MP5

April 28, 2024

1 Deep Q-Learning

Install dependencies for AI gym to run properly (shouldn't take more than a minute). If running on google cloud or running locally, only need to run once. Colab may require installing everytime the vm shuts down.

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

```
Defaulting to user installation because normal site-packages is not writeable
Collecting gym
  Downloading gym-0.26.2.tar.gz (721 kB)
     |                  | 721 kB 5.0 MB/s eta 0:00:01
  Installing build dependencies … done
  Getting requirements to build wheel … done
    Preparing wheel metadata … done
Collecting pyvirtualdisplay
  Downloading PyVirtualDisplay-3.0-py3-none-any.whl (15 kB)
Collecting gym-notices>=0.0.4
  Downloading gym_notices-0.0.8-py3-none-any.whl (3.0 kB)
Requirement already satisfied: cloudpickle>=1.2.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from gym) (2.0.0)
Requirement already satisfied: numpy>=1.18.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from gym)
(1.21.5)
Requirement already satisfied: importlib-metadata>=4.8.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from gym)
(4.11.3)
Requirement already satisfied: zipp>=0.5 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from importlib-
metadata>=4.8.0->gym) (3.7.0)
Building wheels for collected packages: gym
  Building wheel for gym (PEP 517) … done
  Created wheel for gym: filename=gym-0.26.2-py3-none-any.whl size=827623
sha256=8bf773399c1dc258a9bae1d8b7aa153926f9b578220f718a639c20da454f6b04
  Stored in directory: /u/qilong/.cache/pip/wheels/af/2b/30/5e78b8b9599f2a2286a5
82b8da80594f654bf0e18d825a4405
Successfully built gym
```

```
Installing collected packages: gym-notices, pyvirtualdisplay, gym
Successfully installed gym-0.26.2 gym-notices-0.0.8 pyvirtualdisplay-3.0
/bin/bash: /usr/bin/sudo: Permission denied
```

[ ]:
```
!pip install ffmpeg
```

```
Defaulting to user installation because normal site-packages is not writeable
Collecting ffmpeg
  Downloading ffmpeg-1.4.tar.gz (5.1 kB)
Building wheels for collected packages: ffmpeg
  Building wheel for ffmpeg (setup.py) … done
  Created wheel for ffmpeg: filename=ffmpeg-1.4-py3-none-any.whl size=6083
sha256=0d50f6bd0aec9017f112e00d4604a210aeeb99e73d8e8ebcbf7aa2280e93f068
  Stored in directory: /u/qilong/.cache/pip/wheels/1d/57/24/4eff6a03a9ea0e647568
e8a5a0546cdf957e3cf005372c0245
Successfully built ffmpeg
Installing collected packages: ffmpeg
Successfully installed ffmpeg-1.4
```

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

```
Requirement already satisfied: setuptools in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (61.2.0)
Collecting setuptools
  Using cached setuptools-69.5.1-py3-none-any.whl (894 kB)
Installing collected packages: setuptools
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.
spyder 5.1.5 requires pyqt5<5.13, which is not installed.
spyder 5.1.5 requires pyqtwebengine<5.13, which is not installed.
conda-repo-cli 1.0.4 requires pathlib, which is not installed.
Successfully installed setuptools-69.5.1
Defaulting to user installation because normal site-packages is not writeable
Collecting ez_setup
  Downloading ez_setup-0.9.tar.gz (6.6 kB)
Building wheels for collected packages: ez-setup
  Building wheel for ez-setup (setup.py) … done
  Created wheel for ez-setup: filename=ez_setup-0.9-py3-none-any.whl
size=10994
sha256=0e22daeb48edafa38bdb9af1185f4718d74a3c446fdd70fdc2c1740079b98031
  Stored in directory: /u/qilong/.cache/pip/wheels/d7/39/49/e7ce9ce92f074adc6f75
```

```
5a0cc05992407730e14d94ce3f9554
Successfully built ez-setup
Installing collected packages: ez-setup
Successfully installed ez-setup-0.9
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: gym[atari] in
/u/qilong/.local/lib/python3.9/site-packages (0.26.2)
Requirement already satisfied: importlib-metadata>=4.8.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from gym[atari])
(4.11.3)
Requirement already satisfied: gym-notices>=0.0.4 in
/u/qilong/.local/lib/python3.9/site-packages (from gym[atari]) (0.0.8)
Requirement already satisfied: numpy>=1.18.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from gym[atari])
(1.21.5)
Requirement already satisfied: cloudpickle>=1.2.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from gym[atari])
(2.0.0)
Collecting ale-py~=0.8.0
  Downloading
ale_py-0.8.1-cp39-cp39-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.7 MB)
     |                  | 1.7 MB 5.2 MB/s eta 0:00:01
Collecting importlib-resources
  Downloading importlib_resources-6.4.0-py3-none-any.whl (38 kB)
Requirement already satisfied: typing-extensions in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from ale-
py~=0.8.0->gym[atari]) (4.1.1)
Requirement already satisfied: zipp>=0.5 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from importlib-
metadata>=4.8.0->gym[atari]) (3.7.0)
Installing collected packages: importlib-resources, ale-py
Successfully installed ale-py-0.8.1 importlib-resources-6.4.0
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: gym[accept-rom-license] in
/u/qilong/.local/lib/python3.9/site-packages (0.26.2)
Requirement already satisfied: numpy>=1.18.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from gym[accept-
rom-license]) (1.21.5)
Requirement already satisfied: cloudpickle>=1.2.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from gym[accept-
rom-license]) (2.0.0)
Requirement already satisfied: gym-notices>=0.0.4 in
/u/qilong/.local/lib/python3.9/site-packages (from gym[accept-rom-license])
(0.0.8)
Requirement already satisfied: importlib-metadata>=4.8.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from gym[accept-
rom-license]) (4.11.3)
Collecting autorom[accept-rom-license]~=0.4.2
```

```
  Downloading AutoROM-0.4.2-py3-none-any.whl (16 kB)
Requirement already satisfied: click in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license]) (8.0.4)
Requirement already satisfied: requests in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license]) (2.27.1)
Requirement already satisfied: tqdm in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license]) (4.64.0)
Collecting AutoROM.accept-rom-license
  Downloading AutoROM.accept-rom-license-0.6.1.tar.gz (434 kB)
     |                  | 434 kB 5.2 MB/s eta 0:00:01
  Installing build dependencies … done
  Getting requirements to build wheel … done
    Preparing wheel metadata … done
Requirement already satisfied: zipp>=0.5 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from importlib-
metadata>=4.8.0->gym[accept-rom-license]) (3.7.0)
Requirement already satisfied: certifi>=2017.4.17 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
requests->autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license])
(2021.10.8)
Requirement already satisfied: idna<4,>=2.5 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
requests->autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license]) (3.3)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
requests->autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license]) (1.26.9)
Requirement already satisfied: charset-normalizer~=2.0.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
requests->autorom[accept-rom-license]~=0.4.2->gym[accept-rom-license]) (2.0.4)
Building wheels for collected packages: AutoROM.accept-rom-license
  Building wheel for AutoROM.accept-rom-license (PEP 517) … done
  Created wheel for AutoROM.accept-rom-license:
filename=AutoROM.accept_rom_license-0.6.1-py3-none-any.whl size=446659
sha256=c01843e0a36457b7486288738232f06ca733e905a11aaedfef2ba99764295b3c
  Stored in directory: /u/qilong/.cache/pip/wheels/b1/1f/f7/2da07cf4f81ea264bdaf
043028749d88fe0c2227134a22cf80
Successfully built AutoROM.accept-rom-license
Installing collected packages: AutoROM.accept-rom-license, autorom
Successfully installed AutoROM.accept-rom-license-0.6.1 autorom-0.4.2
```

[ ]:
```
!pip install scikit-image
```

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: scikit-image in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (0.19.2)
```

```
Requirement already satisfied: numpy>=1.17.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from scikit-
image) (1.21.5)
Requirement already satisfied: scipy>=1.4.1 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from scikit-
image) (1.7.3)
Requirement already satisfied: networkx>=2.2 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from scikit-
image) (2.7.1)
Requirement already satisfied: PyWavelets>=1.1.1 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from scikit-
image) (1.3.0)
Requirement already satisfied: tifffile>=2019.7.26 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from scikit-
image) (2021.7.2)
Requirement already satisfied: pillow!=7.1.0,!=7.1.1,!=8.3.0,>=6.1.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from scikit-
image) (9.0.1)
Requirement already satisfied: imageio>=2.4.1 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from scikit-
image) (2.9.0)
Requirement already satisfied: packaging>=20.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from scikit-
image) (21.3)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
packaging>=20.0->scikit-image) (3.0.4)
```

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

[ ]:
```
%matplotlib inline

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

```
from model import DQN, DQN_LSTM
from config import *

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

```
The autoreload extension is already loaded. To reload it, use:
  %reload_ext autoreload
```

1.1 Understanding the environment

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

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

[ ]:
```
env = gym.make('BreakoutDeterministic-v4')
state = env.reset()
print(state[0].shape)
```

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

```
(210, 160, 3)
```

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

```
5
(210, 160, 3)
```

1.2 Creating a DQN Agent

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

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

Frame : Number of frames processed in total.

Memory Size : The current size of the replay memory.

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

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

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

1.2.1 Main Training Loop

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

[ ]:
```
rewards, episodes = [], []
best_eval_reward = 0
halt_flag = 0

for e in range(EPISODES):
    done = False
    score = 0

    history = np.zeros([5, 84, 84], dtype=np.uint8)
    step = 0
    state, _ = env.reset()
    next_state = state
    life = number_lives

    get_init_state(history, state, HISTORY_SIZE)

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

        # Perform a fire action if ball is no longer on screen to continue onto␣
 ↪next life
        if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
            action = torch.tensor([[0]]).cuda()
        else:
            action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
#             print(action)
        state = next_state
#         next_state, reward, done, info = env.step(action + 1)
        next_state, reward, terminated, truncated, info = env.step(action + 1)
        done = terminated or truncated
```

```
        frame_next_state = get_frame(next_state)
        history[4, :, :] = frame_next_state
        terminal_state = check_live(life, info['lives'])

        life = info['lives']
        r = reward

        # Store the transition in memory
        agent.memory.push(deepcopy(frame_next_state), action.cpu(), r,␣
 ↪terminal_state)
#         agent.memory.push(deepcopy(frame_next_state), action, r,␣
 ↪terminal_state)
        # Start training after random sample generation
        if(frame >= train_frame): # You can set train_frame to a lower value␣
 ↪while testing your starts training earlier
#         if frame >= 200:
            agent.train_policy_net(frame)
            # Update the target network only for Double DQN only
            if double_dqn and (frame % update_target_network_frequency)== 0:
                agent.update_target_net()
        score += reward
        history[:4, :, :] = history[1:, :, :]

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

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

            # if the mean of scores of last 100 episode is bigger than 5 save␣
 ↪model
            ### Change this save condition to whatever you prefer ###
            if np.mean(evaluation_reward) > 8.0 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)
                if np.mean(evaluation_reward) > 8.2:
                    halt_flag = 1
                    break
    if halt_flag == 1:
        break
```

```
episode: 0   score: 2.0   memory length: 200   epsilon: 1.0    steps: 200    lr:
0.0001     evaluation reward: 2.0
episode: 1   score: 2.0   memory length: 398   epsilon: 1.0    steps: 198    lr:
0.0001     evaluation reward: 2.0
episode: 2   score: 0.0   memory length: 521   epsilon: 1.0    steps: 123    lr:
0.0001     evaluation reward: 1.3333333333333333
episode: 3   score: 0.0   memory length: 644   epsilon: 1.0    steps: 123    lr:
0.0001     evaluation reward: 1.0
episode: 4   score: 0.0   memory length: 766   epsilon: 1.0    steps: 122    lr:
0.0001     evaluation reward: 0.8
episode: 5   score: 0.0   memory length: 889   epsilon: 1.0    steps: 123    lr:
0.0001     evaluation reward: 0.6666666666666666
episode: 6   score: 1.0   memory length: 1042   epsilon: 1.0    steps: 153
lr: 0.0001     evaluation reward: 0.7142857142857143
episode: 7   score: 0.0   memory length: 1165   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 0.625
episode: 8   score: 3.0   memory length: 1412   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 0.8888888888888888
episode: 9   score: 0.0   memory length: 1535   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 0.8
episode: 10   score: 0.0   memory length: 1658   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 0.7272727272727273
episode: 11   score: 3.0   memory length: 1905   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 0.9166666666666666
episode: 12   score: 0.0   memory length: 2028   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 0.8461538461538461
episode: 13   score: 1.0   memory length: 2197   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 0.8571428571428571
episode: 14   score: 0.0   memory length: 2320   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 0.8
episode: 15   score: 3.0   memory length: 2547   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 0.9375
episode: 16   score: 2.0   memory length: 2747   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.0
episode: 17   score: 0.0   memory length: 2870   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 0.9444444444444444
episode: 18   score: 2.0   memory length: 3089   epsilon: 1.0    steps: 219
lr: 0.0001     evaluation reward: 1.0
episode: 19   score: 2.0   memory length: 3305   epsilon: 1.0    steps: 216
```

```
lr: 0.0001     evaluation reward: 1.05
episode: 20   score: 1.0   memory length: 3474   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.0476190476190477
episode: 21   score: 3.0   memory length: 3724   epsilon: 1.0    steps: 250
lr: 0.0001     evaluation reward: 1.1363636363636365
episode: 22   score: 1.0   memory length: 3875   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.1304347826086956
episode: 23   score: 1.0   memory length: 4046   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.125
episode: 24   score: 5.0   memory length: 4390   epsilon: 1.0    steps: 344
lr: 0.0001     evaluation reward: 1.28
episode: 25   score: 3.0   memory length: 4638   epsilon: 1.0    steps: 248
lr: 0.0001     evaluation reward: 1.3461538461538463
episode: 26   score: 0.0   memory length: 4761   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.2962962962962963
episode: 27   score: 2.0   memory length: 4978   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.3214285714285714
episode: 28   score: 0.0   memory length: 5100   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.2758620689655173
episode: 29   score: 3.0   memory length: 5326   epsilon: 1.0    steps: 226
lr: 0.0001     evaluation reward: 1.3333333333333333
episode: 30   score: 1.0   memory length: 5476   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.3225806451612903
episode: 31   score: 1.0   memory length: 5626   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.3125
episode: 32   score: 0.0   memory length: 5749   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.2727272727272727
episode: 33   score: 0.0   memory length: 5872   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.2352941176470589
episode: 34   score: 1.0   memory length: 6041   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.2285714285714286
episode: 35   score: 1.0   memory length: 6210   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.2222222222222223
episode: 36   score: 3.0   memory length: 6456   epsilon: 1.0    steps: 246
lr: 0.0001     evaluation reward: 1.2702702702702702
episode: 37   score: 1.0   memory length: 6607   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.263157894736842
episode: 38   score: 5.0   memory length: 6933   epsilon: 1.0    steps: 326
lr: 0.0001     evaluation reward: 1.358974358974359
episode: 39   score: 0.0   memory length: 7056   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.325
episode: 40   score: 0.0   memory length: 7179   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.2926829268292683
episode: 41   score: 1.0   memory length: 7330   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.2857142857142858
episode: 42   score: 0.0   memory length: 7453   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.255813953488372
episode: 43   score: 5.0   memory length: 7761   epsilon: 1.0    steps: 308
```

```
lr: 0.0001     evaluation reward: 1.3409090909090908
episode: 44   score: 3.0   memory length: 7992   epsilon: 1.0    steps: 231
lr: 0.0001     evaluation reward: 1.3777777777777778
episode: 45   score: 3.0   memory length: 8242   epsilon: 1.0    steps: 250
lr: 0.0001     evaluation reward: 1.4130434782608696
episode: 46   score: 0.0   memory length: 8365   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3829787234042554
episode: 47   score: 1.0   memory length: 8515   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.375
episode: 48   score: 1.0   memory length: 8684   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.3673469387755102
episode: 49   score: 3.0   memory length: 8928   epsilon: 1.0    steps: 244
lr: 0.0001     evaluation reward: 1.4
episode: 50   score: 0.0   memory length: 9050   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.3725490196078431
episode: 51   score: 1.0   memory length: 9200   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.3653846153846154
episode: 52   score: 0.0   memory length: 9323   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3396226415094339
episode: 53   score: 1.0   memory length: 9492   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.3333333333333333
episode: 54   score: 1.0   memory length: 9662   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.3272727272727274
episode: 55   score: 2.0   memory length: 9859   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.3392857142857142
episode: 56   score: 1.0   memory length: 10027   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.3333333333333333
episode: 57   score: 3.0   memory length: 10292   epsilon: 1.0    steps: 265
lr: 0.0001     evaluation reward: 1.3620689655172413
episode: 58   score: 3.0   memory length: 10520   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.3898305084745763
episode: 59   score: 2.0   memory length: 10736   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.4
episode: 60   score: 2.0   memory length: 10956   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.4098360655737705
episode: 61   score: 2.0   memory length: 11154   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.4193548387096775
episode: 62   score: 1.0   memory length: 11305   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.4126984126984128
episode: 63   score: 2.0   memory length: 11502   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.421875
episode: 64   score: 0.0   memory length: 11624   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.4
episode: 65   score: 2.0   memory length: 11841   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.4090909090909092
episode: 66   score: 0.0   memory length: 11964   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3880597014925373
episode: 67   score: 2.0   memory length: 12180   epsilon: 1.0    steps: 216
```

```
lr: 0.0001     evaluation reward: 1.3970588235294117
episode: 68   score: 2.0   memory length: 12397   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.4057971014492754
episode: 69   score: 0.0   memory length: 12519   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.3857142857142857
episode: 70   score: 2.0   memory length: 12717   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.3943661971830985
episode: 71   score: 4.0   memory length: 13015   epsilon: 1.0    steps: 298
lr: 0.0001     evaluation reward: 1.4305555555555556
episode: 72   score: 2.0   memory length: 13216   epsilon: 1.0    steps: 201
lr: 0.0001     evaluation reward: 1.4383561643835616
episode: 73   score: 1.0   memory length: 13367   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.4324324324324325
episode: 74   score: 2.0   memory length: 13569   epsilon: 1.0    steps: 202
lr: 0.0001     evaluation reward: 1.44
episode: 75   score: 3.0   memory length: 13794   epsilon: 1.0    steps: 225
lr: 0.0001     evaluation reward: 1.4605263157894737
episode: 76   score: 2.0   memory length: 14012   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.4675324675324675
episode: 77   score: 1.0   memory length: 14180   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.4615384615384615
episode: 78   score: 2.0   memory length: 14398   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.4683544303797469
episode: 79   score: 0.0   memory length: 14521   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.45
episode: 80   score: 1.0   memory length: 14691   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.4444444444444444
episode: 81   score: 2.0   memory length: 14909   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.451219512195122
episode: 82   score: 4.0   memory length: 15228   epsilon: 1.0    steps: 319
lr: 0.0001     evaluation reward: 1.4819277108433735
episode: 83   score: 0.0   memory length: 15351   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4642857142857142
episode: 84   score: 1.0   memory length: 15502   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.4588235294117646
episode: 85   score: 4.0   memory length: 15797   epsilon: 1.0    steps: 295
lr: 0.0001     evaluation reward: 1.4883720930232558
episode: 86   score: 1.0   memory length: 15948   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.4827586206896552
episode: 87   score: 5.0   memory length: 16255   epsilon: 1.0    steps: 307
lr: 0.0001     evaluation reward: 1.5227272727272727
episode: 88   score: 2.0   memory length: 16471   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.5280898876404494
episode: 89   score: 0.0   memory length: 16594   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.511111111111111
episode: 90   score: 3.0   memory length: 16859   epsilon: 1.0    steps: 265
lr: 0.0001     evaluation reward: 1.5274725274725274
episode: 91   score: 3.0   memory length: 17125   epsilon: 1.0    steps: 266
```

```
lr: 0.0001     evaluation reward: 1.5434782608695652
episode: 92   score: 0.0   memory length: 17248   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.5268817204301075
episode: 93   score: 1.0   memory length: 17417   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.5212765957446808
episode: 94   score: 1.0   memory length: 17585   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.5157894736842106
episode: 95   score: 0.0   memory length: 17708   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.5
episode: 96   score: 1.0   memory length: 17859   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.4948453608247423
episode: 97   score: 0.0   memory length: 17982   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4795918367346939
episode: 98   score: 3.0   memory length: 18210   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.494949494949495
episode: 99   score: 0.0   memory length: 18333   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.48
episode: 100   score: 2.0   memory length: 18514   epsilon: 1.0    steps: 181
lr: 0.0001     evaluation reward: 1.48
episode: 101   score: 2.0   memory length: 18730   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.48
episode: 102   score: 0.0   memory length: 18853   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.48
episode: 103   score: 0.0   memory length: 18975   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.48
episode: 104   score: 0.0   memory length: 19098   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.48
episode: 105   score: 1.0   memory length: 19267   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.49
episode: 106   score: 1.0   memory length: 19435   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.49
episode: 107   score: 1.0   memory length: 19586   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.5
episode: 108   score: 0.0   memory length: 19709   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.47
episode: 109   score: 2.0   memory length: 19907   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.49
episode: 110   score: 3.0   memory length: 20135   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.52
episode: 111   score: 2.0   memory length: 20332   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.51
episode: 112   score: 1.0   memory length: 20483   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.52
episode: 113   score: 2.0   memory length: 20662   epsilon: 1.0    steps: 179
lr: 0.0001     evaluation reward: 1.53
episode: 114   score: 1.0   memory length: 20830   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.54
episode: 115   score: 1.0   memory length: 21002   epsilon: 1.0    steps: 172
```

```
lr: 0.0001     evaluation reward: 1.52
episode: 116   score: 0.0   memory length: 21124   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.5
episode: 117   score: 3.0   memory length: 21372   epsilon: 1.0    steps: 248
lr: 0.0001     evaluation reward: 1.53
episode: 118   score: 2.0   memory length: 21570   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.53
episode: 119   score: 0.0   memory length: 21693   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.51
episode: 120   score: 2.0   memory length: 21891   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.52
episode: 121   score: 1.0   memory length: 22041   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.5
episode: 122   score: 2.0   memory length: 22258   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.51
episode: 123   score: 0.0   memory length: 22380   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.5
episode: 124   score: 1.0   memory length: 22549   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.46
episode: 125   score: 1.0   memory length: 22699   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.44
episode: 126   score: 2.0   memory length: 22897   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.46
episode: 127   score: 1.0   memory length: 23048   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.45
episode: 128   score: 0.0   memory length: 23171   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.45
episode: 129   score: 0.0   memory length: 23294   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.42
episode: 130   score: 2.0   memory length: 23514   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.43
episode: 131   score: 0.0   memory length: 23637   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.42
episode: 132   score: 1.0   memory length: 23788   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.43
episode: 133   score: 1.0   memory length: 23960   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.44
episode: 134   score: 3.0   memory length: 24188   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.46
episode: 135   score: 1.0   memory length: 24340   epsilon: 1.0    steps: 152
lr: 0.0001     evaluation reward: 1.46
episode: 136   score: 2.0   memory length: 24558   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.45
episode: 137   score: 1.0   memory length: 24727   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.45
episode: 138   score: 2.0   memory length: 24945   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.42
episode: 139   score: 0.0   memory length: 25068   epsilon: 1.0    steps: 123
```

```
lr: 0.0001     evaluation reward: 1.42
episode: 140   score: 1.0   memory length: 25237   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.43
episode: 141   score: 1.0   memory length: 25407   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.43
episode: 142   score: 1.0   memory length: 25578   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.44
episode: 143   score: 1.0   memory length: 25747   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.4
episode: 144   score: 1.0   memory length: 25897   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.38
episode: 145   score: 1.0   memory length: 26048   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.36
episode: 146   score: 2.0   memory length: 26228   epsilon: 1.0    steps: 180
lr: 0.0001     evaluation reward: 1.38
episode: 147   score: 1.0   memory length: 26379   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.38
episode: 148   score: 3.0   memory length: 26628   epsilon: 1.0    steps: 249
lr: 0.0001     evaluation reward: 1.4
episode: 149   score: 2.0   memory length: 26844   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.39
episode: 150   score: 0.0   memory length: 26966   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.39
episode: 151   score: 0.0   memory length: 27088   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.38
episode: 152   score: 3.0   memory length: 27316   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.41
episode: 153   score: 0.0   memory length: 27439   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4
episode: 154   score: 0.0   memory length: 27562   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.39
episode: 155   score: 1.0   memory length: 27713   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.38
episode: 156   score: 1.0   memory length: 27881   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.38
episode: 157   score: 1.0   memory length: 28050   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.36
episode: 158   score: 1.0   memory length: 28219   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.34
episode: 159   score: 0.0   memory length: 28341   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.32
episode: 160   score: 1.0   memory length: 28509   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.31
episode: 161   score: 1.0   memory length: 28679   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.3
episode: 162   score: 1.0   memory length: 28829   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.3
episode: 163   score: 4.0   memory length: 29105   epsilon: 1.0    steps: 276
```

```
lr: 0.0001     evaluation reward: 1.32
episode: 164   score: 2.0   memory length: 29325   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.34
episode: 165   score: 0.0   memory length: 29448   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.32
episode: 166   score: 0.0   memory length: 29571   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.32
episode: 167   score: 3.0   memory length: 29817   epsilon: 1.0    steps: 246
lr: 0.0001     evaluation reward: 1.33
episode: 168   score: 3.0   memory length: 30066   epsilon: 1.0    steps: 249
lr: 0.0001     evaluation reward: 1.34
episode: 169   score: 2.0   memory length: 30245   epsilon: 1.0    steps: 179
lr: 0.0001     evaluation reward: 1.36
episode: 170   score: 0.0   memory length: 30368   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.34
episode: 171   score: 1.0   memory length: 30538   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.31
episode: 172   score: 2.0   memory length: 30736   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.31
episode: 173   score: 1.0   memory length: 30907   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.31
episode: 174   score: 1.0   memory length: 31077   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.3
episode: 175   score: 2.0   memory length: 31274   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.29
episode: 176   score: 0.0   memory length: 31397   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.27
episode: 177   score: 3.0   memory length: 31648   epsilon: 1.0    steps: 251
lr: 0.0001     evaluation reward: 1.29
episode: 178   score: 3.0   memory length: 31895   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.3
episode: 179   score: 0.0   memory length: 32018   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3
episode: 180   score: 2.0   memory length: 32215   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.31
episode: 181   score: 1.0   memory length: 32387   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.3
episode: 182   score: 1.0   memory length: 32538   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.27
episode: 183   score: 1.0   memory length: 32707   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.28
episode: 184   score: 2.0   memory length: 32927   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.29
episode: 185   score: 0.0   memory length: 33050   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.25
episode: 186   score: 0.0   memory length: 33173   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.24
episode: 187   score: 1.0   memory length: 33323   epsilon: 1.0    steps: 150
```

```
lr: 0.0001     evaluation reward: 1.2
episode: 188   score: 0.0   memory length: 33446   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.18
episode: 189   score: 2.0   memory length: 33662   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.2
episode: 190   score: 3.0   memory length: 33908   epsilon: 1.0    steps: 246
lr: 0.0001     evaluation reward: 1.2
episode: 191   score: 0.0   memory length: 34031   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.17
episode: 192   score: 2.0   memory length: 34251   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.19
episode: 193   score: 0.0   memory length: 34374   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.18
episode: 194   score: 1.0   memory length: 34545   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.18
episode: 195   score: 0.0   memory length: 34668   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.18
episode: 196   score: 3.0   memory length: 34931   epsilon: 1.0    steps: 263
lr: 0.0001     evaluation reward: 1.2
episode: 197   score: 2.0   memory length: 35129   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.22
episode: 198   score: 0.0   memory length: 35252   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.19
episode: 199   score: 1.0   memory length: 35402   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.2
episode: 200   score: 0.0   memory length: 35525   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.18
episode: 201   score: 0.0   memory length: 35648   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.16
episode: 202   score: 1.0   memory length: 35817   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.17
episode: 203   score: 5.0   memory length: 36124   epsilon: 1.0    steps: 307
lr: 0.0001     evaluation reward: 1.22
episode: 204   score: 2.0   memory length: 36340   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.24
episode: 205   score: 1.0   memory length: 36508   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.24
episode: 206   score: 2.0   memory length: 36690   epsilon: 1.0    steps: 182
lr: 0.0001     evaluation reward: 1.25
episode: 207   score: 0.0   memory length: 36813   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.24
episode: 208   score: 2.0   memory length: 37033   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.26
episode: 209   score: 0.0   memory length: 37156   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.24
episode: 210   score: 2.0   memory length: 37372   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.23
episode: 211   score: 0.0   memory length: 37494   epsilon: 1.0    steps: 122
```

```
lr: 0.0001     evaluation reward: 1.21
episode: 212   score: 1.0   memory length: 37646   epsilon: 1.0    steps: 152
lr: 0.0001     evaluation reward: 1.21
episode: 213   score: 5.0   memory length: 37986   epsilon: 1.0    steps: 340
lr: 0.0001     evaluation reward: 1.24
episode: 214   score: 1.0   memory length: 38137   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.24
episode: 215   score: 1.0   memory length: 38305   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.24
episode: 216   score: 0.0   memory length: 38428   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.24
episode: 217   score: 0.0   memory length: 38550   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.21
episode: 218   score: 0.0   memory length: 38673   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.19
episode: 219   score: 0.0   memory length: 38795   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.19
episode: 220   score: 5.0   memory length: 39096   epsilon: 1.0    steps: 301
lr: 0.0001     evaluation reward: 1.22
episode: 221   score: 0.0   memory length: 39219   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.21
episode: 222   score: 2.0   memory length: 39417   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.21
episode: 223   score: 0.0   memory length: 39540   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.21
episode: 224   score: 0.0   memory length: 39662   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.2
episode: 225   score: 0.0   memory length: 39785   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.19
episode: 226   score: 2.0   memory length: 39985   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.19
episode: 227   score: 2.0   memory length: 40204   epsilon: 1.0    steps: 219
lr: 0.0001     evaluation reward: 1.2
episode: 228   score: 3.0   memory length: 40429   epsilon: 1.0    steps: 225
lr: 0.0001     evaluation reward: 1.23
episode: 229   score: 2.0   memory length: 40626   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.25
episode: 230   score: 3.0   memory length: 40854   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.26
episode: 231   score: 2.0   memory length: 41070   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.28
episode: 232   score: 0.0   memory length: 41192   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.27
episode: 233   score: 2.0   memory length: 41411   epsilon: 1.0    steps: 219
lr: 0.0001     evaluation reward: 1.28
episode: 234   score: 2.0   memory length: 41629   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.27
episode: 235   score: 0.0   memory length: 41752   epsilon: 1.0    steps: 123
```

```
lr: 0.0001     evaluation reward: 1.26
episode: 236   score: 0.0   memory length: 41875   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.24
episode: 237   score: 2.0   memory length: 42073   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.25
episode: 238   score: 0.0   memory length: 42195   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.23
episode: 239   score: 1.0   memory length: 42346   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.24
episode: 240   score: 1.0   memory length: 42515   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.24
episode: 241   score: 3.0   memory length: 42762   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.26
episode: 242   score: 0.0   memory length: 42885   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.25
episode: 243   score: 2.0   memory length: 43082   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.26
episode: 244   score: 3.0   memory length: 43329   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.28
episode: 245   score: 1.0   memory length: 43480   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.28
episode: 246   score: 0.0   memory length: 43603   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.26
episode: 247   score: 0.0   memory length: 43725   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.25
episode: 248   score: 1.0   memory length: 43895   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.23
episode: 249   score: 4.0   memory length: 44190   epsilon: 1.0    steps: 295
lr: 0.0001     evaluation reward: 1.25
episode: 250   score: 1.0   memory length: 44360   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.26
episode: 251   score: 0.0   memory length: 44483   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.26
episode: 252   score: 0.0   memory length: 44606   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.23
episode: 253   score: 1.0   memory length: 44757   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.24
episode: 254   score: 4.0   memory length: 45022   epsilon: 1.0    steps: 265
lr: 0.0001     evaluation reward: 1.28
episode: 255   score: 0.0   memory length: 45145   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.27
episode: 256   score: 1.0   memory length: 45313   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.27
episode: 257   score: 2.0   memory length: 45531   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.28
episode: 258   score: 3.0   memory length: 45756   epsilon: 1.0    steps: 225
lr: 0.0001     evaluation reward: 1.3
episode: 259   score: 1.0   memory length: 45926   epsilon: 1.0    steps: 170
```

```
lr: 0.0001     evaluation reward: 1.31
episode: 260   score: 3.0   memory length: 46191   epsilon: 1.0    steps: 265
lr: 0.0001     evaluation reward: 1.33
episode: 261   score: 2.0   memory length: 46371   epsilon: 1.0    steps: 180
lr: 0.0001     evaluation reward: 1.34
episode: 262   score: 0.0   memory length: 46493   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.33
episode: 263   score: 3.0   memory length: 46719   epsilon: 1.0    steps: 226
lr: 0.0001     evaluation reward: 1.32
episode: 264   score: 0.0   memory length: 46842   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3
episode: 265   score: 9.0   memory length: 47353   epsilon: 1.0    steps: 511
lr: 0.0001     evaluation reward: 1.39
episode: 266   score: 0.0   memory length: 47475   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.39
episode: 267   score: 1.0   memory length: 47644   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.37
episode: 268   score: 3.0   memory length: 47872   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.37
episode: 269   score: 3.0   memory length: 48138   epsilon: 1.0    steps: 266
lr: 0.0001     evaluation reward: 1.38
episode: 270   score: 1.0   memory length: 48308   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.39
episode: 271   score: 2.0   memory length: 48526   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.4
episode: 272   score: 0.0   memory length: 48648   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.38
episode: 273   score: 2.0   memory length: 48845   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.39
episode: 274   score: 2.0   memory length: 49063   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.4
episode: 275   score: 3.0   memory length: 49334   epsilon: 1.0    steps: 271
lr: 0.0001     evaluation reward: 1.41
episode: 276   score: 3.0   memory length: 49600   epsilon: 1.0    steps: 266
lr: 0.0001     evaluation reward: 1.44
episode: 277   score: 0.0   memory length: 49722   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.41
episode: 278   score: 2.0   memory length: 49920   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.4
episode: 279   score: 5.0   memory length: 50283   epsilon: 1.0    steps: 363
lr: 0.0001     evaluation reward: 1.45
episode: 280   score: 2.0   memory length: 50500   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.45
episode: 281   score: 0.0   memory length: 50623   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.44
episode: 282   score: 0.0   memory length: 50746   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.43
episode: 283   score: 0.0   memory length: 50868   epsilon: 1.0    steps: 122
```

```
lr: 0.0001     evaluation reward: 1.42
episode: 284   score: 3.0   memory length: 51114   epsilon: 1.0    steps: 246
lr: 0.0001     evaluation reward: 1.43
episode: 285   score: 2.0   memory length: 51312   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.45
episode: 286   score: 3.0   memory length: 51562   epsilon: 1.0    steps: 250
lr: 0.0001     evaluation reward: 1.48
episode: 287   score: 2.0   memory length: 51759   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.49
episode: 288   score: 0.0   memory length: 51882   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.49
episode: 289   score: 0.0   memory length: 52005   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.47
episode: 290   score: 0.0   memory length: 52127   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.44
episode: 291   score: 0.0   memory length: 52250   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.44
episode: 292   score: 2.0   memory length: 52447   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.44
episode: 293   score: 3.0   memory length: 52694   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.47
episode: 294   score: 1.0   memory length: 52844   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.47
episode: 295   score: 2.0   memory length: 53060   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.49
episode: 296   score: 1.0   memory length: 53228   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.47
episode: 297   score: 1.0   memory length: 53399   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.46
episode: 298   score: 0.0   memory length: 53521   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.46
episode: 299   score: 2.0   memory length: 53739   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.47
episode: 300   score: 1.0   memory length: 53891   epsilon: 1.0    steps: 152
lr: 0.0001     evaluation reward: 1.48
episode: 301   score: 0.0   memory length: 54013   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.48
episode: 302   score: 0.0   memory length: 54136   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.47
episode: 303   score: 3.0   memory length: 54379   epsilon: 1.0    steps: 243
lr: 0.0001     evaluation reward: 1.45
episode: 304   score: 1.0   memory length: 54548   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.44
episode: 305   score: 0.0   memory length: 54671   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.43
episode: 306   score: 0.0   memory length: 54794   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.41
episode: 307   score: 4.0   memory length: 55055   epsilon: 1.0    steps: 261
```

