 4e-05 evaluation reward: 4.54		
episode: 1403 score: 12.0 memory length: 295729	epsilon: 0.6124546000084132	
steps: 522 lr: 4e-05 evaluation reward: 4.63		
episode: 1404 score: 5.0 memory length: 296014	epsilon: 0.6118903000084255	s
teps: 285 lr: 4e-05 evaluation reward: 4.63		
episode: 1405 score: 6.0 memory length: 296386	epsilon: 0.6111537400084415	s
teps: 372 lr: 4e-05 evaluation reward: 4.64		
episode: 1406 score: 5.0 memory length: 296731	epsilon: 0.6104706400084563	s
teps: 345 lr: 4e-05 evaluation reward: 4.68		
episode: 1407 score: 4.0 memory length: 296988	epsilon: 0.6099617800084673	s
teps: 257 lr: 4e-05 evaluation reward: 4.66		
episode: 1408 score: 6.0 memory length: 297345	epsilon: 0.6092549200084827	s
teps: 357 lr: 4e-05 evaluation reward: 4.67		
episode: 1409 score: 7.0 memory length: 297718	epsilon: 0.6085163800084987	s
teps: 373 lr: 4e-05 evaluation reward: 4.71		
episode: 1410 score: 5.0 memory length: 298028	epsilon: 0.607902580008512	st
eps: 310 lr: 4e-05 evaluation reward: 4.67		
episode: 1411 score: 3.0 memory length: 298255	epsilon: 0.6074531200085218	s
teps: 227 lr: 4e-05 evaluation reward: 4.67		
episode: 1412 score: 5.0 memory length: 298539	epsilon: 0.606890800008534	st
eps: 284 lr: 4e-05 evaluation reward: 4.69		
episode: 1413 score: 5.0 memory length: 298883	epsilon: 0.6062096800085488	s
teps: 344 lr: 4e-05 evaluation reward: 4.71		

episode: 1414 score: 4.0 memory length: 299125 epsilon: 0.6057305200085592 s
 teps: 242 lr: 4e-05 evaluation reward: 4.73
 episode: 1415 score: 4.0 memory length: 299402 epsilon: 0.6051820600085711 s
 teps: 277 lr: 4e-05 evaluation reward: 4.74
 episode: 1416 score: 8.0 memory length: 299827 epsilon: 0.6043405600085894 s
 teps: 425 lr: 4e-05 evaluation reward: 4.78
 episode: 1417 score: 4.0 memory length: 300083 epsilon: 0.6038336800086004 s
 teps: 256 lr: 1.6000000000000003e-05 evaluation reward: 4.76
 episode: 1418 score: 6.0 memory length: 300420 epsilon: 0.6031664200086149 s
 teps: 337 lr: 1.6000000000000003e-05 evaluation reward: 4.79
 episode: 1419 score: 4.0 memory length: 300678 epsilon: 0.602655580008626 st
 eps: 258 lr: 1.6000000000000003e-05 evaluation reward: 4.8
 episode: 1420 score: 7.0 memory length: 301110 epsilon: 0.6018002200086445 s
 teps: 432 lr: 1.6000000000000003e-05 evaluation reward: 4.82
 episode: 1421 score: 4.0 memory length: 301353 epsilon: 0.601319080008655 st
 eps: 243 lr: 1.6000000000000003e-05 evaluation reward: 4.84
 episode: 1422 score: 13.0 memory length: 301889 epsilon: 0.600257800008678 s
 teps: 536 lr: 1.6000000000000003e-05 evaluation reward: 4.91
 episode: 1423 score: 8.0 memory length: 302314 epsilon: 0.5994163000086963 s
 teps: 425 lr: 1.6000000000000003e-05 evaluation reward: 4.94
 episode: 1424 score: 3.0 memory length: 302541 epsilon: 0.598966840008706 st
 eps: 227 lr: 1.6000000000000003e-05 evaluation reward: 4.92
 episode: 1425 score: 6.0 memory length: 302901 epsilon: 0.5982540400087215 s
 teps: 360 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1426 score: 6.0 memory length: 303255 epsilon: 0.5975531200087367 s
 teps: 354 lr: 1.6000000000000003e-05 evaluation reward: 4.98
 episode: 1427 score: 5.0 memory length: 303581 epsilon: 0.5969076400087507 s
 teps: 326 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1428 score: 2.0 memory length: 303800 epsilon: 0.5964740200087602 s
 teps: 219 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1429 score: 1.0 memory length: 303952 epsilon: 0.5961730600087667 s
 teps: 152 lr: 1.6000000000000003e-05 evaluation reward: 4.88
 episode: 1430 score: 4.0 memory length: 304210 epsilon: 0.5956622200087778 s
 teps: 258 lr: 1.6000000000000003e-05 evaluation reward: 4.89
 episode: 1431 score: 4.0 memory length: 304504 epsilon: 0.5950801000087904 s
 teps: 294 lr: 1.6000000000000003e-05 evaluation reward: 4.85
 episode: 1432 score: 8.0 memory length: 304946 epsilon: 0.5942049400088094 s
 teps: 442 lr: 1.6000000000000003e-05 evaluation reward: 4.89
 episode: 1433 score: 6.0 memory length: 305302 epsilon: 0.5935000600088247 s
 teps: 356 lr: 1.6000000000000003e-05 evaluation reward: 4.9
 episode: 1434 score: 9.0 memory length: 305767 epsilon: 0.5925793600088447 s
 teps: 465 lr: 1.6000000000000003e-05 evaluation reward: 4.97
 episode: 1435 score: 5.0 memory length: 306077 epsilon: 0.591965560008858 st
 eps: 310 lr: 1.6000000000000003e-05 evaluation reward: 4.96
 episode: 1436 score: 4.0 memory length: 306356 epsilon: 0.59141314000887 ste
 ps: 279 lr: 1.6000000000000003e-05 evaluation reward: 4.95
 episode: 1437 score: 9.0 memory length: 306810 epsilon: 0.5905142200088895 s
 teps: 454 lr: 1.6000000000000003e-05 evaluation reward: 5.01
 episode: 1438 score: 5.0 memory length: 307119 epsilon: 0.5899024000089028 s
 teps: 309 lr: 1.6000000000000003e-05 evaluation reward: 5.02
 episode: 1439 score: 4.0 memory length: 307396 epsilon: 0.5893539400089147 s
 teps: 277 lr: 1.6000000000000003e-05 evaluation reward: 5.0
 episode: 1440 score: 9.0 memory length: 307838 epsilon: 0.5884787800089337 s
 teps: 442 lr: 1.6000000000000003e-05 evaluation reward: 5.05
 episode: 1441 score: 4.0 memory length: 308081 epsilon: 0.5879976400089442 s
 teps: 243 lr: 1.6000000000000003e-05 evaluation reward: 5.03
 episode: 1442 score: 8.0 memory length: 308534 epsilon: 0.5871007000089636 s
 teps: 453 lr: 1.6000000000000003e-05 evaluation reward: 5.09
 episode: 1443 score: 5.0 memory length: 308825 epsilon: 0.5865245200089761 s
 teps: 291 lr: 1.6000000000000003e-05 evaluation reward: 5.11

episode: 1444 score: 7.0 memory length: 309231 epsilon: 0.5857206400089936 s
 teps: 406 lr: 1.6000000000000003e-05 evaluation reward: 5.12
 episode: 1445 score: 5.0 memory length: 309538 epsilon: 0.5851127800090068 s
 teps: 307 lr: 1.6000000000000003e-05 evaluation reward: 5.14
 episode: 1446 score: 7.0 memory length: 309948 epsilon: 0.5843009800090244 s
 teps: 410 lr: 1.6000000000000003e-05 evaluation reward: 5.18
 episode: 1447 score: 4.0 memory length: 310243 epsilon: 0.5837168800090371 s
 teps: 295 lr: 1.6000000000000003e-05 evaluation reward: 5.17
 episode: 1448 score: 6.0 memory length: 310614 epsilon: 0.582982300009053 st
 eps: 371 lr: 1.6000000000000003e-05 evaluation reward: 5.16
 episode: 1449 score: 6.0 memory length: 310972 epsilon: 0.5822734600090684 s
 teps: 358 lr: 1.6000000000000003e-05 evaluation reward: 5.18
 episode: 1450 score: 4.0 memory length: 311246 epsilon: 0.5817309400090802 s
 teps: 274 lr: 1.6000000000000003e-05 evaluation reward: 5.16
 episode: 1451 score: 7.0 memory length: 311669 epsilon: 0.5808934000090984 s
 teps: 423 lr: 1.6000000000000003e-05 evaluation reward: 5.18
 episode: 1452 score: 4.0 memory length: 311964 epsilon: 0.5803093000091111 s
 teps: 295 lr: 1.6000000000000003e-05 evaluation reward: 5.19
 episode: 1453 score: 5.0 memory length: 312293 epsilon: 0.5796578800091252 s
 teps: 329 lr: 1.6000000000000003e-05 evaluation reward: 5.2
 episode: 1454 score: 5.0 memory length: 312609 epsilon: 0.5790322000091388 s
 teps: 316 lr: 1.6000000000000003e-05 evaluation reward: 5.19
 episode: 1455 score: 5.0 memory length: 312954 epsilon: 0.5783491000091536 s
 teps: 345 lr: 1.6000000000000003e-05 evaluation reward: 5.21
 episode: 1456 score: 6.0 memory length: 313330 epsilon: 0.5776046200091698 s
 teps: 376 lr: 1.6000000000000003e-05 evaluation reward: 5.21
 episode: 1457 score: 5.0 memory length: 313606 epsilon: 0.5770581400091817 s
 teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 5.23
 episode: 1458 score: 6.0 memory length: 313961 epsilon: 0.5763552400091969 s
 teps: 355 lr: 1.6000000000000003e-05 evaluation reward: 5.23
 episode: 1459 score: 4.0 memory length: 314203 epsilon: 0.5758760800092073 s
 teps: 242 lr: 1.6000000000000003e-05 evaluation reward: 5.24
 episode: 1460 score: 4.0 memory length: 314483 epsilon: 0.5753216800092193 s
 teps: 280 lr: 1.6000000000000003e-05 evaluation reward: 5.24
 episode: 1461 score: 6.0 memory length: 314824 epsilon: 0.574646500009234 st
 eps: 341 lr: 1.6000000000000003e-05 evaluation reward: 5.27
 episode: 1462 score: 12.0 memory length: 315298 epsilon: 0.5737079800092544
 steps: 474 lr: 1.6000000000000003e-05 evaluation reward: 5.37
 episode: 1463 score: 4.0 memory length: 315541 epsilon: 0.5732268400092648 s
 teps: 243 lr: 1.6000000000000003e-05 evaluation reward: 5.36
 episode: 1464 score: 5.0 memory length: 315867 epsilon: 0.5725813600092788 s
 teps: 326 lr: 1.6000000000000003e-05 evaluation reward: 5.37
 episode: 1465 score: 6.0 memory length: 316209 epsilon: 0.5719042000092935 s
 teps: 342 lr: 1.6000000000000003e-05 evaluation reward: 5.41
 episode: 1466 score: 5.0 memory length: 316504 epsilon: 0.5713201000093062 s
 teps: 295 lr: 1.6000000000000003e-05 evaluation reward: 5.38
 episode: 1467 score: 5.0 memory length: 316778 epsilon: 0.570777580009318 st
 eps: 274 lr: 1.6000000000000003e-05 evaluation reward: 5.41
 episode: 1468 score: 9.0 memory length: 317283 epsilon: 0.5697776800093397 s
 teps: 505 lr: 1.6000000000000003e-05 evaluation reward: 5.44
 episode: 1469 score: 7.0 memory length: 317670 epsilon: 0.5690114200093563 s
 teps: 387 lr: 1.6000000000000003e-05 evaluation reward: 5.46
 episode: 1470 score: 4.0 memory length: 317966 epsilon: 0.5684253400093691 s
 teps: 296 lr: 1.6000000000000003e-05 evaluation reward: 5.46
 episode: 1471 score: 8.0 memory length: 318386 epsilon: 0.5675937400093871 s
 teps: 420 lr: 1.6000000000000003e-05 evaluation reward: 5.52
 episode: 1472 score: 9.0 memory length: 318842 epsilon: 0.5666908600094067 s
 teps: 456 lr: 1.6000000000000003e-05 evaluation reward: 5.57
 episode: 1473 score: 8.0 memory length: 319269 epsilon: 0.5658454000094251 s
 teps: 427 lr: 1.6000000000000003e-05 evaluation reward: 5.64

episode: 1474 score: 10.0 memory length: 319783 epsilon: 0.5648276800094472
 steps: 514 lr: 1.6000000000000003e-05 evaluation reward: 5.7
 episode: 1475 score: 6.0 memory length: 320105 epsilon: 0.564190120009461 st
 eps: 322 lr: 1.6000000000000003e-05 evaluation reward: 5.69
 episode: 1476 score: 4.0 memory length: 320363 epsilon: 0.5636792800094721 s
 teps: 258 lr: 1.6000000000000003e-05 evaluation reward: 5.69
 episode: 1477 score: 5.0 memory length: 320648 epsilon: 0.5631149800094843 s
 teps: 285 lr: 1.6000000000000003e-05 evaluation reward: 5.72
 episode: 1478 score: 5.0 memory length: 320937 epsilon: 0.5625427600094968 s
 teps: 289 lr: 1.6000000000000003e-05 evaluation reward: 5.68
 episode: 1479 score: 5.0 memory length: 321286 epsilon: 0.5618517400095118 s
 teps: 349 lr: 1.6000000000000003e-05 evaluation reward: 5.73
 episode: 1480 score: 7.0 memory length: 321690 epsilon: 0.5610518200095291 s
 teps: 404 lr: 1.6000000000000003e-05 evaluation reward: 5.76
 episode: 1481 score: 3.0 memory length: 321901 epsilon: 0.5606340400095382 s
 teps: 211 lr: 1.6000000000000003e-05 evaluation reward: 5.72
 episode: 1482 score: 5.0 memory length: 322247 epsilon: 0.5599489600095531 s
 teps: 346 lr: 1.6000000000000003e-05 evaluation reward: 5.71
 episode: 1483 score: 6.0 memory length: 322639 epsilon: 0.5591728000095699 s
 teps: 392 lr: 1.6000000000000003e-05 evaluation reward: 5.74
 episode: 1484 score: 9.0 memory length: 323104 epsilon: 0.5582521000095899 s
 teps: 465 lr: 1.6000000000000003e-05 evaluation reward: 5.79
 episode: 1485 score: 5.0 memory length: 323414 epsilon: 0.5576383000096032 s
 teps: 310 lr: 1.6000000000000003e-05 evaluation reward: 5.8
 episode: 1486 score: 6.0 memory length: 323741 epsilon: 0.5569908400096173 s
 teps: 327 lr: 1.6000000000000003e-05 evaluation reward: 5.83
 episode: 1487 score: 6.0 memory length: 324133 epsilon: 0.5562146800096341 s
 teps: 392 lr: 1.6000000000000003e-05 evaluation reward: 5.84
 episode: 1488 score: 7.0 memory length: 324495 epsilon: 0.5554979200096497 s
 teps: 362 lr: 1.6000000000000003e-05 evaluation reward: 5.86
 episode: 1489 score: 16.0 memory length: 325149 epsilon: 0.5542030000096778
 steps: 654 lr: 1.6000000000000003e-05 evaluation reward: 5.97
 episode: 1490 score: 6.0 memory length: 325506 epsilon: 0.5534961400096932 s
 teps: 357 lr: 1.6000000000000003e-05 evaluation reward: 5.98
 episode: 1491 score: 9.0 memory length: 325940 epsilon: 0.5526368200097118 s
 teps: 434 lr: 1.6000000000000003e-05 evaluation reward: 6.04
 episode: 1492 score: 4.0 memory length: 326200 epsilon: 0.552122020009723 st
 eps: 260 lr: 1.6000000000000003e-05 evaluation reward: 6.01
 episode: 1493 score: 6.0 memory length: 326523 epsilon: 0.5514824800097369 s
 teps: 323 lr: 1.6000000000000003e-05 evaluation reward: 6.05
 episode: 1494 score: 5.0 memory length: 326835 epsilon: 0.5508647200097503 s
 teps: 312 lr: 1.6000000000000003e-05 evaluation reward: 6.04
 episode: 1495 score: 5.0 memory length: 327179 epsilon: 0.5501836000097651 s
 teps: 344 lr: 1.6000000000000003e-05 evaluation reward: 6.03
 episode: 1496 score: 6.0 memory length: 327533 epsilon: 0.5494826800097803 s
 teps: 354 lr: 1.6000000000000003e-05 evaluation reward: 6.01
 episode: 1497 score: 3.0 memory length: 327780 epsilon: 0.5489936200097909 s
 teps: 247 lr: 1.6000000000000003e-05 evaluation reward: 5.88
 episode: 1498 score: 3.0 memory length: 328007 epsilon: 0.5485441600098007 s
 teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 5.85
 episode: 1499 score: 5.0 memory length: 328333 epsilon: 0.5478986800098147 s
 teps: 326 lr: 1.6000000000000003e-05 evaluation reward: 5.83
 episode: 1500 score: 12.0 memory length: 328953 epsilon: 0.5466710800098413
 steps: 620 lr: 1.6000000000000003e-05 evaluation reward: 5.91
 episode: 1501 score: 5.0 memory length: 329282 epsilon: 0.5460196600098555 s
 teps: 329 lr: 1.6000000000000003e-05 evaluation reward: 5.9
 episode: 1502 score: 7.0 memory length: 329699 epsilon: 0.5451940000098734 s
 teps: 417 lr: 1.6000000000000003e-05 evaluation reward: 5.91
 episode: 1503 score: 3.0 memory length: 329945 epsilon: 0.544706920009884 st
 eps: 246 lr: 1.6000000000000003e-05 evaluation reward: 5.82

