

# 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 [1]:

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

Requirement already satisfied: gym in /opt/conda/lib/python3.7/site-packages (0.23.1)
Requirement already satisfied: pyvirtualdisplay in /opt/conda/lib/python3.7/site-packages (3.0)
Requirement already satisfied: numpy>=1.18.0 in /opt/conda/lib/python3.7/site-packages (from gym) (1.19.5)
Requirement already satisfied:云loudpickle>=1.2.0 in /opt/conda/lib/python3.7/site-packages (from gym) (2.0.0)
Requirement already satisfied: gym-notices>=0.0.4 in /opt/conda/lib/python3.7/site-packages (from gym) (0.0.6)
Requirement already satisfied: importlib-metadata>=4.10.0 in /opt/conda/lib/python3.7/site-packages (from gym) (4.11.3)
Requirement already satisfied: typing-extensions>=3.6.4 in /opt/conda/lib/python3.7/site-packages (from importlib-metadata>=4.10.0->gym) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /opt/conda/lib/python3.7/site-packages (from importlib-metadata>=4.10.0->gym) (3.7.0)
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: 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: gym-notices>=0.0.4 in /opt/conda/lib/python3.7/site-packages (from gym[atari]) (0.0.6)
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: 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: 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: 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: tqdm in /opt/conda/lib/python3.7/site-packages (from autorom[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: 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: 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: 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: 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: 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: 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: 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)
```

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
```

In [1]:

```
!pip3 uninstall gym atari-py ale-py --yes
!pip3 install gym[atari]
```

```

Found existing installation: gym 0.23.1
Uninstalling gym-0.23.1:
  Successfully uninstalled gym-0.23.1
Found existing installation: atari-py 0.2.9
Uninstalling atari-py-0.2.9:
  Successfully uninstalled atari-py-0.2.9
Found existing installation: ale-py 0.7.5
Uninstalling ale-py-0.7.5:
  Successfully uninstalled ale-py-0.7.5
Collecting gym[atari]
  Using cached gym-0.23.1-py3-none-any.whl
Requirement already satisfied: cloudpickle>=1.2.0 in /opt/conda/lib/python3.7/site-packages (from gym[atari]) (2.0.0)
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: gym-notices>=0.0.4 in /opt/conda/lib/python3.7/site-packages (from gym[atari]) (0.0.6)
Collecting ale-py~=0.7.4
  Using cached ale_py-0.7.5-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.6 MB)
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: 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: typing-extensions>=3.6.4 in /opt/conda/lib/python3.7/site-packages (from importlib-metadata>=4.10.0->gym[atari]) (4.1.1)
Installing collected packages: gym, ale-py
Successfully installed ale-py-0.7.5 gym-0.23.1

```

## Understanding the environment

In the following cell, we initialize our game of **Breakout** and you can see how the environment looks like. For further documentation of the 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 = False # set to True if using double DQN agent

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

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

## Main Training Loop

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

```
In [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

# 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)
```

/home/zacharytgeorge/assignment5/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)

```

episode: 0 score: 4.0 memory length: 258 epsilon: 0.9998831800000025 steps:
258 lr: 0.0001 evaluation reward: 4.0
episode: 1 score: 1.0 memory length: 410 epsilon: 0.9995822200000091 steps:
152 lr: 0.0001 evaluation reward: 2.5
episode: 2 score: 2.0 memory length: 610 epsilon: 0.9991862200000177 steps:
200 lr: 0.0001 evaluation reward: 2.333333333333335
episode: 3 score: 0.0 memory length: 734 epsilon: 0.998940700000023 steps: 1
24 lr: 0.0001 evaluation reward: 1.75
episode: 4 score: 2.0 memory length: 950 epsilon: 0.9985130200000323 steps:
216 lr: 0.0001 evaluation reward: 1.8
episode: 5 score: 0.0 memory length: 1074 epsilon: 0.9982675000000376 steps:
124 lr: 0.0001 evaluation reward: 1.5
episode: 6 score: 2.0 memory length: 1272 epsilon: 0.9978754600000461 steps:
198 lr: 0.0001 evaluation reward: 1.5714285714285714
episode: 7 score: 0.0 memory length: 1395 epsilon: 0.9976319200000514 steps:
123 lr: 0.0001 evaluation reward: 1.375
episode: 8 score: 1.0 memory length: 1547 epsilon: 0.9973309600000579 steps:
152 lr: 0.0001 evaluation reward: 1.333333333333333
episode: 9 score: 2.0 memory length: 1747 epsilon: 0.9969349600000665 steps:
200 lr: 0.0001 evaluation reward: 1.4
episode: 10 score: 1.0 memory length: 1898 epsilon: 0.996635980000073 steps:
151 lr: 0.0001 evaluation reward: 1.3636363636363635
episode: 11 score: 1.0 memory length: 2070 epsilon: 0.9962954200000804 step
s: 172 lr: 0.0001 evaluation reward: 1.333333333333333
episode: 12 score: 2.0 memory length: 2292 epsilon: 0.99585586000009 steps:
222 lr: 0.0001 evaluation reward: 1.3846153846153846
episode: 13 score: 1.0 memory length: 2461 epsilon: 0.9955212400000972 step
s: 169 lr: 0.0001 evaluation reward: 1.3571428571428572
episode: 14 score: 1.0 memory length: 2631 epsilon: 0.9951846400001045 step
s: 170 lr: 0.0001 evaluation reward: 1.333333333333333
episode: 15 score: 2.0 memory length: 2811 epsilon: 0.9948282400001123 step
s: 180 lr: 0.0001 evaluation reward: 1.375
episode: 16 score: 4.0 memory length: 3087 epsilon: 0.9942817600001241 step
s: 276 lr: 0.0001 evaluation reward: 1.5294117647058822
episode: 17 score: 1.0 memory length: 3257 epsilon: 0.9939451600001314 step
s: 170 lr: 0.0001 evaluation reward: 1.5
episode: 18 score: 2.0 memory length: 3456 epsilon: 0.99355114000014 steps:
199 lr: 0.0001 evaluation reward: 1.5263157894736843
episode: 19 score: 2.0 memory length: 3654 epsilon: 0.993159100001485 step
s: 198 lr: 0.0001 evaluation reward: 1.55
episode: 20 score: 2.0 memory length: 3852 epsilon: 0.992767060000157 steps:
198 lr: 0.0001 evaluation reward: 1.5714285714285714
episode: 21 score: 3.0 memory length: 4123 epsilon: 0.9922304800001687 step
s: 271 lr: 0.0001 evaluation reward: 1.6363636363636365
episode: 22 score: 1.0 memory length: 4292 epsilon: 0.9918958600001759 step
s: 169 lr: 0.0001 evaluation reward: 1.608695652173913
episode: 23 score: 2.0 memory length: 4511 epsilon: 0.9914622400001853 step
s: 219 lr: 0.0001 evaluation reward: 1.625
episode: 24 score: 2.0 memory length: 4730 epsilon: 0.9910286200001948 step
s: 219 lr: 0.0001 evaluation reward: 1.64
episode: 25 score: 2.0 memory length: 4932 epsilon: 0.9906286600002034 step
s: 202 lr: 0.0001 evaluation reward: 1.6538461538461537
episode: 26 score: 0.0 memory length: 5056 epsilon: 0.9903831400002088 step
s: 124 lr: 0.0001 evaluation reward: 1.5925925925925926
episode: 27 score: 2.0 memory length: 5254 epsilon: 0.9899911000002173 step
s: 198 lr: 0.0001 evaluation reward: 1.6071428571428572
episode: 28 score: 0.0 memory length: 5378 epsilon: 0.9897455800002226 step
s: 124 lr: 0.0001 evaluation reward: 1.5517241379310345
episode: 29 score: 0.0 memory length: 5502 epsilon: 0.989500600002279 step
s: 124 lr: 0.0001 evaluation reward: 1.5

```

```
episode: 30 score: 0.0 memory length: 5625 epsilon: 0.9892565200002332 step
s: 123 lr: 0.0001 evaluation reward: 1.4516129032258065
episode: 31 score: 1.0 memory length: 5797 epsilon: 0.9889159600002406 step
s: 172 lr: 0.0001 evaluation reward: 1.4375
episode: 32 score: 0.0 memory length: 5921 epsilon: 0.988670440000246 steps:
124 lr: 0.0001 evaluation reward: 1.393939393939394
episode: 33 score: 0.0 memory length: 6045 epsilon: 0.9884249200002513 step
s: 124 lr: 0.0001 evaluation reward: 1.3529411764705883
episode: 34 score: 2.0 memory length: 6243 epsilon: 0.9880328800002598 step
s: 198 lr: 0.0001 evaluation reward: 1.3714285714285714
episode: 35 score: 4.0 memory length: 6522 epsilon: 0.9874804600002718 step
s: 279 lr: 0.0001 evaluation reward: 1.4444444444444444
episode: 36 score: 1.0 memory length: 6673 epsilon: 0.9871814800002783 step
s: 151 lr: 0.0001 evaluation reward: 1.4324324324324325
episode: 37 score: 3.0 memory length: 6924 epsilon: 0.9866845000002891 step
s: 251 lr: 0.0001 evaluation reward: 1.4736842105263157
episode: 38 score: 0.0 memory length: 7048 epsilon: 0.9864389800002944 step
s: 124 lr: 0.0001 evaluation reward: 1.435897435897436
episode: 39 score: 1.0 memory length: 7200 epsilon: 0.9861380200003009 step
s: 152 lr: 0.0001 evaluation reward: 1.425
episode: 40 score: 3.0 memory length: 7448 epsilon: 0.9856469800003116 step
s: 248 lr: 0.0001 evaluation reward: 1.4634146341463414
episode: 41 score: 3.0 memory length: 7675 epsilon: 0.9851975200003213 step
s: 227 lr: 0.0001 evaluation reward: 1.5
episode: 42 score: 2.0 memory length: 7874 epsilon: 0.9848035000003299 step
s: 199 lr: 0.0001 evaluation reward: 1.5116279069767442
episode: 43 score: 0.0 memory length: 7997 epsilon: 0.9845599600003352 step
s: 123 lr: 0.0001 evaluation reward: 1.4772727272727273
episode: 44 score: 1.0 memory length: 8149 epsilon: 0.9842590000003417 step
s: 152 lr: 0.0001 evaluation reward: 1.4666666666666666
episode: 45 score: 2.0 memory length: 8348 epsilon: 0.9838649800003503 step
s: 199 lr: 0.0001 evaluation reward: 1.4782608695652173
episode: 46 score: 0.0 memory length: 8471 epsilon: 0.9836214400003556 step
s: 123 lr: 0.0001 evaluation reward: 1.446808510638298
episode: 47 score: 0.0 memory length: 8595 epsilon: 0.9833759200003609 step
s: 124 lr: 0.0001 evaluation reward: 1.4166666666666667
episode: 48 score: 0.0 memory length: 8719 epsilon: 0.9831304000003662 step
s: 124 lr: 0.0001 evaluation reward: 1.3877551020408163
episode: 49 score: 1.0 memory length: 8871 epsilon: 0.9828294400003728 step
s: 152 lr: 0.0001 evaluation reward: 1.38
episode: 50 score: 1.0 memory length: 9023 epsilon: 0.9825284800003793 step
s: 152 lr: 0.0001 evaluation reward: 1.3725490196078431
episode: 51 score: 1.0 memory length: 9193 epsilon: 0.9821918800003866 step
s: 170 lr: 0.0001 evaluation reward: 1.3653846153846154
episode: 52 score: 2.0 memory length: 9394 epsilon: 0.9817939000003952 step
s: 201 lr: 0.0001 evaluation reward: 1.3773584905660377
episode: 53 score: 2.0 memory length: 9592 epsilon: 0.9814018600004037 step
s: 198 lr: 0.0001 evaluation reward: 1.3888888888888888
episode: 54 score: 2.0 memory length: 9811 epsilon: 0.9809682400004132 step
s: 219 lr: 0.0001 evaluation reward: 1.4
episode: 55 score: 0.0 memory length: 9935 epsilon: 0.9807227200004185 step
s: 124 lr: 0.0001 evaluation reward: 1.375
episode: 56 score: 4.0 memory length: 10212 epsilon: 0.9801742600004304 step
s: 277 lr: 0.0001 evaluation reward: 1.4210526315789473
episode: 57 score: 1.0 memory length: 10364 epsilon: 0.9798733000004369 step
s: 152 lr: 0.0001 evaluation reward: 1.4137931034482758
episode: 58 score: 1.0 memory length: 10534 epsilon: 0.9795367000004442 step
s: 170 lr: 0.0001 evaluation reward: 1.4067796610169492
episode: 59 score: 1.0 memory length: 10703 epsilon: 0.9792020800004515 step
s: 169 lr: 0.0001 evaluation reward: 1.4
```

episode: 60 score: 1.0 memory length: 10873 epsilon: 0.9788654800004588 step  
 s: 170 lr: 0.0001 evaluation reward: 1.3934426229508197  
 episode: 61 score: 0.0 memory length: 10997 epsilon: 0.9786199600004641 step  
 s: 124 lr: 0.0001 evaluation reward: 1.3709677419354838  
 episode: 62 score: 3.0 memory length: 11224 epsilon: 0.9781705000004739 step  
 s: 227 lr: 0.0001 evaluation reward: 1.3968253968253967  
 episode: 63 score: 1.0 memory length: 11375 epsilon: 0.9778715200004804 step  
 s: 151 lr: 0.0001 evaluation reward: 1.390625  
 episode: 64 score: 1.0 memory length: 11545 epsilon: 0.9775349200004877 step  
 s: 170 lr: 0.0001 evaluation reward: 1.3846153846153846  
 episode: 65 score: 2.0 memory length: 11747 epsilon: 0.9771349600004964 step  
 s: 202 lr: 0.0001 evaluation reward: 1.393939393939394  
 episode: 66 score: 1.0 memory length: 11899 epsilon: 0.9768340000005029 step  
 s: 152 lr: 0.0001 evaluation reward: 1.3880597014925373  
 episode: 67 score: 0.0 memory length: 12023 epsilon: 0.9765884800005082 step  
 s: 124 lr: 0.0001 evaluation reward: 1.3676470588235294  
 episode: 68 score: 2.0 memory length: 12221 epsilon: 0.9761964400005168 step  
 s: 198 lr: 0.0001 evaluation reward: 1.3768115942028984  
 episode: 69 score: 0.0 memory length: 12345 epsilon: 0.9759509200005221 step  
 s: 124 lr: 0.0001 evaluation reward: 1.3571428571428572  
 episode: 70 score: 3.0 memory length: 12613 epsilon: 0.9754202800005336 step  
 s: 268 lr: 0.0001 evaluation reward: 1.380281690140845  
 episode: 71 score: 2.0 memory length: 12830 epsilon: 0.9749906200005429 step  
 s: 217 lr: 0.0001 evaluation reward: 1.3888888888888888  
 episode: 72 score: 0.0 memory length: 12954 epsilon: 0.9747451000005483 step  
 s: 124 lr: 0.0001 evaluation reward: 1.36986301369863  
 episode: 73 score: 0.0 memory length: 13078 epsilon: 0.9744995800005536 step  
 s: 124 lr: 0.0001 evaluation reward: 1.3513513513513  
 episode: 74 score: 1.0 memory length: 13248 epsilon: 0.9741629800005609 step  
 s: 170 lr: 0.0001 evaluation reward: 1.3466666666666667  
 episode: 75 score: 3.0 memory length: 13496 epsilon: 0.9736719400005716 step  
 s: 248 lr: 0.0001 evaluation reward: 1.368421052631579  
 episode: 76 score: 1.0 memory length: 13648 epsilon: 0.9733709800005781 step  
 s: 152 lr: 0.0001 evaluation reward: 1.3636363636363635  
 episode: 77 score: 1.0 memory length: 13800 epsilon: 0.9730700200005846 step  
 s: 152 lr: 0.0001 evaluation reward: 1.358974358974359  
 episode: 78 score: 2.0 memory length: 13982 epsilon: 0.9727096600005924 step  
 s: 182 lr: 0.0001 evaluation reward: 1.3670886075949367  
 episode: 79 score: 3.0 memory length: 14209 epsilon: 0.9722602000006022 step  
 s: 227 lr: 0.0001 evaluation reward: 1.3875  
 episode: 80 score: 4.0 memory length: 14467 epsilon: 0.9717493600006133 step  
 s: 258 lr: 0.0001 evaluation reward: 1.4197530864197532  
 episode: 81 score: 1.0 memory length: 14639 epsilon: 0.9714088000006207 step  
 s: 172 lr: 0.0001 evaluation reward: 1.4146341463414633  
 episode: 82 score: 0.0 memory length: 14763 epsilon: 0.971163280000626 step  
 s: 124 lr: 0.0001 evaluation reward: 1.3975903614457832  
 episode: 83 score: 2.0 memory length: 14961 epsilon: 0.9707712400006345 step  
 s: 198 lr: 0.0001 evaluation reward: 1.4047619047619047  
 episode: 84 score: 1.0 memory length: 15131 epsilon: 0.9704346400006418 step  
 s: 170 lr: 0.0001 evaluation reward: 1.4  
 episode: 85 score: 4.0 memory length: 15422 epsilon: 0.9698584600006543 step  
 s: 291 lr: 0.0001 evaluation reward: 1.430232558139535  
 episode: 86 score: 0.0 memory length: 15546 epsilon: 0.9696129400006597 step  
 s: 124 lr: 0.0001 evaluation reward: 1.4137931034482758  
 episode: 87 score: 1.0 memory length: 15718 epsilon: 0.9692723800006671 step  
 s: 172 lr: 0.0001 evaluation reward: 1.4090909090909092  
 episode: 88 score: 0.0 memory length: 15841 epsilon: 0.9690288400006724 step  
 s: 123 lr: 0.0001 evaluation reward: 1.3932584269662922  
 episode: 89 score: 4.0 memory length: 16137 epsilon: 0.9684427600006851 step  
 s: 296 lr: 0.0001 evaluation reward: 1.4222222222222223

episode: 90 score: 0.0 memory length: 16261 epsilon: 0.9681972400006904 step  
 s: 124 lr: 0.0001 evaluation reward: 1.4065934065934067  
 episode: 91 score: 0.0 memory length: 16385 epsilon: 0.9679517200006957 step  
 s: 124 lr: 0.0001 evaluation reward: 1.391304347826087  
 episode: 92 score: 2.0 memory length: 16566 epsilon: 0.9675933400007035 step  
 s: 181 lr: 0.0001 evaluation reward: 1.3978494623655915  
 episode: 93 score: 1.0 memory length: 16735 epsilon: 0.9672587200007108 step  
 s: 169 lr: 0.0001 evaluation reward: 1.3936170212765957  
 episode: 94 score: 0.0 memory length: 16859 epsilon: 0.9670132000007161 step  
 s: 124 lr: 0.0001 evaluation reward: 1.3789473684210527  
 episode: 95 score: 4.0 memory length: 17118 epsilon: 0.9665003800007272 step  
 s: 259 lr: 0.0001 evaluation reward: 1.40625  
 episode: 96 score: 2.0 memory length: 17336 epsilon: 0.9660687400007366 step  
 s: 218 lr: 0.0001 evaluation reward: 1.4123711340206186  
 episode: 97 score: 1.0 memory length: 17506 epsilon: 0.9657321400007439 step  
 s: 170 lr: 0.0001 evaluation reward: 1.4081632653061225  
 episode: 98 score: 2.0 memory length: 17705 epsilon: 0.9653381200007525 step  
 s: 199 lr: 0.0001 evaluation reward: 1.4141414141414141  
 episode: 99 score: 1.0 memory length: 17857 epsilon: 0.965037160000759 step  
 s: 152 lr: 0.0001 evaluation reward: 1.41  
 episode: 100 score: 3.0 memory length: 18106 epsilon: 0.9645441400007697 step  
 ps: 249 lr: 0.0001 evaluation reward: 1.4  
 episode: 101 score: 2.0 memory length: 18328 epsilon: 0.9641045800007793 step  
 ps: 222 lr: 0.0001 evaluation reward: 1.41  
 episode: 102 score: 2.0 memory length: 18546 epsilon: 0.9636729400007886 step  
 ps: 218 lr: 0.0001 evaluation reward: 1.41  
 episode: 103 score: 1.0 memory length: 18717 epsilon: 0.963334360000796 step  
 s: 171 lr: 0.0001 evaluation reward: 1.42  
 episode: 104 score: 5.0 memory length: 19044 epsilon: 0.96268690000081 step  
 s: 327 lr: 0.0001 evaluation reward: 1.45  
 episode: 105 score: 0.0 memory length: 19168 epsilon: 0.9624413800008154 step  
 ps: 124 lr: 0.0001 evaluation reward: 1.45  
 episode: 106 score: 1.0 memory length: 19338 epsilon: 0.9621047800008227 step  
 ps: 170 lr: 0.0001 evaluation reward: 1.44  
 episode: 107 score: 1.0 memory length: 19507 epsilon: 0.9617701600008299 step  
 ps: 169 lr: 0.0001 evaluation reward: 1.45  
 episode: 108 score: 4.0 memory length: 19826 epsilon: 0.9611385400008436 step  
 ps: 319 lr: 0.0001 evaluation reward: 1.48  
 episode: 109 score: 2.0 memory length: 20045 epsilon: 0.9607049200008531 step  
 ps: 219 lr: 0.0001 evaluation reward: 1.48  
 episode: 110 score: 2.0 memory length: 20248 epsilon: 0.9603029800008618 step  
 ps: 203 lr: 0.0001 evaluation reward: 1.49  
 episode: 111 score: 3.0 memory length: 20475 epsilon: 0.9598535200008715 step  
 ps: 227 lr: 0.0001 evaluation reward: 1.51  
 episode: 112 score: 0.0 memory length: 20599 epsilon: 0.9596080000008769 step  
 ps: 124 lr: 0.0001 evaluation reward: 1.49  
 episode: 113 score: 2.0 memory length: 20798 epsilon: 0.9592139800008854 step  
 ps: 199 lr: 0.0001 evaluation reward: 1.5  
 episode: 114 score: 5.0 memory length: 21125 epsilon: 0.9585665200008995 step  
 ps: 327 lr: 0.0001 evaluation reward: 1.54  
 episode: 115 score: 0.0 memory length: 21249 epsilon: 0.9583210000009048 step  
 ps: 124 lr: 0.0001 evaluation reward: 1.52  
 episode: 116 score: 2.0 memory length: 21448 epsilon: 0.9579269800009134 step  
 ps: 199 lr: 0.0001 evaluation reward: 1.5  
 episode: 117 score: 1.0 memory length: 21618 epsilon: 0.9575903800009207 step  
 ps: 170 lr: 0.0001 evaluation reward: 1.5  
 episode: 118 score: 2.0 memory length: 21819 epsilon: 0.9571924000009293 step  
 ps: 201 lr: 0.0001 evaluation reward: 1.5  
 episode: 119 score: 0.0 memory length: 21943 epsilon: 0.9569468800009346 step  
 ps: 124 lr: 0.0001 evaluation reward: 1.48

episode: 120	score: 1.0	memory length: 22113	epsilon: 0.956610280000942	step
s: 170	lr: 0.0001	evaluation reward: 1.47	epsilon: 0.9563073400009485	ste
episode: 121	score: 1.0	memory length: 22266	epsilon: 0.9559113400009571	ste
ps: 153	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9556658200009625	ste
episode: 122	score: 2.0	memory length: 22466	epsilon: 0.9554203000009678	ste
ps: 200	lr: 0.0001	evaluation reward: 1.46	epsilon: 0.9549886600009772	ste
episode: 123	score: 0.0	memory length: 22590	epsilon: 0.9546877000009837	ste
ps: 124	lr: 0.0001	evaluation reward: 1.44	epsilon: 0.9542877400009924	ste
episode: 124	score: 0.0	memory length: 22714	epsilon: 0.9538957000010009	ste
ps: 124	lr: 0.0001	evaluation reward: 1.42	epsilon: 0.9534046600010115	ste
episode: 125	score: 2.0	memory length: 22932	epsilon: 0.95301262000102	step
ps: 218	lr: 0.0001	evaluation reward: 1.42	epsilon: 0.9527690800010253	ste
episode: 126	score: 1.0	memory length: 23084	epsilon: 0.9524681200010319	ste
ps: 152	lr: 0.0001	evaluation reward: 1.43	epsilon: 0.9519315400010435	ste
episode: 127	score: 2.0	memory length: 23286	epsilon: 0.9515177200010525	ste
ps: 202	lr: 0.0001	evaluation reward: 1.43	epsilon: 0.9507692800010688	ste
episode: 128	score: 2.0	memory length: 23484	epsilon: 0.9503693200010774	ste
ps: 198	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9497832400010902	ste
episode: 129	score: 0.0	memory length: 23608	epsilon: 0.9493199200011002	ste
ps: 124	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9489199600011089	ste
episode: 130	score: 0.0	memory length: 23732	epsilon: 0.9485279200011174	ste
ps: 124	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9482843800011227	ste
episode: 131	score: 2.0	memory length: 23930	epsilon: 0.948038860001128	step
s: 198	lr: 0.0001	evaluation reward: 1.46	epsilon: 0.9477933400011334	ste
episode: 132	score: 0.0	memory length: 24053	epsilon: 0.9474943600011398	ste
ps: 123	lr: 0.0001	evaluation reward: 1.46	epsilon: 0.9470627200011492	ste
episode: 133	score: 1.0	memory length: 24205	epsilon: 0.946668700011578	ste
ps: 152	lr: 0.0001	evaluation reward: 1.47	epsilon: 0.9464231800011631	ste
episode: 134	score: 3.0	memory length: 24476	epsilon: 0.9460291600011717	ste
ps: 271	lr: 0.0001	evaluation reward: 1.48		
episode: 135	score: 3.0	memory length: 24685		
ps: 209	lr: 0.0001	evaluation reward: 1.47		
episode: 136	score: 6.0	memory length: 25063		
ps: 378	lr: 0.0001	evaluation reward: 1.52		
episode: 137	score: 2.0	memory length: 25265		
ps: 202	lr: 0.0001	evaluation reward: 1.51		
episode: 138	score: 4.0	memory length: 25561		
ps: 296	lr: 0.0001	evaluation reward: 1.55		
episode: 139	score: 3.0	memory length: 25795		
ps: 234	lr: 0.0001	evaluation reward: 1.57		
episode: 140	score: 2.0	memory length: 25997		
ps: 202	lr: 0.0001	evaluation reward: 1.56		
episode: 141	score: 2.0	memory length: 26195		
ps: 198	lr: 0.0001	evaluation reward: 1.55		
episode: 142	score: 0.0	memory length: 26318		
ps: 123	lr: 0.0001	evaluation reward: 1.53		
episode: 143	score: 0.0	memory length: 26442		
s: 124	lr: 0.0001	evaluation reward: 1.53		
episode: 144	score: 0.0	memory length: 26566		
ps: 124	lr: 0.0001	evaluation reward: 1.52		
episode: 145	score: 1.0	memory length: 26717		
ps: 151	lr: 0.0001	evaluation reward: 1.51		
episode: 146	score: 2.0	memory length: 26935		
ps: 218	lr: 0.0001	evaluation reward: 1.53		
episode: 147	score: 2.0	memory length: 27134		
ps: 199	lr: 0.0001	evaluation reward: 1.55		
episode: 148	score: 0.0	memory length: 27258		
ps: 124	lr: 0.0001	evaluation reward: 1.55		
episode: 149	score: 2.0	memory length: 27457		
ps: 199	lr: 0.0001	evaluation reward: 1.56		

episode: 150	score: 3.0	memory length: 27706	epsilon: 0.9455361400011824	ste
ps: 249	lr: 0.0001	evaluation reward: 1.58		
episode: 151	score: 1.0	memory length: 27877	epsilon: 0.9451975600011897	ste
ps: 171	lr: 0.0001	evaluation reward: 1.58		
episode: 152	score: 2.0	memory length: 28076	epsilon: 0.9448035400011983	ste
ps: 199	lr: 0.0001	evaluation reward: 1.58		
episode: 153	score: 0.0	memory length: 28200	epsilon: 0.9445580200012036	ste
ps: 124	lr: 0.0001	evaluation reward: 1.56		
episode: 154	score: 1.0	memory length: 28353	epsilon: 0.9442550800012102	ste
ps: 153	lr: 0.0001	evaluation reward: 1.55		
episode: 155	score: 2.0	memory length: 28554	epsilon: 0.9438571000012188	ste
ps: 201	lr: 0.0001	evaluation reward: 1.57		
episode: 156	score: 4.0	memory length: 28831	epsilon: 0.9433086400012307	ste
ps: 277	lr: 0.0001	evaluation reward: 1.57		
episode: 157	score: 2.0	memory length: 29048	epsilon: 0.94287898000124	step
s: 217	lr: 0.0001	evaluation reward: 1.58		
episode: 158	score: 3.0	memory length: 29277	epsilon: 0.9424255600012499	ste
ps: 229	lr: 0.0001	evaluation reward: 1.6		
episode: 159	score: 0.0	memory length: 29401	epsilon: 0.9421800400012552	ste
ps: 124	lr: 0.0001	evaluation reward: 1.59		
episode: 160	score: 2.0	memory length: 29623	epsilon: 0.9417404800012648	ste
ps: 222	lr: 0.0001	evaluation reward: 1.6		
episode: 161	score: 5.0	memory length: 29925	epsilon: 0.9411425200012777	ste
ps: 302	lr: 0.0001	evaluation reward: 1.65		
episode: 162	score: 1.0	memory length: 30095	epsilon: 0.940805920001285	step
s: 170	lr: 0.0001	evaluation reward: 1.63		
episode: 163	score: 2.0	memory length: 30293	epsilon: 0.9404138800012936	ste
ps: 198	lr: 0.0001	evaluation reward: 1.64		
episode: 164	score: 1.0	memory length: 30463	epsilon: 0.9400772800013009	ste
ps: 170	lr: 0.0001	evaluation reward: 1.64		
episode: 165	score: 2.0	memory length: 30664	epsilon: 0.9396793000013095	ste
ps: 201	lr: 0.0001	evaluation reward: 1.64		
episode: 166	score: 2.0	memory length: 30881	epsilon: 0.9392496400013188	ste
ps: 217	lr: 0.0001	evaluation reward: 1.65		
episode: 167	score: 5.0	memory length: 31208	epsilon: 0.9386021800013329	ste
ps: 327	lr: 0.0001	evaluation reward: 1.7		
episode: 168	score: 3.0	memory length: 31456	epsilon: 0.9381111400013435	ste
ps: 248	lr: 0.0001	evaluation reward: 1.71		
episode: 169	score: 2.0	memory length: 31638	epsilon: 0.9377507800013514	ste
ps: 182	lr: 0.0001	evaluation reward: 1.73		
episode: 170	score: 1.0	memory length: 31808	epsilon: 0.9374141800013587	ste
ps: 170	lr: 0.0001	evaluation reward: 1.71		
episode: 171	score: 1.0	memory length: 31978	epsilon: 0.937077580001366	step
s: 170	lr: 0.0001	evaluation reward: 1.7		
episode: 172	score: 0.0	memory length: 32102	epsilon: 0.9368320600013713	ste
ps: 124	lr: 0.0001	evaluation reward: 1.7		
episode: 173	score: 1.0	memory length: 32254	epsilon: 0.9365311000013778	ste
ps: 152	lr: 0.0001	evaluation reward: 1.71		
episode: 174	score: 1.0	memory length: 32424	epsilon: 0.9361945000013852	ste
ps: 170	lr: 0.0001	evaluation reward: 1.71		
episode: 175	score: 0.0	memory length: 32548	epsilon: 0.9359489800013905	ste
ps: 124	lr: 0.0001	evaluation reward: 1.68		
episode: 176	score: 1.0	memory length: 32718	epsilon: 0.9356123800013978	ste
ps: 170	lr: 0.0001	evaluation reward: 1.68		
episode: 177	score: 0.0	memory length: 32842	epsilon: 0.9353668600014031	ste
ps: 124	lr: 0.0001	evaluation reward: 1.67		
episode: 178	score: 4.0	memory length: 33139	epsilon: 0.9347788000014159	ste
ps: 297	lr: 0.0001	evaluation reward: 1.69		
episode: 179	score: 0.0	memory length: 33263	epsilon: 0.9345332800014212	ste
ps: 124	lr: 0.0001	evaluation reward: 1.66		

episode: 180	score: 3.0	memory length: 33510	epsilon: 0.9340442200014318	ste
ps: 247	lr: 0.0001	evaluation reward: 1.65	epsilon: 0.9336521800014403	ste
episode: 181	score: 2.0	memory length: 33708	epsilon: 0.9334066600014457	ste
ps: 198	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.9329730400014551	ste
episode: 182	score: 0.0	memory length: 33832	epsilon: 0.9325790200014636	ste
ps: 124	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.932333500001469	step
episode: 183	score: 2.0	memory length: 34051	epsilon: 0.9319414600014775	ste
ps: 219	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.9316959400014828	ste
episode: 184	score: 2.0	memory length: 34250	epsilon: 0.9314524000014881	ste
ps: 199	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.9311158000014954	ste
episode: 185	score: 0.0	memory length: 34374	epsilon: 0.930583180001507	step
s: 124	lr: 0.0001	evaluation reward: 1.63	epsilon: 0.9301891600015155	ste
episode: 186	score: 2.0	memory length: 34572	epsilon: 0.9297575200015249	ste
ps: 198	lr: 0.0001	evaluation reward: 1.65	epsilon: 0.9293635000015334	ste
episode: 187	score: 0.0	memory length: 34696	epsilon: 0.9291179800015388	ste
ps: 124	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.9287774200015462	ste
episode: 188	score: 0.0	memory length: 34819	epsilon: 0.9285338800015515	ste
ps: 123	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.9281972800015588	ste
episode: 189	score: 1.0	memory length: 34989	epsilon: 0.9279517600015641	ste
ps: 170	lr: 0.0001	evaluation reward: 1.61	epsilon: 0.9276508000015706	ste
episode: 190	score: 3.0	memory length: 35258	epsilon: 0.9273142000015779	ste
s: 269	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.9269221600015864	ste
episode: 191	score: 2.0	memory length: 35457	epsilon: 0.926619220001593	step
ps: 199	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.925791580001611	ste
episode: 192	score: 2.0	memory length: 35675	epsilon: 0.9253599400016204	ste
ps: 218	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.9249679000016289	ste
episode: 193	score: 2.0	memory length: 35874	epsilon: 0.9246253600016363	ste
ps: 199	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.9242333200016448	ste
episode: 194	score: 0.0	memory length: 35998	epsilon: 0.9233335000016542	ste
ps: 124	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.9230001600016642	ste
episode: 195	score: 1.0	memory length: 36170	epsilon: 0.92266192200016732	ste
ps: 172	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.9223335000016822	ste
episode: 196	score: 0.0	memory length: 36293	epsilon: 0.9220001600016912	ste
ps: 123	lr: 0.0001	evaluation reward: 1.62	epsilon: 0.921666192200016998	ste
episode: 197	score: 1.0	memory length: 36463	epsilon: 0.9213335000017088	ste
ps: 170	lr: 0.0001	evaluation reward: 1.62	epsilon: 0.9210001600017178	ste
episode: 198	score: 0.0	memory length: 36587	epsilon: 0.920666192200017268	ste
ps: 124	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.9203335000017358	ste
episode: 199	score: 1.0	memory length: 36739	epsilon: 0.9200001600017448	ste
ps: 152	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.919666192200017538	ste
episode: 200	score: 1.0	memory length: 36909	epsilon: 0.9193335000017628	ste
ps: 170	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.9190001600017718	ste
episode: 201	score: 2.0	memory length: 37107	epsilon: 0.918666192200017808	ste
ps: 198	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.9183335000017898	ste
episode: 202	score: 1.0	memory length: 37260	epsilon: 0.9180001600017988	ste
s: 153	lr: 0.0001	evaluation reward: 1.57	epsilon: 0.917666192200018078	ste
episode: 203	score: 1.0	memory length: 37431	epsilon: 0.9173335000018168	ste
ps: 171	lr: 0.0001	evaluation reward: 1.57	epsilon: 0.9170001600018258	ste
episode: 204	score: 0.0	memory length: 37554	epsilon: 0.916666192200018348	ste
ps: 123	lr: 0.0001	evaluation reward: 1.52	epsilon: 0.9163335000018438	ste
episode: 205	score: 0.0	memory length: 37678	epsilon: 0.9160001600018528	ste
s: 124	lr: 0.0001	evaluation reward: 1.52	epsilon: 0.915666192200018618	ste
episode: 206	score: 2.0	memory length: 37896	epsilon: 0.9153335000018708	ste
ps: 218	lr: 0.0001	evaluation reward: 1.53	epsilon: 0.9150001600018798	ste
episode: 207	score: 2.0	memory length: 38094	epsilon: 0.914666192200018888	ste
ps: 198	lr: 0.0001	evaluation reward: 1.54	epsilon: 0.9143335000018978	ste
episode: 208	score: 1.0	memory length: 38267	epsilon: 0.9140001600019068	ste
ps: 173	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.913666192200019158	ste
episode: 209	score: 2.0	memory length: 38465	epsilon: 0.9133335000019248	ste
ps: 198	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.9130001600019338	ste

episode: 210	score: 0.0	memory length: 38589	epsilon: 0.923987800016502	ste
ps: 124	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.9234987400016608	ste
episode: 211	score: 3.0	memory length: 38836	epsilon: 0.9231621400016681	ste
ps: 247	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.9227641600016767	ste
episode: 212	score: 1.0	memory length: 39006	epsilon: 0.922518640001682	step
ps: 170	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9219266200016949	ste
episode: 213	score: 2.0	memory length: 39207	epsilon: 0.9214930000017043	ste
ps: 201	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9210593800017137	ste
episode: 214	score: 0.0	memory length: 39331	epsilon: 0.920722780001721	step
s: 124	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.9203861800017283	ste
episode: 215	score: 4.0	memory length: 39630	epsilon: 0.9200495800017356	ste
ps: 299	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.9192239200017536	ste
episode: 216	score: 2.0	memory length: 39849	epsilon: 0.9187744600017633	ste
ps: 219	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.9184735000017699	ste
episode: 217	score: 2.0	memory length: 40068	epsilon: 0.9182279800017752	ste
ps: 219	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9177884200017847	ste
episode: 218	score: 1.0	memory length: 40238	epsilon: 0.9173904400017934	ste
s: 170	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.9169548400018028	ste
episode: 219	score: 1.0	memory length: 40408	epsilon: 0.9166142800018102	ste
ps: 170	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9162796600018175	ste
episode: 220	score: 1.0	memory length: 40578	epsilon: 0.9156322000018315	ste
ps: 170	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9153886600018368	ste
episode: 221	score: 1.0	memory length: 40748	epsilon: 0.9151451200018421	ste
s: 170	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.9148996000018474	ste
episode: 222	score: 3.0	memory length: 40995	epsilon: 0.9145649800018547	ste
ps: 247	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.91431946000186	step
episode: 223	score: 3.0	memory length: 41222	epsilon: 0.9140759200018653	ste
ps: 227	lr: 0.0001	evaluation reward: 1.54	epsilon: 0.9138304000018707	ste
episode: 224	score: 1.0	memory length: 41374	epsilon: 0.9135868600018759	ste
ps: 152	lr: 0.0001	evaluation reward: 1.55		
episode: 225	score: 0.0	memory length: 41498		
ps: 124	lr: 0.0001	evaluation reward: 1.53		
episode: 226	score: 2.0	memory length: 41720		
ps: 222	lr: 0.0001	evaluation reward: 1.54		
episode: 227	score: 2.0	memory length: 41921		
ps: 201	lr: 0.0001	evaluation reward: 1.54		
episode: 228	score: 2.0	memory length: 42141		
ps: 220	lr: 0.0001	evaluation reward: 1.54		
episode: 229	score: 1.0	memory length: 42313		
ps: 172	lr: 0.0001	evaluation reward: 1.55		
episode: 230	score: 1.0	memory length: 42482		
ps: 169	lr: 0.0001	evaluation reward: 1.56		
episode: 231	score: 5.0	memory length: 42809		
ps: 327	lr: 0.0001	evaluation reward: 1.59		
episode: 232	score: 0.0	memory length: 42932		
ps: 123	lr: 0.0001	evaluation reward: 1.59		
episode: 233	score: 0.0	memory length: 43055		
ps: 123	lr: 0.0001	evaluation reward: 1.58		
episode: 234	score: 0.0	memory length: 43179		
ps: 124	lr: 0.0001	evaluation reward: 1.55		
episode: 235	score: 1.0	memory length: 43348		
ps: 169	lr: 0.0001	evaluation reward: 1.53		
episode: 236	score: 0.0	memory length: 43472		
s: 124	lr: 0.0001	evaluation reward: 1.47		
episode: 237	score: 0.0	memory length: 43595		
ps: 123	lr: 0.0001	evaluation reward: 1.45		
episode: 238	score: 0.0	memory length: 43719		
ps: 124	lr: 0.0001	evaluation reward: 1.41		
episode: 239	score: 0.0	memory length: 43842		
ps: 123	lr: 0.0001	evaluation reward: 1.38		

episode: 240	score: 0.0	memory length: 43966	epsilon: 0.9133413400018813	ste
ps: 124	lr: 0.0001	evaluation reward: 1.36	epsilon: 0.9130007800018887	ste
episode: 241	score: 1.0	memory length: 44138	epsilon: 0.912662200001896	step
ps: 172	lr: 0.0001	evaluation reward: 1.35	epsilon: 0.9122701600019045	ste
episode: 242	score: 1.0	memory length: 44309	epsilon: 0.9119692000019111	ste
s: 171	lr: 0.0001	evaluation reward: 1.36	epsilon: 0.9114385600019226	ste
episode: 243	score: 2.0	memory length: 44507	epsilon: 0.9110762200019304	ste
ps: 198	lr: 0.0001	evaluation reward: 1.38	epsilon: 0.9098644600019568	ste
episode: 244	score: 1.0	memory length: 44659	epsilon: 0.909620920001962	step
ps: 152	lr: 0.0001	evaluation reward: 1.39	epsilon: 0.9084131200019883	ste
episode: 245	score: 3.0	memory length: 44927	epsilon: 0.9078686200020001	ste
ps: 268	lr: 0.0001	evaluation reward: 1.41	epsilon: 0.9074765800020086	ste
episode: 246	score: 2.0	memory length: 45110	epsilon: 0.9070251400020184	ste
ps: 183	lr: 0.0001	evaluation reward: 1.41	epsilon: 0.9067796200020237	ste
episode: 247	score: 0.0	memory length: 45234	epsilon: 0.9064806400020302	ste
ps: 124	lr: 0.0001	evaluation reward: 1.39	epsilon: 0.9062351200020355	ste
episode: 248	score: 5.0	memory length: 45551	epsilon: 0.9058985200020429	ste
ps: 317	lr: 0.0001	evaluation reward: 1.44	epsilon: 0.9055995400020493	ste
episode: 249	score: 1.0	memory length: 45722	epsilon: 0.9051025600020601	ste
ps: 171	lr: 0.0001	evaluation reward: 1.43	epsilon: 0.9047679400020674	ste
episode: 250	score: 0.0	memory length: 45845	epsilon: 0.9042749200020781	ste
s: 123	lr: 0.0001	evaluation reward: 1.4	epsilon: 0.9038809000020867	ste
episode: 251	score: 0.0	memory length: 45968	epsilon: 0.903449260002096	step
ps: 123	lr: 0.0001	evaluation reward: 1.39	epsilon: 0.9032057200021013	ste
episode: 252	score: 3.0	memory length: 46236	epsilon: 0.9028473400021091	ste
ps: 268	lr: 0.0001	evaluation reward: 1.4		
episode: 253	score: 2.0	memory length: 46455		
ps: 219	lr: 0.0001	evaluation reward: 1.42		
episode: 254	score: 1.0	memory length: 46607		
ps: 152	lr: 0.0001	evaluation reward: 1.42		
episode: 255	score: 0.0	memory length: 46730		
ps: 123	lr: 0.0001	evaluation reward: 1.4		
episode: 256	score: 2.0	memory length: 46928		
ps: 198	lr: 0.0001	evaluation reward: 1.38		
episode: 257	score: 3.0	memory length: 47156		
ps: 228	lr: 0.0001	evaluation reward: 1.39		
episode: 258	score: 0.0	memory length: 47280		
ps: 124	lr: 0.0001	evaluation reward: 1.36		
episode: 259	score: 1.0	memory length: 47431		
ps: 151	lr: 0.0001	evaluation reward: 1.37		
episode: 260	score: 0.0	memory length: 47555		
ps: 124	lr: 0.0001	evaluation reward: 1.35		
episode: 261	score: 1.0	memory length: 47725		
ps: 170	lr: 0.0001	evaluation reward: 1.31		
episode: 262	score: 1.0	memory length: 47876		
ps: 151	lr: 0.0001	evaluation reward: 1.31		
episode: 263	score: 3.0	memory length: 48127		
ps: 251	lr: 0.0001	evaluation reward: 1.32		
episode: 264	score: 1.0	memory length: 48296		
ps: 169	lr: 0.0001	evaluation reward: 1.32		
episode: 265	score: 3.0	memory length: 48545		
ps: 249	lr: 0.0001	evaluation reward: 1.33		
episode: 266	score: 2.0	memory length: 48744		
ps: 199	lr: 0.0001	evaluation reward: 1.33		
episode: 267	score: 2.0	memory length: 48962		
s: 218	lr: 0.0001	evaluation reward: 1.3		
episode: 268	score: 0.0	memory length: 49085		
ps: 123	lr: 0.0001	evaluation reward: 1.27		
episode: 269	score: 2.0	memory length: 49266		
ps: 181	lr: 0.0001	evaluation reward: 1.27		

episode: 270	score: 0.0	memory length: 49390	epsilon: 0.9026018200021144	ste
ps: 124	lr: 0.0001	evaluation reward: 1.26	epsilon: 0.902300860002121	step
episode: 271	score: 1.0	memory length: 49542	epsilon: 0.9020018800021274	ste
s: 152	lr: 0.0001	evaluation reward: 1.26	epsilon: 0.9015662800021369	ste
episode: 272	score: 1.0	memory length: 49693	epsilon: 0.9008178400021531	ste
ps: 151	lr: 0.0001	evaluation reward: 1.27	epsilon: 0.9004495600021611	ste
episode: 273	score: 2.0	memory length: 49913	epsilon: 0.8999644600021717	ste
ps: 220	lr: 0.0001	evaluation reward: 1.28	epsilon: 0.8995130200021815	ste
episode: 274	score: 7.0	memory length: 50291	epsilon: 0.8988734800021954	ste
ps: 378	lr: 0.0001	evaluation reward: 1.34	epsilon: 0.8985388600022026	ste
episode: 275	score: 2.0	memory length: 50477	epsilon: 0.898293340002208	step
ps: 186	lr: 0.0001	evaluation reward: 1.36	epsilon: 0.8980478200022133	ste
episode: 276	score: 3.0	memory length: 50722	epsilon: 0.8975508400022241	ste
ps: 245	lr: 0.0001	evaluation reward: 1.38	epsilon: 0.8973053200022294	ste
episode: 277	score: 3.0	memory length: 50950	epsilon: 0.8970598000022347	ste
ps: 228	lr: 0.0001	evaluation reward: 1.41	epsilon: 0.8966657800022433	ste
episode: 278	score: 0.0	memory length: 51073	epsilon: 0.8964202600022486	ste
ps: 123	lr: 0.0001	evaluation reward: 1.37	epsilon: 0.89589358000226	step
episode: 279	score: 2.0	memory length: 51273	epsilon: 0.8955569800022674	ste
ps: 200	lr: 0.0001	evaluation reward: 1.39	epsilon: 0.8953114600022727	ste
episode: 280	score: 1.0	memory length: 51442	epsilon: 0.8950105000022792	ste
ps: 169	lr: 0.0001	evaluation reward: 1.37	epsilon: 0.8945491600022892	ste
episode: 281	score: 0.0	memory length: 51566	epsilon: 0.8942462200022958	ste
s: 124	lr: 0.0001	evaluation reward: 1.35	epsilon: 0.8937532000023065	ste
episode: 282	score: 0.0	memory length: 51690	epsilon: 0.893361160002315	step
ps: 124	lr: 0.0001	evaluation reward: 1.35	epsilon: 0.8930602000023216	ste
episode: 283	score: 3.0	memory length: 51941	epsilon: 0.8926028200023315	ste
ps: 251	lr: 0.0001	evaluation reward: 1.36	epsilon: 0.892303840002338	step
episode: 284	score: 0.0	memory length: 52065	epsilon: 0.8920583200023433	ste
ps: 124	lr: 0.0001	evaluation reward: 1.34		
episode: 285	score: 0.0	memory length: 52189		
ps: 124	lr: 0.0001	evaluation reward: 1.34		
episode: 286	score: 2.0	memory length: 52388		
ps: 199	lr: 0.0001	evaluation reward: 1.34		
episode: 287	score: 0.0	memory length: 52512		
ps: 124	lr: 0.0001	evaluation reward: 1.34		
episode: 288	score: 3.0	memory length: 52778		
s: 266	lr: 0.0001	evaluation reward: 1.37		
episode: 289	score: 1.0	memory length: 52948		
ps: 170	lr: 0.0001	evaluation reward: 1.37		
episode: 290	score: 0.0	memory length: 53072		
ps: 124	lr: 0.0001	evaluation reward: 1.34		
episode: 291	score: 1.0	memory length: 53224		
ps: 152	lr: 0.0001	evaluation reward: 1.33		
episode: 292	score: 3.0	memory length: 53457		
ps: 233	lr: 0.0001	evaluation reward: 1.34		
episode: 293	score: 1.0	memory length: 53610		
ps: 153	lr: 0.0001	evaluation reward: 1.33		
episode: 294	score: 3.0	memory length: 53859		
ps: 249	lr: 0.0001	evaluation reward: 1.36		
episode: 295	score: 2.0	memory length: 54057		
s: 198	lr: 0.0001	evaluation reward: 1.37		
episode: 296	score: 1.0	memory length: 54209		
ps: 152	lr: 0.0001	evaluation reward: 1.38		
episode: 297	score: 3.0	memory length: 54440		
ps: 231	lr: 0.0001	evaluation reward: 1.4		
episode: 298	score: 1.0	memory length: 54591		
s: 151	lr: 0.0001	evaluation reward: 1.41		
episode: 299	score: 0.0	memory length: 54715		
ps: 124	lr: 0.0001	evaluation reward: 1.4		