```
lr: 0.0001     evaluation reward: 1.45
episode: 308   score: 1.0   memory length: 55206   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.44
episode: 309   score: 0.0   memory length: 55329   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.44
episode: 310   score: 0.0   memory length: 55452   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.42
episode: 311   score: 2.0   memory length: 55650   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.44
episode: 312   score: 1.0   memory length: 55801   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.44
episode: 313   score: 1.0   memory length: 55970   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.4
episode: 314   score: 0.0   memory length: 56093   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.39
episode: 315   score: 0.0   memory length: 56215   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.38
episode: 316   score: 0.0   memory length: 56338   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.38
episode: 317   score: 0.0   memory length: 56461   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.38
episode: 318   score: 0.0   memory length: 56583   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.38
episode: 319   score: 2.0   memory length: 56803   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.4
episode: 320   score: 2.0   memory length: 57002   epsilon: 1.0    steps: 199
lr: 0.0001     evaluation reward: 1.37
episode: 321   score: 1.0   memory length: 57153   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.38
episode: 322   score: 0.0   memory length: 57276   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.36
episode: 323   score: 2.0   memory length: 57494   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.38
episode: 324   score: 2.0   memory length: 57694   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.4
episode: 325   score: 2.0   memory length: 57911   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.42
episode: 326   score: 2.0   memory length: 58126   epsilon: 1.0    steps: 215
lr: 0.0001     evaluation reward: 1.42
episode: 327   score: 0.0   memory length: 58249   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4
episode: 328   score: 0.0   memory length: 58372   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.37
episode: 329   score: 0.0   memory length: 58495   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.35
episode: 330   score: 0.0   memory length: 58618   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.32
episode: 331   score: 0.0   memory length: 58741   epsilon: 1.0    steps: 123
```

```
lr: 0.0001     evaluation reward: 1.3
episode: 332   score: 2.0   memory length: 58957   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.32
episode: 333   score: 2.0   memory length: 59157   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.32
episode: 334   score: 2.0   memory length: 59375   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.32
episode: 335   score: 1.0   memory length: 59544   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.33
episode: 336   score: 1.0   memory length: 59712   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.34
episode: 337   score: 0.0   memory length: 59834   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.32
episode: 338   score: 1.0   memory length: 60006   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.33
episode: 339   score: 1.0   memory length: 60175   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.33
episode: 340   score: 2.0   memory length: 60392   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.34
episode: 341   score: 2.0   memory length: 60608   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.33
episode: 342   score: 1.0   memory length: 60777   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.34
episode: 343   score: 0.0   memory length: 60899   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.32
episode: 344   score: 0.0   memory length: 61022   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.29
episode: 345   score: 0.0   memory length: 61145   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.28
episode: 346   score: 2.0   memory length: 61363   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.3
episode: 347   score: 1.0   memory length: 61533   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.31
episode: 348   score: 2.0   memory length: 61731   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.32
episode: 349   score: 0.0   memory length: 61854   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.28
episode: 350   score: 2.0   memory length: 62052   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.29
episode: 351   score: 0.0   memory length: 62175   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.29
episode: 352   score: 1.0   memory length: 62344   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.3
episode: 353   score: 4.0   memory length: 62602   epsilon: 1.0    steps: 258
lr: 0.0001     evaluation reward: 1.33
episode: 354   score: 2.0   memory length: 62818   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.31
episode: 355   score: 1.0   memory length: 62970   epsilon: 1.0    steps: 152
```

```
lr: 0.0001     evaluation reward: 1.32
episode: 356   score: 2.0   memory length: 63171   epsilon: 1.0    steps: 201
lr: 0.0001     evaluation reward: 1.33
episode: 357   score: 1.0   memory length: 63322   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.32
episode: 358   score: 0.0   memory length: 63445   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.29
episode: 359   score: 0.0   memory length: 63567   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.28
episode: 360   score: 4.0   memory length: 63861   epsilon: 1.0    steps: 294
lr: 0.0001     evaluation reward: 1.29
episode: 361   score: 3.0   memory length: 64108   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.3
episode: 362   score: 0.0   memory length: 64231   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3
episode: 363   score: 2.0   memory length: 64446   epsilon: 1.0    steps: 215
lr: 0.0001     evaluation reward: 1.29
episode: 364   score: 0.0   memory length: 64569   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.29
episode: 365   score: 0.0   memory length: 64691   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.2
episode: 366   score: 0.0   memory length: 64814   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.2
episode: 367   score: 2.0   memory length: 65034   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.21
episode: 368   score: 0.0   memory length: 65157   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.18
episode: 369   score: 1.0   memory length: 65328   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.16
episode: 370   score: 0.0   memory length: 65450   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.15
episode: 371   score: 2.0   memory length: 65670   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.15
episode: 372   score: 2.0   memory length: 65868   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.17
episode: 373   score: 2.0   memory length: 66066   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.17
episode: 374   score: 1.0   memory length: 66238   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.16
episode: 375   score: 3.0   memory length: 66508   epsilon: 1.0    steps: 270
lr: 0.0001     evaluation reward: 1.16
episode: 376   score: 1.0   memory length: 66677   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.14
episode: 377   score: 2.0   memory length: 66877   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.16
episode: 378   score: 1.0   memory length: 67028   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.15
episode: 379   score: 6.0   memory length: 67411   epsilon: 1.0    steps: 383
```

```
lr: 0.0001     evaluation reward: 1.16
episode: 380   score: 2.0   memory length: 67609   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.16
episode: 381   score: 3.0   memory length: 67836   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.19
episode: 382   score: 3.0   memory length: 68083   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.22
episode: 383   score: 1.0   memory length: 68254   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.23
episode: 384   score: 1.0   memory length: 68426   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.21
episode: 385   score: 0.0   memory length: 68549   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.19
episode: 386   score: 3.0   memory length: 68795   epsilon: 1.0    steps: 246
lr: 0.0001     evaluation reward: 1.19
episode: 387   score: 3.0   memory length: 69065   epsilon: 1.0    steps: 270
lr: 0.0001     evaluation reward: 1.2
episode: 388   score: 3.0   memory length: 69327   epsilon: 1.0    steps: 262
lr: 0.0001     evaluation reward: 1.23
episode: 389   score: 3.0   memory length: 69574   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.26
episode: 390   score: 1.0   memory length: 69725   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.27
episode: 391   score: 2.0   memory length: 69943   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.29
episode: 392   score: 3.0   memory length: 70206   epsilon: 1.0    steps: 263
lr: 0.0001     evaluation reward: 1.3
episode: 393   score: 1.0   memory length: 70375   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.28
episode: 394   score: 1.0   memory length: 70526   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.28
episode: 395   score: 2.0   memory length: 70742   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.28
episode: 396   score: 3.0   memory length: 71008   epsilon: 1.0    steps: 266
lr: 0.0001     evaluation reward: 1.3
episode: 397   score: 4.0   memory length: 71286   epsilon: 1.0    steps: 278
lr: 0.0001     evaluation reward: 1.33
episode: 398   score: 0.0   memory length: 71409   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.33
episode: 399   score: 2.0   memory length: 71611   epsilon: 1.0    steps: 202
lr: 0.0001     evaluation reward: 1.33
episode: 400   score: 0.0   memory length: 71734   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.32
episode: 401   score: 1.0   memory length: 71902   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.33
episode: 402   score: 1.0   memory length: 72074   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.34
episode: 403   score: 0.0   memory length: 72197   epsilon: 1.0    steps: 123
```

```
lr: 0.0001     evaluation reward: 1.31
episode: 404   score: 0.0   memory length: 72320   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3
episode: 405   score: 1.0   memory length: 72471   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.31
episode: 406   score: 1.0   memory length: 72622   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.32
episode: 407   score: 1.0   memory length: 72791   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.29
episode: 408   score: 3.0   memory length: 73037   epsilon: 1.0    steps: 246
lr: 0.0001     evaluation reward: 1.31
episode: 409   score: 2.0   memory length: 73235   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.33
episode: 410   score: 3.0   memory length: 73478   epsilon: 1.0    steps: 243
lr: 0.0001     evaluation reward: 1.36
episode: 411   score: 2.0   memory length: 73659   epsilon: 1.0    steps: 181
lr: 0.0001     evaluation reward: 1.36
episode: 412   score: 1.0   memory length: 73810   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.36
episode: 413   score: 1.0   memory length: 73979   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.36
episode: 414   score: 0.0   memory length: 74102   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.36
episode: 415   score: 2.0   memory length: 74318   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.38
episode: 416   score: 2.0   memory length: 74518   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.4
episode: 417   score: 2.0   memory length: 74699   epsilon: 1.0    steps: 181
lr: 0.0001     evaluation reward: 1.42
episode: 418   score: 0.0   memory length: 74821   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.42
episode: 419   score: 4.0   memory length: 75135   epsilon: 1.0    steps: 314
lr: 0.0001     evaluation reward: 1.44
episode: 420   score: 0.0   memory length: 75258   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.42
episode: 421   score: 1.0   memory length: 75427   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.42
episode: 422   score: 1.0   memory length: 75596   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.43
episode: 423   score: 4.0   memory length: 75872   epsilon: 1.0    steps: 276
lr: 0.0001     evaluation reward: 1.45
episode: 424   score: 1.0   memory length: 76023   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.44
episode: 425   score: 3.0   memory length: 76249   epsilon: 1.0    steps: 226
lr: 0.0001     evaluation reward: 1.45
episode: 426   score: 1.0   memory length: 76421   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.44
episode: 427   score: 0.0   memory length: 76544   epsilon: 1.0    steps: 123
```

```
lr: 0.0001     evaluation reward: 1.44
episode: 428   score: 0.0   memory length: 76667   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.44
episode: 429   score: 2.0   memory length: 76867   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.46
episode: 430   score: 2.0   memory length: 77069   epsilon: 1.0    steps: 202
lr: 0.0001     evaluation reward: 1.48
episode: 431   score: 0.0   memory length: 77192   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.48
episode: 432   score: 0.0   memory length: 77315   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.46
episode: 433   score: 4.0   memory length: 77607   epsilon: 1.0    steps: 292
lr: 0.0001     evaluation reward: 1.48
episode: 434   score: 0.0   memory length: 77730   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.46
episode: 435   score: 2.0   memory length: 77948   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.47
episode: 436   score: 1.0   memory length: 78117   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.47
episode: 437   score: 1.0   memory length: 78289   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.48
episode: 438   score: 0.0   memory length: 78412   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.47
episode: 439   score: 2.0   memory length: 78614   epsilon: 1.0    steps: 202
lr: 0.0001     evaluation reward: 1.48
episode: 440   score: 2.0   memory length: 78833   epsilon: 1.0    steps: 219
lr: 0.0001     evaluation reward: 1.48
episode: 441   score: 2.0   memory length: 79053   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.48
episode: 442   score: 0.0   memory length: 79175   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.47
episode: 443   score: 2.0   memory length: 79373   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.49
episode: 444   score: 2.0   memory length: 79571   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.51
episode: 445   score: 1.0   memory length: 79740   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.52
episode: 446   score: 0.0   memory length: 79863   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.5
episode: 447   score: 0.0   memory length: 79986   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.49
episode: 448   score: 1.0   memory length: 80155   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.48
episode: 449   score: 1.0   memory length: 80306   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.49
episode: 450   score: 1.0   memory length: 80474   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.48
episode: 451   score: 0.0   memory length: 80596   epsilon: 1.0    steps: 122
```

```
lr: 0.0001     evaluation reward: 1.48
episode: 452   score: 0.0   memory length: 80719   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.47
episode: 453   score: 2.0   memory length: 80936   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.45
episode: 454   score: 0.0   memory length: 81059   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.43
episode: 455   score: 4.0   memory length: 81357   epsilon: 1.0    steps: 298
lr: 0.0001     evaluation reward: 1.46
episode: 456   score: 0.0   memory length: 81479   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.44
episode: 457   score: 2.0   memory length: 81676   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.45
episode: 458   score: 3.0   memory length: 81926   epsilon: 1.0    steps: 250
lr: 0.0001     evaluation reward: 1.48
episode: 459   score: 1.0   memory length: 82078   epsilon: 1.0    steps: 152
lr: 0.0001     evaluation reward: 1.49
episode: 460   score: 0.0   memory length: 82201   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.45
episode: 461   score: 0.0   memory length: 82324   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.42
episode: 462   score: 0.0   memory length: 82447   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.42
episode: 463   score: 1.0   memory length: 82598   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.41
episode: 464   score: 1.0   memory length: 82767   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.42
episode: 465   score: 1.0   memory length: 82918   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.43
episode: 466   score: 0.0   memory length: 83041   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.43
episode: 467   score: 4.0   memory length: 83357   epsilon: 1.0    steps: 316
lr: 0.0001     evaluation reward: 1.45
episode: 468   score: 0.0   memory length: 83479   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.45
episode: 469   score: 0.0   memory length: 83602   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.44
episode: 470   score: 2.0   memory length: 83800   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.46
episode: 471   score: 0.0   memory length: 83923   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.44
episode: 472   score: 1.0   memory length: 84092   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.43
episode: 473   score: 2.0   memory length: 84308   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.43
episode: 474   score: 1.0   memory length: 84476   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.43
episode: 475   score: 2.0   memory length: 84673   epsilon: 1.0    steps: 197
```

```
lr: 0.0001     evaluation reward: 1.42
episode: 476   score: 0.0   memory length: 84795   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.41
episode: 477   score: 3.0   memory length: 85042   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.42
episode: 478   score: 3.0   memory length: 85288   epsilon: 1.0    steps: 246
lr: 0.0001     evaluation reward: 1.44
episode: 479   score: 0.0   memory length: 85411   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.38
episode: 480   score: 2.0   memory length: 85608   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.38
episode: 481   score: 0.0   memory length: 85730   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.35
episode: 482   score: 4.0   memory length: 85987   epsilon: 1.0    steps: 257
lr: 0.0001     evaluation reward: 1.36
episode: 483   score: 0.0   memory length: 86110   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.35
episode: 484   score: 1.0   memory length: 86261   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.35
episode: 485   score: 2.0   memory length: 86459   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.37
episode: 486   score: 3.0   memory length: 86685   epsilon: 1.0    steps: 226
lr: 0.0001     evaluation reward: 1.37
episode: 487   score: 4.0   memory length: 86977   epsilon: 1.0    steps: 292
lr: 0.0001     evaluation reward: 1.38
episode: 488   score: 0.0   memory length: 87100   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.35
episode: 489   score: 1.0   memory length: 87269   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.33
episode: 490   score: 1.0   memory length: 87438   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.33
episode: 491   score: 2.0   memory length: 87654   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.33
episode: 492   score: 0.0   memory length: 87777   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3
episode: 493   score: 2.0   memory length: 87975   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.31
episode: 494   score: 0.0   memory length: 88098   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3
episode: 495   score: 0.0   memory length: 88220   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.28
episode: 496   score: 1.0   memory length: 88371   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.26
episode: 497   score: 1.0   memory length: 88542   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.23
episode: 498   score: 2.0   memory length: 88740   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.25
episode: 499   score: 5.0   memory length: 89073   epsilon: 1.0    steps: 333
```

```
lr: 0.0001     evaluation reward: 1.28
episode: 500   score: 1.0   memory length: 89244   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.29
episode: 501   score: 3.0   memory length: 89491   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.31
episode: 502   score: 3.0   memory length: 89760   epsilon: 1.0    steps: 269
lr: 0.0001     evaluation reward: 1.33
episode: 503   score: 4.0   memory length: 90073   epsilon: 1.0    steps: 313
lr: 0.0001     evaluation reward: 1.37
episode: 504   score: 2.0   memory length: 90294   epsilon: 1.0    steps: 221
lr: 0.0001     evaluation reward: 1.39
episode: 505   score: 1.0   memory length: 90445   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.39
episode: 506   score: 3.0   memory length: 90676   epsilon: 1.0    steps: 231
lr: 0.0001     evaluation reward: 1.41
episode: 507   score: 0.0   memory length: 90798   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.4
episode: 508   score: 2.0   memory length: 90997   epsilon: 1.0    steps: 199
lr: 0.0001     evaluation reward: 1.39
episode: 509   score: 3.0   memory length: 91264   epsilon: 1.0    steps: 267
lr: 0.0001     evaluation reward: 1.4
episode: 510   score: 1.0   memory length: 91434   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.38
episode: 511   score: 3.0   memory length: 91662   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.39
episode: 512   score: 2.0   memory length: 91862   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.4
episode: 513   score: 1.0   memory length: 92031   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.4
episode: 514   score: 0.0   memory length: 92153   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.4
episode: 515   score: 0.0   memory length: 92275   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.38
episode: 516   score: 2.0   memory length: 92473   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.38
episode: 517   score: 1.0   memory length: 92641   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.37
episode: 518   score: 5.0   memory length: 93008   epsilon: 1.0    steps: 367
lr: 0.0001     evaluation reward: 1.42
episode: 519   score: 0.0   memory length: 93131   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.38
episode: 520   score: 1.0   memory length: 93303   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.39
episode: 521   score: 5.0   memory length: 93627   epsilon: 1.0    steps: 324
lr: 0.0001     evaluation reward: 1.43
episode: 522   score: 2.0   memory length: 93844   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.44
episode: 523   score: 0.0   memory length: 93966   epsilon: 1.0    steps: 122
```

```
lr: 0.0001     evaluation reward: 1.4
episode: 524   score: 0.0   memory length: 94089   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.39
episode: 525   score: 0.0   memory length: 94212   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.36
episode: 526   score: 0.0   memory length: 94335   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.35
episode: 527   score: 2.0   memory length: 94550   epsilon: 1.0    steps: 215
lr: 0.0001     evaluation reward: 1.37
episode: 528   score: 1.0   memory length: 94719   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.38
episode: 529   score: 0.0   memory length: 94841   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.36
episode: 530   score: 2.0   memory length: 95039   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.36
episode: 531   score: 0.0   memory length: 95161   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.36
episode: 532   score: 3.0   memory length: 95388   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.39
episode: 533   score: 3.0   memory length: 95615   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.38
episode: 534   score: 3.0   memory length: 95861   epsilon: 1.0    steps: 246
lr: 0.0001     evaluation reward: 1.41
episode: 535   score: 1.0   memory length: 96012   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.4
episode: 536   score: 0.0   memory length: 96135   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.39
episode: 537   score: 2.0   memory length: 96333   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.4
episode: 538   score: 1.0   memory length: 96504   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.41
episode: 539   score: 1.0   memory length: 96655   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.4
episode: 540   score: 2.0   memory length: 96873   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.4
episode: 541   score: 1.0   memory length: 97044   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.39
episode: 542   score: 1.0   memory length: 97213   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.4
episode: 543   score: 1.0   memory length: 97364   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.39
episode: 544   score: 1.0   memory length: 97533   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.38
episode: 545   score: 2.0   memory length: 97748   epsilon: 1.0    steps: 215
lr: 0.0001     evaluation reward: 1.39
episode: 546   score: 0.0   memory length: 97870   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.39
episode: 547   score: 0.0   memory length: 97992   epsilon: 1.0    steps: 122
```

```
lr: 0.0001     evaluation reward: 1.39
episode: 548   score: 2.0   memory length: 98209   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.4
episode: 549   score: 2.0   memory length: 98406   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.41
episode: 550   score: 2.0   memory length: 98624   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.42
episode: 551   score: 2.0   memory length: 98821   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.44
episode: 552   score: 4.0   memory length: 99139   epsilon: 1.0    steps: 318
lr: 0.0001     evaluation reward: 1.48
episode: 553   score: 1.0   memory length: 99309   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.47
episode: 554   score: 3.0   memory length: 99553   epsilon: 1.0    steps: 244
lr: 0.0001     evaluation reward: 1.5
episode: 555   score: 1.0   memory length: 99721   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.47
episode: 556   score: 3.0   memory length: 99948   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.5
```

```
/u/qilong/QA_generator/CS444/mp5/memory.py:30: FutureWarning: The input object
of type 'Tensor' is an array-like implementing one of the corresponding
protocols (`__array__`, `__array_interface__` or `__array_struct__`); but not a
sequence (or 0-D). In the future, this object will be coerced as if it was first
converted using `np.array(obj)`. To retain the old behaviour, you have to either
modify the type 'Tensor', or assign to an empty array created with
`np.empty(correct_shape, dtype=object)`.
  sample = np.array(sample, dtype=object)
/u/qilong/QA_generator/CS444/mp5/agent.py:63: FutureWarning: The input object of
type 'Tensor' is an array-like implementing one of the corresponding protocols
(`__array__`, `__array_interface__` or `__array_struct__`); but not a sequence
(or 0-D). In the future, this object will be coerced as if it was first
converted using `np.array(obj)`. To retain the old behaviour, you have to either
modify the type 'Tensor', or assign to an empty array created with
`np.empty(correct_shape, dtype=object)`.
  mini_batch = np.array(mini_batch, dtype=object).transpose()
```

```
episode: 557   score: 0.0   memory length: 100071   epsilon: 0.9998574400000031
steps: 123    lr: 0.0001     evaluation reward: 1.48
episode: 558   score: 0.0   memory length: 100193   epsilon: 0.9996158800000083
steps: 122    lr: 0.0001     evaluation reward: 1.45
episode: 559   score: 5.0   memory length: 100499   epsilon: 0.9990100000000215
steps: 306    lr: 0.0001     evaluation reward: 1.49
episode: 560   score: 0.0   memory length: 100622   epsilon: 0.9987664600000268
steps: 123    lr: 0.0001     evaluation reward: 1.49
episode: 561   score: 0.0   memory length: 100744   epsilon: 0.998524900000032
steps: 122    lr: 0.0001     evaluation reward: 1.49
episode: 562   score: 0.0   memory length: 100866   epsilon: 0.9982833400000373
steps: 122    lr: 0.0001     evaluation reward: 1.49
```

```
episode: 563   score: 3.0   memory length: 101098   epsilon: 0.9978239800000472
steps: 232    lr: 0.0001     evaluation reward: 1.51
episode: 564   score: 0.0   memory length: 101221   epsilon: 0.9975804400000525
steps: 123    lr: 0.0001     evaluation reward: 1.5
episode: 565   score: 1.0   memory length: 101391   epsilon: 0.9972438400000598
steps: 170    lr: 0.0001     evaluation reward: 1.5
episode: 566   score: 1.0   memory length: 101561   epsilon: 0.9969072400000671
steps: 170    lr: 0.0001     evaluation reward: 1.51
episode: 567   score: 2.0   memory length: 101759   epsilon: 0.9965152000000757
steps: 198    lr: 0.0001     evaluation reward: 1.49
episode: 568   score: 3.0   memory length: 102022   epsilon: 0.995994460000087
steps: 263    lr: 0.0001     evaluation reward: 1.52
episode: 569   score: 2.0   memory length: 102239   epsilon: 0.9955648000000963
steps: 217    lr: 0.0001     evaluation reward: 1.54
episode: 570   score: 3.0   memory length: 102487   epsilon: 0.9950737600001069
steps: 248    lr: 0.0001     evaluation reward: 1.55
episode: 571   score: 2.0   memory length: 102685   epsilon: 0.9946817200001155
steps: 198    lr: 0.0001     evaluation reward: 1.57
episode: 572   score: 2.0   memory length: 102902   epsilon: 0.9942520600001248
steps: 217    lr: 0.0001     evaluation reward: 1.58
episode: 573   score: 1.0   memory length: 103071   epsilon: 0.993917440000132
steps: 169    lr: 0.0001     evaluation reward: 1.57
episode: 574   score: 2.0   memory length: 103271   epsilon: 0.9935214400001406
steps: 200    lr: 0.0001     evaluation reward: 1.58
episode: 575   score: 1.0   memory length: 103443   epsilon: 0.993180880000148
steps: 172    lr: 0.0001     evaluation reward: 1.57
episode: 576   score: 1.0   memory length: 103612   epsilon: 0.9928462600001553
steps: 169    lr: 0.0001     evaluation reward: 1.58
episode: 577   score: 3.0   memory length: 103858   epsilon: 0.9923591800001659
steps: 246    lr: 0.0001     evaluation reward: 1.58
episode: 578   score: 1.0   memory length: 104026   epsilon: 0.9920265400001731
steps: 168    lr: 0.0001     evaluation reward: 1.56
episode: 579   score: 2.0   memory length: 104224   epsilon: 0.9916345000001816
steps: 198    lr: 0.0001     evaluation reward: 1.58
episode: 580   score: 0.0   memory length: 104346   epsilon: 0.9913929400001869
steps: 122    lr: 0.0001     evaluation reward: 1.56
episode: 581   score: 5.0   memory length: 104686   epsilon: 0.9907197400002015
steps: 340    lr: 0.0001     evaluation reward: 1.61
episode: 582   score: 4.0   memory length: 104982   epsilon: 0.9901336600002142
steps: 296    lr: 0.0001     evaluation reward: 1.61
episode: 583   score: 1.0   memory length: 105151   epsilon: 0.9897990400002215
steps: 169    lr: 0.0001     evaluation reward: 1.62
episode: 584   score: 1.0   memory length: 105302   epsilon: 0.9895000600002279
steps: 151    lr: 0.0001     evaluation reward: 1.62
episode: 585   score: 2.0   memory length: 105518   epsilon: 0.9890723800002372
steps: 216    lr: 0.0001     evaluation reward: 1.62
episode: 586   score: 1.0   memory length: 105669   epsilon: 0.9887734000002437
steps: 151    lr: 0.0001     evaluation reward: 1.6
```

```
episode: 587   score: 2.0   memory length: 105887   epsilon: 0.9883417600002531
steps: 218    lr: 0.0001     evaluation reward: 1.58
episode: 588   score: 2.0   memory length: 106085   epsilon: 0.9879497200002616
steps: 198    lr: 0.0001     evaluation reward: 1.6
episode: 589   score: 1.0   memory length: 106256   epsilon: 0.987611140000269
steps: 171    lr: 0.0001     evaluation reward: 1.6
episode: 590   score: 0.0   memory length: 106378   epsilon: 0.9873695800002742
steps: 122    lr: 0.0001     evaluation reward: 1.59
episode: 591   score: 1.0   memory length: 106546   epsilon: 0.9870369400002814
steps: 168    lr: 0.0001     evaluation reward: 1.58
episode: 592   score: 2.0   memory length: 106744   epsilon: 0.9866449000002899
steps: 198    lr: 0.0001     evaluation reward: 1.6
episode: 593   score: 6.0   memory length: 107122   epsilon: 0.9858964600003062
steps: 378    lr: 0.0001     evaluation reward: 1.64
episode: 594   score: 0.0   memory length: 107245   epsilon: 0.9856529200003115
steps: 123    lr: 0.0001     evaluation reward: 1.64
episode: 595   score: 2.0   memory length: 107464   epsilon: 0.9852193000003209
steps: 219    lr: 0.0001     evaluation reward: 1.66
episode: 596   score: 0.0   memory length: 107587   epsilon: 0.9849757600003262
steps: 123    lr: 0.0001     evaluation reward: 1.65
episode: 597   score: 2.0   memory length: 107785   epsilon: 0.9845837200003347
steps: 198    lr: 0.0001     evaluation reward: 1.66
episode: 598   score: 1.0   memory length: 107957   epsilon: 0.9842431600003421
steps: 172    lr: 0.0001     evaluation reward: 1.65
episode: 599   score: 1.0   memory length: 108126   epsilon: 0.9839085400003493
steps: 169    lr: 0.0001     evaluation reward: 1.61
episode: 600   score: 7.0   memory length: 108549   epsilon: 0.9830710000003675
steps: 423    lr: 0.0001     evaluation reward: 1.67
episode: 601   score: 0.0   memory length: 108672   epsilon: 0.9828274600003728
steps: 123    lr: 0.0001     evaluation reward: 1.64
episode: 602   score: 0.0   memory length: 108795   epsilon: 0.9825839200003781
steps: 123    lr: 0.0001     evaluation reward: 1.61
episode: 603   score: 0.0   memory length: 108918   epsilon: 0.9823403800003834
steps: 123    lr: 0.0001     evaluation reward: 1.57
episode: 604   score: 2.0   memory length: 109116   epsilon: 0.9819483400003919
steps: 198    lr: 0.0001     evaluation reward: 1.57
episode: 605   score: 0.0   memory length: 109239   epsilon: 0.9817048000003972
steps: 123    lr: 0.0001     evaluation reward: 1.56
episode: 606   score: 2.0   memory length: 109441   epsilon: 0.9813048400004059
steps: 202    lr: 0.0001     evaluation reward: 1.55
episode: 607   score: 3.0   memory length: 109693   epsilon: 0.9808058800004167
steps: 252    lr: 0.0001     evaluation reward: 1.58
episode: 608   score: 2.0   memory length: 109893   epsilon: 0.9804098800004253
steps: 200    lr: 0.0001     evaluation reward: 1.58
episode: 609   score: 1.0   memory length: 110062   epsilon: 0.9800752600004325
steps: 169    lr: 0.0001     evaluation reward: 1.56
episode: 610   score: 0.0   memory length: 110185   epsilon: 0.9798317200004378
steps: 123    lr: 0.0001     evaluation reward: 1.55
```

```
episode: 611   score: 1.0   memory length: 110335   epsilon: 0.9795347200004443
steps: 150    lr: 0.0001     evaluation reward: 1.53
episode: 612   score: 2.0   memory length: 110532   epsilon: 0.9791446600004527
steps: 197    lr: 0.0001     evaluation reward: 1.53
episode: 613   score: 5.0   memory length: 110873   epsilon: 0.9784694800004674
steps: 341    lr: 0.0001     evaluation reward: 1.57
episode: 614   score: 2.0   memory length: 111071   epsilon: 0.9780774400004759
steps: 198    lr: 0.0001     evaluation reward: 1.59
episode: 615   score: 1.0   memory length: 111239   epsilon: 0.9777448000004831
steps: 168    lr: 0.0001     evaluation reward: 1.6
episode: 616   score: 2.0   memory length: 111436   epsilon: 0.9773547400004916
steps: 197    lr: 0.0001     evaluation reward: 1.6
episode: 617   score: 2.0   memory length: 111634   epsilon: 0.9769627000005001
steps: 198    lr: 0.0001     evaluation reward: 1.61
episode: 618   score: 3.0   memory length: 111860   epsilon: 0.9765152200005098
steps: 226    lr: 0.0001     evaluation reward: 1.59
episode: 619   score: 2.0   memory length: 112083   epsilon: 0.9760736800005194
steps: 223    lr: 0.0001     evaluation reward: 1.61
episode: 620   score: 0.0   memory length: 112206   epsilon: 0.9758301400005247
steps: 123    lr: 0.0001     evaluation reward: 1.6
episode: 621   score: 0.0   memory length: 112329   epsilon: 0.97558660000053
steps: 123    lr: 0.0001     evaluation reward: 1.55
episode: 622   score: 1.0   memory length: 112498   epsilon: 0.9752519800005373
steps: 169    lr: 0.0001     evaluation reward: 1.54
episode: 623   score: 0.0   memory length: 112620   epsilon: 0.9750104200005425
steps: 122    lr: 0.0001     evaluation reward: 1.54
episode: 624   score: 1.0   memory length: 112789   epsilon: 0.9746758000005498
steps: 169    lr: 0.0001     evaluation reward: 1.55
episode: 625   score: 1.0   memory length: 112939   epsilon: 0.9743788000005562
steps: 150    lr: 0.0001     evaluation reward: 1.56
episode: 626   score: 1.0   memory length: 113108   epsilon: 0.9740441800005635
steps: 169    lr: 0.0001     evaluation reward: 1.57
episode: 627   score: 0.0   memory length: 113231   epsilon: 0.9738006400005688
steps: 123    lr: 0.0001     evaluation reward: 1.55
episode: 628   score: 6.0   memory length: 113570   epsilon: 0.9731294200005833
steps: 339    lr: 0.0001     evaluation reward: 1.6
episode: 629   score: 1.0   memory length: 113738   epsilon: 0.9727967800005906
steps: 168    lr: 0.0001     evaluation reward: 1.61
episode: 630   score: 0.0   memory length: 113860   epsilon: 0.9725552200005958
steps: 122    lr: 0.0001     evaluation reward: 1.59
episode: 631   score: 2.0   memory length: 114078   epsilon: 0.9721235800006052
steps: 218    lr: 0.0001     evaluation reward: 1.61
episode: 632   score: 1.0   memory length: 114229   epsilon: 0.9718246000006117
steps: 151    lr: 0.0001     evaluation reward: 1.59
episode: 633   score: 0.0   memory length: 114352   epsilon: 0.971581060000617
steps: 123    lr: 0.0001     evaluation reward: 1.56
episode: 634   score: 2.0   memory length: 114550   epsilon: 0.9711890200006255
steps: 198    lr: 0.0001     evaluation reward: 1.55
```

```
episode: 635   score: 0.0   memory length: 114672   epsilon: 0.9709474600006307
steps: 122    lr: 0.0001     evaluation reward: 1.54
episode: 636   score: 0.0   memory length: 114794   epsilon: 0.970705900000636
steps: 122    lr: 0.0001     evaluation reward: 1.54
episode: 637   score: 2.0   memory length: 115010   epsilon: 0.9702782200006452
steps: 216    lr: 0.0001     evaluation reward: 1.54
episode: 638   score: 0.0   memory length: 115133   epsilon: 0.9700346800006505
steps: 123    lr: 0.0001     evaluation reward: 1.53
episode: 639   score: 0.0   memory length: 115256   epsilon: 0.9697911400006558
steps: 123    lr: 0.0001     evaluation reward: 1.52
episode: 640   score: 2.0   memory length: 115454   epsilon: 0.9693991000006643
steps: 198    lr: 0.0001     evaluation reward: 1.52
episode: 641   score: 2.0   memory length: 115652   epsilon: 0.9690070600006728
steps: 198    lr: 0.0001     evaluation reward: 1.53
episode: 642   score: 0.0   memory length: 115775   epsilon: 0.9687635200006781
steps: 123    lr: 0.0001     evaluation reward: 1.52
episode: 643   score: 0.0   memory length: 115898   epsilon: 0.9685199800006834
steps: 123    lr: 0.0001     evaluation reward: 1.51
episode: 644   score: 0.0   memory length: 116021   epsilon: 0.9682764400006887
steps: 123    lr: 0.0001     evaluation reward: 1.5
episode: 645   score: 3.0   memory length: 116267   epsilon: 0.9677893600006993
steps: 246    lr: 0.0001     evaluation reward: 1.51
episode: 646   score: 2.0   memory length: 116465   epsilon: 0.9673973200007078
steps: 198    lr: 0.0001     evaluation reward: 1.53
episode: 647   score: 0.0   memory length: 116588   epsilon: 0.9671537800007131
steps: 123    lr: 0.0001     evaluation reward: 1.53
episode: 648   score: 2.0   memory length: 116785   epsilon: 0.9667637200007215
steps: 197    lr: 0.0001     evaluation reward: 1.53
episode: 649   score: 0.0   memory length: 116908   epsilon: 0.9665201800007268
steps: 123    lr: 0.0001     evaluation reward: 1.51
episode: 650   score: 0.0   memory length: 117031   epsilon: 0.9662766400007321
steps: 123    lr: 0.0001     evaluation reward: 1.49
episode: 651   score: 0.0   memory length: 117154   epsilon: 0.9660331000007374
steps: 123    lr: 0.0001     evaluation reward: 1.47
episode: 652   score: 0.0   memory length: 117277   epsilon: 0.9657895600007427
steps: 123    lr: 0.0001     evaluation reward: 1.43
episode: 653   score: 2.0   memory length: 117494   epsilon: 0.965359900000752
steps: 217    lr: 0.0001     evaluation reward: 1.44
episode: 654   score: 0.0   memory length: 117617   epsilon: 0.9651163600007573
steps: 123    lr: 0.0001     evaluation reward: 1.41
episode: 655   score: 2.0   memory length: 117836   epsilon: 0.9646827400007667
steps: 219    lr: 0.0001     evaluation reward: 1.42
episode: 656   score: 2.0   memory length: 118053   epsilon: 0.964253080000776
steps: 217    lr: 0.0001     evaluation reward: 1.41
episode: 657   score: 0.0   memory length: 118176   epsilon: 0.9640095400007813
steps: 123    lr: 0.0001     evaluation reward: 1.41
episode: 658   score: 2.0   memory length: 118374   epsilon: 0.9636175000007898
steps: 198    lr: 0.0001     evaluation reward: 1.43
```