episode: 1504 score: 2.0 memory length: 330127 epsilon: 0.5443465600098918 s
 teps: 182 lr: 1.6000000000000003e-05 evaluation reward: 5.79
 episode: 1505 score: 7.0 memory length: 330517 epsilon: 0.5435743600099086 s
 teps: 390 lr: 1.6000000000000003e-05 evaluation reward: 5.8
 episode: 1506 score: 1.0 memory length: 330689 epsilon: 0.543233800009916 st
 eps: 172 lr: 1.6000000000000003e-05 evaluation reward: 5.76
 episode: 1507 score: 5.0 memory length: 331014 epsilon: 0.5425903000099299 s
 teps: 325 lr: 1.6000000000000003e-05 evaluation reward: 5.77
 episode: 1508 score: 5.0 memory length: 331322 epsilon: 0.5419804600099432 s
 teps: 308 lr: 1.6000000000000003e-05 evaluation reward: 5.76
 episode: 1509 score: 6.0 memory length: 331718 epsilon: 0.5411963800099602 s
 teps: 396 lr: 1.6000000000000003e-05 evaluation reward: 5.75
 episode: 1510 score: 4.0 memory length: 331995 epsilon: 0.5406479200099721 s
 teps: 277 lr: 1.6000000000000003e-05 evaluation reward: 5.74
 episode: 1511 score: 1.0 memory length: 332167 epsilon: 0.5403073600099795 s
 teps: 172 lr: 1.6000000000000003e-05 evaluation reward: 5.72
 episode: 1512 score: 6.0 memory length: 332526 epsilon: 0.5395965400099949 s
 teps: 359 lr: 1.6000000000000003e-05 evaluation reward: 5.73
 episode: 1513 score: 2.0 memory length: 332708 epsilon: 0.5392361800100027 s
 teps: 182 lr: 1.6000000000000003e-05 evaluation reward: 5.7
 episode: 1514 score: 4.0 memory length: 333004 epsilon: 0.5386501000100155 s
 teps: 296 lr: 1.6000000000000003e-05 evaluation reward: 5.7
 episode: 1515 score: 4.0 memory length: 333246 epsilon: 0.5381709400100259 s
 teps: 242 lr: 1.6000000000000003e-05 evaluation reward: 5.7
 episode: 1516 score: 7.0 memory length: 333633 epsilon: 0.5374046800100425 s
 teps: 387 lr: 1.6000000000000003e-05 evaluation reward: 5.69
 episode: 1517 score: 7.0 memory length: 334036 epsilon: 0.5366067400100598 s
 teps: 403 lr: 1.6000000000000003e-05 evaluation reward: 5.72
 episode: 1518 score: 6.0 memory length: 334432 epsilon: 0.5358226600100768 s
 teps: 396 lr: 1.6000000000000003e-05 evaluation reward: 5.72
 episode: 1519 score: 10.0 memory length: 334921 epsilon: 0.5348544400100979 s
 steps: 489 lr: 1.6000000000000003e-05 evaluation reward: 5.78
 episode: 1520 score: 6.0 memory length: 335264 epsilon: 0.5341753000101126 s
 teps: 343 lr: 1.6000000000000003e-05 evaluation reward: 5.77
 episode: 1521 score: 7.0 memory length: 335649 epsilon: 0.5334130000101291 s
 teps: 385 lr: 1.6000000000000003e-05 evaluation reward: 5.8
 episode: 1522 score: 13.0 memory length: 336169 epsilon: 0.5323834000101515 s
 steps: 520 lr: 1.6000000000000003e-05 evaluation reward: 5.8
 episode: 1523 score: 5.0 memory length: 336477 epsilon: 0.5317735600101647 s
 teps: 308 lr: 1.6000000000000003e-05 evaluation reward: 5.77
 episode: 1524 score: 9.0 memory length: 336915 epsilon: 0.5309063200101836 s
 teps: 438 lr: 1.6000000000000003e-05 evaluation reward: 5.83
 episode: 1525 score: 13.0 memory length: 337428 epsilon: 0.5298905800102056 s
 steps: 513 lr: 1.6000000000000003e-05 evaluation reward: 5.9
 episode: 1526 score: 4.0 memory length: 337706 epsilon: 0.5293401400102176 s
 teps: 278 lr: 1.6000000000000003e-05 evaluation reward: 5.88
 episode: 1527 score: 4.0 memory length: 337982 epsilon: 0.5287936600102294 s
 teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 5.87
 episode: 1528 score: 8.0 memory length: 338293 epsilon: 0.5281778800102428 s
 teps: 311 lr: 1.6000000000000003e-05 evaluation reward: 5.93
 episode: 1529 score: 4.0 memory length: 338534 epsilon: 0.5277007000102532 s
 teps: 241 lr: 1.6000000000000003e-05 evaluation reward: 5.96
 episode: 1530 score: 6.0 memory length: 338908 epsilon: 0.5269601800102692 s
 teps: 374 lr: 1.6000000000000003e-05 evaluation reward: 5.98
 episode: 1531 score: 6.0 memory length: 339236 epsilon: 0.5263107400102833 s
 teps: 328 lr: 1.6000000000000003e-05 evaluation reward: 6.0
 episode: 1532 score: 8.0 memory length: 339664 epsilon: 0.5254633000103017 s
 teps: 428 lr: 1.6000000000000003e-05 evaluation reward: 6.0
 episode: 1533 score: 8.0 memory length: 340055 epsilon: 0.5246891200103185 s
 teps: 391 lr: 1.6000000000000003e-05 evaluation reward: 6.02

episode: 1534 score: 5.0 memory length: 340387 epsilon: 0.5240317600103328 s
 teps: 332 lr: 1.6000000000000003e-05 evaluation reward: 5.98
 episode: 1535 score: 9.0 memory length: 340857 epsilon: 0.523101160010353 st
 eps: 470 lr: 1.6000000000000003e-05 evaluation reward: 6.02
 episode: 1536 score: 13.0 memory length: 341415 epsilon: 0.521996320010377 s
 teps: 558 lr: 1.6000000000000003e-05 evaluation reward: 6.11
 episode: 1537 score: 9.0 memory length: 341878 epsilon: 0.5210795800103969 s
 teps: 463 lr: 1.6000000000000003e-05 evaluation reward: 6.11
 episode: 1538 score: 5.0 memory length: 342169 epsilon: 0.5205034000104094 s
 teps: 291 lr: 1.6000000000000003e-05 evaluation reward: 6.11
 episode: 1539 score: 6.0 memory length: 342507 epsilon: 0.5198341600104239 s
 teps: 338 lr: 1.6000000000000003e-05 evaluation reward: 6.13
 episode: 1540 score: 4.0 memory length: 342784 epsilon: 0.5192857000104358 s
 teps: 277 lr: 1.6000000000000003e-05 evaluation reward: 6.08
 episode: 1541 score: 7.0 memory length: 343165 epsilon: 0.5185313200104522 s
 teps: 381 lr: 1.6000000000000003e-05 evaluation reward: 6.11
 episode: 1542 score: 8.0 memory length: 343613 epsilon: 0.5176442800104715 s
 teps: 448 lr: 1.6000000000000003e-05 evaluation reward: 6.11
 episode: 1543 score: 11.0 memory length: 344191 epsilon: 0.5164998400104963 s
 steps: 578 lr: 1.6000000000000003e-05 evaluation reward: 6.17
 episode: 1544 score: 4.0 memory length: 344468 epsilon: 0.5159513800105082 s
 teps: 277 lr: 1.6000000000000003e-05 evaluation reward: 6.14
 episode: 1545 score: 15.0 memory length: 345022 epsilon: 0.514854460010532 s
 teps: 554 lr: 1.6000000000000003e-05 evaluation reward: 6.24
 episode: 1546 score: 4.0 memory length: 345283 epsilon: 0.5143376800105433 s
 teps: 261 lr: 1.6000000000000003e-05 evaluation reward: 6.21
 episode: 1547 score: 6.0 memory length: 345661 epsilon: 0.5135892400105595 s
 teps: 378 lr: 1.6000000000000003e-05 evaluation reward: 6.23
 episode: 1548 score: 6.0 memory length: 346014 epsilon: 0.5128903000105747 s
 teps: 353 lr: 1.6000000000000003e-05 evaluation reward: 6.23
 episode: 1549 score: 6.0 memory length: 346370 epsilon: 0.51218542001059 ste
 ps: 356 lr: 1.6000000000000003e-05 evaluation reward: 6.23
 episode: 1550 score: 10.0 memory length: 346859 epsilon: 0.511217200010611 s
 teps: 489 lr: 1.6000000000000003e-05 evaluation reward: 6.29
 episode: 1551 score: 2.0 memory length: 347041 epsilon: 0.5108568400106188 s
 teps: 182 lr: 1.6000000000000003e-05 evaluation reward: 6.24
 episode: 1552 score: 5.0 memory length: 347338 epsilon: 0.5102687800106316 s
 teps: 297 lr: 1.6000000000000003e-05 evaluation reward: 6.25
 episode: 1553 score: 8.0 memory length: 347733 epsilon: 0.5094866800106486 s
 teps: 395 lr: 1.6000000000000003e-05 evaluation reward: 6.28
 episode: 1554 score: 3.0 memory length: 347962 epsilon: 0.5090332600106584 s
 teps: 229 lr: 1.6000000000000003e-05 evaluation reward: 6.26
 episode: 1555 score: 3.0 memory length: 348173 epsilon: 0.5086154800106675 s
 teps: 211 lr: 1.6000000000000003e-05 evaluation reward: 6.24
 episode: 1556 score: 11.0 memory length: 348719 epsilon: 0.507534400010691 s
 teps: 546 lr: 1.6000000000000003e-05 evaluation reward: 6.29
 episode: 1557 score: 8.0 memory length: 349180 epsilon: 0.5066216200107108 s
 teps: 461 lr: 1.6000000000000003e-05 evaluation reward: 6.32
 episode: 1558 score: 3.0 memory length: 349425 epsilon: 0.5061365200107213 s
 teps: 245 lr: 1.6000000000000003e-05 evaluation reward: 6.29
 episode: 1559 score: 4.0 memory length: 349686 epsilon: 0.5056197400107325 s
 teps: 261 lr: 1.6000000000000003e-05 evaluation reward: 6.29
 episode: 1560 score: 9.0 memory length: 350146 epsilon: 0.5047089400107523 s
 teps: 460 lr: 1.6000000000000003e-05 evaluation reward: 6.34
 episode: 1561 score: 6.0 memory length: 350504 epsilon: 0.5040001000107677 s
 teps: 358 lr: 1.6000000000000003e-05 evaluation reward: 6.34
 episode: 1562 score: 8.0 memory length: 350955 epsilon: 0.5031071200107871 s
 teps: 451 lr: 1.6000000000000003e-05 evaluation reward: 6.3
 episode: 1563 score: 8.0 memory length: 351379 epsilon: 0.5022676000108053 s
 teps: 424 lr: 1.6000000000000003e-05 evaluation reward: 6.34

episode: 1564 score: 5.0 memory length: 351705 epsilon: 0.5016221200108193 s
 teps: 326 lr: 1.6000000000000003e-05 evaluation reward: 6.34
 episode: 1565 score: 6.0 memory length: 352080 epsilon: 0.5008796200108354 s
 teps: 375 lr: 1.6000000000000003e-05 evaluation reward: 6.34
 episode: 1566 score: 10.0 memory length: 352591 epsilon: 0.49986784001085366
 steps: 511 lr: 1.6000000000000003e-05 evaluation reward: 6.39
 episode: 1567 score: 4.0 memory length: 352850 epsilon: 0.4993550200108504 s
 teps: 259 lr: 1.6000000000000003e-05 evaluation reward: 6.38
 episode: 1568 score: 8.0 memory length: 353308 epsilon: 0.4984481800108447 s
 teps: 458 lr: 1.6000000000000003e-05 evaluation reward: 6.37
 episode: 1569 score: 6.0 memory length: 353668 epsilon: 0.49773538001084017
 steps: 360 lr: 1.6000000000000003e-05 evaluation reward: 6.36
 episode: 1570 score: 10.0 memory length: 354239 epsilon: 0.496604800010833 s
 teps: 571 lr: 1.6000000000000003e-05 evaluation reward: 6.42
 episode: 1571 score: 6.0 memory length: 354615 epsilon: 0.4958603200108283 s
 teps: 376 lr: 1.6000000000000003e-05 evaluation reward: 6.4
 episode: 1572 score: 8.0 memory length: 355034 epsilon: 0.49503070001082305
 steps: 419 lr: 1.6000000000000003e-05 evaluation reward: 6.39
 episode: 1573 score: 5.0 memory length: 355359 epsilon: 0.494387200010819 st
 eps: 325 lr: 1.6000000000000003e-05 evaluation reward: 6.36
 episode: 1574 score: 9.0 memory length: 355776 epsilon: 0.49356154001081376
 steps: 417 lr: 1.6000000000000003e-05 evaluation reward: 6.35
 episode: 1575 score: 7.0 memory length: 356200 epsilon: 0.49272202001080845
 steps: 424 lr: 1.6000000000000003e-05 evaluation reward: 6.36
 episode: 1576 score: 8.0 memory length: 356626 epsilon: 0.4918785400108031 s
 teps: 426 lr: 1.6000000000000003e-05 evaluation reward: 6.4
 episode: 1577 score: 3.0 memory length: 356853 epsilon: 0.49142908001080027
 steps: 227 lr: 1.6000000000000003e-05 evaluation reward: 6.38
 episode: 1578 score: 9.0 memory length: 357296 epsilon: 0.4905519400107947 s
 teps: 443 lr: 1.6000000000000003e-05 evaluation reward: 6.42
 episode: 1579 score: 8.0 memory length: 357734 epsilon: 0.48968470001078923
 steps: 438 lr: 1.6000000000000003e-05 evaluation reward: 6.45
 episode: 1580 score: 6.0 memory length: 358041 epsilon: 0.4890768400107854 s
 teps: 307 lr: 1.6000000000000003e-05 evaluation reward: 6.44
 episode: 1581 score: 6.0 memory length: 358376 epsilon: 0.4884135400107812 s
 teps: 335 lr: 1.6000000000000003e-05 evaluation reward: 6.47
 episode: 1582 score: 7.0 memory length: 358784 epsilon: 0.4876057000107761 s
 teps: 408 lr: 1.6000000000000003e-05 evaluation reward: 6.49
 episode: 1583 score: 9.0 memory length: 359288 epsilon: 0.48660778001076976
 steps: 504 lr: 1.6000000000000003e-05 evaluation reward: 6.52
 episode: 1584 score: 3.0 memory length: 359500 epsilon: 0.4861880200107671 s
 teps: 212 lr: 1.6000000000000003e-05 evaluation reward: 6.46
 episode: 1585 score: 8.0 memory length: 359968 epsilon: 0.48526138001076125
 steps: 468 lr: 1.6000000000000003e-05 evaluation reward: 6.49
 episode: 1586 score: 4.0 memory length: 360230 epsilon: 0.48474262001075796
 steps: 262 lr: 1.6000000000000003e-05 evaluation reward: 6.47
 episode: 1587 score: 6.0 memory length: 360574 epsilon: 0.48406150001075365
 steps: 344 lr: 1.6000000000000003e-05 evaluation reward: 6.47
 episode: 1588 score: 9.0 memory length: 361035 epsilon: 0.4831487200107479 s
 teps: 461 lr: 1.6000000000000003e-05 evaluation reward: 6.49
 episode: 1589 score: 5.0 memory length: 361342 epsilon: 0.48254086001074403
 steps: 307 lr: 1.6000000000000003e-05 evaluation reward: 6.38
 episode: 1590 score: 5.0 memory length: 361672 epsilon: 0.4818874600107399 s
 teps: 330 lr: 1.6000000000000003e-05 evaluation reward: 6.37
 episode: 1591 score: 5.0 memory length: 361978 epsilon: 0.48128158001073607
 steps: 306 lr: 1.6000000000000003e-05 evaluation reward: 6.33
 episode: 1592 score: 6.0 memory length: 362343 epsilon: 0.4805588800107315 s
 teps: 365 lr: 1.6000000000000003e-05 evaluation reward: 6.35
 episode: 1593 score: 5.0 memory length: 362670 epsilon: 0.4799114200107274 s
 teps: 327 lr: 1.6000000000000003e-05 evaluation reward: 6.34

episode: 1594 score: 8.0 memory length: 363141 epsilon: 0.4789788400107215 s
 teps: 471 lr: 1.6000000000000003e-05 evaluation reward: 6.37
 episode: 1595 score: 15.0 memory length: 363752 epsilon: 0.47776906001071384
 steps: 611 lr: 1.6000000000000003e-05 evaluation reward: 6.47
 episode: 1596 score: 5.0 memory length: 364058 epsilon: 0.47716318001071 ste
 ps: 306 lr: 1.6000000000000003e-05 evaluation reward: 6.46
 episode: 1597 score: 5.0 memory length: 364367 epsilon: 0.47655136001070614
 steps: 309 lr: 1.6000000000000003e-05 evaluation reward: 6.48
 episode: 1598 score: 9.0 memory length: 364826 epsilon: 0.4756425400107004 s
 teps: 459 lr: 1.6000000000000003e-05 evaluation reward: 6.54
 episode: 1599 score: 10.0 memory length: 365369 epsilon: 0.4745674000106936
 steps: 543 lr: 1.6000000000000003e-05 evaluation reward: 6.59
 episode: 1600 score: 7.0 memory length: 365758 epsilon: 0.4737971800106887 s
 teps: 389 lr: 1.6000000000000003e-05 evaluation reward: 6.54
 episode: 1601 score: 12.0 memory length: 366208 epsilon: 0.4729061800106831
 steps: 450 lr: 1.6000000000000003e-05 evaluation reward: 6.61
 episode: 1602 score: 9.0 memory length: 366616 epsilon: 0.47209834001067796
 steps: 408 lr: 1.6000000000000003e-05 evaluation reward: 6.63
 episode: 1603 score: 2.0 memory length: 366815 epsilon: 0.47170432001067547
 steps: 199 lr: 1.6000000000000003e-05 evaluation reward: 6.62
 episode: 1604 score: 4.0 memory length: 367072 epsilon: 0.47119546001067225
 steps: 257 lr: 1.6000000000000003e-05 evaluation reward: 6.64
 episode: 1605 score: 6.0 memory length: 367376 epsilon: 0.47059354001066844
 steps: 304 lr: 1.6000000000000003e-05 evaluation reward: 6.63
 episode: 1606 score: 3.0 memory length: 367590 epsilon: 0.47016982001066576
 steps: 214 lr: 1.6000000000000003e-05 evaluation reward: 6.65
 episode: 1607 score: 9.0 memory length: 368066 epsilon: 0.4692273400106598 s
 teps: 476 lr: 1.6000000000000003e-05 evaluation reward: 6.69
 episode: 1608 score: 7.0 memory length: 368471 epsilon: 0.4684254400106547 s
 teps: 405 lr: 1.6000000000000003e-05 evaluation reward: 6.71
 episode: 1609 score: 8.0 memory length: 368875 epsilon: 0.46762552001064966
 steps: 404 lr: 1.6000000000000003e-05 evaluation reward: 6.73
 episode: 1610 score: 6.0 memory length: 369251 epsilon: 0.46688104001064495
 steps: 376 lr: 1.6000000000000003e-05 evaluation reward: 6.75
 episode: 1611 score: 7.0 memory length: 369666 epsilon: 0.46605934001063976
 steps: 415 lr: 1.6000000000000003e-05 evaluation reward: 6.81
 episode: 1612 score: 7.0 memory length: 370037 epsilon: 0.4653247600106351 s
 teps: 371 lr: 1.6000000000000003e-05 evaluation reward: 6.82
 episode: 1613 score: 7.0 memory length: 370408 epsilon: 0.46459018001063046
 steps: 371 lr: 1.6000000000000003e-05 evaluation reward: 6.87
 episode: 1614 score: 8.0 memory length: 370850 epsilon: 0.4637150200106249 s
 teps: 442 lr: 1.6000000000000003e-05 evaluation reward: 6.91
 episode: 1615 score: 4.0 memory length: 371091 epsilon: 0.4632378400106219 s
 teps: 241 lr: 1.6000000000000003e-05 evaluation reward: 6.91
 episode: 1616 score: 5.0 memory length: 371398 epsilon: 0.46262998001061806
 steps: 307 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1617 score: 14.0 memory length: 371923 epsilon: 0.4615904800106115
 steps: 525 lr: 1.6000000000000003e-05 evaluation reward: 6.96
 episode: 1618 score: 7.0 memory length: 372334 epsilon: 0.46077670001060633
 steps: 411 lr: 1.6000000000000003e-05 evaluation reward: 6.97
 episode: 1619 score: 11.0 memory length: 372866 epsilon: 0.45972334001059967
 steps: 532 lr: 1.6000000000000003e-05 evaluation reward: 6.98
 episode: 1620 score: 4.0 memory length: 373142 epsilon: 0.4591768600105962 s
 teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 6.96
 episode: 1621 score: 7.0 memory length: 373528 epsilon: 0.4584125800105914 s
 teps: 386 lr: 1.6000000000000003e-05 evaluation reward: 6.96
 episode: 1622 score: 6.0 memory length: 373903 epsilon: 0.4576700800105867 s
 teps: 375 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1623 score: 4.0 memory length: 374146 epsilon: 0.45718894001058363
 steps: 243 lr: 1.6000000000000003e-05 evaluation reward: 6.88