episode: 300	score: 3.0	memory length: 54945	epsilon: 0.8916029200023532	ste
ps: 230	lr: 0.0001	evaluation reward: 1.42	epsilon: 0.8912108800023617	ste
episode: 301	score: 2.0	memory length: 55143	epsilon: 0.8908525000023695	ste
ps: 198	lr: 0.0001	evaluation reward: 1.42	epsilon: 0.8903634400023801	ste
episode: 302	score: 2.0	memory length: 55324	epsilon: 0.8899318000023895	ste
ps: 181	lr: 0.0001	evaluation reward: 1.43	epsilon: 0.8894783800023993	ste
episode: 303	score: 3.0	memory length: 55571	epsilon: 0.8890447600024087	ste
ps: 247	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.8886091600024182	ste
episode: 304	score: 2.0	memory length: 55789	epsilon: 0.8882151400024267	ste
ps: 218	lr: 0.0001	evaluation reward: 1.47	epsilon: 0.8877815200024362	ste
episode: 305	score: 3.0	memory length: 56018	epsilon: 0.8875379800024414	ste
ps: 229	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.88714396000245	step
episode: 306	score: 2.0	memory length: 56237	epsilon: 0.8867162800024593	ste
ps: 219	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.8863796800024666	ste
episode: 307	score: 2.0	memory length: 56457	epsilon: 0.8861361400024719	ste
ps: 220	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.8858906200024772	ste
episode: 308	score: 2.0	memory length: 56656	epsilon: 0.8853639400024886	ste
ps: 199	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.885118420002494	step
episode: 309	score: 2.0	memory length: 56875	epsilon: 0.8848729000024993	ste
ps: 219	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.8846293600025046	ste
episode: 310	score: 0.0	memory length: 56998	epsilon: 0.8842670200025124	ste
ps: 123	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.8840215000025178	ste
episode: 311	score: 2.0	memory length: 57197	epsilon: 0.8837759800025231	ste
s: 199	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.8833839400025316	ste
episode: 312	score: 2.0	memory length: 57413	epsilon: 0.8831404000025369	ste
ps: 216	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.8826909400025467	ste
episode: 313	score: 1.0	memory length: 57583	epsilon: 0.882354340002554	step
ps: 170	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.8820177400025613	ste
episode: 314	score: 0.0	memory length: 57706	epsilon: 0.8817187600025678	ste
ps: 123	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.8812336600025783	ste
episode: 315	score: 0.0	memory length: 57830	epsilon: 0.8808525000025851	ste
ps: 124	lr: 0.0001	evaluation reward: 1.46	epsilon: 0.8805118420002593	ste
episode: 316	score: 3.0	memory length: 58096	epsilon: 0.8802670200026046	ste
ps: 266	lr: 0.0001	evaluation reward: 1.47	epsilon: 0.8800233660002613	ste
episode: 317	score: 0.0	memory length: 58220	epsilon: 0.8797815200026214	ste
s: 124	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.8795434000026311	step
episode: 318	score: 0.0	memory length: 58344	epsilon: 0.8793018420002641	ste
ps: 124	lr: 0.0001	evaluation reward: 1.44	epsilon: 0.8790604000026518	ste
episode: 319	score: 0.0	memory length: 58467	epsilon: 0.8788187600026615	ste
ps: 123	lr: 0.0001	evaluation reward: 1.43	epsilon: 0.8785775980002672	ste
episode: 320	score: 2.0	memory length: 58650	epsilon: 0.8783363940002689	ste
ps: 183	lr: 0.0001	evaluation reward: 1.44	epsilon: 0.878095118420002706	ste
episode: 321	score: 0.0	memory length: 58774	epsilon: 0.8778539400027163	ste
ps: 124	lr: 0.0001	evaluation reward: 1.43	epsilon: 0.87761267020002723	ste
episode: 322	score: 0.0	memory length: 58898	epsilon: 0.8773714040002730	ste
ps: 124	lr: 0.0001	evaluation reward: 1.4	epsilon: 0.87713018760002737	ste
episode: 323	score: 2.0	memory length: 59096	epsilon: 0.8768889400027464	ste
ps: 198	lr: 0.0001	evaluation reward: 1.39	epsilon: 0.87664767020002753	ste
episode: 324	score: 0.0	memory length: 59219	epsilon: 0.8764064040002760	ste
ps: 123	lr: 0.0001	evaluation reward: 1.38	epsilon: 0.87616518760002767	ste
episode: 325	score: 3.0	memory length: 59446	epsilon: 0.8759239400027734	ste
ps: 227	lr: 0.0001	evaluation reward: 1.41	epsilon: 0.875682690940002781	ste
episode: 326	score: 1.0	memory length: 59616	epsilon: 0.8754414340002788	step
s: 170	lr: 0.0001	evaluation reward: 1.4	epsilon: 0.8751999400027955	ste
episode: 327	score: 1.0	memory length: 59786	epsilon: 0.87495857020002802	ste
ps: 170	lr: 0.0001	evaluation reward: 1.39	epsilon: 0.8747173040002809	ste
episode: 328	score: 1.0	memory length: 59937	epsilon: 0.87447593760002816	ste
ps: 151	lr: 0.0001	evaluation reward: 1.38	epsilon: 0.87423467020002823	ste
episode: 329	score: 3.0	memory length: 60182	epsilon: 0.8739934040002830	ste
ps: 245	lr: 0.0001	evaluation reward: 1.4	epsilon: 0.87375213760002837	ste

episode: 330	score: 2.0	memory length: 60404	epsilon: 0.8807941000025878	ste
ps: 222	lr: 0.0001	evaluation reward: 1.41	epsilon: 0.8805485800025932	ste
episode: 331	score: 0.0	memory length: 60528	epsilon: 0.8801565400026017	ste
ps: 124	lr: 0.0001	evaluation reward: 1.36	epsilon: 0.8798140000026091	ste
episode: 332	score: 2.0	memory length: 60726	epsilon: 0.8794774000026164	ste
ps: 198	lr: 0.0001	evaluation reward: 1.38	epsilon: 0.8790418000026259	ste
episode: 333	score: 1.0	memory length: 60899	epsilon: 0.8784537400026386	ste
ps: 173	lr: 0.0001	evaluation reward: 1.39	epsilon: 0.878208220002644	step
episode: 334	score: 1.0	memory length: 61069	epsilon: 0.8779627000026493	ste
ps: 170	lr: 0.0001	evaluation reward: 1.4	epsilon: 0.8773409800026628	ste
episode: 335	score: 2.0	memory length: 61289	epsilon: 0.8769469600026714	ste
ps: 220	lr: 0.0001	evaluation reward: 1.41	epsilon: 0.8767014400026767	ste
episode: 336	score: 4.0	memory length: 61586	epsilon: 0.8761510000026886	ste
ps: 297	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.8756560000026994	ste
episode: 337	score: 0.0	memory length: 61710	epsilon: 0.8750164600027133	ste
s: 124	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.8747709400027186	ste
episode: 338	score: 0.0	memory length: 61834	epsilon: 0.8744343400027259	ste
ps: 124	lr: 0.0001	evaluation reward: 1.45	epsilon: 0.8741908000027312	ste
episode: 339	score: 4.0	memory length: 62148	epsilon: 0.8737571800027406	ste
ps: 314	lr: 0.0001	evaluation reward: 1.49	epsilon: 0.8733631600027492	ste
episode: 340	score: 2.0	memory length: 62347	epsilon: 0.8729691400027577	ste
ps: 199	lr: 0.0001	evaluation reward: 1.51	epsilon: 0.8724424600027691	ste
episode: 341	score: 0.0	memory length: 62471	epsilon: 0.8720425000027778	ste
ps: 124	lr: 0.0001	evaluation reward: 1.5	epsilon: 0.8717059000027851	ste
episode: 342	score: 4.0	memory length: 62749	epsilon: 0.8714603800027905	ste
ps: 278	lr: 0.0001	evaluation reward: 1.53	epsilon: 0.870975280002801	step
episode: 343	score: 3.0	memory length: 62999	epsilon: 0.8706763000028075	ste
ps: 250	lr: 0.0001	evaluation reward: 1.54	epsilon: 0.8704307800028128	ste
episode: 344	score: 0.0	memory length: 63123	epsilon: 0.8700902200028202	ste
ps: 124	lr: 0.0001	evaluation reward: 1.53		
episode: 345	score: 2.0	memory length: 63322		
ps: 199	lr: 0.0001	evaluation reward: 1.52		
episode: 346	score: 0.0	memory length: 63446		
ps: 124	lr: 0.0001	evaluation reward: 1.5		
episode: 347	score: 1.0	memory length: 63616		
ps: 170	lr: 0.0001	evaluation reward: 1.51		
episode: 348	score: 0.0	memory length: 63739		
ps: 123	lr: 0.0001	evaluation reward: 1.46		
episode: 349	score: 2.0	memory length: 63958		
ps: 219	lr: 0.0001	evaluation reward: 1.47		
episode: 350	score: 2.0	memory length: 64157		
ps: 199	lr: 0.0001	evaluation reward: 1.49		
episode: 351	score: 2.0	memory length: 64356		
ps: 199	lr: 0.0001	evaluation reward: 1.51		
episode: 352	score: 3.0	memory length: 64622		
ps: 266	lr: 0.0001	evaluation reward: 1.51		
episode: 353	score: 2.0	memory length: 64824		
ps: 202	lr: 0.0001	evaluation reward: 1.51		
episode: 354	score: 1.0	memory length: 64994		
ps: 170	lr: 0.0001	evaluation reward: 1.51		
episode: 355	score: 0.0	memory length: 65118		
ps: 124	lr: 0.0001	evaluation reward: 1.51		
episode: 356	score: 3.0	memory length: 65363		
s: 245	lr: 0.0001	evaluation reward: 1.52		
episode: 357	score: 1.0	memory length: 65514		
ps: 151	lr: 0.0001	evaluation reward: 1.5		
episode: 358	score: 0.0	memory length: 65638		
ps: 124	lr: 0.0001	evaluation reward: 1.5		
episode: 359	score: 1.0	memory length: 65810		
ps: 172	lr: 0.0001	evaluation reward: 1.5		

episode: 360	score: 4.0	memory length: 66086	epsilon: 0.8695437400028321	ste
ps: 276	lr: 0.0001	evaluation reward: 1.54		
episode: 361	score: 0.0	memory length: 66210	epsilon: 0.8692982200028374	ste
ps: 124	lr: 0.0001	evaluation reward: 1.53		
episode: 362	score: 0.0	memory length: 66334	epsilon: 0.8690527000028427	ste
ps: 124	lr: 0.0001	evaluation reward: 1.52		
episode: 363	score: 1.0	memory length: 66504	epsilon: 0.86871610000285	step
s: 170	lr: 0.0001	evaluation reward: 1.5		
episode: 364	score: 2.0	memory length: 66705	epsilon: 0.8683181200028587	ste
ps: 201	lr: 0.0001	evaluation reward: 1.51		
episode: 365	score: 1.0	memory length: 66857	epsilon: 0.8680171600028652	ste
ps: 152	lr: 0.0001	evaluation reward: 1.49		
episode: 366	score: 4.0	memory length: 67175	epsilon: 0.8673875200028789	ste
ps: 318	lr: 0.0001	evaluation reward: 1.51		
episode: 367	score: 6.0	memory length: 67593	epsilon: 0.8665598800028969	ste
ps: 418	lr: 0.0001	evaluation reward: 1.55		
episode: 368	score: 2.0	memory length: 67792	epsilon: 0.8661658600029054	ste
ps: 199	lr: 0.0001	evaluation reward: 1.57		
episode: 369	score: 2.0	memory length: 67991	epsilon: 0.865771840002914	step
s: 199	lr: 0.0001	evaluation reward: 1.57		
episode: 370	score: 2.0	memory length: 68190	epsilon: 0.8653778200029225	ste
ps: 199	lr: 0.0001	evaluation reward: 1.59		
episode: 371	score: 4.0	memory length: 68489	epsilon: 0.8647858000029354	ste
ps: 299	lr: 0.0001	evaluation reward: 1.62		
episode: 372	score: 2.0	memory length: 68687	epsilon: 0.8643937600029439	ste
ps: 198	lr: 0.0001	evaluation reward: 1.63		
episode: 373	score: 3.0	memory length: 68938	epsilon: 0.8638967800029547	ste
ps: 251	lr: 0.0001	evaluation reward: 1.64		
episode: 374	score: 0.0	memory length: 69062	epsilon: 0.86365126000296	step
s: 124	lr: 0.0001	evaluation reward: 1.57		
episode: 375	score: 1.0	memory length: 69232	epsilon: 0.8633146600029673	ste
ps: 170	lr: 0.0001	evaluation reward: 1.56		
episode: 376	score: 1.0	memory length: 69386	epsilon: 0.8630097400029739	ste
ps: 154	lr: 0.0001	evaluation reward: 1.54		
episode: 377	score: 1.0	memory length: 69556	epsilon: 0.8626731400029812	ste
ps: 170	lr: 0.0001	evaluation reward: 1.52		
episode: 378	score: 4.0	memory length: 69848	epsilon: 0.8620949800029938	ste
ps: 292	lr: 0.0001	evaluation reward: 1.56		
episode: 379	score: 0.0	memory length: 69972	epsilon: 0.8618494600029991	ste
ps: 124	lr: 0.0001	evaluation reward: 1.54		
episode: 380	score: 2.0	memory length: 70170	epsilon: 0.8614574200030076	ste
ps: 198	lr: 0.0001	evaluation reward: 1.55		
episode: 381	score: 1.0	memory length: 70339	epsilon: 0.8611228000030149	ste
ps: 169	lr: 0.0001	evaluation reward: 1.56		
episode: 382	score: 0.0	memory length: 70463	epsilon: 0.8608772800030202	ste
ps: 124	lr: 0.0001	evaluation reward: 1.56		
episode: 383	score: 0.0	memory length: 70586	epsilon: 0.8606337400030255	ste
ps: 123	lr: 0.0001	evaluation reward: 1.53		
episode: 384	score: 0.0	memory length: 70710	epsilon: 0.8603882200030308	ste
ps: 124	lr: 0.0001	evaluation reward: 1.53		
episode: 385	score: 3.0	memory length: 70978	epsilon: 0.8598575800030424	ste
ps: 268	lr: 0.0001	evaluation reward: 1.56		
episode: 386	score: 0.0	memory length: 71102	epsilon: 0.8596120600030477	ste
ps: 124	lr: 0.0001	evaluation reward: 1.54		
episode: 387	score: 3.0	memory length: 71348	epsilon: 0.8591249800030583	ste
ps: 246	lr: 0.0001	evaluation reward: 1.57		
episode: 388	score: 0.0	memory length: 71472	epsilon: 0.8588794600030636	ste
ps: 124	lr: 0.0001	evaluation reward: 1.54		
episode: 389	score: 3.0	memory length: 71700	epsilon: 0.8584280200030734	ste
ps: 228	lr: 0.0001	evaluation reward: 1.56		

episode: 390	score: 3.0	memory length: 71929	epsilon: 0.8579746000030832	ste
ps: 229	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.8577290800030886	ste
episode: 391	score: 0.0	memory length: 72053	epsilon: 0.8573350600030971	ste
ps: 124	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.8570341000031036	ste
episode: 392	score: 2.0	memory length: 72252	epsilon: 0.8566400800031122	ste
ps: 199	lr: 0.0001	evaluation reward: 1.57	epsilon: 0.8563965400031175	ste
episode: 393	score: 1.0	memory length: 72404	epsilon: 0.8559055000031282	ste
ps: 152	lr: 0.0001	evaluation reward: 1.57	epsilon: 0.8554144600031388	ste
episode: 394	score: 2.0	memory length: 72603	epsilon: 0.8550204400031474	ste
ps: 199	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.8546264200031559	ste
episode: 395	score: 0.0	memory length: 72726	epsilon: 0.8543828800031612	ste
ps: 123	lr: 0.0001	evaluation reward: 1.54	epsilon: 0.8539492600031706	ste
episode: 396	score: 3.0	memory length: 72974	epsilon: 0.8533869400031828	ste
ps: 248	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.8528008600031955	ste
episode: 397	score: 3.0	memory length: 73222	epsilon: 0.852367240003205	step
ps: 248	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.8519118400032148	ste
episode: 398	score: 2.0	memory length: 73421	epsilon: 0.8515178200032234	ste
ps: 199	lr: 0.0001	evaluation reward: 1.57	epsilon: 0.8512742800032287	ste
episode: 399	score: 2.0	memory length: 73620	epsilon: 0.851028760003234	step
ps: 199	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.8505951400032434	ste
episode: 400	score: 0.0	memory length: 73743	epsilon: 0.8500229200032559	ste
ps: 123	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.8496843400032632	ste
episode: 401	score: 2.0	memory length: 73962	epsilon: 0.8491358800032751	ste
ps: 219	lr: 0.0001	evaluation reward: 1.56	epsilon: 0.8486092000032865	ste
episode: 402	score: 4.0	memory length: 74246	epsilon: 0.8483636800032919	ste
ps: 284	lr: 0.0001	evaluation reward: 1.58	epsilon: 0.8479696600033004	ste
episode: 403	score: 4.0	memory length: 74542	epsilon: 0.8475340600033099	ste
ps: 296	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.8472885400033152	ste
episode: 404	score: 2.0	memory length: 74761	epsilon: 0.8467994800033258	ste
s: 219	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.8463084400033365	ste
episode: 405	score: 3.0	memory length: 74991	epsilon: 0.84599118400033365	ste
ps: 230	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.8456264200033418	ste
episode: 406	score: 2.0	memory length: 75190	epsilon: 0.8452800860003348	ste
ps: 199	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.8449135880003358	ste
episode: 407	score: 0.0	memory length: 75313	epsilon: 0.8445626420003368	ste
ps: 123	lr: 0.0001	evaluation reward: 1.57	epsilon: 0.8442178200033778	ste
episode: 408	score: 0.0	memory length: 75437	epsilon: 0.8438694800033888	ste
s: 124	lr: 0.0001	evaluation reward: 1.55	epsilon: 0.843512742800033998	ste
episode: 409	score: 2.0	memory length: 75656	epsilon: 0.843165624000340098	ste
ps: 219	lr: 0.0001	evaluation reward: 1.55	epsilon: 0.842815178200034108	ste
episode: 410	score: 5.0	memory length: 75945	epsilon: 0.842468434000341198	ste
ps: 289	lr: 0.0001	evaluation reward: 1.6	epsilon: 0.84212127428000342288	ste
episode: 411	score: 1.0	memory length: 76116	epsilon: 0.841774860003433898	ste
ps: 171	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.84142857140003434898	ste
episode: 412	score: 4.0	memory length: 76393	epsilon: 0.841081184000343598	ste
ps: 277	lr: 0.0001	evaluation reward: 1.61	epsilon: 0.840734624000343788	ste
episode: 413	score: 3.0	memory length: 76659	epsilon: 0.840388151400034878	ste
ps: 266	lr: 0.0001	evaluation reward: 1.63	epsilon: 0.840041681700034968	ste
episode: 414	score: 0.0	memory length: 76783	epsilon: 0.839695212000035058	ste
ps: 124	lr: 0.0001	evaluation reward: 1.63	epsilon: 0.839349742300035148	ste
episode: 415	score: 2.0	memory length: 76982	epsilon: 0.838904272600035238	ste
ps: 199	lr: 0.0001	evaluation reward: 1.65	epsilon: 0.838558802900035328	ste
episode: 416	score: 2.0	memory length: 77202	epsilon: 0.838213333200035418	ste
ps: 220	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.837867863500035508	ste
episode: 417	score: 0.0	memory length: 77326	epsilon: 0.837532403800035598	ste
ps: 124	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.837186034100035688	ste
episode: 418	score: 3.0	memory length: 77573	epsilon: 0.836840664400035778	ste
ps: 247	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.836494294700035868	ste
episode: 419	score: 3.0	memory length: 77821	epsilon: 0.836148925000035958	ste
ps: 248	lr: 0.0001	evaluation reward: 1.7	epsilon: 0.835792555300036048	ste

episode: 420	score: 1.0	memory length: 77991	epsilon: 0.8459718400033438	ste
ps: 170	lr: 0.0001	evaluation reward: 1.69	epsilon: 0.8457263200033491	ste
episode: 421	score: 0.0	memory length: 78115	epsilon: 0.8452412200033597	ste
ps: 124	lr: 0.0001	evaluation reward: 1.69	epsilon: 0.8447521600033703	ste
episode: 422	score: 3.0	memory length: 78360	epsilon: 0.8443581400033788	ste
ps: 245	lr: 0.0001	evaluation reward: 1.72	epsilon: 0.8438690800033894	ste
episode: 423	score: 3.0	memory length: 78607	epsilon: 0.8433166600034014	ste
ps: 247	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.8430711400034068	ste
episode: 424	score: 2.0	memory length: 78806	epsilon: 0.8428256200034121	ste
ps: 199	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.8421266800034273	ste
episode: 425	score: 3.0	memory length: 79053	epsilon: 0.8425266400034186	ste
ps: 247	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.8413723000034437	ste
episode: 426	score: 4.0	memory length: 79332	epsilon: 0.8409386800034531	ste
ps: 279	lr: 0.0001	evaluation reward: 1.78	epsilon: 0.8404456600034638	ste
episode: 427	score: 0.0	memory length: 79456	epsilon: 0.8401447000034703	ste
ps: 124	lr: 0.0001	evaluation reward: 1.77	epsilon: 0.8397526600034788	ste
episode: 428	score: 0.0	memory length: 79580	epsilon: 0.8393606200034873	ste
ps: 124	lr: 0.0001	evaluation reward: 1.76	epsilon: 0.8391151000034927	ste
episode: 429	score: 1.0	memory length: 79731	epsilon: 0.8387785000035	steps:
ps: 151	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8383448800035094	ste
episode: 430	score: 2.0	memory length: 79933	epsilon: 0.8380993600035147	ste
ps: 202	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8378003800035212	ste
episode: 431	score: 6.0	memory length: 80314	epsilon: 0.837442000003529	step
ps: 381	lr: 0.0001	evaluation reward: 1.8	epsilon: 0.8371054000035363	ste
episode: 432	score: 2.0	memory length: 80533	epsilon: 0.8367628600035437	ste
ps: 219	lr: 0.0001	evaluation reward: 1.8	epsilon: 0.8363153800035534	ste
episode: 433	score: 3.0	memory length: 80782	epsilon: 0.8359570000035612	ste
ps: 249	lr: 0.0001	evaluation reward: 1.82	epsilon: 0.8357114800035665	ste
episode: 434	score: 1.0	memory length: 80934	epsilon: 0.8354659600035719	ste
ps: 152	lr: 0.0001	evaluation reward: 1.82	epsilon: 0.8348323600035856	ste
episode: 435	score: 2.0	memory length: 81132		
ps: 198	lr: 0.0001	evaluation reward: 1.82		
episode: 436	score: 2.0	memory length: 81330		
ps: 198	lr: 0.0001	evaluation reward: 1.8		
episode: 437	score: 0.0	memory length: 81454		
ps: 124	lr: 0.0001	evaluation reward: 1.8		
episode: 438	score: 1.0	memory length: 81624		
170	lr: 0.0001	evaluation reward: 1.81		
episode: 439	score: 2.0	memory length: 81843		
ps: 219	lr: 0.0001	evaluation reward: 1.79		
episode: 440	score: 0.0	memory length: 81967		
ps: 124	lr: 0.0001	evaluation reward: 1.77		
episode: 441	score: 1.0	memory length: 82118		
ps: 151	lr: 0.0001	evaluation reward: 1.78		
episode: 442	score: 2.0	memory length: 82299		
s: 181	lr: 0.0001	evaluation reward: 1.76		
episode: 443	score: 1.0	memory length: 82469		
ps: 170	lr: 0.0001	evaluation reward: 1.74		
episode: 444	score: 1.0	memory length: 82642		
ps: 173	lr: 0.0001	evaluation reward: 1.75		
episode: 445	score: 3.0	memory length: 82868		
ps: 226	lr: 0.0001	evaluation reward: 1.76		
episode: 446	score: 2.0	memory length: 83049		
ps: 181	lr: 0.0001	evaluation reward: 1.78		
episode: 447	score: 0.0	memory length: 83173		
ps: 124	lr: 0.0001	evaluation reward: 1.77		
episode: 448	score: 0.0	memory length: 83297		
ps: 124	lr: 0.0001	evaluation reward: 1.77		
episode: 449	score: 4.0	memory length: 83617		
ps: 320	lr: 0.0001	evaluation reward: 1.79		

episode: 450	score: 2.0	memory length: 83816	epsilon: 0.8344383400035942	ste
ps: 199	lr: 0.0001	evaluation reward: 1.79		
episode: 451	score: 0.0	memory length: 83940	epsilon: 0.8341928200035995	ste
ps: 124	lr: 0.0001	evaluation reward: 1.77		
episode: 452	score: 1.0	memory length: 84094	epsilon: 0.8338879000036061	ste
ps: 154	lr: 0.0001	evaluation reward: 1.75		
episode: 453	score: 0.0	memory length: 84218	epsilon: 0.8336423800036115	ste
ps: 124	lr: 0.0001	evaluation reward: 1.73		
episode: 454	score: 0.0	memory length: 84341	epsilon: 0.8333988400036167	ste
ps: 123	lr: 0.0001	evaluation reward: 1.72		
episode: 455	score: 2.0	memory length: 84560	epsilon: 0.8329652200036262	ste
ps: 219	lr: 0.0001	evaluation reward: 1.74		
episode: 456	score: 3.0	memory length: 84792	epsilon: 0.8325058600036361	ste
ps: 232	lr: 0.0001	evaluation reward: 1.74		
episode: 457	score: 0.0	memory length: 84916	epsilon: 0.8322603400036415	ste
ps: 124	lr: 0.0001	evaluation reward: 1.73		
episode: 458	score: 0.0	memory length: 85040	epsilon: 0.8320148200036468	ste
ps: 124	lr: 0.0001	evaluation reward: 1.73		
episode: 459	score: 2.0	memory length: 85239	epsilon: 0.8316208000036553	ste
ps: 199	lr: 0.0001	evaluation reward: 1.74		
episode: 460	score: 3.0	memory length: 85467	epsilon: 0.8311693600036651	ste
ps: 228	lr: 0.0001	evaluation reward: 1.73		
episode: 461	score: 2.0	memory length: 85666	epsilon: 0.8307753400036737	ste
ps: 199	lr: 0.0001	evaluation reward: 1.75		
episode: 462	score: 0.0	memory length: 85790	epsilon: 0.830529820003679	step
s: 124	lr: 0.0001	evaluation reward: 1.75		
episode: 463	score: 3.0	memory length: 86038	epsilon: 0.8300387800036897	ste
ps: 248	lr: 0.0001	evaluation reward: 1.77		
episode: 464	score: 0.0	memory length: 86162	epsilon: 0.829793260003695	step
s: 124	lr: 0.0001	evaluation reward: 1.75		
episode: 465	score: 0.0	memory length: 86285	epsilon: 0.8295497200037003	ste
ps: 123	lr: 0.0001	evaluation reward: 1.74		
episode: 466	score: 5.0	memory length: 86628	epsilon: 0.828870580003715	step
s: 343	lr: 0.0001	evaluation reward: 1.75		
episode: 467	score: 3.0	memory length: 86875	epsilon: 0.8283815200037257	ste
ps: 247	lr: 0.0001	evaluation reward: 1.72		
episode: 468	score: 1.0	memory length: 87027	epsilon: 0.8280805600037322	ste
ps: 152	lr: 0.0001	evaluation reward: 1.71		
episode: 469	score: 4.0	memory length: 87342	epsilon: 0.8274568600037457	ste
ps: 315	lr: 0.0001	evaluation reward: 1.73		
episode: 470	score: 0.0	memory length: 87466	epsilon: 0.8272113400037511	ste
ps: 124	lr: 0.0001	evaluation reward: 1.71		
episode: 471	score: 0.0	memory length: 87590	epsilon: 0.8269658200037564	ste
ps: 124	lr: 0.0001	evaluation reward: 1.67		
episode: 472	score: 2.0	memory length: 87788	epsilon: 0.8265737800037649	ste
ps: 198	lr: 0.0001	evaluation reward: 1.67		
episode: 473	score: 2.0	memory length: 87987	epsilon: 0.8261797600037735	ste
ps: 199	lr: 0.0001	evaluation reward: 1.66		
episode: 474	score: 2.0	memory length: 88207	epsilon: 0.8257441600037829	ste
ps: 220	lr: 0.0001	evaluation reward: 1.68		
episode: 475	score: 0.0	memory length: 88330	epsilon: 0.8255006200037882	ste
ps: 123	lr: 0.0001	evaluation reward: 1.67		
episode: 476	score: 2.0	memory length: 88552	epsilon: 0.8250610600037978	ste
ps: 222	lr: 0.0001	evaluation reward: 1.68		
episode: 477	score: 1.0	memory length: 88722	epsilon: 0.8247244600038051	ste
ps: 170	lr: 0.0001	evaluation reward: 1.68		
episode: 478	score: 0.0	memory length: 88846	epsilon: 0.8244789400038104	ste
ps: 124	lr: 0.0001	evaluation reward: 1.64		
episode: 479	score: 2.0	memory length: 89045	epsilon: 0.8240849200038189	ste
ps: 199	lr: 0.0001	evaluation reward: 1.66		

episode: 480	score: 0.0	memory length: 89169	epsilon: 0.8238394000038243	ste
ps: 124	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.8235384400038308	ste
episode: 481	score: 1.0	memory length: 89321	epsilon: 0.8231444200038394	ste
ps: 152	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.8226949600038491	ste
episode: 482	score: 2.0	memory length: 89520	epsilon: 0.8224494400038544	ste
ps: 199	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.8221504600038609	ste
episode: 483	score: 3.0	memory length: 89747	epsilon: 0.8218138600038682	ste
ps: 227	lr: 0.0001	evaluation reward: 1.69	epsilon: 0.8215703200038735	ste
episode: 484	score: 0.0	memory length: 89871	epsilon: 0.8207842600038906	ste
ps: 124	lr: 0.0001	evaluation reward: 1.69	epsilon: 0.8202714400039017	ste
episode: 485	score: 1.0	memory length: 90022	epsilon: 0.8198794000039102	ste
ps: 151	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.8193804400039211	ste
episode: 486	score: 1.0	memory length: 90192	epsilon: 0.8189309800039308	ste
ps: 170	lr: 0.0001	evaluation reward: 1.68	epsilon: 0.8183845000039427	ste
episode: 487	score: 0.0	memory length: 90315	epsilon: 0.8179489000039522	ste
ps: 123	lr: 0.0001	evaluation reward: 1.65	epsilon: 0.8177033800039575	ste
episode: 488	score: 2.0	memory length: 90513	epsilon: 0.8172103600039682	ste
s: 198	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.8168163400039767	ste
episode: 489	score: 2.0	memory length: 90712	epsilon: 0.8164757800039841	ste
ps: 199	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.8160817600039927	ste
episode: 490	score: 4.0	memory length: 90971	epsilon: 0.8156877400040012	ste
ps: 259	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.8152957000040097	ste
episode: 491	score: 2.0	memory length: 91169	epsilon: 0.8149947400040163	ste
ps: 198	lr: 0.0001	evaluation reward: 1.69	epsilon: 0.814547260004026	step
episode: 492	score: 4.0	memory length: 91421	epsilon: 0.8142463000040325	ste
ps: 252	lr: 0.0001	evaluation reward: 1.71	epsilon: 0.8138522800040411	ste
episode: 493	score: 3.0	memory length: 91648	epsilon: 0.8136067600040464	ste
ps: 227	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.8131573000040562	ste
episode: 494	score: 4.0	memory length: 91924	epsilon: 0.8128187200040635	ste
ps: 276	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.8125957000040722	ste
episode: 495	score: 2.0	memory length: 92144	epsilon: 0.81236067600040811	ste
ps: 220	lr: 0.0001	evaluation reward: 1.77	epsilon: 0.8121444200040909	ste
episode: 496	score: 0.0	memory length: 92268	epsilon: 0.8119194400041007	ste
ps: 124	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8117033800041095	ste
episode: 497	score: 3.0	memory length: 92517	epsilon: 0.8114757800041183	ste
ps: 249	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8112444200041271	ste
episode: 498	score: 2.0	memory length: 92716	epsilon: 0.8110194400041359	ste
ps: 199	lr: 0.0001	evaluation reward: 1.74	epsilon: 0.8107938000041447	ste
episode: 499	score: 1.0	memory length: 92888	epsilon: 0.8105687740004535	ste
ps: 172	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.8103434400045423	ste
episode: 500	score: 2.0	memory length: 93087	epsilon: 0.81011819440004621	ste
ps: 199	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.8098930000046309	ste
episode: 501	score: 2.0	memory length: 93286	epsilon: 0.8096677740004707	ste
ps: 199	lr: 0.0001	evaluation reward: 1.75	epsilon: 0.8094424630004795	ste
episode: 502	score: 2.0	memory length: 93484	epsilon: 0.8092171340004883	ste
ps: 198	lr: 0.0001	evaluation reward: 1.73	epsilon: 0.8090028000049071	ste
episode: 503	score: 1.0	memory length: 93636	epsilon: 0.8087774740004979	ste
ps: 152	lr: 0.0001	evaluation reward: 1.7	epsilon: 0.8085522400049867	ste
episode: 504	score: 3.0	memory length: 93862	epsilon: 0.8083270000049955	step
s: 226	lr: 0.0001	evaluation reward: 1.71	epsilon: 0.8081017600049943	ste
episode: 505	score: 1.0	memory length: 94014	epsilon: 0.8078764300049931	ste
ps: 152	lr: 0.0001	evaluation reward: 1.69	epsilon: 0.8076512000049919	ste
episode: 506	score: 2.0	memory length: 94213	epsilon: 0.8074259700049907	ste
ps: 199	lr: 0.0001	evaluation reward: 1.69	epsilon: 0.8072017400049905	ste
episode: 507	score: 0.0	memory length: 94337	epsilon: 0.8069765100049903	ste
ps: 124	lr: 0.0001	evaluation reward: 1.69	epsilon: 0.8067522800049901	ste
episode: 508	score: 3.0	memory length: 94564	epsilon: 0.8065270500049901	ste
ps: 227	lr: 0.0001	evaluation reward: 1.72	epsilon: 0.8063028200049901	ste
episode: 509	score: 1.0	memory length: 94735	epsilon: 0.8060775900049901	ste
ps: 171	lr: 0.0001	evaluation reward: 1.71	epsilon: 0.8058523600049901	ste

episode: 510	score: 2.0	memory length: 94936	epsilon: 0.8124207400040722	ste
ps: 201	lr: 0.0001	evaluation reward: 1.68	epsilon: 0.8121752200040775	ste
episode: 511	score: 0.0	memory length: 95060	epsilon: 0.8115238000040916	ste
ps: 124	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.8111297800041002	ste
episode: 512	score: 5.0	memory length: 95389	epsilon: 0.8108842600041055	ste
ps: 329	lr: 0.0001	evaluation reward: 1.68	epsilon: 0.810583300004112	step
episode: 513	score: 2.0	memory length: 95588	epsilon: 0.8103377800041174	ste
ps: 199	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.8099239600041264	ste
episode: 514	score: 0.0	memory length: 95712	epsilon: 0.8094250000041372	ste
ps: 124	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.8091814600041425	ste
episode: 515	score: 1.0	memory length: 95864	epsilon: 0.8089359400041478	ste
s: 152	lr: 0.0001	evaluation reward: 1.66	epsilon: 0.808052860004167	step
episode: 516	score: 0.0	memory length: 95988	epsilon: 0.8078073400041723	ste
ps: 124	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.8073757000041817	ste
episode: 517	score: 3.0	memory length: 96197	epsilon: 0.8068846600041923	ste
ps: 209	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.8066411200041976	ste
episode: 518	score: 3.0	memory length: 96449	epsilon: 0.8061540400042082	ste
ps: 252	lr: 0.0001	evaluation reward: 1.67	epsilon: 0.8057600200042168	ste
episode: 519	score: 0.0	memory length: 96572	epsilon: 0.8054234200042241	ste
ps: 123	lr: 0.0001	evaluation reward: 1.64	epsilon: 0.8051244400042306	ste
episode: 520	score: 0.0	memory length: 96696	epsilon: 0.8047304200042391	ste
ps: 124	lr: 0.0001	evaluation reward: 1.63	epsilon: 0.8041027600042527	ste
episode: 521	score: 2.0	memory length: 96894	epsilon: 0.803859220004258	step
ps: 198	lr: 0.0001	evaluation reward: 1.65	epsilon: 0.8036156800042633	ste
episode: 522	score: 0.0	memory length: 97018	epsilon: 0.8030256400042761	ste
ps: 124	lr: 0.0001	evaluation reward: 1.62	epsilon: 0.8025959800042854	ste
episode: 523	score: 0.0	memory length: 97142	epsilon: 0.802203940004294	ste
s: 124	lr: 0.0001	evaluation reward: 1.59	epsilon: 0.8019584200042993	st
episode: 524	score: 0.0	memory length: 97266		
ps: 124	lr: 0.0001	evaluation reward: 1.57		
episode: 525	score: 2.0	memory length: 97484		
ps: 218	lr: 0.0001	evaluation reward: 1.56		
episode: 526	score: 3.0	memory length: 97732		
ps: 248	lr: 0.0001	evaluation reward: 1.55		
episode: 527	score: 0.0	memory length: 97855		
ps: 123	lr: 0.0001	evaluation reward: 1.55		
episode: 528	score: 3.0	memory length: 98101		
ps: 246	lr: 0.0001	evaluation reward: 1.58		
episode: 529	score: 2.0	memory length: 98300		
ps: 199	lr: 0.0001	evaluation reward: 1.59		
episode: 530	score: 1.0	memory length: 98470		
ps: 170	lr: 0.0001	evaluation reward: 1.58		
episode: 531	score: 1.0	memory length: 98621		
ps: 151	lr: 0.0001	evaluation reward: 1.53		
episode: 532	score: 2.0	memory length: 98820		
ps: 199	lr: 0.0001	evaluation reward: 1.53		
episode: 533	score: 4.0	memory length: 99137		
ps: 317	lr: 0.0001	evaluation reward: 1.54		
episode: 534	score: 0.0	memory length: 99260		
s: 123	lr: 0.0001	evaluation reward: 1.53		
episode: 535	score: 0.0	memory length: 99383		
ps: 123	lr: 0.0001	evaluation reward: 1.51		
episode: 536	score: 4.0	memory length: 99681		
ps: 298	lr: 0.0001	evaluation reward: 1.53		
episode: 537	score: 2.0	memory length: 99898		
ps: 217	lr: 0.0001	evaluation reward: 1.55		
episode: 538	score: 2.0	memory length: 100096		
ps: 198	lr: 0.0001	evaluation reward: 1.56		
episode: 539	score: 0.0	memory length: 100220		
eps: 124	lr: 4e-05	evaluation reward: 1.54		

episode: 540 score: 2.0 memory length: 100419	epsilon: 0.8015644000043078	st
eps: 199 lr: 4e-05 evaluation reward: 1.56	epsilon: 0.8013188800043132	st
episode: 541 score: 0.0 memory length: 100543	epsilon: 0.8007882400043247	st
eps: 124 lr: 4e-05 evaluation reward: 1.55	epsilon: 0.80054272000433	step
episode: 542 score: 3.0 memory length: 100811	epsilon: 0.8002972000043354	st
eps: 268 lr: 4e-05 evaluation reward: 1.56	epsilon: 0.7998437800043452	st
episode: 543 score: 0.0 memory length: 100935	epsilon: 0.7994873800043529	st
s: 124 lr: 4e-05 evaluation reward: 1.55	epsilon: 0.7992418600043583	st
episode: 544 score: 0.0 memory length: 101059	epsilon: 0.7989963400043636	st
eps: 124 lr: 4e-05 evaluation reward: 1.54	epsilon: 0.7986023200043721	st
episode: 545 score: 3.0 memory length: 101288	epsilon: 0.7982023600043808	st
eps: 229 lr: 4e-05 evaluation reward: 1.54	epsilon: 0.7976182600043935	st
episode: 546 score: 2.0 memory length: 101468	epsilon: 0.7972816600044008	st
eps: 180 lr: 4e-05 evaluation reward: 1.54	epsilon: 0.7970361400044061	st
episode: 547 score: 0.0 memory length: 101592	epsilon: 0.7967926000044114	st
eps: 124 lr: 4e-05 evaluation reward: 1.54	epsilon: 0.7962065200044242	st
episode: 548 score: 0.0 memory length: 101716	epsilon: 0.7956917200044353	st
eps: 124 lr: 4e-05 evaluation reward: 1.54	epsilon: 0.7951789000044465	st
episode: 549 score: 2.0 memory length: 101915	epsilon: 0.794741320004456	ste
eps: 199 lr: 4e-05 evaluation reward: 1.52	epsilon: 0.7943057200044654	st
episode: 550 score: 2.0 memory length: 102117	epsilon: 0.793818640004476	ste
eps: 202 lr: 4e-05 evaluation reward: 1.52	epsilon: 0.7933295800044866	st
episode: 551 score: 4.0 memory length: 102412	epsilon: 0.7930840600044919	st
eps: 295 lr: 4e-05 evaluation reward: 1.56	epsilon: 0.7925910400045026	st
episode: 552 score: 1.0 memory length: 102582	epsilon: 0.7921356400045125	st
eps: 170 lr: 4e-05 evaluation reward: 1.56	epsilon: 0.791743600004521	ste
episode: 553 score: 0.0 memory length: 102706	epsilon: 0.7913515600045296	st
eps: 124 lr: 4e-05 evaluation reward: 1.56	epsilon: 0.7909516000045382	st
episode: 554 score: 0.0 memory length: 102829	epsilon: 0.7907080600045435	st
eps: 123 lr: 4e-05 evaluation reward: 1.56	epsilon: 0.7904645200045488	st
episode: 555 score: 4.0 memory length: 103125		
eps: 296 lr: 4e-05 evaluation reward: 1.58		
episode: 556 score: 4.0 memory length: 103385		
eps: 260 lr: 4e-05 evaluation reward: 1.59		
episode: 557 score: 4.0 memory length: 103644		
eps: 259 lr: 4e-05 evaluation reward: 1.63		
episode: 558 score: 2.0 memory length: 103865		
ps: 221 lr: 4e-05 evaluation reward: 1.65		
episode: 559 score: 2.0 memory length: 104085		
eps: 220 lr: 4e-05 evaluation reward: 1.65		
episode: 560 score: 3.0 memory length: 104331		
ps: 246 lr: 4e-05 evaluation reward: 1.65		
episode: 561 score: 3.0 memory length: 104578		
eps: 247 lr: 4e-05 evaluation reward: 1.66		
episode: 562 score: 0.0 memory length: 104702		
eps: 124 lr: 4e-05 evaluation reward: 1.66		
episode: 563 score: 3.0 memory length: 104951		
eps: 249 lr: 4e-05 evaluation reward: 1.66		
episode: 564 score: 3.0 memory length: 105181		
eps: 230 lr: 4e-05 evaluation reward: 1.69		
episode: 565 score: 2.0 memory length: 105379		
ps: 198 lr: 4e-05 evaluation reward: 1.71		
episode: 566 score: 2.0 memory length: 105577		
eps: 198 lr: 4e-05 evaluation reward: 1.68		
episode: 567 score: 2.0 memory length: 105779		
eps: 202 lr: 4e-05 evaluation reward: 1.67		
episode: 568 score: 0.0 memory length: 105902		
eps: 123 lr: 4e-05 evaluation reward: 1.66		
episode: 569 score: 0.0 memory length: 106025		
eps: 123 lr: 4e-05 evaluation reward: 1.62		