```
episode: 659   score: 0.0   memory length: 118497   epsilon: 0.9633739600007951
steps: 123    lr: 0.0001     evaluation reward: 1.38
episode: 660   score: 5.0   memory length: 118823   epsilon: 0.9627284800008091
steps: 326    lr: 0.0001     evaluation reward: 1.43
episode: 661   score: 0.0   memory length: 118946   epsilon: 0.9624849400008144
steps: 123    lr: 0.0001     evaluation reward: 1.43
episode: 662   score: 0.0   memory length: 119068   epsilon: 0.9622433800008197
steps: 122    lr: 0.0001     evaluation reward: 1.43
episode: 663   score: 0.0   memory length: 119191   epsilon: 0.961999840000825
steps: 123    lr: 0.0001     evaluation reward: 1.4
episode: 664   score: 1.0   memory length: 119361   epsilon: 0.9616632400008323
steps: 170    lr: 0.0001     evaluation reward: 1.41
episode: 665   score: 1.0   memory length: 119531   epsilon: 0.9613266400008396
steps: 170    lr: 0.0001     evaluation reward: 1.41
episode: 666   score: 3.0   memory length: 119794   epsilon: 0.9608059000008509
steps: 263    lr: 0.0001     evaluation reward: 1.43
episode: 667   score: 2.0   memory length: 120012   epsilon: 0.9603742600008602
steps: 218    lr: 0.0001     evaluation reward: 1.43
episode: 668   score: 0.0   memory length: 120134   epsilon: 0.9601327000008655
steps: 122    lr: 0.0001     evaluation reward: 1.4
episode: 669   score: 0.0   memory length: 120256   epsilon: 0.9598911400008707
steps: 122    lr: 0.0001     evaluation reward: 1.38
episode: 670   score: 0.0   memory length: 120379   epsilon: 0.959647600000876
steps: 123    lr: 0.0001     evaluation reward: 1.35
episode: 671   score: 1.0   memory length: 120548   epsilon: 0.9593129800008833
steps: 169    lr: 0.0001     evaluation reward: 1.34
episode: 672   score: 1.0   memory length: 120698   epsilon: 0.9590159800008897
steps: 150    lr: 0.0001     evaluation reward: 1.33
episode: 673   score: 0.0   memory length: 120820   epsilon: 0.958774420000895
steps: 122    lr: 0.0001     evaluation reward: 1.32
episode: 674   score: 2.0   memory length: 121017   epsilon: 0.9583843600009034
steps: 197    lr: 0.0001     evaluation reward: 1.32
episode: 675   score: 0.0   memory length: 121140   epsilon: 0.9581408200009087
steps: 123    lr: 0.0001     evaluation reward: 1.31
episode: 676   score: 0.0   memory length: 121263   epsilon: 0.957897280000914
steps: 123    lr: 0.0001     evaluation reward: 1.3
episode: 677   score: 0.0   memory length: 121386   epsilon: 0.9576537400009193
steps: 123    lr: 0.0001     evaluation reward: 1.27
episode: 678   score: 1.0   memory length: 121537   epsilon: 0.9573547600009258
steps: 151    lr: 0.0001     evaluation reward: 1.27
episode: 679   score: 3.0   memory length: 121762   epsilon: 0.9569092600009355
steps: 225    lr: 0.0001     evaluation reward: 1.28
episode: 680   score: 5.0   memory length: 122106   epsilon: 0.9562281400009502
steps: 344    lr: 0.0001     evaluation reward: 1.33
episode: 681   score: 2.0   memory length: 122304   epsilon: 0.9558361000009588
steps: 198    lr: 0.0001     evaluation reward: 1.3
episode: 682   score: 2.0   memory length: 122502   epsilon: 0.9554440600009673
steps: 198    lr: 0.0001     evaluation reward: 1.28
```

```
episode: 683   score: 1.0   memory length: 122655   epsilon: 0.9551411200009738
steps: 153    lr: 0.0001     evaluation reward: 1.28
episode: 684   score: 2.0   memory length: 122852   epsilon: 0.9547510600009823
steps: 197    lr: 0.0001     evaluation reward: 1.29
episode: 685   score: 0.0   memory length: 122974   epsilon: 0.9545095000009876
steps: 122    lr: 0.0001     evaluation reward: 1.27
episode: 686   score: 2.0   memory length: 123174   epsilon: 0.9541135000009962
steps: 200    lr: 0.0001     evaluation reward: 1.28
episode: 687   score: 0.0   memory length: 123297   epsilon: 0.9538699600010014
steps: 123    lr: 0.0001     evaluation reward: 1.26
episode: 688   score: 3.0   memory length: 123522   epsilon: 0.9534244600010111
steps: 225    lr: 0.0001     evaluation reward: 1.27
episode: 689   score: 1.0   memory length: 123693   epsilon: 0.9530858800010185
steps: 171    lr: 0.0001     evaluation reward: 1.27
episode: 690   score: 2.0   memory length: 123891   epsilon: 0.952693840001027
steps: 198    lr: 0.0001     evaluation reward: 1.29
episode: 691   score: 0.0   memory length: 124013   epsilon: 0.9524522800010322
steps: 122    lr: 0.0001     evaluation reward: 1.28
episode: 692   score: 1.0   memory length: 124182   epsilon: 0.9521176600010395
steps: 169    lr: 0.0001     evaluation reward: 1.27
episode: 693   score: 3.0   memory length: 124429   epsilon: 0.9516286000010501
steps: 247    lr: 0.0001     evaluation reward: 1.24
episode: 694   score: 3.0   memory length: 124675   epsilon: 0.9511415200010607
steps: 246    lr: 0.0001     evaluation reward: 1.27
episode: 695   score: 0.0   memory length: 124798   epsilon: 0.950897980001066
steps: 123    lr: 0.0001     evaluation reward: 1.25
episode: 696   score: 0.0   memory length: 124920   epsilon: 0.9506564200010712
steps: 122    lr: 0.0001     evaluation reward: 1.25
episode: 697   score: 1.0   memory length: 125088   epsilon: 0.9503237800010784
steps: 168    lr: 0.0001     evaluation reward: 1.24
episode: 698   score: 0.0   memory length: 125211   epsilon: 0.9500802400010837
steps: 123    lr: 0.0001     evaluation reward: 1.23
episode: 699   score: 2.0   memory length: 125431   epsilon: 0.9496446400010932
steps: 220    lr: 0.0001     evaluation reward: 1.24
episode: 700   score: 2.0   memory length: 125649   epsilon: 0.9492130000011025
steps: 218    lr: 0.0001     evaluation reward: 1.19
episode: 701   score: 1.0   memory length: 125818   epsilon: 0.9488783800011098
steps: 169    lr: 0.0001     evaluation reward: 1.2
episode: 702   score: 2.0   memory length: 126019   epsilon: 0.9484804000011184
steps: 201    lr: 0.0001     evaluation reward: 1.22
episode: 703   score: 1.0   memory length: 126189   epsilon: 0.9481438000011257
steps: 170    lr: 0.0001     evaluation reward: 1.23
episode: 704   score: 2.0   memory length: 126387   epsilon: 0.9477517600011343
steps: 198    lr: 0.0001     evaluation reward: 1.23
episode: 705   score: 1.0   memory length: 126538   epsilon: 0.9474527800011407
steps: 151    lr: 0.0001     evaluation reward: 1.24
episode: 706   score: 2.0   memory length: 126735   epsilon: 0.9470627200011492
steps: 197    lr: 0.0001     evaluation reward: 1.24
```

```
episode: 707   score: 3.0   memory length: 126988   epsilon: 0.9465617800011601
steps: 253    lr: 0.0001     evaluation reward: 1.24
episode: 708   score: 0.0   memory length: 127111   epsilon: 0.9463182400011654
steps: 123    lr: 0.0001     evaluation reward: 1.22
episode: 709   score: 4.0   memory length: 127398   epsilon: 0.9457499800011777
steps: 287    lr: 0.0001     evaluation reward: 1.25
episode: 710   score: 2.0   memory length: 127615   epsilon: 0.945320320001187
steps: 217    lr: 0.0001     evaluation reward: 1.27
episode: 711   score: 1.0   memory length: 127783   epsilon: 0.9449876800011943
steps: 168    lr: 0.0001     evaluation reward: 1.27
episode: 712   score: 3.0   memory length: 128014   epsilon: 0.9445303000012042
steps: 231    lr: 0.0001     evaluation reward: 1.28
episode: 713   score: 3.0   memory length: 128262   epsilon: 0.9440392600012149
steps: 248    lr: 0.0001     evaluation reward: 1.26
episode: 714   score: 1.0   memory length: 128433   epsilon: 0.9437006800012222
steps: 171    lr: 0.0001     evaluation reward: 1.25
episode: 715   score: 2.0   memory length: 128651   epsilon: 0.9432690400012316
steps: 218    lr: 0.0001     evaluation reward: 1.26
episode: 716   score: 3.0   memory length: 128900   epsilon: 0.9427760200012423
steps: 249    lr: 0.0001     evaluation reward: 1.27
episode: 717   score: 1.0   memory length: 129068   epsilon: 0.9424433800012495
steps: 168    lr: 0.0001     evaluation reward: 1.26
episode: 718   score: 2.0   memory length: 129266   epsilon: 0.942051340001258
steps: 198    lr: 0.0001     evaluation reward: 1.25
episode: 719   score: 0.0   memory length: 129388   epsilon: 0.9418097800012633
steps: 122    lr: 0.0001     evaluation reward: 1.23
episode: 720   score: 1.0   memory length: 129539   epsilon: 0.9415108000012697
steps: 151    lr: 0.0001     evaluation reward: 1.24
episode: 721   score: 3.0   memory length: 129786   epsilon: 0.9410217400012804
steps: 247    lr: 0.0001     evaluation reward: 1.27
episode: 722   score: 3.0   memory length: 130054   epsilon: 0.9404911000012919
steps: 268    lr: 0.0001     evaluation reward: 1.29
episode: 723   score: 0.0   memory length: 130177   epsilon: 0.9402475600012972
steps: 123    lr: 0.0001     evaluation reward: 1.29
episode: 724   score: 0.0   memory length: 130299   epsilon: 0.9400060000013024
steps: 122    lr: 0.0001     evaluation reward: 1.28
episode: 725   score: 2.0   memory length: 130501   epsilon: 0.9396060400013111
steps: 202    lr: 0.0001     evaluation reward: 1.29
episode: 726   score: 2.0   memory length: 130680   epsilon: 0.9392516200013188
steps: 179    lr: 0.0001     evaluation reward: 1.3
episode: 727   score: 3.0   memory length: 130947   epsilon: 0.9387229600013303
steps: 267    lr: 0.0001     evaluation reward: 1.33
episode: 728   score: 1.0   memory length: 131098   epsilon: 0.9384239800013368
steps: 151    lr: 0.0001     evaluation reward: 1.28
episode: 729   score: 3.0   memory length: 131344   epsilon: 0.9379369000013473
steps: 246    lr: 0.0001     evaluation reward: 1.3
episode: 730   score: 2.0   memory length: 131560   epsilon: 0.9375092200013566
steps: 216    lr: 0.0001     evaluation reward: 1.32
```

```
episode: 731   score: 1.0   memory length: 131710   epsilon: 0.9372122200013631
steps: 150    lr: 0.0001     evaluation reward: 1.31
episode: 732   score: 0.0   memory length: 131833   epsilon: 0.9369686800013683
steps: 123    lr: 0.0001     evaluation reward: 1.3
episode: 733   score: 3.0   memory length: 132081   epsilon: 0.936477640001379
steps: 248    lr: 0.0001     evaluation reward: 1.33
episode: 734   score: 0.0   memory length: 132204   epsilon: 0.9362341000013843
steps: 123    lr: 0.0001     evaluation reward: 1.31
episode: 735   score: 3.0   memory length: 132451   epsilon: 0.9357450400013949
steps: 247    lr: 0.0001     evaluation reward: 1.34
episode: 736   score: 1.0   memory length: 132620   epsilon: 0.9354104200014022
steps: 169    lr: 0.0001     evaluation reward: 1.35
episode: 737   score: 2.0   memory length: 132838   epsilon: 0.9349787800014115
steps: 218    lr: 0.0001     evaluation reward: 1.35
episode: 738   score: 1.0   memory length: 133006   epsilon: 0.9346461400014188
steps: 168    lr: 0.0001     evaluation reward: 1.36
episode: 739   score: 3.0   memory length: 133234   epsilon: 0.9341947000014286
steps: 228    lr: 0.0001     evaluation reward: 1.39
episode: 740   score: 0.0   memory length: 133356   epsilon: 0.9339531400014338
steps: 122    lr: 0.0001     evaluation reward: 1.37
episode: 741   score: 2.0   memory length: 133554   epsilon: 0.9335611000014423
steps: 198    lr: 0.0001     evaluation reward: 1.37
episode: 742   score: 2.0   memory length: 133772   epsilon: 0.9331294600014517
steps: 218    lr: 0.0001     evaluation reward: 1.39
episode: 743   score: 0.0   memory length: 133895   epsilon: 0.932885920001457
steps: 123    lr: 0.0001     evaluation reward: 1.39
episode: 744   score: 0.0   memory length: 134018   epsilon: 0.9326423800014623
steps: 123    lr: 0.0001     evaluation reward: 1.39
episode: 745   score: 2.0   memory length: 134235   epsilon: 0.9322127200014716
steps: 217    lr: 0.0001     evaluation reward: 1.38
episode: 746   score: 4.0   memory length: 134512   epsilon: 0.9316642600014835
steps: 277    lr: 0.0001     evaluation reward: 1.4
episode: 747   score: 1.0   memory length: 134684   epsilon: 0.9313237000014909
steps: 172    lr: 0.0001     evaluation reward: 1.41
episode: 748   score: 1.0   memory length: 134856   epsilon: 0.9309831400014983
steps: 172    lr: 0.0001     evaluation reward: 1.4
episode: 749   score: 1.0   memory length: 135007   epsilon: 0.9306841600015048
steps: 151    lr: 0.0001     evaluation reward: 1.41
episode: 750   score: 2.0   memory length: 135224   epsilon: 0.9302545000015141
steps: 217    lr: 0.0001     evaluation reward: 1.43
episode: 751   score: 0.0   memory length: 135347   epsilon: 0.9300109600015194
steps: 123    lr: 0.0001     evaluation reward: 1.43
episode: 752   score: 0.0   memory length: 135470   epsilon: 0.9297674200015247
steps: 123    lr: 0.0001     evaluation reward: 1.43
episode: 753   score: 0.0   memory length: 135593   epsilon: 0.92952388000153
steps: 123    lr: 0.0001     evaluation reward: 1.41
episode: 754   score: 5.0   memory length: 135927   epsilon: 0.9288625600015443
steps: 334    lr: 0.0001     evaluation reward: 1.46
```

```
episode: 755   score: 1.0   memory length: 136097   epsilon: 0.9285259600015516
steps: 170    lr: 0.0001     evaluation reward: 1.45
episode: 756   score: 0.0   memory length: 136220   epsilon: 0.9282824200015569
steps: 123    lr: 0.0001     evaluation reward: 1.43
episode: 757   score: 1.0   memory length: 136371   epsilon: 0.9279834400015634
steps: 151    lr: 0.0001     evaluation reward: 1.44
episode: 758   score: 1.0   memory length: 136540   epsilon: 0.9276488200015707
steps: 169    lr: 0.0001     evaluation reward: 1.43
episode: 759   score: 2.0   memory length: 136755   epsilon: 0.9272231200015799
steps: 215    lr: 0.0001     evaluation reward: 1.45
episode: 760   score: 1.0   memory length: 136905   epsilon: 0.9269261200015864
steps: 150    lr: 0.0001     evaluation reward: 1.41
episode: 761   score: 3.0   memory length: 137131   epsilon: 0.9264786400015961
steps: 226    lr: 0.0001     evaluation reward: 1.44
episode: 762   score: 3.0   memory length: 137378   epsilon: 0.9259895800016067
steps: 247    lr: 0.0001     evaluation reward: 1.47
episode: 763   score: 2.0   memory length: 137575   epsilon: 0.9255995200016152
steps: 197    lr: 0.0001     evaluation reward: 1.49
episode: 764   score: 3.0   memory length: 137800   epsilon: 0.9251540200016248
steps: 225    lr: 0.0001     evaluation reward: 1.51
episode: 765   score: 3.0   memory length: 138029   epsilon: 0.9247006000016347
steps: 229    lr: 0.0001     evaluation reward: 1.53
episode: 766   score: 1.0   memory length: 138197   epsilon: 0.9243679600016419
steps: 168    lr: 0.0001     evaluation reward: 1.51
episode: 767   score: 3.0   memory length: 138426   epsilon: 0.9239145400016517
steps: 229    lr: 0.0001     evaluation reward: 1.52
episode: 768   score: 4.0   memory length: 138737   epsilon: 0.9232987600016651
steps: 311    lr: 0.0001     evaluation reward: 1.56
episode: 769   score: 2.0   memory length: 138956   epsilon: 0.9228651400016745
steps: 219    lr: 0.0001     evaluation reward: 1.58
episode: 770   score: 3.0   memory length: 139201   epsilon: 0.922380040001685
steps: 245    lr: 0.0001     evaluation reward: 1.61
episode: 771   score: 2.0   memory length: 139420   epsilon: 0.9219464200016945
steps: 219    lr: 0.0001     evaluation reward: 1.62
episode: 772   score: 4.0   memory length: 139695   epsilon: 0.9214019200017063
steps: 275    lr: 0.0001     evaluation reward: 1.65
episode: 773   score: 4.0   memory length: 139973   epsilon: 0.9208514800017182
steps: 278    lr: 0.0001     evaluation reward: 1.69
episode: 774   score: 3.0   memory length: 140219   epsilon: 0.9203644000017288
steps: 246    lr: 0.0001     evaluation reward: 1.7
episode: 775   score: 2.0   memory length: 140435   epsilon: 0.9199367200017381
steps: 216    lr: 0.0001     evaluation reward: 1.72
episode: 776   score: 0.0   memory length: 140558   epsilon: 0.9196931800017434
steps: 123    lr: 0.0001     evaluation reward: 1.72
episode: 777   score: 2.0   memory length: 140776   epsilon: 0.9192615400017528
steps: 218    lr: 0.0001     evaluation reward: 1.74
episode: 778   score: 1.0   memory length: 140946   epsilon: 0.9189249400017601
steps: 170    lr: 0.0001     evaluation reward: 1.74
```

```
episode: 779   score: 1.0   memory length: 141117   epsilon: 0.9185863600017674
steps: 171    lr: 0.0001     evaluation reward: 1.72
episode: 780   score: 1.0   memory length: 141286   epsilon: 0.9182517400017747
steps: 169    lr: 0.0001     evaluation reward: 1.68
episode: 781   score: 3.0   memory length: 141533   epsilon: 0.9177626800017853
steps: 247    lr: 0.0001     evaluation reward: 1.69
episode: 782   score: 0.0   memory length: 141656   epsilon: 0.9175191400017906
steps: 123    lr: 0.0001     evaluation reward: 1.67
episode: 783   score: 4.0   memory length: 141954   epsilon: 0.9169291000018034
steps: 298    lr: 0.0001     evaluation reward: 1.7
episode: 784   score: 4.0   memory length: 142230   epsilon: 0.9163826200018153
steps: 276    lr: 0.0001     evaluation reward: 1.72
episode: 785   score: 3.0   memory length: 142457   epsilon: 0.915933160001825
steps: 227    lr: 0.0001     evaluation reward: 1.75
episode: 786   score: 0.0   memory length: 142580   epsilon: 0.9156896200018303
steps: 123    lr: 0.0001     evaluation reward: 1.73
episode: 787   score: 2.0   memory length: 142778   epsilon: 0.9152975800018388
steps: 198    lr: 0.0001     evaluation reward: 1.75
episode: 788   score: 3.0   memory length: 143030   epsilon: 0.9147986200018496
steps: 252    lr: 0.0001     evaluation reward: 1.75
episode: 789   score: 5.0   memory length: 143327   epsilon: 0.9142105600018624
steps: 297    lr: 0.0001     evaluation reward: 1.79
episode: 790   score: 2.0   memory length: 143545   epsilon: 0.9137789200018718
steps: 218    lr: 0.0001     evaluation reward: 1.79
episode: 791   score: 0.0   memory length: 143667   epsilon: 0.913537360001877
steps: 122    lr: 0.0001     evaluation reward: 1.79
episode: 792   score: 2.0   memory length: 143847   epsilon: 0.9131809600018848
steps: 180    lr: 0.0001     evaluation reward: 1.8
episode: 793   score: 1.0   memory length: 143998   epsilon: 0.9128819800018912
steps: 151    lr: 0.0001     evaluation reward: 1.78
episode: 794   score: 1.0   memory length: 144169   epsilon: 0.9125434000018986
steps: 171    lr: 0.0001     evaluation reward: 1.76
episode: 795   score: 1.0   memory length: 144320   epsilon: 0.9122444200019051
steps: 151    lr: 0.0001     evaluation reward: 1.77
episode: 796   score: 2.0   memory length: 144518   epsilon: 0.9118523800019136
steps: 198    lr: 0.0001     evaluation reward: 1.79
episode: 797   score: 3.0   memory length: 144744   epsilon: 0.9114049000019233
steps: 226    lr: 0.0001     evaluation reward: 1.81
episode: 798   score: 0.0   memory length: 144867   epsilon: 0.9111613600019286
steps: 123    lr: 0.0001     evaluation reward: 1.81
episode: 799   score: 2.0   memory length: 145065   epsilon: 0.9107693200019371
steps: 198    lr: 0.0001     evaluation reward: 1.81
episode: 800   score: 2.0   memory length: 145284   epsilon: 0.9103357000019465
steps: 219    lr: 0.0001     evaluation reward: 1.81
episode: 801   score: 3.0   memory length: 145530   epsilon: 0.9098486200019571
steps: 246    lr: 0.0001     evaluation reward: 1.83
episode: 802   score: 0.0   memory length: 145653   epsilon: 0.9096050800019624
steps: 123    lr: 0.0001     evaluation reward: 1.81
```

```
episode: 803   score: 1.0   memory length: 145822   epsilon: 0.9092704600019696
steps: 169    lr: 0.0001     evaluation reward: 1.81
episode: 804   score: 0.0   memory length: 145944   epsilon: 0.9090289000019749
steps: 122    lr: 0.0001     evaluation reward: 1.79
episode: 805   score: 0.0   memory length: 146067   epsilon: 0.9087853600019802
steps: 123    lr: 0.0001     evaluation reward: 1.78
episode: 806   score: 0.0   memory length: 146190   epsilon: 0.9085418200019855
steps: 123    lr: 0.0001     evaluation reward: 1.76
episode: 807   score: 3.0   memory length: 146416   epsilon: 0.9080943400019952
steps: 226    lr: 0.0001     evaluation reward: 1.76
episode: 808   score: 0.0   memory length: 146539   epsilon: 0.9078508000020005
steps: 123    lr: 0.0001     evaluation reward: 1.76
episode: 809   score: 1.0   memory length: 146707   epsilon: 0.9075181600020077
steps: 168    lr: 0.0001     evaluation reward: 1.73
episode: 810   score: 0.0   memory length: 146830   epsilon: 0.907274620002013
steps: 123    lr: 0.0001     evaluation reward: 1.71
episode: 811   score: 2.0   memory length: 147028   epsilon: 0.9068825800020215
steps: 198    lr: 0.0001     evaluation reward: 1.72
episode: 812   score: 1.0   memory length: 147179   epsilon: 0.906583600002028
steps: 151    lr: 0.0001     evaluation reward: 1.7
episode: 813   score: 1.0   memory length: 147332   epsilon: 0.9062806600020346
steps: 153    lr: 0.0001     evaluation reward: 1.68
episode: 814   score: 2.0   memory length: 147548   epsilon: 0.9058529800020438
steps: 216    lr: 0.0001     evaluation reward: 1.69
episode: 815   score: 3.0   memory length: 147797   epsilon: 0.9053599600020545
steps: 249    lr: 0.0001     evaluation reward: 1.7
episode: 816   score: 4.0   memory length: 148072   epsilon: 0.9048154600020664
steps: 275    lr: 0.0001     evaluation reward: 1.71
episode: 817   score: 2.0   memory length: 148289   epsilon: 0.9043858000020757
steps: 217    lr: 0.0001     evaluation reward: 1.72
episode: 818   score: 2.0   memory length: 148507   epsilon: 0.9039541600020851
steps: 218    lr: 0.0001     evaluation reward: 1.72
episode: 819   score: 1.0   memory length: 148658   epsilon: 0.9036551800020916
steps: 151    lr: 0.0001     evaluation reward: 1.73
episode: 820   score: 3.0   memory length: 148924   epsilon: 0.903128500002103
steps: 266    lr: 0.0001     evaluation reward: 1.75
episode: 821   score: 1.0   memory length: 149093   epsilon: 0.9027938800021102
steps: 169    lr: 0.0001     evaluation reward: 1.73
episode: 822   score: 0.0   memory length: 149215   epsilon: 0.9025523200021155
steps: 122    lr: 0.0001     evaluation reward: 1.7
episode: 823   score: 1.0   memory length: 149366   epsilon: 0.902253340002122
steps: 151    lr: 0.0001     evaluation reward: 1.71
episode: 824   score: 1.0   memory length: 149537   epsilon: 0.9019147600021293
steps: 171    lr: 0.0001     evaluation reward: 1.72
episode: 825   score: 0.0   memory length: 149660   epsilon: 0.9016712200021346
steps: 123    lr: 0.0001     evaluation reward: 1.7
episode: 826   score: 2.0   memory length: 149878   epsilon: 0.901239580002144
steps: 218    lr: 0.0001     evaluation reward: 1.7
```

```
episode: 827   score: 4.0   memory length: 150195   epsilon: 0.9006119200021576
steps: 317    lr: 0.0001     evaluation reward: 1.71
episode: 828   score: 2.0   memory length: 150413   epsilon: 0.900180280002167
steps: 218    lr: 0.0001     evaluation reward: 1.72
episode: 829   score: 0.0   memory length: 150536   epsilon: 0.8999367400021723
steps: 123    lr: 0.0001     evaluation reward: 1.69
episode: 830   score: 0.0   memory length: 150659   epsilon: 0.8996932000021776
steps: 123    lr: 0.0001     evaluation reward: 1.67
episode: 831   score: 2.0   memory length: 150839   epsilon: 0.8993368000021853
steps: 180    lr: 0.0001     evaluation reward: 1.68
episode: 832   score: 0.0   memory length: 150961   epsilon: 0.8990952400021905
steps: 122    lr: 0.0001     evaluation reward: 1.68
episode: 833   score: 2.0   memory length: 151177   epsilon: 0.8986675600021998
steps: 216    lr: 0.0001     evaluation reward: 1.67
episode: 834   score: 3.0   memory length: 151386   epsilon: 0.8982537400022088
steps: 209    lr: 0.0001     evaluation reward: 1.7
episode: 835   score: 2.0   memory length: 151584   epsilon: 0.8978617000022173
steps: 198    lr: 0.0001     evaluation reward: 1.69
episode: 836   score: 0.0   memory length: 151707   epsilon: 0.8976181600022226
steps: 123    lr: 0.0001     evaluation reward: 1.68
episode: 837   score: 0.0   memory length: 151830   epsilon: 0.8973746200022279
steps: 123    lr: 0.0001     evaluation reward: 1.66
episode: 838   score: 4.0   memory length: 152127   epsilon: 0.8967865600022407
steps: 297    lr: 0.0001     evaluation reward: 1.69
episode: 839   score: 0.0   memory length: 152250   epsilon: 0.896543020002246
steps: 123    lr: 0.0001     evaluation reward: 1.66
episode: 840   score: 1.0   memory length: 152419   epsilon: 0.8962084000022532
steps: 169    lr: 0.0001     evaluation reward: 1.67
episode: 841   score: 4.0   memory length: 152691   epsilon: 0.8956698400022649
steps: 272    lr: 0.0001     evaluation reward: 1.69
episode: 842   score: 3.0   memory length: 152919   epsilon: 0.8952184000022747
steps: 228    lr: 0.0001     evaluation reward: 1.7
episode: 843   score: 1.0   memory length: 153089   epsilon: 0.894881800002282
steps: 170    lr: 0.0001     evaluation reward: 1.71
episode: 844   score: 3.0   memory length: 153320   epsilon: 0.8944244200022919
steps: 231    lr: 0.0001     evaluation reward: 1.74
episode: 845   score: 0.0   memory length: 153442   epsilon: 0.8941828600022972
steps: 122    lr: 0.0001     evaluation reward: 1.72
episode: 846   score: 2.0   memory length: 153639   epsilon: 0.8937928000023057
steps: 197    lr: 0.0001     evaluation reward: 1.7
episode: 847   score: 1.0   memory length: 153808   epsilon: 0.8934581800023129
steps: 169    lr: 0.0001     evaluation reward: 1.7
episode: 848   score: 3.0   memory length: 154056   epsilon: 0.8929671400023236
steps: 248    lr: 0.0001     evaluation reward: 1.72
episode: 849   score: 1.0   memory length: 154228   epsilon: 0.892626580002331
steps: 172    lr: 0.0001     evaluation reward: 1.72
episode: 850   score: 1.0   memory length: 154397   epsilon: 0.8922919600023382
steps: 169    lr: 0.0001     evaluation reward: 1.71
```

```
episode: 851   score: 2.0   memory length: 154595   epsilon: 0.8918999200023467
steps: 198    lr: 0.0001     evaluation reward: 1.73
episode: 852   score: 0.0   memory length: 154718   epsilon: 0.891656380002352
steps: 123    lr: 0.0001     evaluation reward: 1.73
episode: 853   score: 0.0   memory length: 154840   epsilon: 0.8914148200023573
steps: 122    lr: 0.0001     evaluation reward: 1.73
episode: 854   score: 3.0   memory length: 155088   epsilon: 0.8909237800023679
steps: 248    lr: 0.0001     evaluation reward: 1.71
episode: 855   score: 0.0   memory length: 155211   epsilon: 0.8906802400023732
steps: 123    lr: 0.0001     evaluation reward: 1.7
episode: 856   score: 5.0   memory length: 155554   epsilon: 0.890001100002388
steps: 343    lr: 0.0001     evaluation reward: 1.75
episode: 857   score: 0.0   memory length: 155677   epsilon: 0.8897575600023933
steps: 123    lr: 0.0001     evaluation reward: 1.74
episode: 858   score: 1.0   memory length: 155848   epsilon: 0.8894189800024006
steps: 171    lr: 0.0001     evaluation reward: 1.74
episode: 859   score: 1.0   memory length: 155999   epsilon: 0.8891200000024071
steps: 151    lr: 0.0001     evaluation reward: 1.73
episode: 860   score: 4.0   memory length: 156276   epsilon: 0.888571540002419
steps: 277    lr: 0.0001     evaluation reward: 1.76
episode: 861   score: 4.0   memory length: 156593   epsilon: 0.8879438800024326
steps: 317    lr: 0.0001     evaluation reward: 1.77
episode: 862   score: 1.0   memory length: 156765   epsilon: 0.88760332000244
steps: 172    lr: 0.0001     evaluation reward: 1.75
episode: 863   score: 2.0   memory length: 156985   epsilon: 0.8871677200024495
steps: 220    lr: 0.0001     evaluation reward: 1.75
episode: 864   score: 1.0   memory length: 157137   epsilon: 0.886866760002456
steps: 152    lr: 0.0001     evaluation reward: 1.73
episode: 865   score: 1.0   memory length: 157308   epsilon: 0.8865281800024634
steps: 171    lr: 0.0001     evaluation reward: 1.71
episode: 866   score: 3.0   memory length: 157556   epsilon: 0.886037140002474
steps: 248    lr: 0.0001     evaluation reward: 1.73
episode: 867   score: 2.0   memory length: 157736   epsilon: 0.8856807400024818
steps: 180    lr: 0.0001     evaluation reward: 1.72
episode: 868   score: 3.0   memory length: 157965   epsilon: 0.8852273200024916
steps: 229    lr: 0.0001     evaluation reward: 1.71
episode: 869   score: 3.0   memory length: 158191   epsilon: 0.8847798400025013
steps: 226    lr: 0.0001     evaluation reward: 1.72
episode: 870   score: 0.0   memory length: 158313   epsilon: 0.8845382800025066
steps: 122    lr: 0.0001     evaluation reward: 1.69
episode: 871   score: 1.0   memory length: 158482   epsilon: 0.8842036600025138
steps: 169    lr: 0.0001     evaluation reward: 1.68
episode: 872   score: 2.0   memory length: 158663   epsilon: 0.8838452800025216
steps: 181    lr: 0.0001     evaluation reward: 1.66
episode: 873   score: 1.0   memory length: 158815   epsilon: 0.8835443200025281
steps: 152    lr: 0.0001     evaluation reward: 1.63
episode: 874   score: 0.0   memory length: 158938   epsilon: 0.8833007800025334
steps: 123    lr: 0.0001     evaluation reward: 1.6
```

```
episode: 875   score: 2.0   memory length: 159136   epsilon: 0.8829087400025419
steps: 198    lr: 0.0001     evaluation reward: 1.6
episode: 876   score: 4.0   memory length: 159423   epsilon: 0.8823404800025543
steps: 287    lr: 0.0001     evaluation reward: 1.64
episode: 877   score: 0.0   memory length: 159545   epsilon: 0.8820989200025595
steps: 122    lr: 0.0001     evaluation reward: 1.62
episode: 878   score: 5.0   memory length: 159826   epsilon: 0.8815425400025716
steps: 281    lr: 0.0001     evaluation reward: 1.66
episode: 879   score: 2.0   memory length: 160045   epsilon: 0.881108920002581
steps: 219    lr: 0.0001     evaluation reward: 1.67
episode: 880   score: 4.0   memory length: 160324   epsilon: 0.880556500002593
steps: 279    lr: 0.0001     evaluation reward: 1.7
episode: 881   score: 1.0   memory length: 160492   epsilon: 0.8802238600026002
steps: 168    lr: 0.0001     evaluation reward: 1.68
episode: 882   score: 2.0   memory length: 160693   epsilon: 0.8798258800026089
steps: 201    lr: 0.0001     evaluation reward: 1.7
episode: 883   score: 0.0   memory length: 160816   epsilon: 0.8795823400026141
steps: 123    lr: 0.0001     evaluation reward: 1.66
episode: 884   score: 2.0   memory length: 161014   epsilon: 0.8791903000026227
steps: 198    lr: 0.0001     evaluation reward: 1.64
episode: 885   score: 1.0   memory length: 161165   epsilon: 0.8788913200026292
steps: 151    lr: 0.0001     evaluation reward: 1.62
episode: 886   score: 3.0   memory length: 161432   epsilon: 0.8783626600026406
steps: 267    lr: 0.0001     evaluation reward: 1.65
episode: 887   score: 3.0   memory length: 161695   epsilon: 0.8778419200026519
steps: 263    lr: 0.0001     evaluation reward: 1.66
episode: 888   score: 4.0   memory length: 161990   epsilon: 0.8772578200026646
steps: 295    lr: 0.0001     evaluation reward: 1.67
episode: 889   score: 2.0   memory length: 162188   epsilon: 0.8768657800026731
steps: 198    lr: 0.0001     evaluation reward: 1.64
episode: 890   score: 4.0   memory length: 162450   epsilon: 0.8763470200026844
steps: 262    lr: 0.0001     evaluation reward: 1.66
episode: 891   score: 1.0   memory length: 162619   epsilon: 0.8760124000026916
steps: 169    lr: 0.0001     evaluation reward: 1.67
episode: 892   score: 2.0   memory length: 162817   epsilon: 0.8756203600027002
steps: 198    lr: 0.0001     evaluation reward: 1.67
episode: 893   score: 1.0   memory length: 162968   epsilon: 0.8753213800027067
steps: 151    lr: 0.0001     evaluation reward: 1.67
episode: 894   score: 2.0   memory length: 163166   epsilon: 0.8749293400027152
steps: 198    lr: 0.0001     evaluation reward: 1.68
episode: 895   score: 0.0   memory length: 163289   epsilon: 0.8746858000027204
steps: 123    lr: 0.0001     evaluation reward: 1.67
episode: 896   score: 1.0   memory length: 163440   epsilon: 0.8743868200027269
steps: 151    lr: 0.0001     evaluation reward: 1.66
episode: 897   score: 2.0   memory length: 163638   epsilon: 0.8739947800027354
steps: 198    lr: 0.0001     evaluation reward: 1.65
episode: 898   score: 3.0   memory length: 163887   epsilon: 0.8735017600027462
steps: 249    lr: 0.0001     evaluation reward: 1.68
```