episode: 1624 score: 7.0 memory length: 374552 epsilon: 0.45638506001057855
 steps: 406 lr: 1.6000000000000003e-05 evaluation reward: 6.86
 episode: 1625 score: 11.0 memory length: 375078 epsilon: 0.45534358001057196
 steps: 526 lr: 1.6000000000000003e-05 evaluation reward: 6.84
 episode: 1626 score: 9.0 memory length: 375547 epsilon: 0.4544149600105661 s
 teps: 469 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1627 score: 6.0 memory length: 375905 epsilon: 0.4537061200105616 s
 teps: 358 lr: 1.6000000000000003e-05 evaluation reward: 6.91
 episode: 1628 score: 6.0 memory length: 376238 epsilon: 0.4530467800105574 s
 teps: 333 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1629 score: 4.0 memory length: 376482 epsilon: 0.45256366001055437
 steps: 244 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1630 score: 7.0 memory length: 376831 epsilon: 0.45187264001055 st
 ps: 349 lr: 1.6000000000000003e-05 evaluation reward: 6.9
 episode: 1631 score: 4.0 memory length: 377130 epsilon: 0.45128062001054625
 steps: 299 lr: 1.6000000000000003e-05 evaluation reward: 6.88
 episode: 1632 score: 7.0 memory length: 377568 epsilon: 0.45041338001054076
 steps: 438 lr: 1.6000000000000003e-05 evaluation reward: 6.87
 episode: 1633 score: 7.0 memory length: 377974 epsilon: 0.4496095000105357 s
 teps: 406 lr: 1.6000000000000003e-05 evaluation reward: 6.86
 episode: 1634 score: 4.0 memory length: 378271 epsilon: 0.44902144001053196
 steps: 297 lr: 1.6000000000000003e-05 evaluation reward: 6.85
 episode: 1635 score: 3.0 memory length: 378485 epsilon: 0.4485977200105293 s
 teps: 214 lr: 1.6000000000000003e-05 evaluation reward: 6.79
 episode: 1636 score: 6.0 memory length: 378826 epsilon: 0.447922540010525 st
 eps: 341 lr: 1.6000000000000003e-05 evaluation reward: 6.72
 episode: 1637 score: 10.0 memory length: 379342 epsilon: 0.44690086001051854
 steps: 516 lr: 1.6000000000000003e-05 evaluation reward: 6.73
 episode: 1638 score: 10.0 memory length: 379882 epsilon: 0.4458316600105118
 steps: 540 lr: 1.6000000000000003e-05 evaluation reward: 6.78
 episode: 1639 score: 6.0 memory length: 380243 epsilon: 0.44511688001050725
 steps: 361 lr: 1.6000000000000003e-05 evaluation reward: 6.78
 episode: 1640 score: 9.0 memory length: 380673 epsilon: 0.44426548001050187
 steps: 430 lr: 1.6000000000000003e-05 evaluation reward: 6.83
 episode: 1641 score: 8.0 memory length: 381129 epsilon: 0.44336260001049616
 steps: 456 lr: 1.6000000000000003e-05 evaluation reward: 6.84
 episode: 1642 score: 8.0 memory length: 381530 epsilon: 0.44256862001049113
 steps: 401 lr: 1.6000000000000003e-05 evaluation reward: 6.84
 episode: 1643 score: 7.0 memory length: 381921 epsilon: 0.44179444001048623
 steps: 391 lr: 1.6000000000000003e-05 evaluation reward: 6.8
 episode: 1644 score: 5.0 memory length: 382214 epsilon: 0.44121430001048256
 steps: 293 lr: 1.6000000000000003e-05 evaluation reward: 6.81
 episode: 1645 score: 8.0 memory length: 382491 epsilon: 0.4406658400104791 s
 teps: 277 lr: 1.6000000000000003e-05 evaluation reward: 6.74
 episode: 1646 score: 9.0 memory length: 382998 epsilon: 0.43966198001047274
 steps: 507 lr: 1.6000000000000003e-05 evaluation reward: 6.79
 episode: 1647 score: 8.0 memory length: 383433 epsilon: 0.4388006800104673 s
 teps: 435 lr: 1.6000000000000003e-05 evaluation reward: 6.81
 episode: 1648 score: 9.0 memory length: 383932 epsilon: 0.43781266001046104
 steps: 499 lr: 1.6000000000000003e-05 evaluation reward: 6.84
 episode: 1649 score: 5.0 memory length: 384206 epsilon: 0.4372701400104576 s
 teps: 274 lr: 1.6000000000000003e-05 evaluation reward: 6.83
 episode: 1650 score: 10.0 memory length: 384695 epsilon: 0.4363019200104515
 steps: 489 lr: 1.6000000000000003e-05 evaluation reward: 6.83
 episode: 1651 score: 8.0 memory length: 385132 epsilon: 0.435436660010446 st
 eps: 437 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1652 score: 5.0 memory length: 385443 epsilon: 0.4348208800104421 s
 teps: 311 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1653 score: 10.0 memory length: 385981 epsilon: 0.4337556400104354
 steps: 538 lr: 1.6000000000000003e-05 evaluation reward: 6.91

episode: 1654 score: 9.0 memory length: 386442 epsilon: 0.4328428600104296 s
 teps: 461 lr: 1.6000000000000003e-05 evaluation reward: 6.97
 episode: 1655 score: 4.0 memory length: 386716 epsilon: 0.43230034001042617
 steps: 274 lr: 1.6000000000000003e-05 evaluation reward: 6.98
 episode: 1656 score: 5.0 memory length: 387042 epsilon: 0.4316548600104221 s
 teps: 326 lr: 1.6000000000000003e-05 evaluation reward: 6.92
 episode: 1657 score: 5.0 memory length: 387352 epsilon: 0.4310410600104182 s
 teps: 310 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1658 score: 4.0 memory length: 387593 epsilon: 0.4305638800104152 s
 teps: 241 lr: 1.6000000000000003e-05 evaluation reward: 6.9
 episode: 1659 score: 17.0 memory length: 388245 epsilon: 0.429272920010407 s
 teps: 652 lr: 1.6000000000000003e-05 evaluation reward: 7.03
 episode: 1660 score: 8.0 memory length: 388691 epsilon: 0.4283898400104014 s
 teps: 446 lr: 1.6000000000000003e-05 evaluation reward: 7.02
 episode: 1661 score: 5.0 memory length: 388998 epsilon: 0.4277819800103976 s
 teps: 307 lr: 1.6000000000000003e-05 evaluation reward: 7.01
 episode: 1662 score: 8.0 memory length: 389453 epsilon: 0.4268810800103919 s
 teps: 455 lr: 1.6000000000000003e-05 evaluation reward: 7.01
 episode: 1663 score: 6.0 memory length: 389789 epsilon: 0.42621580001038767
 steps: 336 lr: 1.6000000000000003e-05 evaluation reward: 6.99
 episode: 1664 score: 4.0 memory length: 390048 epsilon: 0.4257029800103844 s
 teps: 259 lr: 1.6000000000000003e-05 evaluation reward: 6.98
 episode: 1665 score: 7.0 memory length: 390436 epsilon: 0.42493474001037956
 steps: 388 lr: 1.6000000000000003e-05 evaluation reward: 6.99
 episode: 1666 score: 10.0 memory length: 390789 epsilon: 0.42423580001037514
 steps: 353 lr: 1.6000000000000003e-05 evaluation reward: 6.99
 episode: 1667 score: 5.0 memory length: 391118 epsilon: 0.423584380010371 st
 eps: 329 lr: 1.6000000000000003e-05 evaluation reward: 7.0
 episode: 1668 score: 8.0 memory length: 391548 epsilon: 0.42273298001036563
 steps: 430 lr: 1.6000000000000003e-05 evaluation reward: 7.0
 episode: 1669 score: 9.0 memory length: 392056 epsilon: 0.42172714001035927
 steps: 508 lr: 1.6000000000000003e-05 evaluation reward: 7.03
 episode: 1670 score: 12.0 memory length: 392511 epsilon: 0.42082624001035357
 steps: 455 lr: 1.6000000000000003e-05 evaluation reward: 7.05
 episode: 1671 score: 4.0 memory length: 392774 epsilon: 0.4203055000103503 s
 teps: 263 lr: 1.6000000000000003e-05 evaluation reward: 7.03
 episode: 1672 score: 9.0 memory length: 393220 epsilon: 0.4194224200103447 s
 teps: 446 lr: 1.6000000000000003e-05 evaluation reward: 7.04
 episode: 1673 score: 10.0 memory length: 393754 epsilon: 0.418365100010338 s
 teps: 534 lr: 1.6000000000000003e-05 evaluation reward: 7.09
 episode: 1674 score: 3.0 memory length: 393981 epsilon: 0.41791564001033515
 steps: 227 lr: 1.6000000000000003e-05 evaluation reward: 7.03
 episode: 1675 score: 6.0 memory length: 394340 epsilon: 0.41720482001033066
 steps: 359 lr: 1.6000000000000003e-05 evaluation reward: 7.02
 episode: 1676 score: 9.0 memory length: 394806 epsilon: 0.4162821400103248 s
 teps: 466 lr: 1.6000000000000003e-05 evaluation reward: 7.03
 episode: 1677 score: 7.0 memory length: 395208 epsilon: 0.4154861800103198 s
 teps: 402 lr: 1.6000000000000003e-05 evaluation reward: 7.07
 episode: 1678 score: 3.0 memory length: 395420 epsilon: 0.4150664200103171 s
 teps: 212 lr: 1.6000000000000003e-05 evaluation reward: 7.01
 episode: 1679 score: 19.0 memory length: 396037 epsilon: 0.4138447600103094
 steps: 617 lr: 1.6000000000000003e-05 evaluation reward: 7.12
 episode: 1680 score: 7.0 memory length: 396457 epsilon: 0.41301316001030414
 steps: 420 lr: 1.6000000000000003e-05 evaluation reward: 7.13
 episode: 1681 score: 8.0 memory length: 396885 epsilon: 0.4121657200102988 s
 teps: 428 lr: 1.6000000000000003e-05 evaluation reward: 7.15
 episode: 1682 score: 5.0 memory length: 397160 epsilon: 0.41162122001029533
 steps: 275 lr: 1.6000000000000003e-05 evaluation reward: 7.13
 episode: 1683 score: 11.0 memory length: 397537 epsilon: 0.4108747600102906
 steps: 377 lr: 1.6000000000000003e-05 evaluation reward: 7.15

```
episode: 1684 score: 6.0 memory length: 397933 epsilon: 0.41009068001028565
  steps: 396 lr: 1.6000000000000003e-05 evaluation reward: 7.18
episode: 1685 score: 8.0 memory length: 398367 epsilon: 0.4092313600102802 s
  teps: 434 lr: 1.6000000000000003e-05 evaluation reward: 7.18
episode: 1686 score: 11.0 memory length: 398970 epsilon: 0.40803742001027266
  steps: 603 lr: 1.6000000000000003e-05 evaluation reward: 7.25
episode: 1687 score: 7.0 memory length: 399339 epsilon: 0.40730680001026803
  steps: 369 lr: 1.6000000000000003e-05 evaluation reward: 7.26
episode: 1688 score: 4.0 memory length: 399584 epsilon: 0.40682170001026496
  steps: 245 lr: 1.6000000000000003e-05 evaluation reward: 7.21
episode: 1689 score: 5.0 memory length: 399892 epsilon: 0.4062118600102611 s
  teps: 308 lr: 1.6000000000000003e-05 evaluation reward: 7.21
episode: 1690 score: 16.0 memory length: 400363 epsilon: 0.4052792800102552
  steps: 471 lr: 6.400000000000001e-06 evaluation reward: 7.32
episode: 1691 score: 7.0 memory length: 400763 epsilon: 0.4044872800102502 s
  teps: 400 lr: 6.400000000000001e-06 evaluation reward: 7.34
episode: 1692 score: 8.0 memory length: 401180 epsilon: 0.40366162001024497
  steps: 417 lr: 6.400000000000001e-06 evaluation reward: 7.36
episode: 1693 score: 9.0 memory length: 401646 epsilon: 0.40273894001023913
  steps: 466 lr: 6.400000000000001e-06 evaluation reward: 7.4
episode: 1694 score: 10.0 memory length: 402111 epsilon: 0.4018182400102333
  steps: 465 lr: 6.400000000000001e-06 evaluation reward: 7.42
episode: 1695 score: 3.0 memory length: 402325 epsilon: 0.4013945200102306 s
  teps: 214 lr: 6.400000000000001e-06 evaluation reward: 7.3
episode: 1696 score: 3.0 memory length: 402539 epsilon: 0.40097080001022795
  steps: 214 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1697 score: 5.0 memory length: 402862 epsilon: 0.4003312600102239 s
  teps: 323 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1698 score: 8.0 memory length: 403297 epsilon: 0.39946996001021845
  steps: 435 lr: 6.400000000000001e-06 evaluation reward: 7.27
episode: 1699 score: 9.0 memory length: 403750 epsilon: 0.3985730200102128 s
  teps: 453 lr: 6.400000000000001e-06 evaluation reward: 7.26
episode: 1700 score: 4.0 memory length: 404049 epsilon: 0.39798100001020903
  steps: 299 lr: 6.400000000000001e-06 evaluation reward: 7.23
episode: 1701 score: 7.0 memory length: 404457 epsilon: 0.3971731600102039 s
  teps: 408 lr: 6.400000000000001e-06 evaluation reward: 7.18
episode: 1702 score: 8.0 memory length: 404912 epsilon: 0.3962722600101982 s
  teps: 455 lr: 6.400000000000001e-06 evaluation reward: 7.17
episode: 1703 score: 5.0 memory length: 405223 epsilon: 0.3956564800101943 s
  teps: 311 lr: 6.400000000000001e-06 evaluation reward: 7.2
episode: 1704 score: 9.0 memory length: 405670 epsilon: 0.3947714200101887 s
  teps: 447 lr: 6.400000000000001e-06 evaluation reward: 7.25
episode: 1705 score: 8.0 memory length: 406074 epsilon: 0.39397150001018366
  steps: 404 lr: 6.400000000000001e-06 evaluation reward: 7.27
episode: 1706 score: 6.0 memory length: 406427 epsilon: 0.39327256001017924
  steps: 353 lr: 6.400000000000001e-06 evaluation reward: 7.3
episode: 1707 score: 11.0 memory length: 406971 epsilon: 0.3921954400101724
  steps: 544 lr: 6.400000000000001e-06 evaluation reward: 7.32
episode: 1708 score: 6.0 memory length: 407341 epsilon: 0.3914628400101678 s
  teps: 370 lr: 6.400000000000001e-06 evaluation reward: 7.31
episode: 1709 score: 7.0 memory length: 407771 epsilon: 0.3906114400101624 s
  teps: 430 lr: 6.400000000000001e-06 evaluation reward: 7.3
episode: 1710 score: 9.0 memory length: 408228 epsilon: 0.3897065800101567 s
  teps: 457 lr: 6.400000000000001e-06 evaluation reward: 7.33
episode: 1711 score: 9.0 memory length: 408687 epsilon: 0.38879776001015093
  steps: 459 lr: 6.400000000000001e-06 evaluation reward: 7.35
episode: 1712 score: 12.0 memory length: 409257 epsilon: 0.3876691600101438
  steps: 570 lr: 6.400000000000001e-06 evaluation reward: 7.4
episode: 1713 score: 14.0 memory length: 409765 epsilon: 0.3866633200101374
  steps: 508 lr: 6.400000000000001e-06 evaluation reward: 7.47
```