episode: 570	score: 2.0	memory length: 106224	epsilon: 0.7900705000045574	st
eps: 199	lr: 4e-05	evaluation reward: 1.64	epsilon: 0.789579460004568	ste
episode: 571	score: 3.0	memory length: 106472	epsilon: 0.7888924000045829	st
ps: 248	lr: 4e-05	evaluation reward: 1.67	epsilon: 0.7886468800045883	st
episode: 572	score: 5.0	memory length: 106819	epsilon: 0.7883102800045956	st
eps: 347	lr: 4e-05	evaluation reward: 1.7	epsilon: 0.7880667400046009	st
episode: 573	score: 0.0	memory length: 106943	epsilon: 0.7878232000046062	st
eps: 124	lr: 4e-05	evaluation reward: 1.68	epsilon: 0.7873757200046159	st
episode: 574	score: 1.0	memory length: 107113	epsilon: 0.7871302000046212	st
eps: 170	lr: 4e-05	evaluation reward: 1.67	epsilon: 0.7866827200046309	st
episode: 575	score: 0.0	memory length: 107236	epsilon: 0.7863837400046374	st
eps: 123	lr: 4e-05	evaluation reward: 1.67	epsilon: 0.7859917000046459	st
episode: 576	score: 0.0	memory length: 107359	epsilon: 0.7855541200046554	st
eps: 123	lr: 4e-05	evaluation reward: 1.65	epsilon: 0.7852531600046619	st
episode: 577	score: 3.0	memory length: 107585	epsilon: 0.7846571800046749	st
eps: 226	lr: 4e-05	evaluation reward: 1.67	epsilon: 0.7842077200046846	st
episode: 578	score: 0.0	memory length: 107709	epsilon: 0.7838156800046931	st
eps: 124	lr: 4e-05	evaluation reward: 1.67	epsilon: 0.7833583000047031	st
episode: 579	score: 3.0	memory length: 107935	epsilon: 0.7827682600047159	st
eps: 226	lr: 4e-05	evaluation reward: 1.68	epsilon: 0.7823405800047252	st
episode: 580	score: 1.0	memory length: 108086	epsilon: 0.7818495400047358	st
eps: 151	lr: 4e-05	evaluation reward: 1.69	epsilon: 0.7814020600047455	st
episode: 581	score: 2.0	memory length: 108284	epsilon: 0.7809110200047562	st
eps: 198	lr: 4e-05	evaluation reward: 1.7	epsilon: 0.7806674800047615	st
episode: 582	score: 2.0	memory length: 108505	epsilon: 0.7804239400047668	st
eps: 221	lr: 4e-05	evaluation reward: 1.7	epsilon: 0.7800616000047746	st
episode: 583	score: 1.0	memory length: 108657	epsilon: 0.7798180600047799	st
eps: 152	lr: 4e-05	evaluation reward: 1.68	epsilon: 0.7793705800047896	st
episode: 584	score: 4.0	memory length: 108958	epsilon: 0.7789369600047991	st
eps: 301	lr: 4e-05	evaluation reward: 1.72	epsilon: 0.7786914400048044	st
episode: 585	score: 3.0	memory length: 109185		
eps: 227	lr: 4e-05	evaluation reward: 1.74		
episode: 586	score: 2.0	memory length: 109383		
eps: 198	lr: 4e-05	evaluation reward: 1.75		
episode: 587	score: 3.0	memory length: 109614		
eps: 231	lr: 4e-05	evaluation reward: 1.78		
episode: 588	score: 4.0	memory length: 109912		
eps: 298	lr: 4e-05	evaluation reward: 1.8		
episode: 589	score: 2.0	memory length: 110128		
eps: 216	lr: 4e-05	evaluation reward: 1.8		
episode: 590	score: 3.0	memory length: 110376		
eps: 248	lr: 4e-05	evaluation reward: 1.79		
episode: 591	score: 3.0	memory length: 110602		
eps: 226	lr: 4e-05	evaluation reward: 1.8		
episode: 592	score: 3.0	memory length: 110850		
eps: 248	lr: 4e-05	evaluation reward: 1.79		
episode: 593	score: 0.0	memory length: 110973		
eps: 123	lr: 4e-05	evaluation reward: 1.76		
episode: 594	score: 0.0	memory length: 111096		
eps: 123	lr: 4e-05	evaluation reward: 1.72		
episode: 595	score: 2.0	memory length: 111279		
eps: 183	lr: 4e-05	evaluation reward: 1.72		
episode: 596	score: 0.0	memory length: 111402		
eps: 123	lr: 4e-05	evaluation reward: 1.72		
episode: 597	score: 3.0	memory length: 111628		
eps: 226	lr: 4e-05	evaluation reward: 1.72		
episode: 598	score: 2.0	memory length: 111847		
eps: 219	lr: 4e-05	evaluation reward: 1.72		
episode: 599	score: 0.0	memory length: 111971		
eps: 124	lr: 4e-05	evaluation reward: 1.71		

episode: 600	score: 0.0	memory length: 112095	epsilon: 0.7784459200048097	st
eps: 124	lr: 4e-05	evaluation reward: 1.69	epsilon: 0.7777628200048246	st
episode: 601	score: 5.0	memory length: 112440	epsilon: 0.777325240004834	ste
eps: 345	lr: 4e-05	evaluation reward: 1.72	epsilon: 0.7770817000048393	st
episode: 602	score: 2.0	memory length: 112661	epsilon: 0.7766322400048491	st
ps: 221	lr: 4e-05	evaluation reward: 1.72	epsilon: 0.7763312800048556	st
episode: 603	score: 0.0	memory length: 112784	epsilon: 0.776085760004861	ste
eps: 123	lr: 4e-05	evaluation reward: 1.71	epsilon: 0.7758402400048663	st
episode: 604	score: 3.0	memory length: 113011	epsilon: 0.7755967000048716	st
eps: 227	lr: 4e-05	evaluation reward: 1.71	epsilon: 0.7753511800048769	st
episode: 605	score: 1.0	memory length: 113163	epsilon: 0.7750502200048834	st
eps: 152	lr: 4e-05	evaluation reward: 1.71	epsilon: 0.7748047000048888	st
episode: 606	score: 0.0	memory length: 113287	epsilon: 0.7745591800048941	st
ps: 124	lr: 4e-05	evaluation reward: 1.69	epsilon: 0.7743156400048994	st
episode: 607	score: 0.0	memory length: 113411	epsilon: 0.773733520004912	ste
eps: 124	lr: 4e-05	evaluation reward: 1.69	epsilon: 0.7734880000049174	st
episode: 608	score: 0.0	memory length: 113534	epsilon: 0.7731514000049247	st
eps: 123	lr: 4e-05	evaluation reward: 1.66	epsilon: 0.7726207600049362	st
episode: 609	score: 0.0	memory length: 113658	epsilon: 0.7721871400049456	st
eps: 124	lr: 4e-05	evaluation reward: 1.65	epsilon: 0.7719416200049509	st
episode: 610	score: 1.0	memory length: 113810	epsilon: 0.7716961000049563	st
eps: 152	lr: 4e-05	evaluation reward: 1.64	epsilon: 0.7712387200049662	st
episode: 611	score: 0.0	memory length: 113934	epsilon: 0.7707912400049759	st
eps: 124	lr: 4e-05	evaluation reward: 1.64	epsilon: 0.7703437600049856	st
episode: 612	score: 0.0	memory length: 114058	epsilon: 0.7699517200049941	st
eps: 124	lr: 4e-05	evaluation reward: 1.59	epsilon: 0.7694606800050048	st
episode: 613	score: 0.0	memory length: 114181	epsilon: 0.7690666600050133	st
eps: 123	lr: 4e-05	evaluation reward: 1.57	epsilon: 0.76857760005024	ste
episode: 614	score: 4.0	memory length: 114475	epsilon: 0.7681420000050334	st
ps: 294	lr: 4e-05	evaluation reward: 1.61	epsilon: 0.7678073800050407	st
episode: 615	score: 0.0	memory length: 114599		
eps: 124	lr: 4e-05	evaluation reward: 1.6		
episode: 616	score: 1.0	memory length: 114769		
eps: 170	lr: 4e-05	evaluation reward: 1.61		
episode: 617	score: 3.0	memory length: 115037		
eps: 268	lr: 4e-05	evaluation reward: 1.61		
episode: 618	score: 2.0	memory length: 115256		
eps: 219	lr: 4e-05	evaluation reward: 1.6		
episode: 619	score: 0.0	memory length: 115380		
eps: 124	lr: 4e-05	evaluation reward: 1.6		
episode: 620	score: 0.0	memory length: 115504		
eps: 124	lr: 4e-05	evaluation reward: 1.6		
episode: 621	score: 3.0	memory length: 115735		
eps: 231	lr: 4e-05	evaluation reward: 1.61		
episode: 622	score: 3.0	memory length: 115961		
eps: 226	lr: 4e-05	evaluation reward: 1.64		
episode: 623	score: 3.0	memory length: 116187		
eps: 226	lr: 4e-05	evaluation reward: 1.67		
episode: 624	score: 2.0	memory length: 116385		
eps: 198	lr: 4e-05	evaluation reward: 1.69		
episode: 625	score: 3.0	memory length: 116633		
eps: 248	lr: 4e-05	evaluation reward: 1.7		
episode: 626	score: 2.0	memory length: 116832		
eps: 199	lr: 4e-05	evaluation reward: 1.69		
episode: 627	score: 3.0	memory length: 117079		
ps: 247	lr: 4e-05	evaluation reward: 1.72		
episode: 628	score: 2.0	memory length: 117299		
eps: 220	lr: 4e-05	evaluation reward: 1.71		
episode: 629	score: 1.0	memory length: 117468		
eps: 169	lr: 4e-05	evaluation reward: 1.7		

episode: 630	score: 2.0	memory length: 117666	epsilon: 0.7674153400050492	st
eps: 198	lr: 4e-05	evaluation reward: 1.71	epsilon: 0.7670213200050577	st
episode: 631	score: 2.0	memory length: 117865	epsilon: 0.7667758000050631	st
eps: 199	lr: 4e-05	evaluation reward: 1.72	epsilon: 0.7663837600050716	st
episode: 632	score: 0.0	memory length: 117989	epsilon: 0.7661382400050769	st
eps: 124	lr: 4e-05	evaluation reward: 1.7	epsilon: 0.7658947000050822	st
episode: 633	score: 2.0	memory length: 118187	epsilon: 0.7655581000050895	st
eps: 198	lr: 4e-05	evaluation reward: 1.68	epsilon: 0.7649641000051024	st
episode: 634	score: 0.0	memory length: 118311	epsilon: 0.7643364400051116	ste
eps: 124	lr: 4e-05	evaluation reward: 1.68	epsilon: 0.7639444000051245	st
episode: 635	score: 0.0	memory length: 118434	epsilon: 0.763415740005136	ste
eps: 123	lr: 4e-05	evaluation reward: 1.68	epsilon: 0.7630573600051438	st
episode: 636	score: 1.0	memory length: 118604	epsilon: 0.7624693000051566	st
eps: 170	lr: 4e-05	evaluation reward: 1.65	epsilon: 0.762128740005164	ste
episode: 637	score: 4.0	memory length: 118904	epsilon: 0.7616792800051737	st
eps: 300	lr: 4e-05	evaluation reward: 1.67	epsilon: 0.7612318000051834	st
episode: 638	score: 4.0	memory length: 119221	epsilon: 0.7609862800051888	st
ps: 317	lr: 4e-05	evaluation reward: 1.69	epsilon: 0.7607407600051941	st
episode: 639	score: 2.0	memory length: 119419	epsilon: 0.7602913000052038	st
eps: 198	lr: 4e-05	evaluation reward: 1.71	epsilon: 0.7600477600052091	st
episode: 640	score: 3.0	memory length: 119686	epsilon: 0.7598022400052145	st
ps: 267	lr: 4e-05	evaluation reward: 1.72	epsilon: 0.759408220005223	ste
episode: 641	score: 2.0	memory length: 119867	epsilon: 0.7591646800052283	st
eps: 181	lr: 4e-05	evaluation reward: 1.74	epsilon: 0.758669680005239	ste
episode: 642	score: 5.0	memory length: 120164	epsilon: 0.7584241600052444	st
eps: 297	lr: 4e-05	evaluation reward: 1.76	epsilon: 0.7581251800052509	st
episode: 643	score: 1.0	memory length: 120336	epsilon: 0.7576777000052606	st
ps: 172	lr: 4e-05	evaluation reward: 1.77	epsilon: 0.7573411000052679	st
episode: 644	score: 3.0	memory length: 120563	epsilon: 0.7570955800052732	st
eps: 227	lr: 4e-05	evaluation reward: 1.8	epsilon: 0.7567966000052797	st
episode: 645	score: 3.0	memory length: 120789		
eps: 226	lr: 4e-05	evaluation reward: 1.8		
episode: 646	score: 0.0	memory length: 120913		
eps: 124	lr: 4e-05	evaluation reward: 1.78		
episode: 647	score: 0.0	memory length: 121037		
eps: 124	lr: 4e-05	evaluation reward: 1.78		
episode: 648	score: 3.0	memory length: 121264		
eps: 227	lr: 4e-05	evaluation reward: 1.81		
episode: 649	score: 0.0	memory length: 121387		
eps: 123	lr: 4e-05	evaluation reward: 1.79		
episode: 650	score: 0.0	memory length: 121511		
eps: 124	lr: 4e-05	evaluation reward: 1.77		
episode: 651	score: 2.0	memory length: 121710		
ps: 199	lr: 4e-05	evaluation reward: 1.75		
episode: 652	score: 0.0	memory length: 121833		
eps: 123	lr: 4e-05	evaluation reward: 1.74		
episode: 653	score: 3.0	memory length: 122083		
ps: 250	lr: 4e-05	evaluation reward: 1.77		
episode: 654	score: 0.0	memory length: 122207		
eps: 124	lr: 4e-05	evaluation reward: 1.77		
episode: 655	score: 1.0	memory length: 122358		
eps: 151	lr: 4e-05	evaluation reward: 1.74		
episode: 656	score: 3.0	memory length: 122584		
eps: 226	lr: 4e-05	evaluation reward: 1.73		
episode: 657	score: 1.0	memory length: 122754		
eps: 170	lr: 4e-05	evaluation reward: 1.7		
episode: 658	score: 0.0	memory length: 122878		
eps: 124	lr: 4e-05	evaluation reward: 1.68		
episode: 659	score: 1.0	memory length: 123029		
eps: 151	lr: 4e-05	evaluation reward: 1.67		

episode: 660	score: 5.0	memory length: 123374	epsilon: 0.7561135000052945	st
eps: 345	lr: 4e-05	evaluation reward: 1.69	epsilon: 0.7556600800053044	st
episode: 661	score: 3.0	memory length: 123603	epsilon: 0.7550165800053183	st
eps: 229	lr: 4e-05	evaluation reward: 1.69	epsilon: 0.7546225600053269	st
episode: 662	score: 5.0	memory length: 123928	epsilon: 0.7540107400053402	st
eps: 325	lr: 4e-05	evaluation reward: 1.74	epsilon: 0.7535791000053496	st
episode: 663	score: 2.0	memory length: 124127	epsilon: 0.7530880600053602	st
eps: 199	lr: 4e-05	evaluation reward: 1.73	epsilon: 0.7526603800053695	st
episode: 664	score: 5.0	memory length: 124436	epsilon: 0.7520347000053831	st
eps: 309	lr: 4e-05	evaluation reward: 1.75	epsilon: 0.7513892200053971	st
episode: 665	score: 2.0	memory length: 124654	epsilon: 0.7506150400054139	st
eps: 218	lr: 4e-05	evaluation reward: 1.75	epsilon: 0.7500448000054263	st
episode: 666	score: 3.0	memory length: 124902	epsilon: 0.7495616800054368	st
eps: 248	lr: 4e-05	evaluation reward: 1.76	epsilon: 0.7491617200054455	st
episode: 667	score: 2.0	memory length: 125118	epsilon: 0.7487281000054549	st
eps: 216	lr: 4e-05	evaluation reward: 1.76	epsilon: 0.7481875600054666	st
episode: 668	score: 4.0	memory length: 125434	epsilon: 0.7477361200054764	st
eps: 316	lr: 4e-05	evaluation reward: 1.8	epsilon: 0.7472074600054879	st
episode: 669	score: 5.0	memory length: 125760	epsilon: 0.7467778000054972	st
eps: 326	lr: 4e-05	evaluation reward: 1.85	epsilon: 0.7463818000055058	st
episode: 670	score: 6.0	memory length: 126151	epsilon: 0.7461362800055111	st
eps: 391	lr: 4e-05	evaluation reward: 1.89	epsilon: 0.7457442400055196	st
episode: 671	score: 4.0	memory length: 126439	epsilon: 0.7453482400055282	st
eps: 288	lr: 4e-05	evaluation reward: 1.9	epsilon: 0.7449542200055368	st
episode: 672	score: 3.0	memory length: 126683	epsilon: 0.7445225800055462	st
eps: 244	lr: 4e-05	evaluation reward: 1.88	epsilon: 0.7441879600055534	st
episode: 673	score: 2.0	memory length: 126885	epsilon: 0.7434375400055697	st
eps: 202	lr: 4e-05	evaluation reward: 1.9	epsilon: 0.7430455000055782	st
episode: 674	score: 2.0	memory length: 127104	epsilon: 0.7425980200055879	st
eps: 219	lr: 4e-05	evaluation reward: 1.91		
episode: 675	score: 4.0	memory length: 127377		
eps: 273	lr: 4e-05	evaluation reward: 1.95		
episode: 676	score: 3.0	memory length: 127605		
eps: 228	lr: 4e-05	evaluation reward: 1.98		
episode: 677	score: 3.0	memory length: 127872		
eps: 267	lr: 4e-05	evaluation reward: 1.98		
episode: 678	score: 2.0	memory length: 128089		
eps: 217	lr: 4e-05	evaluation reward: 2.0		
episode: 679	score: 2.0	memory length: 128289		
eps: 200	lr: 4e-05	evaluation reward: 1.99		
episode: 680	score: 0.0	memory length: 128413		
eps: 124	lr: 4e-05	evaluation reward: 1.98		
episode: 681	score: 2.0	memory length: 128611		
eps: 198	lr: 4e-05	evaluation reward: 1.98		
episode: 682	score: 2.0	memory length: 128811		
eps: 200	lr: 4e-05	evaluation reward: 1.98		
episode: 683	score: 2.0	memory length: 129010		
eps: 199	lr: 4e-05	evaluation reward: 1.99		
episode: 684	score: 2.0	memory length: 129228		
eps: 218	lr: 4e-05	evaluation reward: 1.97		
episode: 685	score: 1.0	memory length: 129397		
eps: 169	lr: 4e-05	evaluation reward: 1.95		
episode: 686	score: 3.0	memory length: 129624		
eps: 227	lr: 4e-05	evaluation reward: 1.96		
episode: 687	score: 1.0	memory length: 129776		
eps: 152	lr: 4e-05	evaluation reward: 1.94		
episode: 688	score: 2.0	memory length: 129974		
eps: 198	lr: 4e-05	evaluation reward: 1.92		
episode: 689	score: 3.0	memory length: 130200		
eps: 226	lr: 4e-05	evaluation reward: 1.93		

episode: 690 score: 2.0 memory length: 130398	epsilon: 0.7422059800055965	st
eps: 198 lr: 4e-05 evaluation reward: 1.92	epsilon: 0.741811960005605	ste
episode: 691 score: 2.0 memory length: 130597	epsilon: 0.7413625000056148	st
ps: 199 lr: 4e-05 evaluation reward: 1.91	epsilon: 0.7408140400056267	st
episode: 692 score: 3.0 memory length: 130824	epsilon: 0.740477440005634	ste
eps: 227 lr: 4e-05 evaluation reward: 1.91	epsilon: 0.7399903600056446	st
episode: 693 score: 4.0 memory length: 131101	epsilon: 0.7396537600056519	st
eps: 277 lr: 4e-05 evaluation reward: 1.95	epsilon: 0.7392617200056604	st
episode: 694 score: 1.0 memory length: 131271	epsilon: 0.7386221800056743	st
ps: 170 lr: 4e-05 evaluation reward: 1.96	epsilon: 0.7382301400056828	st
episode: 695 score: 3.0 memory length: 131517	epsilon: 0.7378361200056913	st
eps: 246 lr: 4e-05 evaluation reward: 1.97	epsilon: 0.7375906000056967	st
episode: 696 score: 1.0 memory length: 131687	epsilon: 0.7371431200057064	st
eps: 170 lr: 4e-05 evaluation reward: 1.98	epsilon: 0.736654060005717	ste
episode: 697 score: 2.0 memory length: 131885	epsilon: 0.7361056000057289	st
eps: 198 lr: 4e-05 evaluation reward: 1.97	epsilon: 0.7358620600057342	st
episode: 698 score: 0.0 memory length: 132009	epsilon: 0.7353730000057448	st
eps: 124 lr: 4e-05 evaluation reward: 1.95	epsilon: 0.7349750200057534	st
episode: 699 score: 2.0 memory length: 132208	epsilon: 0.7345829800057619	st
eps: 199 lr: 4e-05 evaluation reward: 1.97	epsilon: 0.7342463800057693	st
episode: 700 score: 2.0 memory length: 132406	epsilon: 0.7340008600057746	st
eps: 198 lr: 4e-05 evaluation reward: 1.99	epsilon: 0.7337553400057799	st
episode: 701 score: 2.0 memory length: 132605	epsilon: 0.7331672800057927	st
eps: 199 lr: 4e-05 evaluation reward: 1.96	epsilon: 0.7327752400058012	st
episode: 702 score: 0.0 memory length: 132729	epsilon: 0.7325317000058065	st
eps: 124 lr: 4e-05 evaluation reward: 1.94	epsilon: 0.732232720005813	ste
episode: 703 score: 3.0 memory length: 132955	epsilon: 0.7318406800058215	st
eps: 226 lr: 4e-05 evaluation reward: 1.97	epsilon: 0.7314446800058301	st
episode: 704 score: 3.0 memory length: 133202	epsilon: 0.7307992000058441	st
ps: 247 lr: 4e-05 evaluation reward: 1.97		
episode: 705 score: 4.0 memory length: 133479		
eps: 277 lr: 4e-05 evaluation reward: 2.0		
episode: 706 score: 0.0 memory length: 133602		
eps: 123 lr: 4e-05 evaluation reward: 2.0		
episode: 707 score: 3.0 memory length: 133849		
eps: 247 lr: 4e-05 evaluation reward: 2.03		
episode: 708 score: 2.0 memory length: 134050		
eps: 201 lr: 4e-05 evaluation reward: 2.05		
episode: 709 score: 2.0 memory length: 134248		
eps: 198 lr: 4e-05 evaluation reward: 2.07		
episode: 710 score: 1.0 memory length: 134418		
eps: 170 lr: 4e-05 evaluation reward: 2.07		
episode: 711 score: 0.0 memory length: 134542		
eps: 124 lr: 4e-05 evaluation reward: 2.07		
episode: 712 score: 0.0 memory length: 134666		
eps: 124 lr: 4e-05 evaluation reward: 2.07		
episode: 713 score: 4.0 memory length: 134963		
eps: 297 lr: 4e-05 evaluation reward: 2.11		
episode: 714 score: 2.0 memory length: 135161		
eps: 198 lr: 4e-05 evaluation reward: 2.09		
episode: 715 score: 0.0 memory length: 135284		
eps: 123 lr: 4e-05 evaluation reward: 2.09		
episode: 716 score: 1.0 memory length: 135435		
ps: 151 lr: 4e-05 evaluation reward: 2.09		
episode: 717 score: 2.0 memory length: 135633		
eps: 198 lr: 4e-05 evaluation reward: 2.08		
episode: 718 score: 2.0 memory length: 135833		
eps: 200 lr: 4e-05 evaluation reward: 2.08		
episode: 719 score: 5.0 memory length: 136159		
eps: 326 lr: 4e-05 evaluation reward: 2.13		

episode: 720 score: 0.0 memory length: 136282	epsilon: 0.7305556600058494	st
eps: 123 lr: 4e-05 evaluation reward: 2.13	epsilon: 0.7301616400058579	st
episode: 721 score: 2.0 memory length: 136481	epsilon: 0.7299181000058632	st
eps: 199 lr: 4e-05 evaluation reward: 2.12	epsilon: 0.7296171400058697	st
episode: 722 score: 0.0 memory length: 136604	epsilon: 0.7291280800058804	st
eps: 123 lr: 4e-05 evaluation reward: 2.09	epsilon: 0.7288271200058869	st
episode: 723 score: 1.0 memory length: 136756	epsilon: 0.7285816000058922	st
eps: 152 lr: 4e-05 evaluation reward: 2.07	epsilon: 0.7281875800059008	st
episode: 724 score: 3.0 memory length: 137003	epsilon: 0.7276371400059127	st
eps: 247 lr: 4e-05 evaluation reward: 2.08	epsilon: 0.7270312600059259	st
episode: 725 score: 1.0 memory length: 137155	epsilon: 0.7266392200059344	st
eps: 152 lr: 4e-05 evaluation reward: 2.06	epsilon: 0.7263382600059409	st
episode: 726 score: 0.0 memory length: 137279	epsilon: 0.7259462200059494	st
eps: 124 lr: 4e-05 evaluation reward: 2.04	epsilon: 0.725552200005958	ste
episode: 727 score: 2.0 memory length: 137478	epsilon: 0.7253086600059633	st
eps: 199 lr: 4e-05 evaluation reward: 2.03	epsilon: 0.7245919000059788	st
episode: 728 score: 4.0 memory length: 137756	epsilon: 0.7241087800059893	st
eps: 278 lr: 4e-05 evaluation reward: 2.05	epsilon: 0.7232712400060075	st
episode: 729 score: 5.0 memory length: 138062	epsilon: 0.722972260006014	ste
eps: 306 lr: 4e-05 evaluation reward: 2.09	epsilon: 0.7225782400060226	st
episode: 730 score: 2.0 memory length: 138260	epsilon: 0.7223347000060278	st
eps: 198 lr: 4e-05 evaluation reward: 2.09	epsilon: 0.7218080200060393	st
episode: 731 score: 1.0 memory length: 138412	epsilon: 0.7213565800060491	st
eps: 152 lr: 4e-05 evaluation reward: 2.08	epsilon: 0.7208695000060596	st
episode: 732 score: 2.0 memory length: 138610	epsilon: 0.720623980006065	ste
eps: 198 lr: 4e-05 evaluation reward: 2.1	epsilon: 0.7203804400060703	st
episode: 733 score: 2.0 memory length: 138809	epsilon: 0.7198953400060808	st
ps: 199 lr: 4e-05 evaluation reward: 2.1	epsilon: 0.7194973600060894	st
episode: 734 score: 0.0 memory length: 138932	epsilon: 0.7190637400060989	st
eps: 123 lr: 4e-05 evaluation reward: 2.1	epsilon: 0.7185390400061102	st
episode: 735 score: 6.0 memory length: 139294		
eps: 362 lr: 4e-05 evaluation reward: 2.16		
episode: 736 score: 3.0 memory length: 139538		
eps: 244 lr: 4e-05 evaluation reward: 2.18		
episode: 737 score: 7.0 memory length: 139961		
eps: 423 lr: 4e-05 evaluation reward: 2.21		
episode: 738 score: 1.0 memory length: 140112		
ps: 151 lr: 4e-05 evaluation reward: 2.18		
episode: 739 score: 2.0 memory length: 140311		
eps: 199 lr: 4e-05 evaluation reward: 2.18		
episode: 740 score: 0.0 memory length: 140434		
eps: 123 lr: 4e-05 evaluation reward: 2.15		
episode: 741 score: 3.0 memory length: 140700		
eps: 266 lr: 4e-05 evaluation reward: 2.16		
episode: 742 score: 3.0 memory length: 140928		
eps: 228 lr: 4e-05 evaluation reward: 2.14		
episode: 743 score: 3.0 memory length: 141174		
eps: 246 lr: 4e-05 evaluation reward: 2.16		
episode: 744 score: 0.0 memory length: 141298		
ps: 124 lr: 4e-05 evaluation reward: 2.13		
episode: 745 score: 0.0 memory length: 141421		
eps: 123 lr: 4e-05 evaluation reward: 2.1		
episode: 746 score: 3.0 memory length: 141666		
eps: 245 lr: 4e-05 evaluation reward: 2.13		
episode: 747 score: 2.0 memory length: 141867		
eps: 201 lr: 4e-05 evaluation reward: 2.15		
episode: 748 score: 2.0 memory length: 142086		
eps: 219 lr: 4e-05 evaluation reward: 2.14		
episode: 749 score: 3.0 memory length: 142351		
eps: 265 lr: 4e-05 evaluation reward: 2.17		

episode: 750	score: 0.0	memory length: 142475	epsilon: 0.7182935200061156	st
eps: 124	lr: 4e-05	evaluation reward: 2.17	epsilon: 0.7178044600061262	st
episode: 751	score: 3.0	memory length: 142722	epsilon: 0.7175609200061315	st
eps: 247	lr: 4e-05	evaluation reward: 2.18	epsilon: 0.7169609800061445	st
episode: 752	score: 0.0	memory length: 142845	epsilon: 0.7164224200061562	st
eps: 123	lr: 4e-05	evaluation reward: 2.18	epsilon: 0.7159650400061661	st
episode: 753	score: 5.0	memory length: 143148	epsilon: 0.7155314200061755	st
eps: 303	lr: 4e-05	evaluation reward: 2.2	epsilon: 0.7152859000061809	st
episode: 754	score: 3.0	memory length: 143420	epsilon: 0.7148918800061894	st
eps: 272	lr: 4e-05	evaluation reward: 2.23	epsilon: 0.7144998400061979	st
episode: 755	score: 3.0	memory length: 143651	epsilon: 0.7142563000062032	st
eps: 231	lr: 4e-05	evaluation reward: 2.25	epsilon: 0.7138642600062117	st
episode: 756	score: 2.0	memory length: 143870	epsilon: 0.7133732200062224	st
eps: 219	lr: 4e-05	evaluation reward: 2.24	epsilon: 0.7129811800062309	st
episode: 757	score: 0.0	memory length: 143994	epsilon: 0.7124960800062414	st
eps: 124	lr: 4e-05	evaluation reward: 2.23	epsilon: 0.7119020800062543	st
episode: 758	score: 2.0	memory length: 144193	epsilon: 0.711316000006267	ste
eps: 199	lr: 4e-05	evaluation reward: 2.25	epsilon: 0.7107873400062785	st
episode: 759	score: 2.0	memory length: 144391	epsilon: 0.7103933200062871	st
eps: 198	lr: 4e-05	evaluation reward: 2.26	epsilon: 0.7099022800062977	st
episode: 760	score: 0.0	memory length: 144514	epsilon: 0.709565680006305	ste
eps: 123	lr: 4e-05	evaluation reward: 2.21	epsilon: 0.7093201600063104	st
episode: 761	score: 2.0	memory length: 144712	epsilon: 0.7088707000063201	st
eps: 198	lr: 4e-05	evaluation reward: 2.2	epsilon: 0.7082628400063333	st
episode: 762	score: 3.0	memory length: 144960	epsilon: 0.7077282400063449	st
eps: 248	lr: 4e-05	evaluation reward: 2.18	epsilon: 0.7074847000063502	st
episode: 763	score: 2.0	memory length: 145158	epsilon: 0.7072391800063556	st
eps: 198	lr: 4e-05	evaluation reward: 2.18	epsilon: 0.7068412000063642	st
episode: 764	score: 3.0	memory length: 145403	epsilon: 0.7064471800063727	st
eps: 245	lr: 4e-05	evaluation reward: 2.16	epsilon: 0.706203640006378	ste
episode: 765	score: 5.0	memory length: 145703	epsilon: 0.706203640006378	ste
eps: 300	lr: 4e-05	evaluation reward: 2.19		
episode: 766	score: 4.0	memory length: 145999		
ps: 296	lr: 4e-05	evaluation reward: 2.2		
episode: 767	score: 3.0	memory length: 146266		
eps: 267	lr: 4e-05	evaluation reward: 2.21		
episode: 768	score: 2.0	memory length: 146465		
eps: 199	lr: 4e-05	evaluation reward: 2.19		
episode: 769	score: 3.0	memory length: 146713		
eps: 248	lr: 4e-05	evaluation reward: 2.17		
episode: 770	score: 1.0	memory length: 146883		
ps: 170	lr: 4e-05	evaluation reward: 2.12		
episode: 771	score: 0.0	memory length: 147007		
eps: 124	lr: 4e-05	evaluation reward: 2.08		
episode: 772	score: 3.0	memory length: 147234		
eps: 227	lr: 4e-05	evaluation reward: 2.08		
episode: 773	score: 5.0	memory length: 147541		
eps: 307	lr: 4e-05	evaluation reward: 2.11		
episode: 774	score: 3.0	memory length: 147811		
eps: 270	lr: 4e-05	evaluation reward: 2.12		
episode: 775	score: 0.0	memory length: 147934		
eps: 123	lr: 4e-05	evaluation reward: 2.08		
episode: 776	score: 0.0	memory length: 148058		
eps: 124	lr: 4e-05	evaluation reward: 2.05		
episode: 777	score: 2.0	memory length: 148259		
eps: 201	lr: 4e-05	evaluation reward: 2.04		
episode: 778	score: 2.0	memory length: 148458		
eps: 199	lr: 4e-05	evaluation reward: 2.04		
episode: 779	score: 0.0	memory length: 148581		
ps: 123	lr: 4e-05	evaluation reward: 2.02		

episode: 780	score: 2.0	memory length: 148781	epsilon: 0.7058076400063866	st
eps: 200	lr: 4e-05	evaluation reward: 2.04	epsilon: 0.7052077000063997	st
episode: 781	score: 5.0	memory length: 149084	epsilon: 0.704871100006407	ste
eps: 303	lr: 4e-05	evaluation reward: 2.07	epsilon: 0.704318680006419	ste
episode: 782	score: 1.0	memory length: 149254	epsilon: 0.7039266400064275	st
ps: 170	lr: 4e-05	evaluation reward: 2.06	epsilon: 0.7035880600064348	st
episode: 783	score: 4.0	memory length: 149533	epsilon: 0.7031960200064433	st
ps: 279	lr: 4e-05	evaluation reward: 2.08	epsilon: 0.7029524800064486	st
episode: 784	score: 2.0	memory length: 149731	epsilon: 0.7024634200064592	st
eps: 198	lr: 4e-05	evaluation reward: 2.08	epsilon: 0.7021624600064658	st
episode: 785	score: 1.0	memory length: 149902	epsilon: 0.7017664600064744	st
eps: 171	lr: 4e-05	evaluation reward: 2.08	epsilon: 0.7015229200064796	st
episode: 786	score: 2.0	memory length: 150100	epsilon: 0.7011289000064882	st
eps: 198	lr: 4e-05	evaluation reward: 2.07	epsilon: 0.7008833800064935	st
episode: 787	score: 0.0	memory length: 150223	epsilon: 0.7002953200065063	st
eps: 123	lr: 4e-05	evaluation reward: 2.06	epsilon: 0.6997448800065182	st
episode: 788	score: 3.0	memory length: 150470	epsilon: 0.6994459000065247	st
eps: 247	lr: 4e-05	evaluation reward: 2.07	epsilon: 0.69920236000653	step
episode: 789	score: 1.0	memory length: 150622	epsilon: 0.6987529000065398	st
eps: 152	lr: 4e-05	evaluation reward: 2.05	epsilon: 0.6982618600065504	st
episode: 790	score: 2.0	memory length: 150822	epsilon: 0.6978282400065599	st
eps: 200	lr: 4e-05	evaluation reward: 2.05	epsilon: 0.6973411600065704	st
episode: 791	score: 0.0	memory length: 150945	epsilon: 0.6970956400065758	st
eps: 123	lr: 4e-05	evaluation reward: 2.03	epsilon: 0.6966046000065864	st
episode: 792	score: 2.0	memory length: 151144	epsilon: 0.696210580006595	ste
eps: 199	lr: 4e-05	evaluation reward: 2.02	epsilon: 0.6958165600066035	st
episode: 793	score: 0.0	memory length: 151268	epsilon: 0.6953591800066135	st
eps: 124	lr: 4e-05	evaluation reward: 1.98	epsilon: 0.694965160006622	ste
episode: 794	score: 4.0	memory length: 151565	epsilon: 0.6945157000066318	st
eps: 297	lr: 4e-05	evaluation reward: 2.01	epsilon: 0.6939335800066444	st
episode: 795	score: 4.0	memory length: 151843		
eps: 278	lr: 4e-05	evaluation reward: 2.02		
episode: 796	score: 1.0	memory length: 151994		
eps: 151	lr: 4e-05	evaluation reward: 2.02		
episode: 797	score: 0.0	memory length: 152117		
s: 123	lr: 4e-05	evaluation reward: 2.0		
episode: 798	score: 3.0	memory length: 152344		
eps: 227	lr: 4e-05	evaluation reward: 2.03		
episode: 799	score: 3.0	memory length: 152592		
eps: 248	lr: 4e-05	evaluation reward: 2.04		
episode: 800	score: 2.0	memory length: 152811		
eps: 219	lr: 4e-05	evaluation reward: 2.04		
episode: 801	score: 3.0	memory length: 153057		
eps: 246	lr: 4e-05	evaluation reward: 2.05		
episode: 802	score: 0.0	memory length: 153181		
eps: 124	lr: 4e-05	evaluation reward: 2.05		
episode: 803	score: 3.0	memory length: 153429		
eps: 248	lr: 4e-05	evaluation reward: 2.05		
episode: 804	score: 2.0	memory length: 153628		
ps: 199	lr: 4e-05	evaluation reward: 2.04		
episode: 805	score: 2.0	memory length: 153827		
eps: 199	lr: 4e-05	evaluation reward: 2.02		
episode: 806	score: 3.0	memory length: 154058		
eps: 231	lr: 4e-05	evaluation reward: 2.05		
episode: 807	score: 2.0	memory length: 154257		
ps: 199	lr: 4e-05	evaluation reward: 2.04		
episode: 808	score: 3.0	memory length: 154484		
eps: 227	lr: 4e-05	evaluation reward: 2.05		
episode: 809	score: 4.0	memory length: 154778		
eps: 294	lr: 4e-05	evaluation reward: 2.07		

episode: 810	score: 2.0	memory length: 154976	epsilon: 0.6935415400066529	st
eps: 198	lr: 4e-05	evaluation reward: 2.08	epsilon: 0.6931039600066624	st
episode: 811	score: 2.0	memory length: 155197	epsilon: 0.6927119200066709	st
eps: 221	lr: 4e-05	evaluation reward: 2.1	epsilon: 0.6923773000066782	st
episode: 812	score: 2.0	memory length: 155395	epsilon: 0.6917357800066921	st
eps: 198	lr: 4e-05	evaluation reward: 2.12	epsilon: 0.6912467200067027	st
episode: 813	score: 1.0	memory length: 155564	epsilon: 0.6907616200067133	st
eps: 169	lr: 4e-05	evaluation reward: 2.09	epsilon: 0.6905161000067186	st
episode: 814	score: 5.0	memory length: 155888	epsilon: 0.6902705800067239	st
eps: 324	lr: 4e-05	evaluation reward: 2.12	epsilon: 0.6898211200067337	st
episode: 815	score: 3.0	memory length: 156135	epsilon: 0.6893063200067449	st
eps: 247	lr: 4e-05	evaluation reward: 2.15	epsilon: 0.6890608000067502	st
episode: 816	score: 3.0	memory length: 156380	epsilon: 0.6887202400067576	st
eps: 245	lr: 4e-05	evaluation reward: 2.17	epsilon: 0.688379680006765	ste
episode: 817	score: 0.0	memory length: 156504	epsilon: 0.6880391200067724	st
eps: 124	lr: 4e-05	evaluation reward: 2.15	epsilon: 0.6874946200067842	st
episode: 818	score: 0.0	memory length: 156628	epsilon: 0.686902600006797	ste
eps: 124	lr: 4e-05	evaluation reward: 2.13	epsilon: 0.6866570800068024	st
episode: 819	score: 3.0	memory length: 156855	epsilon: 0.6862650400068109	st
eps: 227	lr: 4e-05	evaluation reward: 2.11	epsilon: 0.6859284400068182	st
episode: 820	score: 4.0	memory length: 157115	epsilon: 0.6852809800068322	st
eps: 260	lr: 4e-05	evaluation reward: 2.15	epsilon: 0.6847879600068429	st
episode: 821	score: 0.0	memory length: 157239	epsilon: 0.6845444200068482	st
eps: 124	lr: 4e-05	evaluation reward: 2.13	epsilon: 0.684096940006858	ste
episode: 822	score: 1.0	memory length: 157411	epsilon: 0.683405920006873	ste
eps: 172	lr: 4e-05	evaluation reward: 2.14	epsilon: 0.6829762600068823	st
episode: 823	score: 1.0	memory length: 157583	epsilon: 0.6827327200068876	st
ps: 172	lr: 4e-05	evaluation reward: 2.14	epsilon: 0.6824891800068928	st
episode: 824	score: 1.0	memory length: 157755	epsilon: 0.6822456400068981	st
eps: 172	lr: 4e-05	evaluation reward: 2.12	epsilon: 0.6817981600069078	st
episode: 825	score: 4.0	memory length: 158030		
eps: 275	lr: 4e-05	evaluation reward: 2.15		
episode: 826	score: 4.0	memory length: 158329		
ps: 299	lr: 4e-05	evaluation reward: 2.19		
episode: 827	score: 0.0	memory length: 158453		
eps: 124	lr: 4e-05	evaluation reward: 2.17		
episode: 828	score: 2.0	memory length: 158651		
eps: 198	lr: 4e-05	evaluation reward: 2.15		
episode: 829	score: 1.0	memory length: 158821		
eps: 170	lr: 4e-05	evaluation reward: 2.11		
episode: 830	score: 5.0	memory length: 159148		
eps: 327	lr: 4e-05	evaluation reward: 2.14		
episode: 831	score: 3.0	memory length: 159397		
eps: 249	lr: 4e-05	evaluation reward: 2.16		
episode: 832	score: 0.0	memory length: 159520		
eps: 123	lr: 4e-05	evaluation reward: 2.14		
episode: 833	score: 3.0	memory length: 159746		
ps: 226	lr: 4e-05	evaluation reward: 2.15		
episode: 834	score: 8.0	memory length: 160095		
ps: 349	lr: 4e-05	evaluation reward: 2.23		
episode: 835	score: 2.0	memory length: 160312		
eps: 217	lr: 4e-05	evaluation reward: 2.19		
episode: 836	score: 0.0	memory length: 160435		
eps: 123	lr: 4e-05	evaluation reward: 2.16		
episode: 837	score: 0.0	memory length: 160558		
eps: 123	lr: 4e-05	evaluation reward: 2.09		
episode: 838	score: 0.0	memory length: 160681		
eps: 123	lr: 4e-05	evaluation reward: 2.08		
episode: 839	score: 3.0	memory length: 160907		
eps: 226	lr: 4e-05	evaluation reward: 2.09		

episode: 840	score: 9.0	memory length: 161266	epsilon: 0.6810873400069233	st
eps: 359	lr: 4e-05	evaluation reward: 2.18	epsilon: 0.680637880006933	ste
episode: 841	score: 3.0	memory length: 161493	epsilon: 0.6801468400069437	st
ps: 227	lr: 4e-05	evaluation reward: 2.18	epsilon: 0.6796043200069555	st
episode: 842	score: 3.0	memory length: 161741	epsilon: 0.6789647800069694	st
eps: 248	lr: 4e-05	evaluation reward: 2.18	epsilon: 0.6785727400069779	st
episode: 843	score: 4.0	memory length: 162015	epsilon: 0.6781807000069864	st
eps: 274	lr: 4e-05	evaluation reward: 2.19	epsilon: 0.6777312400069961	st
episode: 844	score: 5.0	memory length: 162338	epsilon: 0.677184760007008	ste
eps: 323	lr: 4e-05	evaluation reward: 2.24	epsilon: 0.6765828400070211	st
episode: 845	score: 2.0	memory length: 162536	epsilon: 0.6763393000070264	st
eps: 198	lr: 4e-05	evaluation reward: 2.26	epsilon: 0.6757136200070399	st
episode: 846	score: 2.0	memory length: 162734	epsilon: 0.6751295200070526	st
eps: 198	lr: 4e-05	evaluation reward: 2.25	epsilon: 0.6744800800070667	st
episode: 847	score: 3.0	memory length: 162961	epsilon: 0.6740880400070752	st
eps: 227	lr: 4e-05	evaluation reward: 2.26	epsilon: 0.6737514400070825	st
episode: 848	score: 4.0	memory length: 163237	epsilon: 0.6733574200070911	st
ps: 276	lr: 4e-05	evaluation reward: 2.28	epsilon: 0.6729634000070996	st
episode: 849	score: 5.0	memory length: 163541	epsilon: 0.6725713600071082	st
eps: 304	lr: 4e-05	evaluation reward: 2.3	epsilon: 0.6719912200071207	st
episode: 850	score: 0.0	memory length: 163664	epsilon: 0.6715061200071313	st
eps: 123	lr: 4e-05	evaluation reward: 2.3	epsilon: 0.6711655600071387	st
episode: 851	score: 5.0	memory length: 163980	epsilon: 0.6706804600071492	st
eps: 316	lr: 4e-05	evaluation reward: 2.32	epsilon: 0.6702884200071577	st
episode: 852	score: 5.0	memory length: 164275	epsilon: 0.6697993600071683	st
eps: 295	lr: 4e-05	evaluation reward: 2.37	epsilon: 0.6691855600071817	st
episode: 853	score: 5.0	memory length: 164603	epsilon: 0.6686430400071934	st
eps: 328	lr: 4e-05	evaluation reward: 2.37	epsilon: 0.6681935800072032	st
episode: 854	score: 2.0	memory length: 164801	epsilon: 0.6678015400072117	st
eps: 198	lr: 4e-05	evaluation reward: 2.36	epsilon: 0.6673461400072216	st
episode: 855	score: 1.0	memory length: 164971		
eps: 170	lr: 4e-05	evaluation reward: 2.34		
episode: 856	score: 2.0	memory length: 165170		
eps: 199	lr: 4e-05	evaluation reward: 2.34		
episode: 857	score: 2.0	memory length: 165369		
eps: 199	lr: 4e-05	evaluation reward: 2.36		
episode: 858	score: 2.0	memory length: 165567		
eps: 198	lr: 4e-05	evaluation reward: 2.36		
episode: 859	score: 5.0	memory length: 165860		
eps: 293	lr: 4e-05	evaluation reward: 2.39		
episode: 860	score: 3.0	memory length: 166105		
eps: 245	lr: 4e-05	evaluation reward: 2.42		
episode: 861	score: 1.0	memory length: 166277		
eps: 172	lr: 4e-05	evaluation reward: 2.41		
episode: 862	score: 3.0	memory length: 166522		
eps: 245	lr: 4e-05	evaluation reward: 2.41		
episode: 863	score: 2.0	memory length: 166720		
eps: 198	lr: 4e-05	evaluation reward: 2.41		
episode: 864	score: 3.0	memory length: 166967		
eps: 247	lr: 4e-05	evaluation reward: 2.41		
episode: 865	score: 5.0	memory length: 167277		
eps: 310	lr: 4e-05	evaluation reward: 2.41		
episode: 866	score: 4.0	memory length: 167551		
eps: 274	lr: 4e-05	evaluation reward: 2.41		
episode: 867	score: 3.0	memory length: 167778		
eps: 227	lr: 4e-05	evaluation reward: 2.41		
episode: 868	score: 2.0	memory length: 167976		
eps: 198	lr: 4e-05	evaluation reward: 2.41		
episode: 869	score: 3.0	memory length: 168206		
eps: 230	lr: 4e-05	evaluation reward: 2.41		

episode: 870	score: 4.0	memory length: 168502	epsilon: 0.6667600600072343	st
eps: 296	lr: 4e-05	evaluation reward: 2.44	epsilon: 0.6664254400072416	st
episode: 871	score: 1.0	memory length: 168671	epsilon: 0.6661819000072469	st
eps: 169	lr: 4e-05	evaluation reward: 2.45	epsilon: 0.6657284800072567	st
episode: 872	score: 0.0	memory length: 168794	epsilon: 0.665482960007262	ste
eps: 123	lr: 4e-05	evaluation reward: 2.42	epsilon: 0.6650473600072715	st
episode: 873	score: 3.0	memory length: 169023	epsilon: 0.6646137400072809	st
eps: 229	lr: 4e-05	evaluation reward: 2.4	epsilon: 0.6640692400072927	st
episode: 874	score: 0.0	memory length: 169147	epsilon: 0.6635247400073045	st
ps: 124	lr: 4e-05	evaluation reward: 2.37	epsilon: 0.663225760007311	ste
episode: 875	score: 2.0	memory length: 169367	epsilon: 0.6629802400073164	st
eps: 220	lr: 4e-05	evaluation reward: 2.39	epsilon: 0.6627347200073217	st
episode: 876	score: 2.0	memory length: 169586	epsilon: 0.6621902200073335	st
eps: 219	lr: 4e-05	evaluation reward: 2.41	epsilon: 0.6613091200073526	st
episode: 877	score: 4.0	memory length: 169861	epsilon: 0.6608180800073633	st
eps: 275	lr: 4e-05	evaluation reward: 2.43	epsilon: 0.6602676400073753	st
episode: 878	score: 4.0	memory length: 170136	epsilon: 0.6596221600073893	st
eps: 275	lr: 4e-05	evaluation reward: 2.45	epsilon: 0.659036080007402	ste
episode: 879	score: 1.0	memory length: 170287	epsilon: 0.6587925400074073	st
ps: 151	lr: 4e-05	evaluation reward: 2.46	epsilon: 0.6584519800074147	st
episode: 880	score: 0.0	memory length: 170411	epsilon: 0.6580579600074232	st
eps: 124	lr: 4e-05	evaluation reward: 2.44	epsilon: 0.6576599800074319	st
episode: 881	score: 0.0	memory length: 170535	epsilon: 0.6572402200074411	ste
eps: 124	lr: 4e-05	evaluation reward: 2.39	epsilon: 0.6568481800074495	st
episode: 882	score: 4.0	memory length: 170810	epsilon: 0.6563333800074607	st
eps: 275	lr: 4e-05	evaluation reward: 2.42	epsilon: 0.6558819400074705	st
episode: 883	score: 8.0	memory length: 171255	epsilon: 0.6552721000074837	st
eps: 445	lr: 4e-05	evaluation reward: 2.46	epsilon: 0.6548800600074922	st
episode: 884	score: 3.0	memory length: 171503	epsilon: 0.6543335800075041	st
eps: 248	lr: 4e-05	evaluation reward: 2.47	epsilon: 0.6537455200075168	st
episode: 885	score: 4.0	memory length: 171781		
eps: 278	lr: 4e-05	evaluation reward: 2.5		
episode: 886	score: 5.0	memory length: 172107		
eps: 326	lr: 4e-05	evaluation reward: 2.53		
episode: 887	score: 4.0	memory length: 172403		
ps: 296	lr: 4e-05	evaluation reward: 2.57		
episode: 888	score: 0.0	memory length: 172526		
eps: 123	lr: 4e-05	evaluation reward: 2.54		
episode: 889	score: 1.0	memory length: 172698		
eps: 172	lr: 4e-05	evaluation reward: 2.54		
episode: 890	score: 2.0	memory length: 172897		
eps: 199	lr: 4e-05	evaluation reward: 2.54		
episode: 891	score: 2.0	memory length: 173098		
eps: 201	lr: 4e-05	evaluation reward: 2.56		
episode: 892	score: 3.0	memory length: 173310		
ps: 212	lr: 4e-05	evaluation reward: 2.57		
episode: 893	score: 2.0	memory length: 173508		
eps: 198	lr: 4e-05	evaluation reward: 2.59		
episode: 894	score: 4.0	memory length: 173768		
eps: 260	lr: 4e-05	evaluation reward: 2.59		
episode: 895	score: 3.0	memory length: 173996		
eps: 228	lr: 4e-05	evaluation reward: 2.58		
episode: 896	score: 5.0	memory length: 174304		
eps: 308	lr: 4e-05	evaluation reward: 2.62		
episode: 897	score: 2.0	memory length: 174502		
eps: 198	lr: 4e-05	evaluation reward: 2.64		
episode: 898	score: 4.0	memory length: 174778		
eps: 276	lr: 4e-05	evaluation reward: 2.65		
episode: 899	score: 4.0	memory length: 175075		
eps: 297	lr: 4e-05	evaluation reward: 2.66		