```
episode: 899   score: 4.0   memory length: 164164   epsilon: 0.8729533000027581
steps: 277    lr: 0.0001     evaluation reward: 1.7
episode: 900   score: 0.0   memory length: 164287   epsilon: 0.8727097600027633
steps: 123    lr: 0.0001     evaluation reward: 1.68
episode: 901   score: 1.0   memory length: 164458   epsilon: 0.8723711800027707
steps: 171    lr: 0.0001     evaluation reward: 1.66
episode: 902   score: 0.0   memory length: 164581   epsilon: 0.872127640002776
steps: 123    lr: 0.0001     evaluation reward: 1.66
episode: 903   score: 3.0   memory length: 164809   epsilon: 0.8716762000027858
steps: 228    lr: 0.0001     evaluation reward: 1.68
episode: 904   score: 1.0   memory length: 164979   epsilon: 0.8713396000027931
steps: 170    lr: 0.0001     evaluation reward: 1.69
episode: 905   score: 0.0   memory length: 165102   epsilon: 0.8710960600027984
steps: 123    lr: 0.0001     evaluation reward: 1.69
episode: 906   score: 5.0   memory length: 165428   epsilon: 0.8704505800028124
steps: 326    lr: 0.0001     evaluation reward: 1.74
episode: 907   score: 3.0   memory length: 165675   epsilon: 0.869961520002823
steps: 247    lr: 0.0001     evaluation reward: 1.74
episode: 908   score: 1.0   memory length: 165843   epsilon: 0.8696288800028302
steps: 168    lr: 0.0001     evaluation reward: 1.75
episode: 909   score: 2.0   memory length: 166043   epsilon: 0.8692328800028388
steps: 200    lr: 0.0001     evaluation reward: 1.76
episode: 910   score: 4.0   memory length: 166340   epsilon: 0.8686448200028516
steps: 297    lr: 0.0001     evaluation reward: 1.8
episode: 911   score: 3.0   memory length: 166548   epsilon: 0.8682329800028605
steps: 208    lr: 0.0001     evaluation reward: 1.81
episode: 912   score: 0.0   memory length: 166670   epsilon: 0.8679914200028658
steps: 122    lr: 0.0001     evaluation reward: 1.8
episode: 913   score: 2.0   memory length: 166868   epsilon: 0.8675993800028743
steps: 198    lr: 0.0001     evaluation reward: 1.81
episode: 914   score: 2.0   memory length: 167069   epsilon: 0.8672014000028829
steps: 201    lr: 0.0001     evaluation reward: 1.81
episode: 915   score: 3.0   memory length: 167297   epsilon: 0.8667499600028927
steps: 228    lr: 0.0001     evaluation reward: 1.81
episode: 916   score: 1.0   memory length: 167469   epsilon: 0.8664094000029001
steps: 172    lr: 0.0001     evaluation reward: 1.78
episode: 917   score: 2.0   memory length: 167667   epsilon: 0.8660173600029086
steps: 198    lr: 0.0001     evaluation reward: 1.78
episode: 918   score: 4.0   memory length: 167963   epsilon: 0.8654312800029214
steps: 296    lr: 0.0001     evaluation reward: 1.8
episode: 919   score: 3.0   memory length: 168208   epsilon: 0.8649461800029319
steps: 245    lr: 0.0001     evaluation reward: 1.82
episode: 920   score: 6.0   memory length: 168534   epsilon: 0.8643007000029459
steps: 326    lr: 0.0001     evaluation reward: 1.85
episode: 921   score: 3.0   memory length: 168803   epsilon: 0.8637680800029575
steps: 269    lr: 0.0001     evaluation reward: 1.87
episode: 922   score: 1.0   memory length: 168954   epsilon: 0.863469100002964
steps: 151    lr: 0.0001     evaluation reward: 1.88
```

```
episode: 923   score: 1.0   memory length: 169126   epsilon: 0.8631285400029713
steps: 172    lr: 0.0001     evaluation reward: 1.88
episode: 924   score: 0.0   memory length: 169249   epsilon: 0.8628850000029766
steps: 123    lr: 0.0001     evaluation reward: 1.87
episode: 925   score: 0.0   memory length: 169371   epsilon: 0.8626434400029819
steps: 122    lr: 0.0001     evaluation reward: 1.87
episode: 926   score: 2.0   memory length: 169592   epsilon: 0.8622058600029914
steps: 221    lr: 0.0001     evaluation reward: 1.87
episode: 927   score: 2.0   memory length: 169792   epsilon: 0.861809860003
steps: 200    lr: 0.0001     evaluation reward: 1.85
episode: 928   score: 3.0   memory length: 170062   epsilon: 0.8612752600030116
steps: 270    lr: 0.0001     evaluation reward: 1.86
episode: 929   score: 1.0   memory length: 170212   epsilon: 0.860978260003018
steps: 150    lr: 0.0001     evaluation reward: 1.87
episode: 930   score: 5.0   memory length: 170517   epsilon: 0.8603743600030311
steps: 305    lr: 0.0001     evaluation reward: 1.92
episode: 931   score: 2.0   memory length: 170720   epsilon: 0.8599724200030399
steps: 203    lr: 0.0001     evaluation reward: 1.92
episode: 932   score: 5.0   memory length: 171027   epsilon: 0.8593645600030531
steps: 307    lr: 0.0001     evaluation reward: 1.97
episode: 933   score: 2.0   memory length: 171247   epsilon: 0.8589289600030625
steps: 220    lr: 0.0001     evaluation reward: 1.97
episode: 934   score: 1.0   memory length: 171397   epsilon: 0.858631960003069
steps: 150    lr: 0.0001     evaluation reward: 1.95
episode: 935   score: 3.0   memory length: 171641   epsilon: 0.8581488400030794
steps: 244    lr: 0.0001     evaluation reward: 1.96
episode: 936   score: 3.0   memory length: 171908   epsilon: 0.8576201800030909
steps: 267    lr: 0.0001     evaluation reward: 1.99
episode: 937   score: 4.0   memory length: 172204   epsilon: 0.8570341000031036
steps: 296    lr: 0.0001     evaluation reward: 2.03
episode: 938   score: 0.0   memory length: 172327   epsilon: 0.8567905600031089
steps: 123    lr: 0.0001     evaluation reward: 1.99
episode: 939   score: 2.0   memory length: 172525   epsilon: 0.8563985200031174
steps: 198    lr: 0.0001     evaluation reward: 2.01
episode: 940   score: 3.0   memory length: 172773   epsilon: 0.8559074800031281
steps: 248    lr: 0.0001     evaluation reward: 2.03
episode: 941   score: 2.0   memory length: 172991   epsilon: 0.8554758400031375
steps: 218    lr: 0.0001     evaluation reward: 2.01
episode: 942   score: 1.0   memory length: 173142   epsilon: 0.855176860003144
steps: 151    lr: 0.0001     evaluation reward: 1.99
episode: 943   score: 2.0   memory length: 173321   epsilon: 0.8548224400031517
steps: 179    lr: 0.0001     evaluation reward: 2.0
episode: 944   score: 2.0   memory length: 173520   epsilon: 0.8544284200031602
steps: 199    lr: 0.0001     evaluation reward: 1.99
episode: 945   score: 3.0   memory length: 173746   epsilon: 0.8539809400031699
steps: 226    lr: 0.0001     evaluation reward: 2.02
episode: 946   score: 3.0   memory length: 173975   epsilon: 0.8535275200031798
steps: 229    lr: 0.0001     evaluation reward: 2.03
```

```
episode: 947   score: 2.0   memory length: 174192   epsilon: 0.8530978600031891
steps: 217    lr: 0.0001     evaluation reward: 2.04
episode: 948   score: 0.0   memory length: 174315   epsilon: 0.8528543200031944
steps: 123    lr: 0.0001     evaluation reward: 2.01
episode: 949   score: 2.0   memory length: 174515   epsilon: 0.852458320003203
steps: 200    lr: 0.0001     evaluation reward: 2.02
episode: 950   score: 1.0   memory length: 174684   epsilon: 0.8521237000032102
steps: 169    lr: 0.0001     evaluation reward: 2.02
episode: 951   score: 0.0   memory length: 174807   epsilon: 0.8518801600032155
steps: 123    lr: 0.0001     evaluation reward: 2.0
episode: 952   score: 2.0   memory length: 175004   epsilon: 0.851490100003224
steps: 197    lr: 0.0001     evaluation reward: 2.02
episode: 953   score: 2.0   memory length: 175221   epsilon: 0.8510604400032333
steps: 217    lr: 0.0001     evaluation reward: 2.04
episode: 954   score: 2.0   memory length: 175438   epsilon: 0.8506307800032427
steps: 217    lr: 0.0001     evaluation reward: 2.03
episode: 955   score: 1.0   memory length: 175588   epsilon: 0.8503337800032491
steps: 150    lr: 0.0001     evaluation reward: 2.04
episode: 956   score: 1.0   memory length: 175757   epsilon: 0.8499991600032564
steps: 169    lr: 0.0001     evaluation reward: 2.0
episode: 957   score: 3.0   memory length: 175983   epsilon: 0.8495516800032661
steps: 226    lr: 0.0001     evaluation reward: 2.03
episode: 958   score: 4.0   memory length: 176283   epsilon: 0.848957680003279
steps: 300    lr: 0.0001     evaluation reward: 2.06
episode: 959   score: 4.0   memory length: 176581   epsilon: 0.8483676400032918
steps: 298    lr: 0.0001     evaluation reward: 2.09
episode: 960   score: 1.0   memory length: 176749   epsilon: 0.848035000003299
steps: 168    lr: 0.0001     evaluation reward: 2.06
episode: 961   score: 3.0   memory length: 176993   epsilon: 0.8475518800033095
steps: 244    lr: 0.0001     evaluation reward: 2.05
episode: 962   score: 0.0   memory length: 177116   epsilon: 0.8473083400033148
steps: 123    lr: 0.0001     evaluation reward: 2.04
episode: 963   score: 2.0   memory length: 177313   epsilon: 0.8469182800033233
steps: 197    lr: 0.0001     evaluation reward: 2.04
episode: 964   score: 2.0   memory length: 177514   epsilon: 0.8465203000033319
steps: 201    lr: 0.0001     evaluation reward: 2.05
episode: 965   score: 1.0   memory length: 177682   epsilon: 0.8461876600033391
steps: 168    lr: 0.0001     evaluation reward: 2.05
episode: 966   score: 1.0   memory length: 177851   epsilon: 0.8458530400033464
steps: 169    lr: 0.0001     evaluation reward: 2.03
episode: 967   score: 4.0   memory length: 178107   epsilon: 0.8453461600033574
steps: 256    lr: 0.0001     evaluation reward: 2.05
episode: 968   score: 0.0   memory length: 178230   epsilon: 0.8451026200033627
steps: 123    lr: 0.0001     evaluation reward: 2.02
episode: 969   score: 3.0   memory length: 178455   epsilon: 0.8446571200033723
steps: 225    lr: 0.0001     evaluation reward: 2.02
episode: 970   score: 1.0   memory length: 178623   epsilon: 0.8443244800033796
steps: 168    lr: 0.0001     evaluation reward: 2.03
```

```
episode: 971   score: 4.0   memory length: 178897   epsilon: 0.8437819600033913
steps: 274    lr: 0.0001     evaluation reward: 2.06
episode: 972   score: 3.0   memory length: 179145   epsilon: 0.843290920003402
steps: 248    lr: 0.0001     evaluation reward: 2.07
episode: 973   score: 3.0   memory length: 179373   epsilon: 0.8428394800034118
steps: 228    lr: 0.0001     evaluation reward: 2.09
episode: 974   score: 0.0   memory length: 179496   epsilon: 0.8425959400034171
steps: 123    lr: 0.0001     evaluation reward: 2.09
episode: 975   score: 4.0   memory length: 179775   epsilon: 0.8420435200034291
steps: 279    lr: 0.0001     evaluation reward: 2.11
episode: 976   score: 4.0   memory length: 180089   epsilon: 0.8414218000034426
steps: 314    lr: 0.0001     evaluation reward: 2.11
episode: 977   score: 2.0   memory length: 180286   epsilon: 0.841031740003451
steps: 197    lr: 0.0001     evaluation reward: 2.13
episode: 978   score: 3.0   memory length: 180533   epsilon: 0.8405426800034617
steps: 247    lr: 0.0001     evaluation reward: 2.11
episode: 979   score: 2.0   memory length: 180734   epsilon: 0.8401447000034703
steps: 201    lr: 0.0001     evaluation reward: 2.11
episode: 980   score: 0.0   memory length: 180856   epsilon: 0.8399031400034755
steps: 122    lr: 0.0001     evaluation reward: 2.07
episode: 981   score: 4.0   memory length: 181134   epsilon: 0.8393527000034875
steps: 278    lr: 0.0001     evaluation reward: 2.1
episode: 982   score: 0.0   memory length: 181256   epsilon: 0.8391111400034927
steps: 122    lr: 0.0001     evaluation reward: 2.08
episode: 983   score: 2.0   memory length: 181454   epsilon: 0.8387191000035013
steps: 198    lr: 0.0001     evaluation reward: 2.1
episode: 984   score: 0.0   memory length: 181577   epsilon: 0.8384755600035065
steps: 123    lr: 0.0001     evaluation reward: 2.08
episode: 985   score: 1.0   memory length: 181728   epsilon: 0.838176580003513
steps: 151    lr: 0.0001     evaluation reward: 2.08
episode: 986   score: 1.0   memory length: 181879   epsilon: 0.8378776000035195
steps: 151    lr: 0.0001     evaluation reward: 2.06
episode: 987   score: 3.0   memory length: 182147   epsilon: 0.837346960003531
steps: 268    lr: 0.0001     evaluation reward: 2.06
episode: 988   score: 3.0   memory length: 182375   epsilon: 0.8368955200035408
steps: 228    lr: 0.0001     evaluation reward: 2.05
episode: 989   score: 1.0   memory length: 182544   epsilon: 0.8365609000035481
steps: 169    lr: 0.0001     evaluation reward: 2.04
episode: 990   score: 3.0   memory length: 182790   epsilon: 0.8360738200035587
steps: 246    lr: 0.0001     evaluation reward: 2.03
episode: 991   score: 3.0   memory length: 183016   epsilon: 0.8356263400035684
steps: 226    lr: 0.0001     evaluation reward: 2.05
episode: 992   score: 1.0   memory length: 183185   epsilon: 0.8352917200035757
steps: 169    lr: 0.0001     evaluation reward: 2.04
episode: 993   score: 1.0   memory length: 183336   epsilon: 0.8349927400035821
steps: 151    lr: 0.0001     evaluation reward: 2.04
episode: 994   score: 3.0   memory length: 183565   epsilon: 0.834539320003592
steps: 229    lr: 0.0001     evaluation reward: 2.05
```

```
episode: 995   score: 0.0   memory length: 183688   epsilon: 0.8342957800035973
steps: 123    lr: 0.0001     evaluation reward: 2.05
episode: 996   score: 3.0   memory length: 183918   epsilon: 0.8338403800036072
steps: 230    lr: 0.0001     evaluation reward: 2.07
episode: 997   score: 2.0   memory length: 184139   epsilon: 0.8334028000036167
steps: 221    lr: 0.0001     evaluation reward: 2.07
episode: 998   score: 4.0   memory length: 184414   epsilon: 0.8328583000036285
steps: 275    lr: 0.0001     evaluation reward: 2.08
episode: 999   score: 1.0   memory length: 184564   epsilon: 0.8325613000036349
steps: 150    lr: 0.0001     evaluation reward: 2.05
episode: 1000   score: 4.0   memory length: 184822   epsilon: 0.832050460003646
steps: 258    lr: 0.0001     evaluation reward: 2.09
episode: 1001   score: 2.0   memory length: 185040   epsilon: 0.8316188200036554
steps: 218    lr: 0.0001     evaluation reward: 2.1
episode: 1002   score: 2.0   memory length: 185240   epsilon: 0.831222820003664
steps: 200    lr: 0.0001     evaluation reward: 2.12
episode: 1003   score: 2.0   memory length: 185458   epsilon: 0.8307911800036734
steps: 218    lr: 0.0001     evaluation reward: 2.11
episode: 1004   score: 3.0   memory length: 185687   epsilon: 0.8303377600036832
steps: 229    lr: 0.0001     evaluation reward: 2.13
episode: 1005   score: 1.0   memory length: 185838   epsilon: 0.8300387800036897
steps: 151    lr: 0.0001     evaluation reward: 2.14
episode: 1006   score: 3.0   memory length: 186064   epsilon: 0.8295913000036994
steps: 226    lr: 0.0001     evaluation reward: 2.12
episode: 1007   score: 0.0   memory length: 186187   epsilon: 0.8293477600037047
steps: 123    lr: 0.0001     evaluation reward: 2.09
episode: 1008   score: 2.0   memory length: 186385   epsilon: 0.8289557200037132
steps: 198    lr: 0.0001     evaluation reward: 2.1
episode: 1009   score: 4.0   memory length: 186702   epsilon: 0.8283280600037268
steps: 317    lr: 0.0001     evaluation reward: 2.12
episode: 1010   score: 2.0   memory length: 186899   epsilon: 0.8279380000037353
steps: 197    lr: 0.0001     evaluation reward: 2.1
episode: 1011   score: 1.0   memory length: 187071   epsilon: 0.8275974400037427
steps: 172    lr: 0.0001     evaluation reward: 2.08
episode: 1012   score: 2.0   memory length: 187288   epsilon: 0.827167780003752
steps: 217    lr: 0.0001     evaluation reward: 2.1
episode: 1013   score: 2.0   memory length: 187485   epsilon: 0.8267777200037605
steps: 197    lr: 0.0001     evaluation reward: 2.1
episode: 1014   score: 2.0   memory length: 187683   epsilon: 0.826385680003769
steps: 198    lr: 0.0001     evaluation reward: 2.1
episode: 1015   score: 2.0   memory length: 187880   epsilon: 0.8259956200037775
steps: 197    lr: 0.0001     evaluation reward: 2.09
episode: 1016   score: 3.0   memory length: 188106   epsilon: 0.8255481400037872
steps: 226    lr: 0.0001     evaluation reward: 2.11
episode: 1017   score: 0.0   memory length: 188228   epsilon: 0.8253065800037924
steps: 122    lr: 0.0001     evaluation reward: 2.09
episode: 1018   score: 0.0   memory length: 188351   epsilon: 0.8250630400037977
steps: 123    lr: 0.0001     evaluation reward: 2.05
```

```
episode: 1019   score: 1.0   memory length: 188501   epsilon: 0.8247660400038042
steps: 150    lr: 0.0001     evaluation reward: 2.03
episode: 1020   score: 2.0   memory length: 188717   epsilon: 0.8243383600038134
steps: 216    lr: 0.0001     evaluation reward: 1.99
episode: 1021   score: 4.0   memory length: 189012   epsilon: 0.8237542600038261
steps: 295    lr: 0.0001     evaluation reward: 2.0
episode: 1022   score: 0.0   memory length: 189135   epsilon: 0.8235107200038314
steps: 123    lr: 0.0001     evaluation reward: 1.99
episode: 1023   score: 1.0   memory length: 189305   epsilon: 0.8231741200038387
steps: 170    lr: 0.0001     evaluation reward: 1.99
episode: 1024   score: 2.0   memory length: 189522   epsilon: 0.822744460003848
steps: 217    lr: 0.0001     evaluation reward: 2.01
episode: 1025   score: 3.0   memory length: 189750   epsilon: 0.8222930200038578
steps: 228    lr: 0.0001     evaluation reward: 2.04
episode: 1026   score: 2.0   memory length: 189948   epsilon: 0.8219009800038664
steps: 198    lr: 0.0001     evaluation reward: 2.04
episode: 1027   score: 3.0   memory length: 190159   epsilon: 0.8214832000038754
steps: 211    lr: 0.0001     evaluation reward: 2.05
episode: 1028   score: 3.0   memory length: 190369   epsilon: 0.8210674000038845
steps: 210    lr: 0.0001     evaluation reward: 2.05
episode: 1029   score: 1.0   memory length: 190538   epsilon: 0.8207327800038917
steps: 169    lr: 0.0001     evaluation reward: 2.05
episode: 1030   score: 0.0   memory length: 190660   epsilon: 0.820491220003897
steps: 122    lr: 0.0001     evaluation reward: 2.0
episode: 1031   score: 2.0   memory length: 190860   epsilon: 0.8200952200039056
steps: 200    lr: 0.0001     evaluation reward: 2.0
episode: 1032   score: 0.0   memory length: 190982   epsilon: 0.8198536600039108
steps: 122    lr: 0.0001     evaluation reward: 1.95
episode: 1033   score: 0.0   memory length: 191104   epsilon: 0.819612100003916
steps: 122    lr: 0.0001     evaluation reward: 1.93
episode: 1034   score: 1.0   memory length: 191255   epsilon: 0.8193131200039225
steps: 151    lr: 0.0001     evaluation reward: 1.93
episode: 1035   score: 1.0   memory length: 191423   epsilon: 0.8189804800039298
steps: 168    lr: 0.0001     evaluation reward: 1.91
episode: 1036   score: 4.0   memory length: 191721   epsilon: 0.8183904400039426
steps: 298    lr: 0.0001     evaluation reward: 1.92
episode: 1037   score: 3.0   memory length: 191952   epsilon: 0.8179330600039525
steps: 231    lr: 0.0001     evaluation reward: 1.91
episode: 1038   score: 3.0   memory length: 192217   epsilon: 0.8174083600039639
steps: 265    lr: 0.0001     evaluation reward: 1.94
episode: 1039   score: 3.0   memory length: 192443   epsilon: 0.8169608800039736
steps: 226    lr: 0.0001     evaluation reward: 1.95
episode: 1040   score: 1.0   memory length: 192614   epsilon: 0.816622300003981
steps: 171    lr: 0.0001     evaluation reward: 1.93
episode: 1041   score: 0.0   memory length: 192736   epsilon: 0.8163807400039862
steps: 122    lr: 0.0001     evaluation reward: 1.91
episode: 1042   score: 2.0   memory length: 192954   epsilon: 0.8159491000039956
steps: 218    lr: 0.0001     evaluation reward: 1.92
```

```
episode: 1043   score: 2.0   memory length: 193154   epsilon: 0.8155531000040042
steps: 200    lr: 0.0001     evaluation reward: 1.92
episode: 1044   score: 0.0   memory length: 193277   epsilon: 0.8153095600040094
steps: 123    lr: 0.0001     evaluation reward: 1.9
episode: 1045   score: 8.0   memory length: 193628   epsilon: 0.8146145800040245
steps: 351    lr: 0.0001     evaluation reward: 1.95
episode: 1046   score: 7.0   memory length: 194022   epsilon: 0.8138344600040415
steps: 394    lr: 0.0001     evaluation reward: 1.99
episode: 1047   score: 2.0   memory length: 194220   epsilon: 0.81344242000405
steps: 198    lr: 0.0001     evaluation reward: 1.99
episode: 1048   score: 4.0   memory length: 194497   epsilon: 0.8128939600040619
steps: 277    lr: 0.0001     evaluation reward: 2.03
episode: 1049   score: 1.0   memory length: 194648   epsilon: 0.8125949800040684
steps: 151    lr: 0.0001     evaluation reward: 2.02
episode: 1050   score: 1.0   memory length: 194798   epsilon: 0.8122979800040748
steps: 150    lr: 0.0001     evaluation reward: 2.02
episode: 1051   score: 0.0   memory length: 194921   epsilon: 0.8120544400040801
steps: 123    lr: 0.0001     evaluation reward: 2.02
episode: 1052   score: 3.0   memory length: 195149   epsilon: 0.8116030000040899
steps: 228    lr: 0.0001     evaluation reward: 2.03
episode: 1053   score: 1.0   memory length: 195299   epsilon: 0.8113060000040964
steps: 150    lr: 0.0001     evaluation reward: 2.02
episode: 1054   score: 0.0   memory length: 195422   epsilon: 0.8110624600041016
steps: 123    lr: 0.0001     evaluation reward: 2.0
episode: 1055   score: 4.0   memory length: 195703   epsilon: 0.8105060800041137
steps: 281    lr: 0.0001     evaluation reward: 2.03
episode: 1056   score: 2.0   memory length: 195904   epsilon: 0.8101081000041224
steps: 201    lr: 0.0001     evaluation reward: 2.04
episode: 1057   score: 4.0   memory length: 196225   epsilon: 0.8094725200041362
steps: 321    lr: 0.0001     evaluation reward: 2.05
episode: 1058   score: 1.0   memory length: 196394   epsilon: 0.8091379000041434
steps: 169    lr: 0.0001     evaluation reward: 2.02
episode: 1059   score: 1.0   memory length: 196545   epsilon: 0.8088389200041499
steps: 151    lr: 0.0001     evaluation reward: 1.99
episode: 1060   score: 1.0   memory length: 196717   epsilon: 0.8084983600041573
steps: 172    lr: 0.0001     evaluation reward: 1.99
episode: 1061   score: 3.0   memory length: 196964   epsilon: 0.8080093000041679
steps: 247    lr: 0.0001     evaluation reward: 1.99
episode: 1062   score: 0.0   memory length: 197087   epsilon: 0.8077657600041732
steps: 123    lr: 0.0001     evaluation reward: 1.99
episode: 1063   score: 1.0   memory length: 197258   epsilon: 0.8074271800041806
steps: 171    lr: 0.0001     evaluation reward: 1.98
episode: 1064   score: 1.0   memory length: 197427   epsilon: 0.8070925600041878
steps: 169    lr: 0.0001     evaluation reward: 1.97
episode: 1065   score: 3.0   memory length: 197652   epsilon: 0.8066470600041975
steps: 225    lr: 0.0001     evaluation reward: 1.99
episode: 1066   score: 1.0   memory length: 197824   epsilon: 0.8063065000042049
steps: 172    lr: 0.0001     evaluation reward: 1.99
```

```
episode: 1067   score: 1.0   memory length: 197974   epsilon: 0.8060095000042113
steps: 150    lr: 0.0001     evaluation reward: 1.96
episode: 1068   score: 2.0   memory length: 198172   epsilon: 0.8056174600042199
steps: 198    lr: 0.0001     evaluation reward: 1.98
episode: 1069   score: 4.0   memory length: 198472   epsilon: 0.8050234600042327
steps: 300    lr: 0.0001     evaluation reward: 1.99
episode: 1070   score: 4.0   memory length: 198766   epsilon: 0.8044413400042454
steps: 294    lr: 0.0001     evaluation reward: 2.02
episode: 1071   score: 0.0   memory length: 198889   epsilon: 0.8041978000042507
steps: 123    lr: 0.0001     evaluation reward: 1.98
episode: 1072   score: 1.0   memory length: 199059   epsilon: 0.803861200004258
steps: 170    lr: 0.0001     evaluation reward: 1.96
episode: 1073   score: 0.0   memory length: 199182   epsilon: 0.8036176600042633
steps: 123    lr: 0.0001     evaluation reward: 1.93
episode: 1074   score: 4.0   memory length: 199440   epsilon: 0.8031068200042744
steps: 258    lr: 0.0001     evaluation reward: 1.97
episode: 1075   score: 2.0   memory length: 199657   epsilon: 0.8026771600042837
steps: 217    lr: 0.0001     evaluation reward: 1.95
episode: 1076   score: 2.0   memory length: 199855   epsilon: 0.8022851200042922
steps: 198    lr: 0.0001     evaluation reward: 1.93
episode: 1077   score: 4.0   memory length: 200151   epsilon: 0.8016990400043049
steps: 296    lr: 4e-05     evaluation reward: 1.95
episode: 1078   score: 5.0   memory length: 200475   epsilon: 0.8010575200043188
steps: 324    lr: 4e-05     evaluation reward: 1.97
episode: 1079   score: 3.0   memory length: 200742   epsilon: 0.8005288600043303
steps: 267    lr: 4e-05     evaluation reward: 1.98
episode: 1080   score: 3.0   memory length: 200952   epsilon: 0.8001130600043393
steps: 210    lr: 4e-05     evaluation reward: 2.01
episode: 1081   score: 2.0   memory length: 201150   epsilon: 0.7997210200043479
steps: 198    lr: 4e-05     evaluation reward: 1.99
episode: 1082   score: 1.0   memory length: 201320   epsilon: 0.7993844200043552
steps: 170    lr: 4e-05     evaluation reward: 2.0
episode: 1083   score: 0.0   memory length: 201443   epsilon: 0.7991408800043605
steps: 123    lr: 4e-05     evaluation reward: 1.98
episode: 1084   score: 6.0   memory length: 201813   epsilon: 0.7984082800043764
steps: 370    lr: 4e-05     evaluation reward: 2.04
episode: 1085   score: 5.0   memory length: 202139   epsilon: 0.7977628000043904
steps: 326    lr: 4e-05     evaluation reward: 2.08
episode: 1086   score: 4.0   memory length: 202433   epsilon: 0.797180680004403
steps: 294    lr: 4e-05     evaluation reward: 2.11
episode: 1087   score: 1.0   memory length: 202583   epsilon: 0.7968836800044095
steps: 150    lr: 4e-05     evaluation reward: 2.09
episode: 1088   score: 1.0   memory length: 202753   epsilon: 0.7965470800044168
steps: 170    lr: 4e-05     evaluation reward: 2.07
episode: 1089   score: 4.0   memory length: 203065   epsilon: 0.7959293200044302
steps: 312    lr: 4e-05     evaluation reward: 2.1
episode: 1090   score: 2.0   memory length: 203264   epsilon: 0.7955353000044387
steps: 199    lr: 4e-05     evaluation reward: 2.09
```

```
episode: 1091   score: 1.0   memory length: 203414   epsilon: 0.7952383000044452
steps: 150    lr: 4e-05     evaluation reward: 2.07
episode: 1092   score: 3.0   memory length: 203647   epsilon: 0.7947769600044552
steps: 233    lr: 4e-05     evaluation reward: 2.09
episode: 1093   score: 2.0   memory length: 203866   epsilon: 0.7943433400044646
steps: 219    lr: 4e-05     evaluation reward: 2.1
episode: 1094   score: 5.0   memory length: 204173   epsilon: 0.7937354800044778
steps: 307    lr: 4e-05     evaluation reward: 2.12
episode: 1095   score: 7.0   memory length: 204570   epsilon: 0.7929494200044949
steps: 397    lr: 4e-05     evaluation reward: 2.19
episode: 1096   score: 6.0   memory length: 204947   epsilon: 0.7922029600045111
steps: 377    lr: 4e-05     evaluation reward: 2.22
episode: 1097   score: 8.0   memory length: 205405   epsilon: 0.7912961200045308
steps: 458    lr: 4e-05     evaluation reward: 2.28
episode: 1098   score: 2.0   memory length: 205602   epsilon: 0.7909060600045392
steps: 197    lr: 4e-05     evaluation reward: 2.26
episode: 1099   score: 2.0   memory length: 205821   epsilon: 0.7904724400045486
steps: 219    lr: 4e-05     evaluation reward: 2.27
episode: 1100   score: 1.0   memory length: 205992   epsilon: 0.790133860004556
steps: 171    lr: 4e-05     evaluation reward: 2.24
episode: 1101   score: 4.0   memory length: 206284   epsilon: 0.7895557000045685
steps: 292    lr: 4e-05     evaluation reward: 2.26
episode: 1102   score: 2.0   memory length: 206482   epsilon: 0.789163660004577
steps: 198    lr: 4e-05     evaluation reward: 2.26
episode: 1103   score: 3.0   memory length: 206708   epsilon: 0.7887161800045868
steps: 226    lr: 4e-05     evaluation reward: 2.27
episode: 1104   score: 4.0   memory length: 207003   epsilon: 0.7881320800045994
steps: 295    lr: 4e-05     evaluation reward: 2.28
episode: 1105   score: 4.0   memory length: 207247   epsilon: 0.7876489600046099
steps: 244    lr: 4e-05     evaluation reward: 2.31
episode: 1106   score: 10.0   memory length: 207679   epsilon:
0.7867936000046285    steps: 432    lr: 4e-05     evaluation reward: 2.38
episode: 1107   score: 1.0   memory length: 207830   epsilon: 0.786494620004635
steps: 151    lr: 4e-05     evaluation reward: 2.39
episode: 1108   score: 4.0   memory length: 208088   epsilon: 0.7859837800046461
steps: 258    lr: 4e-05     evaluation reward: 2.41
episode: 1109   score: 3.0   memory length: 208355   epsilon: 0.7854551200046576
steps: 267    lr: 4e-05     evaluation reward: 2.4
episode: 1110   score: 1.0   memory length: 208523   epsilon: 0.7851224800046648
steps: 168    lr: 4e-05     evaluation reward: 2.39
episode: 1111   score: 1.0   memory length: 208694   epsilon: 0.7847839000046721
steps: 171    lr: 4e-05     evaluation reward: 2.39
episode: 1112   score: 1.0   memory length: 208866   epsilon: 0.7844433400046795
steps: 172    lr: 4e-05     evaluation reward: 2.38
episode: 1113   score: 2.0   memory length: 209048   epsilon: 0.7840829800046873
steps: 182    lr: 4e-05     evaluation reward: 2.38
episode: 1114   score: 2.0   memory length: 209248   epsilon: 0.7836869800046959
steps: 200    lr: 4e-05     evaluation reward: 2.38
```

```
episode: 1115   score: 4.0   memory length: 209564   epsilon: 0.7830613000047095
steps: 316    lr: 4e-05     evaluation reward: 2.4
episode: 1116   score: 3.0   memory length: 209807   epsilon: 0.78258016000472
steps: 243    lr: 4e-05     evaluation reward: 2.4
episode: 1117   score: 4.0   memory length: 210086   epsilon: 0.782027740004732
steps: 279    lr: 4e-05     evaluation reward: 2.44
episode: 1118   score: 3.0   memory length: 210333   epsilon: 0.7815386800047426
steps: 247    lr: 4e-05     evaluation reward: 2.47
episode: 1119   score: 3.0   memory length: 210577   epsilon: 0.7810555600047531
steps: 244    lr: 4e-05     evaluation reward: 2.49
episode: 1120   score: 1.0   memory length: 210728   epsilon: 0.7807565800047596
steps: 151    lr: 4e-05     evaluation reward: 2.48
episode: 1121   score: 4.0   memory length: 211024   epsilon: 0.7801705000047723
steps: 296    lr: 4e-05     evaluation reward: 2.48
episode: 1122   score: 2.0   memory length: 211243   epsilon: 0.7797368800047817
steps: 219    lr: 4e-05     evaluation reward: 2.5
episode: 1123   score: 2.0   memory length: 211461   epsilon: 0.7793052400047911
steps: 218    lr: 4e-05     evaluation reward: 2.51
episode: 1124   score: 2.0   memory length: 211643   epsilon: 0.7789448800047989
steps: 182    lr: 4e-05     evaluation reward: 2.51
episode: 1125   score: 5.0   memory length: 211929   epsilon: 0.7783786000048112
steps: 286    lr: 4e-05     evaluation reward: 2.53
episode: 1126   score: 3.0   memory length: 212175   epsilon: 0.7778915200048218
steps: 246    lr: 4e-05     evaluation reward: 2.54
episode: 1127   score: 2.0   memory length: 212373   epsilon: 0.7774994800048303
steps: 198    lr: 4e-05     evaluation reward: 2.53
episode: 1128   score: 3.0   memory length: 212598   epsilon: 0.7770539800048399
steps: 225    lr: 4e-05     evaluation reward: 2.53
episode: 1129   score: 4.0   memory length: 212875   epsilon: 0.7765055200048518
steps: 277    lr: 4e-05     evaluation reward: 2.56
episode: 1130   score: 2.0   memory length: 213073   epsilon: 0.7761134800048604
steps: 198    lr: 4e-05     evaluation reward: 2.58
episode: 1131   score: 3.0   memory length: 213321   epsilon: 0.775622440004871
steps: 248    lr: 4e-05     evaluation reward: 2.59
episode: 1132   score: 1.0   memory length: 213471   epsilon: 0.7753254400048775
steps: 150    lr: 4e-05     evaluation reward: 2.6
episode: 1133   score: 2.0   memory length: 213689   epsilon: 0.7748938000048868
steps: 218    lr: 4e-05     evaluation reward: 2.62
episode: 1134   score: 1.0   memory length: 213859   epsilon: 0.7745572000048941
steps: 170    lr: 4e-05     evaluation reward: 2.62
episode: 1135   score: 5.0   memory length: 214201   epsilon: 0.7738800400049088
steps: 342    lr: 4e-05     evaluation reward: 2.66
episode: 1136   score: 1.0   memory length: 214371   epsilon: 0.7735434400049161
steps: 170    lr: 4e-05     evaluation reward: 2.63
episode: 1137   score: 3.0   memory length: 214618   epsilon: 0.7730543800049268
steps: 247    lr: 4e-05     evaluation reward: 2.63
episode: 1138   score: 2.0   memory length: 214815   epsilon: 0.7726643200049352
steps: 197    lr: 4e-05     evaluation reward: 2.62
```