episode: 1714 score: 7.0 memory length: 410152 epsilon: 0.3858970600101326 s
 teps: 387 lr: 6.400000000000000e-06 evaluation reward: 7.46
 episode: 1715 score: 11.0 memory length: 410762 epsilon: 0.38468926001012493
 steps: 610 lr: 6.400000000000000e-06 evaluation reward: 7.53
 episode: 1716 score: 6.0 memory length: 411108 epsilon: 0.3840041800101206 s
 teps: 346 lr: 6.400000000000000e-06 evaluation reward: 7.54
 episode: 1717 score: 9.0 memory length: 411598 epsilon: 0.38303398001011446
 steps: 490 lr: 6.400000000000000e-06 evaluation reward: 7.49
 episode: 1718 score: 6.0 memory length: 411945 epsilon: 0.3823469200101101 s
 teps: 347 lr: 6.400000000000000e-06 evaluation reward: 7.48
 episode: 1719 score: 7.0 memory length: 412336 epsilon: 0.3815727400101052 s
 teps: 391 lr: 6.400000000000000e-06 evaluation reward: 7.44
 episode: 1720 score: 5.0 memory length: 412663 epsilon: 0.3809252800101011 s
 teps: 327 lr: 6.400000000000000e-06 evaluation reward: 7.45
 episode: 1721 score: 4.0 memory length: 412942 epsilon: 0.3803728600100976 s
 teps: 279 lr: 6.400000000000000e-06 evaluation reward: 7.42
 episode: 1722 score: 7.0 memory length: 413352 epsilon: 0.3795610600100925 s
 teps: 410 lr: 6.400000000000000e-06 evaluation reward: 7.43
 episode: 1723 score: 10.0 memory length: 413744 epsilon: 0.3787849000100876
 steps: 392 lr: 6.400000000000000e-06 evaluation reward: 7.49
 episode: 1724 score: 10.0 memory length: 414102 epsilon: 0.3780760600100831
 steps: 358 lr: 6.400000000000000e-06 evaluation reward: 7.52
 episode: 1725 score: 9.0 memory length: 414532 epsilon: 0.3772246600100777 s
 teps: 430 lr: 6.400000000000000e-06 evaluation reward: 7.5
 episode: 1726 score: 15.0 memory length: 415253 epsilon: 0.3757970800100687
 steps: 721 lr: 6.400000000000000e-06 evaluation reward: 7.56
 episode: 1727 score: 7.0 memory length: 415643 epsilon: 0.3750248800100638 s
 teps: 390 lr: 6.400000000000000e-06 evaluation reward: 7.57
 episode: 1728 score: 7.0 memory length: 416057 epsilon: 0.3742051600100586 s
 teps: 414 lr: 6.400000000000000e-06 evaluation reward: 7.58
 episode: 1729 score: 9.0 memory length: 416507 epsilon: 0.37331416001005296
 steps: 450 lr: 6.400000000000000e-06 evaluation reward: 7.63
 episode: 1730 score: 9.0 memory length: 417011 epsilon: 0.37231624001004665
 steps: 504 lr: 6.400000000000000e-06 evaluation reward: 7.65
 episode: 1731 score: 7.0 memory length: 417418 epsilon: 0.37151038001004155
 steps: 407 lr: 6.400000000000000e-06 evaluation reward: 7.68
 episode: 1732 score: 9.0 memory length: 417874 epsilon: 0.37060750001003584
 steps: 456 lr: 6.400000000000000e-06 evaluation reward: 7.7
 episode: 1733 score: 10.0 memory length: 418429 epsilon: 0.3695086000100289
 steps: 555 lr: 6.400000000000000e-06 evaluation reward: 7.73
 episode: 1734 score: 7.0 memory length: 418816 epsilon: 0.36874234001002404
 steps: 387 lr: 6.400000000000000e-06 evaluation reward: 7.76
 episode: 1735 score: 7.0 memory length: 419187 epsilon: 0.3680077600100194 s
 teps: 371 lr: 6.400000000000000e-06 evaluation reward: 7.8
 episode: 1736 score: 4.0 memory length: 419443 epsilon: 0.3675008800100162 s
 teps: 256 lr: 6.400000000000000e-06 evaluation reward: 7.78
 episode: 1737 score: 7.0 memory length: 419815 epsilon: 0.3667643200100115 s
 teps: 372 lr: 6.400000000000000e-06 evaluation reward: 7.75
 episode: 1738 score: 9.0 memory length: 420275 epsilon: 0.36585352001000576
 steps: 460 lr: 6.400000000000000e-06 evaluation reward: 7.74
 episode: 1739 score: 9.0 memory length: 420709 epsilon: 0.3649942000100003 s
 teps: 434 lr: 6.400000000000000e-06 evaluation reward: 7.77
 episode: 1740 score: 18.0 memory length: 421333 epsilon: 0.3637586800099925
 steps: 624 lr: 6.400000000000000e-06 evaluation reward: 7.86
 episode: 1741 score: 6.0 memory length: 421753 epsilon: 0.36292708000998725
 steps: 420 lr: 6.400000000000000e-06 evaluation reward: 7.84
 episode: 1742 score: 8.0 memory length: 422175 epsilon: 0.36209152000998196
 steps: 422 lr: 6.400000000000000e-06 evaluation reward: 7.84
 episode: 1743 score: 6.0 memory length: 422514 epsilon: 0.3614203000099777 s
 teps: 339 lr: 6.400000000000000e-06 evaluation reward: 7.83

episode: 1744 score: 8.0 memory length: 422940 epsilon: 0.3605768200099724 s
 teps: 426 lr: 6.400000000000001e-06 evaluation reward: 7.86
 episode: 1745 score: 5.0 memory length: 423269 epsilon: 0.35992540000996825
 steps: 329 lr: 6.400000000000001e-06 evaluation reward: 7.83
 episode: 1746 score: 13.0 memory length: 423776 epsilon: 0.3589215400099619
 steps: 507 lr: 6.400000000000001e-06 evaluation reward: 7.87
 episode: 1747 score: 8.0 memory length: 424253 epsilon: 0.35797708000995593
 steps: 477 lr: 6.400000000000001e-06 evaluation reward: 7.87
 episode: 1748 score: 6.0 memory length: 424557 epsilon: 0.3573751600099521 s
 teps: 304 lr: 6.400000000000001e-06 evaluation reward: 7.84
 episode: 1749 score: 10.0 memory length: 425083 epsilon: 0.35633368000994553
 steps: 526 lr: 6.400000000000001e-06 evaluation reward: 7.89
 episode: 1750 score: 6.0 memory length: 425442 epsilon: 0.35562286000994103
 steps: 359 lr: 6.400000000000001e-06 evaluation reward: 7.85
 episode: 1751 score: 10.0 memory length: 425925 epsilon: 0.354666520009935 s
 teps: 483 lr: 6.400000000000001e-06 evaluation reward: 7.87
 episode: 1752 score: 10.0 memory length: 426457 epsilon: 0.3536131600099283
 steps: 532 lr: 6.400000000000001e-06 evaluation reward: 7.92
 episode: 1753 score: 11.0 memory length: 426866 epsilon: 0.3528033400099232
 steps: 409 lr: 6.400000000000001e-06 evaluation reward: 7.93
 episode: 1754 score: 9.0 memory length: 427177 epsilon: 0.3521875600099193 s
 teps: 311 lr: 6.400000000000001e-06 evaluation reward: 7.93
 episode: 1755 score: 12.0 memory length: 427640 epsilon: 0.3512708200099135
 steps: 463 lr: 6.400000000000001e-06 evaluation reward: 8.01
 episode: 1756 score: 8.0 memory length: 428093 epsilon: 0.3503738800099078 s
 teps: 453 lr: 6.400000000000001e-06 evaluation reward: 8.04
 episode: 1757 score: 11.0 memory length: 428603 epsilon: 0.34936408000990143
 steps: 510 lr: 6.400000000000001e-06 evaluation reward: 8.1
 episode: 1758 score: 7.0 memory length: 428988 epsilon: 0.3486017800098966 s
 teps: 385 lr: 6.400000000000001e-06 evaluation reward: 8.13
 episode: 1759 score: 8.0 memory length: 429390 epsilon: 0.3478058200098916 s
 teps: 402 lr: 6.400000000000001e-06 evaluation reward: 8.04
 episode: 1760 score: 4.0 memory length: 429650 epsilon: 0.3472910200098883 s
 teps: 260 lr: 6.400000000000001e-06 evaluation reward: 8.0
 episode: 1761 score: 8.0 memory length: 430019 epsilon: 0.3465604000098837 s
 teps: 369 lr: 6.400000000000001e-06 evaluation reward: 8.03
 episode: 1762 score: 7.0 memory length: 430426 epsilon: 0.3457545400098786 s
 teps: 407 lr: 6.400000000000001e-06 evaluation reward: 8.02
 episode: 1763 score: 9.0 memory length: 430910 epsilon: 0.34479622000987253
 steps: 484 lr: 6.400000000000001e-06 evaluation reward: 8.05
 episode: 1764 score: 14.0 memory length: 431488 epsilon: 0.3436517800098653
 steps: 578 lr: 6.400000000000001e-06 evaluation reward: 8.15
 episode: 1765 score: 7.0 memory length: 431895 epsilon: 0.3428459200098602 s
 teps: 407 lr: 6.400000000000001e-06 evaluation reward: 8.15
 episode: 1766 score: 15.0 memory length: 432479 epsilon: 0.3416896000098529
 steps: 584 lr: 6.400000000000001e-06 evaluation reward: 8.2
 episode: 1767 score: 10.0 memory length: 433001 epsilon: 0.34065604000984634
 steps: 522 lr: 6.400000000000001e-06 evaluation reward: 8.25
 episode: 1768 score: 6.0 memory length: 433345 epsilon: 0.33997492000984203
 steps: 344 lr: 6.400000000000001e-06 evaluation reward: 8.23
 episode: 1769 score: 4.0 memory length: 433604 epsilon: 0.3394621000098388 s
 teps: 259 lr: 6.400000000000001e-06 evaluation reward: 8.18
 episode: 1770 score: 6.0 memory length: 433945 epsilon: 0.3387869200098345 s
 teps: 341 lr: 6.400000000000001e-06 evaluation reward: 8.12
 episode: 1771 score: 11.0 memory length: 434517 epsilon: 0.33765436000982735
 steps: 572 lr: 6.400000000000001e-06 evaluation reward: 8.19
 episode: 1772 score: 7.0 memory length: 434766 epsilon: 0.33716134000982423
 steps: 249 lr: 6.400000000000001e-06 evaluation reward: 8.17
 episode: 1773 score: 12.0 memory length: 435293 epsilon: 0.3361178800098176
 steps: 527 lr: 6.400000000000001e-06 evaluation reward: 8.19

episode: 1774 score: 5.0 memory length: 435588 epsilon: 0.33553378000981393
 steps: 295 lr: 6.400000000000001e-06 evaluation reward: 8.21
 episode: 1775 score: 9.0 memory length: 436097 epsilon: 0.33452596000980755
 steps: 509 lr: 6.400000000000001e-06 evaluation reward: 8.24
 episode: 1776 score: 8.0 memory length: 436518 epsilon: 0.3336923800098023 s
 teps: 421 lr: 6.400000000000001e-06 evaluation reward: 8.23
 episode: 1777 score: 8.0 memory length: 436933 epsilon: 0.3328706800097971 s
 teps: 415 lr: 6.400000000000001e-06 evaluation reward: 8.24
 episode: 1778 score: 11.0 memory length: 437447 epsilon: 0.33185296000979064
 steps: 514 lr: 6.400000000000001e-06 evaluation reward: 8.32
 episode: 1779 score: 11.0 memory length: 437857 epsilon: 0.3310411600097855
 steps: 410 lr: 6.400000000000001e-06 evaluation reward: 8.24
 episode: 1780 score: 13.0 memory length: 438446 epsilon: 0.32987494000977813
 steps: 589 lr: 6.400000000000001e-06 evaluation reward: 8.3
 episode: 1781 score: 7.0 memory length: 438853 epsilon: 0.32906908000977303
 steps: 407 lr: 6.400000000000001e-06 evaluation reward: 8.29
 episode: 1782 score: 10.0 memory length: 439372 epsilon: 0.32804146000976653
 steps: 519 lr: 6.400000000000001e-06 evaluation reward: 8.34
 episode: 1783 score: 8.0 memory length: 439841 epsilon: 0.32711284000976065
 steps: 469 lr: 6.400000000000001e-06 evaluation reward: 8.31
 episode: 1784 score: 10.0 memory length: 440394 epsilon: 0.3260179000097537
 steps: 553 lr: 6.400000000000001e-06 evaluation reward: 8.35
 episode: 1785 score: 6.0 memory length: 440733 epsilon: 0.3253466800097495 s
 teps: 339 lr: 6.400000000000001e-06 evaluation reward: 8.33
 episode: 1786 score: 12.0 memory length: 441379 epsilon: 0.3240676000097414
 steps: 646 lr: 6.400000000000001e-06 evaluation reward: 8.34
 episode: 1787 score: 9.0 memory length: 441686 epsilon: 0.32345974000973754
 steps: 307 lr: 6.400000000000001e-06 evaluation reward: 8.36
 episode: 1788 score: 10.0 memory length: 442194 epsilon: 0.3224539000097312
 steps: 508 lr: 6.400000000000001e-06 evaluation reward: 8.42
 episode: 1789 score: 9.0 memory length: 442689 epsilon: 0.321473800009725 st
 eps: 495 lr: 6.400000000000001e-06 evaluation reward: 8.46
 episode: 1790 score: 5.0 memory length: 442999 epsilon: 0.3208600000097211 s
 teps: 310 lr: 6.400000000000001e-06 evaluation reward: 8.35
 episode: 1791 score: 17.0 memory length: 443547 epsilon: 0.3197749600097142
 steps: 548 lr: 6.400000000000001e-06 evaluation reward: 8.45
 episode: 1792 score: 7.0 memory length: 443913 epsilon: 0.31905028000970964
 steps: 366 lr: 6.400000000000001e-06 evaluation reward: 8.44
 episode: 1793 score: 8.0 memory length: 444351 epsilon: 0.31818304000970415
 steps: 438 lr: 6.400000000000001e-06 evaluation reward: 8.43
 episode: 1794 score: 7.0 memory length: 444746 epsilon: 0.3174009400096992 s
 teps: 395 lr: 6.400000000000001e-06 evaluation reward: 8.4
 episode: 1795 score: 9.0 memory length: 445221 epsilon: 0.31646044000969326
 steps: 475 lr: 6.400000000000001e-06 evaluation reward: 8.46
 episode: 1796 score: 5.0 memory length: 445533 epsilon: 0.31584268000968935
 steps: 312 lr: 6.400000000000001e-06 evaluation reward: 8.48
 episode: 1797 score: 10.0 memory length: 446033 epsilon: 0.3148526800096831
 steps: 500 lr: 6.400000000000001e-06 evaluation reward: 8.53
 episode: 1798 score: 8.0 memory length: 446450 epsilon: 0.31402702000967786
 steps: 417 lr: 6.400000000000001e-06 evaluation reward: 8.53
 episode: 1799 score: 8.0 memory length: 446877 epsilon: 0.3131815600096725 s
 teps: 427 lr: 6.400000000000001e-06 evaluation reward: 8.52
 episode: 1800 score: 7.0 memory length: 447281 epsilon: 0.31238164000966745
 steps: 404 lr: 6.400000000000001e-06 evaluation reward: 8.55
 episode: 1801 score: 8.0 memory length: 447698 epsilon: 0.3115559800096622 s
 teps: 417 lr: 6.400000000000001e-06 evaluation reward: 8.56
 episode: 1802 score: 12.0 memory length: 448134 epsilon: 0.31069270000965676
 steps: 436 lr: 6.400000000000001e-06 evaluation reward: 8.6
 episode: 1803 score: 12.0 memory length: 448672 epsilon: 0.30962746000965 st
 eps: 538 lr: 6.400000000000001e-06 evaluation reward: 8.67

episode: 1804 score: 11.0 memory length: 449121 epsilon: 0.3087384400096444
 steps: 449 lr: 6.400000000000000e-06 evaluation reward: 8.69
 episode: 1805 score: 7.0 memory length: 449370 epsilon: 0.3082454200096413 s
 teps: 249 lr: 6.400000000000000e-06 evaluation reward: 8.68
 episode: 1806 score: 5.0 memory length: 449662 epsilon: 0.3076672600096376 s
 teps: 292 lr: 6.400000000000000e-06 evaluation reward: 8.67
 episode: 1807 score: 17.0 memory length: 450310 epsilon: 0.3063842200096295
 steps: 648 lr: 6.400000000000000e-06 evaluation reward: 8.73
 episode: 1808 score: 8.0 memory length: 450751 epsilon: 0.305511040009624 st
 eps: 441 lr: 6.400000000000000e-06 evaluation reward: 8.75
 episode: 1809 score: 5.0 memory length: 451040 epsilon: 0.30493882000962036
 steps: 289 lr: 6.400000000000000e-06 evaluation reward: 8.73
 episode: 1810 score: 5.0 memory length: 451348 epsilon: 0.3043289800096165 s
 teps: 308 lr: 6.400000000000000e-06 evaluation reward: 8.69
 episode: 1811 score: 11.0 memory length: 451899 epsilon: 0.3032380000096096
 steps: 551 lr: 6.400000000000000e-06 evaluation reward: 8.71
 episode: 1812 score: 9.0 memory length: 452364 epsilon: 0.3023173000096038 s
 teps: 465 lr: 6.400000000000000e-06 evaluation reward: 8.68
 episode: 1813 score: 7.0 memory length: 452773 epsilon: 0.30150748000959865
 steps: 409 lr: 6.400000000000000e-06 evaluation reward: 8.61
 episode: 1814 score: 12.0 memory length: 453320 epsilon: 0.3004244200095918
 steps: 547 lr: 6.400000000000000e-06 evaluation reward: 8.66
 episode: 1815 score: 7.0 memory length: 453712 epsilon: 0.2996482600095869 s
 teps: 392 lr: 6.400000000000000e-06 evaluation reward: 8.62
 episode: 1816 score: 10.0 memory length: 454211 epsilon: 0.29866024000958064
 steps: 499 lr: 6.400000000000000e-06 evaluation reward: 8.66
 episode: 1817 score: 7.0 memory length: 454638 epsilon: 0.2978147800095753 s
 teps: 427 lr: 6.400000000000000e-06 evaluation reward: 8.64
 episode: 1818 score: 8.0 memory length: 455061 epsilon: 0.29697724000957 ste
 ps: 423 lr: 6.400000000000000e-06 evaluation reward: 8.66
 episode: 1819 score: 6.0 memory length: 455395 epsilon: 0.2963159200095658 s
 teps: 334 lr: 6.400000000000000e-06 evaluation reward: 8.65
 episode: 1820 score: 11.0 memory length: 455970 epsilon: 0.2951774200095586
 steps: 575 lr: 6.400000000000000e-06 evaluation reward: 8.71
 episode: 1821 score: 9.0 memory length: 456427 epsilon: 0.2942725600095529 s
 teps: 457 lr: 6.400000000000000e-06 evaluation reward: 8.76
 episode: 1822 score: 12.0 memory length: 457038 epsilon: 0.2930627800095452
 steps: 611 lr: 6.400000000000000e-06 evaluation reward: 8.81
 episode: 1823 score: 14.0 memory length: 457645 epsilon: 0.2918609200095376
 steps: 607 lr: 6.400000000000000e-06 evaluation reward: 8.85
 episode: 1824 score: 10.0 memory length: 458132 epsilon: 0.2908966600095315
 steps: 487 lr: 6.400000000000000e-06 evaluation reward: 8.85
 episode: 1825 score: 7.0 memory length: 458510 epsilon: 0.2901482200095268 s
 teps: 378 lr: 6.400000000000000e-06 evaluation reward: 8.83
 episode: 1826 score: 7.0 memory length: 458886 epsilon: 0.28940374000952207
 steps: 376 lr: 6.400000000000000e-06 evaluation reward: 8.75
 episode: 1827 score: 9.0 memory length: 459374 epsilon: 0.28843750000951596
 steps: 488 lr: 6.400000000000000e-06 evaluation reward: 8.77
 episode: 1828 score: 6.0 memory length: 459738 epsilon: 0.2877167800095114 s
 teps: 364 lr: 6.400000000000000e-06 evaluation reward: 8.76
 episode: 1829 score: 4.0 memory length: 460001 epsilon: 0.2871960400095081 s
 teps: 263 lr: 6.400000000000000e-06 evaluation reward: 8.71
 episode: 1830 score: 12.0 memory length: 460604 epsilon: 0.28600210000950055
 steps: 603 lr: 6.400000000000000e-06 evaluation reward: 8.74
 episode: 1831 score: 7.0 memory length: 460992 epsilon: 0.2852338600094957 s
 teps: 388 lr: 6.400000000000000e-06 evaluation reward: 8.74
 episode: 1832 score: 8.0 memory length: 461444 epsilon: 0.28433890000949 st
 ps: 452 lr: 6.400000000000000e-06 evaluation reward: 8.73
 episode: 1833 score: 5.0 memory length: 461769 epsilon: 0.28369540000948595
 steps: 325 lr: 6.400000000000000e-06 evaluation reward: 8.68

```

episode: 1834 score: 6.0 memory length: 462107 epsilon: 0.2830261600094817 s
teps: 338 lr: 6.400000000000001e-06 evaluation reward: 8.67
episode: 1835 score: 11.0 memory length: 462665 epsilon: 0.28192132000947473
steps: 558 lr: 6.400000000000001e-06 evaluation reward: 8.71
episode: 1836 score: 15.0 memory length: 463212 epsilon: 0.2808382600094679
steps: 547 lr: 6.400000000000001e-06 evaluation reward: 8.82
episode: 1837 score: 10.0 memory length: 463740 epsilon: 0.27979282000946126
steps: 528 lr: 6.400000000000001e-06 evaluation reward: 8.85
episode: 1838 score: 7.0 memory length: 464164 epsilon: 0.27895330000945595
steps: 424 lr: 6.400000000000001e-06 evaluation reward: 8.83
episode: 1839 score: 9.0 memory length: 464669 epsilon: 0.2779534000094496 s
teps: 505 lr: 6.400000000000001e-06 evaluation reward: 8.83
episode: 1840 score: 2.0 memory length: 464852 epsilon: 0.27759106000944733
steps: 183 lr: 6.400000000000001e-06 evaluation reward: 8.67
episode: 1841 score: 12.0 memory length: 465471 epsilon: 0.2763654400094396
steps: 619 lr: 6.400000000000001e-06 evaluation reward: 8.73
episode: 1842 score: 11.0 memory length: 466037 epsilon: 0.2752447600094325
steps: 566 lr: 6.400000000000001e-06 evaluation reward: 8.76
episode: 1843 score: 4.0 memory length: 466282 epsilon: 0.2747596600094294 s
teps: 245 lr: 6.400000000000001e-06 evaluation reward: 8.74
episode: 1844 score: 6.0 memory length: 466622 epsilon: 0.27408646000942516
steps: 340 lr: 6.400000000000001e-06 evaluation reward: 8.72
episode: 1845 score: 3.0 memory length: 466836 epsilon: 0.2736627400094225 s
teps: 214 lr: 6.400000000000001e-06 evaluation reward: 8.7
episode: 1846 score: 12.0 memory length: 467457 epsilon: 0.2724331600094147
steps: 621 lr: 6.400000000000001e-06 evaluation reward: 8.69
episode: 1847 score: 7.0 memory length: 467826 epsilon: 0.2717025400094101 s
teps: 369 lr: 6.400000000000001e-06 evaluation reward: 8.68
episode: 1848 score: 4.0 memory length: 468089 epsilon: 0.2711818000094068 s
teps: 263 lr: 6.400000000000001e-06 evaluation reward: 8.66
episode: 1849 score: 9.0 memory length: 468545 epsilon: 0.27027892000940107
steps: 456 lr: 6.400000000000001e-06 evaluation reward: 8.65
episode: 1850 score: 9.0 memory length: 469023 epsilon: 0.2693324800093951 s
teps: 478 lr: 6.400000000000001e-06 evaluation reward: 8.68
episode: 1851 score: 12.0 memory length: 469608 epsilon: 0.26817418000938775
steps: 585 lr: 6.400000000000001e-06 evaluation reward: 8.7
episode: 1852 score: 8.0 memory length: 470027 epsilon: 0.2673445600093825 s
teps: 419 lr: 6.400000000000001e-06 evaluation reward: 8.68
episode: 1853 score: 10.0 memory length: 470520 epsilon: 0.2663684200093763
steps: 493 lr: 6.400000000000001e-06 evaluation reward: 8.67
episode: 1854 score: 9.0 memory length: 470971 epsilon: 0.2654754400093707 s
teps: 451 lr: 6.400000000000001e-06 evaluation reward: 8.67
episode: 1855 score: 16.0 memory length: 471618 epsilon: 0.2641943800093626
steps: 647 lr: 6.400000000000001e-06 evaluation reward: 8.71
episode: 1856 score: 7.0 memory length: 471989 epsilon: 0.2634598000093579 s
teps: 371 lr: 6.400000000000001e-06 evaluation reward: 8.7
episode: 1857 score: 6.0 memory length: 472332 epsilon: 0.26278066000935363
steps: 343 lr: 6.400000000000001e-06 evaluation reward: 8.65
episode: 1858 score: 15.0 memory length: 473015 epsilon: 0.26142832000934507
steps: 683 lr: 6.400000000000001e-06 evaluation reward: 8.73
episode: 1859 score: 8.0 memory length: 473484 epsilon: 0.2604997000093392 s
teps: 469 lr: 6.400000000000001e-06 evaluation reward: 8.73
episode: 1860 score: 8.0 memory length: 473874 epsilon: 0.2597275000093343 s
teps: 390 lr: 6.400000000000001e-06 evaluation reward: 8.77
episode: 1861 score: 7.0 memory length: 474282 epsilon: 0.2589196600093292 s
teps: 408 lr: 6.400000000000001e-06 evaluation reward: 8.76
episode: 1862 score: 7.0 memory length: 474693 epsilon: 0.25810588000932405
steps: 411 lr: 6.400000000000001e-06 evaluation reward: 8.76
episode: 1863 score: 11.0 memory length: 475193 epsilon: 0.2571158800093178
steps: 500 lr: 6.400000000000001e-06 evaluation reward: 8.78