episode: 900	score: 0.0	memory length: 175198	epsilon: 0.6535019800075221	st
eps: 123	lr: 4e-05	evaluation reward: 2.64	epsilon: 0.6528723400075358	st
episode: 901	score: 4.0	memory length: 175516	epsilon: 0.6522882400075485	st
eps: 318	lr: 4e-05	evaluation reward: 2.65	epsilon: 0.6520447000075538	st
episode: 902	score: 5.0	memory length: 175811	epsilon: 0.6515972200075635	st
eps: 295	lr: 4e-05	evaluation reward: 2.7	epsilon: 0.651203200007572	ste
episode: 903	score: 0.0	memory length: 175934	epsilon: 0.6505438600075863	st
eps: 123	lr: 4e-05	evaluation reward: 2.67	epsilon: 0.6500607400075968	st
episode: 904	score: 3.0	memory length: 176160	epsilon: 0.6493954600076113	st
eps: 226	lr: 4e-05	evaluation reward: 2.68	epsilon: 0.6488747200076226	st
episode: 905	score: 2.0	memory length: 176359	epsilon: 0.6486292000076279	st
ps: 199	lr: 4e-05	evaluation reward: 2.68	epsilon: 0.6481005400076394	st
episode: 906	score: 5.0	memory length: 176692	epsilon: 0.6474530800076534	st
eps: 333	lr: 4e-05	evaluation reward: 2.7	epsilon: 0.6470590600076662	ste
episode: 907	score: 3.0	memory length: 176936	epsilon: 0.6467224600076693	st
eps: 244	lr: 4e-05	evaluation reward: 2.71	epsilon: 0.6461502400076817	st
episode: 908	score: 5.0	memory length: 177272	epsilon: 0.6456215800076932	st
eps: 336	lr: 4e-05	evaluation reward: 2.73	epsilon: 0.6453760600076985	st
episode: 909	score: 4.0	memory length: 177535	epsilon: 0.6451305400077039	st
eps: 263	lr: 4e-05	evaluation reward: 2.73	epsilon: 0.644708800007713	ste
episode: 910	score: 0.0	memory length: 177659	epsilon: 0.6442613200077227	st
eps: 124	lr: 4e-05	evaluation reward: 2.71	epsilon: 0.6437702800077334	st
episode: 911	score: 3.0	memory length: 177926	epsilon: 0.6430317400077494	st
eps: 267	lr: 4e-05	evaluation reward: 2.72	epsilon: 0.6424258600077626	st
episode: 912	score: 5.0	memory length: 178253	epsilon: 0.6420338200077711	st
eps: 327	lr: 4e-05	evaluation reward: 2.75	epsilon: 0.6414873400077783	ste
episode: 913	score: 2.0	memory length: 178452	epsilon: 0.6408418600077797	ste
ps: 199	lr: 4e-05	evaluation reward: 2.76	epsilon: 0.6403528000078076	st
episode: 914	score: 1.0	memory length: 178622	epsilon: 0.6398340400078188	st
eps: 170	lr: 4e-05	evaluation reward: 2.72		
episode: 915	score: 4.0	memory length: 178911		
eps: 289	lr: 4e-05	evaluation reward: 2.73		
episode: 916	score: 3.0	memory length: 179178		
eps: 267	lr: 4e-05	evaluation reward: 2.73		
episode: 917	score: 0.0	memory length: 179302		
eps: 124	lr: 4e-05	evaluation reward: 2.73		
episode: 918	score: 0.0	memory length: 179426		
eps: 124	lr: 4e-05	evaluation reward: 2.73		
episode: 919	score: 3.0	memory length: 179639		
ps: 213	lr: 4e-05	evaluation reward: 2.73		
episode: 920	score: 3.0	memory length: 179865		
eps: 226	lr: 4e-05	evaluation reward: 2.72		
episode: 921	score: 3.0	memory length: 180113		
eps: 248	lr: 4e-05	evaluation reward: 2.75		
episode: 922	score: 6.0	memory length: 180486		
eps: 373	lr: 4e-05	evaluation reward: 2.8		
episode: 923	score: 5.0	memory length: 180792		
eps: 306	lr: 4e-05	evaluation reward: 2.84		
episode: 924	score: 2.0	memory length: 180990		
eps: 198	lr: 4e-05	evaluation reward: 2.85		
episode: 925	score: 4.0	memory length: 181266		
ps: 276	lr: 4e-05	evaluation reward: 2.85		
episode: 926	score: 0.0	memory length: 181389		
eps: 123	lr: 4e-05	evaluation reward: 2.81		
episode: 927	score: 2.0	memory length: 181592		
ps: 203	lr: 4e-05	evaluation reward: 2.83		
episode: 928	score: 3.0	memory length: 181839		
eps: 247	lr: 4e-05	evaluation reward: 2.84		
episode: 929	score: 3.0	memory length: 182101		
eps: 262	lr: 4e-05	evaluation reward: 2.86		

episode: 930 score: 2.0 memory length: 182319	epsilon: 0.6394024000078282	st
eps: 218 lr: 4e-05 evaluation reward: 2.83	epsilon: 0.6387628600078421	st
episode: 931 score: 5.0 memory length: 182642	epsilon: 0.638216380007854	ste
eps: 323 lr: 4e-05 evaluation reward: 2.85	epsilon: 0.6377213800078647	st
episode: 932 score: 4.0 memory length: 182918	epsilon: 0.6372719200078745	st
ps: 276 lr: 4e-05 evaluation reward: 2.89	epsilon: 0.636741280007886	ste
episode: 933 score: 3.0 memory length: 183168	epsilon: 0.6363492400078945	st
eps: 250 lr: 4e-05 evaluation reward: 2.89	epsilon: 0.6358601800079051	st
episode: 934 score: 3.0 memory length: 183395	epsilon: 0.635313700007917	ste
eps: 227 lr: 4e-05 evaluation reward: 2.84	epsilon: 0.634670200007931	ste
episode: 935 score: 3.0 memory length: 183663	epsilon: 0.6341811400079416	st
ps: 268 lr: 4e-05 evaluation reward: 2.85	epsilon: 0.6337871200079501	st
episode: 936 score: 2.0 memory length: 183861	epsilon: 0.6333396400079598	st
eps: 198 lr: 4e-05 evaluation reward: 2.87	epsilon: 0.6329456200079684	st
episode: 937 score: 3.0 memory length: 184108	epsilon: 0.6324862600079784	st
eps: 247 lr: 4e-05 evaluation reward: 2.9	epsilon: 0.6320942200079869	st
episode: 938 score: 4.0 memory length: 184384	epsilon: 0.6311873800080066	st
ps: 276 lr: 4e-05 evaluation reward: 2.94	epsilon: 0.6307953400080151	st
episode: 939 score: 5.0 memory length: 184709	epsilon: 0.6303043000080257	st
ps: 325 lr: 4e-05 evaluation reward: 2.96	epsilon: 0.6297023800080388	st
episode: 940 score: 3.0 memory length: 184956	epsilon: 0.6292469800080487	st
eps: 247 lr: 4e-05 evaluation reward: 2.9	epsilon: 0.6287955400080585	st
episode: 941 score: 2.0 memory length: 185155	epsilon: 0.628496560008065	ste
eps: 199 lr: 4e-05 evaluation reward: 2.89	epsilon: 0.6281025400080735	st
episode: 942 score: 3.0 memory length: 185381	epsilon: 0.6274511200080877	st
eps: 226 lr: 4e-05 evaluation reward: 2.89	epsilon: 0.6269046400080995	st
episode: 943 score: 2.0 memory length: 185580	epsilon: 0.6264551800081093	st
eps: 199 lr: 4e-05 evaluation reward: 2.87	epsilon: 0.6256552600081267	st
episode: 944 score: 3.0 memory length: 185812	epsilon: 0.6251662000081373	st
eps: 232 lr: 4e-05 evaluation reward: 2.85	epsilon: 0.6247721800081458	st
episode: 945 score: 2.0 memory length: 186010		
eps: 198 lr: 4e-05 evaluation reward: 2.85		
episode: 946 score: 8.0 memory length: 186468		
eps: 458 lr: 4e-05 evaluation reward: 2.91		
episode: 947 score: 2.0 memory length: 186666		
eps: 198 lr: 4e-05 evaluation reward: 2.9		
episode: 948 score: 3.0 memory length: 186914		
eps: 248 lr: 4e-05 evaluation reward: 2.89		
episode: 949 score: 5.0 memory length: 187218		
eps: 304 lr: 4e-05 evaluation reward: 2.89		
episode: 950 score: 3.0 memory length: 187448		
eps: 230 lr: 4e-05 evaluation reward: 2.92		
episode: 951 score: 3.0 memory length: 187676		
eps: 228 lr: 4e-05 evaluation reward: 2.9		
episode: 952 score: 1.0 memory length: 187827		
ps: 151 lr: 4e-05 evaluation reward: 2.86		
episode: 953 score: 2.0 memory length: 188026		
eps: 199 lr: 4e-05 evaluation reward: 2.83		
episode: 954 score: 5.0 memory length: 188355		
eps: 329 lr: 4e-05 evaluation reward: 2.86		
episode: 955 score: 4.0 memory length: 188631		
eps: 276 lr: 4e-05 evaluation reward: 2.89		
episode: 956 score: 3.0 memory length: 188858		
eps: 227 lr: 4e-05 evaluation reward: 2.9		
episode: 957 score: 9.0 memory length: 189262		
eps: 404 lr: 4e-05 evaluation reward: 2.97		
episode: 958 score: 3.0 memory length: 189509		
eps: 247 lr: 4e-05 evaluation reward: 2.98		
episode: 959 score: 2.0 memory length: 189708		
eps: 199 lr: 4e-05 evaluation reward: 2.95		

episode: 960 score: 6.0 memory length: 190062	epsilon: 0.624071260008161	ste
ps: 354 lr: 4e-05 evaluation reward: 2.98	epsilon: 0.6232673800081785	st
episode: 961 score: 7.0 memory length: 190468	epsilon: 0.6228337600081879	st
eps: 406 lr: 4e-05 evaluation reward: 3.04	epsilon: 0.6222853000081998	st
episode: 962 score: 2.0 memory length: 190687	epsilon: 0.6217388200082117	st
eps: 219 lr: 4e-05 evaluation reward: 3.03	epsilon: 0.6212517400082223	st
episode: 963 score: 4.0 memory length: 190964	epsilon: 0.6208181200082317	st
eps: 277 lr: 4e-05 evaluation reward: 3.05	epsilon: 0.6202716400082435	st
episode: 964 score: 4.0 memory length: 191240	epsilon: 0.6198538600082526	st
eps: 276 lr: 4e-05 evaluation reward: 3.06	epsilon: 0.6194618200082611	st
episode: 965 score: 3.0 memory length: 191486	epsilon: 0.6190381000082703	st
eps: 246 lr: 4e-05 evaluation reward: 3.04	epsilon: 0.6186421000082789	st
episode: 966 score: 2.0 memory length: 191705	epsilon: 0.6181946200082886	st
eps: 219 lr: 4e-05 evaluation reward: 3.02	epsilon: 0.6177431800082984	st
episode: 967 score: 4.0 memory length: 191981	epsilon: 0.6169531600083156	st
eps: 276 lr: 4e-05 evaluation reward: 3.03	epsilon: 0.6166145800083229	st
episode: 968 score: 3.0 memory length: 192192	epsilon: 0.6161809600083323	st
eps: 211 lr: 4e-05 evaluation reward: 3.04	epsilon: 0.6155374600083463	st
episode: 969 score: 2.0 memory length: 192390	epsilon: 0.6150226600083575	st
eps: 198 lr: 4e-05 evaluation reward: 3.03	epsilon: 0.6145138000083685	st
episode: 970 score: 3.0 memory length: 192604	epsilon: 0.6138722800083825	st
eps: 214 lr: 4e-05 evaluation reward: 3.02	epsilon: 0.613571320008389	ste
episode: 971 score: 2.0 memory length: 192804	epsilon: 0.6132723400083955	st
eps: 200 lr: 4e-05 evaluation reward: 3.03	epsilon: 0.6128209000084053	st
episode: 972 score: 3.0 memory length: 193030	epsilon: 0.6124565800084132	st
eps: 226 lr: 4e-05 evaluation reward: 3.06	epsilon: 0.6118507000084263	st
episode: 973 score: 3.0 memory length: 193258	epsilon: 0.611361640008437	ste
eps: 228 lr: 4e-05 evaluation reward: 3.06	epsilon: 0.6104627200084565	st
episode: 974 score: 6.0 memory length: 193657	epsilon: 0.6100033600084664	st
eps: 399 lr: 4e-05 evaluation reward: 3.12	epsilon: 0.6091143400084857	st
episode: 975 score: 1.0 memory length: 193828		
eps: 171 lr: 4e-05 evaluation reward: 3.11		
episode: 976 score: 2.0 memory length: 194047		
eps: 219 lr: 4e-05 evaluation reward: 3.11		
episode: 977 score: 5.0 memory length: 194372		
eps: 325 lr: 4e-05 evaluation reward: 3.12		
episode: 978 score: 4.0 memory length: 194632		
eps: 260 lr: 4e-05 evaluation reward: 3.12		
episode: 979 score: 4.0 memory length: 194889		
eps: 257 lr: 4e-05 evaluation reward: 3.15		
episode: 980 score: 5.0 memory length: 195213		
eps: 324 lr: 4e-05 evaluation reward: 3.2		
episode: 981 score: 1.0 memory length: 195365		
ps: 152 lr: 4e-05 evaluation reward: 3.21		
episode: 982 score: 1.0 memory length: 195516		
eps: 151 lr: 4e-05 evaluation reward: 3.18		
episode: 983 score: 3.0 memory length: 195744		
eps: 228 lr: 4e-05 evaluation reward: 3.13		
episode: 984 score: 2.0 memory length: 195928		
eps: 184 lr: 4e-05 evaluation reward: 3.12		
episode: 985 score: 5.0 memory length: 196234		
eps: 306 lr: 4e-05 evaluation reward: 3.13		
episode: 986 score: 3.0 memory length: 196481		
ps: 247 lr: 4e-05 evaluation reward: 3.11		
episode: 987 score: 8.0 memory length: 196935		
eps: 454 lr: 4e-05 evaluation reward: 3.15		
episode: 988 score: 3.0 memory length: 197167		
eps: 232 lr: 4e-05 evaluation reward: 3.18		
episode: 989 score: 8.0 memory length: 197616		
eps: 449 lr: 4e-05 evaluation reward: 3.25		

episode: 990	score: 0.0	memory length: 197740	epsilon: 0.6088688200084911	st
eps: 124	lr: 4e-05	evaluation reward: 3.23	epsilon: 0.6084708400084997	st
episode: 991	score: 2.0	memory length: 197941	epsilon: 0.6080768200085083	st
eps: 201	lr: 4e-05	evaluation reward: 3.23	epsilon: 0.6075679600085193	st
episode: 992	score: 2.0	memory length: 198140	epsilon: 0.6071145400085292	st
eps: 199	lr: 4e-05	evaluation reward: 3.22	epsilon: 0.606568060008541	ste
episode: 993	score: 4.0	memory length: 198397	epsilon: 0.6060908800085514	st
eps: 257	lr: 4e-05	evaluation reward: 3.24	epsilon: 0.6056335000085613	st
episode: 994	score: 3.0	memory length: 198626	epsilon: 0.6050731600085735	st
eps: 229	lr: 4e-05	evaluation reward: 3.23	epsilon: 0.6044653000085867	st
episode: 995	score: 4.0	memory length: 198902	epsilon: 0.604128700008594	st
ps: 276	lr: 4e-05	evaluation reward: 3.24	epsilon: 0.6036376600086046	s
episode: 996	score: 4.0	memory length: 199143	epsilon: 0.6030476200086174	s
eps: 241	lr: 4e-05	evaluation reward: 3.23	epsilon: 0.6025546000086281	s
episode: 997	score: 3.0	memory length: 199374	epsilon: 0.6021308800086373	s
eps: 231	lr: 4e-05	evaluation reward: 3.24	epsilon: 0.60214834200086514	s
episode: 998	score: 4.0	memory length: 199657	epsilon: 0.60108940000866	ste
eps: 283	lr: 4e-05	evaluation reward: 3.24	epsilon: 0.6016000000000000	s
episode: 999	score: 5.0	memory length: 199964	epsilon: 0.6016000000000000	s
eps: 307	lr: 4e-05	evaluation reward: 3.25	epsilon: 0.6016000000000000	s
episode: 1000	score: 1.0	memory length: 200134	epsilon: 0.6016000000000000	s
eps: 170	lr: 4e-05	evaluation reward: 3.26	epsilon: 0.6016000000000000	s
episode: 1001	score: 3.0	memory length: 200382	epsilon: 0.6016000000000000	s
teps: 248	lr: 1.6000000000000000e-05	evaluation reward: 3.25	epsilon: 0.6016000000000000	s
episode: 1002	score: 4.0	memory length: 200680	epsilon: 0.6016000000000000	s
teps: 298	lr: 1.6000000000000000e-05	evaluation reward: 3.24	epsilon: 0.6016000000000000	s
episode: 1003	score: 3.0	memory length: 200929	epsilon: 0.6016000000000000	s
teps: 249	lr: 1.6000000000000000e-05	evaluation reward: 3.27	epsilon: 0.6016000000000000	s
episode: 1004	score: 3.0	memory length: 201143	epsilon: 0.6016000000000000	s
teps: 214	lr: 1.6000000000000000e-05	evaluation reward: 3.27	epsilon: 0.6016000000000000	s
episode: 1005	score: 5.0	memory length: 201470	epsilon: 0.6016000000000000	s
teps: 327	lr: 1.6000000000000000e-05	evaluation reward: 3.3	epsilon: 0.6016000000000000	s
episode: 1006	score: 2.0	memory length: 201669	epsilon: 0.6016000000000000	ste
ps: 199	lr: 1.6000000000000000e-05	evaluation reward: 3.27	epsilon: 0.6016000000000000	s
episode: 1007	score: 2.0	memory length: 201867	epsilon: 0.6016000000000000	s
teps: 198	lr: 1.6000000000000000e-05	evaluation reward: 3.26	epsilon: 0.6016000000000000	s
episode: 1008	score: 0.0	memory length: 201991	epsilon: 0.6016000000000000	s
teps: 124	lr: 1.6000000000000000e-05	evaluation reward: 3.21	epsilon: 0.6016000000000000	s
episode: 1009	score: 5.0	memory length: 202339	epsilon: 0.5997628000086888	s
teps: 348	lr: 1.6000000000000000e-05	evaluation reward: 3.22	epsilon: 0.5997628000086888	s
episode: 1010	score: 2.0	memory length: 202521	epsilon: 0.5994024400086966	s
teps: 182	lr: 1.6000000000000000e-05	evaluation reward: 3.24	epsilon: 0.5994024400086966	s
episode: 1011	score: 3.0	memory length: 202788	epsilon: 0.5988737800087081	s
teps: 267	lr: 1.6000000000000000e-05	evaluation reward: 3.24	epsilon: 0.5988737800087081	s
episode: 1012	score: 2.0	memory length: 203007	epsilon: 0.5984401600087175	s
teps: 219	lr: 1.6000000000000000e-05	evaluation reward: 3.21	epsilon: 0.5984401600087175	s
episode: 1013	score: 0.0	memory length: 203131	epsilon: 0.5981946400087228	s
teps: 124	lr: 1.6000000000000000e-05	evaluation reward: 3.19	epsilon: 0.5981946400087228	s
episode: 1014	score: 6.0	memory length: 203509	epsilon: 0.597446200008739	st
eps: 378	lr: 1.6000000000000000e-05	evaluation reward: 3.24	epsilon: 0.597446200008739	s
episode: 1015	score: 8.0	memory length: 203885	epsilon: 0.5967017200087552	s
teps: 376	lr: 1.6000000000000000e-05	evaluation reward: 3.28	epsilon: 0.5967017200087552	s
episode: 1016	score: 3.0	memory length: 204112	epsilon: 0.596252260008765	st
eps: 227	lr: 1.6000000000000000e-05	evaluation reward: 3.28	epsilon: 0.596252260008765	s
episode: 1017	score: 3.0	memory length: 204326	epsilon: 0.5958285400087742	s
teps: 214	lr: 1.6000000000000000e-05	evaluation reward: 3.31	epsilon: 0.5958285400087742	s
episode: 1018	score: 3.0	memory length: 204557	epsilon: 0.5953711600087841	s
teps: 231	lr: 1.6000000000000000e-05	evaluation reward: 3.34	epsilon: 0.5953711600087841	s
episode: 1019	score: 2.0	memory length: 204774	epsilon: 0.5949415000087934	s
teps: 217	lr: 1.6000000000000000e-05	evaluation reward: 3.33	epsilon: 0.5949415000087934	s

episode: 1020 score: 3.0 memory length: 205001 epsilon: 0.5944920400088032 s  
 teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.33  
 episode: 1021 score: 3.0 memory length: 205248 epsilon: 0.5940029800088138 s  
 teps: 247 lr: 1.6000000000000003e-05 evaluation reward: 3.33  
 episode: 1022 score: 13.0 memory length: 205693 epsilon: 0.5931218800088329  
 steps: 445 lr: 1.6000000000000003e-05 evaluation reward: 3.4  
 episode: 1023 score: 1.0 memory length: 205844 epsilon: 0.5928229000088394 s  
 teps: 151 lr: 1.6000000000000003e-05 evaluation reward: 3.36  
 episode: 1024 score: 3.0 memory length: 206073 epsilon: 0.5923694800088493 s  
 teps: 229 lr: 1.6000000000000003e-05 evaluation reward: 3.37  
 episode: 1025 score: 2.0 memory length: 206271 epsilon: 0.5919774400088578 s  
 teps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.35  
 episode: 1026 score: 9.0 memory length: 206784 epsilon: 0.5909617000088798 s  
 teps: 513 lr: 1.6000000000000003e-05 evaluation reward: 3.44  
 episode: 1027 score: 3.0 memory length: 207011 epsilon: 0.5905122400088896 s  
 teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.45  
 episode: 1028 score: 3.0 memory length: 207261 epsilon: 0.5900172400089003 s  
 teps: 250 lr: 1.6000000000000003e-05 evaluation reward: 3.45  
 episode: 1029 score: 2.0 memory length: 207443 epsilon: 0.5896568800089081 s  
 teps: 182 lr: 1.6000000000000003e-05 evaluation reward: 3.44  
 episode: 1030 score: 2.0 memory length: 207642 epsilon: 0.5892628600089167 s  
 teps: 199 lr: 1.6000000000000003e-05 evaluation reward: 3.44  
 episode: 1031 score: 2.0 memory length: 207823 epsilon: 0.5889044800089245 s  
 teps: 181 lr: 1.6000000000000003e-05 evaluation reward: 3.41  
 episode: 1032 score: 3.0 memory length: 208054 epsilon: 0.5884471000089344 s  
 teps: 231 lr: 1.6000000000000003e-05 evaluation reward: 3.4  
 episode: 1033 score: 3.0 memory length: 208281 epsilon: 0.5879976400089442 s  
 teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.4  
 episode: 1034 score: 7.0 memory length: 208712 epsilon: 0.5871442600089627 s  
 teps: 431 lr: 1.6000000000000003e-05 evaluation reward: 3.44  
 episode: 1035 score: 3.0 memory length: 208960 epsilon: 0.5866532200089734 s  
 teps: 248 lr: 1.6000000000000003e-05 evaluation reward: 3.44  
 episode: 1036 score: 3.0 memory length: 209207 epsilon: 0.586164160008984 st  
 eps: 247 lr: 1.6000000000000003e-05 evaluation reward: 3.45  
 episode: 1037 score: 2.0 memory length: 209405 epsilon: 0.5857721200089925 s  
 teps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.44  
 episode: 1038 score: 1.0 memory length: 209574 epsilon: 0.5854375000089997 s  
 teps: 169 lr: 1.6000000000000003e-05 evaluation reward: 3.41  
 episode: 1039 score: 2.0 memory length: 209757 epsilon: 0.5850751600090076 s  
 teps: 183 lr: 1.6000000000000003e-05 evaluation reward: 3.38  
 episode: 1040 score: 2.0 memory length: 209939 epsilon: 0.5847148000090154 s  
 teps: 182 lr: 1.6000000000000003e-05 evaluation reward: 3.37  
 episode: 1041 score: 3.0 memory length: 210184 epsilon: 0.584229700009026 st  
 eps: 245 lr: 1.6000000000000003e-05 evaluation reward: 3.38  
 episode: 1042 score: 4.0 memory length: 210423 epsilon: 0.5837564800090362 s  
 teps: 239 lr: 1.6000000000000003e-05 evaluation reward: 3.39  
 episode: 1043 score: 5.0 memory length: 210741 epsilon: 0.5831268400090499 s  
 teps: 318 lr: 1.6000000000000003e-05 evaluation reward: 3.42  
 episode: 1044 score: 5.0 memory length: 211062 epsilon: 0.5824912600090637 s  
 teps: 321 lr: 1.6000000000000003e-05 evaluation reward: 3.44  
 episode: 1045 score: 4.0 memory length: 211360 epsilon: 0.5819012200090765 s  
 teps: 298 lr: 1.6000000000000003e-05 evaluation reward: 3.46  
 episode: 1046 score: 3.0 memory length: 211587 epsilon: 0.5814517600090863 s  
 teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.41  
 episode: 1047 score: 4.0 memory length: 211864 epsilon: 0.5809033000090982 s  
 teps: 277 lr: 1.6000000000000003e-05 evaluation reward: 3.43  
 episode: 1048 score: 1.0 memory length: 212033 epsilon: 0.5805686800091054 s  
 teps: 169 lr: 1.6000000000000003e-05 evaluation reward: 3.41  
 episode: 1049 score: 3.0 memory length: 212302 epsilon: 0.580036060009117 st  
 eps: 269 lr: 1.6000000000000003e-05 evaluation reward: 3.39

episode: 1050 score: 1.0 memory length: 212454 epsilon: 0.5797351000091235 s  
teps: 152 lr: 1.6000000000000003e-05 evaluation reward: 3.37  
episode: 1051 score: 1.0 memory length: 212626 epsilon: 0.5793945400091309 s  
teps: 172 lr: 1.6000000000000003e-05 evaluation reward: 3.35  
episode: 1052 score: 4.0 memory length: 212945 epsilon: 0.5787629200091446 s  
teps: 319 lr: 1.6000000000000003e-05 evaluation reward: 3.38  
episode: 1053 score: 3.0 memory length: 213172 epsilon: 0.5783134600091544 s  
teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.39  
episode: 1054 score: 9.0 memory length: 213545 epsilon: 0.5775749200091704 s  
teps: 373 lr: 1.6000000000000003e-05 evaluation reward: 3.43  
episode: 1055 score: 3.0 memory length: 213772 epsilon: 0.5771254600091802 s  
teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.42  
episode: 1056 score: 8.0 memory length: 214085 epsilon: 0.5765057200091936 s  
teps: 313 lr: 1.6000000000000003e-05 evaluation reward: 3.47  
episode: 1057 score: 2.0 memory length: 214267 epsilon: 0.5761453600092015 s  
teps: 182 lr: 1.6000000000000003e-05 evaluation reward: 3.4  
episode: 1058 score: 4.0 memory length: 214543 epsilon: 0.5755988800092133 s  
teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.41  
episode: 1059 score: 2.0 memory length: 214742 epsilon: 0.5752048600092219 s  
teps: 199 lr: 1.6000000000000003e-05 evaluation reward: 3.41  
episode: 1060 score: 7.0 memory length: 215115 epsilon: 0.5744663200092379 s  
teps: 373 lr: 1.6000000000000003e-05 evaluation reward: 3.42  
episode: 1061 score: 4.0 memory length: 215375 epsilon: 0.5739515200092491 s  
teps: 260 lr: 1.6000000000000003e-05 evaluation reward: 3.39  
episode: 1062 score: 3.0 memory length: 215604 epsilon: 0.5734981000092589 s  
teps: 229 lr: 1.6000000000000003e-05 evaluation reward: 3.4  
episode: 1063 score: 4.0 memory length: 215879 epsilon: 0.5729536000092708 s  
teps: 275 lr: 1.6000000000000003e-05 evaluation reward: 3.4  
episode: 1064 score: 0.0 memory length: 216003 epsilon: 0.5727080800092761 s  
teps: 124 lr: 1.6000000000000003e-05 evaluation reward: 3.36  
episode: 1065 score: 6.0 memory length: 216314 epsilon: 0.5720923000092895 s  
teps: 311 lr: 1.6000000000000003e-05 evaluation reward: 3.39  
episode: 1066 score: 5.0 memory length: 216620 epsilon: 0.5714864200093026 s  
teps: 306 lr: 1.6000000000000003e-05 evaluation reward: 3.42  
episode: 1067 score: 1.0 memory length: 216772 epsilon: 0.5711854600093091 s  
teps: 152 lr: 1.6000000000000003e-05 evaluation reward: 3.39  
episode: 1068 score: 3.0 memory length: 216983 epsilon: 0.5707676800093182 s  
teps: 211 lr: 1.6000000000000003e-05 evaluation reward: 3.39  
episode: 1069 score: 3.0 memory length: 217197 epsilon: 0.5703439600093274 s  
teps: 214 lr: 1.6000000000000003e-05 evaluation reward: 3.4  
episode: 1070 score: 4.0 memory length: 217473 epsilon: 0.5697974800093393 s  
teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.41  
episode: 1071 score: 3.0 memory length: 217683 epsilon: 0.5693816800093483 s  
teps: 210 lr: 1.6000000000000003e-05 evaluation reward: 3.42  
episode: 1072 score: 3.0 memory length: 217912 epsilon: 0.5689282600093581 s  
teps: 229 lr: 1.6000000000000003e-05 evaluation reward: 3.42  
episode: 1073 score: 7.0 memory length: 218320 epsilon: 0.5681204200093757 s  
teps: 408 lr: 1.6000000000000003e-05 evaluation reward: 3.46  
episode: 1074 score: 4.0 memory length: 218596 epsilon: 0.5675739400093875 s  
teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.44  
episode: 1075 score: 8.0 memory length: 218929 epsilon: 0.5669146000094019 s  
teps: 333 lr: 1.6000000000000003e-05 evaluation reward: 3.51  
episode: 1076 score: 3.0 memory length: 219177 epsilon: 0.5664235600094125 s  
teps: 248 lr: 1.6000000000000003e-05 evaluation reward: 3.52  
episode: 1077 score: 1.0 memory length: 219348 epsilon: 0.5660849800094199 s  
teps: 171 lr: 1.6000000000000003e-05 evaluation reward: 3.48  
episode: 1078 score: 3.0 memory length: 219595 epsilon: 0.5655959200094305 s  
teps: 247 lr: 1.6000000000000003e-05 evaluation reward: 3.47  
episode: 1079 score: 5.0 memory length: 219920 epsilon: 0.5649524200094445 s  
teps: 325 lr: 1.6000000000000003e-05 evaluation reward: 3.48

episode: 1080 score: 8.0 memory length: 220318 epsilon: 0.5641643800094616 s  
 teps: 398 lr: 1.6000000000000003e-05 evaluation reward: 3.51  
 episode: 1081 score: 3.0 memory length: 220545 epsilon: 0.5637149200094713 s  
 teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.53  
 episode: 1082 score: 6.0 memory length: 220881 epsilon: 0.5630496400094858 s  
 teps: 336 lr: 1.6000000000000003e-05 evaluation reward: 3.58  
 episode: 1083 score: 6.0 memory length: 221255 epsilon: 0.5623091200095018 s  
 teps: 374 lr: 1.6000000000000003e-05 evaluation reward: 3.61  
 episode: 1084 score: 2.0 memory length: 221472 epsilon: 0.5618794600095112 s  
 teps: 217 lr: 1.6000000000000003e-05 evaluation reward: 3.61  
 episode: 1085 score: 5.0 memory length: 221780 epsilon: 0.5612696200095244 s  
 teps: 308 lr: 1.6000000000000003e-05 evaluation reward: 3.61  
 episode: 1086 score: 4.0 memory length: 222040 epsilon: 0.5607548200095356 s  
 teps: 260 lr: 1.6000000000000003e-05 evaluation reward: 3.62  
 episode: 1087 score: 3.0 memory length: 222287 epsilon: 0.5602657600095462 s  
 teps: 247 lr: 1.6000000000000003e-05 evaluation reward: 3.57  
 episode: 1088 score: 4.0 memory length: 222580 epsilon: 0.5596856200095588 s  
 teps: 293 lr: 1.6000000000000003e-05 evaluation reward: 3.58  
 episode: 1089 score: 5.0 memory length: 222908 epsilon: 0.5590361800095729 s  
 teps: 328 lr: 1.6000000000000003e-05 evaluation reward: 3.55  
 episode: 1090 score: 3.0 memory length: 223121 epsilon: 0.558614440009582 st  
 eps: 213 lr: 1.6000000000000003e-05 evaluation reward: 3.58  
 episode: 1091 score: 6.0 memory length: 223481 epsilon: 0.5579016400095975 s  
 teps: 360 lr: 1.6000000000000003e-05 evaluation reward: 3.62  
 episode: 1092 score: 3.0 memory length: 223713 epsilon: 0.5574422800096075 s  
 teps: 232 lr: 1.6000000000000003e-05 evaluation reward: 3.63  
 episode: 1093 score: 3.0 memory length: 223924 epsilon: 0.5570245000096166 s  
 teps: 211 lr: 1.6000000000000003e-05 evaluation reward: 3.62  
 episode: 1094 score: 8.0 memory length: 224396 epsilon: 0.5560899400096369 s  
 teps: 472 lr: 1.6000000000000003e-05 evaluation reward: 3.67  
 episode: 1095 score: 5.0 memory length: 224705 epsilon: 0.5554781200096501 s  
 teps: 309 lr: 1.6000000000000003e-05 evaluation reward: 3.68  
 episode: 1096 score: 4.0 memory length: 224983 epsilon: 0.5549276800096621 s  
 teps: 278 lr: 1.6000000000000003e-05 evaluation reward: 3.68  
 episode: 1097 score: 1.0 memory length: 225134 epsilon: 0.5546287000096686 s  
 teps: 151 lr: 1.6000000000000003e-05 evaluation reward: 3.66  
 episode: 1098 score: 5.0 memory length: 225435 epsilon: 0.5540327200096815 s  
 teps: 301 lr: 1.6000000000000003e-05 evaluation reward: 3.67  
 episode: 1099 score: 5.0 memory length: 225743 epsilon: 0.5534228800096948 s  
 teps: 308 lr: 1.6000000000000003e-05 evaluation reward: 3.67  
 episode: 1100 score: 3.0 memory length: 225991 epsilon: 0.5529318400097054 s  
 teps: 248 lr: 1.6000000000000003e-05 evaluation reward: 3.69  
 episode: 1101 score: 4.0 memory length: 226267 epsilon: 0.5523853600097173 s  
 teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.7  
 episode: 1102 score: 4.0 memory length: 226543 epsilon: 0.5518388800097291 s  
 teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.7  
 episode: 1103 score: 7.0 memory length: 226973 epsilon: 0.5509874800097476 s  
 teps: 430 lr: 1.6000000000000003e-05 evaluation reward: 3.74  
 episode: 1104 score: 2.0 memory length: 227174 epsilon: 0.5505895000097563 s  
 teps: 201 lr: 1.6000000000000003e-05 evaluation reward: 3.73  
 episode: 1105 score: 1.0 memory length: 227326 epsilon: 0.5502885400097628 s  
 teps: 152 lr: 1.6000000000000003e-05 evaluation reward: 3.69  
 episode: 1106 score: 1.0 memory length: 227478 epsilon: 0.5499875800097693 s  
 teps: 152 lr: 1.6000000000000003e-05 evaluation reward: 3.68  
 episode: 1107 score: 3.0 memory length: 227705 epsilon: 0.5495381200097791 s  
 teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.69  
 episode: 1108 score: 1.0 memory length: 227875 epsilon: 0.5492015200097864 s  
 teps: 170 lr: 1.6000000000000003e-05 evaluation reward: 3.7  
 episode: 1109 score: 5.0 memory length: 228167 epsilon: 0.5486233600097989 s  
 teps: 292 lr: 1.6000000000000003e-05 evaluation reward: 3.7

episode: 1110 score: 6.0 memory length: 228523 epsilon: 0.5479184800098142 s  
 teps: 356 lr: 1.6000000000000003e-05 evaluation reward: 3.74  
 episode: 1111 score: 1.0 memory length: 228695 epsilon: 0.5475779200098216 s  
 teps: 172 lr: 1.6000000000000003e-05 evaluation reward: 3.72  
 episode: 1112 score: 2.0 memory length: 228896 epsilon: 0.5471799400098303 s  
 teps: 201 lr: 1.6000000000000003e-05 evaluation reward: 3.72  
 episode: 1113 score: 2.0 memory length: 229095 epsilon: 0.5467859200098388 s  
 teps: 199 lr: 1.6000000000000003e-05 evaluation reward: 3.74  
 episode: 1114 score: 3.0 memory length: 229306 epsilon: 0.5463681400098479 s  
 teps: 211 lr: 1.6000000000000003e-05 evaluation reward: 3.71  
 episode: 1115 score: 5.0 memory length: 229597 epsilon: 0.5457919600098604 s  
 teps: 291 lr: 1.6000000000000003e-05 evaluation reward: 3.68  
 episode: 1116 score: 7.0 memory length: 229953 epsilon: 0.5450870800098757 s  
 teps: 356 lr: 1.6000000000000003e-05 evaluation reward: 3.72  
 episode: 1117 score: 5.0 memory length: 230280 epsilon: 0.5444396200098898 s  
 teps: 327 lr: 1.6000000000000003e-05 evaluation reward: 3.74  
 episode: 1118 score: 4.0 memory length: 230538 epsilon: 0.5439287800099009 s  
 teps: 258 lr: 1.6000000000000003e-05 evaluation reward: 3.75  
 episode: 1119 score: 4.0 memory length: 230816 epsilon: 0.5433783400099128 s  
 teps: 278 lr: 1.6000000000000003e-05 evaluation reward: 3.77  
 episode: 1120 score: 4.0 memory length: 231090 epsilon: 0.5428358200099246 s  
 teps: 274 lr: 1.6000000000000003e-05 evaluation reward: 3.78  
 episode: 1121 score: 4.0 memory length: 231388 epsilon: 0.5422457800099374 s  
 teps: 298 lr: 1.6000000000000003e-05 evaluation reward: 3.79  
 episode: 1122 score: 4.0 memory length: 231664 epsilon: 0.5416993000099493 s  
 teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.7  
 episode: 1123 score: 3.0 memory length: 231891 epsilon: 0.541249840009959 st  
 eps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.72  
 episode: 1124 score: 2.0 memory length: 232107 epsilon: 0.5408221600099683 s  
 teps: 216 lr: 1.6000000000000003e-05 evaluation reward: 3.71  
 episode: 1125 score: 7.0 memory length: 232533 epsilon: 0.5399786800099866 s  
 teps: 426 lr: 1.6000000000000003e-05 evaluation reward: 3.76  
 episode: 1126 score: 5.0 memory length: 232841 epsilon: 0.5393688400099999 s  
 teps: 308 lr: 1.6000000000000003e-05 evaluation reward: 3.72  
 episode: 1127 score: 7.0 memory length: 233243 epsilon: 0.5385728800100171 s  
 teps: 402 lr: 1.6000000000000003e-05 evaluation reward: 3.76  
 episode: 1128 score: 4.0 memory length: 233520 epsilon: 0.538024420010029 st  
 eps: 277 lr: 1.6000000000000003e-05 evaluation reward: 3.77  
 episode: 1129 score: 8.0 memory length: 233997 epsilon: 0.5370799600100495 s  
 teps: 477 lr: 1.6000000000000003e-05 evaluation reward: 3.83  
 episode: 1130 score: 2.0 memory length: 234196 epsilon: 0.5366859400100581 s  
 teps: 199 lr: 1.6000000000000003e-05 evaluation reward: 3.83  
 episode: 1131 score: 3.0 memory length: 234423 epsilon: 0.5362364800100679 s  
 teps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.84  
 episode: 1132 score: 4.0 memory length: 234679 epsilon: 0.5357296000100789 s  
 teps: 256 lr: 1.6000000000000003e-05 evaluation reward: 3.85  
 episode: 1133 score: 3.0 memory length: 234908 epsilon: 0.5352761800100887 s  
 teps: 229 lr: 1.6000000000000003e-05 evaluation reward: 3.85  
 episode: 1134 score: 2.0 memory length: 235111 epsilon: 0.5348742400100974 s  
 teps: 203 lr: 1.6000000000000003e-05 evaluation reward: 3.8  
 episode: 1135 score: 1.0 memory length: 235281 epsilon: 0.5345376400101047 s  
 teps: 170 lr: 1.6000000000000003e-05 evaluation reward: 3.78  
 episode: 1136 score: 4.0 memory length: 235540 epsilon: 0.5340248200101159 s  
 teps: 259 lr: 1.6000000000000003e-05 evaluation reward: 3.79  
 episode: 1137 score: 5.0 memory length: 235859 epsilon: 0.5333932000101296 s  
 teps: 319 lr: 1.6000000000000003e-05 evaluation reward: 3.82  
 episode: 1138 score: 6.0 memory length: 236224 epsilon: 0.5326705000101453 s  
 teps: 365 lr: 1.6000000000000003e-05 evaluation reward: 3.87  
 episode: 1139 score: 8.0 memory length: 236699 epsilon: 0.5317300000101657 s  
 teps: 475 lr: 1.6000000000000003e-05 evaluation reward: 3.93

episode: 1140 score: 3.0 memory length: 236947 epsilon: 0.5312389600101763 s  
 teps: 248 lr: 1.6000000000000003e-05 evaluation reward: 3.94  
 episode: 1141 score: 2.0 memory length: 237130 epsilon: 0.5308766200101842 s  
 teps: 183 lr: 1.6000000000000003e-05 evaluation reward: 3.93  
 episode: 1142 score: 4.0 memory length: 237406 epsilon: 0.5303301400101961 s  
 teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.93  
 episode: 1143 score: 3.0 memory length: 237617 epsilon: 0.5299123600102051 s  
 teps: 211 lr: 1.6000000000000003e-05 evaluation reward: 3.91  
 episode: 1144 score: 4.0 memory length: 237879 epsilon: 0.5293936000102164 s  
 teps: 262 lr: 1.6000000000000003e-05 evaluation reward: 3.9  
 episode: 1145 score: 1.0 memory length: 238049 epsilon: 0.5290570000102237 s  
 teps: 170 lr: 1.6000000000000003e-05 evaluation reward: 3.87  
 episode: 1146 score: 8.0 memory length: 238432 epsilon: 0.5282986600102402 s  
 teps: 383 lr: 1.6000000000000003e-05 evaluation reward: 3.92  
 episode: 1147 score: 1.0 memory length: 238584 epsilon: 0.5279977000102467 s  
 teps: 152 lr: 1.6000000000000003e-05 evaluation reward: 3.89  
 episode: 1148 score: 5.0 memory length: 238910 epsilon: 0.5273522200102607 s  
 teps: 326 lr: 1.6000000000000003e-05 evaluation reward: 3.93  
 episode: 1149 score: 5.0 memory length: 239228 epsilon: 0.5267225800102744 s  
 teps: 318 lr: 1.6000000000000003e-05 evaluation reward: 3.95  
 episode: 1150 score: 13.0 memory length: 239771 epsilon: 0.5256474400102977 s  
 steps: 543 lr: 1.6000000000000003e-05 evaluation reward: 4.07  
 episode: 1151 score: 7.0 memory length: 240188 epsilon: 0.5248217800103157 s  
 teps: 417 lr: 1.6000000000000003e-05 evaluation reward: 4.13  
 episode: 1152 score: 5.0 memory length: 240532 epsilon: 0.5241406600103304 s  
 teps: 344 lr: 1.6000000000000003e-05 evaluation reward: 4.14  
 episode: 1153 score: 5.0 memory length: 240846 epsilon: 0.5235189400103439 s  
 teps: 314 lr: 1.6000000000000003e-05 evaluation reward: 4.16  
 episode: 1154 score: 5.0 memory length: 241197 epsilon: 0.522823960010359 st  
 eps: 351 lr: 1.6000000000000003e-05 evaluation reward: 4.12  
 episode: 1155 score: 4.0 memory length: 241458 epsilon: 0.5223071800103702 s  
 teps: 261 lr: 1.6000000000000003e-05 evaluation reward: 4.13  
 episode: 1156 score: 2.0 memory length: 241677 epsilon: 0.5218735600103797 s  
 teps: 219 lr: 1.6000000000000003e-05 evaluation reward: 4.07  
 episode: 1157 score: 6.0 memory length: 242025 epsilon: 0.5211845200103946 s  
 teps: 348 lr: 1.6000000000000003e-05 evaluation reward: 4.11  
 episode: 1158 score: 5.0 memory length: 242369 epsilon: 0.5205034000104094 s  
 teps: 344 lr: 1.6000000000000003e-05 evaluation reward: 4.12  
 episode: 1159 score: 7.0 memory length: 242770 epsilon: 0.5197094200104266 s  
 teps: 401 lr: 1.6000000000000003e-05 evaluation reward: 4.17  
 episode: 1160 score: 3.0 memory length: 242996 epsilon: 0.5192619400104364 s  
 teps: 226 lr: 1.6000000000000003e-05 evaluation reward: 4.13  
 episode: 1161 score: 5.0 memory length: 243321 epsilon: 0.5186184400104503 s  
 teps: 325 lr: 1.6000000000000003e-05 evaluation reward: 4.14  
 episode: 1162 score: 4.0 memory length: 243576 epsilon: 0.5181135400104613 s  
 teps: 255 lr: 1.6000000000000003e-05 evaluation reward: 4.15  
 episode: 1163 score: 1.0 memory length: 243746 epsilon: 0.5177769400104686 s  
 teps: 170 lr: 1.6000000000000003e-05 evaluation reward: 4.12  
 episode: 1164 score: 5.0 memory length: 244093 epsilon: 0.5170898800104835 s  
 teps: 347 lr: 1.6000000000000003e-05 evaluation reward: 4.17  
 episode: 1165 score: 5.0 memory length: 244404 epsilon: 0.5164741000104969 s  
 teps: 311 lr: 1.6000000000000003e-05 evaluation reward: 4.16  
 episode: 1166 score: 3.0 memory length: 244614 epsilon: 0.5160583000105059 s  
 teps: 210 lr: 1.6000000000000003e-05 evaluation reward: 4.14  
 episode: 1167 score: 4.0 memory length: 244888 epsilon: 0.5155157800105177 s  
 teps: 274 lr: 1.6000000000000003e-05 evaluation reward: 4.17  
 episode: 1168 score: 4.0 memory length: 245144 epsilon: 0.5150089000105287 s  
 teps: 256 lr: 1.6000000000000003e-05 evaluation reward: 4.18  
 episode: 1169 score: 5.0 memory length: 245472 epsilon: 0.5143594600105428 s  
 teps: 328 lr: 1.6000000000000003e-05 evaluation reward: 4.2