```
episode: 1139   score: 3.0   memory length: 215063   epsilon: 0.7721732800049459
steps: 248    lr: 4e-05     evaluation reward: 2.62
episode: 1140   score: 2.0   memory length: 215281   epsilon: 0.7717416400049553
steps: 218    lr: 4e-05     evaluation reward: 2.63
episode: 1141   score: 5.0   memory length: 215605   epsilon: 0.7711001200049692
steps: 324    lr: 4e-05     evaluation reward: 2.68
episode: 1142   score: 4.0   memory length: 215873   epsilon: 0.7705694800049807
steps: 268    lr: 4e-05     evaluation reward: 2.7
episode: 1143   score: 4.0   memory length: 216169   epsilon: 0.7699834000049934
steps: 296    lr: 4e-05     evaluation reward: 2.72
episode: 1144   score: 4.0   memory length: 216410   epsilon: 0.7695062200050038
steps: 241    lr: 4e-05     evaluation reward: 2.76
episode: 1145   score: 5.0   memory length: 216737   epsilon: 0.7688587600050178
steps: 327    lr: 4e-05     evaluation reward: 2.73
episode: 1146   score: 3.0   memory length: 216965   epsilon: 0.7684073200050276
steps: 228    lr: 4e-05     evaluation reward: 2.69
episode: 1147   score: 1.0   memory length: 217133   epsilon: 0.7680746800050349
steps: 168    lr: 4e-05     evaluation reward: 2.68
episode: 1148   score: 1.0   memory length: 217284   epsilon: 0.7677757000050414
steps: 151    lr: 4e-05     evaluation reward: 2.65
episode: 1149   score: 2.0   memory length: 217482   epsilon: 0.7673836600050499
steps: 198    lr: 4e-05     evaluation reward: 2.66
episode: 1150   score: 3.0   memory length: 217727   epsilon: 0.7668985600050604
steps: 245    lr: 4e-05     evaluation reward: 2.68
episode: 1151   score: 2.0   memory length: 217907   epsilon: 0.7665421600050681
steps: 180    lr: 4e-05     evaluation reward: 2.7
episode: 1152   score: 4.0   memory length: 218218   epsilon: 0.7659263800050815
steps: 311    lr: 4e-05     evaluation reward: 2.71
episode: 1153   score: 1.0   memory length: 218387   epsilon: 0.7655917600050888
steps: 169    lr: 4e-05     evaluation reward: 2.71
episode: 1154   score: 2.0   memory length: 218585   epsilon: 0.7651997200050973
steps: 198    lr: 4e-05     evaluation reward: 2.73
episode: 1155   score: 5.0   memory length: 218890   epsilon: 0.7645958200051104
steps: 305    lr: 4e-05     evaluation reward: 2.74
episode: 1156   score: 2.0   memory length: 219087   epsilon: 0.7642057600051189
steps: 197    lr: 4e-05     evaluation reward: 2.74
episode: 1157   score: 3.0   memory length: 219312   epsilon: 0.7637602600051285
steps: 225    lr: 4e-05     evaluation reward: 2.73
episode: 1158   score: 4.0   memory length: 219608   epsilon: 0.7631741800051413
steps: 296    lr: 4e-05     evaluation reward: 2.76
episode: 1159   score: 2.0   memory length: 219805   epsilon: 0.7627841200051497
steps: 197    lr: 4e-05     evaluation reward: 2.77
episode: 1160   score: 2.0   memory length: 220021   epsilon: 0.762356440005159
steps: 216    lr: 4e-05     evaluation reward: 2.78
episode: 1161   score: 2.0   memory length: 220221   epsilon: 0.7619604400051676
steps: 200    lr: 4e-05     evaluation reward: 2.77
episode: 1162   score: 4.0   memory length: 220485   epsilon: 0.761437720005179
steps: 264    lr: 4e-05     evaluation reward: 2.81
```

```
episode: 1163   score: 3.0   memory length: 220693   epsilon: 0.7610258800051879
steps: 208    lr: 4e-05     evaluation reward: 2.83
episode: 1164   score: 4.0   memory length: 220968   epsilon: 0.7604813800051997
steps: 275    lr: 4e-05     evaluation reward: 2.86
episode: 1165   score: 4.0   memory length: 221240   epsilon: 0.7599428200052114
steps: 272    lr: 4e-05     evaluation reward: 2.87
episode: 1166   score: 3.0   memory length: 221466   epsilon: 0.7594953400052211
steps: 226    lr: 4e-05     evaluation reward: 2.89
episode: 1167   score: 4.0   memory length: 221721   epsilon: 0.7589904400052321
steps: 255    lr: 4e-05     evaluation reward: 2.92
episode: 1168   score: 4.0   memory length: 222017   epsilon: 0.7584043600052448
steps: 296    lr: 4e-05     evaluation reward: 2.94
episode: 1169   score: 7.0   memory length: 222368   epsilon: 0.7577093800052599
steps: 351    lr: 4e-05     evaluation reward: 2.97
episode: 1170   score: 5.0   memory length: 222691   epsilon: 0.7570698400052738
steps: 323    lr: 4e-05     evaluation reward: 2.98
episode: 1171   score: 5.0   memory length: 223016   epsilon: 0.7564263400052877
steps: 325    lr: 4e-05     evaluation reward: 3.03
episode: 1172   score: 3.0   memory length: 223282   epsilon: 0.7558996600052992
steps: 266    lr: 4e-05     evaluation reward: 3.05
episode: 1173   score: 10.0   memory length: 223838   epsilon:
0.7547987800053231    steps: 556    lr: 4e-05     evaluation reward: 3.15
episode: 1174   score: 5.0   memory length: 224160   epsilon: 0.7541612200053369
steps: 322    lr: 4e-05     evaluation reward: 3.16
episode: 1175   score: 6.0   memory length: 224550   epsilon: 0.7533890200053537
steps: 390    lr: 4e-05     evaluation reward: 3.2
episode: 1176   score: 2.0   memory length: 224751   epsilon: 0.7529910400053623
steps: 201    lr: 4e-05     evaluation reward: 3.2
episode: 1177   score: 3.0   memory length: 224976   epsilon: 0.752545540005372
steps: 225    lr: 4e-05     evaluation reward: 3.19
episode: 1178   score: 2.0   memory length: 225199   epsilon: 0.7521040000053816
steps: 223    lr: 4e-05     evaluation reward: 3.16
episode: 1179   score: 3.0   memory length: 225429   epsilon: 0.7516486000053915
steps: 230    lr: 4e-05     evaluation reward: 3.16
episode: 1180   score: 2.0   memory length: 225611   epsilon: 0.7512882400053993
steps: 182    lr: 4e-05     evaluation reward: 3.15
episode: 1181   score: 3.0   memory length: 225854   epsilon: 0.7508071000054097
steps: 243    lr: 4e-05     evaluation reward: 3.16
episode: 1182   score: 2.0   memory length: 226052   epsilon: 0.7504150600054182
steps: 198    lr: 4e-05     evaluation reward: 3.17
episode: 1183   score: 5.0   memory length: 226368   epsilon: 0.7497893800054318
steps: 316    lr: 4e-05     evaluation reward: 3.22
episode: 1184   score: 5.0   memory length: 226690   epsilon: 0.7491518200054457
steps: 322    lr: 4e-05     evaluation reward: 3.21
episode: 1185   score: 3.0   memory length: 226919   epsilon: 0.7486984000054555
steps: 229    lr: 4e-05     evaluation reward: 3.19
episode: 1186   score: 2.0   memory length: 227117   epsilon: 0.748306360005464
steps: 198    lr: 4e-05     evaluation reward: 3.17
```

```
episode: 1187   score: 5.0   memory length: 227425   epsilon: 0.7476965200054773
steps: 308    lr: 4e-05     evaluation reward: 3.21
episode: 1188   score: 1.0   memory length: 227594   epsilon: 0.7473619000054845
steps: 169    lr: 4e-05     evaluation reward: 3.21
episode: 1189   score: 2.0   memory length: 227774   epsilon: 0.7470055000054923
steps: 180    lr: 4e-05     evaluation reward: 3.19
episode: 1190   score: 0.0   memory length: 227897   epsilon: 0.7467619600054975
steps: 123    lr: 4e-05     evaluation reward: 3.17
episode: 1191   score: 3.0   memory length: 228122   epsilon: 0.7463164600055072
steps: 225    lr: 4e-05     evaluation reward: 3.19
episode: 1192   score: 1.0   memory length: 228293   epsilon: 0.7459778800055146
steps: 171    lr: 4e-05     evaluation reward: 3.17
episode: 1193   score: 3.0   memory length: 228537   epsilon: 0.7454947600055251
steps: 244    lr: 4e-05     evaluation reward: 3.18
episode: 1194   score: 4.0   memory length: 228814   epsilon: 0.744946300005537
steps: 277    lr: 4e-05     evaluation reward: 3.17
episode: 1195   score: 1.0   memory length: 228983   epsilon: 0.7446116800055442
steps: 169    lr: 4e-05     evaluation reward: 3.11
episode: 1196   score: 1.0   memory length: 229134   epsilon: 0.7443127000055507
steps: 151    lr: 4e-05     evaluation reward: 3.06
episode: 1197   score: 2.0   memory length: 229331   epsilon: 0.7439226400055592
steps: 197    lr: 4e-05     evaluation reward: 3.0
episode: 1198   score: 8.0   memory length: 229780   epsilon: 0.7430336200055785
steps: 449    lr: 4e-05     evaluation reward: 3.06
episode: 1199   score: 2.0   memory length: 229999   epsilon: 0.7426000000055879
steps: 219    lr: 4e-05     evaluation reward: 3.06
episode: 1200   score: 3.0   memory length: 230245   epsilon: 0.7421129200055985
steps: 246    lr: 4e-05     evaluation reward: 3.08
episode: 1201   score: 5.0   memory length: 230539   epsilon: 0.7415308000056111
steps: 294    lr: 4e-05     evaluation reward: 3.09
episode: 1202   score: 4.0   memory length: 230831   epsilon: 0.7409526400056237
steps: 292    lr: 4e-05     evaluation reward: 3.11
episode: 1203   score: 7.0   memory length: 231199   epsilon: 0.7402240000056395
steps: 368    lr: 4e-05     evaluation reward: 3.15
episode: 1204   score: 4.0   memory length: 231511   epsilon: 0.7396062400056529
steps: 312    lr: 4e-05     evaluation reward: 3.15
episode: 1205   score: 5.0   memory length: 231854   epsilon: 0.7389271000056676
steps: 343    lr: 4e-05     evaluation reward: 3.16
episode: 1206   score: 2.0   memory length: 232051   epsilon: 0.7385370400056761
steps: 197    lr: 4e-05     evaluation reward: 3.08
episode: 1207   score: 2.0   memory length: 232271   epsilon: 0.7381014400056856
steps: 220    lr: 4e-05     evaluation reward: 3.09
episode: 1208   score: 1.0   memory length: 232440   epsilon: 0.7377668200056928
steps: 169    lr: 4e-05     evaluation reward: 3.06
episode: 1209   score: 4.0   memory length: 232713   epsilon: 0.7372262800057046
steps: 273    lr: 4e-05     evaluation reward: 3.07
episode: 1210   score: 1.0   memory length: 232882   epsilon: 0.7368916600057118
steps: 169    lr: 4e-05     evaluation reward: 3.07
```

```
episode: 1211   score: 4.0   memory length: 233158   epsilon: 0.7363451800057237
steps: 276    lr: 4e-05     evaluation reward: 3.1
episode: 1212   score: 2.0   memory length: 233356   epsilon: 0.7359531400057322
steps: 198    lr: 4e-05     evaluation reward: 3.11
episode: 1213   score: 5.0   memory length: 233661   epsilon: 0.7353492400057453
steps: 305    lr: 4e-05     evaluation reward: 3.14
episode: 1214   score: 3.0   memory length: 233905   epsilon: 0.7348661200057558
steps: 244    lr: 4e-05     evaluation reward: 3.15
episode: 1215   score: 6.0   memory length: 234280   epsilon: 0.7341236200057719
steps: 375    lr: 4e-05     evaluation reward: 3.17
episode: 1216   score: 4.0   memory length: 234574   epsilon: 0.7335415000057846
steps: 294    lr: 4e-05     evaluation reward: 3.18
episode: 1217   score: 5.0   memory length: 234875   epsilon: 0.7329455200057975
steps: 301    lr: 4e-05     evaluation reward: 3.19
episode: 1218   score: 4.0   memory length: 235154   epsilon: 0.7323931000058095
steps: 279    lr: 4e-05     evaluation reward: 3.2
episode: 1219   score: 4.0   memory length: 235398   epsilon: 0.73190998000582
steps: 244    lr: 4e-05     evaluation reward: 3.21
episode: 1220   score: 4.0   memory length: 235696   epsilon: 0.7313199400058328
steps: 298    lr: 4e-05     evaluation reward: 3.24
episode: 1221   score: 3.0   memory length: 235921   epsilon: 0.7308744400058425
steps: 225    lr: 4e-05     evaluation reward: 3.23
episode: 1222   score: 3.0   memory length: 236167   epsilon: 0.730387360005853
steps: 246    lr: 4e-05     evaluation reward: 3.24
episode: 1223   score: 3.0   memory length: 236376   epsilon: 0.729973540005862
steps: 209    lr: 4e-05     evaluation reward: 3.25
episode: 1224   score: 2.0   memory length: 236595   epsilon: 0.7295399200058714
steps: 219    lr: 4e-05     evaluation reward: 3.25
episode: 1225   score: 2.0   memory length: 236794   epsilon: 0.72914590000588
steps: 199    lr: 4e-05     evaluation reward: 3.22
episode: 1226   score: 3.0   memory length: 237043   epsilon: 0.7286528800058907
steps: 249    lr: 4e-05     evaluation reward: 3.22
episode: 1227   score: 5.0   memory length: 237390   epsilon: 0.7279658200059056
steps: 347    lr: 4e-05     evaluation reward: 3.25
episode: 1228   score: 4.0   memory length: 237649   epsilon: 0.7274530000059167
steps: 259    lr: 4e-05     evaluation reward: 3.26
episode: 1229   score: 3.0   memory length: 237876   epsilon: 0.7270035400059265
steps: 227    lr: 4e-05     evaluation reward: 3.25
episode: 1230   score: 5.0   memory length: 238163   epsilon: 0.7264352800059388
steps: 287    lr: 4e-05     evaluation reward: 3.28
episode: 1231   score: 4.0   memory length: 238436   epsilon: 0.7258947400059506
steps: 273    lr: 4e-05     evaluation reward: 3.29
episode: 1232   score: 7.0   memory length: 238817   epsilon: 0.7251403600059669
steps: 381    lr: 4e-05     evaluation reward: 3.35
episode: 1233   score: 3.0   memory length: 239065   epsilon: 0.7246493200059776
steps: 248    lr: 4e-05     evaluation reward: 3.36
episode: 1234   score: 3.0   memory length: 239293   epsilon: 0.7241978800059874
steps: 228    lr: 4e-05     evaluation reward: 3.38
```

```
episode: 1235   score: 3.0   memory length: 239519   epsilon: 0.7237504000059971
steps: 226    lr: 4e-05     evaluation reward: 3.36
episode: 1236   score: 3.0   memory length: 239765   epsilon: 0.7232633200060077
steps: 246    lr: 4e-05     evaluation reward: 3.38
episode: 1237   score: 3.0   memory length: 240009   epsilon: 0.7227802000060182
steps: 244    lr: 4e-05     evaluation reward: 3.38
episode: 1238   score: 3.0   memory length: 240258   epsilon: 0.7222871800060289
steps: 249    lr: 4e-05     evaluation reward: 3.39
episode: 1239   score: 6.0   memory length: 240610   epsilon: 0.721590220006044
steps: 352    lr: 4e-05     evaluation reward: 3.42
episode: 1240   score: 1.0   memory length: 240779   epsilon: 0.7212556000060513
steps: 169    lr: 4e-05     evaluation reward: 3.41
episode: 1241   score: 5.0   memory length: 241067   epsilon: 0.7206853600060636
steps: 288    lr: 4e-05     evaluation reward: 3.41
episode: 1242   score: 1.0   memory length: 241218   epsilon: 0.7203863800060701
steps: 151    lr: 4e-05     evaluation reward: 3.38
episode: 1243   score: 0.0   memory length: 241341   epsilon: 0.7201428400060754
steps: 123    lr: 4e-05     evaluation reward: 3.34
episode: 1244   score: 4.0   memory length: 241638   epsilon: 0.7195547800060882
steps: 297    lr: 4e-05     evaluation reward: 3.34
episode: 1245   score: 2.0   memory length: 241836   epsilon: 0.7191627400060967
steps: 198    lr: 4e-05     evaluation reward: 3.31
episode: 1246   score: 4.0   memory length: 242093   epsilon: 0.7186538800061077
steps: 257    lr: 4e-05     evaluation reward: 3.32
episode: 1247   score: 2.0   memory length: 242290   epsilon: 0.7182638200061162
steps: 197    lr: 4e-05     evaluation reward: 3.33
episode: 1248   score: 4.0   memory length: 242566   epsilon: 0.7177173400061281
steps: 276    lr: 4e-05     evaluation reward: 3.36
episode: 1249   score: 3.0   memory length: 242792   epsilon: 0.7172698600061378
steps: 226    lr: 4e-05     evaluation reward: 3.37
episode: 1250   score: 4.0   memory length: 243049   epsilon: 0.7167610000061488
steps: 257    lr: 4e-05     evaluation reward: 3.38
episode: 1251   score: 3.0   memory length: 243294   epsilon: 0.7162759000061594
steps: 245    lr: 4e-05     evaluation reward: 3.39
episode: 1252   score: 2.0   memory length: 243494   epsilon: 0.715879900006168
steps: 200    lr: 4e-05     evaluation reward: 3.37
episode: 1253   score: 2.0   memory length: 243692   epsilon: 0.7154878600061765
steps: 198    lr: 4e-05     evaluation reward: 3.38
episode: 1254   score: 6.0   memory length: 244061   epsilon: 0.7147572400061923
steps: 369    lr: 4e-05     evaluation reward: 3.42
episode: 1255   score: 4.0   memory length: 244357   epsilon: 0.7141711600062051
steps: 296    lr: 4e-05     evaluation reward: 3.41
episode: 1256   score: 4.0   memory length: 244629   epsilon: 0.7136326000062168
steps: 272    lr: 4e-05     evaluation reward: 3.43
episode: 1257   score: 2.0   memory length: 244830   epsilon: 0.7132346200062254
steps: 201    lr: 4e-05     evaluation reward: 3.42
episode: 1258   score: 2.0   memory length: 245047   epsilon: 0.7128049600062347
steps: 217    lr: 4e-05     evaluation reward: 3.4
```

```
episode: 1259   score: 5.0   memory length: 245372   epsilon: 0.7121614600062487
steps: 325    lr: 4e-05     evaluation reward: 3.43
episode: 1260   score: 2.0   memory length: 245570   epsilon: 0.7117694200062572
steps: 198    lr: 4e-05     evaluation reward: 3.43
episode: 1261   score: 4.0   memory length: 245810   epsilon: 0.7112942200062675
steps: 240    lr: 4e-05     evaluation reward: 3.45
episode: 1262   score: 4.0   memory length: 246085   epsilon: 0.7107497200062793
steps: 275    lr: 4e-05     evaluation reward: 3.45
episode: 1263   score: 5.0   memory length: 246413   epsilon: 0.7101002800062934
steps: 328    lr: 4e-05     evaluation reward: 3.47
episode: 1264   score: 6.0   memory length: 246756   epsilon: 0.7094211400063082
steps: 343    lr: 4e-05     evaluation reward: 3.49
episode: 1265   score: 3.0   memory length: 246969   epsilon: 0.7089994000063173
steps: 213    lr: 4e-05     evaluation reward: 3.48
episode: 1266   score: 2.0   memory length: 247167   epsilon: 0.7086073600063258
steps: 198    lr: 4e-05     evaluation reward: 3.47
episode: 1267   score: 3.0   memory length: 247416   epsilon: 0.7081143400063366
steps: 249    lr: 4e-05     evaluation reward: 3.46
episode: 1268   score: 4.0   memory length: 247689   epsilon: 0.7075738000063483
steps: 273    lr: 4e-05     evaluation reward: 3.46
episode: 1269   score: 8.0   memory length: 248089   epsilon: 0.7067818000063655
steps: 400    lr: 4e-05     evaluation reward: 3.47
episode: 1270   score: 4.0   memory length: 248369   epsilon: 0.7062274000063775
steps: 280    lr: 4e-05     evaluation reward: 3.46
episode: 1271   score: 4.0   memory length: 248662   epsilon: 0.7056472600063901
steps: 293    lr: 4e-05     evaluation reward: 3.45
episode: 1272   score: 4.0   memory length: 248956   epsilon: 0.7050651400064027
steps: 294    lr: 4e-05     evaluation reward: 3.46
episode: 1273   score: 3.0   memory length: 249167   epsilon: 0.7046473600064118
steps: 211    lr: 4e-05     evaluation reward: 3.39
episode: 1274   score: 5.0   memory length: 249475   epsilon: 0.704037520006425
steps: 308    lr: 4e-05     evaluation reward: 3.39
episode: 1275   score: 5.0   memory length: 249782   epsilon: 0.7034296600064383
steps: 307    lr: 4e-05     evaluation reward: 3.38
episode: 1276   score: 3.0   memory length: 250012   epsilon: 0.7029742600064481
steps: 230    lr: 4e-05     evaluation reward: 3.39
episode: 1277   score: 4.0   memory length: 250286   epsilon: 0.7024317400064599
steps: 274    lr: 4e-05     evaluation reward: 3.4
episode: 1278   score: 2.0   memory length: 250484   epsilon: 0.7020397000064684
steps: 198    lr: 4e-05     evaluation reward: 3.4
episode: 1279   score: 5.0   memory length: 250828   epsilon: 0.7013585800064832
steps: 344    lr: 4e-05     evaluation reward: 3.42
episode: 1280   score: 4.0   memory length: 251102   epsilon: 0.700816060006495
steps: 274    lr: 4e-05     evaluation reward: 3.44
episode: 1281   score: 1.0   memory length: 251271   epsilon: 0.7004814400065023
steps: 169    lr: 4e-05     evaluation reward: 3.42
episode: 1282   score: 4.0   memory length: 251570   epsilon: 0.6998894200065151
steps: 299    lr: 4e-05     evaluation reward: 3.44
```

```
episode: 1283   score: 3.0   memory length: 251781   epsilon: 0.6994716400065242
steps: 211    lr: 4e-05     evaluation reward: 3.42
episode: 1284   score: 7.0   memory length: 252185   epsilon: 0.6986717200065415
steps: 404    lr: 4e-05     evaluation reward: 3.44
episode: 1285   score: 5.0   memory length: 252531   epsilon: 0.6979866400065564
steps: 346    lr: 4e-05     evaluation reward: 3.46
episode: 1286   score: 3.0   memory length: 252777   epsilon: 0.697499560006567
steps: 246    lr: 4e-05     evaluation reward: 3.47
episode: 1287   score: 2.0   memory length: 252976   epsilon: 0.6971055400065755
steps: 199    lr: 4e-05     evaluation reward: 3.44
episode: 1288   score: 8.0   memory length: 253439   epsilon: 0.6961888000065954
steps: 463    lr: 4e-05     evaluation reward: 3.51
episode: 1289   score: 5.0   memory length: 253762   epsilon: 0.6955492600066093
steps: 323    lr: 4e-05     evaluation reward: 3.54
episode: 1290   score: 5.0   memory length: 254048   epsilon: 0.6949829800066216
steps: 286    lr: 4e-05     evaluation reward: 3.59
episode: 1291   score: 4.0   memory length: 254344   epsilon: 0.6943969000066343
steps: 296    lr: 4e-05     evaluation reward: 3.6
episode: 1292   score: 1.0   memory length: 254495   epsilon: 0.6940979200066408
steps: 151    lr: 4e-05     evaluation reward: 3.6
episode: 1293   score: 2.0   memory length: 254693   epsilon: 0.6937058800066493
steps: 198    lr: 4e-05     evaluation reward: 3.59
episode: 1294   score: 2.0   memory length: 254912   epsilon: 0.6932722600066588
steps: 219    lr: 4e-05     evaluation reward: 3.57
episode: 1295   score: 3.0   memory length: 255138   epsilon: 0.6928247800066685
steps: 226    lr: 4e-05     evaluation reward: 3.59
episode: 1296   score: 2.0   memory length: 255335   epsilon: 0.6924347200066769
steps: 197    lr: 4e-05     evaluation reward: 3.6
episode: 1297   score: 0.0   memory length: 255457   epsilon: 0.6921931600066822
steps: 122    lr: 4e-05     evaluation reward: 3.58
episode: 1298   score: 2.0   memory length: 255639   epsilon: 0.69183280000669
steps: 182    lr: 4e-05     evaluation reward: 3.52
episode: 1299   score: 6.0   memory length: 256017   epsilon: 0.6910843600067063
steps: 378    lr: 4e-05     evaluation reward: 3.56
episode: 1300   score: 7.0   memory length: 256398   epsilon: 0.6903299800067226
steps: 381    lr: 4e-05     evaluation reward: 3.6
episode: 1301   score: 4.0   memory length: 256673   epsilon: 0.6897854800067345
steps: 275    lr: 4e-05     evaluation reward: 3.59
episode: 1302   score: 4.0   memory length: 256934   epsilon: 0.6892687000067457
steps: 261    lr: 4e-05     evaluation reward: 3.59
episode: 1303   score: 5.0   memory length: 257260   epsilon: 0.6886232200067597
steps: 326    lr: 4e-05     evaluation reward: 3.57
episode: 1304   score: 5.0   memory length: 257551   epsilon: 0.6880470400067722
steps: 291    lr: 4e-05     evaluation reward: 3.58
episode: 1305   score: 4.0   memory length: 257827   epsilon: 0.6875005600067841
steps: 276    lr: 4e-05     evaluation reward: 3.57
episode: 1306   score: 3.0   memory length: 258076   epsilon: 0.6870075400067948
steps: 249    lr: 4e-05     evaluation reward: 3.58
```

```
episode: 1307   score: 3.0   memory length: 258327   epsilon: 0.6865105600068055
steps: 251    lr: 4e-05     evaluation reward: 3.59
episode: 1308   score: 5.0   memory length: 258656   epsilon: 0.6858591400068197
steps: 329    lr: 4e-05     evaluation reward: 3.63
episode: 1309   score: 5.0   memory length: 259003   epsilon: 0.6851720800068346
steps: 347    lr: 4e-05     evaluation reward: 3.64
episode: 1310   score: 4.0   memory length: 259278   epsilon: 0.6846275800068464
steps: 275    lr: 4e-05     evaluation reward: 3.67
episode: 1311   score: 4.0   memory length: 259556   epsilon: 0.6840771400068584
steps: 278    lr: 4e-05     evaluation reward: 3.67
episode: 1312   score: 2.0   memory length: 259774   epsilon: 0.6836455000068677
steps: 218    lr: 4e-05     evaluation reward: 3.67
episode: 1313   score: 7.0   memory length: 260190   epsilon: 0.6828218200068856
steps: 416    lr: 4e-05     evaluation reward: 3.69
episode: 1314   score: 3.0   memory length: 260403   epsilon: 0.6824000800068948
steps: 213    lr: 4e-05     evaluation reward: 3.69
episode: 1315   score: 4.0   memory length: 260675   epsilon: 0.6818615200069065
steps: 272    lr: 4e-05     evaluation reward: 3.67
episode: 1316   score: 8.0   memory length: 261114   epsilon: 0.6809923000069253
steps: 439    lr: 4e-05     evaluation reward: 3.71
episode: 1317   score: 5.0   memory length: 261463   epsilon: 0.6803012800069403
steps: 349    lr: 4e-05     evaluation reward: 3.71
episode: 1318   score: 1.0   memory length: 261615   epsilon: 0.6800003200069469
steps: 152    lr: 4e-05     evaluation reward: 3.68
episode: 1319   score: 4.0   memory length: 261874   epsilon: 0.679487500006958
steps: 259    lr: 4e-05     evaluation reward: 3.68
episode: 1320   score: 5.0   memory length: 262188   epsilon: 0.6788657800069715
steps: 314    lr: 4e-05     evaluation reward: 3.69
episode: 1321   score: 3.0   memory length: 262435   epsilon: 0.6783767200069821
steps: 247    lr: 4e-05     evaluation reward: 3.69
episode: 1322   score: 2.0   memory length: 262654   epsilon: 0.6779431000069915
steps: 219    lr: 4e-05     evaluation reward: 3.68
episode: 1323   score: 4.0   memory length: 262934   epsilon: 0.6773887000070036
steps: 280    lr: 4e-05     evaluation reward: 3.69
episode: 1324   score: 2.0   memory length: 263131   epsilon: 0.676998640007012
steps: 197    lr: 4e-05     evaluation reward: 3.69
episode: 1325   score: 4.0   memory length: 263430   epsilon: 0.6764066200070249
steps: 299    lr: 4e-05     evaluation reward: 3.71
episode: 1326   score: 4.0   memory length: 263705   epsilon: 0.6758621200070367
steps: 275    lr: 4e-05     evaluation reward: 3.72
episode: 1327   score: 2.0   memory length: 263887   epsilon: 0.6755017600070445
steps: 182    lr: 4e-05     evaluation reward: 3.69
episode: 1328   score: 3.0   memory length: 264132   epsilon: 0.6750166600070551
steps: 245    lr: 4e-05     evaluation reward: 3.68
episode: 1329   score: 4.0   memory length: 264407   epsilon: 0.6744721600070669
steps: 275    lr: 4e-05     evaluation reward: 3.69
episode: 1330   score: 3.0   memory length: 264652   epsilon: 0.6739870600070774
steps: 245    lr: 4e-05     evaluation reward: 3.67
```

```
episode: 1331   score: 5.0   memory length: 264983   epsilon: 0.6733316800070916
steps: 331    lr: 4e-05     evaluation reward: 3.68
episode: 1332   score: 2.0   memory length: 265181   epsilon: 0.6729396400071002
steps: 198    lr: 4e-05     evaluation reward: 3.63
episode: 1333   score: 1.0   memory length: 265332   epsilon: 0.6726406600071066
steps: 151    lr: 4e-05     evaluation reward: 3.61
episode: 1334   score: 7.0   memory length: 265739   epsilon: 0.6718348000071241
steps: 407    lr: 4e-05     evaluation reward: 3.65
episode: 1335   score: 7.0   memory length: 266145   epsilon: 0.6710309200071416
steps: 406    lr: 4e-05     evaluation reward: 3.69
episode: 1336   score: 3.0   memory length: 266391   epsilon: 0.6705438400071522
steps: 246    lr: 4e-05     evaluation reward: 3.69
episode: 1337   score: 1.0   memory length: 266542   epsilon: 0.6702448600071587
steps: 151    lr: 4e-05     evaluation reward: 3.67
episode: 1338   score: 4.0   memory length: 266817   epsilon: 0.6697003600071705
steps: 275    lr: 4e-05     evaluation reward: 3.68
episode: 1339   score: 4.0   memory length: 267092   epsilon: 0.6691558600071823
steps: 275    lr: 4e-05     evaluation reward: 3.66
episode: 1340   score: 7.0   memory length: 267520   epsilon: 0.6683084200072007
steps: 428    lr: 4e-05     evaluation reward: 3.72
episode: 1341   score: 3.0   memory length: 267750   epsilon: 0.6678530200072106
steps: 230    lr: 4e-05     evaluation reward: 3.7
episode: 1342   score: 5.0   memory length: 268091   epsilon: 0.6671778400072252
steps: 341    lr: 4e-05     evaluation reward: 3.74
episode: 1343   score: 6.0   memory length: 268426   epsilon: 0.6665145400072396
steps: 335    lr: 4e-05     evaluation reward: 3.8
episode: 1344   score: 5.0   memory length: 268726   epsilon: 0.6659205400072525
steps: 300    lr: 4e-05     evaluation reward: 3.81
episode: 1345   score: 5.0   memory length: 269014   epsilon: 0.6653503000072649
steps: 288    lr: 4e-05     evaluation reward: 3.84
episode: 1346   score: 3.0   memory length: 269259   epsilon: 0.6648652000072754
steps: 245    lr: 4e-05     evaluation reward: 3.83
episode: 1347   score: 5.0   memory length: 269568   epsilon: 0.6642533800072887
steps: 309    lr: 4e-05     evaluation reward: 3.86
episode: 1348   score: 3.0   memory length: 269816   epsilon: 0.6637623400072994
steps: 248    lr: 4e-05     evaluation reward: 3.85
episode: 1349   score: 3.0   memory length: 270046   epsilon: 0.6633069400073093
steps: 230    lr: 4e-05     evaluation reward: 3.85
episode: 1350   score: 6.0   memory length: 270389   epsilon: 0.662627800007324
steps: 343    lr: 4e-05     evaluation reward: 3.87
episode: 1351   score: 3.0   memory length: 270615   epsilon: 0.6621803200073337
steps: 226    lr: 4e-05     evaluation reward: 3.87
episode: 1352   score: 5.0   memory length: 270957   epsilon: 0.6615031600073484
steps: 342    lr: 4e-05     evaluation reward: 3.9
episode: 1353   score: 1.0   memory length: 271129   epsilon: 0.6611626000073558
steps: 172    lr: 4e-05     evaluation reward: 3.89
episode: 1354   score: 2.0   memory length: 271330   epsilon: 0.6607646200073645
steps: 201    lr: 4e-05     evaluation reward: 3.85
```

```
episode: 1355   score: 4.0   memory length: 271627   epsilon: 0.6601765600073772
steps: 297    lr: 4e-05     evaluation reward: 3.85
episode: 1356   score: 2.0   memory length: 271844   epsilon: 0.6597469000073866
steps: 217    lr: 4e-05     evaluation reward: 3.83
episode: 1357   score: 1.0   memory length: 271995   epsilon: 0.659447920007393
steps: 151    lr: 4e-05     evaluation reward: 3.82
episode: 1358   score: 0.0   memory length: 272118   epsilon: 0.6592043800073983
steps: 123    lr: 4e-05     evaluation reward: 3.8
episode: 1359   score: 1.0   memory length: 272288   epsilon: 0.6588677800074056
steps: 170    lr: 4e-05     evaluation reward: 3.76
episode: 1360   score: 5.0   memory length: 272596   epsilon: 0.6582579400074189
steps: 308    lr: 4e-05     evaluation reward: 3.79
episode: 1361   score: 2.0   memory length: 272778   epsilon: 0.6578975800074267
steps: 182    lr: 4e-05     evaluation reward: 3.77
episode: 1362   score: 5.0   memory length: 273122   epsilon: 0.6572164600074415
steps: 344    lr: 4e-05     evaluation reward: 3.78
episode: 1363   score: 4.0   memory length: 273420   epsilon: 0.6566264200074543
steps: 298    lr: 4e-05     evaluation reward: 3.77
episode: 1364   score: 4.0   memory length: 273697   epsilon: 0.6560779600074662
steps: 277    lr: 4e-05     evaluation reward: 3.75
episode: 1365   score: 5.0   memory length: 274024   epsilon: 0.6554305000074803
steps: 327    lr: 4e-05     evaluation reward: 3.77
episode: 1366   score: 2.0   memory length: 274221   epsilon: 0.6550404400074887
steps: 197    lr: 4e-05     evaluation reward: 3.77
episode: 1367   score: 6.0   memory length: 274591   epsilon: 0.6543078400075046
steps: 370    lr: 4e-05     evaluation reward: 3.8
episode: 1368   score: 2.0   memory length: 274775   epsilon: 0.6539435200075125
steps: 184    lr: 4e-05     evaluation reward: 3.78
episode: 1369   score: 4.0   memory length: 275068   epsilon: 0.6533633800075251
steps: 293    lr: 4e-05     evaluation reward: 3.74
episode: 1370   score: 6.0   memory length: 275407   epsilon: 0.6526921600075397
steps: 339    lr: 4e-05     evaluation reward: 3.76
episode: 1371   score: 4.0   memory length: 275683   epsilon: 0.6521456800075516
steps: 276    lr: 4e-05     evaluation reward: 3.76
episode: 1372   score: 4.0   memory length: 275976   epsilon: 0.6515655400075642
steps: 293    lr: 4e-05     evaluation reward: 3.76
episode: 1373   score: 4.0   memory length: 276233   epsilon: 0.6510566800075752
steps: 257    lr: 4e-05     evaluation reward: 3.77
episode: 1374   score: 3.0   memory length: 276498   epsilon: 0.6505319800075866
steps: 265    lr: 4e-05     evaluation reward: 3.75
episode: 1375   score: 4.0   memory length: 276795   epsilon: 0.6499439200075994
steps: 297    lr: 4e-05     evaluation reward: 3.74
episode: 1376   score: 4.0   memory length: 277090   epsilon: 0.649359820007612
steps: 295    lr: 4e-05     evaluation reward: 3.75
episode: 1377   score: 5.0   memory length: 277432   epsilon: 0.6486826600076268
steps: 342    lr: 4e-05     evaluation reward: 3.76
episode: 1378   score: 2.0   memory length: 277648   epsilon: 0.648254980007636
steps: 216    lr: 4e-05     evaluation reward: 3.76
```