```

episode: 1864 score: 8.0 memory length: 475647 epsilon: 0.2562169600093121 s
 teps: 454 lr: 6.400000000000001e-06 evaluation reward: 8.72
 episode: 1865 score: 7.0 memory length: 476017 epsilon: 0.25548436000930747
 steps: 370 lr: 6.400000000000001e-06 evaluation reward: 8.72
 episode: 1866 score: 9.0 memory length: 476469 epsilon: 0.2545894000093018 s
 teps: 452 lr: 6.400000000000001e-06 evaluation reward: 8.66
 episode: 1867 score: 9.0 memory length: 476935 epsilon: 0.25366672000929597
 steps: 466 lr: 6.400000000000001e-06 evaluation reward: 8.65
 episode: 1868 score: 9.0 memory length: 477386 epsilon: 0.2527737400092903 s
 teps: 451 lr: 6.400000000000001e-06 evaluation reward: 8.68
 episode: 1869 score: 7.0 memory length: 477809 epsilon: 0.251936200009285 st
 eps: 423 lr: 6.400000000000001e-06 evaluation reward: 8.71
 episode: 1870 score: 10.0 memory length: 478206 epsilon: 0.25115014000928004
 steps: 397 lr: 6.400000000000001e-06 evaluation reward: 8.75
 episode: 1871 score: 17.0 memory length: 478862 epsilon: 0.24985126000927182
 steps: 656 lr: 6.400000000000001e-06 evaluation reward: 8.81
 episode: 1872 score: 9.0 memory length: 479329 epsilon: 0.24892660000926597
 steps: 467 lr: 6.400000000000001e-06 evaluation reward: 8.83
 episode: 1873 score: 13.0 memory length: 479864 epsilon: 0.24786730000925927
 steps: 535 lr: 6.400000000000001e-06 evaluation reward: 8.84
 episode: 1874 score: 10.0 memory length: 480403 epsilon: 0.24680008000925252
 steps: 539 lr: 6.400000000000001e-06 evaluation reward: 8.89
 episode: 1875 score: 14.0 memory length: 481072 epsilon: 0.24547546000924414
 steps: 669 lr: 6.400000000000001e-06 evaluation reward: 8.94
 episode: 1876 score: 7.0 memory length: 481450 epsilon: 0.2447270200092394 s
 teps: 378 lr: 6.400000000000001e-06 evaluation reward: 8.93
 episode: 1877 score: 9.0 memory length: 481925 epsilon: 0.24378652000923345
 steps: 475 lr: 6.400000000000001e-06 evaluation reward: 8.94
 episode: 1878 score: 7.0 memory length: 482311 epsilon: 0.24302224000922862
 steps: 386 lr: 6.400000000000001e-06 evaluation reward: 8.9
 episode: 1879 score: 13.0 memory length: 482820 epsilon: 0.24201442000922224
 steps: 509 lr: 6.400000000000001e-06 evaluation reward: 8.92
 episode: 1880 score: 12.0 memory length: 483422 epsilon: 0.2408224600092147
 steps: 602 lr: 6.400000000000001e-06 evaluation reward: 8.91
 episode: 1881 score: 11.0 memory length: 483977 epsilon: 0.23972356000920775
 steps: 555 lr: 6.400000000000001e-06 evaluation reward: 8.95
 episode: 1882 score: 7.0 memory length: 484400 epsilon: 0.23888602000920245
 steps: 423 lr: 6.400000000000001e-06 evaluation reward: 8.92
 episode: 1883 score: 7.0 memory length: 484771 epsilon: 0.2381514400091978 s
 teps: 371 lr: 6.400000000000001e-06 evaluation reward: 8.91
 episode: 1884 score: 13.0 memory length: 485299 epsilon: 0.2371060000091912
 steps: 528 lr: 6.400000000000001e-06 evaluation reward: 8.94
 episode: 1885 score: 11.0 memory length: 485816 epsilon: 0.2360823400091847
 steps: 517 lr: 6.400000000000001e-06 evaluation reward: 8.99
 episode: 1886 score: 11.0 memory length: 486374 epsilon: 0.23497750000917772
 steps: 558 lr: 6.400000000000001e-06 evaluation reward: 8.98
 episode: 1887 score: 8.0 memory length: 486796 epsilon: 0.23414194000917243
 steps: 422 lr: 6.400000000000001e-06 evaluation reward: 8.97
 episode: 1888 score: 12.0 memory length: 487388 epsilon: 0.23296978000916502
 steps: 592 lr: 6.400000000000001e-06 evaluation reward: 8.99
 episode: 1889 score: 10.0 memory length: 487890 epsilon: 0.23197582000915873
 steps: 502 lr: 6.400000000000001e-06 evaluation reward: 9.0
 episode: 1890 score: 6.0 memory length: 488246 epsilon: 0.23127094000915427
 steps: 356 lr: 6.400000000000001e-06 evaluation reward: 9.01
 episode: 1891 score: 9.0 memory length: 488707 epsilon: 0.2303581600091485 s
 teps: 461 lr: 6.400000000000001e-06 evaluation reward: 8.93
 episode: 1892 score: 9.0 memory length: 489193 epsilon: 0.2293958800091424 s
 teps: 486 lr: 6.400000000000001e-06 evaluation reward: 8.95
 episode: 1893 score: 8.0 memory length: 489629 epsilon: 0.22853260000913694
 steps: 436 lr: 6.400000000000001e-06 evaluation reward: 8.95

episode: 1894 score: 9.0 memory length: 490097 epsilon: 0.22760596000913108
 steps: 468 lr: 6.400000000000000e-06 evaluation reward: 8.97
 episode: 1895 score: 11.0 memory length: 490613 epsilon: 0.22658428000912462
 steps: 516 lr: 6.400000000000000e-06 evaluation reward: 8.99
 episode: 1896 score: 7.0 memory length: 491019 epsilon: 0.22578040000911953
 steps: 406 lr: 6.400000000000000e-06 evaluation reward: 9.01
 episode: 1897 score: 5.0 memory length: 491325 epsilon: 0.2251745200091157 s
 teps: 306 lr: 6.400000000000000e-06 evaluation reward: 8.96
 episode: 1898 score: 7.0 memory length: 491713 epsilon: 0.22440628000911084
 steps: 388 lr: 6.400000000000000e-06 evaluation reward: 8.95
 episode: 1899 score: 9.0 memory length: 492209 epsilon: 0.22342420000910462
 steps: 496 lr: 6.400000000000000e-06 evaluation reward: 8.96
 episode: 1900 score: 9.0 memory length: 492668 epsilon: 0.22251538000909887
 steps: 459 lr: 6.400000000000000e-06 evaluation reward: 8.98
 episode: 1901 score: 9.0 memory length: 493101 epsilon: 0.22165804000909345
 steps: 433 lr: 6.400000000000000e-06 evaluation reward: 8.99
 episode: 1902 score: 8.0 memory length: 493525 epsilon: 0.22081852000908814
 steps: 424 lr: 6.400000000000000e-06 evaluation reward: 8.95
 episode: 1903 score: 16.0 memory length: 494201 epsilon: 0.21948004000907967
 steps: 676 lr: 6.400000000000000e-06 evaluation reward: 8.99
 episode: 1904 score: 10.0 memory length: 494756 epsilon: 0.21838114000907272
 steps: 555 lr: 6.400000000000000e-06 evaluation reward: 8.98
 episode: 1905 score: 8.0 memory length: 495228 epsilon: 0.2174465800090668 s
 teps: 472 lr: 6.400000000000000e-06 evaluation reward: 8.99
 episode: 1906 score: 6.0 memory length: 495566 epsilon: 0.21677734000906257
 steps: 338 lr: 6.400000000000000e-06 evaluation reward: 9.0
 episode: 1907 score: 6.0 memory length: 495923 epsilon: 0.2160704800090581 s
 teps: 357 lr: 6.400000000000000e-06 evaluation reward: 8.89
 episode: 1908 score: 11.0 memory length: 496361 epsilon: 0.2152032400090526
 steps: 438 lr: 6.400000000000000e-06 evaluation reward: 8.92
 episode: 1909 score: 8.0 memory length: 496798 epsilon: 0.21433798000904714
 steps: 437 lr: 6.400000000000000e-06 evaluation reward: 8.95
 episode: 1910 score: 6.0 memory length: 497157 epsilon: 0.21362716000904264
 steps: 359 lr: 6.400000000000000e-06 evaluation reward: 8.96
 episode: 1911 score: 9.0 memory length: 497587 epsilon: 0.21277576000903725
 steps: 430 lr: 6.400000000000000e-06 evaluation reward: 8.94
 episode: 1912 score: 6.0 memory length: 497912 epsilon: 0.21213226000903318
 steps: 325 lr: 6.400000000000000e-06 evaluation reward: 8.91
 episode: 1913 score: 15.0 memory length: 498549 epsilon: 0.210871000090252
 steps: 637 lr: 6.400000000000000e-06 evaluation reward: 8.99
 episode: 1914 score: 13.0 memory length: 499189 epsilon: 0.20960380000901718
 steps: 640 lr: 6.400000000000000e-06 evaluation reward: 9.0
 episode: 1915 score: 14.0 memory length: 499710 epsilon: 0.20857222000901066
 steps: 521 lr: 6.400000000000000e-06 evaluation reward: 9.07
 episode: 1916 score: 14.0 memory length: 500255 epsilon: 0.20749312000900383
 steps: 545 lr: 2.560000000000000e-06 evaluation reward: 9.11
 episode: 1917 score: 7.0 memory length: 500647 epsilon: 0.20671696000899892
 steps: 392 lr: 2.560000000000000e-06 evaluation reward: 9.11
 episode: 1918 score: 6.0 memory length: 500987 epsilon: 0.20604376000899466
 steps: 340 lr: 2.560000000000000e-06 evaluation reward: 9.09
 episode: 1919 score: 13.0 memory length: 501630 epsilon: 0.2047706200089866
 steps: 643 lr: 2.560000000000000e-06 evaluation reward: 9.16
 episode: 1920 score: 8.0 memory length: 502065 epsilon: 0.20390932000898115
 steps: 435 lr: 2.560000000000000e-06 evaluation reward: 9.13
 episode: 1921 score: 7.0 memory length: 502419 epsilon: 0.20320840000897672
 steps: 354 lr: 2.560000000000000e-06 evaluation reward: 9.11
 episode: 1922 score: 21.0 memory length: 503012 epsilon: 0.2020342600089693
 steps: 593 lr: 2.560000000000000e-06 evaluation reward: 9.2
 episode: 1923 score: 9.0 memory length: 503493 epsilon: 0.20108188000896327
 steps: 481 lr: 2.560000000000000e-06 evaluation reward: 9.15

```
episode: 1924 score: 12.0 memory length: 504055 epsilon: 0.19996912000895622
steps: 562 lr: 2.5600000000000001e-06 evaluation reward: 9.17
episode: 1925 score: 6.0 memory length: 504397 epsilon: 0.19929196000895194
steps: 342 lr: 2.5600000000000001e-06 evaluation reward: 9.16
episode: 1926 score: 14.0 memory length: 505005 epsilon: 0.19808812000894432
steps: 608 lr: 2.5600000000000001e-06 evaluation reward: 9.23
episode: 1927 score: 11.0 memory length: 505531 epsilon: 0.19704664000893773
steps: 526 lr: 2.5600000000000001e-06 evaluation reward: 9.25
episode: 1928 score: 9.0 memory length: 506010 epsilon: 0.19609822000893173
steps: 479 lr: 2.5600000000000001e-06 evaluation reward: 9.28
episode: 1929 score: 11.0 memory length: 506555 epsilon: 0.1950191200089249
steps: 545 lr: 2.5600000000000001e-06 evaluation reward: 9.35
episode: 1930 score: 9.0 memory length: 507052 epsilon: 0.19403506000891868
steps: 497 lr: 2.5600000000000001e-06 evaluation reward: 9.32
episode: 1931 score: 11.0 memory length: 507627 epsilon: 0.19289656000891148
steps: 575 lr: 2.5600000000000001e-06 evaluation reward: 9.36
episode: 1932 score: 8.0 memory length: 508063 epsilon: 0.19203328000890602
steps: 436 lr: 2.5600000000000001e-06 evaluation reward: 9.36
episode: 1933 score: 4.0 memory length: 508327 epsilon: 0.1915105600089027 s
teps: 264 lr: 2.5600000000000001e-06 evaluation reward: 9.35
episode: 1934 score: 9.0 memory length: 508790 epsilon: 0.1905938200088969 s
teps: 463 lr: 2.5600000000000001e-06 evaluation reward: 9.38
episode: 1935 score: 13.0 memory length: 509331 epsilon: 0.18952264000889013
steps: 541 lr: 2.5600000000000001e-06 evaluation reward: 9.4
episode: 1936 score: 11.0 memory length: 509746 epsilon: 0.18870094000888493
steps: 415 lr: 2.5600000000000001e-06 evaluation reward: 9.36
episode: 1937 score: 6.0 memory length: 510102 epsilon: 0.18799606000888047
steps: 356 lr: 2.5600000000000001e-06 evaluation reward: 9.32
episode: 1938 score: 11.0 memory length: 510506 epsilon: 0.1871961400088754
steps: 404 lr: 2.5600000000000001e-06 evaluation reward: 9.36
episode: 1939 score: 11.0 memory length: 511067 epsilon: 0.18608536000886838
steps: 561 lr: 2.5600000000000001e-06 evaluation reward: 9.38
episode: 1940 score: 10.0 memory length: 511567 epsilon: 0.18509536000886212
steps: 500 lr: 2.5600000000000001e-06 evaluation reward: 9.46
episode: 1941 score: 13.0 memory length: 512056 epsilon: 0.184127140008856 s
teps: 489 lr: 2.5600000000000001e-06 evaluation reward: 9.47
episode: 1942 score: 13.0 memory length: 512568 epsilon: 0.18311338000884958
steps: 512 lr: 2.5600000000000001e-06 evaluation reward: 9.49
episode: 1943 score: 9.0 memory length: 513045 epsilon: 0.1821689200088436 s
teps: 477 lr: 2.5600000000000001e-06 evaluation reward: 9.54
episode: 1944 score: 9.0 memory length: 513486 epsilon: 0.18129574000883808
steps: 441 lr: 2.5600000000000001e-06 evaluation reward: 9.57
episode: 1945 score: 6.0 memory length: 513824 epsilon: 0.18062650000883385
steps: 338 lr: 2.5600000000000001e-06 evaluation reward: 9.6
episode: 1946 score: 7.0 memory length: 514200 epsilon: 0.17988202000882914
steps: 376 lr: 2.5600000000000001e-06 evaluation reward: 9.55
episode: 1947 score: 11.0 memory length: 514764 epsilon: 0.17876530000882207
steps: 564 lr: 2.5600000000000001e-06 evaluation reward: 9.59
episode: 1948 score: 6.0 memory length: 515084 epsilon: 0.17813170000881806
steps: 320 lr: 2.5600000000000001e-06 evaluation reward: 9.61
episode: 1949 score: 5.0 memory length: 515412 epsilon: 0.17748226000881395
steps: 328 lr: 2.5600000000000001e-06 evaluation reward: 9.57
episode: 1950 score: 8.0 memory length: 515864 epsilon: 0.1765873000088083 s
teps: 452 lr: 2.5600000000000001e-06 evaluation reward: 9.56
episode: 1951 score: 22.0 memory length: 516590 epsilon: 0.1751498200087992
steps: 726 lr: 2.5600000000000001e-06 evaluation reward: 9.66
episode: 1952 score: 8.0 memory length: 516990 epsilon: 0.17435782000879418
steps: 400 lr: 2.5600000000000001e-06 evaluation reward: 9.66
episode: 1953 score: 9.0 memory length: 517494 epsilon: 0.17335990000878787
steps: 504 lr: 2.5600000000000001e-06 evaluation reward: 9.65
```