episode: 1170 score: 9.0 memory length: 245927 epsilon: 0.5134585600105623 s  
 teps: 455 lr: 1.6000000000000003e-05 evaluation reward: 4.25  
 episode: 1171 score: 3.0 memory length: 246140 epsilon: 0.5130368200105715 s  
 teps: 213 lr: 1.6000000000000003e-05 evaluation reward: 4.25  
 episode: 1172 score: 6.0 memory length: 246479 epsilon: 0.5123656000105861 s  
 teps: 339 lr: 1.6000000000000003e-05 evaluation reward: 4.28  
 episode: 1173 score: 5.0 memory length: 246788 epsilon: 0.5117537800105993 s  
 teps: 309 lr: 1.6000000000000003e-05 evaluation reward: 4.26  
 episode: 1174 score: 5.0 memory length: 247077 epsilon: 0.5111815600106118 s  
 teps: 289 lr: 1.6000000000000003e-05 evaluation reward: 4.27  
 episode: 1175 score: 2.0 memory length: 247259 epsilon: 0.5108212000106196 s  
 teps: 182 lr: 1.6000000000000003e-05 evaluation reward: 4.21  
 episode: 1176 score: 4.0 memory length: 247558 epsilon: 0.5102291800106324 s  
 teps: 299 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1177 score: 3.0 memory length: 247804 epsilon: 0.509742100010643 st  
 eps: 246 lr: 1.6000000000000003e-05 evaluation reward: 4.24  
 episode: 1178 score: 5.0 memory length: 248129 epsilon: 0.509098600010657 st  
 eps: 325 lr: 1.6000000000000003e-05 evaluation reward: 4.26  
 episode: 1179 score: 3.0 memory length: 248374 epsilon: 0.5086135000106675 s  
 teps: 245 lr: 1.6000000000000003e-05 evaluation reward: 4.24  
 episode: 1180 score: 7.0 memory length: 248759 epsilon: 0.5078512000106841 s  
 teps: 385 lr: 1.6000000000000003e-05 evaluation reward: 4.23  
 episode: 1181 score: 1.0 memory length: 248911 epsilon: 0.5075502400106906 s  
 teps: 152 lr: 1.6000000000000003e-05 evaluation reward: 4.21  
 episode: 1182 score: 7.0 memory length: 249277 epsilon: 0.5068255600107063 s  
 teps: 366 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1183 score: 6.0 memory length: 249632 epsilon: 0.5061226600107216 s  
 teps: 355 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1184 score: 2.0 memory length: 249815 epsilon: 0.5057603200107295 s  
 teps: 183 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1185 score: 4.0 memory length: 250074 epsilon: 0.5052475000107406 s  
 teps: 259 lr: 1.6000000000000003e-05 evaluation reward: 4.21  
 episode: 1186 score: 6.0 memory length: 250428 epsilon: 0.5045465800107558 s  
 teps: 354 lr: 1.6000000000000003e-05 evaluation reward: 4.23  
 episode: 1187 score: 1.0 memory length: 250580 epsilon: 0.5042456200107623 s  
 teps: 152 lr: 1.6000000000000003e-05 evaluation reward: 4.21  
 episode: 1188 score: 1.0 memory length: 250731 epsilon: 0.5039466400107688 s  
 teps: 151 lr: 1.6000000000000003e-05 evaluation reward: 4.18  
 episode: 1189 score: 5.0 memory length: 251027 epsilon: 0.5033605600107816 s  
 teps: 296 lr: 1.6000000000000003e-05 evaluation reward: 4.18  
 episode: 1190 score: 6.0 memory length: 251424 epsilon: 0.5025745000107986 s  
 teps: 397 lr: 1.6000000000000003e-05 evaluation reward: 4.21  
 episode: 1191 score: 4.0 memory length: 251723 epsilon: 0.5019824800108115 s  
 teps: 299 lr: 1.6000000000000003e-05 evaluation reward: 4.19  
 episode: 1192 score: 7.0 memory length: 252140 epsilon: 0.5011568200108294 s  
 teps: 417 lr: 1.6000000000000003e-05 evaluation reward: 4.23  
 episode: 1193 score: 2.0 memory length: 252338 epsilon: 0.5007647800108379 s  
 teps: 198 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1194 score: 3.0 memory length: 252587 epsilon: 0.5002717600108486 s  
 teps: 249 lr: 1.6000000000000003e-05 evaluation reward: 4.17  
 episode: 1195 score: 3.0 memory length: 252799 epsilon: 0.49985200001085356 s  
 steps: 212 lr: 1.6000000000000003e-05 evaluation reward: 4.15  
 episode: 1196 score: 8.0 memory length: 253233 epsilon: 0.4989926800108481 s  
 teps: 434 lr: 1.6000000000000003e-05 evaluation reward: 4.19  
 episode: 1197 score: 3.0 memory length: 253443 epsilon: 0.4985768800108455 s  
 teps: 210 lr: 1.6000000000000003e-05 evaluation reward: 4.21  
 episode: 1198 score: 6.0 memory length: 253813 epsilon: 0.49784428001084086 s  
 steps: 370 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1199 score: 9.0 memory length: 254286 epsilon: 0.49690774001083493 s  
 steps: 473 lr: 1.6000000000000003e-05 evaluation reward: 4.26

episode: 1200	score: 4.0	memory length: 254543	epsilon: 0.4963988800108317	s
teps: 257	lr: 1.6000000000000003e-05		evaluation reward: 4.27	
episode: 1201	score: 5.0	memory length: 254857	epsilon: 0.4957771600108278	s
teps: 314	lr: 1.6000000000000003e-05		evaluation reward: 4.28	
episode: 1202	score: 5.0	memory length: 255177	epsilon: 0.49514356001082377	
steps: 320	lr: 1.6000000000000003e-05		evaluation reward: 4.29	
episode: 1203	score: 6.0	memory length: 255553	epsilon: 0.49439908001081906	
steps: 376	lr: 1.6000000000000003e-05		evaluation reward: 4.28	
episode: 1204	score: 4.0	memory length: 255830	epsilon: 0.4938506200108156	s
teps: 277	lr: 1.6000000000000003e-05		evaluation reward: 4.3	
episode: 1205	score: 6.0	memory length: 256192	epsilon: 0.49313386001081105	
steps: 362	lr: 1.6000000000000003e-05		evaluation reward: 4.35	
episode: 1206	score: 1.0	memory length: 256343	epsilon: 0.49283488001080916	
steps: 151	lr: 1.6000000000000003e-05		evaluation reward: 4.35	
episode: 1207	score: 5.0	memory length: 256632	epsilon: 0.49226266001080554	
steps: 289	lr: 1.6000000000000003e-05		evaluation reward: 4.37	
episode: 1208	score: 4.0	memory length: 256909	epsilon: 0.49171420001080207	
steps: 277	lr: 1.6000000000000003e-05		evaluation reward: 4.4	
episode: 1209	score: 6.0	memory length: 257261	epsilon: 0.49101724001079766	
steps: 352	lr: 1.6000000000000003e-05		evaluation reward: 4.41	
episode: 1210	score: 6.0	memory length: 257602	epsilon: 0.4903420600107934	s
teps: 341	lr: 1.6000000000000003e-05		evaluation reward: 4.41	
episode: 1211	score: 4.0	memory length: 257881	epsilon: 0.4897896400107899	s
teps: 279	lr: 1.6000000000000003e-05		evaluation reward: 4.44	
episode: 1212	score: 2.0	memory length: 258082	epsilon: 0.4893916600107874	s
teps: 201	lr: 1.6000000000000003e-05		evaluation reward: 4.44	
episode: 1213	score: 4.0	memory length: 258363	epsilon: 0.48883528001078386	
steps: 281	lr: 1.6000000000000003e-05		evaluation reward: 4.46	
episode: 1214	score: 4.0	memory length: 258639	epsilon: 0.4882888000107804	s
teps: 276	lr: 1.6000000000000003e-05		evaluation reward: 4.47	
episode: 1215	score: 5.0	memory length: 258928	epsilon: 0.4877165800107768	s
teps: 289	lr: 1.6000000000000003e-05		evaluation reward: 4.47	
episode: 1216	score: 2.0	memory length: 259111	epsilon: 0.4873542400107745	s
teps: 183	lr: 1.6000000000000003e-05		evaluation reward: 4.42	
episode: 1217	score: 4.0	memory length: 259370	epsilon: 0.48684142001077124	
steps: 259	lr: 1.6000000000000003e-05		evaluation reward: 4.41	
episode: 1218	score: 2.0	memory length: 259551	epsilon: 0.486483040010769	st
eps: 181	lr: 1.6000000000000003e-05		evaluation reward: 4.39	
episode: 1219	score: 6.0	memory length: 259889	epsilon: 0.48581380001076474	
steps: 338	lr: 1.6000000000000003e-05		evaluation reward: 4.41	
episode: 1220	score: 3.0	memory length: 260115	epsilon: 0.4853663200107619	s
teps: 226	lr: 1.6000000000000003e-05		evaluation reward: 4.4	
episode: 1221	score: 3.0	memory length: 260364	epsilon: 0.4848733000107588	s
teps: 249	lr: 1.6000000000000003e-05		evaluation reward: 4.39	
episode: 1222	score: 2.0	memory length: 260563	epsilon: 0.4844792800107563	s
teps: 199	lr: 1.6000000000000003e-05		evaluation reward: 4.37	
episode: 1223	score: 3.0	memory length: 260794	epsilon: 0.4840219000107534	s
teps: 231	lr: 1.6000000000000003e-05		evaluation reward: 4.37	
episode: 1224	score: 7.0	memory length: 261221	epsilon: 0.48317644001074805	
steps: 427	lr: 1.6000000000000003e-05		evaluation reward: 4.42	
episode: 1225	score: 2.0	memory length: 261401	epsilon: 0.4828200400107458	s
teps: 180	lr: 1.6000000000000003e-05		evaluation reward: 4.37	
episode: 1226	score: 6.0	memory length: 261788	epsilon: 0.48205378001074095	
steps: 387	lr: 1.6000000000000003e-05		evaluation reward: 4.38	
episode: 1227	score: 6.0	memory length: 262118	epsilon: 0.4814003800107368	s
teps: 330	lr: 1.6000000000000003e-05		evaluation reward: 4.37	
episode: 1228	score: 3.0	memory length: 262387	epsilon: 0.48086776001073345	
steps: 269	lr: 1.6000000000000003e-05		evaluation reward: 4.36	
episode: 1229	score: 2.0	memory length: 262588	epsilon: 0.48046978001073093	
steps: 201	lr: 1.6000000000000003e-05		evaluation reward: 4.3	

episode: 1230 score: 5.0 memory length: 262862 epsilon: 0.4799272600107275 s  
 teps: 274 lr: 1.6000000000000003e-05 evaluation reward: 4.33  
 episode: 1231 score: 2.0 memory length: 263043 epsilon: 0.47956888001072523  
 steps: 181 lr: 1.6000000000000003e-05 evaluation reward: 4.32  
 episode: 1232 score: 4.0 memory length: 263319 epsilon: 0.47902240001072177  
 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 4.32  
 episode: 1233 score: 7.0 memory length: 263764 epsilon: 0.4781413000107162 s  
 teps: 445 lr: 1.6000000000000003e-05 evaluation reward: 4.36  
 episode: 1234 score: 3.0 memory length: 263997 epsilon: 0.4776799600107133 s  
 teps: 233 lr: 1.6000000000000003e-05 evaluation reward: 4.37  
 episode: 1235 score: 2.0 memory length: 264196 epsilon: 0.4772859400107108 s  
 teps: 199 lr: 1.6000000000000003e-05 evaluation reward: 4.38  
 episode: 1236 score: 3.0 memory length: 264441 epsilon: 0.4768008400107077 s  
 teps: 245 lr: 1.6000000000000003e-05 evaluation reward: 4.37  
 episode: 1237 score: 5.0 memory length: 264787 epsilon: 0.4761157600107034 s  
 teps: 346 lr: 1.6000000000000003e-05 evaluation reward: 4.37  
 episode: 1238 score: 2.0 memory length: 264968 epsilon: 0.4757573800107011 s  
 teps: 181 lr: 1.6000000000000003e-05 evaluation reward: 4.33  
 episode: 1239 score: 0.0 memory length: 265091 epsilon: 0.4755138400106996 s  
 teps: 123 lr: 1.6000000000000003e-05 evaluation reward: 4.25  
 episode: 1240 score: 7.0 memory length: 265476 epsilon: 0.47475154001069475  
 steps: 385 lr: 1.6000000000000003e-05 evaluation reward: 4.29  
 episode: 1241 score: 4.0 memory length: 265754 epsilon: 0.47420110001069127  
 steps: 278 lr: 1.6000000000000003e-05 evaluation reward: 4.31  
 episode: 1242 score: 2.0 memory length: 265936 epsilon: 0.473840740010689 st  
 eps: 182 lr: 1.6000000000000003e-05 evaluation reward: 4.29  
 episode: 1243 score: 4.0 memory length: 266211 epsilon: 0.47329624001068554  
 steps: 275 lr: 1.6000000000000003e-05 evaluation reward: 4.3  
 episode: 1244 score: 2.0 memory length: 266393 epsilon: 0.47293588001068326  
 steps: 182 lr: 1.6000000000000003e-05 evaluation reward: 4.28  
 episode: 1245 score: 6.0 memory length: 266768 epsilon: 0.47219338001067857  
 steps: 375 lr: 1.6000000000000003e-05 evaluation reward: 4.33  
 episode: 1246 score: 6.0 memory length: 267104 epsilon: 0.47152810001067436  
 steps: 336 lr: 1.6000000000000003e-05 evaluation reward: 4.31  
 episode: 1247 score: 5.0 memory length: 267427 epsilon: 0.4708885600106703 s  
 teps: 323 lr: 1.6000000000000003e-05 evaluation reward: 4.35  
 episode: 1248 score: 8.0 memory length: 267830 epsilon: 0.47009062001066526  
 steps: 403 lr: 1.6000000000000003e-05 evaluation reward: 4.38  
 episode: 1249 score: 4.0 memory length: 268105 epsilon: 0.4695461200106618 s  
 teps: 275 lr: 1.6000000000000003e-05 evaluation reward: 4.37  
 episode: 1250 score: 3.0 memory length: 268354 epsilon: 0.4690531000106587 s  
 teps: 249 lr: 1.6000000000000003e-05 evaluation reward: 4.27  
 episode: 1251 score: 2.0 memory length: 268573 epsilon: 0.46861948001065595  
 steps: 219 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1252 score: 3.0 memory length: 268785 epsilon: 0.4681997200106533 s  
 teps: 212 lr: 1.6000000000000003e-05 evaluation reward: 4.2  
 episode: 1253 score: 4.0 memory length: 269046 epsilon: 0.46768294001065003  
 steps: 261 lr: 1.6000000000000003e-05 evaluation reward: 4.19  
 episode: 1254 score: 1.0 memory length: 269197 epsilon: 0.46738396001064814  
 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 4.15  
 episode: 1255 score: 4.0 memory length: 269492 epsilon: 0.46679986001064444  
 steps: 295 lr: 1.6000000000000003e-05 evaluation reward: 4.15  
 episode: 1256 score: 3.0 memory length: 269722 epsilon: 0.46634446001064156  
 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 4.16  
 episode: 1257 score: 3.0 memory length: 269948 epsilon: 0.46589698001063873  
 steps: 226 lr: 1.6000000000000003e-05 evaluation reward: 4.13  
 episode: 1258 score: 8.0 memory length: 270401 epsilon: 0.46500004001063305  
 steps: 453 lr: 1.6000000000000003e-05 evaluation reward: 4.16  
 episode: 1259 score: 2.0 memory length: 270599 epsilon: 0.4646080000106306 s  
 teps: 198 lr: 1.6000000000000003e-05 evaluation reward: 4.11

episode: 1260 score: 6.0 memory length: 270935 epsilon: 0.46394272001062636  
 steps: 336 lr: 1.6000000000000003e-05 evaluation reward: 4.14  
 episode: 1261 score: 4.0 memory length: 271197 epsilon: 0.4634239600106231 s  
 teps: 262 lr: 1.6000000000000003e-05 evaluation reward: 4.13  
 episode: 1262 score: 3.0 memory length: 271445 epsilon: 0.46293292001062 ste  
 ps: 248 lr: 1.6000000000000003e-05 evaluation reward: 4.12  
 episode: 1263 score: 6.0 memory length: 271787 epsilon: 0.4622557600106157 s  
 teps: 342 lr: 1.6000000000000003e-05 evaluation reward: 4.17  
 episode: 1264 score: 6.0 memory length: 272180 epsilon: 0.46147762001061077  
 steps: 393 lr: 1.6000000000000003e-05 evaluation reward: 4.18  
 episode: 1265 score: 4.0 memory length: 272455 epsilon: 0.4609331200106073 s  
 teps: 275 lr: 1.6000000000000003e-05 evaluation reward: 4.17  
 episode: 1266 score: 8.0 memory length: 272897 epsilon: 0.4600579600106018 s  
 teps: 442 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1267 score: 4.0 memory length: 273154 epsilon: 0.45954910001059857  
 steps: 257 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1268 score: 4.0 memory length: 273413 epsilon: 0.4590362800105953 s  
 teps: 259 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1269 score: 3.0 memory length: 273661 epsilon: 0.4585452400105922 s  
 teps: 248 lr: 1.6000000000000003e-05 evaluation reward: 4.2  
 episode: 1270 score: 7.0 memory length: 274047 epsilon: 0.4577809600105874 s  
 teps: 386 lr: 1.6000000000000003e-05 evaluation reward: 4.18  
 episode: 1271 score: 0.0 memory length: 274170 epsilon: 0.45753742001058584  
 steps: 123 lr: 1.6000000000000003e-05 evaluation reward: 4.15  
 episode: 1272 score: 7.0 memory length: 274563 epsilon: 0.4567592800105809 s  
 teps: 393 lr: 1.6000000000000003e-05 evaluation reward: 4.16  
 episode: 1273 score: 6.0 memory length: 274924 epsilon: 0.4560445000105764 s  
 teps: 361 lr: 1.6000000000000003e-05 evaluation reward: 4.17  
 episode: 1274 score: 2.0 memory length: 275107 epsilon: 0.4556821600105741 s  
 teps: 183 lr: 1.6000000000000003e-05 evaluation reward: 4.14  
 episode: 1275 score: 2.0 memory length: 275289 epsilon: 0.4553218000105718 s  
 teps: 182 lr: 1.6000000000000003e-05 evaluation reward: 4.14  
 episode: 1276 score: 5.0 memory length: 275597 epsilon: 0.45471196001056796  
 steps: 308 lr: 1.6000000000000003e-05 evaluation reward: 4.15  
 episode: 1277 score: 1.0 memory length: 275766 epsilon: 0.45437734001056584  
 steps: 169 lr: 1.6000000000000003e-05 evaluation reward: 4.13  
 episode: 1278 score: 3.0 memory length: 275992 epsilon: 0.453929860010563 st  
 eps: 226 lr: 1.6000000000000003e-05 evaluation reward: 4.11  
 episode: 1279 score: 2.0 memory length: 276175 epsilon: 0.4535675200105607 s  
 teps: 183 lr: 1.6000000000000003e-05 evaluation reward: 4.1  
 episode: 1280 score: 4.0 memory length: 276434 epsilon: 0.4530547000105575 s  
 teps: 259 lr: 1.6000000000000003e-05 evaluation reward: 4.07  
 episode: 1281 score: 4.0 memory length: 276711 epsilon: 0.452506240010554 st  
 eps: 277 lr: 1.6000000000000003e-05 evaluation reward: 4.1  
 episode: 1282 score: 3.0 memory length: 276957 epsilon: 0.4520191600105509 s  
 teps: 246 lr: 1.6000000000000003e-05 evaluation reward: 4.06  
 episode: 1283 score: 6.0 memory length: 277279 epsilon: 0.4513816000105469 s  
 teps: 322 lr: 1.6000000000000003e-05 evaluation reward: 4.06  
 episode: 1284 score: 6.0 memory length: 277616 epsilon: 0.45071434001054267  
 steps: 337 lr: 1.6000000000000003e-05 evaluation reward: 4.1  
 episode: 1285 score: 2.0 memory length: 277816 epsilon: 0.45031834001054016  
 steps: 200 lr: 1.6000000000000003e-05 evaluation reward: 4.08  
 episode: 1286 score: 6.0 memory length: 278170 epsilon: 0.44961742001053573  
 steps: 354 lr: 1.6000000000000003e-05 evaluation reward: 4.08  
 episode: 1287 score: 3.0 memory length: 278382 epsilon: 0.4491976600105331 s  
 teps: 212 lr: 1.6000000000000003e-05 evaluation reward: 4.1  
 episode: 1288 score: 3.0 memory length: 278611 epsilon: 0.4487442400105302 s  
 teps: 229 lr: 1.6000000000000003e-05 evaluation reward: 4.12  
 episode: 1289 score: 5.0 memory length: 278959 epsilon: 0.44805520001052584  
 steps: 348 lr: 1.6000000000000003e-05 evaluation reward: 4.12

episode: 1290 score: 7.0 memory length: 279347 epsilon: 0.447286960010521 st  
 eps: 388 lr: 1.6000000000000003e-05 evaluation reward: 4.13  
 episode: 1291 score: 8.0 memory length: 279753 epsilon: 0.4464830800105159 s  
 teps: 406 lr: 1.6000000000000003e-05 evaluation reward: 4.17  
 episode: 1292 score: 2.0 memory length: 279934 epsilon: 0.44612470001051363  
 steps: 181 lr: 1.6000000000000003e-05 evaluation reward: 4.12  
 episode: 1293 score: 5.0 memory length: 280225 epsilon: 0.44554852001051 st  
 ps: 291 lr: 1.6000000000000003e-05 evaluation reward: 4.15  
 episode: 1294 score: 3.0 memory length: 280453 epsilon: 0.44509708001050713  
 steps: 228 lr: 1.6000000000000003e-05 evaluation reward: 4.15  
 episode: 1295 score: 6.0 memory length: 280809 epsilon: 0.44439220001050267  
 steps: 356 lr: 1.6000000000000003e-05 evaluation reward: 4.18  
 episode: 1296 score: 3.0 memory length: 281056 epsilon: 0.4439031400104996 s  
 teps: 247 lr: 1.6000000000000003e-05 evaluation reward: 4.13  
 episode: 1297 score: 5.0 memory length: 281401 epsilon: 0.44322004001049525  
 steps: 345 lr: 1.6000000000000003e-05 evaluation reward: 4.15  
 episode: 1298 score: 4.0 memory length: 281694 epsilon: 0.4426399000104916 s  
 teps: 293 lr: 1.6000000000000003e-05 evaluation reward: 4.13  
 episode: 1299 score: 4.0 memory length: 281975 epsilon: 0.44208352001048806  
 steps: 281 lr: 1.6000000000000003e-05 evaluation reward: 4.08  
 episode: 1300 score: 4.0 memory length: 282236 epsilon: 0.4415667400104848 s  
 teps: 261 lr: 1.6000000000000003e-05 evaluation reward: 4.08  
 episode: 1301 score: 9.0 memory length: 282708 epsilon: 0.4406321800104789 s  
 teps: 472 lr: 1.6000000000000003e-05 evaluation reward: 4.12  
 episode: 1302 score: 5.0 memory length: 282991 epsilon: 0.44007184001047533  
 steps: 283 lr: 1.6000000000000003e-05 evaluation reward: 4.12  
 episode: 1303 score: 3.0 memory length: 283238 epsilon: 0.43958278001047224  
 steps: 247 lr: 1.6000000000000003e-05 evaluation reward: 4.09  
 episode: 1304 score: 6.0 memory length: 283637 epsilon: 0.43879276001046724  
 steps: 399 lr: 1.6000000000000003e-05 evaluation reward: 4.11  
 episode: 1305 score: 5.0 memory length: 283957 epsilon: 0.43815916001046323  
 steps: 320 lr: 1.6000000000000003e-05 evaluation reward: 4.1  
 episode: 1306 score: 2.0 memory length: 284176 epsilon: 0.4377255400104605 s  
 teps: 219 lr: 1.6000000000000003e-05 evaluation reward: 4.11  
 episode: 1307 score: 6.0 memory length: 284549 epsilon: 0.4369870000104558 s  
 teps: 373 lr: 1.6000000000000003e-05 evaluation reward: 4.12  
 episode: 1308 score: 3.0 memory length: 284798 epsilon: 0.4364939800104527 s  
 teps: 249 lr: 1.6000000000000003e-05 evaluation reward: 4.11  
 episode: 1309 score: 3.0 memory length: 285027 epsilon: 0.43604056001044983  
 steps: 229 lr: 1.6000000000000003e-05 evaluation reward: 4.08  
 episode: 1310 score: 4.0 memory length: 285269 epsilon: 0.4355614000104468 s  
 teps: 242 lr: 1.6000000000000003e-05 evaluation reward: 4.06  
 episode: 1311 score: 14.0 memory length: 285723 epsilon: 0.4346624800104411  
 steps: 454 lr: 1.6000000000000003e-05 evaluation reward: 4.16  
 episode: 1312 score: 4.0 memory length: 285965 epsilon: 0.4341833200104381 s  
 teps: 242 lr: 1.6000000000000003e-05 evaluation reward: 4.18  
 episode: 1313 score: 3.0 memory length: 286175 epsilon: 0.43376752001043545  
 steps: 210 lr: 1.6000000000000003e-05 evaluation reward: 4.17  
 episode: 1314 score: 6.0 memory length: 286530 epsilon: 0.433064620010431 st  
 eps: 355 lr: 1.6000000000000003e-05 evaluation reward: 4.19  
 episode: 1315 score: 8.0 memory length: 287021 epsilon: 0.43209244001042485  
 steps: 491 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1316 score: 3.0 memory length: 287234 epsilon: 0.4316707000104222 s  
 teps: 213 lr: 1.6000000000000003e-05 evaluation reward: 4.23  
 episode: 1317 score: 3.0 memory length: 287501 epsilon: 0.43114204001041884  
 steps: 267 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1318 score: 2.0 memory length: 287699 epsilon: 0.43075000001041636  
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 4.22  
 episode: 1319 score: 6.0 memory length: 288058 epsilon: 0.43003918001041186  
 steps: 359 lr: 1.6000000000000003e-05 evaluation reward: 4.22

episode: 1320 score: 6.0 memory length: 288423 epsilon: 0.4293164800104073 s  
 teps: 365 lr: 1.6000000000000003e-05 evaluation reward: 4.25  
 episode: 1321 score: 3.0 memory length: 288654 epsilon: 0.4288591000104044 s  
 teps: 231 lr: 1.6000000000000003e-05 evaluation reward: 4.25  
 episode: 1322 score: 6.0 memory length: 289007 epsilon: 0.42816016001039997  
 steps: 353 lr: 1.6000000000000003e-05 evaluation reward: 4.29  
 episode: 1323 score: 6.0 memory length: 289323 epsilon: 0.427534480010396 st  
 eps: 316 lr: 1.6000000000000003e-05 evaluation reward: 4.32  
 episode: 1324 score: 6.0 memory length: 289673 epsilon: 0.4268414800103916 s  
 teps: 350 lr: 1.6000000000000003e-05 evaluation reward: 4.31  
 episode: 1325 score: 7.0 memory length: 290080 epsilon: 0.42603562001038653  
 steps: 407 lr: 1.6000000000000003e-05 evaluation reward: 4.36  
 episode: 1326 score: 6.0 memory length: 290416 epsilon: 0.4253703400103823 s  
 teps: 336 lr: 1.6000000000000003e-05 evaluation reward: 4.36  
 episode: 1327 score: 4.0 memory length: 290693 epsilon: 0.42482188001037885  
 steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 4.34  
 episode: 1328 score: 4.0 memory length: 290972 epsilon: 0.42426946001037535  
 steps: 279 lr: 1.6000000000000003e-05 evaluation reward: 4.35  
 episode: 1329 score: 7.0 memory length: 291340 epsilon: 0.42354082001037074  
 steps: 368 lr: 1.6000000000000003e-05 evaluation reward: 4.4  
 episode: 1330 score: 4.0 memory length: 291580 epsilon: 0.42306562001036774  
 steps: 240 lr: 1.6000000000000003e-05 evaluation reward: 4.39  
 episode: 1331 score: 8.0 memory length: 292021 epsilon: 0.4221924400103622 s  
 teps: 441 lr: 1.6000000000000003e-05 evaluation reward: 4.45  
 episode: 1332 score: 6.0 memory length: 292376 epsilon: 0.42148954001035777  
 steps: 355 lr: 1.6000000000000003e-05 evaluation reward: 4.47  
 episode: 1333 score: 4.0 memory length: 292652 epsilon: 0.4209430600103543 s  
 teps: 276 lr: 1.6000000000000003e-05 evaluation reward: 4.44  
 episode: 1334 score: 6.0 memory length: 292988 epsilon: 0.4202777800103501 s  
 teps: 336 lr: 1.6000000000000003e-05 evaluation reward: 4.47  
 episode: 1335 score: 4.0 memory length: 293248 epsilon: 0.41976298001034684  
 steps: 260 lr: 1.6000000000000003e-05 evaluation reward: 4.49  
 episode: 1336 score: 6.0 memory length: 293641 epsilon: 0.4189848400103419 s  
 teps: 393 lr: 1.6000000000000003e-05 evaluation reward: 4.52  
 episode: 1337 score: 4.0 memory length: 293883 epsilon: 0.4185056800103389 s  
 teps: 242 lr: 1.6000000000000003e-05 evaluation reward: 4.51  
 episode: 1338 score: 7.0 memory length: 294256 epsilon: 0.4177671400103342 s  
 teps: 373 lr: 1.6000000000000003e-05 evaluation reward: 4.56  
 episode: 1339 score: 8.0 memory length: 294678 epsilon: 0.41693158001032893  
 steps: 422 lr: 1.6000000000000003e-05 evaluation reward: 4.64  
 episode: 1340 score: 15.0 memory length: 295108 epsilon: 0.41608018001032354  
 steps: 430 lr: 1.6000000000000003e-05 evaluation reward: 4.72  
 episode: 1341 score: 5.0 memory length: 295400 epsilon: 0.4155020200103199 s  
 teps: 292 lr: 1.6000000000000003e-05 evaluation reward: 4.73  
 episode: 1342 score: 5.0 memory length: 295689 epsilon: 0.41492980001031626  
 steps: 289 lr: 1.6000000000000003e-05 evaluation reward: 4.76  
 episode: 1343 score: 3.0 memory length: 295915 epsilon: 0.41448232001031343  
 steps: 226 lr: 1.6000000000000003e-05 evaluation reward: 4.75  
 episode: 1344 score: 8.0 memory length: 296335 epsilon: 0.41365072001030817  
 steps: 420 lr: 1.6000000000000003e-05 evaluation reward: 4.81  
 episode: 1345 score: 5.0 memory length: 296657 epsilon: 0.41301316001030414  
 steps: 322 lr: 1.6000000000000003e-05 evaluation reward: 4.8  
 episode: 1346 score: 2.0 memory length: 296837 epsilon: 0.4126567600103019 s  
 teps: 180 lr: 1.6000000000000003e-05 evaluation reward: 4.76  
 episode: 1347 score: 9.0 memory length: 297310 epsilon: 0.41172022001029596  
 steps: 473 lr: 1.6000000000000003e-05 evaluation reward: 4.8  
 episode: 1348 score: 6.0 memory length: 297703 epsilon: 0.41094208001029103  
 steps: 393 lr: 1.6000000000000003e-05 evaluation reward: 4.78  
 episode: 1349 score: 7.0 memory length: 298109 epsilon: 0.41013820001028595  
 steps: 406 lr: 1.6000000000000003e-05 evaluation reward: 4.81

episode: 1350 score: 5.0 memory length: 298437 epsilon: 0.40948876001028184  
 steps: 328 lr: 1.6000000000000003e-05 evaluation reward: 4.83  
 episode: 1351 score: 6.0 memory length: 298792 epsilon: 0.4087858600102774 s  
 teps: 355 lr: 1.6000000000000003e-05 evaluation reward: 4.87  
 episode: 1352 score: 4.0 memory length: 299068 epsilon: 0.40823938001027393  
 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 4.88  
 episode: 1353 score: 3.0 memory length: 299314 epsilon: 0.40775230001027085  
 steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 4.87  
 episode: 1354 score: 6.0 memory length: 299669 epsilon: 0.4070494000102664 s  
 teps: 355 lr: 1.6000000000000003e-05 evaluation reward: 4.92  
 episode: 1355 score: 5.0 memory length: 299957 epsilon: 0.4064791600102628 s  
 teps: 288 lr: 1.6000000000000003e-05 evaluation reward: 4.93  
 episode: 1356 score: 4.0 memory length: 300215 epsilon: 0.40596832001025956  
 steps: 258 lr: 6.400000000000001e-06 evaluation reward: 4.94  
 episode: 1357 score: 5.0 memory length: 300506 epsilon: 0.4053921400102559 s  
 teps: 291 lr: 6.400000000000001e-06 evaluation reward: 4.96  
 episode: 1358 score: 3.0 memory length: 300735 epsilon: 0.40493872001025305  
 steps: 229 lr: 6.400000000000001e-06 evaluation reward: 4.91  
 episode: 1359 score: 5.0 memory length: 301058 epsilon: 0.404299180010249 st  
 eps: 323 lr: 6.400000000000001e-06 evaluation reward: 4.94  
 episode: 1360 score: 6.0 memory length: 301373 epsilon: 0.40367548001024506  
 steps: 315 lr: 6.400000000000001e-06 evaluation reward: 4.94  
 episode: 1361 score: 9.0 memory length: 301815 epsilon: 0.4028003200102395 s  
 teps: 442 lr: 6.400000000000001e-06 evaluation reward: 4.99  
 episode: 1362 score: 3.0 memory length: 302045 epsilon: 0.40234492001023664  
 steps: 230 lr: 6.400000000000001e-06 evaluation reward: 4.99  
 episode: 1363 score: 2.0 memory length: 302228 epsilon: 0.40198258001023435  
 steps: 183 lr: 6.400000000000001e-06 evaluation reward: 4.95  
 episode: 1364 score: 2.0 memory length: 302428 epsilon: 0.40158658001023184  
 steps: 200 lr: 6.400000000000001e-06 evaluation reward: 4.91  
 episode: 1365 score: 7.0 memory length: 302814 epsilon: 0.400822300010227 st  
 eps: 386 lr: 6.400000000000001e-06 evaluation reward: 4.94  
 episode: 1366 score: 4.0 memory length: 303109 epsilon: 0.4002382000102233 s  
 teps: 295 lr: 6.400000000000001e-06 evaluation reward: 4.9  
 episode: 1367 score: 4.0 memory length: 303405 epsilon: 0.3996521200102196 s  
 teps: 296 lr: 6.400000000000001e-06 evaluation reward: 4.9  
 episode: 1368 score: 6.0 memory length: 303742 epsilon: 0.3989848600102154 s  
 teps: 337 lr: 6.400000000000001e-06 evaluation reward: 4.92  
 episode: 1369 score: 9.0 memory length: 304207 epsilon: 0.39806416001020956  
 steps: 465 lr: 6.400000000000001e-06 evaluation reward: 4.98  
 episode: 1370 score: 5.0 memory length: 304533 epsilon: 0.39741868001020547  
 steps: 326 lr: 6.400000000000001e-06 evaluation reward: 4.96  
 episode: 1371 score: 5.0 memory length: 304860 epsilon: 0.3967712200102014 s  
 teps: 327 lr: 6.400000000000001e-06 evaluation reward: 5.01  
 episode: 1372 score: 3.0 memory length: 305073 epsilon: 0.3963494800101987 s  
 teps: 213 lr: 6.400000000000001e-06 evaluation reward: 4.97  
 episode: 1373 score: 9.0 memory length: 305520 epsilon: 0.3954644200101931 s  
 teps: 447 lr: 6.400000000000001e-06 evaluation reward: 5.0  
 episode: 1374 score: 8.0 memory length: 305942 epsilon: 0.3946288600101878 s  
 teps: 422 lr: 6.400000000000001e-06 evaluation reward: 5.06  
 episode: 1375 score: 5.0 memory length: 306252 epsilon: 0.39401506001018394  
 steps: 310 lr: 6.400000000000001e-06 evaluation reward: 5.09  
 episode: 1376 score: 4.0 memory length: 306508 epsilon: 0.39350818001018073  
 steps: 256 lr: 6.400000000000001e-06 evaluation reward: 5.08  
 episode: 1377 score: 8.0 memory length: 306902 epsilon: 0.3927280600101758 s  
 teps: 394 lr: 6.400000000000001e-06 evaluation reward: 5.15  
 episode: 1378 score: 14.0 memory length: 307330 epsilon: 0.39188062001017043  
 steps: 428 lr: 6.400000000000001e-06 evaluation reward: 5.26  
 episode: 1379 score: 5.0 memory length: 307654 epsilon: 0.3912391000101664 s  
 teps: 324 lr: 6.400000000000001e-06 evaluation reward: 5.29

episode: 1380 score: 6.0 memory length: 308034 epsilon: 0.3904867000101616 s  
 teps: 380 lr: 6.400000000000001e-06 evaluation reward: 5.31  
 episode: 1381 score: 4.0 memory length: 308297 epsilon: 0.3899659600101583 s  
 teps: 263 lr: 6.400000000000001e-06 evaluation reward: 5.31  
 episode: 1382 score: 7.0 memory length: 308701 epsilon: 0.38916604001015326  
 steps: 404 lr: 6.400000000000001e-06 evaluation reward: 5.35  
 episode: 1383 score: 4.0 memory length: 308999 epsilon: 0.3885760000101495 s  
 teps: 298 lr: 6.400000000000001e-06 evaluation reward: 5.33  
 episode: 1384 score: 4.0 memory length: 309242 epsilon: 0.3880948600101465 s  
 teps: 243 lr: 6.400000000000001e-06 evaluation reward: 5.31  
 episode: 1385 score: 5.0 memory length: 309550 epsilon: 0.3874850200101426 s  
 teps: 308 lr: 6.400000000000001e-06 evaluation reward: 5.34  
 episode: 1386 score: 4.0 memory length: 309807 epsilon: 0.3869761600101394 s  
 teps: 257 lr: 6.400000000000001e-06 evaluation reward: 5.32  
 episode: 1387 score: 7.0 memory length: 310218 epsilon: 0.38616238001013425  
 steps: 411 lr: 6.400000000000001e-06 evaluation reward: 5.36  
 episode: 1388 score: 6.0 memory length: 310590 epsilon: 0.3854258200101296 s  
 teps: 372 lr: 6.400000000000001e-06 evaluation reward: 5.39  
 episode: 1389 score: 5.0 memory length: 310897 epsilon: 0.38481796001012575  
 steps: 307 lr: 6.400000000000001e-06 evaluation reward: 5.39  
 episode: 1390 score: 7.0 memory length: 311314 epsilon: 0.3839923000101205 s  
 teps: 417 lr: 6.400000000000001e-06 evaluation reward: 5.39  
 episode: 1391 score: 3.0 memory length: 311543 epsilon: 0.38353888001011766  
 steps: 229 lr: 6.400000000000001e-06 evaluation reward: 5.34  
 episode: 1392 score: 6.0 memory length: 311915 epsilon: 0.382802320010113 st  
 eps: 372 lr: 6.400000000000001e-06 evaluation reward: 5.38  
 episode: 1393 score: 5.0 memory length: 312223 epsilon: 0.38219248001010914  
 steps: 308 lr: 6.400000000000001e-06 evaluation reward: 5.38  
 episode: 1394 score: 3.0 memory length: 312467 epsilon: 0.3817093600101061 s  
 teps: 244 lr: 6.400000000000001e-06 evaluation reward: 5.38  
 episode: 1395 score: 5.0 memory length: 312773 epsilon: 0.38110348001010225  
 steps: 306 lr: 6.400000000000001e-06 evaluation reward: 5.37  
 episode: 1396 score: 3.0 memory length: 313000 epsilon: 0.3806540200100994 s  
 teps: 227 lr: 6.400000000000001e-06 evaluation reward: 5.37  
 episode: 1397 score: 3.0 memory length: 313228 epsilon: 0.38020258001009655  
 steps: 228 lr: 6.400000000000001e-06 evaluation reward: 5.35  
 episode: 1398 score: 7.0 memory length: 313635 epsilon: 0.37939672001009145  
 steps: 407 lr: 6.400000000000001e-06 evaluation reward: 5.38  
 episode: 1399 score: 4.0 memory length: 313892 epsilon: 0.37888786001008823  
 steps: 257 lr: 6.400000000000001e-06 evaluation reward: 5.38  
 episode: 1400 score: 5.0 memory length: 314184 epsilon: 0.37830970001008457  
 steps: 292 lr: 6.400000000000001e-06 evaluation reward: 5.39  
 episode: 1401 score: 3.0 memory length: 314413 epsilon: 0.3778562800100817 s  
 teps: 229 lr: 6.400000000000001e-06 evaluation reward: 5.33  
 episode: 1402 score: 7.0 memory length: 314799 epsilon: 0.37709200001007687  
 steps: 386 lr: 6.400000000000001e-06 evaluation reward: 5.35  
 episode: 1403 score: 5.0 memory length: 315105 epsilon: 0.37648612001007303  
 steps: 306 lr: 6.400000000000001e-06 evaluation reward: 5.37  
 episode: 1404 score: 3.0 memory length: 315318 epsilon: 0.37606438001007036  
 steps: 213 lr: 6.400000000000001e-06 evaluation reward: 5.34  
 episode: 1405 score: 7.0 memory length: 315743 epsilon: 0.37522288001006504  
 steps: 425 lr: 6.400000000000001e-06 evaluation reward: 5.36  
 episode: 1406 score: 7.0 memory length: 316152 epsilon: 0.3744130600100599 s  
 teps: 409 lr: 6.400000000000001e-06 evaluation reward: 5.41  
 episode: 1407 score: 5.0 memory length: 316460 epsilon: 0.37380322001005606  
 steps: 308 lr: 6.400000000000001e-06 evaluation reward: 5.4  
 episode: 1408 score: 5.0 memory length: 316764 epsilon: 0.37320130001005225  
 steps: 304 lr: 6.400000000000001e-06 evaluation reward: 5.42  
 episode: 1409 score: 3.0 memory length: 316975 epsilon: 0.3727835200100496 s  
 teps: 211 lr: 6.400000000000001e-06 evaluation reward: 5.42

episode: 1410 score: 4.0 memory length: 317251 epsilon: 0.37223704001004615  
 steps: 276 lr: 6.400000000000000e-06 evaluation reward: 5.42  
 episode: 1411 score: 6.0 memory length: 317623 epsilon: 0.3715004800100415 s  
 teps: 372 lr: 6.400000000000000e-06 evaluation reward: 5.34  
 episode: 1412 score: 3.0 memory length: 317850 epsilon: 0.37105102001003865  
 steps: 227 lr: 6.400000000000000e-06 evaluation reward: 5.33  
 episode: 1413 score: 7.0 memory length: 318257 epsilon: 0.37024516001003355  
 steps: 407 lr: 6.400000000000000e-06 evaluation reward: 5.37  
 episode: 1414 score: 7.0 memory length: 318632 epsilon: 0.36950266001002885  
 steps: 375 lr: 6.400000000000000e-06 evaluation reward: 5.38  
 episode: 1415 score: 3.0 memory length: 318862 epsilon: 0.36904726001002597  
 steps: 230 lr: 6.400000000000000e-06 evaluation reward: 5.33  
 episode: 1416 score: 7.0 memory length: 319284 epsilon: 0.3682117000100207 s  
 teps: 422 lr: 6.400000000000000e-06 evaluation reward: 5.37  
 episode: 1417 score: 4.0 memory length: 319583 epsilon: 0.36761968001001694  
 steps: 299 lr: 6.400000000000000e-06 evaluation reward: 5.38  
 episode: 1418 score: 4.0 memory length: 319878 epsilon: 0.36703558001001324  
 steps: 295 lr: 6.400000000000000e-06 evaluation reward: 5.4  
 episode: 1419 score: 4.0 memory length: 320158 epsilon: 0.36648118001000973  
 steps: 280 lr: 6.400000000000000e-06 evaluation reward: 5.38  
 episode: 1420 score: 4.0 memory length: 320418 epsilon: 0.3659663800100065 s  
 teps: 260 lr: 6.400000000000000e-06 evaluation reward: 5.36  
 episode: 1421 score: 5.0 memory length: 320744 epsilon: 0.3653209000100024 s  
 teps: 326 lr: 6.400000000000000e-06 evaluation reward: 5.38  
 episode: 1422 score: 3.0 memory length: 320974 epsilon: 0.364865500009995 s  
 teps: 230 lr: 6.400000000000000e-06 evaluation reward: 5.35  
 episode: 1423 score: 8.0 memory length: 321404 epsilon: 0.3640141000099941 s  
 teps: 430 lr: 6.400000000000000e-06 evaluation reward: 5.37  
 episode: 1424 score: 4.0 memory length: 321650 epsilon: 0.36352702000999104  
 steps: 246 lr: 6.400000000000000e-06 evaluation reward: 5.35  
 episode: 1425 score: 6.0 memory length: 322006 epsilon: 0.3628221400099866 s  
 teps: 356 lr: 6.400000000000000e-06 evaluation reward: 5.34  
 episode: 1426 score: 4.0 memory length: 322282 epsilon: 0.3622756600099831 s  
 teps: 276 lr: 6.400000000000000e-06 evaluation reward: 5.32  
 episode: 1427 score: 3.0 memory length: 322508 epsilon: 0.3618281800099803 s  
 teps: 226 lr: 6.400000000000000e-06 evaluation reward: 5.31  
 episode: 1428 score: 6.0 memory length: 322853 epsilon: 0.36114508000997597  
 steps: 345 lr: 6.400000000000000e-06 evaluation reward: 5.33  
 episode: 1429 score: 5.0 memory length: 323180 epsilon: 0.3604976200099719 s  
 teps: 327 lr: 6.400000000000000e-06 evaluation reward: 5.31  
 episode: 1430 score: 6.0 memory length: 323516 epsilon: 0.35983234000996767  
 steps: 336 lr: 6.400000000000000e-06 evaluation reward: 5.33  
 episode: 1431 score: 5.0 memory length: 323827 epsilon: 0.35921656000996377  
 steps: 311 lr: 6.400000000000000e-06 evaluation reward: 5.3  
 episode: 1432 score: 4.0 memory length: 324069 epsilon: 0.35873740000996074  
 steps: 242 lr: 6.400000000000000e-06 evaluation reward: 5.28  
 episode: 1433 score: 3.0 memory length: 324279 epsilon: 0.3583216000099581 s  
 teps: 210 lr: 6.400000000000000e-06 evaluation reward: 5.27  
 episode: 1434 score: 5.0 memory length: 324608 epsilon: 0.357670180009954 st  
 eps: 329 lr: 6.400000000000000e-06 evaluation reward: 5.26  
 episode: 1435 score: 7.0 memory length: 324977 epsilon: 0.35693956000994936  
 steps: 369 lr: 6.400000000000000e-06 evaluation reward: 5.29  
 episode: 1436 score: 9.0 memory length: 325435 epsilon: 0.3560327200099436 s  
 teps: 458 lr: 6.400000000000000e-06 evaluation reward: 5.32  
 episode: 1437 score: 7.0 memory length: 325782 epsilon: 0.3553456600099393 s  
 teps: 347 lr: 6.400000000000000e-06 evaluation reward: 5.35  
 episode: 1438 score: 4.0 memory length: 326059 epsilon: 0.3547972000099358 s  
 teps: 277 lr: 6.400000000000000e-06 evaluation reward: 5.32  
 episode: 1439 score: 4.0 memory length: 326319 epsilon: 0.35428240000993255  
 steps: 260 lr: 6.400000000000000e-06 evaluation reward: 5.28