```
episode: 1379   score: 1.0   memory length: 277817   epsilon: 0.6479203600076433
steps: 169    lr: 4e-05     evaluation reward: 3.72
episode: 1380   score: 4.0   memory length: 278072   epsilon: 0.6474154600076543
steps: 255    lr: 4e-05     evaluation reward: 3.72
episode: 1381   score: 4.0   memory length: 278347   epsilon: 0.6468709600076661
steps: 275    lr: 4e-05     evaluation reward: 3.75
episode: 1382   score: 6.0   memory length: 278720   epsilon: 0.6461324200076821
steps: 373    lr: 4e-05     evaluation reward: 3.77
episode: 1383   score: 8.0   memory length: 279167   epsilon: 0.6452473600077013
steps: 447    lr: 4e-05     evaluation reward: 3.82
episode: 1384   score: 8.0   memory length: 279622   epsilon: 0.6443464600077209
steps: 455    lr: 4e-05     evaluation reward: 3.83
episode: 1385   score: 5.0   memory length: 279949   epsilon: 0.643699000007735
steps: 327    lr: 4e-05     evaluation reward: 3.83
episode: 1386   score: 2.0   memory length: 280131   epsilon: 0.6433386400077428
steps: 182    lr: 4e-05     evaluation reward: 3.82
episode: 1387   score: 3.0   memory length: 280360   epsilon: 0.6428852200077526
steps: 229    lr: 4e-05     evaluation reward: 3.83
episode: 1388   score: 4.0   memory length: 280655   epsilon: 0.6423011200077653
steps: 295    lr: 4e-05     evaluation reward: 3.79
episode: 1389   score: 4.0   memory length: 280899   epsilon: 0.6418180000077758
steps: 244    lr: 4e-05     evaluation reward: 3.78
episode: 1390   score: 5.0   memory length: 281241   epsilon: 0.6411408400077905
steps: 342    lr: 4e-05     evaluation reward: 3.78
episode: 1391   score: 10.0   memory length: 281777   epsilon:
0.6400795600078135    steps: 536    lr: 4e-05     evaluation reward: 3.84
episode: 1392   score: 8.0   memory length: 282229   epsilon: 0.639184600007833
steps: 452    lr: 4e-05     evaluation reward: 3.91
episode: 1393   score: 5.0   memory length: 282573   epsilon: 0.6385034800078477
steps: 344    lr: 4e-05     evaluation reward: 3.94
episode: 1394   score: 5.0   memory length: 282886   epsilon: 0.6378837400078612
steps: 313    lr: 4e-05     evaluation reward: 3.97
episode: 1395   score: 8.0   memory length: 283333   epsilon: 0.6369986800078804
steps: 447    lr: 4e-05     evaluation reward: 4.02
episode: 1396   score: 6.0   memory length: 283690   epsilon: 0.6362918200078957
steps: 357    lr: 4e-05     evaluation reward: 4.06
episode: 1397   score: 5.0   memory length: 283983   epsilon: 0.6357116800079083
steps: 293    lr: 4e-05     evaluation reward: 4.11
episode: 1398   score: 3.0   memory length: 284232   epsilon: 0.635218660007919
steps: 249    lr: 4e-05     evaluation reward: 4.12
episode: 1399   score: 1.0   memory length: 284401   epsilon: 0.6348840400079263
steps: 169    lr: 4e-05     evaluation reward: 4.07
episode: 1400   score: 4.0   memory length: 284695   epsilon: 0.634301920007939
steps: 294    lr: 4e-05     evaluation reward: 4.04
episode: 1401   score: 9.0   memory length: 285062   epsilon: 0.6335752600079547
steps: 367    lr: 4e-05     evaluation reward: 4.09
episode: 1402   score: 6.0   memory length: 285416   epsilon: 0.6328743400079699
steps: 354    lr: 4e-05     evaluation reward: 4.11
```

```
episode: 1403   score: 5.0   memory length: 285727   epsilon: 0.6322585600079833
steps: 311    lr: 4e-05     evaluation reward: 4.11
episode: 1404   score: 7.0   memory length: 286131   epsilon: 0.6314586400080007
steps: 404    lr: 4e-05     evaluation reward: 4.13
episode: 1405   score: 6.0   memory length: 286485   epsilon: 0.6307577200080159
steps: 354    lr: 4e-05     evaluation reward: 4.15
episode: 1406   score: 2.0   memory length: 286701   epsilon: 0.6303300400080252
steps: 216    lr: 4e-05     evaluation reward: 4.14
episode: 1407   score: 3.0   memory length: 286949   epsilon: 0.6298390000080358
steps: 248    lr: 4e-05     evaluation reward: 4.14
episode: 1408   score: 1.0   memory length: 287100   epsilon: 0.6295400200080423
steps: 151    lr: 4e-05     evaluation reward: 4.1
episode: 1409   score: 8.0   memory length: 287532   epsilon: 0.6286846600080609
steps: 432    lr: 4e-05     evaluation reward: 4.13
episode: 1410   score: 3.0   memory length: 287758   epsilon: 0.6282371800080706
steps: 226    lr: 4e-05     evaluation reward: 4.12
episode: 1411   score: 7.0   memory length: 288163   epsilon: 0.627435280008088
steps: 405    lr: 4e-05     evaluation reward: 4.15
episode: 1412   score: 3.0   memory length: 288408   epsilon: 0.6269501800080985
steps: 245    lr: 4e-05     evaluation reward: 4.16
episode: 1413   score: 2.0   memory length: 288606   epsilon: 0.6265581400081071
steps: 198    lr: 4e-05     evaluation reward: 4.11
episode: 1414   score: 5.0   memory length: 288930   epsilon: 0.625916620008121
steps: 324    lr: 4e-05     evaluation reward: 4.13
episode: 1415   score: 5.0   memory length: 289258   epsilon: 0.6252671800081351
steps: 328    lr: 4e-05     evaluation reward: 4.14
episode: 1416   score: 5.0   memory length: 289547   epsilon: 0.6246949600081475
steps: 289    lr: 4e-05     evaluation reward: 4.11
episode: 1417   score: 3.0   memory length: 289774   epsilon: 0.6242455000081573
steps: 227    lr: 4e-05     evaluation reward: 4.09
episode: 1418   score: 10.0   memory length: 290156   epsilon:
0.6234891400081737    steps: 382    lr: 4e-05     evaluation reward: 4.18
episode: 1419   score: 7.0   memory length: 290608   epsilon: 0.6225941800081931
steps: 452    lr: 4e-05     evaluation reward: 4.21
episode: 1420   score: 5.0   memory length: 290957   epsilon: 0.6219031600082081
steps: 349    lr: 4e-05     evaluation reward: 4.21
episode: 1421   score: 3.0   memory length: 291190   epsilon: 0.6214418200082181
steps: 233    lr: 4e-05     evaluation reward: 4.21
episode: 1422   score: 7.0   memory length: 291610   epsilon: 0.6206102200082362
steps: 420    lr: 4e-05     evaluation reward: 4.26
episode: 1423   score: 5.0   memory length: 291898   epsilon: 0.6200399800082486
steps: 288    lr: 4e-05     evaluation reward: 4.27
episode: 1424   score: 2.0   memory length: 292097   epsilon: 0.6196459600082571
steps: 199    lr: 4e-05     evaluation reward: 4.27
episode: 1425   score: 4.0   memory length: 292339   epsilon: 0.6191668000082675
steps: 242    lr: 4e-05     evaluation reward: 4.27
episode: 1426   score: 5.0   memory length: 292671   epsilon: 0.6185094400082818
steps: 332    lr: 4e-05     evaluation reward: 4.28
```

```
episode: 1427   score: 5.0   memory length: 292975   epsilon: 0.6179075200082949
steps: 304    lr: 4e-05     evaluation reward: 4.31
episode: 1428   score: 4.0   memory length: 293270   epsilon: 0.6173234200083075
steps: 295    lr: 4e-05     evaluation reward: 4.32
episode: 1429   score: 5.0   memory length: 293575   epsilon: 0.6167195200083206
steps: 305    lr: 4e-05     evaluation reward: 4.33
episode: 1430   score: 6.0   memory length: 293945   epsilon: 0.6159869200083365
steps: 370    lr: 4e-05     evaluation reward: 4.36
episode: 1431   score: 3.0   memory length: 294208   epsilon: 0.6154661800083479
steps: 263    lr: 4e-05     evaluation reward: 4.34
episode: 1432   score: 3.0   memory length: 294457   epsilon: 0.6149731600083586
steps: 249    lr: 4e-05     evaluation reward: 4.35
episode: 1433   score: 6.0   memory length: 294813   epsilon: 0.6142682800083739
steps: 356    lr: 4e-05     evaluation reward: 4.4
episode: 1434   score: 4.0   memory length: 295087   epsilon: 0.6137257600083856
steps: 274    lr: 4e-05     evaluation reward: 4.37
episode: 1435   score: 4.0   memory length: 295352   epsilon: 0.613201060008397
steps: 265    lr: 4e-05     evaluation reward: 4.34
episode: 1436   score: 4.0   memory length: 295629   epsilon: 0.6126526000084089
steps: 277    lr: 4e-05     evaluation reward: 4.35
episode: 1437   score: 3.0   memory length: 295875   epsilon: 0.6121655200084195
steps: 246    lr: 4e-05     evaluation reward: 4.37
episode: 1438   score: 3.0   memory length: 296122   epsilon: 0.6116764600084301
steps: 247    lr: 4e-05     evaluation reward: 4.36
episode: 1439   score: 4.0   memory length: 296382   epsilon: 0.6111616600084413
steps: 260    lr: 4e-05     evaluation reward: 4.36
episode: 1440   score: 5.0   memory length: 296697   epsilon: 0.6105379600084548
steps: 315    lr: 4e-05     evaluation reward: 4.34
episode: 1441   score: 3.0   memory length: 296942   epsilon: 0.6100528600084654
steps: 245    lr: 4e-05     evaluation reward: 4.34
episode: 1442   score: 4.0   memory length: 297237   epsilon: 0.609468760008478
steps: 295    lr: 4e-05     evaluation reward: 4.33
episode: 1443   score: 5.0   memory length: 297525   epsilon: 0.6088985200084904
steps: 288    lr: 4e-05     evaluation reward: 4.32
episode: 1444   score: 0.0   memory length: 297648   epsilon: 0.6086549800084957
steps: 123    lr: 4e-05     evaluation reward: 4.27
episode: 1445   score: 6.0   memory length: 298012   epsilon: 0.6079342600085114
steps: 364    lr: 4e-05     evaluation reward: 4.28
episode: 1446   score: 6.0   memory length: 298386   epsilon: 0.6071937400085274
steps: 374    lr: 4e-05     evaluation reward: 4.31
episode: 1447   score: 4.0   memory length: 298665   epsilon: 0.6066413200085394
steps: 279    lr: 4e-05     evaluation reward: 4.3
episode: 1448   score: 3.0   memory length: 298876   epsilon: 0.6062235400085485
steps: 211    lr: 4e-05     evaluation reward: 4.3
episode: 1449   score: 4.0   memory length: 299151   epsilon: 0.6056790400085603
steps: 275    lr: 4e-05     evaluation reward: 4.31
episode: 1450   score: 4.0   memory length: 299427   epsilon: 0.6051325600085722
steps: 276    lr: 4e-05     evaluation reward: 4.29
```

```
episode: 1451   score: 3.0   memory length: 299652   epsilon: 0.6046870600085819
steps: 225    lr: 4e-05     evaluation reward: 4.29
episode: 1452   score: 6.0   memory length: 299988   epsilon: 0.6040217800085963
steps: 336    lr: 4e-05     evaluation reward: 4.3
episode: 1453   score: 1.0   memory length: 300139   epsilon: 0.6037228000086028
steps: 151    lr: 1.6000000000000003e-05     evaluation reward: 4.3
episode: 1454   score: 2.0   memory length: 300355   epsilon: 0.6032951200086121
steps: 216    lr: 1.6000000000000003e-05     evaluation reward: 4.3
episode: 1455   score: 6.0   memory length: 300710   epsilon: 0.6025922200086273
steps: 355    lr: 1.6000000000000003e-05     evaluation reward: 4.32
episode: 1456   score: 6.0   memory length: 301082   epsilon: 0.6018556600086433
steps: 372    lr: 1.6000000000000003e-05     evaluation reward: 4.36
episode: 1457   score: 6.0   memory length: 301436   epsilon: 0.6011547400086585
steps: 354    lr: 1.6000000000000003e-05     evaluation reward: 4.41
episode: 1458   score: 3.0   memory length: 301648   epsilon: 0.6007349800086677
steps: 212    lr: 1.6000000000000003e-05     evaluation reward: 4.44
episode: 1459   score: 9.0   memory length: 302153   epsilon: 0.5997350800086894
steps: 505    lr: 1.6000000000000003e-05     evaluation reward: 4.52
episode: 1460   score: 1.0   memory length: 302324   epsilon: 0.5993965000086967
steps: 171    lr: 1.6000000000000003e-05     evaluation reward: 4.48
episode: 1461   score: 6.0   memory length: 302661   epsilon: 0.5987292400087112
steps: 337    lr: 1.6000000000000003e-05     evaluation reward: 4.52
episode: 1462   score: 3.0   memory length: 302873   epsilon: 0.5983094800087203
steps: 212    lr: 1.6000000000000003e-05     evaluation reward: 4.5
episode: 1463   score: 11.0   memory length: 303345   epsilon:
0.5973749200087406    steps: 472    lr: 1.6000000000000003e-05     evaluation
reward: 4.57
episode: 1464   score: 7.0   memory length: 303749   epsilon: 0.596575000008758
steps: 404    lr: 1.6000000000000003e-05     evaluation reward: 4.6
episode: 1465   score: 4.0   memory length: 304025   epsilon: 0.5960285200087698
steps: 276    lr: 1.6000000000000003e-05     evaluation reward: 4.59
episode: 1466   score: 2.0   memory length: 304225   epsilon: 0.5956325200087784
steps: 200    lr: 1.6000000000000003e-05     evaluation reward: 4.59
episode: 1467   score: 2.0   memory length: 304404   epsilon: 0.5952781000087861
steps: 179    lr: 1.6000000000000003e-05     evaluation reward: 4.55
episode: 1468   score: 7.0   memory length: 304802   epsilon: 0.5944900600088032
steps: 398    lr: 1.6000000000000003e-05     evaluation reward: 4.6
episode: 1469   score: 5.0   memory length: 305143   epsilon: 0.5938148800088179
steps: 341    lr: 1.6000000000000003e-05     evaluation reward: 4.61
episode: 1470   score: 11.0   memory length: 305688   epsilon:
0.5927357800088413    steps: 545    lr: 1.6000000000000003e-05     evaluation
reward: 4.66
episode: 1471   score: 7.0   memory length: 306087   epsilon: 0.5919457600088585
steps: 399    lr: 1.6000000000000003e-05     evaluation reward: 4.69
episode: 1472   score: 3.0   memory length: 306316   epsilon: 0.5914923400088683
steps: 229    lr: 1.6000000000000003e-05     evaluation reward: 4.68
episode: 1473   score: 9.0   memory length: 306790   epsilon: 0.5905538200088887
steps: 474    lr: 1.6000000000000003e-05     evaluation reward: 4.73
```

```
episode: 1474   score: 3.0   memory length: 307003   epsilon: 0.5901320800088978
steps: 213    lr: 1.6000000000000003e-05     evaluation reward: 4.73
episode: 1475   score: 9.0   memory length: 307466   epsilon: 0.5892153400089177
steps: 463    lr: 1.6000000000000003e-05     evaluation reward: 4.78
episode: 1476   score: 6.0   memory length: 307825   epsilon: 0.5885045200089332
steps: 359    lr: 1.6000000000000003e-05     evaluation reward: 4.8
episode: 1477   score: 7.0   memory length: 308226   epsilon: 0.5877105400089504
steps: 401    lr: 1.6000000000000003e-05     evaluation reward: 4.82
episode: 1478   score: 4.0   memory length: 308499   epsilon: 0.5871700000089621
steps: 273    lr: 1.6000000000000003e-05     evaluation reward: 4.84
episode: 1479   score: 9.0   memory length: 309023   epsilon: 0.5861324800089847
steps: 524    lr: 1.6000000000000003e-05     evaluation reward: 4.92
episode: 1480   score: 5.0   memory length: 309374   epsilon: 0.5854375000089997
steps: 351    lr: 1.6000000000000003e-05     evaluation reward: 4.93
episode: 1481   score: 4.0   memory length: 309655   epsilon: 0.5848811200090118
steps: 281    lr: 1.6000000000000003e-05     evaluation reward: 4.93
episode: 1482   score: 3.0   memory length: 309886   epsilon: 0.5844237400090218
steps: 231    lr: 1.6000000000000003e-05     evaluation reward: 4.9
episode: 1483   score: 1.0   memory length: 310037   epsilon: 0.5841247600090282
steps: 151    lr: 1.6000000000000003e-05     evaluation reward: 4.83
episode: 1484   score: 4.0   memory length: 310299   epsilon: 0.5836060000090395
steps: 262    lr: 1.6000000000000003e-05     evaluation reward: 4.79
episode: 1485   score: 5.0   memory length: 310625   epsilon: 0.5829605200090535
steps: 326    lr: 1.6000000000000003e-05     evaluation reward: 4.79
episode: 1486   score: 6.0   memory length: 310948   epsilon: 0.5823209800090674
steps: 323    lr: 1.6000000000000003e-05     evaluation reward: 4.83
episode: 1487   score: 10.0   memory length: 311475   epsilon: 0.58127752000909
steps: 527    lr: 1.6000000000000003e-05     evaluation reward: 4.9
episode: 1488   score: 3.0   memory length: 311688   epsilon: 0.5808557800090992
steps: 213    lr: 1.6000000000000003e-05     evaluation reward: 4.89
episode: 1489   score: 5.0   memory length: 312012   epsilon: 0.5802142600091131
steps: 324    lr: 1.6000000000000003e-05     evaluation reward: 4.9
episode: 1490   score: 2.0   memory length: 312191   epsilon: 0.5798598400091208
steps: 179    lr: 1.6000000000000003e-05     evaluation reward: 4.87
episode: 1491   score: 6.0   memory length: 312528   epsilon: 0.5791925800091353
steps: 337    lr: 1.6000000000000003e-05     evaluation reward: 4.83
episode: 1492   score: 5.0   memory length: 312801   epsilon: 0.578652040009147
steps: 273    lr: 1.6000000000000003e-05     evaluation reward: 4.8
episode: 1493   score: 3.0   memory length: 313048   epsilon: 0.5781629800091577
steps: 247    lr: 1.6000000000000003e-05     evaluation reward: 4.78
episode: 1494   score: 8.0   memory length: 313478   epsilon: 0.5773115800091762
steps: 430    lr: 1.6000000000000003e-05     evaluation reward: 4.81
episode: 1495   score: 3.0   memory length: 313725   epsilon: 0.5768225200091868
steps: 247    lr: 1.6000000000000003e-05     evaluation reward: 4.76
episode: 1496   score: 4.0   memory length: 314002   epsilon: 0.5762740600091987
steps: 277    lr: 1.6000000000000003e-05     evaluation reward: 4.74
episode: 1497   score: 5.0   memory length: 314312   epsilon: 0.575660260009212
steps: 310    lr: 1.6000000000000003e-05     evaluation reward: 4.74
```