episode: 1954 score: 4.0 memory length: 517739 epsilon: 0.1728748000087848 s
 teps: 245 lr: 2.5600000000000001e-06 evaluation reward: 9.6
 episode: 1955 score: 12.0 memory length: 518295 epsilon: 0.17177392000877784
 steps: 556 lr: 2.5600000000000001e-06 evaluation reward: 9.56
 episode: 1956 score: 11.0 memory length: 518837 epsilon: 0.17070076000877105
 steps: 542 lr: 2.5600000000000001e-06 evaluation reward: 9.6
 episode: 1957 score: 13.0 memory length: 519321 epsilon: 0.16974244000876498
 steps: 484 lr: 2.5600000000000001e-06 evaluation reward: 9.67
 episode: 1958 score: 10.0 memory length: 519805 epsilon: 0.16878412000875892
 steps: 484 lr: 2.5600000000000001e-06 evaluation reward: 9.62
 episode: 1959 score: 5.0 memory length: 520103 epsilon: 0.1681940800087552 s
 teps: 298 lr: 2.5600000000000001e-06 evaluation reward: 9.59
 episode: 1960 score: 18.0 memory length: 520813 epsilon: 0.1667882800087463
 steps: 710 lr: 2.5600000000000001e-06 evaluation reward: 9.69
 episode: 1961 score: 12.0 memory length: 521404 epsilon: 0.1656181000087389
 steps: 591 lr: 2.5600000000000001e-06 evaluation reward: 9.74
 episode: 1962 score: 14.0 memory length: 521947 epsilon: 0.1645429600087321
 steps: 543 lr: 2.5600000000000001e-06 evaluation reward: 9.81
 episode: 1963 score: 10.0 memory length: 522426 epsilon: 0.16359454000872609
 steps: 479 lr: 2.5600000000000001e-06 evaluation reward: 9.8
 episode: 1964 score: 14.0 memory length: 523101 epsilon: 0.16225804000871763
 steps: 675 lr: 2.5600000000000001e-06 evaluation reward: 9.86
 episode: 1965 score: 5.0 memory length: 523426 epsilon: 0.16161454000871356
 steps: 325 lr: 2.5600000000000001e-06 evaluation reward: 9.84
 episode: 1966 score: 6.0 memory length: 523731 epsilon: 0.16101064000870974
 steps: 305 lr: 2.5600000000000001e-06 evaluation reward: 9.81
 episode: 1967 score: 12.0 memory length: 524173 epsilon: 0.1601354800087042
 steps: 442 lr: 2.5600000000000001e-06 evaluation reward: 9.84
 episode: 1968 score: 8.0 memory length: 524573 epsilon: 0.1593434800086992 s
 teps: 400 lr: 2.5600000000000001e-06 evaluation reward: 9.83
 episode: 1969 score: 11.0 memory length: 525147 epsilon: 0.158206960008692 s
 teps: 574 lr: 2.5600000000000001e-06 evaluation reward: 9.87
 episode: 1970 score: 10.0 memory length: 525630 epsilon: 0.15725062000868595
 steps: 483 lr: 2.5600000000000001e-06 evaluation reward: 9.87
 episode: 1971 score: 9.0 memory length: 526138 epsilon: 0.15624478000867958
 steps: 508 lr: 2.5600000000000001e-06 evaluation reward: 9.79
 episode: 1972 score: 6.0 memory length: 526476 epsilon: 0.15557554000867535
 steps: 338 lr: 2.5600000000000001e-06 evaluation reward: 9.76
 episode: 1973 score: 11.0 memory length: 527088 epsilon: 0.15436378000866768
 steps: 612 lr: 2.5600000000000001e-06 evaluation reward: 9.74
 episode: 1974 score: 9.0 memory length: 527587 epsilon: 0.15337576000866143
 steps: 499 lr: 2.5600000000000001e-06 evaluation reward: 9.73
 episode: 1975 score: 8.0 memory length: 527980 epsilon: 0.1525976200086565 s
 teps: 393 lr: 2.5600000000000001e-06 evaluation reward: 9.67
 episode: 1976 score: 6.0 memory length: 528300 epsilon: 0.1519640200086525 s
 teps: 320 lr: 2.5600000000000001e-06 evaluation reward: 9.66
 episode: 1977 score: 8.0 memory length: 528706 epsilon: 0.15116014000864741
 steps: 406 lr: 2.5600000000000001e-06 evaluation reward: 9.65
 episode: 1978 score: 10.0 memory length: 529221 epsilon: 0.15014044000864096
 steps: 515 lr: 2.5600000000000001e-06 evaluation reward: 9.68
 episode: 1979 score: 8.0 memory length: 529681 epsilon: 0.1492296400086352 s
 teps: 460 lr: 2.5600000000000001e-06 evaluation reward: 9.63
 episode: 1980 score: 6.0 memory length: 530022 epsilon: 0.14855446000863093
 steps: 341 lr: 2.5600000000000001e-06 evaluation reward: 9.57
 episode: 1981 score: 9.0 memory length: 530512 epsilon: 0.1475842600086248 s
 teps: 490 lr: 2.5600000000000001e-06 evaluation reward: 9.55
 episode: 1982 score: 9.0 memory length: 530982 epsilon: 0.1466536600086189 s
 teps: 470 lr: 2.5600000000000001e-06 evaluation reward: 9.57
 episode: 1983 score: 8.0 memory length: 531387 epsilon: 0.14585176000861383
 steps: 405 lr: 2.5600000000000001e-06 evaluation reward: 9.58

episode: 1984 score: 10.0 memory length: 531881 epsilon: 0.14487364000860764
steps: 494 lr: 2.5600000000000001e-06 evaluation reward: 9.55
episode: 1985 score: 12.0 memory length: 532284 epsilon: 0.1440757000086026
steps: 403 lr: 2.5600000000000001e-06 evaluation reward: 9.56
episode: 1986 score: 14.0 memory length: 532828 epsilon: 0.14299858000859578
steps: 544 lr: 2.5600000000000001e-06 evaluation reward: 9.59
episode: 1987 score: 5.0 memory length: 533139 epsilon: 0.14238280000859188
steps: 311 lr: 2.5600000000000001e-06 evaluation reward: 9.56
episode: 1988 score: 10.0 memory length: 533655 epsilon: 0.14136112000858542
steps: 516 lr: 2.5600000000000001e-06 evaluation reward: 9.54
episode: 1989 score: 7.0 memory length: 534026 epsilon: 0.14062654000858077
steps: 371 lr: 2.5600000000000001e-06 evaluation reward: 9.51
episode: 1990 score: 8.0 memory length: 534465 epsilon: 0.13975732000857527
steps: 439 lr: 2.5600000000000001e-06 evaluation reward: 9.53
episode: 1991 score: 11.0 memory length: 535014 epsilon: 0.1386703000085684
steps: 549 lr: 2.5600000000000001e-06 evaluation reward: 9.55
episode: 1992 score: 6.0 memory length: 535375 epsilon: 0.13795552000856387
steps: 361 lr: 2.5600000000000001e-06 evaluation reward: 9.52
episode: 1993 score: 9.0 memory length: 535832 epsilon: 0.13705066000855815
steps: 457 lr: 2.5600000000000001e-06 evaluation reward: 9.53
episode: 1994 score: 15.0 memory length: 536346 epsilon: 0.1360329400085517
steps: 514 lr: 2.5600000000000001e-06 evaluation reward: 9.59
episode: 1995 score: 13.0 memory length: 536834 epsilon: 0.1350667000085456
steps: 488 lr: 2.5600000000000001e-06 evaluation reward: 9.61
episode: 1996 score: 7.0 memory length: 537208 epsilon: 0.1343261800085409 s
steps: 374 lr: 2.5600000000000001e-06 evaluation reward: 9.61
episode: 1997 score: 6.0 memory length: 537573 epsilon: 0.13360348000853633
steps: 365 lr: 2.5600000000000001e-06 evaluation reward: 9.62
episode: 1998 score: 6.0 memory length: 537931 epsilon: 0.13289464000853185
steps: 358 lr: 2.5600000000000001e-06 evaluation reward: 9.61
episode: 1999 score: 9.0 memory length: 538418 epsilon: 0.13193038000852575
steps: 487 lr: 2.5600000000000001e-06 evaluation reward: 9.61
episode: 2000 score: 11.0 memory length: 539025 epsilon: 0.13072852000851815
steps: 607 lr: 2.5600000000000001e-06 evaluation reward: 9.63
episode: 2001 score: 12.0 memory length: 539569 epsilon: 0.12965140000851133
steps: 544 lr: 2.5600000000000001e-06 evaluation reward: 9.66
episode: 2002 score: 6.0 memory length: 539912 epsilon: 0.12897226000850703
steps: 343 lr: 2.5600000000000001e-06 evaluation reward: 9.64
episode: 2003 score: 11.0 memory length: 540418 epsilon: 0.1279703800085007
steps: 506 lr: 2.5600000000000001e-06 evaluation reward: 9.59
episode: 2004 score: 11.0 memory length: 540919 epsilon: 0.12697840000849442
steps: 501 lr: 2.5600000000000001e-06 evaluation reward: 9.6
episode: 2005 score: 7.0 memory length: 541270 epsilon: 0.12628342000849002
steps: 351 lr: 2.5600000000000001e-06 evaluation reward: 9.59
episode: 2006 score: 11.0 memory length: 541791 epsilon: 0.1252518400084835
steps: 521 lr: 2.5600000000000001e-06 evaluation reward: 9.64
episode: 2007 score: 17.0 memory length: 542466 epsilon: 0.12391534000848264
steps: 675 lr: 2.5600000000000001e-06 evaluation reward: 9.75
episode: 2008 score: 17.0 memory length: 543086 epsilon: 0.12268774000848348
steps: 620 lr: 2.5600000000000001e-06 evaluation reward: 9.81
episode: 2009 score: 10.0 memory length: 543601 epsilon: 0.12166804000848418
steps: 515 lr: 2.5600000000000001e-06 evaluation reward: 9.83
episode: 2010 score: 9.0 memory length: 544083 epsilon: 0.12071368000848483
steps: 482 lr: 2.5600000000000001e-06 evaluation reward: 9.86
episode: 2011 score: 9.0 memory length: 544517 epsilon: 0.11985436000848541
steps: 434 lr: 2.5600000000000001e-06 evaluation reward: 9.86
episode: 2012 score: 14.0 memory length: 545057 epsilon: 0.11878516000848614
steps: 540 lr: 2.5600000000000001e-06 evaluation reward: 9.94
episode: 2013 score: 14.0 memory length: 545590 epsilon: 0.11772982000848686
steps: 533 lr: 2.5600000000000001e-06 evaluation reward: 9.93