episode: 1440 score: 3.0 memory length: 326532 epsilon: 0.3538606600099299 s  
 teps: 213 lr: 6.400000000000001e-06 evaluation reward: 5.16  
 episode: 1441 score: 5.0 memory length: 326821 epsilon: 0.35328844000992626  
 steps: 289 lr: 6.400000000000001e-06 evaluation reward: 5.16  
 episode: 1442 score: 2.0 memory length: 327021 epsilon: 0.35289244000992376  
 steps: 200 lr: 6.400000000000001e-06 evaluation reward: 5.13  
 episode: 1443 score: 4.0 memory length: 327298 epsilon: 0.3523439800099203 s  
 teps: 277 lr: 6.400000000000001e-06 evaluation reward: 5.14  
 episode: 1444 score: 4.0 memory length: 327592 epsilon: 0.3517618600099166 s  
 teps: 294 lr: 6.400000000000001e-06 evaluation reward: 5.1  
 episode: 1445 score: 4.0 memory length: 327865 epsilon: 0.3512213200099132 s  
 teps: 273 lr: 6.400000000000001e-06 evaluation reward: 5.09  
 episode: 1446 score: 4.0 memory length: 328122 epsilon: 0.35071246000990997  
 steps: 257 lr: 6.400000000000001e-06 evaluation reward: 5.11  
 episode: 1447 score: 4.0 memory length: 328398 epsilon: 0.3501659800099065 s  
 teps: 276 lr: 6.400000000000001e-06 evaluation reward: 5.06  
 episode: 1448 score: 4.0 memory length: 328657 epsilon: 0.34965316000990326  
 steps: 259 lr: 6.400000000000001e-06 evaluation reward: 5.04  
 episode: 1449 score: 8.0 memory length: 329091 epsilon: 0.3487938400098978 s  
 teps: 434 lr: 6.400000000000001e-06 evaluation reward: 5.05  
 episode: 1450 score: 6.0 memory length: 329447 epsilon: 0.34808896000989337  
 steps: 356 lr: 6.400000000000001e-06 evaluation reward: 5.06  
 episode: 1451 score: 4.0 memory length: 329704 epsilon: 0.34758010000989015  
 steps: 257 lr: 6.400000000000001e-06 evaluation reward: 5.04  
 episode: 1452 score: 3.0 memory length: 329917 epsilon: 0.3471583600098875 s  
 teps: 213 lr: 6.400000000000001e-06 evaluation reward: 5.03  
 episode: 1453 score: 8.0 memory length: 330350 epsilon: 0.34630102000988205  
 steps: 433 lr: 6.400000000000001e-06 evaluation reward: 5.08  
 episode: 1454 score: 6.0 memory length: 330712 epsilon: 0.3455842600098775 s  
 teps: 362 lr: 6.400000000000001e-06 evaluation reward: 5.08  
 episode: 1455 score: 7.0 memory length: 331098 epsilon: 0.3448199800098727 s  
 teps: 386 lr: 6.400000000000001e-06 evaluation reward: 5.1  
 episode: 1456 score: 5.0 memory length: 331406 epsilon: 0.3442101400098688 s  
 teps: 308 lr: 6.400000000000001e-06 evaluation reward: 5.11  
 episode: 1457 score: 9.0 memory length: 331905 epsilon: 0.3432221200098626 s  
 teps: 499 lr: 6.400000000000001e-06 evaluation reward: 5.15  
 episode: 1458 score: 10.0 memory length: 332411 epsilon: 0.34222024000985624  
 steps: 506 lr: 6.400000000000001e-06 evaluation reward: 5.22  
 episode: 1459 score: 3.0 memory length: 332641 epsilon: 0.34176484000985335  
 steps: 230 lr: 6.400000000000001e-06 evaluation reward: 5.2  
 episode: 1460 score: 7.0 memory length: 333046 epsilon: 0.3409629400098483 s  
 teps: 405 lr: 6.400000000000001e-06 evaluation reward: 5.21  
 episode: 1461 score: 8.0 memory length: 333488 epsilon: 0.34008778000984274  
 steps: 442 lr: 6.400000000000001e-06 evaluation reward: 5.2  
 episode: 1462 score: 7.0 memory length: 333872 epsilon: 0.33932746000983793  
 steps: 384 lr: 6.400000000000001e-06 evaluation reward: 5.24  
 episode: 1463 score: 4.0 memory length: 334149 epsilon: 0.33877900000983446  
 steps: 277 lr: 6.400000000000001e-06 evaluation reward: 5.26  
 episode: 1464 score: 4.0 memory length: 334404 epsilon: 0.33827410000983127  
 steps: 255 lr: 6.400000000000001e-06 evaluation reward: 5.28  
 episode: 1465 score: 7.0 memory length: 334790 epsilon: 0.33750982000982643  
 steps: 386 lr: 6.400000000000001e-06 evaluation reward: 5.28  
 episode: 1466 score: 5.0 memory length: 335095 epsilon: 0.3369059200098226 s  
 teps: 305 lr: 6.400000000000001e-06 evaluation reward: 5.29  
 episode: 1467 score: 4.0 memory length: 335356 epsilon: 0.33638914000981934  
 steps: 261 lr: 6.400000000000001e-06 evaluation reward: 5.29  
 episode: 1468 score: 5.0 memory length: 335686 epsilon: 0.3357357400098152 s  
 teps: 330 lr: 6.400000000000001e-06 evaluation reward: 5.28  
 episode: 1469 score: 4.0 memory length: 335985 epsilon: 0.33514372000981146  
 steps: 299 lr: 6.400000000000001e-06 evaluation reward: 5.23

episode: 1470 score: 5.0 memory length: 336314 epsilon: 0.33449230000980734  
 steps: 329 lr: 6.400000000000000e-06 evaluation reward: 5.23  
 episode: 1471 score: 7.0 memory length: 336741 epsilon: 0.333646840009802 st  
 eps: 427 lr: 6.400000000000000e-06 evaluation reward: 5.25  
 episode: 1472 score: 6.0 memory length: 337113 epsilon: 0.33291028000979733  
 steps: 372 lr: 6.400000000000000e-06 evaluation reward: 5.28  
 episode: 1473 score: 4.0 memory length: 337372 epsilon: 0.3323974600097941 s  
 teps: 259 lr: 6.400000000000000e-06 evaluation reward: 5.23  
 episode: 1474 score: 12.0 memory length: 337719 epsilon: 0.33171040000978974  
 steps: 347 lr: 6.400000000000000e-06 evaluation reward: 5.27  
 episode: 1475 score: 4.0 memory length: 337978 epsilon: 0.3311975800097865 s  
 teps: 259 lr: 6.400000000000000e-06 evaluation reward: 5.26  
 episode: 1476 score: 7.0 memory length: 338382 epsilon: 0.33039766000978144  
 steps: 404 lr: 6.400000000000000e-06 evaluation reward: 5.29  
 episode: 1477 score: 6.0 memory length: 338757 epsilon: 0.32965516000977674  
 steps: 375 lr: 6.400000000000000e-06 evaluation reward: 5.27  
 episode: 1478 score: 3.0 memory length: 339005 epsilon: 0.32916412000977363  
 steps: 248 lr: 6.400000000000000e-06 evaluation reward: 5.16  
 episode: 1479 score: 4.0 memory length: 339299 epsilon: 0.32858200000976995  
 steps: 294 lr: 6.400000000000000e-06 evaluation reward: 5.15  
 episode: 1480 score: 7.0 memory length: 339739 epsilon: 0.32771080000976444  
 steps: 440 lr: 6.400000000000000e-06 evaluation reward: 5.16  
 episode: 1481 score: 2.0 memory length: 339938 epsilon: 0.32731678000976194  
 steps: 199 lr: 6.400000000000000e-06 evaluation reward: 5.14  
 episode: 1482 score: 6.0 memory length: 340272 epsilon: 0.32665546000975776  
 steps: 334 lr: 6.400000000000000e-06 evaluation reward: 5.13  
 episode: 1483 score: 4.0 memory length: 340565 epsilon: 0.3260753200097541 s  
 teps: 293 lr: 6.400000000000000e-06 evaluation reward: 5.13  
 episode: 1484 score: 6.0 memory length: 340920 epsilon: 0.32537242000974964  
 steps: 355 lr: 6.400000000000000e-06 evaluation reward: 5.15  
 episode: 1485 score: 4.0 memory length: 341179 epsilon: 0.3248596000097464 s  
 teps: 259 lr: 6.400000000000000e-06 evaluation reward: 5.14  
 episode: 1486 score: 6.0 memory length: 341535 epsilon: 0.32415472000974194  
 steps: 356 lr: 6.400000000000000e-06 evaluation reward: 5.16  
 episode: 1487 score: 11.0 memory length: 341983 epsilon: 0.3232676800097363  
 steps: 448 lr: 6.400000000000000e-06 evaluation reward: 5.2  
 episode: 1488 score: 4.0 memory length: 342242 epsilon: 0.3227548600097331 s  
 teps: 259 lr: 6.400000000000000e-06 evaluation reward: 5.18  
 episode: 1489 score: 6.0 memory length: 342575 epsilon: 0.3220955200097289 s  
 teps: 333 lr: 6.400000000000000e-06 evaluation reward: 5.19  
 episode: 1490 score: 6.0 memory length: 342930 epsilon: 0.32139262000972446  
 steps: 355 lr: 6.400000000000000e-06 evaluation reward: 5.18  
 episode: 1491 score: 5.0 memory length: 343236 epsilon: 0.32078674000972063  
 steps: 306 lr: 6.400000000000000e-06 evaluation reward: 5.2  
 episode: 1492 score: 5.0 memory length: 343561 epsilon: 0.32014324000971656  
 steps: 325 lr: 6.400000000000000e-06 evaluation reward: 5.19  
 episode: 1493 score: 7.0 memory length: 343977 epsilon: 0.31931956000971135  
 steps: 416 lr: 6.400000000000000e-06 evaluation reward: 5.21  
 episode: 1494 score: 6.0 memory length: 344315 epsilon: 0.3186503200097071 s  
 teps: 338 lr: 6.400000000000000e-06 evaluation reward: 5.24  
 episode: 1495 score: 10.0 memory length: 344822 epsilon: 0.31764646000970076  
 steps: 507 lr: 6.400000000000000e-06 evaluation reward: 5.29  
 episode: 1496 score: 8.0 memory length: 345258 epsilon: 0.3167831800096953 s  
 teps: 436 lr: 6.400000000000000e-06 evaluation reward: 5.34  
 episode: 1497 score: 5.0 memory length: 345547 epsilon: 0.3162109600096917 s  
 teps: 289 lr: 6.400000000000000e-06 evaluation reward: 5.36  
 episode: 1498 score: 6.0 memory length: 345878 epsilon: 0.31555558000968753  
 steps: 331 lr: 6.400000000000000e-06 evaluation reward: 5.35  
 episode: 1499 score: 3.0 memory length: 346108 epsilon: 0.31510018000968465  
 steps: 230 lr: 6.400000000000000e-06 evaluation reward: 5.34

episode: 1500 score: 3.0 memory length: 346318 epsilon: 0.314684380009682 st  
 eps: 210 lr: 6.400000000000000e-06 evaluation reward: 5.32  
 episode: 1501 score: 4.0 memory length: 346597 epsilon: 0.3141319600096785 s  
 teps: 279 lr: 6.400000000000000e-06 evaluation reward: 5.33  
 episode: 1502 score: 7.0 memory length: 346989 epsilon: 0.3133558000096736 s  
 teps: 392 lr: 6.400000000000000e-06 evaluation reward: 5.33  
 episode: 1503 score: 6.0 memory length: 347365 epsilon: 0.3126113200096689 s  
 teps: 376 lr: 6.400000000000000e-06 evaluation reward: 5.34  
 episode: 1504 score: 2.0 memory length: 347545 epsilon: 0.31225492000966665  
 steps: 180 lr: 6.400000000000000e-06 evaluation reward: 5.33  
 episode: 1505 score: 4.0 memory length: 347801 epsilon: 0.31174804000966344 s  
 steps: 256 lr: 6.400000000000000e-06 evaluation reward: 5.3  
 episode: 1506 score: 5.0 memory length: 348090 epsilon: 0.3111758200096598 s  
 teps: 289 lr: 6.400000000000000e-06 evaluation reward: 5.28  
 episode: 1507 score: 8.0 memory length: 348535 epsilon: 0.31029472000965425  
 steps: 445 lr: 6.400000000000000e-06 evaluation reward: 5.31  
 episode: 1508 score: 6.0 memory length: 348910 epsilon: 0.30955222000964955 s  
 steps: 375 lr: 6.400000000000000e-06 evaluation reward: 5.32  
 episode: 1509 score: 4.0 memory length: 349166 epsilon: 0.30904534000964634  
 steps: 256 lr: 6.400000000000000e-06 evaluation reward: 5.33  
 episode: 1510 score: 5.0 memory length: 349453 epsilon: 0.30847708000964275  
 steps: 287 lr: 6.400000000000000e-06 evaluation reward: 5.34  
 episode: 1511 score: 6.0 memory length: 349775 epsilon: 0.3078395200096387 s  
 teps: 322 lr: 6.400000000000000e-06 evaluation reward: 5.34  
 episode: 1512 score: 6.0 memory length: 350171 epsilon: 0.30705544000963375 s  
 steps: 396 lr: 6.400000000000000e-06 evaluation reward: 5.37  
 episode: 1513 score: 6.0 memory length: 350526 epsilon: 0.3063525400096293 s  
 teps: 355 lr: 6.400000000000000e-06 evaluation reward: 5.36  
 episode: 1514 score: 5.0 memory length: 350848 epsilon: 0.30571498000962527  
 steps: 322 lr: 6.400000000000000e-06 evaluation reward: 5.34  
 episode: 1515 score: 4.0 memory length: 351143 epsilon: 0.3051308800096216 s  
 teps: 295 lr: 6.400000000000000e-06 evaluation reward: 5.35  
 episode: 1516 score: 6.0 memory length: 351520 epsilon: 0.30438442000961685  
 steps: 377 lr: 6.400000000000000e-06 evaluation reward: 5.34  
 episode: 1517 score: 8.0 memory length: 351954 epsilon: 0.3035251000096114 s  
 teps: 434 lr: 6.400000000000000e-06 evaluation reward: 5.38  
 episode: 1518 score: 4.0 memory length: 352212 epsilon: 0.3030142600096082 s  
 teps: 258 lr: 6.400000000000000e-06 evaluation reward: 5.38  
 episode: 1519 score: 6.0 memory length: 352531 epsilon: 0.3023826400096042 s  
 teps: 319 lr: 6.400000000000000e-06 evaluation reward: 5.4  
 episode: 1520 score: 3.0 memory length: 352779 epsilon: 0.3018916000096011 s  
 teps: 248 lr: 6.400000000000000e-06 evaluation reward: 5.39  
 episode: 1521 score: 8.0 memory length: 353251 epsilon: 0.30095704000959517  
 steps: 472 lr: 6.400000000000000e-06 evaluation reward: 5.42  
 episode: 1522 score: 12.0 memory length: 353702 epsilon: 0.3000640600095895 s  
 steps: 451 lr: 6.400000000000000e-06 evaluation reward: 5.51  
 episode: 1523 score: 8.0 memory length: 354087 epsilon: 0.2993017600095847 s  
 teps: 385 lr: 6.400000000000000e-06 evaluation reward: 5.51  
 episode: 1524 score: 11.0 memory length: 354621 epsilon: 0.298244440009578 s  
 teps: 534 lr: 6.400000000000000e-06 evaluation reward: 5.58  
 episode: 1525 score: 5.0 memory length: 354928 epsilon: 0.29763658000957416  
 steps: 307 lr: 6.400000000000000e-06 evaluation reward: 5.57  
 episode: 1526 score: 4.0 memory length: 355187 epsilon: 0.2971237600095709 s  
 teps: 259 lr: 6.400000000000000e-06 evaluation reward: 5.57  
 episode: 1527 score: 9.0 memory length: 355520 epsilon: 0.29646442000956674  
 steps: 333 lr: 6.400000000000000e-06 evaluation reward: 5.63  
 episode: 1528 score: 5.0 memory length: 355865 epsilon: 0.2957813200095624 s  
 teps: 345 lr: 6.400000000000000e-06 evaluation reward: 5.62  
 episode: 1529 score: 3.0 memory length: 356091 epsilon: 0.2953338400095596 s  
 teps: 226 lr: 6.400000000000000e-06 evaluation reward: 5.6

episode: 1530 score: 4.0 memory length: 356371 epsilon: 0.2947794400095561 s  
 teps: 280 lr: 6.400000000000000e-06 evaluation reward: 5.58  
 episode: 1531 score: 4.0 memory length: 356610 epsilon: 0.2943062200095531 s  
 teps: 239 lr: 6.400000000000000e-06 evaluation reward: 5.57  
 episode: 1532 score: 6.0 memory length: 356951 epsilon: 0.2936310400095488 s  
 teps: 341 lr: 6.400000000000000e-06 evaluation reward: 5.59  
 episode: 1533 score: 8.0 memory length: 357362 epsilon: 0.29281726000954367 s  
 steps: 411 lr: 6.400000000000000e-06 evaluation reward: 5.64  
 episode: 1534 score: 5.0 memory length: 357650 epsilon: 0.29224702000954006 s  
 steps: 288 lr: 6.400000000000000e-06 evaluation reward: 5.64  
 episode: 1535 score: 3.0 memory length: 357876 epsilon: 0.29179954000953723 s  
 steps: 226 lr: 6.400000000000000e-06 evaluation reward: 5.6  
 episode: 1536 score: 4.0 memory length: 358136 epsilon: 0.29128474000953397 s  
 steps: 260 lr: 6.400000000000000e-06 evaluation reward: 5.55  
 episode: 1537 score: 2.0 memory length: 358318 epsilon: 0.2909243800095317 s  
 teps: 182 lr: 6.400000000000000e-06 evaluation reward: 5.5  
 episode: 1538 score: 6.0 memory length: 358657 epsilon: 0.29025316000952744 s  
 steps: 339 lr: 6.400000000000000e-06 evaluation reward: 5.52  
 episode: 1539 score: 7.0 memory length: 359029 epsilon: 0.2895166000095228 s  
 teps: 372 lr: 6.400000000000000e-06 evaluation reward: 5.55  
 episode: 1540 score: 6.0 memory length: 359385 epsilon: 0.2888117200095183 s  
 teps: 356 lr: 6.400000000000000e-06 evaluation reward: 5.58  
 episode: 1541 score: 6.0 memory length: 359761 epsilon: 0.2880672400095136 s  
 teps: 376 lr: 6.400000000000000e-06 evaluation reward: 5.59  
 episode: 1542 score: 11.0 memory length: 360295 epsilon: 0.2870099200095069 s  
 steps: 534 lr: 6.400000000000000e-06 evaluation reward: 5.68  
 episode: 1543 score: 2.0 memory length: 360475 epsilon: 0.28665352000950467 s  
 steps: 180 lr: 6.400000000000000e-06 evaluation reward: 5.66  
 episode: 1544 score: 4.0 memory length: 360735 epsilon: 0.2861387200095014 s  
 teps: 260 lr: 6.400000000000000e-06 evaluation reward: 5.66  
 episode: 1545 score: 3.0 memory length: 360983 epsilon: 0.2856476800094983 s  
 teps: 248 lr: 6.400000000000000e-06 evaluation reward: 5.65  
 episode: 1546 score: 5.0 memory length: 361269 epsilon: 0.2850814000094947 s  
 teps: 286 lr: 6.400000000000000e-06 evaluation reward: 5.66  
 episode: 1547 score: 7.0 memory length: 361650 epsilon: 0.28432702000948995 s  
 steps: 381 lr: 6.400000000000000e-06 evaluation reward: 5.69  
 episode: 1548 score: 5.0 memory length: 361935 epsilon: 0.2837627200094864 s  
 teps: 285 lr: 6.400000000000000e-06 evaluation reward: 5.7  
 episode: 1549 score: 7.0 memory length: 362321 epsilon: 0.28299844000948154 s  
 steps: 386 lr: 6.400000000000000e-06 evaluation reward: 5.69  
 episode: 1550 score: 7.0 memory length: 362699 epsilon: 0.2822500000094768 s  
 teps: 378 lr: 6.400000000000000e-06 evaluation reward: 5.7  
 episode: 1551 score: 4.0 memory length: 362940 epsilon: 0.2817728200094738 s  
 teps: 241 lr: 6.400000000000000e-06 evaluation reward: 5.7  
 episode: 1552 score: 3.0 memory length: 363189 epsilon: 0.28127980000947067 s  
 steps: 249 lr: 6.400000000000000e-06 evaluation reward: 5.7  
 episode: 1553 score: 6.0 memory length: 363571 epsilon: 0.2805234400094659 s  
 teps: 382 lr: 6.400000000000000e-06 evaluation reward: 5.68  
 episode: 1554 score: 9.0 memory length: 364090 epsilon: 0.2794958200094594 s  
 teps: 519 lr: 6.400000000000000e-06 evaluation reward: 5.71  
 episode: 1555 score: 6.0 memory length: 364443 epsilon: 0.27879688000945496 s  
 steps: 353 lr: 6.400000000000000e-06 evaluation reward: 5.7  
 episode: 1556 score: 6.0 memory length: 364763 epsilon: 0.27816328000945095 s  
 steps: 320 lr: 6.400000000000000e-06 evaluation reward: 5.71  
 episode: 1557 score: 3.0 memory length: 364992 epsilon: 0.2777098600094481 s  
 teps: 229 lr: 6.400000000000000e-06 evaluation reward: 5.65  
 episode: 1558 score: 5.0 memory length: 365335 epsilon: 0.2770307200094438 s  
 teps: 343 lr: 6.400000000000000e-06 evaluation reward: 5.6  
 episode: 1559 score: 4.0 memory length: 365595 epsilon: 0.27651592000944053 s  
 steps: 260 lr: 6.400000000000000e-06 evaluation reward: 5.61

episode: 1560 score: 4.0 memory length: 365872 epsilon: 0.27596746000943706  
 steps: 277 lr: 6.400000000000000e-06 evaluation reward: 5.58  
 episode: 1561 score: 7.0 memory length: 366257 epsilon: 0.27520516000943224  
 steps: 385 lr: 6.400000000000000e-06 evaluation reward: 5.57  
 episode: 1562 score: 4.0 memory length: 366533 epsilon: 0.2746586800094288 s  
 teps: 276 lr: 6.400000000000000e-06 evaluation reward: 5.54  
 episode: 1563 score: 5.0 memory length: 366840 epsilon: 0.27405082000942493  
 steps: 307 lr: 6.400000000000000e-06 evaluation reward: 5.55  
 episode: 1564 score: 5.0 memory length: 367167 epsilon: 0.27340336000942084  
 steps: 327 lr: 6.400000000000000e-06 evaluation reward: 5.56  
 episode: 1565 score: 8.0 memory length: 367601 epsilon: 0.2725440400094154 s  
 teps: 434 lr: 6.400000000000000e-06 evaluation reward: 5.57  
 episode: 1566 score: 5.0 memory length: 367887 epsilon: 0.2719777600094118 s  
 teps: 286 lr: 6.400000000000000e-06 evaluation reward: 5.57  
 episode: 1567 score: 6.0 memory length: 368227 epsilon: 0.27130456000940756  
 steps: 340 lr: 6.400000000000000e-06 evaluation reward: 5.59  
 episode: 1568 score: 11.0 memory length: 368687 epsilon: 0.2703937600094018  
 steps: 460 lr: 6.400000000000000e-06 evaluation reward: 5.65  
 episode: 1569 score: 6.0 memory length: 369041 epsilon: 0.26969284000939736  
 steps: 354 lr: 6.400000000000000e-06 evaluation reward: 5.67  
 episode: 1570 score: 7.0 memory length: 369408 epsilon: 0.26896618000939276  
 steps: 367 lr: 6.400000000000000e-06 evaluation reward: 5.69  
 episode: 1571 score: 6.0 memory length: 369745 epsilon: 0.26829892000938854  
 steps: 337 lr: 6.400000000000000e-06 evaluation reward: 5.68  
 episode: 1572 score: 6.0 memory length: 370065 epsilon: 0.26766532000938453  
 steps: 320 lr: 6.400000000000000e-06 evaluation reward: 5.68  
 episode: 1573 score: 4.0 memory length: 370304 epsilon: 0.26719210000938154  
 steps: 239 lr: 6.400000000000000e-06 evaluation reward: 5.68  
 episode: 1574 score: 5.0 memory length: 370628 epsilon: 0.2665505800093775 s  
 teps: 324 lr: 6.400000000000000e-06 evaluation reward: 5.61  
 episode: 1575 score: 6.0 memory length: 370998 epsilon: 0.26581798000937285  
 steps: 370 lr: 6.400000000000000e-06 evaluation reward: 5.63  
 episode: 1576 score: 3.0 memory length: 371212 epsilon: 0.26539426000937016  
 steps: 214 lr: 6.400000000000000e-06 evaluation reward: 5.59  
 episode: 1577 score: 6.0 memory length: 371534 epsilon: 0.26475670000936613  
 steps: 322 lr: 6.400000000000000e-06 evaluation reward: 5.59  
 episode: 1578 score: 4.0 memory length: 371793 epsilon: 0.2642438800093629 s  
 teps: 259 lr: 6.400000000000000e-06 evaluation reward: 5.6  
 episode: 1579 score: 4.0 memory length: 372086 epsilon: 0.2636637400093592 s  
 teps: 293 lr: 6.400000000000000e-06 evaluation reward: 5.6  
 episode: 1580 score: 4.0 memory length: 372344 epsilon: 0.263152900009356 st  
 eps: 258 lr: 6.400000000000000e-06 evaluation reward: 5.57  
 episode: 1581 score: 16.0 memory length: 372935 epsilon: 0.2619827200093486  
 steps: 591 lr: 6.400000000000000e-06 evaluation reward: 5.71  
 episode: 1582 score: 7.0 memory length: 373348 epsilon: 0.2611649800093434 s  
 teps: 413 lr: 6.400000000000000e-06 evaluation reward: 5.72  
 episode: 1583 score: 10.0 memory length: 373851 epsilon: 0.2601690400093371  
 steps: 503 lr: 6.400000000000000e-06 evaluation reward: 5.78  
 episode: 1584 score: 4.0 memory length: 374110 epsilon: 0.25965622000933386  
 steps: 259 lr: 6.400000000000000e-06 evaluation reward: 5.76  
 episode: 1585 score: 4.0 memory length: 374352 epsilon: 0.25917706000933083  
 steps: 242 lr: 6.400000000000000e-06 evaluation reward: 5.76  
 episode: 1586 score: 8.0 memory length: 374789 epsilon: 0.25831180000932535  
 steps: 437 lr: 6.400000000000000e-06 evaluation reward: 5.78  
 episode: 1587 score: 5.0 memory length: 375114 epsilon: 0.2576683000093213 s  
 teps: 325 lr: 6.400000000000000e-06 evaluation reward: 5.72  
 episode: 1588 score: 8.0 memory length: 375586 epsilon: 0.25673374000931537  
 steps: 472 lr: 6.400000000000000e-06 evaluation reward: 5.76  
 episode: 1589 score: 3.0 memory length: 375797 epsilon: 0.2563159600093127 s  
 teps: 211 lr: 6.400000000000000e-06 evaluation reward: 5.73

```

episode: 1590    score: 12.0    memory length: 376406    epsilon: 0.2551101400093051
steps: 609      lr: 6.400000000000001e-06    evaluation reward: 5.79
episode: 1591    score: 5.0     memory length: 376730    epsilon: 0.25446862000930104
steps: 324      lr: 6.400000000000001e-06    evaluation reward: 5.79
episode: 1592    score: 8.0     memory length: 377152    epsilon: 0.25363306000929575
steps: 422      lr: 6.400000000000001e-06    evaluation reward: 5.82
episode: 1593    score: 6.0     memory length: 377509    epsilon: 0.2529262000092913    s
teps: 357      lr: 6.400000000000001e-06    evaluation reward: 5.81
episode: 1594    score: 4.0     memory length: 377802    epsilon: 0.2523460600092876    s
teps: 293      lr: 6.400000000000001e-06    evaluation reward: 5.79
episode: 1595    score: 4.0     memory length: 378080    epsilon: 0.2517956200092841    s
teps: 278      lr: 6.400000000000001e-06    evaluation reward: 5.73
episode: 1596    score: 4.0     memory length: 378335    epsilon: 0.25129072000928093
steps: 255      lr: 6.400000000000001e-06    evaluation reward: 5.69
episode: 1597    score: 4.0     memory length: 378610    epsilon: 0.2507462200092775    s
teps: 275      lr: 6.400000000000001e-06    evaluation reward: 5.68
episode: 1598    score: 7.0     memory length: 379033    epsilon: 0.2499086800092722    s
teps: 423      lr: 6.400000000000001e-06    evaluation reward: 5.69
episode: 1599    score: 3.0     memory length: 379243    epsilon: 0.24949288000926956
steps: 210      lr: 6.400000000000001e-06    evaluation reward: 5.69
episode: 1600    score: 5.0     memory length: 379548    epsilon: 0.24888898000926574
steps: 305      lr: 6.400000000000001e-06    evaluation reward: 5.71
episode: 1601    score: 8.0     memory length: 379925    epsilon: 0.24814252000926101
steps: 377      lr: 6.400000000000001e-06    evaluation reward: 5.75
episode: 1602    score: 5.0     memory length: 380252    epsilon: 0.24749506000925692
steps: 327      lr: 6.400000000000001e-06    evaluation reward: 5.73
episode: 1603    score: 6.0     memory length: 380630    epsilon: 0.24674662000925218
steps: 378      lr: 6.400000000000001e-06    evaluation reward: 5.73
episode: 1604    score: 6.0     memory length: 380987    epsilon: 0.2460397600092477    s
teps: 357      lr: 6.400000000000001e-06    evaluation reward: 5.77
episode: 1605    score: 10.0    memory length: 381494    epsilon: 0.24503590000924136
steps: 507      lr: 6.400000000000001e-06    evaluation reward: 5.83
episode: 1606    score: 7.0     memory length: 381905    epsilon: 0.2442221200092362    s
teps: 411      lr: 6.400000000000001e-06    evaluation reward: 5.85
episode: 1607    score: 9.0     memory length: 382408    epsilon: 0.2432261800092299    s
teps: 503      lr: 6.400000000000001e-06    evaluation reward: 5.86
episode: 1608    score: 7.0     memory length: 382828    epsilon: 0.24239458000922465
steps: 420      lr: 6.400000000000001e-06    evaluation reward: 5.87
episode: 1609    score: 4.0     memory length: 383086    epsilon: 0.24188374000922142
steps: 258      lr: 6.400000000000001e-06    evaluation reward: 5.87
episode: 1610    score: 7.0     memory length: 383480    epsilon: 0.24110362000921648
steps: 394      lr: 6.400000000000001e-06    evaluation reward: 5.89
episode: 1611    score: 5.0     memory length: 383789    epsilon: 0.2404918000092126    s
teps: 309      lr: 6.400000000000001e-06    evaluation reward: 5.88
episode: 1612    score: 4.0     memory length: 384068    epsilon: 0.2399393800092091    s
teps: 279      lr: 6.400000000000001e-06    evaluation reward: 5.86
episode: 1613    score: 7.0     memory length: 384480    epsilon: 0.23912362000920395
steps: 412      lr: 6.400000000000001e-06    evaluation reward: 5.87
episode: 1614    score: 4.0     memory length: 384721    epsilon: 0.23864644000920093
steps: 241      lr: 6.400000000000001e-06    evaluation reward: 5.86
episode: 1615    score: 10.0    memory length: 385098    epsilon: 0.2378999800091962
steps: 377      lr: 6.400000000000001e-06    evaluation reward: 5.92
episode: 1616    score: 12.0    memory length: 385597    epsilon: 0.23691196000918996
steps: 499      lr: 6.400000000000001e-06    evaluation reward: 5.98
episode: 1617    score: 4.0     memory length: 385839    epsilon: 0.23643280000918693
steps: 242      lr: 6.400000000000001e-06    evaluation reward: 5.94
episode: 1618    score: 4.0     memory length: 386115    epsilon: 0.23588632000918347
steps: 276      lr: 6.400000000000001e-06    evaluation reward: 5.94
episode: 1619    score: 4.0     memory length: 386370    epsilon: 0.23538142000918028
steps: 255      lr: 6.400000000000001e-06    evaluation reward: 5.92

```

episode: 1620 score: 7.0 memory length: 386759 epsilon: 0.2346112000091754 s  
 teps: 389 lr: 6.400000000000000e-06 evaluation reward: 5.96  
 episode: 1621 score: 6.0 memory length: 387080 epsilon: 0.23397562000917138  
 steps: 321 lr: 6.400000000000000e-06 evaluation reward: 5.94  
 episode: 1622 score: 4.0 memory length: 387335 epsilon: 0.2334707200091682 s  
 teps: 255 lr: 6.400000000000000e-06 evaluation reward: 5.86  
 episode: 1623 score: 6.0 memory length: 387711 epsilon: 0.23272624000916348  
 steps: 376 lr: 6.400000000000000e-06 evaluation reward: 5.84  
 episode: 1624 score: 12.0 memory length: 388260 epsilon: 0.2316392200091566  
 steps: 549 lr: 6.400000000000000e-06 evaluation reward: 5.85  
 episode: 1625 score: 7.0 memory length: 388688 epsilon: 0.23079178000915124  
 steps: 428 lr: 6.400000000000000e-06 evaluation reward: 5.87  
 episode: 1626 score: 6.0 memory length: 389031 epsilon: 0.23011264000914694  
 steps: 343 lr: 6.400000000000000e-06 evaluation reward: 5.89  
 episode: 1627 score: 4.0 memory length: 389308 epsilon: 0.22956418000914347  
 steps: 277 lr: 6.400000000000000e-06 evaluation reward: 5.84  
 episode: 1628 score: 5.0 memory length: 389619 epsilon: 0.22894840000913957  
 steps: 311 lr: 6.400000000000000e-06 evaluation reward: 5.84  
 episode: 1629 score: 5.0 memory length: 389908 epsilon: 0.22837618000913595  
 steps: 289 lr: 6.400000000000000e-06 evaluation reward: 5.86  
 episode: 1630 score: 5.0 memory length: 390235 epsilon: 0.22772872000913186  
 steps: 327 lr: 6.400000000000000e-06 evaluation reward: 5.87  
 episode: 1631 score: 10.0 memory length: 390705 epsilon: 0.22679812000912597  
 steps: 470 lr: 6.400000000000000e-06 evaluation reward: 5.93  
 episode: 1632 score: 4.0 memory length: 390982 epsilon: 0.2262496600091225 s  
 teps: 277 lr: 6.400000000000000e-06 evaluation reward: 5.91  
 episode: 1633 score: 7.0 memory length: 391370 epsilon: 0.22548142000911764  
 steps: 388 lr: 6.400000000000000e-06 evaluation reward: 5.9  
 episode: 1634 score: 4.0 memory length: 391647 epsilon: 0.22493296000911417  
 steps: 277 lr: 6.400000000000000e-06 evaluation reward: 5.89  
 episode: 1635 score: 4.0 memory length: 391905 epsilon: 0.22442212000911094  
 steps: 258 lr: 6.400000000000000e-06 evaluation reward: 5.9  
 episode: 1636 score: 6.0 memory length: 392223 epsilon: 0.22379248000910695  
 steps: 318 lr: 6.400000000000000e-06 evaluation reward: 5.92  
 episode: 1637 score: 4.0 memory length: 392522 epsilon: 0.2232004600091032 s  
 teps: 299 lr: 6.400000000000000e-06 evaluation reward: 5.94  
 episode: 1638 score: 8.0 memory length: 392967 epsilon: 0.22231936000909763  
 steps: 445 lr: 6.400000000000000e-06 evaluation reward: 5.96  
 episode: 1639 score: 7.0 memory length: 393372 epsilon: 0.22151746000909256  
 steps: 405 lr: 6.400000000000000e-06 evaluation reward: 5.96  
 episode: 1640 score: 8.0 memory length: 393789 epsilon: 0.22069180000908734  
 steps: 417 lr: 6.400000000000000e-06 evaluation reward: 5.98  
 episode: 1641 score: 7.0 memory length: 394195 epsilon: 0.21988792000908225  
 steps: 406 lr: 6.400000000000000e-06 evaluation reward: 5.99  
 episode: 1642 score: 6.0 memory length: 394588 epsilon: 0.21910978000907733  
 steps: 393 lr: 6.400000000000000e-06 evaluation reward: 5.94  
 episode: 1643 score: 6.0 memory length: 394926 epsilon: 0.2184405400090731 s  
 teps: 338 lr: 6.400000000000000e-06 evaluation reward: 5.98  
 episode: 1644 score: 5.0 memory length: 395217 epsilon: 0.21786436000906945  
 steps: 291 lr: 6.400000000000000e-06 evaluation reward: 5.99  
 episode: 1645 score: 8.0 memory length: 395657 epsilon: 0.21699316000906393  
 steps: 440 lr: 6.400000000000000e-06 evaluation reward: 6.04  
 episode: 1646 score: 5.0 memory length: 395961 epsilon: 0.21639124000906013  
 steps: 304 lr: 6.400000000000000e-06 evaluation reward: 6.04  
 episode: 1647 score: 6.0 memory length: 396315 epsilon: 0.2156903200090557 s  
 teps: 354 lr: 6.400000000000000e-06 evaluation reward: 6.03  
 episode: 1648 score: 4.0 memory length: 396573 epsilon: 0.21517948000905246  
 steps: 258 lr: 6.400000000000000e-06 evaluation reward: 6.02  
 episode: 1649 score: 5.0 memory length: 396916 epsilon: 0.21450034000904816  
 steps: 343 lr: 6.400000000000000e-06 evaluation reward: 6.0

episode: 1650 score: 8.0 memory length: 397332 epsilon: 0.21367666000904295  
 steps: 416 lr: 6.400000000000000e-06 evaluation reward: 6.01  
 episode: 1651 score: 6.0 memory length: 397662 epsilon: 0.21302326000903882  
 steps: 330 lr: 6.400000000000000e-06 evaluation reward: 6.03  
 episode: 1652 score: 6.0 memory length: 398058 epsilon: 0.21223918000903386  
 steps: 396 lr: 6.400000000000000e-06 evaluation reward: 6.06  
 episode: 1653 score: 10.0 memory length: 398465 epsilon: 0.21143332000902876  
 steps: 407 lr: 6.400000000000000e-06 evaluation reward: 6.1  
 episode: 1654 score: 4.0 memory length: 398725 epsilon: 0.2109185200090255 s  
 teps: 260 lr: 6.400000000000000e-06 evaluation reward: 6.05  
 episode: 1655 score: 7.0 memory length: 399122 epsilon: 0.21013246000902053  
 steps: 397 lr: 6.400000000000000e-06 evaluation reward: 6.06  
 episode: 1656 score: 6.0 memory length: 399455 epsilon: 0.20947312000901636  
 steps: 333 lr: 6.400000000000000e-06 evaluation reward: 6.06  
 episode: 1657 score: 11.0 memory length: 399978 epsilon: 0.2084375800090098  
 steps: 523 lr: 6.400000000000000e-06 evaluation reward: 6.14  
 episode: 1658 score: 5.0 memory length: 400261 epsilon: 0.20787724000900626  
 steps: 283 lr: 2.560000000000000e-06 evaluation reward: 6.14  
 episode: 1659 score: 7.0 memory length: 400687 epsilon: 0.20703376000900092  
 steps: 426 lr: 2.560000000000000e-06 evaluation reward: 6.17  
 episode: 1660 score: 5.0 memory length: 400995 epsilon: 0.20642392000899706  
 steps: 308 lr: 2.560000000000000e-06 evaluation reward: 6.18  
 episode: 1661 score: 6.0 memory length: 401330 epsilon: 0.20576062000899287  
 steps: 335 lr: 2.560000000000000e-06 evaluation reward: 6.17  
 episode: 1662 score: 5.0 memory length: 401613 epsilon: 0.20520028000898932  
 steps: 283 lr: 2.560000000000000e-06 evaluation reward: 6.18  
 episode: 1663 score: 6.0 memory length: 401935 epsilon: 0.2045627200089853 s  
 teps: 322 lr: 2.560000000000000e-06 evaluation reward: 6.19  
 episode: 1664 score: 5.0 memory length: 402278 epsilon: 0.203883580008981 st  
 eps: 343 lr: 2.560000000000000e-06 evaluation reward: 6.19  
 episode: 1665 score: 4.0 memory length: 402537 epsilon: 0.20337076000897775  
 steps: 259 lr: 2.560000000000000e-06 evaluation reward: 6.15  
 episode: 1666 score: 7.0 memory length: 402927 epsilon: 0.20259856000897286  
 steps: 390 lr: 2.560000000000000e-06 evaluation reward: 6.17  
 episode: 1667 score: 4.0 memory length: 403185 epsilon: 0.20208772000896963  
 steps: 258 lr: 2.560000000000000e-06 evaluation reward: 6.15  
 episode: 1668 score: 6.0 memory length: 403530 epsilon: 0.2014046200089653 s  
 teps: 345 lr: 2.560000000000000e-06 evaluation reward: 6.1  
 episode: 1669 score: 13.0 memory length: 404045 epsilon: 0.20038492000895886  
 steps: 515 lr: 2.560000000000000e-06 evaluation reward: 6.17  
 episode: 1670 score: 9.0 memory length: 404522 epsilon: 0.19944046000895288  
 steps: 477 lr: 2.560000000000000e-06 evaluation reward: 6.19  
 episode: 1671 score: 9.0 memory length: 404996 epsilon: 0.19850194000894694  
 steps: 474 lr: 2.560000000000000e-06 evaluation reward: 6.22  
 episode: 1672 score: 5.0 memory length: 405302 epsilon: 0.1978960600089431 s  
 teps: 306 lr: 2.560000000000000e-06 evaluation reward: 6.21  
 episode: 1673 score: 7.0 memory length: 405685 epsilon: 0.1971377200089383 s  
 teps: 383 lr: 2.560000000000000e-06 evaluation reward: 6.24  
 episode: 1674 score: 8.0 memory length: 406143 epsilon: 0.19623088000893257  
 steps: 458 lr: 2.560000000000000e-06 evaluation reward: 6.27  
 episode: 1675 score: 4.0 memory length: 406404 epsilon: 0.1957141000089293 s  
 teps: 261 lr: 2.560000000000000e-06 evaluation reward: 6.25  
 episode: 1676 score: 7.0 memory length: 406794 epsilon: 0.19494190000892442  
 steps: 390 lr: 2.560000000000000e-06 evaluation reward: 6.29  
 episode: 1677 score: 6.0 memory length: 407113 epsilon: 0.19431028000892042  
 steps: 319 lr: 2.560000000000000e-06 evaluation reward: 6.29  
 episode: 1678 score: 13.0 memory length: 407580 epsilon: 0.19338562000891457  
 steps: 467 lr: 2.560000000000000e-06 evaluation reward: 6.38  
 episode: 1679 score: 7.0 memory length: 407987 epsilon: 0.19257976000890947  
 steps: 407 lr: 2.560000000000000e-06 evaluation reward: 6.41

episode: 1680 score: 6.0 memory length: 408323 epsilon: 0.19191448000890526  
 steps: 336 lr: 2.5600000000000001e-06 evaluation reward: 6.43  
 episode: 1681 score: 9.0 memory length: 408683 epsilon: 0.19120168000890075  
 steps: 360 lr: 2.5600000000000001e-06 evaluation reward: 6.36  
 episode: 1682 score: 6.0 memory length: 409054 epsilon: 0.1904671000088961 s  
 teps: 371 lr: 2.5600000000000001e-06 evaluation reward: 6.35  
 episode: 1683 score: 4.0 memory length: 409349 epsilon: 0.1898830000088924 s  
 teps: 295 lr: 2.5600000000000001e-06 evaluation reward: 6.29  
 episode: 1684 score: 10.0 memory length: 409857 epsilon: 0.18887716000888605  
 steps: 508 lr: 2.5600000000000001e-06 evaluation reward: 6.35  
 episode: 1685 score: 8.0 memory length: 410279 epsilon: 0.18804160000888076  
 steps: 422 lr: 2.5600000000000001e-06 evaluation reward: 6.39  
 episode: 1686 score: 7.0 memory length: 410704 epsilon: 0.18720010000887544  
 steps: 425 lr: 2.5600000000000001e-06 evaluation reward: 6.38  
 episode: 1687 score: 6.0 memory length: 411042 epsilon: 0.1865308600088712 s  
 teps: 338 lr: 2.5600000000000001e-06 evaluation reward: 6.39  
 episode: 1688 score: 4.0 memory length: 411338 epsilon: 0.1859447800088675 s  
 teps: 296 lr: 2.5600000000000001e-06 evaluation reward: 6.35  
 episode: 1689 score: 11.0 memory length: 411824 epsilon: 0.1849825000088614  
 steps: 486 lr: 2.5600000000000001e-06 evaluation reward: 6.43  
 episode: 1690 score: 5.0 memory length: 412113 epsilon: 0.18441028000885779  
 steps: 289 lr: 2.5600000000000001e-06 evaluation reward: 6.36  
 episode: 1691 score: 14.0 memory length: 412646 epsilon: 0.1833549400088511  
 steps: 533 lr: 2.5600000000000001e-06 evaluation reward: 6.45  
 episode: 1692 score: 4.0 memory length: 412906 epsilon: 0.18284014000884785  
 steps: 260 lr: 2.5600000000000001e-06 evaluation reward: 6.41  
 episode: 1693 score: 7.0 memory length: 413289 epsilon: 0.18208180000884305  
 steps: 383 lr: 2.5600000000000001e-06 evaluation reward: 6.42  
 episode: 1694 score: 5.0 memory length: 413580 epsilon: 0.18150562000088394 s  
 teps: 291 lr: 2.5600000000000001e-06 evaluation reward: 6.43  
 episode: 1695 score: 6.0 memory length: 413937 epsilon: 0.18079876000883494  
 steps: 357 lr: 2.5600000000000001e-06 evaluation reward: 6.45  
 episode: 1696 score: 6.0 memory length: 414309 epsilon: 0.18006220000883028  
 steps: 372 lr: 2.5600000000000001e-06 evaluation reward: 6.47  
 episode: 1697 score: 14.0 memory length: 414879 epsilon: 0.17893360000882313  
 steps: 570 lr: 2.5600000000000001e-06 evaluation reward: 6.57  
 episode: 1698 score: 7.0 memory length: 415250 epsilon: 0.1781990200088185 s  
 teps: 371 lr: 2.5600000000000001e-06 evaluation reward: 6.57  
 episode: 1699 score: 7.0 memory length: 415603 epsilon: 0.17750008000881407  
 steps: 353 lr: 2.5600000000000001e-06 evaluation reward: 6.61  
 episode: 1700 score: 5.0 memory length: 415946 epsilon: 0.17682094000880977  
 steps: 343 lr: 2.5600000000000001e-06 evaluation reward: 6.61  
 episode: 1701 score: 5.0 memory length: 416249 epsilon: 0.17622100000880597  
 steps: 303 lr: 2.5600000000000001e-06 evaluation reward: 6.58  
 episode: 1702 score: 7.0 memory length: 416622 epsilon: 0.1754824600088013 s  
 teps: 373 lr: 2.5600000000000001e-06 evaluation reward: 6.6  
 episode: 1703 score: 5.0 memory length: 416947 epsilon: 0.17483896000879723  
 steps: 325 lr: 2.5600000000000001e-06 evaluation reward: 6.59  
 episode: 1704 score: 4.0 memory length: 417205 epsilon: 0.174328120008794 st  
 eps: 258 lr: 2.5600000000000001e-06 evaluation reward: 6.57  
 episode: 1705 score: 7.0 memory length: 417589 epsilon: 0.17356780000878919  
 steps: 384 lr: 2.5600000000000001e-06 evaluation reward: 6.54  
 episode: 1706 score: 7.0 memory length: 417972 epsilon: 0.1728094600087844 s  
 teps: 383 lr: 2.5600000000000001e-06 evaluation reward: 6.54  
 episode: 1707 score: 5.0 memory length: 418259 epsilon: 0.1722412000087808 s  
 teps: 287 lr: 2.5600000000000001e-06 evaluation reward: 6.5  
 episode: 1708 score: 5.0 memory length: 418567 epsilon: 0.17163136000877693  
 steps: 308 lr: 2.5600000000000001e-06 evaluation reward: 6.48  
 episode: 1709 score: 7.0 memory length: 418971 epsilon: 0.17083144000877187  
 steps: 404 lr: 2.5600000000000001e-06 evaluation reward: 6.51