```
episode: 1498   score: 6.0   memory length: 314673   epsilon: 0.5749454800092275
steps: 361    lr: 1.6000000000000003e-05     evaluation reward: 4.77
episode: 1499   score: 6.0   memory length: 315044   epsilon: 0.5742109000092435
steps: 371    lr: 1.6000000000000003e-05     evaluation reward: 4.82
episode: 1500   score: 5.0   memory length: 315324   epsilon: 0.5736565000092555
steps: 280    lr: 1.6000000000000003e-05     evaluation reward: 4.83
episode: 1501   score: 9.0   memory length: 315788   epsilon: 0.5727377800092754
steps: 464    lr: 1.6000000000000003e-05     evaluation reward: 4.83
episode: 1502   score: 4.0   memory length: 316062   epsilon: 0.5721952600092872
steps: 274    lr: 1.6000000000000003e-05     evaluation reward: 4.81
episode: 1503   score: 4.0   memory length: 316321   epsilon: 0.5716824400092984
steps: 259    lr: 1.6000000000000003e-05     evaluation reward: 4.8
episode: 1504   score: 4.0   memory length: 316579   epsilon: 0.5711716000093094
steps: 258    lr: 1.6000000000000003e-05     evaluation reward: 4.77
episode: 1505   score: 4.0   memory length: 316855   epsilon: 0.5706251200093213
steps: 276    lr: 1.6000000000000003e-05     evaluation reward: 4.75
episode: 1506   score: 7.0   memory length: 317263   epsilon: 0.5698172800093388
steps: 408    lr: 1.6000000000000003e-05     evaluation reward: 4.8
episode: 1507   score: 8.0   memory length: 317679   epsilon: 0.5689936000093567
steps: 416    lr: 1.6000000000000003e-05     evaluation reward: 4.85
episode: 1508   score: 2.0   memory length: 317877   epsilon: 0.5686015600093652
steps: 198    lr: 1.6000000000000003e-05     evaluation reward: 4.86
episode: 1509   score: 2.0   memory length: 318075   epsilon: 0.5682095200093737
steps: 198    lr: 1.6000000000000003e-05     evaluation reward: 4.8
episode: 1510   score: 8.0   memory length: 318523   epsilon: 0.567322480009393
steps: 448    lr: 1.6000000000000003e-05     evaluation reward: 4.85
episode: 1511   score: 6.0   memory length: 318901   epsilon: 0.5665740400094093
steps: 378    lr: 1.6000000000000003e-05     evaluation reward: 4.84
episode: 1512   score: 6.0   memory length: 319241   epsilon: 0.5659008400094239
steps: 340    lr: 1.6000000000000003e-05     evaluation reward: 4.87
episode: 1513   score: 8.0   memory length: 319666   epsilon: 0.5650593400094421
steps: 425    lr: 1.6000000000000003e-05     evaluation reward: 4.93
episode: 1514   score: 3.0   memory length: 319879   epsilon: 0.5646376000094513
steps: 213    lr: 1.6000000000000003e-05     evaluation reward: 4.91
episode: 1515   score: 3.0   memory length: 320107   epsilon: 0.5641861600094611
steps: 228    lr: 1.6000000000000003e-05     evaluation reward: 4.89
episode: 1516   score: 3.0   memory length: 320356   epsilon: 0.5636931400094718
steps: 249    lr: 1.6000000000000003e-05     evaluation reward: 4.87
episode: 1517   score: 7.0   memory length: 320727   epsilon: 0.5629585600094877
steps: 371    lr: 1.6000000000000003e-05     evaluation reward: 4.91
episode: 1518   score: 3.0   memory length: 320957   epsilon: 0.5625031600094976
steps: 230    lr: 1.6000000000000003e-05     evaluation reward: 4.84
episode: 1519   score: 5.0   memory length: 321267   epsilon: 0.561889360009511
steps: 310    lr: 1.6000000000000003e-05     evaluation reward: 4.82
episode: 1520   score: 3.0   memory length: 321494   epsilon: 0.5614399000095207
steps: 227    lr: 1.6000000000000003e-05     evaluation reward: 4.8
episode: 1521   score: 5.0   memory length: 321775   epsilon: 0.5608835200095328
steps: 281    lr: 1.6000000000000003e-05     evaluation reward: 4.82
```

episode: 1522   score: 8.0   memory length: 322212   epsilon: 0.5600182600095516
steps: 437   lr: 1.6000000000000003e-05   evaluation reward: 4.83
episode: 1523   score: 7.0   memory length: 322584   epsilon: 0.5592817000095676
steps: 372   lr: 1.6000000000000003e-05   evaluation reward: 4.85
episode: 1524   score: 6.0   memory length: 322924   epsilon: 0.5586085000095822
steps: 340   lr: 1.6000000000000003e-05   evaluation reward: 4.89
episode: 1525   score: 4.0   memory length: 323166   epsilon: 0.5581293400095926
steps: 242   lr: 1.6000000000000003e-05   evaluation reward: 4.89
episode: 1526   score: 4.0   memory length: 323406   epsilon: 0.5576541400096029
steps: 240   lr: 1.6000000000000003e-05   evaluation reward: 4.88
episode: 1527   score: 6.0   memory length: 323764   epsilon: 0.5569453000096183
steps: 358   lr: 1.6000000000000003e-05   evaluation reward: 4.89
episode: 1528   score: 6.0   memory length: 324137   epsilon: 0.5562067600096343
steps: 373   lr: 1.6000000000000003e-05   evaluation reward: 4.91
episode: 1529   score: 4.0   memory length: 324416   epsilon: 0.5556543400096463
steps: 279   lr: 1.6000000000000003e-05   evaluation reward: 4.9
episode: 1530   score: 5.0   memory length: 324724   epsilon: 0.5550445000096595
steps: 308   lr: 1.6000000000000003e-05   evaluation reward: 4.89
episode: 1531   score: 5.0   memory length: 325046   epsilon: 0.5544069400096734
steps: 322   lr: 1.6000000000000003e-05   evaluation reward: 4.91
episode: 1532   score: 4.0   memory length: 325323   epsilon: 0.5538584800096853
steps: 277   lr: 1.6000000000000003e-05   evaluation reward: 4.92
episode: 1533   score: 8.0   memory length: 325783   epsilon: 0.5529476800097051
steps: 460   lr: 1.6000000000000003e-05   evaluation reward: 4.94
episode: 1534   score: 9.0   memory length: 326242   epsilon: 0.5520388600097248
steps: 459   lr: 1.6000000000000003e-05   evaluation reward: 4.99
episode: 1535   score: 6.0   memory length: 326614   epsilon: 0.5513023000097408
steps: 372   lr: 1.6000000000000003e-05   evaluation reward: 5.01
episode: 1536   score: 5.0   memory length: 326902   epsilon: 0.5507320600097532
steps: 288   lr: 1.6000000000000003e-05   evaluation reward: 5.02
episode: 1537   score: 7.0   memory length: 327332   epsilon: 0.5498806600097716
steps: 430   lr: 1.6000000000000003e-05   evaluation reward: 5.06
episode: 1538   score: 4.0   memory length: 327591   epsilon: 0.5493678400097828
steps: 259   lr: 1.6000000000000003e-05   evaluation reward: 5.07
episode: 1539   score: 6.0   memory length: 327949   epsilon: 0.5486590000097982
steps: 358   lr: 1.6000000000000003e-05   evaluation reward: 5.09
episode: 1540   score: 10.0   memory length: 328385   epsilon:
0.5477957200098169   steps: 436   lr: 1.6000000000000003e-05   evaluation
reward: 5.14
episode: 1541   score: 8.0   memory length: 328793   epsilon: 0.5469878800098344
steps: 408   lr: 1.6000000000000003e-05   evaluation reward: 5.19
episode: 1542   score: 7.0   memory length: 329146   epsilon: 0.5462889400098496
steps: 353   lr: 1.6000000000000003e-05   evaluation reward: 5.22
episode: 1543   score: 6.0   memory length: 329489   epsilon: 0.5456098000098644
steps: 343   lr: 1.6000000000000003e-05   evaluation reward: 5.23
episode: 1544   score: 5.0   memory length: 329813   epsilon: 0.5449682800098783
steps: 324   lr: 1.6000000000000003e-05   evaluation reward: 5.28
episode: 1545   score: 6.0   memory length: 330124   epsilon: 0.5443525000098917

steps: 311   lr: 1.6000000000000003e-05   evaluation reward: 5.28
episode: 1546   score: 4.0   memory length: 330390   epsilon: 0.5438258200099031
steps: 266   lr: 1.6000000000000003e-05   evaluation reward: 5.26
episode: 1547   score: 10.0   memory length: 330930   epsilon:
0.5427566200099263   steps: 540   lr: 1.6000000000000003e-05   evaluation
reward: 5.32
episode: 1548   score: 7.0   memory length: 331336   epsilon: 0.5419527400099438
steps: 406   lr: 1.6000000000000003e-05   evaluation reward: 5.36
episode: 1549   score: 6.0   memory length: 331720   epsilon: 0.5411924200099603
steps: 384   lr: 1.6000000000000003e-05   evaluation reward: 5.38
episode: 1550   score: 4.0   memory length: 331978   epsilon: 0.5406815800099714
steps: 258   lr: 1.6000000000000003e-05   evaluation reward: 5.38
episode: 1551   score: 8.0   memory length: 332390   epsilon: 0.5398658200099891
steps: 412   lr: 1.6000000000000003e-05   evaluation reward: 5.43
episode: 1552   score: 6.0   memory length: 332728   epsilon: 0.5391965800100036
steps: 338   lr: 1.6000000000000003e-05   evaluation reward: 5.43
episode: 1553   score: 8.0   memory length: 333126   epsilon: 0.5384085400100207
steps: 398   lr: 1.6000000000000003e-05   evaluation reward: 5.5
episode: 1554   score: 5.0   memory length: 333434   epsilon: 0.5377987000100339
steps: 308   lr: 1.6000000000000003e-05   evaluation reward: 5.53
episode: 1555   score: 4.0   memory length: 333696   epsilon: 0.5372799400100452
steps: 262   lr: 1.6000000000000003e-05   evaluation reward: 5.51
episode: 1556   score: 3.0   memory length: 333943   epsilon: 0.5367908800100558
steps: 247   lr: 1.6000000000000003e-05   evaluation reward: 5.48
episode: 1557   score: 5.0   memory length: 334286   epsilon: 0.5361117400100706
steps: 343   lr: 1.6000000000000003e-05   evaluation reward: 5.47
episode: 1558   score: 3.0   memory length: 334516   epsilon: 0.5356563400100804
steps: 230   lr: 1.6000000000000003e-05   evaluation reward: 5.47
episode: 1559   score: 6.0   memory length: 334856   epsilon: 0.5349831400100951
steps: 340   lr: 1.6000000000000003e-05   evaluation reward: 5.44
episode: 1560   score: 12.0   memory length: 335299   epsilon:
0.5341060000101141   steps: 443   lr: 1.6000000000000003e-05   evaluation
reward: 5.55
episode: 1561   score: 5.0   memory length: 335623   epsilon: 0.533464480010128
steps: 324   lr: 1.6000000000000003e-05   evaluation reward: 5.54
episode: 1562   score: 5.0   memory length: 335934   epsilon: 0.5328487000101414
steps: 311   lr: 1.6000000000000003e-05   evaluation reward: 5.56
episode: 1563   score: 2.0   memory length: 336136   epsilon: 0.5324487400101501
steps: 202   lr: 1.6000000000000003e-05   evaluation reward: 5.47
episode: 1564   score: 1.0   memory length: 336287   epsilon: 0.5321497600101566
steps: 151   lr: 1.6000000000000003e-05   evaluation reward: 5.41
episode: 1565   score: 4.0   memory length: 336532   epsilon: 0.5316646600101671
steps: 245   lr: 1.6000000000000003e-05   evaluation reward: 5.41
episode: 1566   score: 4.0   memory length: 336807   epsilon: 0.5311201600101789
steps: 275   lr: 1.6000000000000003e-05   evaluation reward: 5.43
episode: 1567   score: 5.0   memory length: 337120   epsilon: 0.5305004200101924
steps: 313   lr: 1.6000000000000003e-05   evaluation reward: 5.46
episode: 1568   score: 8.0   memory length: 337576   epsilon: 0.529597540010212

steps: 456 lr: 1.6000000000000003e-05 evaluation reward: 5.47
episode: 1569 score: 5.0 memory length: 337906 epsilon: 0.5289441400102262
steps: 330 lr: 1.6000000000000003e-05 evaluation reward: 5.47
episode: 1570 score: 6.0 memory length: 338277 epsilon: 0.5282095600102421
steps: 371 lr: 1.6000000000000003e-05 evaluation reward: 5.42
episode: 1571 score: 6.0 memory length: 338636 epsilon: 0.5274987400102575
steps: 359 lr: 1.6000000000000003e-05 evaluation reward: 5.41
episode: 1572 score: 5.0 memory length: 338926 epsilon: 0.52692454001027
steps: 290 lr: 1.6000000000000003e-05 evaluation reward: 5.43
episode: 1573 score: 7.0 memory length: 339362 epsilon: 0.5260612600102887
steps: 436 lr: 1.6000000000000003e-05 evaluation reward: 5.41
episode: 1574 score: 5.0 memory length: 339706 epsilon: 0.5253801400103035
steps: 344 lr: 1.6000000000000003e-05 evaluation reward: 5.43
episode: 1575 score: 10.0 memory length: 340222 epsilon:
0.5243584600103257 steps: 516 lr: 1.6000000000000003e-05 evaluation
reward: 5.44
episode: 1576 score: 12.0 memory length: 340683 epsilon:
0.5234456800103455 steps: 461 lr: 1.6000000000000003e-05 evaluation
reward: 5.5
episode: 1577 score: 14.0 memory length: 341302 epsilon:
0.5222200600103721 steps: 619 lr: 1.6000000000000003e-05 evaluation
reward: 5.57
episode: 1578 score: 6.0 memory length: 341652 epsilon: 0.5215270600103872
steps: 350 lr: 1.6000000000000003e-05 evaluation reward: 5.59
episode: 1579 score: 7.0 memory length: 342076 epsilon: 0.5206875400104054
steps: 424 lr: 1.6000000000000003e-05 evaluation reward: 5.57
episode: 1580 score: 4.0 memory length: 342352 epsilon: 0.5201410600104173
steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 5.56
episode: 1581 score: 5.0 memory length: 342695 epsilon: 0.519461920010432
steps: 343 lr: 1.6000000000000003e-05 evaluation reward: 5.57
episode: 1582 score: 8.0 memory length: 342990 epsilon: 0.5188778200104447
steps: 295 lr: 1.6000000000000003e-05 evaluation reward: 5.62
episode: 1583 score: 5.0 memory length: 343316 epsilon: 0.5182323400104587
steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 5.66
episode: 1584 score: 4.0 memory length: 343596 epsilon: 0.5176779400104707
steps: 280 lr: 1.6000000000000003e-05 evaluation reward: 5.66
episode: 1585 score: 9.0 memory length: 344108 epsilon: 0.5166641800104927
steps: 512 lr: 1.6000000000000003e-05 evaluation reward: 5.7
episode: 1586 score: 8.0 memory length: 344509 epsilon: 0.51587020001051
steps: 401 lr: 1.6000000000000003e-05 evaluation reward: 5.72
episode: 1587 score: 3.0 memory length: 344740 epsilon: 0.5154128200105199
steps: 231 lr: 1.6000000000000003e-05 evaluation reward: 5.65
episode: 1588 score: 8.0 memory length: 345034 epsilon: 0.5148307000105325
steps: 294 lr: 1.6000000000000003e-05 evaluation reward: 5.7
episode: 1589 score: 8.0 memory length: 345467 epsilon: 0.5139733600105512
steps: 433 lr: 1.6000000000000003e-05 evaluation reward: 5.73
episode: 1590 score: 3.0 memory length: 345713 epsilon: 0.5134862800105617
steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 5.74

episode: 1591   score: 5.0   memory length: 346041   epsilon: 0.5128368400105758
steps: 328   lr: 1.6000000000000003e-05   evaluation reward: 5.73
episode: 1592   score: 2.0   memory length: 346242   epsilon: 0.5124388600105845
steps: 201   lr: 1.6000000000000003e-05   evaluation reward: 5.7
episode: 1593   score: 6.0   memory length: 346602   epsilon: 0.5117260600106
steps: 360   lr: 1.6000000000000003e-05   evaluation reward: 5.73
episode: 1594   score: 4.0   memory length: 346862   epsilon: 0.5112112600106111
steps: 260   lr: 1.6000000000000003e-05   evaluation reward: 5.69
episode: 1595   score: 5.0   memory length: 347206   epsilon: 0.5105301400106259
steps: 344   lr: 1.6000000000000003e-05   evaluation reward: 5.71
episode: 1596   score: 5.0   memory length: 347517   epsilon: 0.5099143600106393
steps: 311   lr: 1.6000000000000003e-05   evaluation reward: 5.72
episode: 1597   score: 6.0   memory length: 347838   epsilon: 0.5092787800106531
steps: 321   lr: 1.6000000000000003e-05   evaluation reward: 5.73
episode: 1598   score: 7.0   memory length: 348224   epsilon: 0.5085145000106697
steps: 386   lr: 1.6000000000000003e-05   evaluation reward: 5.74
episode: 1599   score: 6.0   memory length: 348562   epsilon: 0.5078452600106842
steps: 338   lr: 1.6000000000000003e-05   evaluation reward: 5.74
episode: 1600   score: 8.0   memory length: 349006   epsilon: 0.5069661400107033
steps: 444   lr: 1.6000000000000003e-05   evaluation reward: 5.77
episode: 1601   score: 5.0   memory length: 349331   epsilon: 0.5063226400107173
steps: 325   lr: 1.6000000000000003e-05   evaluation reward: 5.73
episode: 1602   score: 6.0   memory length: 349706   epsilon: 0.5055801400107334
steps: 375   lr: 1.6000000000000003e-05   evaluation reward: 5.75
episode: 1603   score: 9.0   memory length: 350178   epsilon: 0.5046455800107537
steps: 472   lr: 1.6000000000000003e-05   evaluation reward: 5.8
episode: 1604   score: 10.0   memory length: 350700   epsilon:
0.5036120200107761   steps: 522   lr: 1.6000000000000003e-05   evaluation
reward: 5.86
episode: 1605   score: 7.0   memory length: 351059   epsilon: 0.5029012000107915
steps: 359   lr: 1.6000000000000003e-05   evaluation reward: 5.89
episode: 1606   score: 8.0   memory length: 351516   epsilon: 0.5019963400108112
steps: 457   lr: 1.6000000000000003e-05   evaluation reward: 5.9
episode: 1607   score: 4.0   memory length: 351798   epsilon: 0.5014379800108233
steps: 282   lr: 1.6000000000000003e-05   evaluation reward: 5.86
episode: 1608   score: 6.0   memory length: 352192   epsilon: 0.5006578600108402
steps: 394   lr: 1.6000000000000003e-05   evaluation reward: 5.9
episode: 1609   score: 6.0   memory length: 352506   epsilon: 0.5000361400108537
steps: 314   lr: 1.6000000000000003e-05   evaluation reward: 5.94
episode: 1610   score: 8.0   memory length: 352906   epsilon: 0.4992441400108497
steps: 400   lr: 1.6000000000000003e-05   evaluation reward: 5.94
episode: 1611   score: 1.0   memory length: 353057   epsilon: 0.4989451600108478
steps: 151   lr: 1.6000000000000003e-05   evaluation reward: 5.89
episode: 1612   score: 5.0   memory length: 353327   epsilon:
0.49841056001084444   steps: 270   lr: 1.6000000000000003e-05   evaluation
reward: 5.88
episode: 1613   score: 6.0   memory length: 353665   epsilon: 0.4977413200108402
steps: 338   lr: 1.6000000000000003e-05   evaluation reward: 5.86

episode: 1614   score: 8.0   memory length: 354115   epsilon:
0.49685032001083457   steps: 450   lr: 1.6000000000000003e-05   evaluation
reward: 5.91
episode: 1615   score: 5.0   memory length: 354423   epsilon: 0.4962404800108307
steps: 308   lr: 1.6000000000000003e-05   evaluation reward: 5.93
episode: 1616   score: 8.0   memory length: 354869   epsilon: 0.4953574000108251
steps: 446   lr: 1.6000000000000003e-05   evaluation reward: 5.98
episode: 1617   score: 7.0   memory length: 355290   epsilon:
0.49452382001081985   steps: 421   lr: 1.6000000000000003e-05   evaluation
reward: 5.98
episode: 1618   score: 5.0   memory length: 355637   epsilon: 0.4938367600108155
steps: 347   lr: 1.6000000000000003e-05   evaluation reward: 6.0
episode: 1619   score: 7.0   memory length: 356040   epsilon:
0.49303882001081045   steps: 403   lr: 1.6000000000000003e-05   evaluation
reward: 6.02
episode: 1620   score: 6.0   memory length: 356414   epsilon:
0.49229830001080577   steps: 374   lr: 1.6000000000000003e-05   evaluation
reward: 6.05
episode: 1621   score: 11.0   memory length: 356834   epsilon:
0.4914667000108005   steps: 420   lr: 1.6000000000000003e-05   evaluation
reward: 6.11
episode: 1622   score: 5.0   memory length: 357143   epsilon:
0.49085488001079663   steps: 309   lr: 1.6000000000000003e-05   evaluation
reward: 6.08
episode: 1623   score: 8.0   memory length: 357587   epsilon: 0.4899757600107911
steps: 444   lr: 1.6000000000000003e-05   evaluation reward: 6.09
episode: 1624   score: 6.0   memory length: 357925   epsilon:
0.48930652001078684   steps: 338   lr: 1.6000000000000003e-05   evaluation
reward: 6.09
episode: 1625   score: 6.0   memory length: 358296   epsilon: 0.4885719400107822
steps: 371   lr: 1.6000000000000003e-05   evaluation reward: 6.11
episode: 1626   score: 6.0   memory length: 358654   epsilon: 0.4878631000107777
steps: 358   lr: 1.6000000000000003e-05   evaluation reward: 6.13
episode: 1627   score: 5.0   memory length: 358998   epsilon: 0.4871819800107734
steps: 344   lr: 1.6000000000000003e-05   evaluation reward: 6.12
episode: 1628   score: 10.0   memory length: 359406   epsilon:
0.4863741400107683   steps: 408   lr: 1.6000000000000003e-05   evaluation
reward: 6.16
episode: 1629   score: 3.0   memory length: 359635   epsilon: 0.4859207200107654
steps: 229   lr: 1.6000000000000003e-05   evaluation reward: 6.15
episode: 1630   score: 5.0   memory length: 359943   epsilon:
0.48531088001076156   steps: 308   lr: 1.6000000000000003e-05   evaluation
reward: 6.15
episode: 1631   score: 5.0   memory length: 360270   epsilon:
0.48466342001075746   steps: 327   lr: 1.6000000000000003e-05   evaluation
reward: 6.15
episode: 1632   score: 10.0   memory length: 360742   epsilon:
0.48372886001075155   steps: 472   lr: 1.6000000000000003e-05   evaluation

reward: 6.21
episode: 1633   score: 4.0   memory length: 361054   epsilon:
0.48311110001074764   steps: 312   lr: 1.6000000000000003e-05   evaluation
reward: 6.17
episode: 1634   score: 6.0   memory length: 361408   epsilon: 0.4824101800107432
steps: 354   lr: 1.6000000000000003e-05   evaluation reward: 6.14
episode: 1635   score: 5.0   memory length: 361730   epsilon:
0.48177262001073917   steps: 322   lr: 1.6000000000000003e-05   evaluation
reward: 6.13
episode: 1636   score: 7.0   memory length: 362146   epsilon:
0.48094894001073396   steps: 416   lr: 1.6000000000000003e-05   evaluation
reward: 6.15
episode: 1637   score: 5.0   memory length: 362456   epsilon: 0.4803351400107301
steps: 310   lr: 1.6000000000000003e-05   evaluation reward: 6.13
episode: 1638   score: 5.0   memory length: 362746   epsilon:
0.47976094001072644   steps: 290   lr: 1.6000000000000003e-05   evaluation
reward: 6.14
episode: 1639   score: 4.0   memory length: 363004   epsilon: 0.4792501000107232
steps: 258   lr: 1.6000000000000003e-05   evaluation reward: 6.12
episode: 1640   score: 6.0   memory length: 363351   epsilon:
0.47856304001071887   steps: 347   lr: 1.6000000000000003e-05   evaluation
reward: 6.08
episode: 1641   score: 9.0   memory length: 363801   epsilon:
0.47767204001071323   steps: 450   lr: 1.6000000000000003e-05   evaluation
reward: 6.09
episode: 1642   score: 9.0   memory length: 364275   epsilon: 0.4767335200107073
steps: 474   lr: 1.6000000000000003e-05   evaluation reward: 6.11
episode: 1643   score: 5.0   memory length: 364582   epsilon:
0.47612566001070344   steps: 307   lr: 1.6000000000000003e-05   evaluation
reward: 6.1
episode: 1644   score: 9.0   memory length: 365067   epsilon:
0.47516536001069737   steps: 485   lr: 1.6000000000000003e-05   evaluation
reward: 6.14
episode: 1645   score: 6.0   memory length: 365434   epsilon:
0.47443870001069277   steps: 367   lr: 1.6000000000000003e-05   evaluation
reward: 6.14
episode: 1646   score: 4.0   memory length: 365730   epsilon:
0.47385262001068906   steps: 296   lr: 1.6000000000000003e-05   evaluation
reward: 6.14
episode: 1647   score: 9.0   memory length: 366183   epsilon: 0.4729556800106834
steps: 453   lr: 1.6000000000000003e-05   evaluation reward: 6.13
episode: 1648   score: 2.0   memory length: 366381   epsilon: 0.4725636400106809
steps: 198   lr: 1.6000000000000003e-05   evaluation reward: 6.08
episode: 1649   score: 6.0   memory length: 366742   epsilon: 0.4718488600106764
steps: 361   lr: 1.6000000000000003e-05   evaluation reward: 6.08
episode: 1650   score: 3.0   memory length: 366953   epsilon:
0.47143108001067374   steps: 211   lr: 1.6000000000000003e-05   evaluation
reward: 6.07

episode: 1651   score: 4.0   memory length: 367230   epsilon:
0.47088262001067027   steps: 277   lr: 1.6000000000000003e-05   evaluation
reward: 6.03
episode: 1652   score: 5.0   memory length: 367557   epsilon: 0.4702351600106662
steps: 327   lr: 1.6000000000000003e-05   evaluation reward: 6.02
episode: 1653   score: 8.0   memory length: 367986   epsilon: 0.4693857400106608
steps: 429   lr: 1.6000000000000003e-05   evaluation reward: 6.02
episode: 1654   score: 7.0   memory length: 368392   epsilon: 0.4685818600106557
steps: 406   lr: 1.6000000000000003e-05   evaluation reward: 6.04
episode: 1655   score: 5.0   memory length: 368717   epsilon:
0.46793836001065164   steps: 325   lr: 1.6000000000000003e-05   evaluation
reward: 6.05
episode: 1656   score: 2.0   memory length: 368920   epsilon: 0.4675364200106491
steps: 203   lr: 1.6000000000000003e-05   evaluation reward: 6.04
episode: 1657   score: 5.0   memory length: 369226   epsilon:
0.46693054001064527   steps: 306   lr: 1.6000000000000003e-05   evaluation
reward: 6.04
episode: 1658   score: 7.0   memory length: 369599   epsilon: 0.4661920000106406
steps: 373   lr: 1.6000000000000003e-05   evaluation reward: 6.08
episode: 1659   score: 5.0   memory length: 369907   epsilon:
0.46558216001063674   steps: 308   lr: 1.6000000000000003e-05   evaluation
reward: 6.07
episode: 1660   score: 6.0   memory length: 370260   epsilon: 0.4648832200106323
steps: 353   lr: 1.6000000000000003e-05   evaluation reward: 6.01
episode: 1661   score: 4.0   memory length: 370536   epsilon:
0.46433674001062886   steps: 276   lr: 1.6000000000000003e-05   evaluation
reward: 6.0
episode: 1662   score: 5.0   memory length: 370897   epsilon:
0.46362196001062433   steps: 361   lr: 1.6000000000000003e-05   evaluation
reward: 6.0
episode: 1663   score: 7.0   memory length: 371280   epsilon:
0.46286362001061954   steps: 383   lr: 1.6000000000000003e-05   evaluation
reward: 6.05
episode: 1664   score: 5.0   memory length: 371566   epsilon:
0.46229734001061595   steps: 286   lr: 1.6000000000000003e-05   evaluation
reward: 6.09
episode: 1665   score: 3.0   memory length: 371812   epsilon:
0.46181026001061287   steps: 246   lr: 1.6000000000000003e-05   evaluation
reward: 6.08
episode: 1666   score: 5.0   memory length: 372159   epsilon: 0.4611232000106085
steps: 347   lr: 1.6000000000000003e-05   evaluation reward: 6.09
episode: 1667   score: 7.0   memory length: 372599   epsilon: 0.460252000010603
steps: 440   lr: 1.6000000000000003e-05   evaluation reward: 6.11
episode: 1668   score: 6.0   memory length: 372921   epsilon: 0.459614440010599
steps: 322   lr: 1.6000000000000003e-05   evaluation reward: 6.09
episode: 1669   score: 2.0   memory length: 373119   epsilon: 0.4592224000105965
steps: 198   lr: 1.6000000000000003e-05   evaluation reward: 6.06
episode: 1670   score: 5.0   memory length: 373451   epsilon:

0.45856504001059234 steps: 332 lr: 1.6000000000000003e-05 evaluation reward: 6.05
episode: 1671 score: 5.0 memory length: 373741 epsilon: 0.4579908400105887 steps: 290 lr: 1.6000000000000003e-05 evaluation reward: 6.04
episode: 1672 score: 4.0 memory length: 373999 epsilon: 0.4574800000105855 steps: 258 lr: 1.6000000000000003e-05 evaluation reward: 6.03
episode: 1673 score: 6.0 memory length: 374370 epsilon: 0.4567454200105808 steps: 371 lr: 1.6000000000000003e-05 evaluation reward: 6.02
episode: 1674 score: 5.0 memory length: 374695 epsilon: 0.45610192001057676 steps: 325 lr: 1.6000000000000003e-05 evaluation reward: 6.02
episode: 1675 score: 13.0 memory length: 375156 epsilon: 0.455189140010571 steps: 461 lr: 1.6000000000000003e-05 evaluation reward: 6.05
episode: 1676 score: 6.0 memory length: 375496 epsilon: 0.4545159400105667 steps: 340 lr: 1.6000000000000003e-05 evaluation reward: 5.99
episode: 1677 score: 6.0 memory length: 375852 epsilon: 0.45381106001056226 steps: 356 lr: 1.6000000000000003e-05 evaluation reward: 5.91
episode: 1678 score: 4.0 memory length: 376109 epsilon: 0.45330220001055904 steps: 257 lr: 1.6000000000000003e-05 evaluation reward: 5.89
episode: 1679 score: 7.0 memory length: 376483 epsilon: 0.45256168001055436 steps: 374 lr: 1.6000000000000003e-05 evaluation reward: 5.89
episode: 1680 score: 6.0 memory length: 376846 epsilon: 0.4518429400105498 steps: 363 lr: 1.6000000000000003e-05 evaluation reward: 5.91
episode: 1681 score: 5.0 memory length: 377154 epsilon: 0.45123310001054595 steps: 308 lr: 1.6000000000000003e-05 evaluation reward: 5.91
episode: 1682 score: 10.0 memory length: 377661 epsilon: 0.4502292400105396 steps: 507 lr: 1.6000000000000003e-05 evaluation reward: 5.93
episode: 1683 score: 4.0 memory length: 377939 epsilon: 0.4496788000105361 steps: 278 lr: 1.6000000000000003e-05 evaluation reward: 5.92
episode: 1684 score: 2.0 memory length: 378121 epsilon: 0.44931844001053384 steps: 182 lr: 1.6000000000000003e-05 evaluation reward: 5.9
episode: 1685 score: 6.0 memory length: 378460 epsilon: 0.4486472200105296 steps: 339 lr: 1.6000000000000003e-05 evaluation reward: 5.87
episode: 1686 score: 3.0 memory length: 378691 epsilon: 0.4481898400105267 steps: 231 lr: 1.6000000000000003e-05 evaluation reward: 5.82
episode: 1687 score: 10.0 memory length: 379064 epsilon: 0.447451300010522 steps: 373 lr: 1.6000000000000003e-05 evaluation reward: 5.89
episode: 1688 score: 7.0 memory length: 379449 epsilon: 0.4466890000105172 steps: 385 lr: 1.6000000000000003e-05 evaluation reward: 5.88
episode: 1689 score: 3.0 memory length: 379701 epsilon: 0.44619004001051404 steps: 252 lr: 1.6000000000000003e-05 evaluation reward: 5.83

episode: 1690   score: 2.0   memory length: 379903   epsilon: 0.4457900800105115
steps: 202   lr: 1.6000000000000003e-05   evaluation reward: 5.82
episode: 1691   score: 6.0   memory length: 380246   epsilon: 0.4451109400105072
steps: 343   lr: 1.6000000000000003e-05   evaluation reward: 5.83
episode: 1692   score: 2.0   memory length: 380444   epsilon:
0.44471890001050474   steps: 198   lr: 1.6000000000000003e-05   evaluation
reward: 5.83
episode: 1693   score: 5.0   memory length: 380754   epsilon:
0.44410510001050085   steps: 310   lr: 1.6000000000000003e-05   evaluation
reward: 5.82
episode: 1694   score: 6.0   memory length: 381111   epsilon: 0.4433982400104964
steps: 357   lr: 1.6000000000000003e-05   evaluation reward: 5.84
episode: 1695   score: 5.0   memory length: 381434   epsilon:
0.44275870001049233   steps: 323   lr: 1.6000000000000003e-05   evaluation
reward: 5.84
episode: 1696   score: 3.0   memory length: 381644   epsilon: 0.4423429000104897
steps: 210   lr: 1.6000000000000003e-05   evaluation reward: 5.82
episode: 1697   score: 4.0   memory length: 381921   epsilon:
0.44179444001048623   steps: 277   lr: 1.6000000000000003e-05   evaluation
reward: 5.8
episode: 1698   score: 7.0   memory length: 382326   epsilon:
0.44099254001048116   steps: 405   lr: 1.6000000000000003e-05   evaluation
reward: 5.8
episode: 1699   score: 6.0   memory length: 382650   epsilon: 0.4403510200104771
steps: 324   lr: 1.6000000000000003e-05   evaluation reward: 5.8
episode: 1700   score: 8.0   memory length: 383056   epsilon: 0.439547140010472
steps: 406   lr: 1.6000000000000003e-05   evaluation reward: 5.8
episode: 1701   score: 7.0   memory length: 383460   epsilon:
0.43874722001046695   steps: 404   lr: 1.6000000000000003e-05   evaluation
reward: 5.82
episode: 1702   score: 2.0   memory length: 383640   epsilon: 0.4383908200104647
steps: 180   lr: 1.6000000000000003e-05   evaluation reward: 5.78
episode: 1703   score: 6.0   memory length: 384020   epsilon:
0.43763842001045994   steps: 380   lr: 1.6000000000000003e-05   evaluation
reward: 5.75
episode: 1704   score: 9.0   memory length: 384460   epsilon: 0.4367672200104544
steps: 440   lr: 1.6000000000000003e-05   evaluation reward: 5.74
episode: 1705   score: 6.0   memory length: 384825   epsilon:
0.43604452001044985   steps: 365   lr: 1.6000000000000003e-05   evaluation
reward: 5.73
episode: 1706   score: 8.0   memory length: 385301   epsilon: 0.4351020400104439
steps: 476   lr: 1.6000000000000003e-05   evaluation reward: 5.73
episode: 1707   score: 6.0   memory length: 385636   epsilon: 0.4344387400104397
steps: 335   lr: 1.6000000000000003e-05   evaluation reward: 5.75
episode: 1708   score: 4.0   memory length: 385895   epsilon:
0.43392592001043645   steps: 259   lr: 1.6000000000000003e-05   evaluation
reward: 5.73
episode: 1709   score: 6.0   memory length: 386251   epsilon: 0.433221040010432

steps: 356 lr: 1.6000000000000003e-05 evaluation reward: 5.73
episode: 1710 score: 6.0 memory length: 386609 epsilon: 0.4325122000104275
steps: 358 lr: 1.6000000000000003e-05 evaluation reward: 5.71
episode: 1711 score: 6.0 memory length: 386946 epsilon: 0.4318449400104233
steps: 337 lr: 1.6000000000000003e-05 evaluation reward: 5.76
episode: 1712 score: 8.0 memory length: 387327 epsilon: 0.4310905600104185
steps: 381 lr: 1.6000000000000003e-05 evaluation reward: 5.79
episode: 1713 score: 5.0 memory length: 387669 epsilon: 0.4304134000104142
steps: 342 lr: 1.6000000000000003e-05 evaluation reward: 5.78
episode: 1714 score: 7.0 memory length: 388042 epsilon:
0.42967486001040955 steps: 373 lr: 1.6000000000000003e-05 evaluation
reward: 5.77
episode: 1715 score: 4.0 memory length: 388319 epsilon: 0.4291264000104061
steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 5.76
episode: 1716 score: 7.0 memory length: 388721 epsilon:
0.42833044001040105 steps: 402 lr: 1.6000000000000003e-05 evaluation
reward: 5.75
episode: 1717 score: 5.0 memory length: 389014 epsilon: 0.4277503000103974
steps: 293 lr: 1.6000000000000003e-05 evaluation reward: 5.73
episode: 1718 score: 9.0 memory length: 389517 epsilon: 0.4267543600103911
steps: 503 lr: 1.6000000000000003e-05 evaluation reward: 5.77
episode: 1719 score: 4.0 memory length: 389795 epsilon: 0.4262039200103876
steps: 278 lr: 1.6000000000000003e-05 evaluation reward: 5.74
episode: 1720 score: 5.0 memory length: 390121 epsilon: 0.4255584400103835
steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 5.73
episode: 1721 score: 6.0 memory length: 390495 epsilon: 0.4248179200103788
steps: 374 lr: 1.6000000000000003e-05 evaluation reward: 5.68
episode: 1722 score: 4.0 memory length: 390735 epsilon: 0.4243427200103758
steps: 240 lr: 1.6000000000000003e-05 evaluation reward: 5.67
episode: 1723 score: 4.0 memory length: 390995 epsilon:
0.42382792001037256 steps: 260 lr: 1.6000000000000003e-05 evaluation
reward: 5.63
episode: 1724 score: 10.0 memory length: 391369 epsilon:
0.4230874000103679 steps: 374 lr: 1.6000000000000003e-05 evaluation
reward: 5.67
episode: 1725 score: 5.0 memory length: 391675 epsilon:
0.42248152001036404 steps: 306 lr: 1.6000000000000003e-05 evaluation
reward: 5.66
episode: 1726 score: 7.0 memory length: 392094 epsilon: 0.4216519000103588
steps: 419 lr: 1.6000000000000003e-05 evaluation reward: 5.67
episode: 1727 score: 3.0 memory length: 392345 epsilon:
0.42115492001035565 steps: 251 lr: 1.6000000000000003e-05 evaluation
reward: 5.65
episode: 1728 score: 6.0 memory length: 392724 epsilon: 0.4204045000103509
steps: 379 lr: 1.6000000000000003e-05 evaluation reward: 5.61
episode: 1729 score: 5.0 memory length: 393046 epsilon:
0.41976694001034687 steps: 322 lr: 1.6000000000000003e-05 evaluation
reward: 5.63

episode: 1730 score: 6.0 memory length: 393389 epsilon: 0.41908780001034257 steps: 343 lr: 1.6000000000000003e-05 evaluation reward: 5.64
episode: 1731 score: 7.0 memory length: 393793 epsilon: 0.4182878800103375 steps: 404 lr: 1.6000000000000003e-05 evaluation reward: 5.66
episode: 1732 score: 9.0 memory length: 394258 epsilon: 0.4173671800103317 steps: 465 lr: 1.6000000000000003e-05 evaluation reward: 5.65
episode: 1733 score: 5.0 memory length: 394564 epsilon: 0.41676130001032785 steps: 306 lr: 1.6000000000000003e-05 evaluation reward: 5.66
episode: 1734 score: 10.0 memory length: 395091 epsilon: 0.41571784001032125 steps: 527 lr: 1.6000000000000003e-05 evaluation reward: 5.7
episode: 1735 score: 6.0 memory length: 395467 epsilon: 0.41497336001031654 steps: 376 lr: 1.6000000000000003e-05 evaluation reward: 5.71
episode: 1736 score: 10.0 memory length: 396017 epsilon: 0.41388436001030965 steps: 550 lr: 1.6000000000000003e-05 evaluation reward: 5.74
episode: 1737 score: 7.0 memory length: 396445 epsilon: 0.4130369200103043 steps: 428 lr: 1.6000000000000003e-05 evaluation reward: 5.76
episode: 1738 score: 4.0 memory length: 396705 epsilon: 0.41252212001030103 steps: 260 lr: 1.6000000000000003e-05 evaluation reward: 5.75
episode: 1739 score: 4.0 memory length: 396984 epsilon: 0.41196970001029753 steps: 279 lr: 1.6000000000000003e-05 evaluation reward: 5.75
episode: 1740 score: 5.0 memory length: 397275 epsilon: 0.4113935200102939 steps: 291 lr: 1.6000000000000003e-05 evaluation reward: 5.74
episode: 1741 score: 6.0 memory length: 397635 epsilon: 0.4106807200102894 steps: 360 lr: 1.6000000000000003e-05 evaluation reward: 5.71
episode: 1742 score: 5.0 memory length: 397979 epsilon: 0.40999960001028507 steps: 344 lr: 1.6000000000000003e-05 evaluation reward: 5.67
episode: 1743 score: 13.0 memory length: 398592 epsilon: 0.4087858600102774 steps: 613 lr: 1.6000000000000003e-05 evaluation reward: 5.75
episode: 1744 score: 7.0 memory length: 398981 epsilon: 0.4080156400102725 steps: 389 lr: 1.6000000000000003e-05 evaluation reward: 5.73
episode: 1745 score: 7.0 memory length: 399372 epsilon: 0.4072414600102676 steps: 391 lr: 1.6000000000000003e-05 evaluation reward: 5.74
episode: 1746 score: 3.0 memory length: 399602 epsilon: 0.40678606001026474 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 5.73
episode: 1747 score: 7.0 memory length: 400009 epsilon: 0.40598020001025964 steps: 407 lr: 6.400000000000001e-06 evaluation reward: 5.71
episode: 1748 score: 14.0 memory length: 400577 epsilon:

0.4048555600102525 steps: 568 lr: 6.400000000000001e-06 evaluation reward: 5.83
episode: 1749 score: 4.0 memory length: 400835 epsilon: 0.4043447200102493 steps: 258 lr: 6.400000000000001e-06 evaluation reward: 5.81
episode: 1750 score: 9.0 memory length: 401301 epsilon: 0.40342204001024345 steps: 466 lr: 6.400000000000001e-06 evaluation reward: 5.87
episode: 1751 score: 3.0 memory length: 401533 epsilon: 0.40296268001024055 steps: 232 lr: 6.400000000000001e-06 evaluation reward: 5.86
episode: 1752 score: 3.0 memory length: 401780 epsilon: 0.40247362001023745 steps: 247 lr: 6.400000000000001e-06 evaluation reward: 5.84
episode: 1753 score: 8.0 memory length: 402232 epsilon: 0.4015786600102318 steps: 452 lr: 6.400000000000001e-06 evaluation reward: 5.84
episode: 1754 score: 8.0 memory length: 402663 epsilon: 0.4007252800102264 steps: 431 lr: 6.400000000000001e-06 evaluation reward: 5.85
episode: 1755 score: 5.0 memory length: 402985 epsilon: 0.40008772001022236 steps: 322 lr: 6.400000000000001e-06 evaluation reward: 5.85
episode: 1756 score: 5.0 memory length: 403318 epsilon: 0.3994283800102182 steps: 333 lr: 6.400000000000001e-06 evaluation reward: 5.88
episode: 1757 score: 6.0 memory length: 403666 epsilon: 0.3987393400102138 steps: 348 lr: 6.400000000000001e-06 evaluation reward: 5.89
episode: 1758 score: 4.0 memory length: 403926 epsilon: 0.39822454001021057 steps: 260 lr: 6.400000000000001e-06 evaluation reward: 5.86
episode: 1759 score: 7.0 memory length: 404308 epsilon: 0.3974681800102058 steps: 382 lr: 6.400000000000001e-06 evaluation reward: 5.88
episode: 1760 score: 6.0 memory length: 404686 epsilon: 0.39671974001020105 steps: 378 lr: 6.400000000000001e-06 evaluation reward: 5.88
episode: 1761 score: 8.0 memory length: 405138 epsilon: 0.3958247800101954 steps: 452 lr: 6.400000000000001e-06 evaluation reward: 5.92
episode: 1762 score: 5.0 memory length: 405451 epsilon: 0.39520504001019147 steps: 313 lr: 6.400000000000001e-06 evaluation reward: 5.92
episode: 1763 score: 4.0 memory length: 405730 epsilon: 0.39465262001018797 steps: 279 lr: 6.400000000000001e-06 evaluation reward: 5.89
episode: 1764 score: 5.0 memory length: 406054 epsilon: 0.3940111000101839 steps: 324 lr: 6.400000000000001e-06 evaluation reward: 5.89
episode: 1765 score: 4.0 memory length: 406331 epsilon: 0.39346264001018044 steps: 277 lr: 6.400000000000001e-06 evaluation reward: 5.9
episode: 1766 score: 9.0 memory length: 406835 epsilon: 0.39246472001017413 steps: 504 lr: 6.400000000000001e-06 evaluation reward: 5.94

```
episode: 1767   score: 6.0   memory length: 407190   epsilon: 0.3917618200101697
steps: 355    lr: 6.400000000000001e-06     evaluation reward: 5.93
episode: 1768   score: 7.0   memory length: 407578   epsilon: 0.3909935800101648
steps: 388    lr: 6.400000000000001e-06     evaluation reward: 5.94
episode: 1769   score: 7.0   memory length: 408000   epsilon:
0.39015802001015953    steps: 422    lr: 6.400000000000001e-06     evaluation
reward: 5.99
episode: 1770   score: 10.0   memory length: 408426   epsilon:
0.3893145400101542    steps: 426    lr: 6.400000000000001e-06     evaluation
reward: 6.04
episode: 1771   score: 9.0   memory length: 408918   epsilon:
0.38834038001014803    steps: 492    lr: 6.400000000000001e-06     evaluation
reward: 6.08
episode: 1772   score: 3.0   memory length: 409147   epsilon:
0.38788696001014517    steps: 229    lr: 6.400000000000001e-06     evaluation
reward: 6.07
episode: 1773   score: 6.0   memory length: 409504   epsilon: 0.3871801000101407
steps: 357    lr: 6.400000000000001e-06     evaluation reward: 6.07
episode: 1774   score: 9.0   memory length: 410000   epsilon: 0.3861980200101345
steps: 496    lr: 6.400000000000001e-06     evaluation reward: 6.11
episode: 1775   score: 5.0   memory length: 410331   epsilon:
0.38554264001013033    steps: 331    lr: 6.400000000000001e-06     evaluation
reward: 6.03
episode: 1776   score: 13.0   memory length: 410811   epsilon:
0.3845922400101243    steps: 480    lr: 6.400000000000001e-06     evaluation
reward: 6.1
episode: 1777   score: 4.0   memory length: 411105   epsilon:
0.38401012001012064    steps: 294    lr: 6.400000000000001e-06     evaluation
reward: 6.08
episode: 1778   score: 5.0   memory length: 411402   epsilon: 0.3834220600101169
steps: 297    lr: 6.400000000000001e-06     evaluation reward: 6.09
episode: 1779   score: 7.0   memory length: 411794   epsilon: 0.382645900010112
steps: 392    lr: 6.400000000000001e-06     evaluation reward: 6.09
episode: 1780   score: 7.0   memory length: 412182   epsilon:
0.38187766001010715    steps: 388    lr: 6.400000000000001e-06     evaluation
reward: 6.1
episode: 1781   score: 4.0   memory length: 412460   epsilon:
0.38132722001010366    steps: 278    lr: 6.400000000000001e-06     evaluation
reward: 6.09
episode: 1782   score: 6.0   memory length: 412812   epsilon:
0.38063026001009925    steps: 352    lr: 6.400000000000001e-06     evaluation
reward: 6.05
episode: 1783   score: 6.0   memory length: 413184   epsilon: 0.3798937000100946
steps: 372    lr: 6.400000000000001e-06     evaluation reward: 6.07
episode: 1784   score: 6.0   memory length: 413541   epsilon: 0.3791868400100901
steps: 357    lr: 6.400000000000001e-06     evaluation reward: 6.11
episode: 1785   score: 18.0   memory length: 414047   epsilon:
0.3781849600100838    steps: 506    lr: 6.400000000000001e-06     evaluation
```