episode: 2014 score: 8.0 memory length: 546010 epsilon: 0.11689822000848743
 steps: 420 lr: 2.5600000000000001e-06 evaluation reward: 9.88
 episode: 2015 score: 9.0 memory length: 546498 epsilon: 0.11593198000848809
 steps: 488 lr: 2.5600000000000001e-06 evaluation reward: 9.83
 episode: 2016 score: 6.0 memory length: 546834 epsilon: 0.11526670000848854
 steps: 336 lr: 2.5600000000000001e-06 evaluation reward: 9.75
 episode: 2017 score: 6.0 memory length: 547154 epsilon: 0.11463310000848898
 steps: 320 lr: 2.5600000000000001e-06 evaluation reward: 9.74
 episode: 2018 score: 12.0 memory length: 547772 epsilon: 0.11340946000848981
 steps: 618 lr: 2.5600000000000001e-06 evaluation reward: 9.8
 episode: 2019 score: 9.0 memory length: 548232 epsilon: 0.11249866000849043
 steps: 460 lr: 2.5600000000000001e-06 evaluation reward: 9.76
 episode: 2020 score: 16.0 memory length: 548827 epsilon: 0.11132056000849123
 steps: 595 lr: 2.5600000000000001e-06 evaluation reward: 9.84
 episode: 2021 score: 10.0 memory length: 549359 epsilon: 0.11026720000849195
 steps: 532 lr: 2.5600000000000001e-06 evaluation reward: 9.87
 episode: 2022 score: 10.0 memory length: 549870 epsilon: 0.10925542000849264
 steps: 511 lr: 2.5600000000000001e-06 evaluation reward: 9.76
 episode: 2023 score: 9.0 memory length: 550292 epsilon: 0.10841986000849321
 steps: 422 lr: 2.5600000000000001e-06 evaluation reward: 9.76
 episode: 2024 score: 14.0 memory length: 550947 epsilon: 0.1071229600084941
 steps: 655 lr: 2.5600000000000001e-06 evaluation reward: 9.78
 episode: 2025 score: 8.0 memory length: 551383 epsilon: 0.10625968000849469
 steps: 436 lr: 2.5600000000000001e-06 evaluation reward: 9.8
 episode: 2026 score: 11.0 memory length: 551887 epsilon: 0.10526176000849537
 steps: 504 lr: 2.5600000000000001e-06 evaluation reward: 9.77
 episode: 2027 score: 8.0 memory length: 552289 epsilon: 0.10446580000849591
 steps: 402 lr: 2.5600000000000001e-06 evaluation reward: 9.74
 episode: 2028 score: 12.0 memory length: 552810 epsilon: 0.10343422000849661
 steps: 521 lr: 2.5600000000000001e-06 evaluation reward: 9.77
 episode: 2029 score: 5.0 memory length: 553123 epsilon: 0.10281448000849704
 steps: 313 lr: 2.5600000000000001e-06 evaluation reward: 9.71
 episode: 2030 score: 7.0 memory length: 553509 epsilon: 0.10205020000849756
 steps: 386 lr: 2.5600000000000001e-06 evaluation reward: 9.69
 episode: 2031 score: 8.0 memory length: 553939 epsilon: 0.10119880000849814
 steps: 430 lr: 2.5600000000000001e-06 evaluation reward: 9.66
 episode: 2032 score: 11.0 memory length: 554485 epsilon: 0.10011772000849888
 steps: 546 lr: 2.5600000000000001e-06 evaluation reward: 9.69
 episode: 2033 score: 6.0 memory length: 554825 epsilon: 0.09944452000849933
 steps: 340 lr: 2.5600000000000001e-06 evaluation reward: 9.71
 episode: 2034 score: 11.0 memory length: 555372 epsilon: 0.09836146000850007
 steps: 547 lr: 2.5600000000000001e-06 evaluation reward: 9.73
 episode: 2035 score: 11.0 memory length: 555920 epsilon: 0.09727642000850081
 steps: 548 lr: 2.5600000000000001e-06 evaluation reward: 9.71
 episode: 2036 score: 11.0 memory length: 556473 epsilon: 0.09618148000850156
 steps: 553 lr: 2.5600000000000001e-06 evaluation reward: 9.71
 episode: 2037 score: 8.0 memory length: 556876 epsilon: 0.0953835400085021 s
 steps: 403 lr: 2.5600000000000001e-06 evaluation reward: 9.73
 episode: 2038 score: 6.0 memory length: 557256 epsilon: 0.09463114000850262
 steps: 380 lr: 2.5600000000000001e-06 evaluation reward: 9.68
 episode: 2039 score: 11.0 memory length: 557754 epsilon: 0.09364510000850329
 steps: 498 lr: 2.5600000000000001e-06 evaluation reward: 9.68
 episode: 2040 score: 13.0 memory length: 558378 epsilon: 0.09240958000850413
 steps: 624 lr: 2.5600000000000001e-06 evaluation reward: 9.71
 episode: 2041 score: 11.0 memory length: 558909 epsilon: 0.09135820000850485
 steps: 531 lr: 2.5600000000000001e-06 evaluation reward: 9.69
 episode: 2042 score: 12.0 memory length: 559346 epsilon: 0.09049294000850544
 steps: 437 lr: 2.5600000000000001e-06 evaluation reward: 9.68
 episode: 2043 score: 13.0 memory length: 559954 epsilon: 0.08928910000850626
 steps: 608 lr: 2.5600000000000001e-06 evaluation reward: 9.72

```
episode: 2044 score: 10.0 memory length: 560418 epsilon: 0.08837038000850689
steps: 464 lr: 2.5600000000000001e-06 evaluation reward: 9.73
episode: 2045 score: 10.0 memory length: 560867 epsilon: 0.0874813600085075
steps: 449 lr: 2.5600000000000001e-06 evaluation reward: 9.77
episode: 2046 score: 6.0 memory length: 561223 epsilon: 0.08677648000850798
steps: 356 lr: 2.5600000000000001e-06 evaluation reward: 9.76
episode: 2047 score: 12.0 memory length: 561804 epsilon: 0.08562610000850876
steps: 581 lr: 2.5600000000000001e-06 evaluation reward: 9.77
episode: 2048 score: 10.0 memory length: 562256 epsilon: 0.08473114000850937
steps: 452 lr: 2.5600000000000001e-06 evaluation reward: 9.81
episode: 2049 score: 13.0 memory length: 562818 epsilon: 0.08361838000851013
steps: 562 lr: 2.5600000000000001e-06 evaluation reward: 9.89
episode: 2050 score: 10.0 memory length: 563342 epsilon: 0.08258086000851084
steps: 524 lr: 2.5600000000000001e-06 evaluation reward: 9.91
episode: 2051 score: 19.0 memory length: 564008 epsilon: 0.08126218000851174
steps: 666 lr: 2.5600000000000001e-06 evaluation reward: 9.88
episode: 2052 score: 13.0 memory length: 564646 epsilon: 0.0799989400085126
steps: 638 lr: 2.5600000000000001e-06 evaluation reward: 9.93
episode: 2053 score: 7.0 memory length: 565034 epsilon: 0.07923070000851312
steps: 388 lr: 2.5600000000000001e-06 evaluation reward: 9.91
episode: 2054 score: 7.0 memory length: 565423 epsilon: 0.07846048000851365
steps: 389 lr: 2.5600000000000001e-06 evaluation reward: 9.94
episode: 2055 score: 12.0 memory length: 565955 epsilon: 0.07740712000851437
steps: 532 lr: 2.5600000000000001e-06 evaluation reward: 9.94
episode: 2056 score: 10.0 memory length: 566444 epsilon: 0.07643890000851503
steps: 489 lr: 2.5600000000000001e-06 evaluation reward: 9.93
episode: 2057 score: 10.0 memory length: 567018 epsilon: 0.0753023800085158
steps: 574 lr: 2.5600000000000001e-06 evaluation reward: 9.9
episode: 2058 score: 12.0 memory length: 567569 epsilon: 0.07421140000851655
steps: 551 lr: 2.5600000000000001e-06 evaluation reward: 9.92
episode: 2059 score: 12.0 memory length: 568036 epsilon: 0.07328674000851718
steps: 467 lr: 2.5600000000000001e-06 evaluation reward: 9.99
episode: 2060 score: 9.0 memory length: 568507 epsilon: 0.07235416000851781
steps: 471 lr: 2.5600000000000001e-06 evaluation reward: 9.9
episode: 2061 score: 9.0 memory length: 568938 epsilon: 0.0715007800085184 s
steps: 431 lr: 2.5600000000000001e-06 evaluation reward: 9.87
episode: 2062 score: 8.0 memory length: 569327 epsilon: 0.07073056000851892
steps: 389 lr: 2.5600000000000001e-06 evaluation reward: 9.81
episode: 2063 score: 12.0 memory length: 569905 epsilon: 0.0695861200085197
steps: 578 lr: 2.5600000000000001e-06 evaluation reward: 9.83
episode: 2064 score: 10.0 memory length: 570387 epsilon: 0.06863176000852035
steps: 482 lr: 2.5600000000000001e-06 evaluation reward: 9.79
episode: 2065 score: 11.0 memory length: 570913 epsilon: 0.06759028000852106
steps: 526 lr: 2.5600000000000001e-06 evaluation reward: 9.85
episode: 2066 score: 13.0 memory length: 571553 epsilon: 0.06632308000852193
steps: 640 lr: 2.5600000000000001e-06 evaluation reward: 9.92
episode: 2067 score: 13.0 memory length: 572158 epsilon: 0.06512518000852274
steps: 605 lr: 2.5600000000000001e-06 evaluation reward: 9.93
episode: 2068 score: 11.0 memory length: 572731 epsilon: 0.06399064000852352
steps: 573 lr: 2.5600000000000001e-06 evaluation reward: 9.96
episode: 2069 score: 5.0 memory length: 573042 epsilon: 0.06337486000852394
steps: 311 lr: 2.5600000000000001e-06 evaluation reward: 9.9
episode: 2070 score: 12.0 memory length: 573636 epsilon: 0.06219874000852474
steps: 594 lr: 2.5600000000000001e-06 evaluation reward: 9.92
episode: 2071 score: 12.0 memory length: 574210 epsilon: 0.061062220008525514
steps: 574 lr: 2.5600000000000001e-06 evaluation reward: 9.95
episode: 2072 score: 7.0 memory length: 574627 epsilon: 0.06023656000852608
steps: 417 lr: 2.5600000000000001e-06 evaluation reward: 9.96
episode: 2073 score: 19.0 memory length: 575354 epsilon: 0.05879710000852706
steps: 727 lr: 2.5600000000000001e-06 evaluation reward: 10.04
```

```
episode: 2074 score: 8.0 memory length: 575740 epsilon: 0.05803282000852758
steps: 386 lr: 2.560000000000000e-06 evaluation reward: 10.03
episode: 2075 score: 7.0 memory length: 576145 epsilon: 0.05723092000852813
steps: 405 lr: 2.560000000000000e-06 evaluation reward: 10.02
episode: 2076 score: 10.0 memory length: 576670 epsilon: 0.056191420008528836
steps: 525 lr: 2.560000000000000e-06 evaluation reward: 10.06
episode: 2077 score: 8.0 memory length: 577089 epsilon: 0.0553618000085294 s
steps: 419 lr: 2.560000000000000e-06 evaluation reward: 10.06
episode: 2078 score: 11.0 memory length: 577643 epsilon: 0.05426488000853015
steps: 554 lr: 2.560000000000000e-06 evaluation reward: 10.07
episode: 2079 score: 11.0 memory length: 578190 epsilon: 0.05318182000853089
steps: 547 lr: 2.560000000000000e-06 evaluation reward: 10.1
episode: 2080 score: 5.0 memory length: 578537 epsilon: 0.05249476000853136
steps: 347 lr: 2.560000000000000e-06 evaluation reward: 10.09
episode: 2081 score: 10.0 memory length: 579065 epsilon: 0.05144932000853207
steps: 528 lr: 2.560000000000000e-06 evaluation reward: 10.1
episode: 2082 score: 6.0 memory length: 579423 epsilon: 0.050740480008532554
steps: 358 lr: 2.560000000000000e-06 evaluation reward: 10.07
episode: 2083 score: 16.0 memory length: 580113 epsilon: 0.049374280008533486
steps: 690 lr: 2.560000000000000e-06 evaluation reward: 10.15
episode: 2084 score: 9.0 memory length: 580564 epsilon: 0.048481300008534095
steps: 451 lr: 2.560000000000000e-06 evaluation reward: 10.14
episode: 2085 score: 10.0 memory length: 581047 epsilon: 0.04752496000853475
steps: 483 lr: 2.560000000000000e-06 evaluation reward: 10.12
episode: 2086 score: 12.0 memory length: 581466 epsilon: 0.04669534000853531
steps: 419 lr: 2.560000000000000e-06 evaluation reward: 10.1
episode: 2087 score: 14.0 memory length: 582007 epsilon: 0.045624160008536044
steps: 541 lr: 2.560000000000000e-06 evaluation reward: 10.19
episode: 2088 score: 9.0 memory length: 582418 epsilon: 0.0448103800085366 s
steps: 411 lr: 2.560000000000000e-06 evaluation reward: 10.18
episode: 2089 score: 10.0 memory length: 582883 epsilon: 0.04388968000853723
steps: 465 lr: 2.560000000000000e-06 evaluation reward: 10.21
episode: 2090 score: 8.0 memory length: 583263 epsilon: 0.04313728000853774
steps: 380 lr: 2.560000000000000e-06 evaluation reward: 10.21
episode: 2091 score: 8.0 memory length: 583643 epsilon: 0.04238488000853825
steps: 380 lr: 2.560000000000000e-06 evaluation reward: 10.18
episode: 2092 score: 10.0 memory length: 584122 epsilon: 0.0414364600085389
steps: 479 lr: 2.560000000000000e-06 evaluation reward: 10.22
episode: 2093 score: 7.0 memory length: 584532 epsilon: 0.040624660008539454
steps: 410 lr: 2.560000000000000e-06 evaluation reward: 10.2
episode: 2094 score: 13.0 memory length: 585144 epsilon: 0.03941290000854028
steps: 612 lr: 2.560000000000000e-06 evaluation reward: 10.18
episode: 2095 score: 11.0 memory length: 585643 epsilon: 0.038424880008540954
steps: 499 lr: 2.560000000000000e-06 evaluation reward: 10.16
episode: 2096 score: 8.0 memory length: 586088 epsilon: 0.037543780008541555
steps: 445 lr: 2.560000000000000e-06 evaluation reward: 10.17
episode: 2097 score: 8.0 memory length: 586524 epsilon: 0.036680500008542144
steps: 436 lr: 2.560000000000000e-06 evaluation reward: 10.19
episode: 2098 score: 12.0 memory length: 587027 epsilon: 0.03568456000854282
steps: 503 lr: 2.560000000000000e-06 evaluation reward: 10.25
episode: 2099 score: 16.0 memory length: 587582 epsilon: 0.03458566000854357
steps: 555 lr: 2.560000000000000e-06 evaluation reward: 10.32
episode: 2100 score: 12.0 memory length: 588143 epsilon: 0.03347488000854433
steps: 561 lr: 2.560000000000000e-06 evaluation reward: 10.33
episode: 2101 score: 16.0 memory length: 588709 epsilon: 0.032354200008545095
steps: 566 lr: 2.560000000000000e-06 evaluation reward: 10.37
episode: 2102 score: 6.0 memory length: 589049 epsilon: 0.031681000008545554
steps: 340 lr: 2.560000000000000e-06 evaluation reward: 10.37
episode: 2103 score: 19.0 memory length: 589757 epsilon: 0.03027916000854651
steps: 708 lr: 2.560000000000000e-06 evaluation reward: 10.45
```

```
episode: 2104 score: 19.0 memory length: 590385 epsilon: 0.029035720008547358
steps: 628 lr: 2.5600000000000001e-06 evaluation reward: 10.53
episode: 2105 score: 11.0 memory length: 590994 epsilon: 0.02782990000854818
steps: 609 lr: 2.5600000000000001e-06 evaluation reward: 10.57
episode: 2106 score: 14.0 memory length: 591619 epsilon: 0.026592400008549025
steps: 625 lr: 2.5600000000000001e-06 evaluation reward: 10.6
episode: 2107 score: 16.0 memory length: 592224 epsilon: 0.025394500008549842
steps: 605 lr: 2.5600000000000001e-06 evaluation reward: 10.59
episode: 2108 score: 9.0 memory length: 592703 epsilon: 0.02444608000855049
steps: 479 lr: 2.5600000000000001e-06 evaluation reward: 10.51
episode: 2109 score: 6.0 memory length: 593046 epsilon: 0.023766940008550952
steps: 343 lr: 2.5600000000000001e-06 evaluation reward: 10.47
episode: 2110 score: 9.0 memory length: 593505 epsilon: 0.022858120008551572
steps: 459 lr: 2.5600000000000001e-06 evaluation reward: 10.47
episode: 2111 score: 17.0 memory length: 594124 epsilon: 0.021632500008552408
steps: 619 lr: 2.5600000000000001e-06 evaluation reward: 10.55
episode: 2112 score: 15.0 memory length: 594761 epsilon: 0.020371240008553268
steps: 637 lr: 2.5600000000000001e-06 evaluation reward: 10.56
episode: 2113 score: 17.0 memory length: 595434 epsilon: 0.019038700008554177
steps: 673 lr: 2.5600000000000001e-06 evaluation reward: 10.59
episode: 2114 score: 19.0 memory length: 596145 epsilon: 0.017630920008555137
steps: 711 lr: 2.5600000000000001e-06 evaluation reward: 10.7
episode: 2115 score: 15.0 memory length: 596573 epsilon: 0.016783480008555715
steps: 428 lr: 2.5600000000000001e-06 evaluation reward: 10.76
episode: 2116 score: 17.0 memory length: 597224 epsilon: 0.01549450000855648
steps: 651 lr: 2.5600000000000001e-06 evaluation reward: 10.87
episode: 2117 score: 8.0 memory length: 597674 epsilon: 0.014603500008556307
steps: 450 lr: 2.5600000000000001e-06 evaluation reward: 10.89
episode: 2118 score: 14.0 memory length: 598235 epsilon: 0.013492720008556091
steps: 561 lr: 2.5600000000000001e-06 evaluation reward: 10.91
episode: 2119 score: 12.0 memory length: 598858 epsilon: 0.012259180008555852
steps: 623 lr: 2.5600000000000001e-06 evaluation reward: 10.94
episode: 2120 score: 15.0 memory length: 599607 epsilon: 0.010776160008555564
steps: 749 lr: 2.5600000000000001e-06 evaluation reward: 10.93
episode: 2121 score: 9.0 memory length: 600094 epsilon: 0.009998020008555413
steps: 487 lr: 1.0240000000000005e-06 evaluation reward: 10.92
episode: 2122 score: 9.0 memory length: 600591 epsilon: 0.009998020008555413
steps: 497 lr: 1.0240000000000005e-06 evaluation reward: 10.91
episode: 2123 score: 8.0 memory length: 601027 epsilon: 0.009998020008555413
steps: 436 lr: 1.0240000000000005e-06 evaluation reward: 10.9
episode: 2124 score: 15.0 memory length: 601592 epsilon: 0.009998020008555413
steps: 565 lr: 1.0240000000000005e-06 evaluation reward: 10.91
episode: 2125 score: 11.0 memory length: 602016 epsilon: 0.009998020008555413
steps: 424 lr: 1.0240000000000005e-06 evaluation reward: 10.94
episode: 2126 score: 10.0 memory length: 602568 epsilon: 0.009998020008555413
steps: 552 lr: 1.0240000000000005e-06 evaluation reward: 10.93
episode: 2127 score: 7.0 memory length: 602972 epsilon: 0.009998020008555413
steps: 404 lr: 1.0240000000000005e-06 evaluation reward: 10.92
episode: 2128 score: 15.0 memory length: 603549 epsilon: 0.009998020008555413
steps: 577 lr: 1.0240000000000005e-06 evaluation reward: 10.95
episode: 2129 score: 7.0 memory length: 603953 epsilon: 0.009998020008555413
steps: 404 lr: 1.0240000000000005e-06 evaluation reward: 10.97
episode: 2130 score: 11.0 memory length: 604537 epsilon: 0.009998020008555413
steps: 584 lr: 1.0240000000000005e-06 evaluation reward: 11.01
episode: 2131 score: 12.0 memory length: 605104 epsilon: 0.009998020008555413
steps: 567 lr: 1.0240000000000005e-06 evaluation reward: 11.05
episode: 2132 score: 10.0 memory length: 605599 epsilon: 0.009998020008555413
steps: 495 lr: 1.0240000000000005e-06 evaluation reward: 11.04
episode: 2133 score: 8.0 memory length: 606037 epsilon: 0.009998020008555413
steps: 438 lr: 1.0240000000000005e-06 evaluation reward: 11.06
```

```
episode: 2134 score: 13.0 memory length: 606606 epsilon: 0.009998020008555413
steps: 569 lr: 1.0240000000000005e-06 evaluation reward: 11.08
episode: 2135 score: 10.0 memory length: 607091 epsilon: 0.009998020008555413
steps: 485 lr: 1.0240000000000005e-06 evaluation reward: 11.07
episode: 2136 score: 8.0 memory length: 607509 epsilon: 0.009998020008555413
steps: 418 lr: 1.0240000000000005e-06 evaluation reward: 11.04
episode: 2137 score: 8.0 memory length: 607908 epsilon: 0.009998020008555413
steps: 399 lr: 1.0240000000000005e-06 evaluation reward: 11.04
episode: 2138 score: 22.0 memory length: 608627 epsilon: 0.009998020008555413
steps: 719 lr: 1.0240000000000005e-06 evaluation reward: 11.2
episode: 2139 score: 15.0 memory length: 609166 epsilon: 0.009998020008555413
steps: 539 lr: 1.0240000000000005e-06 evaluation reward: 11.24
episode: 2140 score: 8.0 memory length: 609620 epsilon: 0.009998020008555413
steps: 454 lr: 1.0240000000000005e-06 evaluation reward: 11.19
episode: 2141 score: 12.0 memory length: 610196 epsilon: 0.009998020008555413
steps: 576 lr: 1.0240000000000005e-06 evaluation reward: 11.2
episode: 2142 score: 19.0 memory length: 610872 epsilon: 0.009998020008555413
steps: 676 lr: 1.0240000000000005e-06 evaluation reward: 11.27
episode: 2143 score: 16.0 memory length: 611480 epsilon: 0.009998020008555413
steps: 608 lr: 1.0240000000000005e-06 evaluation reward: 11.3
episode: 2144 score: 20.0 memory length: 612151 epsilon: 0.009998020008555413
steps: 671 lr: 1.0240000000000005e-06 evaluation reward: 11.4
episode: 2145 score: 17.0 memory length: 612817 epsilon: 0.009998020008555413
steps: 666 lr: 1.0240000000000005e-06 evaluation reward: 11.47
episode: 2146 score: 15.0 memory length: 613446 epsilon: 0.009998020008555413
steps: 629 lr: 1.0240000000000005e-06 evaluation reward: 11.56
episode: 2147 score: 13.0 memory length: 614092 epsilon: 0.009998020008555413
steps: 646 lr: 1.0240000000000005e-06 evaluation reward: 11.57
episode: 2148 score: 8.0 memory length: 614509 epsilon: 0.009998020008555413
steps: 417 lr: 1.0240000000000005e-06 evaluation reward: 11.55
episode: 2149 score: 11.0 memory length: 615089 epsilon: 0.009998020008555413
steps: 580 lr: 1.0240000000000005e-06 evaluation reward: 11.53
episode: 2150 score: 13.0 memory length: 615715 epsilon: 0.009998020008555413
steps: 626 lr: 1.0240000000000005e-06 evaluation reward: 11.56
episode: 2151 score: 12.0 memory length: 616324 epsilon: 0.009998020008555413
steps: 609 lr: 1.0240000000000005e-06 evaluation reward: 11.49
episode: 2152 score: 10.0 memory length: 616807 epsilon: 0.009998020008555413
steps: 483 lr: 1.0240000000000005e-06 evaluation reward: 11.46
episode: 2153 score: 12.0 memory length: 617322 epsilon: 0.009998020008555413
steps: 515 lr: 1.0240000000000005e-06 evaluation reward: 11.51
episode: 2154 score: 12.0 memory length: 617873 epsilon: 0.009998020008555413
steps: 551 lr: 1.0240000000000005e-06 evaluation reward: 11.56
episode: 2155 score: 12.0 memory length: 618438 epsilon: 0.009998020008555413
steps: 565 lr: 1.0240000000000005e-06 evaluation reward: 11.56
episode: 2156 score: 6.0 memory length: 618798 epsilon: 0.009998020008555413
steps: 360 lr: 1.0240000000000005e-06 evaluation reward: 11.52
episode: 2157 score: 10.0 memory length: 619316 epsilon: 0.009998020008555413
steps: 518 lr: 1.0240000000000005e-06 evaluation reward: 11.52
episode: 2158 score: 11.0 memory length: 619864 epsilon: 0.009998020008555413
steps: 548 lr: 1.0240000000000005e-06 evaluation reward: 11.51
episode: 2159 score: 12.0 memory length: 620441 epsilon: 0.009998020008555413
steps: 577 lr: 1.0240000000000005e-06 evaluation reward: 11.51
episode: 2160 score: 10.0 memory length: 620958 epsilon: 0.009998020008555413
steps: 517 lr: 1.0240000000000005e-06 evaluation reward: 11.52
episode: 2161 score: 13.0 memory length: 621595 epsilon: 0.009998020008555413
steps: 637 lr: 1.0240000000000005e-06 evaluation reward: 11.56
episode: 2162 score: 12.0 memory length: 622186 epsilon: 0.009998020008555413
steps: 591 lr: 1.0240000000000005e-06 evaluation reward: 11.6
episode: 2163 score: 12.0 memory length: 622791 epsilon: 0.009998020008555413
steps: 605 lr: 1.0240000000000005e-06 evaluation reward: 11.6
```