episode: 1710 score: 7.0 memory length: 419376 epsilon: 0.1700295400087668 s  
teps: 405 lr: 2.5600000000000001e-06 evaluation reward: 6.51  
episode: 1711 score: 5.0 memory length: 419647 epsilon: 0.1694929600087634 s  
teps: 271 lr: 2.5600000000000001e-06 evaluation reward: 6.51  
episode: 1712 score: 6.0 memory length: 419967 epsilon: 0.1688593600087594 s  
teps: 320 lr: 2.5600000000000001e-06 evaluation reward: 6.53  
episode: 1713 score: 7.0 memory length: 420358 epsilon: 0.1680851800087545 s  
teps: 391 lr: 2.5600000000000001e-06 evaluation reward: 6.53  
episode: 1714 score: 5.0 memory length: 420664 epsilon: 0.16747930000875066 s  
steps: 306 lr: 2.5600000000000001e-06 evaluation reward: 6.54  
episode: 1715 score: 6.0 memory length: 421020 epsilon: 0.1667744200087462 s  
teps: 356 lr: 2.5600000000000001e-06 evaluation reward: 6.5  
episode: 1716 score: 6.0 memory length: 421393 epsilon: 0.16603588000874153 s  
steps: 373 lr: 2.5600000000000001e-06 evaluation reward: 6.44  
episode: 1717 score: 5.0 memory length: 421681 epsilon: 0.16546564000873792 s  
steps: 288 lr: 2.5600000000000001e-06 evaluation reward: 6.45  
episode: 1718 score: 5.0 memory length: 422006 epsilon: 0.16482214000873385 s  
steps: 325 lr: 2.5600000000000001e-06 evaluation reward: 6.46  
episode: 1719 score: 7.0 memory length: 422397 epsilon: 0.16404796000872895 s  
steps: 391 lr: 2.5600000000000001e-06 evaluation reward: 6.49  
episode: 1720 score: 6.0 memory length: 422736 epsilon: 0.1633767400087247 s  
teps: 339 lr: 2.5600000000000001e-06 evaluation reward: 6.48  
episode: 1721 score: 5.0 memory length: 423024 epsilon: 0.1628065000087211 s  
teps: 288 lr: 2.5600000000000001e-06 evaluation reward: 6.47  
episode: 1722 score: 4.0 memory length: 423283 epsilon: 0.16229368000871786 s  
steps: 259 lr: 2.5600000000000001e-06 evaluation reward: 6.47  
episode: 1723 score: 8.0 memory length: 423718 epsilon: 0.1614323800087124 s  
teps: 435 lr: 2.5600000000000001e-06 evaluation reward: 6.49  
episode: 1724 score: 5.0 memory length: 424003 epsilon: 0.16086808000870884 s  
steps: 285 lr: 2.5600000000000001e-06 evaluation reward: 6.42  
episode: 1725 score: 7.0 memory length: 424372 epsilon: 0.1601374600087042 s  
teps: 369 lr: 2.5600000000000001e-06 evaluation reward: 6.42  
episode: 1726 score: 4.0 memory length: 424630 epsilon: 0.15962662000870098 s  
steps: 258 lr: 2.5600000000000001e-06 evaluation reward: 6.4  
episode: 1727 score: 8.0 memory length: 425036 epsilon: 0.1588227400086959 s  
teps: 406 lr: 2.5600000000000001e-06 evaluation reward: 6.44  
episode: 1728 score: 4.0 memory length: 425316 epsilon: 0.1582683400086924 s  
teps: 280 lr: 2.5600000000000001e-06 evaluation reward: 6.43  
episode: 1729 score: 14.0 memory length: 425860 epsilon: 0.15719122000868557 s  
steps: 544 lr: 2.5600000000000001e-06 evaluation reward: 6.52  
episode: 1730 score: 8.0 memory length: 426310 epsilon: 0.15630022000867994 s  
steps: 450 lr: 2.5600000000000001e-06 evaluation reward: 6.55  
episode: 1731 score: 4.0 memory length: 426609 epsilon: 0.1557082000086762 s  
teps: 299 lr: 2.5600000000000001e-06 evaluation reward: 6.49  
episode: 1732 score: 6.0 memory length: 426925 epsilon: 0.15508252000867223 s  
steps: 316 lr: 2.5600000000000001e-06 evaluation reward: 6.51  
episode: 1733 score: 7.0 memory length: 427312 epsilon: 0.15431626000866738 s  
steps: 387 lr: 2.5600000000000001e-06 evaluation reward: 6.51  
episode: 1734 score: 12.0 memory length: 427735 epsilon: 0.15347872000866208 s  
steps: 423 lr: 2.5600000000000001e-06 evaluation reward: 6.59  
episode: 1735 score: 5.0 memory length: 428023 epsilon: 0.15290848000865848 s  
steps: 288 lr: 2.5600000000000001e-06 evaluation reward: 6.6  
episode: 1736 score: 10.0 memory length: 428470 epsilon: 0.15202342000865288 s  
steps: 447 lr: 2.5600000000000001e-06 evaluation reward: 6.64  
episode: 1737 score: 4.0 memory length: 428728 epsilon: 0.15151258000864964 s  
steps: 258 lr: 2.5600000000000001e-06 evaluation reward: 6.64  
episode: 1738 score: 4.0 memory length: 428970 epsilon: 0.1510334200086466 s  
teps: 242 lr: 2.5600000000000001e-06 evaluation reward: 6.6  
episode: 1739 score: 6.0 memory length: 429347 epsilon: 0.1502869600086419 s  
teps: 377 lr: 2.5600000000000001e-06 evaluation reward: 6.59

episode: 1740 score: 8.0 memory length: 429763 epsilon: 0.14946328000863668  
 steps: 416 lr: 2.5600000000000001e-06 evaluation reward: 6.59  
 episode: 1741 score: 6.0 memory length: 430116 epsilon: 0.14876434000863226  
 steps: 353 lr: 2.5600000000000001e-06 evaluation reward: 6.58  
 episode: 1742 score: 5.0 memory length: 430440 epsilon: 0.1481228200086282 s  
 teps: 324 lr: 2.5600000000000001e-06 evaluation reward: 6.57  
 episode: 1743 score: 5.0 memory length: 430748 epsilon: 0.14751298000862434  
 steps: 308 lr: 2.5600000000000001e-06 evaluation reward: 6.56  
 episode: 1744 score: 7.0 memory length: 431133 epsilon: 0.14675068000861952  
 steps: 385 lr: 2.5600000000000001e-06 evaluation reward: 6.58  
 episode: 1745 score: 6.0 memory length: 431471 epsilon: 0.14608144000861528  
 steps: 338 lr: 2.5600000000000001e-06 evaluation reward: 6.56  
 episode: 1746 score: 5.0 memory length: 431762 epsilon: 0.14550526000861164  
 steps: 291 lr: 2.5600000000000001e-06 evaluation reward: 6.56  
 episode: 1747 score: 7.0 memory length: 432128 epsilon: 0.14478058000860705  
 steps: 366 lr: 2.5600000000000001e-06 evaluation reward: 6.57  
 episode: 1748 score: 4.0 memory length: 432369 epsilon: 0.14430340000860403  
 steps: 241 lr: 2.5600000000000001e-06 evaluation reward: 6.57  
 episode: 1749 score: 4.0 memory length: 432650 epsilon: 0.1437470200086005 s  
 teps: 281 lr: 2.5600000000000001e-06 evaluation reward: 6.56  
 episode: 1750 score: 4.0 memory length: 432927 epsilon: 0.14319856000859704  
 steps: 277 lr: 2.5600000000000001e-06 evaluation reward: 6.52  
 episode: 1751 score: 6.0 memory length: 433259 epsilon: 0.14254120000859288  
 steps: 332 lr: 2.5600000000000001e-06 evaluation reward: 6.52  
 episode: 1752 score: 6.0 memory length: 433592 epsilon: 0.1418818600085887 s  
 teps: 333 lr: 2.5600000000000001e-06 evaluation reward: 6.52  
 episode: 1753 score: 6.0 memory length: 433936 epsilon: 0.1412007400085844 s  
 teps: 344 lr: 2.5600000000000001e-06 evaluation reward: 6.48  
 episode: 1754 score: 6.0 memory length: 434283 epsilon: 0.14051368000858006  
 steps: 347 lr: 2.5600000000000001e-06 evaluation reward: 6.5  
 episode: 1755 score: 7.0 memory length: 434685 epsilon: 0.13971772000857502  
 steps: 402 lr: 2.5600000000000001e-06 evaluation reward: 6.5  
 episode: 1756 score: 11.0 memory length: 435144 epsilon: 0.13880890000856927  
 steps: 459 lr: 2.5600000000000001e-06 evaluation reward: 6.55  
 episode: 1757 score: 6.0 memory length: 435467 epsilon: 0.13816936000856522  
 steps: 323 lr: 2.5600000000000001e-06 evaluation reward: 6.5  
 episode: 1758 score: 9.0 memory length: 435971 epsilon: 0.1371714400085589 s  
 teps: 504 lr: 2.5600000000000001e-06 evaluation reward: 6.54  
 episode: 1759 score: 10.0 memory length: 436466 epsilon: 0.1361913400085527  
 steps: 495 lr: 2.5600000000000001e-06 evaluation reward: 6.57  
 episode: 1760 score: 11.0 memory length: 436930 epsilon: 0.1352726200085469  
 steps: 464 lr: 2.5600000000000001e-06 evaluation reward: 6.63  
 episode: 1761 score: 7.0 memory length: 437342 epsilon: 0.13445686000854173  
 steps: 412 lr: 2.5600000000000001e-06 evaluation reward: 6.64  
 episode: 1762 score: 10.0 memory length: 437711 epsilon: 0.1337262400085371  
 steps: 369 lr: 2.5600000000000001e-06 evaluation reward: 6.69  
 episode: 1763 score: 5.0 memory length: 438015 epsilon: 0.1331243200085333 s  
 teps: 304 lr: 2.5600000000000001e-06 evaluation reward: 6.68  
 episode: 1764 score: 5.0 memory length: 438303 epsilon: 0.1325540800085297 s  
 teps: 288 lr: 2.5600000000000001e-06 evaluation reward: 6.68  
 episode: 1765 score: 8.0 memory length: 438713 epsilon: 0.13174228000852456  
 steps: 410 lr: 2.5600000000000001e-06 evaluation reward: 6.72  
 episode: 1766 score: 5.0 memory length: 439037 epsilon: 0.1311007600085205 s  
 teps: 324 lr: 2.5600000000000001e-06 evaluation reward: 6.7  
 episode: 1767 score: 5.0 memory length: 439324 epsilon: 0.1305325000085169 s  
 teps: 287 lr: 2.5600000000000001e-06 evaluation reward: 6.71  
 episode: 1768 score: 5.0 memory length: 439596 epsilon: 0.1299939400085135 s  
 teps: 272 lr: 2.5600000000000001e-06 evaluation reward: 6.7  
 episode: 1769 score: 4.0 memory length: 439837 epsilon: 0.12951676000851048  
 steps: 241 lr: 2.5600000000000001e-06 evaluation reward: 6.61

episode: 1770 score: 6.0 memory length: 440178 epsilon: 0.1288415800085062 s  
teps: 341 lr: 2.5600000000000001e-06 evaluation reward: 6.58  
episode: 1771 score: 5.0 memory length: 440466 epsilon: 0.1282713400085026 s  
teps: 288 lr: 2.5600000000000001e-06 evaluation reward: 6.54  
episode: 1772 score: 11.0 memory length: 441022 epsilon: 0.12717046000849563  
steps: 556 lr: 2.5600000000000001e-06 evaluation reward: 6.6  
episode: 1773 score: 6.0 memory length: 441400 epsilon: 0.1264220200084909 s  
teps: 378 lr: 2.5600000000000001e-06 evaluation reward: 6.59  
episode: 1774 score: 7.0 memory length: 441786 epsilon: 0.12565774000848606  
steps: 386 lr: 2.5600000000000001e-06 evaluation reward: 6.58  
episode: 1775 score: 6.0 memory length: 442120 epsilon: 0.1249964200084819 s  
teps: 334 lr: 2.5600000000000001e-06 evaluation reward: 6.6  
episode: 1776 score: 6.0 memory length: 442460 epsilon: 0.12432322000848237  
steps: 340 lr: 2.5600000000000001e-06 evaluation reward: 6.59  
episode: 1777 score: 7.0 memory length: 442862 epsilon: 0.12352726000848291  
steps: 402 lr: 2.5600000000000001e-06 evaluation reward: 6.6  
episode: 1778 score: 5.0 memory length: 443170 epsilon: 0.12291742000848332  
steps: 308 lr: 2.5600000000000001e-06 evaluation reward: 6.52  
episode: 1779 score: 7.0 memory length: 443559 epsilon: 0.12214720000848385  
steps: 389 lr: 2.5600000000000001e-06 evaluation reward: 6.52  
episode: 1780 score: 8.0 memory length: 443990 epsilon: 0.12129382000848443  
steps: 431 lr: 2.5600000000000001e-06 evaluation reward: 6.54  
episode: 1781 score: 8.0 memory length: 444426 epsilon: 0.12043054000848502  
steps: 436 lr: 2.5600000000000001e-06 evaluation reward: 6.53  
episode: 1782 score: 10.0 memory length: 444781 epsilon: 0.1197276400084855  
steps: 355 lr: 2.5600000000000001e-06 evaluation reward: 6.57  
episode: 1783 score: 10.0 memory length: 445335 epsilon: 0.11863072000848625  
steps: 554 lr: 2.5600000000000001e-06 evaluation reward: 6.63  
episode: 1784 score: 9.0 memory length: 445788 epsilon: 0.11773378000848686  
steps: 453 lr: 2.5600000000000001e-06 evaluation reward: 6.62  
episode: 1785 score: 7.0 memory length: 446179 epsilon: 0.11695960000848739  
steps: 391 lr: 2.5600000000000001e-06 evaluation reward: 6.61  
episode: 1786 score: 6.0 memory length: 446535 epsilon: 0.11625472000848787  
steps: 356 lr: 2.5600000000000001e-06 evaluation reward: 6.6  
episode: 1787 score: 9.0 memory length: 447004 epsilon: 0.1153261000084885 s  
teps: 469 lr: 2.5600000000000001e-06 evaluation reward: 6.63  
episode: 1788 score: 4.0 memory length: 447244 epsilon: 0.11485090000848883  
steps: 240 lr: 2.5600000000000001e-06 evaluation reward: 6.63  
episode: 1789 score: 6.0 memory length: 447589 epsilon: 0.11416780000848929  
steps: 345 lr: 2.5600000000000001e-06 evaluation reward: 6.58  
episode: 1790 score: 6.0 memory length: 447944 epsilon: 0.11346490000848977  
steps: 355 lr: 2.5600000000000001e-06 evaluation reward: 6.59  
episode: 1791 score: 4.0 memory length: 448202 epsilon: 0.11295406000849012  
steps: 258 lr: 2.5600000000000001e-06 evaluation reward: 6.49  
episode: 1792 score: 5.0 memory length: 448492 epsilon: 0.11237986000849051  
steps: 290 lr: 2.5600000000000001e-06 evaluation reward: 6.5  
episode: 1793 score: 8.0 memory length: 448910 epsilon: 0.11155222000849108  
steps: 418 lr: 2.5600000000000001e-06 evaluation reward: 6.51  
episode: 1794 score: 5.0 memory length: 449199 epsilon: 0.11098000000849147  
steps: 289 lr: 2.5600000000000001e-06 evaluation reward: 6.51  
episode: 1795 score: 5.0 memory length: 449482 epsilon: 0.11041966000849185  
steps: 283 lr: 2.5600000000000001e-06 evaluation reward: 6.5  
episode: 1796 score: 5.0 memory length: 449770 epsilon: 0.10984942000849224  
steps: 288 lr: 2.5600000000000001e-06 evaluation reward: 6.49  
episode: 1797 score: 9.0 memory length: 450240 epsilon: 0.10891882000849287  
steps: 470 lr: 2.5600000000000001e-06 evaluation reward: 6.44  
episode: 1798 score: 10.0 memory length: 450732 epsilon: 0.10794466000849354  
steps: 492 lr: 2.5600000000000001e-06 evaluation reward: 6.47  
episode: 1799 score: 4.0 memory length: 450990 epsilon: 0.10743382000849389  
steps: 258 lr: 2.5600000000000001e-06 evaluation reward: 6.44

```

episode: 1800    score: 9.0    memory length: 451453    epsilon: 0.10651708000849451
steps: 463      lr: 2.5600000000000001e-06    evaluation reward: 6.48
episode: 1801    score: 5.0    memory length: 451723    epsilon: 0.10598248000849488
steps: 270      lr: 2.5600000000000001e-06    evaluation reward: 6.48
episode: 1802    score: 8.0    memory length: 452177    epsilon: 0.10508356000849549
steps: 454      lr: 2.5600000000000001e-06    evaluation reward: 6.49
episode: 1803    score: 7.0    memory length: 452511    epsilon: 0.10442224000849594
steps: 334      lr: 2.5600000000000001e-06    evaluation reward: 6.51
episode: 1804    score: 10.0   memory length: 453012    epsilon: 0.10343026000849662
steps: 501      lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 1805    score: 8.0    memory length: 453467    epsilon: 0.10252936000849723
steps: 455      lr: 2.5600000000000001e-06    evaluation reward: 6.58
episode: 1806    score: 4.0    memory length: 453709    epsilon: 0.10205020000849756
steps: 242      lr: 2.5600000000000001e-06    evaluation reward: 6.55
episode: 1807    score: 8.0    memory length: 454098    epsilon: 0.10127998000849808
steps: 389      lr: 2.5600000000000001e-06    evaluation reward: 6.58
episode: 1808    score: 5.0    memory length: 454369    epsilon: 0.10074340000849845
steps: 271      lr: 2.5600000000000001e-06    evaluation reward: 6.58
episode: 1809    score: 6.0    memory length: 454681    epsilon: 0.10012564000849887
steps: 312      lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 1810    score: 5.0    memory length: 454985    epsilon: 0.09952372000849928
steps: 304      lr: 2.5600000000000001e-06    evaluation reward: 6.55
episode: 1811    score: 7.0    memory length: 455336    epsilon: 0.09882874000849975
steps: 351      lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 1812    score: 7.0    memory length: 455748    epsilon: 0.09801298000850031
steps: 412      lr: 2.5600000000000001e-06    evaluation reward: 6.58
episode: 1813    score: 6.0    memory length: 456104    epsilon: 0.09730810000850079
steps: 356      lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 1814    score: 7.0    memory length: 456470    epsilon: 0.09658342000850129
steps: 366      lr: 2.5600000000000001e-06    evaluation reward: 6.59
episode: 1815    score: 4.0    memory length: 456712    epsilon: 0.09610426000850161
steps: 242      lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 1816    score: 4.0    memory length: 456953    epsilon: 0.09562708000850194
steps: 241      lr: 2.5600000000000001e-06    evaluation reward: 6.55
episode: 1817    score: 5.0    memory length: 457244    epsilon: 0.09505090000850233
steps: 291      lr: 2.5600000000000001e-06    evaluation reward: 6.55
episode: 1818    score: 5.0    memory length: 457515    epsilon: 0.0945143200085027    s
steps: 271      lr: 2.5600000000000001e-06    evaluation reward: 6.55
episode: 1819    score: 10.0   memory length: 458001    epsilon: 0.09355204000850335
steps: 486      lr: 2.5600000000000001e-06    evaluation reward: 6.58
episode: 1820    score: 4.0    memory length: 458259    epsilon: 0.0930412000085037    s
steps: 258      lr: 2.5600000000000001e-06    evaluation reward: 6.56
episode: 1821    score: 6.0    memory length: 458597    epsilon: 0.09237196000850416
steps: 338      lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 1822    score: 8.0    memory length: 458988    epsilon: 0.09159778000850469
steps: 391      lr: 2.5600000000000001e-06    evaluation reward: 6.61
episode: 1823    score: 9.0    memory length: 459312    epsilon: 0.09095626000850512
steps: 324      lr: 2.5600000000000001e-06    evaluation reward: 6.62
episode: 1824    score: 5.0    memory length: 459623    epsilon: 0.09034048000850554
steps: 311      lr: 2.5600000000000001e-06    evaluation reward: 6.62
episode: 1825    score: 4.0    memory length: 459883    epsilon: 0.0898256800085059    s
steps: 260      lr: 2.5600000000000001e-06    evaluation reward: 6.59
episode: 1826    score: 7.0    memory length: 460234    epsilon: 0.08913070000850637
steps: 351      lr: 2.5600000000000001e-06    evaluation reward: 6.62
episode: 1827    score: 8.0    memory length: 460661    epsilon: 0.08828524000850695
steps: 427      lr: 2.5600000000000001e-06    evaluation reward: 6.62
episode: 1828    score: 13.0   memory length: 461139    epsilon: 0.08733880000850759
steps: 478      lr: 2.5600000000000001e-06    evaluation reward: 6.71
episode: 1829    score: 5.0    memory length: 461410    epsilon: 0.08680222000850796
steps: 271      lr: 2.5600000000000001e-06    evaluation reward: 6.62

```

episode: 1830 score: 5.0 memory length: 461717 epsilon: 0.08619436000850837  
 steps: 307 lr: 2.5600000000000001e-06 evaluation reward: 6.59  
 episode: 1831 score: 6.0 memory length: 462051 epsilon: 0.08553304000850882  
 steps: 334 lr: 2.5600000000000001e-06 evaluation reward: 6.61  
 episode: 1832 score: 11.0 memory length: 462467 epsilon: 0.08470936000850939  
 steps: 416 lr: 2.5600000000000001e-06 evaluation reward: 6.66  
 episode: 1833 score: 8.0 memory length: 462895 epsilon: 0.08386192000850996  
 steps: 428 lr: 2.5600000000000001e-06 evaluation reward: 6.67  
 episode: 1834 score: 8.0 memory length: 463317 epsilon: 0.08302636000851053  
 steps: 422 lr: 2.5600000000000001e-06 evaluation reward: 6.63  
 episode: 1835 score: 8.0 memory length: 463742 epsilon: 0.08218486000851111  
 steps: 425 lr: 2.5600000000000001e-06 evaluation reward: 6.66  
 episode: 1836 score: 4.0 memory length: 463984 epsilon: 0.08170570000851143  
 steps: 242 lr: 2.5600000000000001e-06 evaluation reward: 6.6  
 episode: 1837 score: 6.0 memory length: 464357 epsilon: 0.08096716000851194  
 steps: 373 lr: 2.5600000000000001e-06 evaluation reward: 6.62  
 episode: 1838 score: 10.0 memory length: 464844 epsilon: 0.0800029000085126  
 steps: 487 lr: 2.5600000000000001e-06 evaluation reward: 6.68  
 episode: 1839 score: 4.0 memory length: 465104 epsilon: 0.07948810000851295  
 steps: 260 lr: 2.5600000000000001e-06 evaluation reward: 6.66  
 episode: 1840 score: 7.0 memory length: 465470 epsilon: 0.07876342000851344  
 steps: 366 lr: 2.5600000000000001e-06 evaluation reward: 6.65  
 episode: 1841 score: 7.0 memory length: 465855 epsilon: 0.07800112000851396  
 steps: 385 lr: 2.5600000000000001e-06 evaluation reward: 6.66  
 episode: 1842 score: 6.0 memory length: 466177 epsilon: 0.0773635600085144 s  
 steps: 322 lr: 2.5600000000000001e-06 evaluation reward: 6.67  
 episode: 1843 score: 11.0 memory length: 466710 epsilon: 0.07630822000851512  
 steps: 533 lr: 2.5600000000000001e-06 evaluation reward: 6.73  
 episode: 1844 score: 11.0 memory length: 467188 epsilon: 0.07536178000851576  
 steps: 478 lr: 2.5600000000000001e-06 evaluation reward: 6.77  
 episode: 1845 score: 4.0 memory length: 467448 epsilon: 0.07484698000851611  
 steps: 260 lr: 2.5600000000000001e-06 evaluation reward: 6.75  
 episode: 1846 score: 5.0 memory length: 467721 epsilon: 0.07430644000851648  
 steps: 273 lr: 2.5600000000000001e-06 evaluation reward: 6.75  
 episode: 1847 score: 6.0 memory length: 468060 epsilon: 0.07363522000851694  
 steps: 339 lr: 2.5600000000000001e-06 evaluation reward: 6.74  
 episode: 1848 score: 13.0 memory length: 468506 epsilon: 0.07275214000851754  
 steps: 446 lr: 2.5600000000000001e-06 evaluation reward: 6.83  
 episode: 1849 score: 9.0 memory length: 468950 epsilon: 0.07187302000851814  
 steps: 444 lr: 2.5600000000000001e-06 evaluation reward: 6.88  
 episode: 1850 score: 12.0 memory length: 469437 epsilon: 0.0709087600085188  
 steps: 487 lr: 2.5600000000000001e-06 evaluation reward: 6.96  
 episode: 1851 score: 4.0 memory length: 469695 epsilon: 0.07039792000851915  
 steps: 258 lr: 2.5600000000000001e-06 evaluation reward: 6.94  
 episode: 1852 score: 7.0 memory length: 470079 epsilon: 0.06963760000851967  
 steps: 384 lr: 2.5600000000000001e-06 evaluation reward: 6.95  
 episode: 1853 score: 12.0 memory length: 470489 epsilon: 0.06882580000852022  
 steps: 410 lr: 2.5600000000000001e-06 evaluation reward: 7.01  
 episode: 1854 score: 7.0 memory length: 470893 epsilon: 0.06802588000852076  
 steps: 404 lr: 2.5600000000000001e-06 evaluation reward: 7.02  
 episode: 1855 score: 4.0 memory length: 471134 epsilon: 0.06754870000852109  
 steps: 241 lr: 2.5600000000000001e-06 evaluation reward: 6.99  
 episode: 1856 score: 9.0 memory length: 471577 epsilon: 0.06667156000852169  
 steps: 443 lr: 2.5600000000000001e-06 evaluation reward: 6.97  
 episode: 1857 score: 4.0 memory length: 471855 epsilon: 0.06612112000852206  
 steps: 278 lr: 2.5600000000000001e-06 evaluation reward: 6.95  
 episode: 1858 score: 5.0 memory length: 472142 epsilon: 0.06555286000852245  
 steps: 287 lr: 2.5600000000000001e-06 evaluation reward: 6.91  
 episode: 1859 score: 6.0 memory length: 472483 epsilon: 0.06487768000852291  
 steps: 341 lr: 2.5600000000000001e-06 evaluation reward: 6.87

episode: 1860 score: 8.0 memory length: 472901 epsilon: 0.06405004000852348  
 steps: 418 lr: 2.5600000000000001e-06 evaluation reward: 6.84  
 episode: 1861 score: 10.0 memory length: 473380 epsilon: 0.06310162000852412  
 steps: 479 lr: 2.5600000000000001e-06 evaluation reward: 6.87  
 episode: 1862 score: 8.0 memory length: 473761 epsilon: 0.06234724000852464  
 steps: 381 lr: 2.5600000000000001e-06 evaluation reward: 6.85  
 episode: 1863 score: 5.0 memory length: 474051 epsilon: 0.06177304000852503  
 steps: 290 lr: 2.5600000000000001e-06 evaluation reward: 6.85  
 episode: 1864 score: 10.0 memory length: 474422 epsilon: 0.06103846000852553  
 steps: 371 lr: 2.5600000000000001e-06 evaluation reward: 6.9  
 episode: 1865 score: 7.0 memory length: 474812 epsilon: 0.06026626000852606  
 steps: 390 lr: 2.5600000000000001e-06 evaluation reward: 6.89  
 episode: 1866 score: 4.0 memory length: 475053 epsilon: 0.05978908000852638  
 steps: 241 lr: 2.5600000000000001e-06 evaluation reward: 6.88  
 episode: 1867 score: 3.0 memory length: 475264 epsilon: 0.05937130000852667  
 steps: 211 lr: 2.5600000000000001e-06 evaluation reward: 6.86  
 episode: 1868 score: 4.0 memory length: 475502 epsilon: 0.05890006000852699  
 steps: 238 lr: 2.5600000000000001e-06 evaluation reward: 6.85  
 episode: 1869 score: 7.0 memory length: 475890 epsilon: 0.05813182000852751  
 steps: 388 lr: 2.5600000000000001e-06 evaluation reward: 6.88  
 episode: 1870 score: 5.0 memory length: 476163 epsilon: 0.05759128000852788  
 steps: 273 lr: 2.5600000000000001e-06 evaluation reward: 6.87  
 episode: 1871 score: 7.0 memory length: 476541 epsilon: 0.05684284000852839  
 steps: 378 lr: 2.5600000000000001e-06 evaluation reward: 6.89  
 episode: 1872 score: 5.0 memory length: 476849 epsilon: 0.05623300000852881  
 steps: 308 lr: 2.5600000000000001e-06 evaluation reward: 6.83  
 episode: 1873 score: 8.0 memory length: 477229 epsilon: 0.05548060000852932  
 steps: 380 lr: 2.5600000000000001e-06 evaluation reward: 6.85  
 episode: 1874 score: 7.0 memory length: 477581 epsilon: 0.054783640008529796  
 steps: 352 lr: 2.5600000000000001e-06 evaluation reward: 6.85  
 episode: 1875 score: 7.0 memory length: 477947 epsilon: 0.05405896000853029  
 steps: 366 lr: 2.5600000000000001e-06 evaluation reward: 6.86  
 episode: 1876 score: 5.0 memory length: 478235 epsilon: 0.05348872000853068  
 steps: 288 lr: 2.5600000000000001e-06 evaluation reward: 6.85  
 episode: 1877 score: 7.0 memory length: 478614 epsilon: 0.05273830000853119  
 steps: 379 lr: 2.5600000000000001e-06 evaluation reward: 6.85  
 episode: 1878 score: 6.0 memory length: 478968 epsilon: 0.05203738000853167  
 steps: 354 lr: 2.5600000000000001e-06 evaluation reward: 6.86  
 episode: 1879 score: 6.0 memory length: 479307 epsilon: 0.05136616000853213  
 steps: 339 lr: 2.5600000000000001e-06 evaluation reward: 6.85  
 episode: 1880 score: 7.0 memory length: 479695 epsilon: 0.05059792000853265  
 steps: 388 lr: 2.5600000000000001e-06 evaluation reward: 6.84  
 episode: 1881 score: 6.0 memory length: 480048 epsilon: 0.04989898000853313  
 steps: 353 lr: 2.5600000000000001e-06 evaluation reward: 6.82  
 episode: 1882 score: 8.0 memory length: 480465 epsilon: 0.04907332000853369  
 steps: 417 lr: 2.5600000000000001e-06 evaluation reward: 6.8  
 episode: 1883 score: 9.0 memory length: 480936 epsilon: 0.04814074000853433  
 steps: 471 lr: 2.5600000000000001e-06 evaluation reward: 6.79  
 episode: 1884 score: 13.0 memory length: 481444 epsilon: 0.04713490000853501  
 steps: 508 lr: 2.5600000000000001e-06 evaluation reward: 6.83  
 episode: 1885 score: 8.0 memory length: 481850 epsilon: 0.04633102000853556  
 steps: 406 lr: 2.5600000000000001e-06 evaluation reward: 6.84  
 episode: 1886 score: 8.0 memory length: 482265 epsilon: 0.04550932000853612  
 steps: 415 lr: 2.5600000000000001e-06 evaluation reward: 6.86  
 episode: 1887 score: 4.0 memory length: 482545 epsilon: 0.0449549200085365 s  
 steps: 280 lr: 2.5600000000000001e-06 evaluation reward: 6.81  
 episode: 1888 score: 9.0 memory length: 483004 epsilon: 0.04404610000853712  
 steps: 459 lr: 2.5600000000000001e-06 evaluation reward: 6.86  
 episode: 1889 score: 4.0 memory length: 483244 epsilon: 0.043570900008537444  
 steps: 240 lr: 2.5600000000000001e-06 evaluation reward: 6.84

episode: 1890 score: 8.0 memory length: 483654 epsilon: 0.042759100008538 st  
 eps: 410 lr: 2.5600000000000001e-06 evaluation reward: 6.86  
 episode: 1891 score: 4.0 memory length: 483893 epsilon: 0.04228588000853832  
 steps: 239 lr: 2.5600000000000001e-06 evaluation reward: 6.86  
 episode: 1892 score: 8.0 memory length: 484291 epsilon: 0.04149784000853886  
 steps: 398 lr: 2.5600000000000001e-06 evaluation reward: 6.89  
 episode: 1893 score: 9.0 memory length: 484671 epsilon: 0.04074544000853937  
 steps: 380 lr: 2.5600000000000001e-06 evaluation reward: 6.9  
 episode: 1894 score: 8.0 memory length: 485109 epsilon: 0.03987820000853996  
 steps: 438 lr: 2.5600000000000001e-06 evaluation reward: 6.93  
 episode: 1895 score: 9.0 memory length: 485558 epsilon: 0.03898918000854057  
 steps: 449 lr: 2.5600000000000001e-06 evaluation reward: 6.97  
 episode: 1896 score: 4.0 memory length: 485799 epsilon: 0.038512000008540895  
 steps: 241 lr: 2.5600000000000001e-06 evaluation reward: 6.96  
 episode: 1897 score: 6.0 memory length: 486104 epsilon: 0.03790810000854131  
 steps: 305 lr: 2.5600000000000001e-06 evaluation reward: 6.93  
 episode: 1898 score: 8.0 memory length: 486555 epsilon: 0.037015120008541916  
 steps: 451 lr: 2.5600000000000001e-06 evaluation reward: 6.91  
 episode: 1899 score: 6.0 memory length: 486900 epsilon: 0.03633202000854238  
 steps: 345 lr: 2.5600000000000001e-06 evaluation reward: 6.93  
 episode: 1900 score: 9.0 memory length: 487386 epsilon: 0.03536974000854304  
 steps: 486 lr: 2.5600000000000001e-06 evaluation reward: 6.93  
 episode: 1901 score: 7.0 memory length: 487754 epsilon: 0.034641100008543535  
 steps: 368 lr: 2.5600000000000001e-06 evaluation reward: 6.95  
 episode: 1902 score: 11.0 memory length: 488165 epsilon: 0.03382732000854409  
 steps: 411 lr: 2.5600000000000001e-06 evaluation reward: 6.98  
 episode: 1903 score: 13.0 memory length: 488682 epsilon: 0.03280366000854479  
 steps: 517 lr: 2.5600000000000001e-06 evaluation reward: 7.04  
 episode: 1904 score: 6.0 memory length: 489022 epsilon: 0.03213046000854525  
 steps: 340 lr: 2.5600000000000001e-06 evaluation reward: 7.0  
 episode: 1905 score: 7.0 memory length: 489391 epsilon: 0.031399840008545746  
 steps: 369 lr: 2.5600000000000001e-06 evaluation reward: 6.99  
 episode: 1906 score: 6.0 memory length: 489730 epsilon: 0.030728620008546204  
 steps: 339 lr: 2.5600000000000001e-06 evaluation reward: 7.01  
 episode: 1907 score: 9.0 memory length: 490188 epsilon: 0.029821780008546822  
 steps: 458 lr: 2.5600000000000001e-06 evaluation reward: 7.02  
 episode: 1908 score: 4.0 memory length: 490461 epsilon: 0.02928124000854719  
 steps: 273 lr: 2.5600000000000001e-06 evaluation reward: 7.01  
 episode: 1909 score: 8.0 memory length: 490879 epsilon: 0.028453600008547755  
 steps: 418 lr: 2.5600000000000001e-06 evaluation reward: 7.03  
 episode: 1910 score: 7.0 memory length: 491249 epsilon: 0.027721000008548255  
 steps: 370 lr: 2.5600000000000001e-06 evaluation reward: 7.05  
 episode: 1911 score: 9.0 memory length: 491750 epsilon: 0.02672902000854893  
 steps: 501 lr: 2.5600000000000001e-06 evaluation reward: 7.07  
 episode: 1912 score: 7.0 memory length: 492159 epsilon: 0.025919200008549484  
 steps: 409 lr: 2.5600000000000001e-06 evaluation reward: 7.07  
 episode: 1913 score: 11.0 memory length: 492625 epsilon: 0.024996520008550113  
 steps: 466 lr: 2.5600000000000001e-06 evaluation reward: 7.12  
 episode: 1914 score: 5.0 memory length: 492949 epsilon: 0.02435500000855055  
 steps: 324 lr: 2.5600000000000001e-06 evaluation reward: 7.1  
 episode: 1915 score: 11.0 memory length: 493405 epsilon: 0.023452120008551167  
 steps: 456 lr: 2.5600000000000001e-06 evaluation reward: 7.17  
 episode: 1916 score: 5.0 memory length: 493692 epsilon: 0.022883860008551554  
 steps: 287 lr: 2.5600000000000001e-06 evaluation reward: 7.18  
 episode: 1917 score: 6.0 memory length: 494025 epsilon: 0.022224520008552004  
 steps: 333 lr: 2.5600000000000001e-06 evaluation reward: 7.19  
 episode: 1918 score: 6.0 memory length: 494347 epsilon: 0.02158696000855244  
 steps: 322 lr: 2.5600000000000001e-06 evaluation reward: 7.2  
 episode: 1919 score: 4.0 memory length: 494606 epsilon: 0.02107414000855279  
 steps: 259 lr: 2.5600000000000001e-06 evaluation reward: 7.14

```
episode: 1920 score: 4.0 memory length: 494864 epsilon: 0.020563300008553137
steps: 258 lr: 2.5600000000000001e-06 evaluation reward: 7.14
episode: 1921 score: 7.0 memory length: 495253 epsilon: 0.019793080008553662
steps: 389 lr: 2.5600000000000001e-06 evaluation reward: 7.15
episode: 1922 score: 8.0 memory length: 495658 epsilon: 0.01899118000855421
steps: 405 lr: 2.5600000000000001e-06 evaluation reward: 7.15
episode: 1923 score: 7.0 memory length: 496011 epsilon: 0.018292240008554686
steps: 353 lr: 2.5600000000000001e-06 evaluation reward: 7.13
episode: 1924 score: 6.0 memory length: 496331 epsilon: 0.017658640008555118
steps: 320 lr: 2.5600000000000001e-06 evaluation reward: 7.14
episode: 1925 score: 6.0 memory length: 496651 epsilon: 0.01702504000855555
steps: 320 lr: 2.5600000000000001e-06 evaluation reward: 7.16
episode: 1926 score: 4.0 memory length: 496892 epsilon: 0.016547860008555876
steps: 241 lr: 2.5600000000000001e-06 evaluation reward: 7.13
episode: 1927 score: 10.0 memory length: 497265 epsilon: 0.01580932000855638
steps: 373 lr: 2.5600000000000001e-06 evaluation reward: 7.15
episode: 1928 score: 11.0 memory length: 497822 epsilon: 0.014706460008556327
steps: 557 lr: 2.5600000000000001e-06 evaluation reward: 7.13
episode: 1929 score: 10.0 memory length: 498320 epsilon: 0.013720420008556135
steps: 498 lr: 2.5600000000000001e-06 evaluation reward: 7.18
episode: 1930 score: 10.0 memory length: 498786 epsilon: 0.012797740008555956
steps: 466 lr: 2.5600000000000001e-06 evaluation reward: 7.23
episode: 1931 score: 17.0 memory length: 499250 epsilon: 0.011879020008555778
steps: 464 lr: 2.5600000000000001e-06 evaluation reward: 7.34
episode: 1932 score: 6.0 memory length: 499625 epsilon: 0.011136520008555634
steps: 375 lr: 2.5600000000000001e-06 evaluation reward: 7.29
episode: 1933 score: 10.0 memory length: 500165 epsilon: 0.010067320008555427
steps: 540 lr: 2.5600000000000001e-06 evaluation reward: 7.31
episode: 1934 score: 6.0 memory length: 500485 epsilon: 0.009998020008555413
steps: 320 lr: 1.0240000000000005e-06 evaluation reward: 7.29
episode: 1935 score: 6.0 memory length: 500825 epsilon: 0.009998020008555413
steps: 340 lr: 1.0240000000000005e-06 evaluation reward: 7.27
episode: 1936 score: 6.0 memory length: 501162 epsilon: 0.009998020008555413
steps: 337 lr: 1.0240000000000005e-06 evaluation reward: 7.29
episode: 1937 score: 7.0 memory length: 501506 epsilon: 0.009998020008555413
steps: 344 lr: 1.0240000000000005e-06 evaluation reward: 7.3
episode: 1938 score: 6.0 memory length: 501827 epsilon: 0.009998020008555413
steps: 321 lr: 1.0240000000000005e-06 evaluation reward: 7.26
episode: 1939 score: 6.0 memory length: 502163 epsilon: 0.009998020008555413
steps: 336 lr: 1.0240000000000005e-06 evaluation reward: 7.28
episode: 1940 score: 14.0 memory length: 502692 epsilon: 0.009998020008555413
steps: 529 lr: 1.0240000000000005e-06 evaluation reward: 7.35
episode: 1941 score: 7.0 memory length: 503058 epsilon: 0.009998020008555413
steps: 366 lr: 1.0240000000000005e-06 evaluation reward: 7.35
episode: 1942 score: 7.0 memory length: 503460 epsilon: 0.009998020008555413
steps: 402 lr: 1.0240000000000005e-06 evaluation reward: 7.36
episode: 1943 score: 5.0 memory length: 503731 epsilon: 0.009998020008555413
steps: 271 lr: 1.0240000000000005e-06 evaluation reward: 7.3
episode: 1944 score: 10.0 memory length: 504211 epsilon: 0.009998020008555413
steps: 480 lr: 1.0240000000000005e-06 evaluation reward: 7.29
episode: 1945 score: 8.0 memory length: 504677 epsilon: 0.009998020008555413
steps: 466 lr: 1.0240000000000005e-06 evaluation reward: 7.33
episode: 1946 score: 6.0 memory length: 505028 epsilon: 0.009998020008555413
steps: 351 lr: 1.0240000000000005e-06 evaluation reward: 7.34
episode: 1947 score: 6.0 memory length: 505361 epsilon: 0.009998020008555413
steps: 333 lr: 1.0240000000000005e-06 evaluation reward: 7.34
episode: 1948 score: 11.0 memory length: 505943 epsilon: 0.009998020008555413
steps: 582 lr: 1.0240000000000005e-06 evaluation reward: 7.32
episode: 1949 score: 6.0 memory length: 506276 epsilon: 0.009998020008555413
steps: 333 lr: 1.0240000000000005e-06 evaluation reward: 7.29
```

```
episode: 1950    score: 11.0    memory length: 506759    epsilon: 0.009998020008555413
steps: 483      lr: 1.0240000000000005e-06    evaluation reward: 7.28
episode: 1951    score: 11.0    memory length: 507307    epsilon: 0.009998020008555413
steps: 548      lr: 1.0240000000000005e-06    evaluation reward: 7.35
episode: 1952    score: 8.0     memory length: 507734    epsilon: 0.009998020008555413
steps: 427      lr: 1.0240000000000005e-06    evaluation reward: 7.36
episode: 1953    score: 9.0     memory length: 508110    epsilon: 0.009998020008555413
steps: 376      lr: 1.0240000000000005e-06    evaluation reward: 7.33
episode: 1954    score: 7.0     memory length: 508509    epsilon: 0.009998020008555413
steps: 399      lr: 1.0240000000000005e-06    evaluation reward: 7.33
episode: 1955    score: 10.0    memory length: 508989    epsilon: 0.009998020008555413
steps: 480      lr: 1.0240000000000005e-06    evaluation reward: 7.39
episode: 1956    score: 6.0     memory length: 509309    epsilon: 0.009998020008555413
steps: 320      lr: 1.0240000000000005e-06    evaluation reward: 7.36
episode: 1957    score: 4.0     memory length: 509550    epsilon: 0.009998020008555413
steps: 241      lr: 1.0240000000000005e-06    evaluation reward: 7.36
episode: 1958    score: 4.0     memory length: 509843    epsilon: 0.009998020008555413
steps: 293      lr: 1.0240000000000005e-06    evaluation reward: 7.35
episode: 1959    score: 7.0     memory length: 510241    epsilon: 0.009998020008555413
steps: 398      lr: 1.0240000000000005e-06    evaluation reward: 7.36
episode: 1960    score: 8.0     memory length: 510654    epsilon: 0.009998020008555413
steps: 413      lr: 1.0240000000000005e-06    evaluation reward: 7.36
episode: 1961    score: 7.0     memory length: 511024    epsilon: 0.009998020008555413
steps: 370      lr: 1.0240000000000005e-06    evaluation reward: 7.33
episode: 1962    score: 7.0     memory length: 511412    epsilon: 0.009998020008555413
steps: 388      lr: 1.0240000000000005e-06    evaluation reward: 7.32
episode: 1963    score: 4.0     memory length: 511705    epsilon: 0.009998020008555413
steps: 293      lr: 1.0240000000000005e-06    evaluation reward: 7.31
episode: 1964    score: 7.0     memory length: 512104    epsilon: 0.009998020008555413
steps: 399      lr: 1.0240000000000005e-06    evaluation reward: 7.28
episode: 1965    score: 8.0     memory length: 512534    epsilon: 0.009998020008555413
steps: 430      lr: 1.0240000000000005e-06    evaluation reward: 7.29
episode: 1966    score: 7.0     memory length: 512933    epsilon: 0.009998020008555413
steps: 399      lr: 1.0240000000000005e-06    evaluation reward: 7.32
episode: 1967    score: 6.0     memory length: 513306    epsilon: 0.009998020008555413
steps: 373      lr: 1.0240000000000005e-06    evaluation reward: 7.35
episode: 1968    score: 7.0     memory length: 513730    epsilon: 0.009998020008555413
steps: 424      lr: 1.0240000000000005e-06    evaluation reward: 7.38
episode: 1969    score: 7.0     memory length: 514132    epsilon: 0.009998020008555413
steps: 402      lr: 1.0240000000000005e-06    evaluation reward: 7.38
episode: 1970    score: 12.0    memory length: 514627    epsilon: 0.009998020008555413
steps: 495      lr: 1.0240000000000005e-06    evaluation reward: 7.45
episode: 1971    score: 6.0     memory length: 514978    epsilon: 0.009998020008555413
steps: 351      lr: 1.0240000000000005e-06    evaluation reward: 7.44
episode: 1972    score: 7.0     memory length: 515384    epsilon: 0.009998020008555413
steps: 406      lr: 1.0240000000000005e-06    evaluation reward: 7.46
episode: 1973    score: 9.0     memory length: 515869    epsilon: 0.009998020008555413
steps: 485      lr: 1.0240000000000005e-06    evaluation reward: 7.47
episode: 1974    score: 10.0    memory length: 516331    epsilon: 0.009998020008555413
steps: 462      lr: 1.0240000000000005e-06    evaluation reward: 7.5
episode: 1975    score: 9.0     memory length: 516787    epsilon: 0.009998020008555413
steps: 456      lr: 1.0240000000000005e-06    evaluation reward: 7.52
episode: 1976    score: 9.0     memory length: 517242    epsilon: 0.009998020008555413
steps: 455      lr: 1.0240000000000005e-06    evaluation reward: 7.56
episode: 1977    score: 5.0     memory length: 517568    epsilon: 0.009998020008555413
steps: 326      lr: 1.0240000000000005e-06    evaluation reward: 7.54
episode: 1978    score: 9.0     memory length: 518037    epsilon: 0.009998020008555413
steps: 469      lr: 1.0240000000000005e-06    evaluation reward: 7.57
episode: 1979    score: 7.0     memory length: 518443    epsilon: 0.009998020008555413
steps: 406      lr: 1.0240000000000005e-06    evaluation reward: 7.58
```

episode: 1980 score: 5.0 memory length: 518731 epsilon: 0.009998020008555413  
steps: 288 lr: 1.0240000000000005e-06 evaluation reward: 7.56  
episode: 1981 score: 7.0 memory length: 519153 epsilon: 0.009998020008555413  
steps: 422 lr: 1.0240000000000005e-06 evaluation reward: 7.57  
episode: 1982 score: 5.0 memory length: 519495 epsilon: 0.009998020008555413  
steps: 342 lr: 1.0240000000000005e-06 evaluation reward: 7.54  
episode: 1983 score: 4.0 memory length: 519772 epsilon: 0.009998020008555413  
steps: 277 lr: 1.0240000000000005e-06 evaluation reward: 7.49  
episode: 1984 score: 7.0 memory length: 520193 epsilon: 0.009998020008555413  
steps: 421 lr: 1.0240000000000005e-06 evaluation reward: 7.43  
episode: 1985 score: 9.0 memory length: 520631 epsilon: 0.009998020008555413  
steps: 438 lr: 1.0240000000000005e-06 evaluation reward: 7.44  
episode: 1986 score: 9.0 memory length: 521108 epsilon: 0.009998020008555413  
steps: 477 lr: 1.0240000000000005e-06 evaluation reward: 7.45  
episode: 1987 score: 6.0 memory length: 521464 epsilon: 0.009998020008555413  
steps: 356 lr: 1.0240000000000005e-06 evaluation reward: 7.47  
episode: 1988 score: 5.0 memory length: 521804 epsilon: 0.009998020008555413  
steps: 340 lr: 1.0240000000000005e-06 evaluation reward: 7.43  
episode: 1989 score: 4.0 memory length: 522097 epsilon: 0.009998020008555413  
steps: 293 lr: 1.0240000000000005e-06 evaluation reward: 7.43  
episode: 1990 score: 5.0 memory length: 522440 epsilon: 0.009998020008555413  
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.4  
episode: 1991 score: 5.0 memory length: 522783 epsilon: 0.009998020008555413  
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.41  
episode: 1992 score: 7.0 memory length: 523205 epsilon: 0.009998020008555413  
steps: 422 lr: 1.0240000000000005e-06 evaluation reward: 7.4  
episode: 1993 score: 7.0 memory length: 523575 epsilon: 0.009998020008555413  
steps: 370 lr: 1.0240000000000005e-06 evaluation reward: 7.38  
episode: 1994 score: 6.0 memory length: 523927 epsilon: 0.009998020008555413  
steps: 352 lr: 1.0240000000000005e-06 evaluation reward: 7.36  
episode: 1995 score: 8.0 memory length: 524345 epsilon: 0.009998020008555413  
steps: 418 lr: 1.0240000000000005e-06 evaluation reward: 7.35  
episode: 1996 score: 4.0 memory length: 524641 epsilon: 0.009998020008555413  
steps: 296 lr: 1.0240000000000005e-06 evaluation reward: 7.35  
episode: 1997 score: 5.0 memory length: 524986 epsilon: 0.009998020008555413  
steps: 345 lr: 1.0240000000000005e-06 evaluation reward: 7.34  
episode: 1998 score: 11.0 memory length: 525404 epsilon: 0.009998020008555413  
steps: 418 lr: 1.0240000000000005e-06 evaluation reward: 7.37  
episode: 1999 score: 6.0 memory length: 525737 epsilon: 0.009998020008555413  
steps: 333 lr: 1.0240000000000005e-06 evaluation reward: 7.37  
episode: 2000 score: 10.0 memory length: 526222 epsilon: 0.009998020008555413  
steps: 485 lr: 1.0240000000000005e-06 evaluation reward: 7.38  
episode: 2001 score: 5.0 memory length: 526549 epsilon: 0.009998020008555413  
steps: 327 lr: 1.0240000000000005e-06 evaluation reward: 7.36  
episode: 2002 score: 8.0 memory length: 527007 epsilon: 0.009998020008555413  
steps: 458 lr: 1.0240000000000005e-06 evaluation reward: 7.33  
episode: 2003 score: 7.0 memory length: 527429 epsilon: 0.009998020008555413  
steps: 422 lr: 1.0240000000000005e-06 evaluation reward: 7.27  
episode: 2004 score: 8.0 memory length: 527869 epsilon: 0.009998020008555413  
steps: 440 lr: 1.0240000000000005e-06 evaluation reward: 7.29  
episode: 2005 score: 10.0 memory length: 528349 epsilon: 0.009998020008555413  
steps: 480 lr: 1.0240000000000005e-06 evaluation reward: 7.32  
episode: 2006 score: 7.0 memory length: 528698 epsilon: 0.009998020008555413  
steps: 349 lr: 1.0240000000000005e-06 evaluation reward: 7.33  
episode: 2007 score: 7.0 memory length: 529104 epsilon: 0.009998020008555413  
steps: 406 lr: 1.0240000000000005e-06 evaluation reward: 7.31  
episode: 2008 score: 7.0 memory length: 529485 epsilon: 0.009998020008555413  
steps: 381 lr: 1.0240000000000005e-06 evaluation reward: 7.34  
episode: 2009 score: 9.0 memory length: 529890 epsilon: 0.009998020008555413  
steps: 405 lr: 1.0240000000000005e-06 evaluation reward: 7.35

```
episode: 2010 score: 9.0 memory length: 530374 epsilon: 0.009998020008555413
steps: 484 lr: 1.0240000000000005e-06 evaluation reward: 7.37
episode: 2011 score: 9.0 memory length: 530839 epsilon: 0.009998020008555413
steps: 465 lr: 1.0240000000000005e-06 evaluation reward: 7.37
episode: 2012 score: 6.0 memory length: 531153 epsilon: 0.009998020008555413
steps: 314 lr: 1.0240000000000005e-06 evaluation reward: 7.36
episode: 2013 score: 5.0 memory length: 531436 epsilon: 0.009998020008555413
steps: 283 lr: 1.0240000000000005e-06 evaluation reward: 7.3
episode: 2014 score: 8.0 memory length: 531867 epsilon: 0.009998020008555413
steps: 431 lr: 1.0240000000000005e-06 evaluation reward: 7.33
episode: 2015 score: 7.0 memory length: 532218 epsilon: 0.009998020008555413
steps: 351 lr: 1.0240000000000005e-06 evaluation reward: 7.29
episode: 2016 score: 8.0 memory length: 532630 epsilon: 0.009998020008555413
steps: 412 lr: 1.0240000000000005e-06 evaluation reward: 7.32
episode: 2017 score: 5.0 memory length: 532973 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.31
episode: 2018 score: 4.0 memory length: 533269 epsilon: 0.009998020008555413
steps: 296 lr: 1.0240000000000005e-06 evaluation reward: 7.29
episode: 2019 score: 4.0 memory length: 533565 epsilon: 0.009998020008555413
steps: 296 lr: 1.0240000000000005e-06 evaluation reward: 7.29
episode: 2020 score: 9.0 memory length: 533996 epsilon: 0.009998020008555413
steps: 431 lr: 1.0240000000000005e-06 evaluation reward: 7.34
episode: 2021 score: 13.0 memory length: 534496 epsilon: 0.009998020008555413
steps: 500 lr: 1.0240000000000005e-06 evaluation reward: 7.4
episode: 2022 score: 10.0 memory length: 534966 epsilon: 0.009998020008555413
steps: 470 lr: 1.0240000000000005e-06 evaluation reward: 7.42
episode: 2023 score: 7.0 memory length: 535394 epsilon: 0.009998020008555413
steps: 428 lr: 1.0240000000000005e-06 evaluation reward: 7.42
episode: 2024 score: 11.0 memory length: 535909 epsilon: 0.009998020008555413
steps: 515 lr: 1.0240000000000005e-06 evaluation reward: 7.47
episode: 2025 score: 11.0 memory length: 536429 epsilon: 0.009998020008555413
steps: 520 lr: 1.0240000000000005e-06 evaluation reward: 7.52
episode: 2026 score: 9.0 memory length: 536917 epsilon: 0.009998020008555413
steps: 488 lr: 1.0240000000000005e-06 evaluation reward: 7.57
episode: 2027 score: 7.0 memory length: 537287 epsilon: 0.009998020008555413
steps: 370 lr: 1.0240000000000005e-06 evaluation reward: 7.54
episode: 2028 score: 7.0 memory length: 537709 epsilon: 0.009998020008555413
steps: 422 lr: 1.0240000000000005e-06 evaluation reward: 7.5
episode: 2029 score: 11.0 memory length: 538251 epsilon: 0.009998020008555413
steps: 542 lr: 1.0240000000000005e-06 evaluation reward: 7.51
episode: 2030 score: 14.0 memory length: 538753 epsilon: 0.009998020008555413
steps: 502 lr: 1.0240000000000005e-06 evaluation reward: 7.55
episode: 2031 score: 10.0 memory length: 539262 epsilon: 0.009998020008555413
steps: 509 lr: 1.0240000000000005e-06 evaluation reward: 7.48
episode: 2032 score: 9.0 memory length: 539710 epsilon: 0.009998020008555413
steps: 448 lr: 1.0240000000000005e-06 evaluation reward: 7.51
episode: 2033 score: 7.0 memory length: 540093 epsilon: 0.009998020008555413
steps: 383 lr: 1.0240000000000005e-06 evaluation reward: 7.48
episode: 2034 score: 5.0 memory length: 540414 epsilon: 0.009998020008555413
steps: 321 lr: 1.0240000000000005e-06 evaluation reward: 7.47
episode: 2035 score: 6.0 memory length: 540784 epsilon: 0.009998020008555413
steps: 370 lr: 1.0240000000000005e-06 evaluation reward: 7.47
episode: 2036 score: 5.0 memory length: 541127 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.46
episode: 2037 score: 7.0 memory length: 541509 epsilon: 0.009998020008555413
steps: 382 lr: 1.0240000000000005e-06 evaluation reward: 7.46
episode: 2038 score: 8.0 memory length: 541908 epsilon: 0.009998020008555413
steps: 399 lr: 1.0240000000000005e-06 evaluation reward: 7.48
episode: 2039 score: 10.0 memory length: 542406 epsilon: 0.009998020008555413
steps: 498 lr: 1.0240000000000005e-06 evaluation reward: 7.52
```

```
episode: 2040 score: 9.0 memory length: 542906 epsilon: 0.009998020008555413
steps: 500 lr: 1.0240000000000005e-06 evaluation reward: 7.47
episode: 2041 score: 6.0 memory length: 543231 epsilon: 0.009998020008555413
steps: 325 lr: 1.0240000000000005e-06 evaluation reward: 7.46
episode: 2042 score: 12.0 memory length: 543642 epsilon: 0.009998020008555413
steps: 411 lr: 1.0240000000000005e-06 evaluation reward: 7.51
episode: 2043 score: 7.0 memory length: 543974 epsilon: 0.009998020008555413
steps: 332 lr: 1.0240000000000005e-06 evaluation reward: 7.53
episode: 2044 score: 9.0 memory length: 544384 epsilon: 0.009998020008555413
steps: 410 lr: 1.0240000000000005e-06 evaluation reward: 7.52
episode: 2045 score: 7.0 memory length: 544788 epsilon: 0.009998020008555413
steps: 404 lr: 1.0240000000000005e-06 evaluation reward: 7.51
episode: 2046 score: 9.0 memory length: 545212 epsilon: 0.009998020008555413
steps: 424 lr: 1.0240000000000005e-06 evaluation reward: 7.54
episode: 2047 score: 6.0 memory length: 545568 epsilon: 0.009998020008555413
steps: 356 lr: 1.0240000000000005e-06 evaluation reward: 7.54
episode: 2048 score: 7.0 memory length: 545913 epsilon: 0.009998020008555413
steps: 345 lr: 1.0240000000000005e-06 evaluation reward: 7.5
episode: 2049 score: 5.0 memory length: 546256 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.49
episode: 2050 score: 11.0 memory length: 546752 epsilon: 0.009998020008555413
steps: 496 lr: 1.0240000000000005e-06 evaluation reward: 7.49
episode: 2051 score: 6.0 memory length: 547069 epsilon: 0.009998020008555413
steps: 317 lr: 1.0240000000000005e-06 evaluation reward: 7.44
episode: 2052 score: 17.0 memory length: 547822 epsilon: 0.009998020008555413
steps: 753 lr: 1.0240000000000005e-06 evaluation reward: 7.53
episode: 2053 score: 9.0 memory length: 548272 epsilon: 0.009998020008555413
steps: 450 lr: 1.0240000000000005e-06 evaluation reward: 7.53
episode: 2054 score: 11.0 memory length: 548713 epsilon: 0.009998020008555413
steps: 441 lr: 1.0240000000000005e-06 evaluation reward: 7.57
episode: 2055 score: 5.0 memory length: 549038 epsilon: 0.009998020008555413
steps: 325 lr: 1.0240000000000005e-06 evaluation reward: 7.52
episode: 2056 score: 10.0 memory length: 549426 epsilon: 0.009998020008555413
steps: 388 lr: 1.0240000000000005e-06 evaluation reward: 7.56
episode: 2057 score: 15.0 memory length: 549965 epsilon: 0.009998020008555413
steps: 539 lr: 1.0240000000000005e-06 evaluation reward: 7.67
episode: 2058 score: 7.0 memory length: 550367 epsilon: 0.009998020008555413
steps: 402 lr: 1.0240000000000005e-06 evaluation reward: 7.7
episode: 2059 score: 14.0 memory length: 550881 epsilon: 0.009998020008555413
steps: 514 lr: 1.0240000000000005e-06 evaluation reward: 7.77
episode: 2060 score: 6.0 memory length: 551269 epsilon: 0.009998020008555413
steps: 388 lr: 1.0240000000000005e-06 evaluation reward: 7.75
episode: 2061 score: 8.0 memory length: 551667 epsilon: 0.009998020008555413
steps: 398 lr: 1.0240000000000005e-06 evaluation reward: 7.76
episode: 2062 score: 4.0 memory length: 551943 epsilon: 0.009998020008555413
steps: 276 lr: 1.0240000000000005e-06 evaluation reward: 7.73
episode: 2063 score: 9.0 memory length: 552394 epsilon: 0.009998020008555413
steps: 451 lr: 1.0240000000000005e-06 evaluation reward: 7.78
episode: 2064 score: 4.0 memory length: 552687 epsilon: 0.009998020008555413
steps: 293 lr: 1.0240000000000005e-06 evaluation reward: 7.75
episode: 2065 score: 10.0 memory length: 553192 epsilon: 0.009998020008555413
steps: 505 lr: 1.0240000000000005e-06 evaluation reward: 7.77
episode: 2066 score: 4.0 memory length: 553485 epsilon: 0.009998020008555413
steps: 293 lr: 1.0240000000000005e-06 evaluation reward: 7.74
episode: 2067 score: 4.0 memory length: 553778 epsilon: 0.009998020008555413
steps: 293 lr: 1.0240000000000005e-06 evaluation reward: 7.72
episode: 2068 score: 5.0 memory length: 554121 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.7
episode: 2069 score: 8.0 memory length: 554560 epsilon: 0.009998020008555413
steps: 439 lr: 1.0240000000000005e-06 evaluation reward: 7.71
```

```
episode: 2070 score: 5.0 memory length: 554881 epsilon: 0.009998020008555413
steps: 321 lr: 1.0240000000000005e-06 evaluation reward: 7.64
episode: 2071 score: 7.0 memory length: 555248 epsilon: 0.009998020008555413
steps: 367 lr: 1.0240000000000005e-06 evaluation reward: 7.65
episode: 2072 score: 8.0 memory length: 555638 epsilon: 0.009998020008555413
steps: 390 lr: 1.0240000000000005e-06 evaluation reward: 7.66
episode: 2073 score: 9.0 memory length: 556080 epsilon: 0.009998020008555413
steps: 442 lr: 1.0240000000000005e-06 evaluation reward: 7.66
episode: 2074 score: 6.0 memory length: 556451 epsilon: 0.009998020008555413
steps: 371 lr: 1.0240000000000005e-06 evaluation reward: 7.62
episode: 2075 score: 8.0 memory length: 556834 epsilon: 0.009998020008555413
steps: 383 lr: 1.0240000000000005e-06 evaluation reward: 7.61
episode: 2076 score: 7.0 memory length: 557273 epsilon: 0.009998020008555413
steps: 439 lr: 1.0240000000000005e-06 evaluation reward: 7.59
episode: 2077 score: 4.0 memory length: 557549 epsilon: 0.009998020008555413
steps: 276 lr: 1.0240000000000005e-06 evaluation reward: 7.58
episode: 2078 score: 4.0 memory length: 557842 epsilon: 0.009998020008555413
steps: 293 lr: 1.0240000000000005e-06 evaluation reward: 7.53
episode: 2079 score: 5.0 memory length: 558185 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.51
episode: 2080 score: 5.0 memory length: 558506 epsilon: 0.009998020008555413
steps: 321 lr: 1.0240000000000005e-06 evaluation reward: 7.51
episode: 2081 score: 4.0 memory length: 558799 epsilon: 0.009998020008555413
steps: 293 lr: 1.0240000000000005e-06 evaluation reward: 7.48
episode: 2082 score: 6.0 memory length: 559198 epsilon: 0.009998020008555413
steps: 399 lr: 1.0240000000000005e-06 evaluation reward: 7.49
episode: 2083 score: 5.0 memory length: 559540 epsilon: 0.009998020008555413
steps: 342 lr: 1.0240000000000005e-06 evaluation reward: 7.5
episode: 2084 score: 8.0 memory length: 560000 epsilon: 0.009998020008555413
steps: 460 lr: 1.0240000000000005e-06 evaluation reward: 7.51
episode: 2085 score: 5.0 memory length: 560321 epsilon: 0.009998020008555413
steps: 321 lr: 1.0240000000000005e-06 evaluation reward: 7.47
episode: 2086 score: 6.0 memory length: 560692 epsilon: 0.009998020008555413
steps: 371 lr: 1.0240000000000005e-06 evaluation reward: 7.44
episode: 2087 score: 6.0 memory length: 561086 epsilon: 0.009998020008555413
steps: 394 lr: 1.0240000000000005e-06 evaluation reward: 7.44
episode: 2088 score: 4.0 memory length: 561379 epsilon: 0.009998020008555413
steps: 293 lr: 1.0240000000000005e-06 evaluation reward: 7.43
episode: 2089 score: 7.0 memory length: 561780 epsilon: 0.009998020008555413
steps: 401 lr: 1.0240000000000005e-06 evaluation reward: 7.46
episode: 2090 score: 4.0 memory length: 562073 epsilon: 0.009998020008555413
steps: 293 lr: 1.0240000000000005e-06 evaluation reward: 7.45
episode: 2091 score: 7.0 memory length: 562475 epsilon: 0.009998020008555413
steps: 402 lr: 1.0240000000000005e-06 evaluation reward: 7.47
episode: 2092 score: 8.0 memory length: 562935 epsilon: 0.009998020008555413
steps: 460 lr: 1.0240000000000005e-06 evaluation reward: 7.48
episode: 2093 score: 5.0 memory length: 563256 epsilon: 0.009998020008555413
steps: 321 lr: 1.0240000000000005e-06 evaluation reward: 7.46
episode: 2094 score: 9.0 memory length: 563733 epsilon: 0.009998020008555413
steps: 477 lr: 1.0240000000000005e-06 evaluation reward: 7.49
episode: 2095 score: 5.0 memory length: 564056 epsilon: 0.009998020008555413
steps: 323 lr: 1.0240000000000005e-06 evaluation reward: 7.46
episode: 2096 score: 7.0 memory length: 564411 epsilon: 0.009998020008555413
steps: 355 lr: 1.0240000000000005e-06 evaluation reward: 7.49
episode: 2097 score: 6.0 memory length: 564787 epsilon: 0.009998020008555413
steps: 376 lr: 1.0240000000000005e-06 evaluation reward: 7.5
episode: 2098 score: 10.0 memory length: 565310 epsilon: 0.009998020008555413
steps: 523 lr: 1.0240000000000005e-06 evaluation reward: 7.49
episode: 2099 score: 14.0 memory length: 565836 epsilon: 0.009998020008555413
steps: 526 lr: 1.0240000000000005e-06 evaluation reward: 7.57
```

```
episode: 2100 score: 5.0 memory length: 566179 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.52
episode: 2101 score: 11.0 memory length: 566663 epsilon: 0.009998020008555413
steps: 484 lr: 1.0240000000000005e-06 evaluation reward: 7.58
episode: 2102 score: 4.0 memory length: 566956 epsilon: 0.009998020008555413
steps: 293 lr: 1.0240000000000005e-06 evaluation reward: 7.54
episode: 2103 score: 8.0 memory length: 567413 epsilon: 0.009998020008555413
steps: 457 lr: 1.0240000000000005e-06 evaluation reward: 7.55
episode: 2104 score: 7.0 memory length: 567812 epsilon: 0.009998020008555413
steps: 399 lr: 1.0240000000000005e-06 evaluation reward: 7.54
episode: 2105 score: 10.0 memory length: 568283 epsilon: 0.009998020008555413
steps: 471 lr: 1.0240000000000005e-06 evaluation reward: 7.54
episode: 2106 score: 5.0 memory length: 568626 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.52
episode: 2107 score: 14.0 memory length: 569218 epsilon: 0.009998020008555413
steps: 592 lr: 1.0240000000000005e-06 evaluation reward: 7.59
episode: 2108 score: 6.0 memory length: 569575 epsilon: 0.009998020008555413
steps: 357 lr: 1.0240000000000005e-06 evaluation reward: 7.58
episode: 2109 score: 7.0 memory length: 569974 epsilon: 0.009998020008555413
steps: 399 lr: 1.0240000000000005e-06 evaluation reward: 7.56
episode: 2110 score: 6.0 memory length: 570330 epsilon: 0.009998020008555413
steps: 356 lr: 1.0240000000000005e-06 evaluation reward: 7.53
episode: 2111 score: 7.0 memory length: 570728 epsilon: 0.009998020008555413
steps: 398 lr: 1.0240000000000005e-06 evaluation reward: 7.51
episode: 2112 score: 8.0 memory length: 571127 epsilon: 0.009998020008555413
steps: 399 lr: 1.0240000000000005e-06 evaluation reward: 7.53
episode: 2113 score: 5.0 memory length: 571448 epsilon: 0.009998020008555413
steps: 321 lr: 1.0240000000000005e-06 evaluation reward: 7.53
episode: 2114 score: 4.0 memory length: 571707 epsilon: 0.009998020008555413
steps: 259 lr: 1.0240000000000005e-06 evaluation reward: 7.49
episode: 2115 score: 8.0 memory length: 572128 epsilon: 0.009998020008555413
steps: 421 lr: 1.0240000000000005e-06 evaluation reward: 7.5
episode: 2116 score: 11.0 memory length: 572702 epsilon: 0.009998020008555413
steps: 574 lr: 1.0240000000000005e-06 evaluation reward: 7.53
episode: 2117 score: 10.0 memory length: 573227 epsilon: 0.009998020008555413
steps: 525 lr: 1.0240000000000005e-06 evaluation reward: 7.58
episode: 2118 score: 5.0 memory length: 573510 epsilon: 0.009998020008555413
steps: 283 lr: 1.0240000000000005e-06 evaluation reward: 7.59
episode: 2119 score: 7.0 memory length: 573865 epsilon: 0.009998020008555413
steps: 355 lr: 1.0240000000000005e-06 evaluation reward: 7.62
episode: 2120 score: 10.0 memory length: 574357 epsilon: 0.009998020008555413
steps: 492 lr: 1.0240000000000005e-06 evaluation reward: 7.63
episode: 2121 score: 5.0 memory length: 574665 epsilon: 0.009998020008555413
steps: 308 lr: 1.0240000000000005e-06 evaluation reward: 7.55
episode: 2122 score: 7.0 memory length: 575007 epsilon: 0.009998020008555413
steps: 342 lr: 1.0240000000000005e-06 evaluation reward: 7.52
episode: 2123 score: 9.0 memory length: 575496 epsilon: 0.009998020008555413
steps: 489 lr: 1.0240000000000005e-06 evaluation reward: 7.54
episode: 2124 score: 5.0 memory length: 575839 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.48
episode: 2125 score: 6.0 memory length: 576172 epsilon: 0.009998020008555413
steps: 333 lr: 1.0240000000000005e-06 evaluation reward: 7.43
episode: 2126 score: 5.0 memory length: 576515 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.39
episode: 2127 score: 10.0 memory length: 577001 epsilon: 0.009998020008555413
steps: 486 lr: 1.0240000000000005e-06 evaluation reward: 7.42
episode: 2128 score: 4.0 memory length: 577297 epsilon: 0.009998020008555413
steps: 296 lr: 1.0240000000000005e-06 evaluation reward: 7.39
episode: 2129 score: 8.0 memory length: 577699 epsilon: 0.009998020008555413
steps: 402 lr: 1.0240000000000005e-06 evaluation reward: 7.36
```

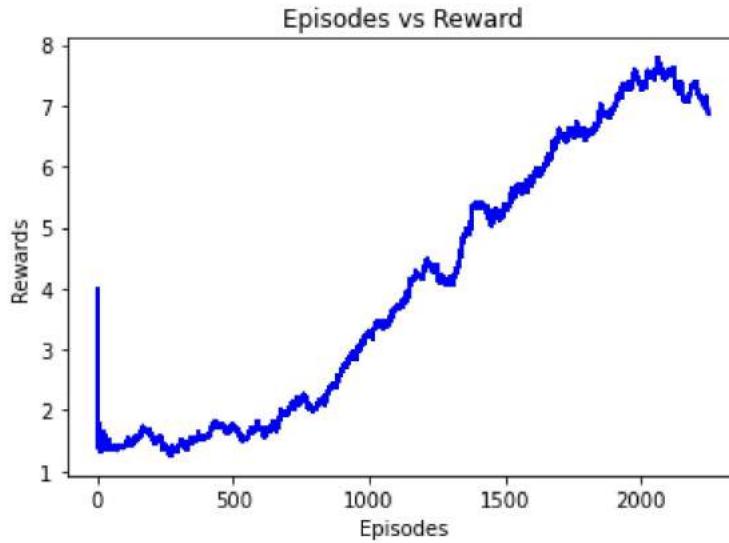
```
episode: 2130 score: 6.0 memory length: 578072 epsilon: 0.009998020008555413
steps: 373 lr: 1.0240000000000005e-06 evaluation reward: 7.28
episode: 2131 score: 10.0 memory length: 578603 epsilon: 0.009998020008555413
steps: 531 lr: 1.0240000000000005e-06 evaluation reward: 7.28
episode: 2132 score: 10.0 memory length: 579065 epsilon: 0.009998020008555413
steps: 462 lr: 1.0240000000000005e-06 evaluation reward: 7.29
episode: 2133 score: 5.0 memory length: 579387 epsilon: 0.009998020008555413
steps: 322 lr: 1.0240000000000005e-06 evaluation reward: 7.27
episode: 2134 score: 13.0 memory length: 579976 epsilon: 0.009998020008555413
steps: 589 lr: 1.0240000000000005e-06 evaluation reward: 7.35
episode: 2135 score: 5.0 memory length: 580259 epsilon: 0.009998020008555413
steps: 283 lr: 1.0240000000000005e-06 evaluation reward: 7.34
episode: 2136 score: 7.0 memory length: 580663 epsilon: 0.009998020008555413
steps: 404 lr: 1.0240000000000005e-06 evaluation reward: 7.36
episode: 2137 score: 6.0 memory length: 581014 epsilon: 0.009998020008555413
steps: 351 lr: 1.0240000000000005e-06 evaluation reward: 7.35
episode: 2138 score: 5.0 memory length: 581337 epsilon: 0.009998020008555413
steps: 323 lr: 1.0240000000000005e-06 evaluation reward: 7.32
episode: 2139 score: 4.0 memory length: 581633 epsilon: 0.009998020008555413
steps: 296 lr: 1.0240000000000005e-06 evaluation reward: 7.26
episode: 2140 score: 4.0 memory length: 581891 epsilon: 0.009998020008555413
steps: 258 lr: 1.0240000000000005e-06 evaluation reward: 7.21
episode: 2141 score: 7.0 memory length: 582248 epsilon: 0.009998020008555413
steps: 357 lr: 1.0240000000000005e-06 evaluation reward: 7.22
episode: 2142 score: 17.0 memory length: 582782 epsilon: 0.009998020008555413
steps: 534 lr: 1.0240000000000005e-06 evaluation reward: 7.27
episode: 2143 score: 5.0 memory length: 583126 epsilon: 0.009998020008555413
steps: 344 lr: 1.0240000000000005e-06 evaluation reward: 7.25
episode: 2144 score: 14.0 memory length: 583631 epsilon: 0.009998020008555413
steps: 505 lr: 1.0240000000000005e-06 evaluation reward: 7.3
episode: 2145 score: 7.0 memory length: 583986 epsilon: 0.009998020008555413
steps: 355 lr: 1.0240000000000005e-06 evaluation reward: 7.3
episode: 2146 score: 7.0 memory length: 584378 epsilon: 0.009998020008555413
steps: 392 lr: 1.0240000000000005e-06 evaluation reward: 7.28
episode: 2147 score: 11.0 memory length: 584881 epsilon: 0.009998020008555413
steps: 503 lr: 1.0240000000000005e-06 evaluation reward: 7.33
episode: 2148 score: 10.0 memory length: 585352 epsilon: 0.009998020008555413
steps: 471 lr: 1.0240000000000005e-06 evaluation reward: 7.36
episode: 2149 score: 7.0 memory length: 585754 epsilon: 0.009998020008555413
steps: 402 lr: 1.0240000000000005e-06 evaluation reward: 7.38
episode: 2150 score: 7.0 memory length: 586156 epsilon: 0.009998020008555413
steps: 402 lr: 1.0240000000000005e-06 evaluation reward: 7.34
episode: 2151 score: 5.0 memory length: 586499 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.33
episode: 2152 score: 8.0 memory length: 586908 epsilon: 0.009998020008555413
steps: 409 lr: 1.0240000000000005e-06 evaluation reward: 7.24
episode: 2153 score: 9.0 memory length: 587408 epsilon: 0.009998020008555413
steps: 500 lr: 1.0240000000000005e-06 evaluation reward: 7.24
episode: 2154 score: 5.0 memory length: 587729 epsilon: 0.009998020008555413
steps: 321 lr: 1.0240000000000005e-06 evaluation reward: 7.18
episode: 2155 score: 7.0 memory length: 588127 epsilon: 0.009998020008555413
steps: 398 lr: 1.0240000000000005e-06 evaluation reward: 7.2
episode: 2156 score: 13.0 memory length: 588523 epsilon: 0.009998020008555413
steps: 396 lr: 1.0240000000000005e-06 evaluation reward: 7.23
episode: 2157 score: 4.0 memory length: 588819 epsilon: 0.009998020008555413
steps: 296 lr: 1.0240000000000005e-06 evaluation reward: 7.12
episode: 2158 score: 7.0 memory length: 589203 epsilon: 0.009998020008555413
steps: 384 lr: 1.0240000000000005e-06 evaluation reward: 7.12
episode: 2159 score: 10.0 memory length: 589664 epsilon: 0.009998020008555413
steps: 461 lr: 1.0240000000000005e-06 evaluation reward: 7.08
```

```
episode: 2160 score: 8.0 memory length: 590077 epsilon: 0.009998020008555413
steps: 413 lr: 1.0240000000000005e-06 evaluation reward: 7.1
episode: 2161 score: 6.0 memory length: 590375 epsilon: 0.009998020008555413
steps: 298 lr: 1.0240000000000005e-06 evaluation reward: 7.08
episode: 2162 score: 5.0 memory length: 590718 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.09
episode: 2163 score: 9.0 memory length: 591143 epsilon: 0.009998020008555413
steps: 425 lr: 1.0240000000000005e-06 evaluation reward: 7.09
episode: 2164 score: 6.0 memory length: 591516 epsilon: 0.009998020008555413
steps: 373 lr: 1.0240000000000005e-06 evaluation reward: 7.11
episode: 2165 score: 5.0 memory length: 591859 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.06
episode: 2166 score: 6.0 memory length: 592216 epsilon: 0.009998020008555413
steps: 357 lr: 1.0240000000000005e-06 evaluation reward: 7.08
episode: 2167 score: 7.0 memory length: 592636 epsilon: 0.009998020008555413
steps: 420 lr: 1.0240000000000005e-06 evaluation reward: 7.11
episode: 2168 score: 7.0 memory length: 593020 epsilon: 0.009998020008555413
steps: 384 lr: 1.0240000000000005e-06 evaluation reward: 7.13
episode: 2169 score: 5.0 memory length: 593363 epsilon: 0.009998020008555413
steps: 343 lr: 1.0240000000000005e-06 evaluation reward: 7.1
episode: 2170 score: 7.0 memory length: 593765 epsilon: 0.009998020008555413
steps: 402 lr: 1.0240000000000005e-06 evaluation reward: 7.12
episode: 2171 score: 9.0 memory length: 594239 epsilon: 0.009998020008555413
steps: 474 lr: 1.0240000000000005e-06 evaluation reward: 7.14
episode: 2172 score: 7.0 memory length: 594659 epsilon: 0.009998020008555413
steps: 420 lr: 1.0240000000000005e-06 evaluation reward: 7.13
episode: 2173 score: 6.0 memory length: 595030 epsilon: 0.009998020008555413
steps: 371 lr: 1.0240000000000005e-06 evaluation reward: 7.1
episode: 2174 score: 7.0 memory length: 595362 epsilon: 0.009998020008555413
steps: 332 lr: 1.0240000000000005e-06 evaluation reward: 7.11
episode: 2175 score: 7.0 memory length: 595693 epsilon: 0.009998020008555413
steps: 331 lr: 1.0240000000000005e-06 evaluation reward: 7.1
episode: 2176 score: 4.0 memory length: 595951 epsilon: 0.009998020008555413
steps: 258 lr: 1.0240000000000005e-06 evaluation reward: 7.07
episode: 2177 score: 8.0 memory length: 596370 epsilon: 0.009998020008555413
steps: 419 lr: 1.0240000000000005e-06 evaluation reward: 7.11
episode: 2178 score: 6.0 memory length: 596687 epsilon: 0.009998020008555413
steps: 317 lr: 1.0240000000000005e-06 evaluation reward: 7.13
episode: 2179 score: 9.0 memory length: 597169 epsilon: 0.009998020008555413
steps: 482 lr: 1.0240000000000005e-06 evaluation reward: 7.17
episode: 2180 score: 6.0 memory length: 597470 epsilon: 0.009998020008555413
steps: 301 lr: 1.0240000000000005e-06 evaluation reward: 7.18
episode: 2181 score: 5.0 memory length: 597815 epsilon: 0.009998020008555413
steps: 345 lr: 1.0240000000000005e-06 evaluation reward: 7.19
episode: 2182 score: 5.0 memory length: 598082 epsilon: 0.009998020008555413
steps: 267 lr: 1.0240000000000005e-06 evaluation reward: 7.18
episode: 2183 score: 8.0 memory length: 598503 epsilon: 0.009998020008555413
steps: 421 lr: 1.0240000000000005e-06 evaluation reward: 7.21
episode: 2184 score: 9.0 memory length: 598968 epsilon: 0.009998020008555413
steps: 465 lr: 1.0240000000000005e-06 evaluation reward: 7.22
episode: 2185 score: 10.0 memory length: 599463 epsilon: 0.009998020008555413
steps: 495 lr: 1.0240000000000005e-06 evaluation reward: 7.27
episode: 2186 score: 8.0 memory length: 599843 epsilon: 0.009998020008555413
steps: 380 lr: 1.0240000000000005e-06 evaluation reward: 7.29
episode: 2187 score: 5.0 memory length: 600164 epsilon: 0.009998020008555413
steps: 321 lr: 1.0240000000000005e-06 evaluation reward: 7.28
episode: 2188 score: 10.0 memory length: 600673 epsilon: 0.009998020008555413
steps: 509 lr: 4.096000000000023e-07 evaluation reward: 7.34
episode: 2189 score: 5.0 memory length: 600994 epsilon: 0.009998020008555413
steps: 321 lr: 4.096000000000023e-07 evaluation reward: 7.32
```

```
episode: 2190 score: 7.0 memory length: 601396 epsilon: 0.009998020008555413
steps: 402 lr: 4.0960000000000023e-07 evaluation reward: 7.35
episode: 2191 score: 5.0 memory length: 601717 epsilon: 0.009998020008555413
steps: 321 lr: 4.0960000000000023e-07 evaluation reward: 7.33
episode: 2192 score: 8.0 memory length: 602151 epsilon: 0.009998020008555413
steps: 434 lr: 4.0960000000000023e-07 evaluation reward: 7.33
episode: 2193 score: 5.0 memory length: 602472 epsilon: 0.009998020008555413
steps: 321 lr: 4.0960000000000023e-07 evaluation reward: 7.33
episode: 2194 score: 10.0 memory length: 602968 epsilon: 0.009998020008555413
steps: 496 lr: 4.0960000000000023e-07 evaluation reward: 7.34
episode: 2195 score: 8.0 memory length: 603387 epsilon: 0.009998020008555413
steps: 419 lr: 4.0960000000000023e-07 evaluation reward: 7.37
episode: 2196 score: 7.0 memory length: 603769 epsilon: 0.009998020008555413
steps: 382 lr: 4.0960000000000023e-07 evaluation reward: 7.37
episode: 2197 score: 7.0 memory length: 604139 epsilon: 0.009998020008555413
steps: 370 lr: 4.0960000000000023e-07 evaluation reward: 7.38
episode: 2198 score: 10.0 memory length: 604614 epsilon: 0.009998020008555413
steps: 475 lr: 4.0960000000000023e-07 evaluation reward: 7.38
episode: 2199 score: 8.0 memory length: 605042 epsilon: 0.009998020008555413
steps: 428 lr: 4.0960000000000023e-07 evaluation reward: 7.32
episode: 2200 score: 10.0 memory length: 605539 epsilon: 0.009998020008555413
steps: 497 lr: 4.0960000000000023e-07 evaluation reward: 7.37
episode: 2201 score: 7.0 memory length: 605884 epsilon: 0.009998020008555413
steps: 345 lr: 4.0960000000000023e-07 evaluation reward: 7.33
episode: 2202 score: 10.0 memory length: 606359 epsilon: 0.009998020008555413
steps: 475 lr: 4.0960000000000023e-07 evaluation reward: 7.39
episode: 2203 score: 7.0 memory length: 606704 epsilon: 0.009998020008555413
steps: 345 lr: 4.0960000000000023e-07 evaluation reward: 7.38
episode: 2204 score: 4.0 memory length: 606997 epsilon: 0.009998020008555413
steps: 293 lr: 4.0960000000000023e-07 evaluation reward: 7.35
episode: 2205 score: 5.0 memory length: 607340 epsilon: 0.009998020008555413
steps: 343 lr: 4.0960000000000023e-07 evaluation reward: 7.3
episode: 2206 score: 5.0 memory length: 607661 epsilon: 0.009998020008555413
steps: 321 lr: 4.0960000000000023e-07 evaluation reward: 7.3
episode: 2207 score: 8.0 memory length: 608080 epsilon: 0.009998020008555413
steps: 419 lr: 4.0960000000000023e-07 evaluation reward: 7.24
episode: 2208 score: 4.0 memory length: 608376 epsilon: 0.009998020008555413
steps: 296 lr: 4.0960000000000023e-07 evaluation reward: 7.22
episode: 2209 score: 4.0 memory length: 608672 epsilon: 0.009998020008555413
steps: 296 lr: 4.0960000000000023e-07 evaluation reward: 7.19
episode: 2210 score: 8.0 memory length: 609087 epsilon: 0.009998020008555413
steps: 415 lr: 4.0960000000000023e-07 evaluation reward: 7.21
episode: 2211 score: 8.0 memory length: 609496 epsilon: 0.009998020008555413
steps: 409 lr: 4.0960000000000023e-07 evaluation reward: 7.22
episode: 2212 score: 5.0 memory length: 609818 epsilon: 0.009998020008555413
steps: 322 lr: 4.0960000000000023e-07 evaluation reward: 7.19
episode: 2213 score: 5.0 memory length: 610145 epsilon: 0.009998020008555413
steps: 327 lr: 4.0960000000000023e-07 evaluation reward: 7.19
episode: 2214 score: 9.0 memory length: 610583 epsilon: 0.009998020008555413
steps: 438 lr: 4.0960000000000023e-07 evaluation reward: 7.24
episode: 2215 score: 6.0 memory length: 610881 epsilon: 0.009998020008555413
steps: 298 lr: 4.0960000000000023e-07 evaluation reward: 7.22
episode: 2216 score: 5.0 memory length: 611224 epsilon: 0.009998020008555413
steps: 343 lr: 4.0960000000000023e-07 evaluation reward: 7.16
episode: 2217 score: 8.0 memory length: 611680 epsilon: 0.009998020008555413
steps: 456 lr: 4.0960000000000023e-07 evaluation reward: 7.14
episode: 2218 score: 6.0 memory length: 612053 epsilon: 0.009998020008555413
steps: 373 lr: 4.0960000000000023e-07 evaluation reward: 7.15
episode: 2219 score: 8.0 memory length: 612486 epsilon: 0.009998020008555413
steps: 433 lr: 4.0960000000000023e-07 evaluation reward: 7.16
```

```
episode: 2220 score: 6.0 memory length: 612803 epsilon: 0.009998020008555413
steps: 317 lr: 4.0960000000000023e-07 evaluation reward: 7.12
episode: 2221 score: 5.0 memory length: 613124 epsilon: 0.009998020008555413
steps: 321 lr: 4.0960000000000023e-07 evaluation reward: 7.12
episode: 2222 score: 5.0 memory length: 613411 epsilon: 0.009998020008555413
steps: 287 lr: 4.0960000000000023e-07 evaluation reward: 7.1
episode: 2223 score: 8.0 memory length: 613812 epsilon: 0.009998020008555413
steps: 401 lr: 4.0960000000000023e-07 evaluation reward: 7.09
episode: 2224 score: 7.0 memory length: 614190 epsilon: 0.009998020008555413
steps: 378 lr: 4.0960000000000023e-07 evaluation reward: 7.11
episode: 2225 score: 6.0 memory length: 614563 epsilon: 0.009998020008555413
steps: 373 lr: 4.0960000000000023e-07 evaluation reward: 7.11
episode: 2226 score: 4.0 memory length: 614856 epsilon: 0.009998020008555413
steps: 293 lr: 4.0960000000000023e-07 evaluation reward: 7.1
episode: 2227 score: 6.0 memory length: 615244 epsilon: 0.009998020008555413
steps: 388 lr: 4.0960000000000023e-07 evaluation reward: 7.06
episode: 2228 score: 5.0 memory length: 615587 epsilon: 0.009998020008555413
steps: 343 lr: 4.0960000000000023e-07 evaluation reward: 7.07
episode: 2229 score: 14.0 memory length: 616093 epsilon: 0.009998020008555413
steps: 506 lr: 4.0960000000000023e-07 evaluation reward: 7.13
episode: 2230 score: 5.0 memory length: 616436 epsilon: 0.009998020008555413
steps: 343 lr: 4.0960000000000023e-07 evaluation reward: 7.12
episode: 2231 score: 7.0 memory length: 616856 epsilon: 0.009998020008555413
steps: 420 lr: 4.0960000000000023e-07 evaluation reward: 7.09
episode: 2232 score: 4.0 memory length: 617152 epsilon: 0.009998020008555413
steps: 296 lr: 4.0960000000000023e-07 evaluation reward: 7.03
episode: 2233 score: 9.0 memory length: 617590 epsilon: 0.009998020008555413
steps: 438 lr: 4.0960000000000023e-07 evaluation reward: 7.07
episode: 2234 score: 4.0 memory length: 617883 epsilon: 0.009998020008555413
steps: 293 lr: 4.0960000000000023e-07 evaluation reward: 6.98
episode: 2235 score: 7.0 memory length: 618306 epsilon: 0.009998020008555413
steps: 423 lr: 4.0960000000000023e-07 evaluation reward: 7.0
episode: 2236 score: 5.0 memory length: 618649 epsilon: 0.009998020008555413
steps: 343 lr: 4.0960000000000023e-07 evaluation reward: 6.98
episode: 2237 score: 7.0 memory length: 619053 epsilon: 0.009998020008555413
steps: 404 lr: 4.0960000000000023e-07 evaluation reward: 6.99
episode: 2238 score: 12.0 memory length: 619510 epsilon: 0.009998020008555413
steps: 457 lr: 4.0960000000000023e-07 evaluation reward: 7.06
episode: 2239 score: 8.0 memory length: 619964 epsilon: 0.009998020008555413
steps: 454 lr: 4.0960000000000023e-07 evaluation reward: 7.1
episode: 2240 score: 5.0 memory length: 620307 epsilon: 0.009998020008555413
steps: 343 lr: 4.0960000000000023e-07 evaluation reward: 7.11
episode: 2241 score: 11.0 memory length: 620833 epsilon: 0.009998020008555413
steps: 526 lr: 4.0960000000000023e-07 evaluation reward: 7.15
episode: 2242 score: 8.0 memory length: 621287 epsilon: 0.009998020008555413
steps: 454 lr: 4.0960000000000023e-07 evaluation reward: 7.06
episode: 2243 score: 7.0 memory length: 621727 epsilon: 0.009998020008555413
steps: 440 lr: 4.0960000000000023e-07 evaluation reward: 7.08
episode: 2244 score: 7.0 memory length: 622150 epsilon: 0.009998020008555413
steps: 423 lr: 4.0960000000000023e-07 evaluation reward: 7.01
episode: 2245 score: 5.0 memory length: 622471 epsilon: 0.009998020008555413
steps: 321 lr: 4.0960000000000023e-07 evaluation reward: 6.99
episode: 2246 score: 5.0 memory length: 622792 epsilon: 0.009998020008555413
steps: 321 lr: 4.0960000000000023e-07 evaluation reward: 6.97
episode: 2247 score: 5.0 memory length: 623136 epsilon: 0.009998020008555413
steps: 344 lr: 4.0960000000000023e-07 evaluation reward: 6.91
episode: 2248 score: 5.0 memory length: 623419 epsilon: 0.009998020008555413
steps: 283 lr: 4.0960000000000023e-07 evaluation reward: 6.86
episode: 2249 score: 10.0 memory length: 623884 epsilon: 0.009998020008555413
steps: 465 lr: 4.0960000000000023e-07 evaluation reward: 6.89
```

```
episode: 2250    score: 5.0    memory length: 624205    epsilon: 0.009998020008555413
steps: 321      lr: 4.0960000000000023e-07    evaluation reward: 6.87
episode: 2251    score: 5.0    memory length: 624526    epsilon: 0.009998020008555413
steps: 321      lr: 4.0960000000000023e-07    evaluation reward: 6.87
episode: 2252    score: 6.0    memory length: 624843    epsilon: 0.009998020008555413
steps: 317      lr: 4.0960000000000023e-07    evaluation reward: 6.85
```



## 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 you reach the last tuple, where you will make Q value predictions. Training loop remains nearly the same.

```
In [ ]: from agent import LSTM_Agent
agent = LSTM_Agent(action_size)
evaluation_reward = deque(maxlen=evaluation_reward_length)
frame = 0
memory_size = 0

HISTORY_SIZE = 1
rewards, episodes = [], []
best_eval_reward = 0
for e in range(EPISODES):
    done = False
    score = 0

    history = np.zeros([HISTORY_SIZE + 1, 84, 84], dtype=np.uint8)
    step = 0
    state = env.reset()
    next_state = state
    life = number_lives
    hidden = None
```