episode: 2164 score: 9.0 memory length: 623259 epsilon: 0.009998020008555413
steps: 468 lr: 1.0240000000000005e-06 evaluation reward: 11.59
episode: 2165 score: 6.0 memory length: 623581 epsilon: 0.009998020008555413
steps: 322 lr: 1.0240000000000005e-06 evaluation reward: 11.54
episode: 2166 score: 11.0 memory length: 624074 epsilon: 0.009998020008555413
steps: 493 lr: 1.0240000000000005e-06 evaluation reward: 11.52
episode: 2167 score: 12.0 memory length: 624630 epsilon: 0.009998020008555413
steps: 556 lr: 1.0240000000000005e-06 evaluation reward: 11.51
episode: 2168 score: 12.0 memory length: 625126 epsilon: 0.009998020008555413
steps: 496 lr: 1.0240000000000005e-06 evaluation reward: 11.52
episode: 2169 score: 10.0 memory length: 625639 epsilon: 0.009998020008555413
steps: 513 lr: 1.0240000000000005e-06 evaluation reward: 11.57
episode: 2170 score: 15.0 memory length: 626319 epsilon: 0.009998020008555413
steps: 680 lr: 1.0240000000000005e-06 evaluation reward: 11.6
episode: 2171 score: 6.0 memory length: 626681 epsilon: 0.009998020008555413
steps: 362 lr: 1.0240000000000005e-06 evaluation reward: 11.54
episode: 2172 score: 5.0 memory length: 626957 epsilon: 0.009998020008555413
steps: 276 lr: 1.0240000000000005e-06 evaluation reward: 11.52
episode: 2173 score: 12.0 memory length: 627494 epsilon: 0.009998020008555413
steps: 537 lr: 1.0240000000000005e-06 evaluation reward: 11.45
episode: 2174 score: 9.0 memory length: 628017 epsilon: 0.009998020008555413
steps: 523 lr: 1.0240000000000005e-06 evaluation reward: 11.46
episode: 2175 score: 8.0 memory length: 628456 epsilon: 0.009998020008555413
steps: 439 lr: 1.0240000000000005e-06 evaluation reward: 11.47
episode: 2176 score: 18.0 memory length: 629125 epsilon: 0.009998020008555413
steps: 669 lr: 1.0240000000000005e-06 evaluation reward: 11.55
episode: 2177 score: 9.0 memory length: 629553 epsilon: 0.009998020008555413
steps: 428 lr: 1.0240000000000005e-06 evaluation reward: 11.56
episode: 2178 score: 9.0 memory length: 629982 epsilon: 0.009998020008555413
steps: 429 lr: 1.0240000000000005e-06 evaluation reward: 11.54
episode: 2179 score: 14.0 memory length: 630508 epsilon: 0.009998020008555413
steps: 526 lr: 1.0240000000000005e-06 evaluation reward: 11.57
episode: 2180 score: 12.0 memory length: 631088 epsilon: 0.009998020008555413
steps: 580 lr: 1.0240000000000005e-06 evaluation reward: 11.64
episode: 2181 score: 9.0 memory length: 631578 epsilon: 0.009998020008555413
steps: 490 lr: 1.0240000000000005e-06 evaluation reward: 11.63
episode: 2182 score: 10.0 memory length: 632114 epsilon: 0.009998020008555413
steps: 536 lr: 1.0240000000000005e-06 evaluation reward: 11.67
episode: 2183 score: 6.0 memory length: 632436 epsilon: 0.009998020008555413
steps: 322 lr: 1.0240000000000005e-06 evaluation reward: 11.57
episode: 2184 score: 8.0 memory length: 632875 epsilon: 0.009998020008555413
steps: 439 lr: 1.0240000000000005e-06 evaluation reward: 11.56
episode: 2185 score: 9.0 memory length: 633327 epsilon: 0.009998020008555413
steps: 452 lr: 1.0240000000000005e-06 evaluation reward: 11.55
episode: 2186 score: 10.0 memory length: 633827 epsilon: 0.009998020008555413
steps: 500 lr: 1.0240000000000005e-06 evaluation reward: 11.53
episode: 2187 score: 11.0 memory length: 634364 epsilon: 0.009998020008555413
steps: 537 lr: 1.0240000000000005e-06 evaluation reward: 11.5
episode: 2188 score: 13.0 memory length: 635021 epsilon: 0.009998020008555413
steps: 657 lr: 1.0240000000000005e-06 evaluation reward: 11.54
episode: 2189 score: 8.0 memory length: 635405 epsilon: 0.009998020008555413
steps: 384 lr: 1.0240000000000005e-06 evaluation reward: 11.52
episode: 2190 score: 7.0 memory length: 635793 epsilon: 0.009998020008555413
steps: 388 lr: 1.0240000000000005e-06 evaluation reward: 11.51
episode: 2191 score: 18.0 memory length: 636360 epsilon: 0.009998020008555413
steps: 567 lr: 1.0240000000000005e-06 evaluation reward: 11.61
episode: 2192 score: 13.0 memory length: 636953 epsilon: 0.009998020008555413
steps: 593 lr: 1.0240000000000005e-06 evaluation reward: 11.64
episode: 2193 score: 19.0 memory length: 637597 epsilon: 0.009998020008555413
steps: 644 lr: 1.0240000000000005e-06 evaluation reward: 11.76

```
episode: 2194 score: 12.0 memory length: 638188 epsilon: 0.009998020008555413
steps: 591 lr: 1.0240000000000005e-06 evaluation reward: 11.75
episode: 2195 score: 8.0 memory length: 638656 epsilon: 0.009998020008555413
steps: 468 lr: 1.0240000000000005e-06 evaluation reward: 11.72
episode: 2196 score: 17.0 memory length: 639311 epsilon: 0.009998020008555413
steps: 655 lr: 1.0240000000000005e-06 evaluation reward: 11.81
episode: 2197 score: 11.0 memory length: 639861 epsilon: 0.009998020008555413
steps: 550 lr: 1.0240000000000005e-06 evaluation reward: 11.84
episode: 2198 score: 14.0 memory length: 640512 epsilon: 0.009998020008555413
steps: 651 lr: 1.0240000000000005e-06 evaluation reward: 11.86
episode: 2199 score: 11.0 memory length: 641115 epsilon: 0.009998020008555413
steps: 603 lr: 1.0240000000000005e-06 evaluation reward: 11.81
episode: 2200 score: 8.0 memory length: 641538 epsilon: 0.009998020008555413
steps: 423 lr: 1.0240000000000005e-06 evaluation reward: 11.77
episode: 2201 score: 9.0 memory length: 642007 epsilon: 0.009998020008555413
steps: 469 lr: 1.0240000000000005e-06 evaluation reward: 11.7
episode: 2202 score: 7.0 memory length: 642389 epsilon: 0.009998020008555413
steps: 382 lr: 1.0240000000000005e-06 evaluation reward: 11.71
episode: 2203 score: 12.0 memory length: 643004 epsilon: 0.009998020008555413
steps: 615 lr: 1.0240000000000005e-06 evaluation reward: 11.64
episode: 2204 score: 11.0 memory length: 643570 epsilon: 0.009998020008555413
steps: 566 lr: 1.0240000000000005e-06 evaluation reward: 11.56
episode: 2205 score: 10.0 memory length: 644083 epsilon: 0.009998020008555413
steps: 513 lr: 1.0240000000000005e-06 evaluation reward: 11.55
episode: 2206 score: 15.0 memory length: 644684 epsilon: 0.009998020008555413
steps: 601 lr: 1.0240000000000005e-06 evaluation reward: 11.56
episode: 2207 score: 13.0 memory length: 645266 epsilon: 0.009998020008555413
steps: 582 lr: 1.0240000000000005e-06 evaluation reward: 11.53
episode: 2208 score: 13.0 memory length: 645843 epsilon: 0.009998020008555413
steps: 577 lr: 1.0240000000000005e-06 evaluation reward: 11.57
episode: 2209 score: 7.0 memory length: 646247 epsilon: 0.009998020008555413
steps: 404 lr: 1.0240000000000005e-06 evaluation reward: 11.58
episode: 2210 score: 13.0 memory length: 646721 epsilon: 0.009998020008555413
steps: 474 lr: 1.0240000000000005e-06 evaluation reward: 11.62
episode: 2211 score: 9.0 memory length: 647203 epsilon: 0.009998020008555413
steps: 482 lr: 1.0240000000000005e-06 evaluation reward: 11.54
episode: 2212 score: 9.0 memory length: 647631 epsilon: 0.009998020008555413
steps: 428 lr: 1.0240000000000005e-06 evaluation reward: 11.48
episode: 2213 score: 7.0 memory length: 648035 epsilon: 0.009998020008555413
steps: 404 lr: 1.0240000000000005e-06 evaluation reward: 11.38
episode: 2214 score: 18.0 memory length: 648635 epsilon: 0.009998020008555413
steps: 600 lr: 1.0240000000000005e-06 evaluation reward: 11.37
episode: 2215 score: 10.0 memory length: 649155 epsilon: 0.009998020008555413
steps: 520 lr: 1.0240000000000005e-06 evaluation reward: 11.32
episode: 2216 score: 13.0 memory length: 649715 epsilon: 0.009998020008555413
steps: 560 lr: 1.0240000000000005e-06 evaluation reward: 11.28
episode: 2217 score: 12.0 memory length: 650293 epsilon: 0.009998020008555413
steps: 578 lr: 1.0240000000000005e-06 evaluation reward: 11.32
episode: 2218 score: 9.0 memory length: 650778 epsilon: 0.009998020008555413
steps: 485 lr: 1.0240000000000005e-06 evaluation reward: 11.27
episode: 2219 score: 10.0 memory length: 651277 epsilon: 0.009998020008555413
steps: 499 lr: 1.0240000000000005e-06 evaluation reward: 11.25
episode: 2220 score: 11.0 memory length: 651830 epsilon: 0.009998020008555413
steps: 553 lr: 1.0240000000000005e-06 evaluation reward: 11.21
episode: 2221 score: 7.0 memory length: 652203 epsilon: 0.009998020008555413
steps: 373 lr: 1.0240000000000005e-06 evaluation reward: 11.19
episode: 2222 score: 8.0 memory length: 652642 epsilon: 0.009998020008555413
steps: 439 lr: 1.0240000000000005e-06 evaluation reward: 11.18
episode: 2223 score: 15.0 memory length: 653216 epsilon: 0.009998020008555413
steps: 574 lr: 1.0240000000000005e-06 evaluation reward: 11.25
```

```
episode: 2224 score: 11.0 memory length: 653750 epsilon: 0.009998020008555413
steps: 534 lr: 1.0240000000000005e-06 evaluation reward: 11.21
episode: 2225 score: 10.0 memory length: 654276 epsilon: 0.009998020008555413
steps: 526 lr: 1.0240000000000005e-06 evaluation reward: 11.2
episode: 2226 score: 14.0 memory length: 654793 epsilon: 0.009998020008555413
steps: 517 lr: 1.0240000000000005e-06 evaluation reward: 11.24
episode: 2227 score: 8.0 memory length: 655215 epsilon: 0.009998020008555413
steps: 422 lr: 1.0240000000000005e-06 evaluation reward: 11.25
episode: 2228 score: 9.0 memory length: 655737 epsilon: 0.009998020008555413
steps: 522 lr: 1.0240000000000005e-06 evaluation reward: 11.19
episode: 2229 score: 14.0 memory length: 656263 epsilon: 0.009998020008555413
steps: 526 lr: 1.0240000000000005e-06 evaluation reward: 11.26
episode: 2230 score: 15.0 memory length: 656853 epsilon: 0.009998020008555413
steps: 590 lr: 1.0240000000000005e-06 evaluation reward: 11.3
episode: 2231 score: 13.0 memory length: 657440 epsilon: 0.009998020008555413
steps: 587 lr: 1.0240000000000005e-06 evaluation reward: 11.31
episode: 2232 score: 12.0 memory length: 658093 epsilon: 0.009998020008555413
steps: 653 lr: 1.0240000000000005e-06 evaluation reward: 11.33
episode: 2233 score: 9.0 memory length: 658561 epsilon: 0.009998020008555413
steps: 468 lr: 1.0240000000000005e-06 evaluation reward: 11.34
episode: 2234 score: 9.0 memory length: 659034 epsilon: 0.009998020008555413
steps: 473 lr: 1.0240000000000005e-06 evaluation reward: 11.3
episode: 2235 score: 13.0 memory length: 659529 epsilon: 0.009998020008555413
steps: 495 lr: 1.0240000000000005e-06 evaluation reward: 11.33
episode: 2236 score: 12.0 memory length: 659986 epsilon: 0.009998020008555413
steps: 457 lr: 1.0240000000000005e-06 evaluation reward: 11.37
episode: 2237 score: 7.0 memory length: 660355 epsilon: 0.009998020008555413
steps: 369 lr: 1.0240000000000005e-06 evaluation reward: 11.36
episode: 2238 score: 11.0 memory length: 660874 epsilon: 0.009998020008555413
steps: 519 lr: 1.0240000000000005e-06 evaluation reward: 11.25
episode: 2239 score: 8.0 memory length: 661329 epsilon: 0.009998020008555413
steps: 455 lr: 1.0240000000000005e-06 evaluation reward: 11.18
episode: 2240 score: 8.0 memory length: 661713 epsilon: 0.009998020008555413
steps: 384 lr: 1.0240000000000005e-06 evaluation reward: 11.18
episode: 2241 score: 14.0 memory length: 662284 epsilon: 0.009998020008555413
steps: 571 lr: 1.0240000000000005e-06 evaluation reward: 11.2
episode: 2242 score: 10.0 memory length: 662856 epsilon: 0.009998020008555413
steps: 572 lr: 1.0240000000000005e-06 evaluation reward: 11.11
episode: 2243 score: 17.0 memory length: 663538 epsilon: 0.009998020008555413
steps: 682 lr: 1.0240000000000005e-06 evaluation reward: 11.12
episode: 2244 score: 16.0 memory length: 664015 epsilon: 0.009998020008555413
steps: 477 lr: 1.0240000000000005e-06 evaluation reward: 11.08
episode: 2245 score: 9.0 memory length: 664518 epsilon: 0.009998020008555413
steps: 503 lr: 1.0240000000000005e-06 evaluation reward: 11.0
episode: 2246 score: 13.0 memory length: 665182 epsilon: 0.009998020008555413
steps: 664 lr: 1.0240000000000005e-06 evaluation reward: 10.98
episode: 2247 score: 12.0 memory length: 665626 epsilon: 0.009998020008555413
steps: 444 lr: 1.0240000000000005e-06 evaluation reward: 10.97
episode: 2248 score: 11.0 memory length: 666158 epsilon: 0.009998020008555413
steps: 532 lr: 1.0240000000000005e-06 evaluation reward: 11.0
episode: 2249 score: 7.0 memory length: 666545 epsilon: 0.009998020008555413
steps: 387 lr: 1.0240000000000005e-06 evaluation reward: 10.96
episode: 2250 score: 11.0 memory length: 667144 epsilon: 0.009998020008555413
steps: 599 lr: 1.0240000000000005e-06 evaluation reward: 10.94
episode: 2251 score: 15.0 memory length: 667696 epsilon: 0.009998020008555413
steps: 552 lr: 1.0240000000000005e-06 evaluation reward: 10.97
episode: 2252 score: 13.0 memory length: 668279 epsilon: 0.009998020008555413
steps: 583 lr: 1.0240000000000005e-06 evaluation reward: 11.0
episode: 2253 score: 10.0 memory length: 668785 epsilon: 0.009998020008555413
steps: 506 lr: 1.0240000000000005e-06 evaluation reward: 10.98
```

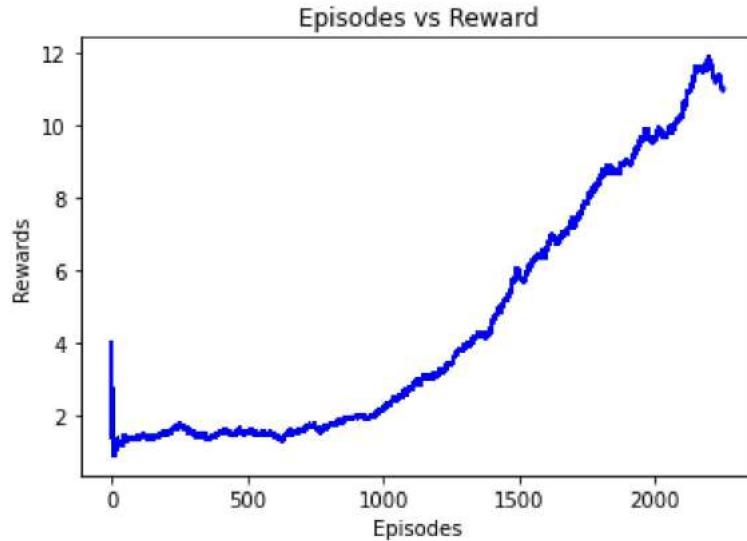
```

-----
KeyboardInterrupt                                     Traceback (most recent call last)
/tmp/ipykernel_11648/1504460345.py in <module>
      36         # Start training after random sample generation
      37         if(frame >= train_frame):
---> 38             agent.train_policy_net(frame)
      39             # Update the target network only for Double DQN only
      40             if double_dqn and (frame % update_target_network_frequency)== 0:

~/assignment5_materials/assignment5_materials/agent_double.py in train_policy_net(self, frame)
      70                 self.epsilon -= self.epsilon_decay
      71
---> 72                 mini_batch = self.memory.sample_mini_batch(frame)
      73                 mini_batch = np.array(mini_batch).transpose()
      74

~/assignment5_materials/assignment5_materials/memory.py in sample_mini_batch(self, frame)
      26             sample = []
      27             for j in range(HISTORY_SIZE + 1):
---> 28                 sample.append(self.memory[i + j])
      29             sample = np.array(sample)
      30             mini_batch.append((np.stack(sample[:, 0], axis=0), sample[3, 1],
sample[3, 2], sample[3, 3]))
```

KeyboardInterrupt:



Creating a DQN LSTM Agent

Now we will create a DQN agent that uses LSTM rather than past frames as history. We augment the experience replay to contain previous few (state, action, reward, next state) tuples rather than just one (state, action, reward, next state) tuple so it can work with LSTMs. Use the previous tuples to generate the current hidden and context vector for LSTM. Essentially, when you get a sample from replay buffer during training, start with the first tuple and generate hidden and context vector from this and pass it onto the next tuple. Do so consequitively till