```
reward: 6.23
episode: 1786   score: 9.0   memory length: 414506   epsilon:
0.37727614001007803    steps: 459    lr: 6.400000000000001e-06     evaluation
reward: 6.29
episode: 1787   score: 6.0   memory length: 414841   epsilon:
0.37661284001007383    steps: 335    lr: 6.400000000000001e-06     evaluation
reward: 6.25
episode: 1788   score: 6.0   memory length: 415200   epsilon:
0.37590202001006934    steps: 359    lr: 6.400000000000001e-06     evaluation
reward: 6.24
episode: 1789   score: 7.0   memory length: 415603   epsilon: 0.3751040800100643
steps: 403    lr: 6.400000000000001e-06     evaluation reward: 6.28
episode: 1790   score: 7.0   memory length: 415994   epsilon: 0.3743299000100594
steps: 391    lr: 6.400000000000001e-06     evaluation reward: 6.33
episode: 1791   score: 6.0   memory length: 416337   epsilon: 0.3736507600100551
steps: 343    lr: 6.400000000000001e-06     evaluation reward: 6.33
episode: 1792   score: 10.0   memory length: 416785   epsilon:
0.3727637200100495    steps: 448    lr: 6.400000000000001e-06     evaluation
reward: 6.41
episode: 1793   score: 3.0   memory length: 417015   epsilon: 0.3723083200100466
steps: 230    lr: 6.400000000000001e-06     evaluation reward: 6.39
episode: 1794   score: 7.0   memory length: 417417   epsilon:
0.37151236001004156    steps: 402    lr: 6.400000000000001e-06     evaluation
reward: 6.4
episode: 1795   score: 8.0   memory length: 417857   epsilon:
0.37064116001003605    steps: 440    lr: 6.400000000000001e-06     evaluation
reward: 6.43
episode: 1796   score: 6.0   memory length: 418197   epsilon: 0.3699679600100318
steps: 340    lr: 6.400000000000001e-06     evaluation reward: 6.46
episode: 1797   score: 8.0   memory length: 418652   epsilon: 0.3690670600100261
steps: 455    lr: 6.400000000000001e-06     evaluation reward: 6.5
episode: 1798   score: 3.0   memory length: 418897   epsilon: 0.368581960010023
steps: 245    lr: 6.400000000000001e-06     evaluation reward: 6.46
episode: 1799   score: 9.0   memory length: 419370   epsilon: 0.3676454200100171
steps: 473    lr: 6.400000000000001e-06     evaluation reward: 6.49
episode: 1800   score: 10.0   memory length: 419732   epsilon:
0.36692866001001256    steps: 362    lr: 6.400000000000001e-06     evaluation
reward: 6.51
episode: 1801   score: 6.0   memory length: 420070   epsilon:
0.36625942001000833    steps: 338    lr: 6.400000000000001e-06     evaluation
reward: 6.5
episode: 1802   score: 6.0   memory length: 420408   epsilon: 0.3655901800100041
steps: 338    lr: 6.400000000000001e-06     evaluation reward: 6.54
episode: 1803   score: 8.0   memory length: 420864   epsilon: 0.3646873000099984
steps: 456    lr: 6.400000000000001e-06     evaluation reward: 6.56
episode: 1804   score: 6.0   memory length: 421187   epsilon:
0.36404776000999434    steps: 323    lr: 6.400000000000001e-06     evaluation
reward: 6.53
```

episode: 1805   score: 7.0   memory length: 421544   epsilon:
0.36334090000998986   steps: 357   lr: 6.400000000000001e-06   evaluation
reward: 6.54
episode: 1806   score: 7.0   memory length: 421918   epsilon: 0.3626003800099852
steps: 374   lr: 6.400000000000001e-06   evaluation reward: 6.53
episode: 1807   score: 6.0   memory length: 422298   epsilon: 0.3618479800099804
steps: 380   lr: 6.400000000000001e-06   evaluation reward: 6.53
episode: 1808   score: 7.0   memory length: 422721   epsilon: 0.3610104400099751
steps: 423   lr: 6.400000000000001e-06   evaluation reward: 6.56
episode: 1809   score: 6.0   memory length: 423059   epsilon: 0.3603412000099709
steps: 338   lr: 6.400000000000001e-06   evaluation reward: 6.56
episode: 1810   score: 7.0   memory length: 423446   epsilon:
0.35957494000996604   steps: 387   lr: 6.400000000000001e-06   evaluation
reward: 6.57
episode: 1811   score: 19.0   memory length: 424204   epsilon:
0.35807410000995654   steps: 758   lr: 6.400000000000001e-06   evaluation
reward: 6.7
episode: 1812   score: 10.0   memory length: 424616   epsilon:
0.3572583400099514   steps: 412   lr: 6.400000000000001e-06   evaluation
reward: 6.72
episode: 1813   score: 6.0   memory length: 424996   epsilon: 0.3565059400099466
steps: 380   lr: 6.400000000000001e-06   evaluation reward: 6.73
episode: 1814   score: 5.0   memory length: 425320   epsilon:
0.35586442000994256   steps: 324   lr: 6.400000000000001e-06   evaluation
reward: 6.71
episode: 1815   score: 5.0   memory length: 425645   epsilon: 0.3552209200099385
steps: 325   lr: 6.400000000000001e-06   evaluation reward: 6.72
episode: 1816   score: 8.0   memory length: 426083   epsilon: 0.354353680009933
steps: 438   lr: 6.400000000000001e-06   evaluation reward: 6.73
episode: 1817   score: 6.0   memory length: 426439   epsilon:
0.35364880000992854   steps: 356   lr: 6.400000000000001e-06   evaluation
reward: 6.74
episode: 1818   score: 3.0   memory length: 426670   epsilon:
0.35319142000992565   steps: 231   lr: 6.400000000000001e-06   evaluation
reward: 6.68
episode: 1819   score: 4.0   memory length: 426932   epsilon:
0.35267266000992237   steps: 262   lr: 6.400000000000001e-06   evaluation
reward: 6.68
episode: 1820   score: 8.0   memory length: 427394   epsilon: 0.3517579000099166
steps: 462   lr: 6.400000000000001e-06   evaluation reward: 6.71
episode: 1821   score: 6.0   memory length: 427751   epsilon: 0.3510510400099121
steps: 357   lr: 6.400000000000001e-06   evaluation reward: 6.71
episode: 1822   score: 5.0   memory length: 428076   epsilon:
0.35040754000990804   steps: 325   lr: 6.400000000000001e-06   evaluation
reward: 6.72
episode: 1823   score: 6.0   memory length: 428404   epsilon: 0.3497581000099039
steps: 328   lr: 6.400000000000001e-06   evaluation reward: 6.74
episode: 1824   score: 8.0   memory length: 428858   epsilon:

0.34885918000989824 steps: 454 lr: 6.400000000000001e-06 evaluation reward: 6.72
episode: 1825 score: 15.0 memory length: 429414 epsilon: 0.3477583000098913 steps: 556 lr: 6.400000000000001e-06 evaluation reward: 6.82
episode: 1826 score: 8.0 memory length: 429830 epsilon: 0.34693462000988606 steps: 416 lr: 6.400000000000001e-06 evaluation reward: 6.83
episode: 1827 score: 9.0 memory length: 430289 epsilon: 0.3460258000098803 steps: 459 lr: 6.400000000000001e-06 evaluation reward: 6.89
episode: 1828 score: 10.0 memory length: 430783 epsilon: 0.3450476800098741 steps: 494 lr: 6.400000000000001e-06 evaluation reward: 6.93
episode: 1829 score: 8.0 memory length: 431229 epsilon: 0.34416460000986854 steps: 446 lr: 6.400000000000001e-06 evaluation reward: 6.96
episode: 1830 score: 7.0 memory length: 431634 epsilon: 0.34336270000986346 steps: 405 lr: 6.400000000000001e-06 evaluation reward: 6.97
episode: 1831 score: 5.0 memory length: 431980 epsilon: 0.34267762000985913 steps: 346 lr: 6.400000000000001e-06 evaluation reward: 6.95
episode: 1832 score: 8.0 memory length: 432428 epsilon: 0.3417905800098535 steps: 448 lr: 6.400000000000001e-06 evaluation reward: 6.94
episode: 1833 score: 10.0 memory length: 432905 epsilon: 0.34084612000984754 steps: 477 lr: 6.400000000000001e-06 evaluation reward: 6.99
episode: 1834 score: 10.0 memory length: 433423 epsilon: 0.33982048000984105 steps: 518 lr: 6.400000000000001e-06 evaluation reward: 6.99
episode: 1835 score: 16.0 memory length: 434158 epsilon: 0.33836518000983185 steps: 735 lr: 6.400000000000001e-06 evaluation reward: 7.09
episode: 1836 score: 8.0 memory length: 434582 epsilon: 0.33752566000982653 steps: 424 lr: 6.400000000000001e-06 evaluation reward: 7.07
episode: 1837 score: 6.0 memory length: 434938 epsilon: 0.3368207800098221 steps: 356 lr: 6.400000000000001e-06 evaluation reward: 7.06
episode: 1838 score: 8.0 memory length: 435410 epsilon: 0.33588622000981616 steps: 472 lr: 6.400000000000001e-06 evaluation reward: 7.1
episode: 1839 score: 8.0 memory length: 435889 epsilon: 0.33493780000981016 steps: 479 lr: 6.400000000000001e-06 evaluation reward: 7.14
episode: 1840 score: 5.0 memory length: 436217 epsilon: 0.33428836000980605 steps: 328 lr: 6.400000000000001e-06 evaluation reward: 7.14
episode: 1841 score: 7.0 memory length: 436586 epsilon:

0.33355774000980143 steps: 369 lr: 6.400000000000001e-06 evaluation reward: 7.15
episode: 1842 score: 7.0 memory length: 437011 epsilon: 0.3327162400097961 steps: 425 lr: 6.400000000000001e-06 evaluation reward: 7.17
episode: 1843 score: 9.0 memory length: 437520 epsilon: 0.33170842000978973 steps: 509 lr: 6.400000000000001e-06 evaluation reward: 7.13
episode: 1844 score: 5.0 memory length: 437864 epsilon: 0.3310273000097854 steps: 344 lr: 6.400000000000001e-06 evaluation reward: 7.11
episode: 1845 score: 10.0 memory length: 438374 epsilon: 0.33001750000977903 steps: 510 lr: 6.400000000000001e-06 evaluation reward: 7.14
episode: 1846 score: 7.0 memory length: 438768 epsilon: 0.3292373800097741 steps: 394 lr: 6.400000000000001e-06 evaluation reward: 7.18
episode: 1847 score: 7.0 memory length: 439199 epsilon: 0.3283840000097687 steps: 431 lr: 6.400000000000001e-06 evaluation reward: 7.18
episode: 1848 score: 15.0 memory length: 439761 epsilon: 0.32727124000976165 steps: 562 lr: 6.400000000000001e-06 evaluation reward: 7.19
episode: 1849 score: 8.0 memory length: 440166 epsilon: 0.3264693400097566 steps: 405 lr: 6.400000000000001e-06 evaluation reward: 7.23
episode: 1850 score: 9.0 memory length: 440657 epsilon: 0.32549716000975043 steps: 491 lr: 6.400000000000001e-06 evaluation reward: 7.23
episode: 1851 score: 5.0 memory length: 440965 epsilon: 0.32488732000974657 steps: 308 lr: 6.400000000000001e-06 evaluation reward: 7.25
episode: 1852 score: 6.0 memory length: 441303 epsilon: 0.32421808000974234 steps: 338 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1853 score: 9.0 memory length: 441811 epsilon: 0.323212240009736 steps: 508 lr: 6.400000000000001e-06 evaluation reward: 7.29
episode: 1854 score: 3.0 memory length: 442056 epsilon: 0.3227271400097329 steps: 245 lr: 6.400000000000001e-06 evaluation reward: 7.24
episode: 1855 score: 8.0 memory length: 442481 epsilon: 0.3218856400097276 steps: 425 lr: 6.400000000000001e-06 evaluation reward: 7.27
episode: 1856 score: 6.0 memory length: 442816 epsilon: 0.3212223400097234 steps: 335 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1857 score: 13.0 memory length: 443290 epsilon: 0.32028382000971745 steps: 474 lr: 6.400000000000001e-06 evaluation reward: 7.35
episode: 1858 score: 7.0 memory length: 443714 epsilon: 0.31944430000971213 steps: 424 lr: 6.400000000000001e-06 evaluation reward: 7.38
episode: 1859 score: 11.0 memory length: 444265 epsilon: 0.31835332000970523 steps: 551 lr: 6.400000000000001e-06 evaluation reward: 7.42
episode: 1860 score: 3.0 memory length: 444496 epsilon:

0.31789594000970234 steps: 231 lr: 6.400000000000001e-06 evaluation reward: 7.39
episode: 1861 score: 4.0 memory length: 444758 epsilon: 0.31737718000969906 steps: 262 lr: 6.400000000000001e-06 evaluation reward: 7.35
episode: 1862 score: 7.0 memory length: 445138 epsilon: 0.3166247800096943 steps: 380 lr: 6.400000000000001e-06 evaluation reward: 7.37
episode: 1863 score: 7.0 memory length: 445546 epsilon: 0.3158169400096892 steps: 408 lr: 6.400000000000001e-06 evaluation reward: 7.4
episode: 1864 score: 6.0 memory length: 445887 epsilon: 0.3151417600096849 steps: 341 lr: 6.400000000000001e-06 evaluation reward: 7.41
episode: 1865 score: 9.0 memory length: 446331 epsilon: 0.31426264000967935 steps: 444 lr: 6.400000000000001e-06 evaluation reward: 7.46
episode: 1866 score: 6.0 memory length: 446689 epsilon: 0.31355380000967487 steps: 358 lr: 6.400000000000001e-06 evaluation reward: 7.43
episode: 1867 score: 7.0 memory length: 447059 epsilon: 0.31282120000967023 steps: 370 lr: 6.400000000000001e-06 evaluation reward: 7.44
episode: 1868 score: 1.0 memory length: 447210 epsilon: 0.31252222000966834 steps: 151 lr: 6.400000000000001e-06 evaluation reward: 7.38
episode: 1869 score: 4.0 memory length: 447475 epsilon: 0.311997520009665 steps: 265 lr: 6.400000000000001e-06 evaluation reward: 7.35
episode: 1870 score: 7.0 memory length: 447861 epsilon: 0.3112332400096602 steps: 386 lr: 6.400000000000001e-06 evaluation reward: 7.32
episode: 1871 score: 9.0 memory length: 448364 epsilon: 0.3102373000096539 steps: 503 lr: 6.400000000000001e-06 evaluation reward: 7.32
episode: 1872 score: 8.0 memory length: 448746 epsilon: 0.3094809400096491 steps: 382 lr: 6.400000000000001e-06 evaluation reward: 7.37
episode: 1873 score: 14.0 memory length: 449238 epsilon: 0.30850678000964293 steps: 492 lr: 6.400000000000001e-06 evaluation reward: 7.45
episode: 1874 score: 8.0 memory length: 449642 epsilon: 0.3077068600096379 steps: 404 lr: 6.400000000000001e-06 evaluation reward: 7.44
episode: 1875 score: 7.0 memory length: 450063 epsilon: 0.3068732800096326 steps: 421 lr: 6.400000000000001e-06 evaluation reward: 7.46
episode: 1876 score: 6.0 memory length: 450453 epsilon: 0.3061010800096277 steps: 390 lr: 6.400000000000001e-06 evaluation reward: 7.39
episode: 1877 score: 6.0 memory length: 450810 epsilon: 0.30539422000962324 steps: 357 lr: 6.400000000000001e-06 evaluation reward: 7.41
episode: 1878 score: 6.0 memory length: 451165 epsilon: 0.3046913200096188 steps: 355 lr: 6.400000000000001e-06 evaluation reward: 7.42
episode: 1879 score: 8.0 memory length: 451550 epsilon: 0.30392902000961397 steps: 385 lr: 6.400000000000001e-06 evaluation reward: 7.43

episode: 1880 score: 9.0 memory length: 452050 epsilon: 0.3029390200096077 steps: 500 lr: 6.400000000000001e-06 evaluation reward: 7.45
episode: 1881 score: 6.0 memory length: 452438 epsilon: 0.30217078000960285 steps: 388 lr: 6.400000000000001e-06 evaluation reward: 7.47
episode: 1882 score: 5.0 memory length: 452742 epsilon: 0.30156886000959904 steps: 304 lr: 6.400000000000001e-06 evaluation reward: 7.46
episode: 1883 score: 5.0 memory length: 453029 epsilon: 0.30100060000959544 steps: 287 lr: 6.400000000000001e-06 evaluation reward: 7.45
episode: 1884 score: 11.0 memory length: 453576 epsilon: 0.2999175400095886 steps: 547 lr: 6.400000000000001e-06 evaluation reward: 7.5
episode: 1885 score: 5.0 memory length: 453886 epsilon: 0.2993037400095847 steps: 310 lr: 6.400000000000001e-06 evaluation reward: 7.37
episode: 1886 score: 3.0 memory length: 454117 epsilon: 0.2988463600095818 steps: 231 lr: 6.400000000000001e-06 evaluation reward: 7.31
episode: 1887 score: 8.0 memory length: 454567 epsilon: 0.2979553600095762 steps: 450 lr: 6.400000000000001e-06 evaluation reward: 7.33
episode: 1888 score: 7.0 memory length: 455000 epsilon: 0.29709802000957075 steps: 433 lr: 6.400000000000001e-06 evaluation reward: 7.34
episode: 1889 score: 8.0 memory length: 455426 epsilon: 0.2962545400095654 steps: 426 lr: 6.400000000000001e-06 evaluation reward: 7.35
episode: 1890 score: 10.0 memory length: 455904 epsilon: 0.2953081000095594 steps: 478 lr: 6.400000000000001e-06 evaluation reward: 7.38
episode: 1891 score: 5.0 memory length: 456193 epsilon: 0.2947358800095558 steps: 289 lr: 6.400000000000001e-06 evaluation reward: 7.37
episode: 1892 score: 4.0 memory length: 456472 epsilon: 0.2941834600095523 steps: 279 lr: 6.400000000000001e-06 evaluation reward: 7.31
episode: 1893 score: 10.0 memory length: 456871 epsilon: 0.2933934400095473 steps: 399 lr: 6.400000000000001e-06 evaluation reward: 7.38
episode: 1894 score: 7.0 memory length: 457259 epsilon: 0.29262520000954245 steps: 388 lr: 6.400000000000001e-06 evaluation reward: 7.38
episode: 1895 score: 7.0 memory length: 457612 epsilon: 0.29192626000953803 steps: 353 lr: 6.400000000000001e-06 evaluation reward: 7.37
episode: 1896 score: 9.0 memory length: 458041 epsilon: 0.29107684000953266 steps: 429 lr: 6.400000000000001e-06 evaluation reward: 7.4
episode: 1897 score: 6.0 memory length: 458362 epsilon: 0.29044126000952863 steps: 321 lr: 6.400000000000001e-06 evaluation reward: 7.38
episode: 1898 score: 7.0 memory length: 458765 epsilon: 0.2896433200095236

steps: 403   lr: 6.400000000000001e-06   evaluation reward: 7.42
episode: 1899   score: 11.0   memory length: 459170   epsilon:
0.2888414200095185   steps: 405   lr: 6.400000000000001e-06   evaluation
reward: 7.44
episode: 1900   score: 4.0   memory length: 459447   epsilon:
0.28829296000951504   steps: 277   lr: 6.400000000000001e-06   evaluation
reward: 7.38
episode: 1901   score: 9.0   memory length: 459891   epsilon: 0.2874138400095095
steps: 444   lr: 6.400000000000001e-06   evaluation reward: 7.41
episode: 1902   score: 11.0   memory length: 460448   epsilon:
0.2863109800095025   steps: 557   lr: 6.400000000000001e-06   evaluation
reward: 7.46
episode: 1903   score: 9.0   memory length: 460933   epsilon: 0.2853506800094964
steps: 485   lr: 6.400000000000001e-06   evaluation reward: 7.47
episode: 1904   score: 5.0   memory length: 461261   epsilon: 0.2847012400094923
steps: 328   lr: 6.400000000000001e-06   evaluation reward: 7.46
episode: 1905   score: 5.0   memory length: 461588   epsilon: 0.2840537800094882
steps: 327   lr: 6.400000000000001e-06   evaluation reward: 7.44
episode: 1906   score: 9.0   memory length: 462035   epsilon: 0.2831687200094826
steps: 447   lr: 6.400000000000001e-06   evaluation reward: 7.46
episode: 1907   score: 3.0   memory length: 462266   epsilon:
0.28271134000947973   steps: 231   lr: 6.400000000000001e-06   evaluation
reward: 7.43
episode: 1908   score: 6.0   memory length: 462620   epsilon: 0.2820104200094753
steps: 354   lr: 6.400000000000001e-06   evaluation reward: 7.42
episode: 1909   score: 7.0   memory length: 463024   epsilon:
0.28121050000947023   steps: 404   lr: 6.400000000000001e-06   evaluation
reward: 7.43
episode: 1910   score: 10.0   memory length: 463522   epsilon: 0.280224460009464
steps: 498   lr: 6.400000000000001e-06   evaluation reward: 7.46
episode: 1911   score: 6.0   memory length: 463860   epsilon:
0.27955522000945976   steps: 338   lr: 6.400000000000001e-06   evaluation
reward: 7.33
episode: 1912   score: 5.0   memory length: 464168   epsilon: 0.2789453800094559
steps: 308   lr: 6.400000000000001e-06   evaluation reward: 7.28
episode: 1913   score: 4.0   memory length: 464426   epsilon:
0.27843454000945267   steps: 258   lr: 6.400000000000001e-06   evaluation
reward: 7.26
episode: 1914   score: 7.0   memory length: 464811   epsilon:
0.27767224000944785   steps: 385   lr: 6.400000000000001e-06   evaluation
reward: 7.28
episode: 1915   score: 9.0   memory length: 465266   epsilon:
0.27677134000944215   steps: 455   lr: 6.400000000000001e-06   evaluation
reward: 7.32
episode: 1916   score: 5.0   memory length: 465593   epsilon:
0.27612388000943805   steps: 327   lr: 6.400000000000001e-06   evaluation
reward: 7.29
episode: 1917   score: 3.0   memory length: 465824   epsilon:

0.27566650000943516 steps: 231 lr: 6.400000000000001e-06 evaluation reward: 7.26
episode: 1918 score: 3.0 memory length: 466055 epsilon: 0.27520912000943226 steps: 231 lr: 6.400000000000001e-06 evaluation reward: 7.26
episode: 1919 score: 9.0 memory length: 466556 epsilon: 0.274217140009426 steps: 501 lr: 6.400000000000001e-06 evaluation reward: 7.31
episode: 1920 score: 5.0 memory length: 466827 epsilon: 0.2736805600094226 steps: 271 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1921 score: 9.0 memory length: 467296 epsilon: 0.2727519400094167 steps: 469 lr: 6.400000000000001e-06 evaluation reward: 7.31
episode: 1922 score: 12.0 memory length: 467736 epsilon: 0.2718807400094112 steps: 440 lr: 6.400000000000001e-06 evaluation reward: 7.38
episode: 1923 score: 11.0 memory length: 468287 epsilon: 0.2707897600094043 steps: 551 lr: 6.400000000000001e-06 evaluation reward: 7.43
episode: 1924 score: 11.0 memory length: 468839 epsilon: 0.2696968000093974 steps: 552 lr: 6.400000000000001e-06 evaluation reward: 7.46
episode: 1925 score: 6.0 memory length: 469181 epsilon: 0.2690196400093931 steps: 342 lr: 6.400000000000001e-06 evaluation reward: 7.37
episode: 1926 score: 6.0 memory length: 469538 epsilon: 0.26831278000938863 steps: 357 lr: 6.400000000000001e-06 evaluation reward: 7.35
episode: 1927 score: 3.0 memory length: 469769 epsilon: 0.26785540000938574 steps: 231 lr: 6.400000000000001e-06 evaluation reward: 7.29
episode: 1928 score: 7.0 memory length: 470173 epsilon: 0.2670554800093807 steps: 404 lr: 6.400000000000001e-06 evaluation reward: 7.26
episode: 1929 score: 6.0 memory length: 470511 epsilon: 0.26638624000937644 steps: 338 lr: 6.400000000000001e-06 evaluation reward: 7.24
episode: 1930 score: 8.0 memory length: 470930 epsilon: 0.2655566200093712 steps: 419 lr: 6.400000000000001e-06 evaluation reward: 7.25
episode: 1931 score: 7.0 memory length: 471335 epsilon: 0.2647547200093661 steps: 405 lr: 6.400000000000001e-06 evaluation reward: 7.27
episode: 1932 score: 11.0 memory length: 471888 epsilon: 0.2636597800093592 steps: 553 lr: 6.400000000000001e-06 evaluation reward: 7.3
episode: 1933 score: 11.0 memory length: 472432 epsilon: 0.2625826600093524 steps: 544 lr: 6.400000000000001e-06 evaluation reward: 7.31
episode: 1934 score: 4.0 memory length: 472710 epsilon: 0.2620322200093489 steps: 278 lr: 6.400000000000001e-06 evaluation reward: 7.25
episode: 1935 score: 2.0 memory length: 472892 epsilon: 0.2616718600093466 steps: 182 lr: 6.400000000000001e-06 evaluation reward: 7.11
episode: 1936 score: 6.0 memory length: 473267 epsilon: 0.2609293600093419

steps: 375   lr: 6.400000000000001e-06   evaluation reward: 7.09
episode: 1937   score: 10.0   memory length: 473779   epsilon:
0.2599156000093355   steps: 512   lr: 6.400000000000001e-06   evaluation
reward: 7.13
episode: 1938   score: 4.0   memory length: 474077   epsilon:
0.25932556000933177   steps: 298   lr: 6.400000000000001e-06   evaluation
reward: 7.09
episode: 1939   score: 9.0   memory length: 474520   epsilon: 0.2584484200093262
steps: 443   lr: 6.400000000000001e-06   evaluation reward: 7.1
episode: 1940   score: 8.0   memory length: 474996   epsilon:
0.25750594000932026   steps: 476   lr: 6.400000000000001e-06   evaluation
reward: 7.13
episode: 1941   score: 3.0   memory length: 475227   epsilon:
0.25704856000931736   steps: 231   lr: 6.400000000000001e-06   evaluation
reward: 7.09
episode: 1942   score: 6.0   memory length: 475561   epsilon: 0.2563872400093132
steps: 334   lr: 6.400000000000001e-06   evaluation reward: 7.08
episode: 1943   score: 10.0   memory length: 476075   epsilon:
0.25536952000930674   steps: 514   lr: 6.400000000000001e-06   evaluation
reward: 7.09
episode: 1944   score: 3.0   memory length: 476304   epsilon:
0.25491610000930387   steps: 229   lr: 6.400000000000001e-06   evaluation
reward: 7.07
episode: 1945   score: 8.0   memory length: 476771   epsilon: 0.253991440009298
steps: 467   lr: 6.400000000000001e-06   evaluation reward: 7.05
episode: 1946   score: 8.0   memory length: 477212   epsilon: 0.2531182600092925
steps: 441   lr: 6.400000000000001e-06   evaluation reward: 7.06
episode: 1947   score: 5.0   memory length: 477500   epsilon: 0.2525480200092889
steps: 288   lr: 6.400000000000001e-06   evaluation reward: 7.04
episode: 1948   score: 2.0   memory length: 477682   epsilon: 0.2521876600092866
steps: 182   lr: 6.400000000000001e-06   evaluation reward: 6.91
episode: 1949   score: 5.0   memory length: 477988   epsilon: 0.2515817800092828
steps: 306   lr: 6.400000000000001e-06   evaluation reward: 6.88
episode: 1950   score: 6.0   memory length: 478344   epsilon: 0.2508769000092783
steps: 356   lr: 6.400000000000001e-06   evaluation reward: 6.85
episode: 1951   score: 11.0   memory length: 478910   epsilon:
0.24975622000927122   steps: 566   lr: 6.400000000000001e-06   evaluation
reward: 6.91
episode: 1952   score: 9.0   memory length: 479358   epsilon: 0.2488691800092656
steps: 448   lr: 6.400000000000001e-06   evaluation reward: 6.94
episode: 1953   score: 6.0   memory length: 479702   epsilon: 0.2481880600092613
steps: 344   lr: 6.400000000000001e-06   evaluation reward: 6.91
episode: 1954   score: 10.0   memory length: 480112   epsilon:
0.24737626000925617   steps: 410   lr: 6.400000000000001e-06   evaluation
reward: 6.98
episode: 1955   score: 10.0   memory length: 480543   epsilon:
0.24652288000925077   steps: 431   lr: 6.400000000000001e-06   evaluation
reward: 7.0

episode: 1956 score: 16.0 memory length: 481265 epsilon: 0.24509332000924172 steps: 722 lr: 6.400000000000001e-06 evaluation reward: 7.1
episode: 1957 score: 5.0 memory length: 481576 epsilon: 0.24447754000923783 steps: 311 lr: 6.400000000000001e-06 evaluation reward: 7.02
episode: 1958 score: 2.0 memory length: 481758 epsilon: 0.24411718000923555 steps: 182 lr: 6.400000000000001e-06 evaluation reward: 6.97
episode: 1959 score: 17.0 memory length: 482368 epsilon: 0.2429093800092279 steps: 610 lr: 6.400000000000001e-06 evaluation reward: 7.03
episode: 1960 score: 8.0 memory length: 482773 epsilon: 0.24210748000922283 steps: 405 lr: 6.400000000000001e-06 evaluation reward: 7.08
episode: 1961 score: 4.0 memory length: 483050 epsilon: 0.24155902000921936 steps: 277 lr: 6.400000000000001e-06 evaluation reward: 7.08
episode: 1962 score: 6.0 memory length: 483426 epsilon: 0.24081454000921465 steps: 376 lr: 6.400000000000001e-06 evaluation reward: 7.07
episode: 1963 score: 13.0 memory length: 483980 epsilon: 0.2397176200092077 steps: 554 lr: 6.400000000000001e-06 evaluation reward: 7.13
episode: 1964 score: 14.0 memory length: 484500 epsilon: 0.2386880200092012 steps: 520 lr: 6.400000000000001e-06 evaluation reward: 7.21
episode: 1965 score: 10.0 memory length: 485020 epsilon: 0.23765842000919468 steps: 520 lr: 6.400000000000001e-06 evaluation reward: 7.22
episode: 1966 score: 13.0 memory length: 485573 epsilon: 0.23656348000918775 steps: 553 lr: 6.400000000000001e-06 evaluation reward: 7.29
episode: 1967 score: 11.0 memory length: 486116 epsilon: 0.23548834000918095 steps: 543 lr: 6.400000000000001e-06 evaluation reward: 7.33
episode: 1968 score: 11.0 memory length: 486640 epsilon: 0.2344508200091744 steps: 524 lr: 6.400000000000001e-06 evaluation reward: 7.43
episode: 1969 score: 7.0 memory length: 487062 epsilon: 0.2336152600091691 steps: 422 lr: 6.400000000000001e-06 evaluation reward: 7.46
episode: 1970 score: 9.0 memory length: 487543 epsilon: 0.23266288000916308 steps: 481 lr: 6.400000000000001e-06 evaluation reward: 7.48
episode: 1971 score: 10.0 memory length: 488015 epsilon: 0.23172832000915716 steps: 472 lr: 6.400000000000001e-06 evaluation reward: 7.49
episode: 1972 score: 7.0 memory length: 488422 epsilon:

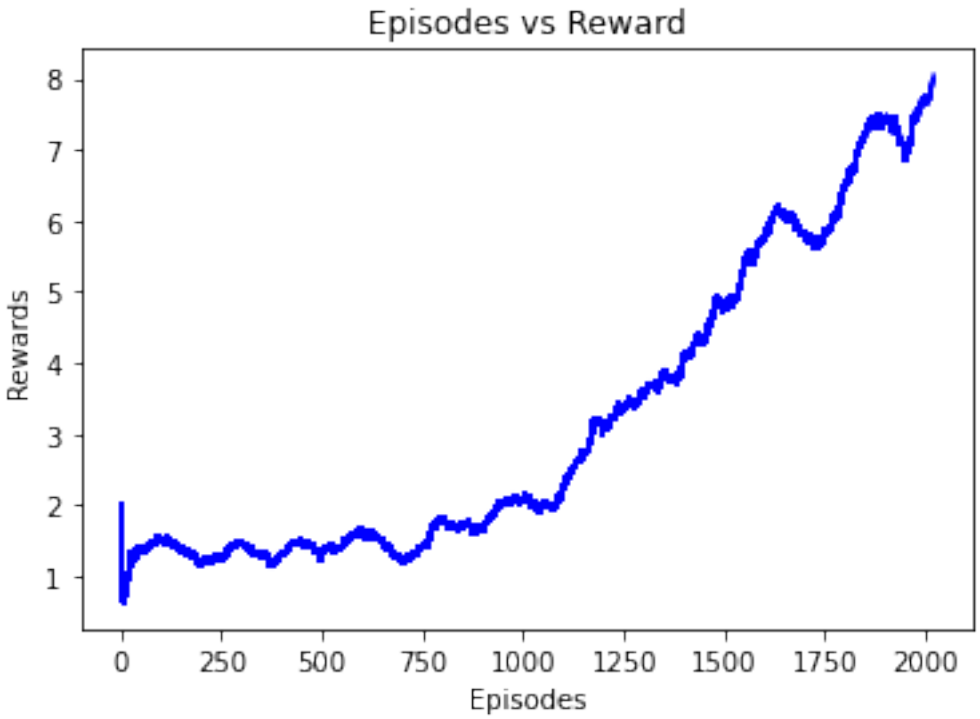
```
0.23092246000915206    steps: 407    lr: 6.400000000000001e-06     evaluation
reward: 7.48
episode: 1973   score: 5.0   memory length: 488729   epsilon:
0.23031460000914822    steps: 307    lr: 6.400000000000001e-06     evaluation
reward: 7.39
episode: 1974   score: 8.0   memory length: 489117   epsilon:
0.22954636000914336    steps: 388    lr: 6.400000000000001e-06     evaluation
reward: 7.39
episode: 1975   score: 8.0   memory length: 489570   epsilon:
0.22864942000913768    steps: 453    lr: 6.400000000000001e-06     evaluation
reward: 7.4
episode: 1976   score: 13.0   memory length: 490080   epsilon:
0.2276396200091313    steps: 510    lr: 6.400000000000001e-06     evaluation
reward: 7.47
episode: 1977   score: 9.0   memory length: 490533   epsilon:
0.22674268000912562    steps: 453    lr: 6.400000000000001e-06     evaluation
reward: 7.5
episode: 1978   score: 14.0   memory length: 491066   epsilon:
0.22568734000911894    steps: 533    lr: 6.400000000000001e-06     evaluation
reward: 7.58
episode: 1979   score: 3.0   memory length: 491298   epsilon:
0.22522798000911604    steps: 232    lr: 6.400000000000001e-06     evaluation
reward: 7.53
episode: 1980   score: 6.0   memory length: 491651   epsilon: 0.2245290400091116
steps: 353    lr: 6.400000000000001e-06     evaluation reward: 7.5
episode: 1981   score: 5.0   memory length: 491997   epsilon:
0.22384396000910728    steps: 346    lr: 6.400000000000001e-06     evaluation
reward: 7.49
episode: 1982   score: 8.0   memory length: 492431   epsilon:
0.22298464000910184    steps: 434    lr: 6.400000000000001e-06     evaluation
reward: 7.52
episode: 1983   score: 11.0   memory length: 492986   epsilon:
0.2218857400090949    steps: 555    lr: 6.400000000000001e-06     evaluation
reward: 7.58
episode: 1984   score: 10.0   memory length: 493506   epsilon:
0.22085614000908838    steps: 520    lr: 6.400000000000001e-06     evaluation
reward: 7.57
episode: 1985   score: 5.0   memory length: 493833   epsilon:
0.22020868000908428    steps: 327    lr: 6.400000000000001e-06     evaluation
reward: 7.57
episode: 1986   score: 8.0   memory length: 494252   epsilon:
0.21937906000907903    steps: 419    lr: 6.400000000000001e-06     evaluation
reward: 7.62
episode: 1987   score: 5.0   memory length: 494598   epsilon: 0.2186939800090747
steps: 346    lr: 6.400000000000001e-06     evaluation reward: 7.59
episode: 1988   score: 9.0   memory length: 495100   epsilon: 0.2177000200090684
steps: 502    lr: 6.400000000000001e-06     evaluation reward: 7.61
episode: 1989   score: 9.0   memory length: 495546   epsilon:
```

0.21681694000906282   steps: 446   lr: 6.400000000000001e-06   evaluation
reward: 7.62
episode: 1990   score: 16.0   memory length: 496182   epsilon:
0.21555766000905485   steps: 636   lr: 6.400000000000001e-06   evaluation
reward: 7.68
episode: 1991   score: 8.0   memory length: 496634   epsilon: 0.2146627000090492
steps: 452   lr: 6.400000000000001e-06   evaluation reward: 7.71
episode: 1992   score: 3.0   memory length: 496866   epsilon:
0.21420334000904628   steps: 232   lr: 6.400000000000001e-06   evaluation
reward: 7.7
episode: 1993   score: 5.0   memory length: 497156   epsilon:
0.21362914000904265   steps: 290   lr: 6.400000000000001e-06   evaluation
reward: 7.65
episode: 1994   score: 8.0   memory length: 497615   epsilon: 0.2127203200090369
steps: 459   lr: 6.400000000000001e-06   evaluation reward: 7.66
episode: 1995   score: 8.0   memory length: 497997   epsilon:
0.21196396000903212   steps: 382   lr: 6.400000000000001e-06   evaluation
reward: 7.67
episode: 1996   score: 12.0   memory length: 498505   epsilon:
0.21095812000902575   steps: 508   lr: 6.400000000000001e-06   evaluation
reward: 7.7
episode: 1997   score: 9.0   memory length: 499000   epsilon:
0.20997802000901955   steps: 495   lr: 6.400000000000001e-06   evaluation
reward: 7.73
episode: 1998   score: 8.0   memory length: 499404   epsilon: 0.2091781000090145
steps: 404   lr: 6.400000000000001e-06   evaluation reward: 7.74
episode: 1999   score: 7.0   memory length: 499811   epsilon: 0.2083722400090094
steps: 407   lr: 6.400000000000001e-06   evaluation reward: 7.7
episode: 2000   score: 3.0   memory length: 500042   epsilon: 0.2079148600090065
steps: 231   lr: 2.560000000000001e-06   evaluation reward: 7.69
episode: 2001   score: 5.0   memory length: 500367   epsilon:
0.20727136000900243   steps: 325   lr: 2.560000000000001e-06   evaluation
reward: 7.65
episode: 2002   score: 11.0   memory length: 500871   epsilon:
0.2062734400089961   steps: 504   lr: 2.560000000000001e-06   evaluation
reward: 7.65
episode: 2003   score: 11.0   memory length: 501369   epsilon:
0.20528740000898987   steps: 498   lr: 2.560000000000001e-06   evaluation
reward: 7.67
episode: 2004   score: 10.0   memory length: 501840   epsilon:
0.20435482000898397   steps: 471   lr: 2.560000000000001e-06   evaluation
reward: 7.72
episode: 2005   score: 6.0   memory length: 502183   epsilon:
0.20367568000897968   steps: 343   lr: 2.560000000000001e-06   evaluation
reward: 7.73
episode: 2006   score: 8.0   memory length: 502588   epsilon: 0.2028737800089746
steps: 405   lr: 2.560000000000001e-06   evaluation reward: 7.72
episode: 2007   score: 5.0   memory length: 502916   epsilon: 0.2022243400089705

```
steps: 328    lr: 2.560000000000001e-06     evaluation reward: 7.74
episode: 2008   score: 4.0   memory length: 503194   epsilon: 0.201673900008967
steps: 278    lr: 2.560000000000001e-06     evaluation reward: 7.72
episode: 2009   score: 7.0   memory length: 503547   epsilon: 0.2009749600089626
steps: 353    lr: 2.560000000000001e-06     evaluation reward: 7.72
episode: 2010   score: 11.0   memory length: 504122   epsilon:
0.19983646000895539    steps: 575    lr: 2.560000000000001e-06     evaluation
reward: 7.73
episode: 2011   score: 15.0   memory length: 504681   epsilon:
0.19872964000894838    steps: 559    lr: 2.560000000000001e-06     evaluation
reward: 7.82
episode: 2012   score: 11.0   memory length: 505100   epsilon:
0.19790002000894313    steps: 419    lr: 2.560000000000001e-06     evaluation
reward: 7.88
episode: 2013   score: 3.0   memory length: 505349   epsilon:
0.19740700000894001    steps: 249    lr: 2.560000000000001e-06     evaluation
reward: 7.87
episode: 2014   score: 7.0   memory length: 505700   epsilon:
0.19671202000893562    steps: 351    lr: 2.560000000000001e-06     evaluation
reward: 7.87
episode: 2015   score: 12.0   memory length: 506209   epsilon:
0.19570420000892924    steps: 509    lr: 2.560000000000001e-06     evaluation
reward: 7.9
episode: 2016   score: 10.0   memory length: 506727   epsilon:
0.19467856000892275    steps: 518    lr: 2.560000000000001e-06     evaluation
reward: 7.95
episode: 2017   score: 5.0   memory length: 507015   epsilon:
0.19410832000891914    steps: 288    lr: 2.560000000000001e-06     evaluation
reward: 7.97
episode: 2018   score: 5.0   memory length: 507298   epsilon: 0.1935479800089156
steps: 283    lr: 2.560000000000001e-06     evaluation reward: 7.99
episode: 2019   score: 8.0   memory length: 507696   epsilon: 0.1927599400089106
steps: 398    lr: 2.560000000000001e-06     evaluation reward: 7.98
episode: 2020   score: 12.0   memory length: 508158   epsilon:
0.19184518000890483    steps: 462    lr: 2.560000000000001e-06     evaluation
reward: 8.05
```

```
---------------------------------------------------------------------------
RuntimeError                              Traceback (most recent call last)
Input In [5], in <cell line: 5>()
     67 # if the mean of scores of last 100 episode is bigger than 5 save model
     68 ### Change this save condition to whatever you prefer ###
     69 if np.mean(evaluation_reward) > 8.0 and np.mean(evaluation_reward) >␣
 ↪best_eval_reward:
---> 70     torch.save(agent.policy_net, "./save_model/breakout_dqn.pth")
     71     best_eval_reward = np.mean(evaluation_reward)
     72     if np.mean(evaluation_reward) > 8.2:
```

```
File /sw/external/python/anaconda3_gpu/lib/python3.9/site-packages/torch/
 ↪serialization.py:421, in save(obj, f, pickle_module, pickle_protocol,␣
 ↪_use_new_zipfile_serialization)
    418 _check_dill_version(pickle_module)
    420 if _use_new_zipfile_serialization:
--> 421     with _open_zipfile_writer(f) as opened_zipfile:
    422         _save(obj, opened_zipfile, pickle_module, pickle_protocol)
    423         return

File /sw/external/python/anaconda3_gpu/lib/python3.9/site-packages/torch/
 ↪serialization.py:308, in _open_zipfile_writer(name_or_buffer)
    306 else:
    307     container = _open_zipfile_writer_buffer
--> 308 return container(name_or_buffer)

File /sw/external/python/anaconda3_gpu/lib/python3.9/site-packages/torch/
 ↪serialization.py:286, in _open_zipfile_writer_file.__init__(self, name)
    285 def __init__(self, name) -> None:
--> 286     super(_open_zipfile_writer_file, self).
 ↪__init__(torch._C.PyTorchFileWriter(str(name)))

RuntimeError: Parent directory ./save_model does not exist.
```

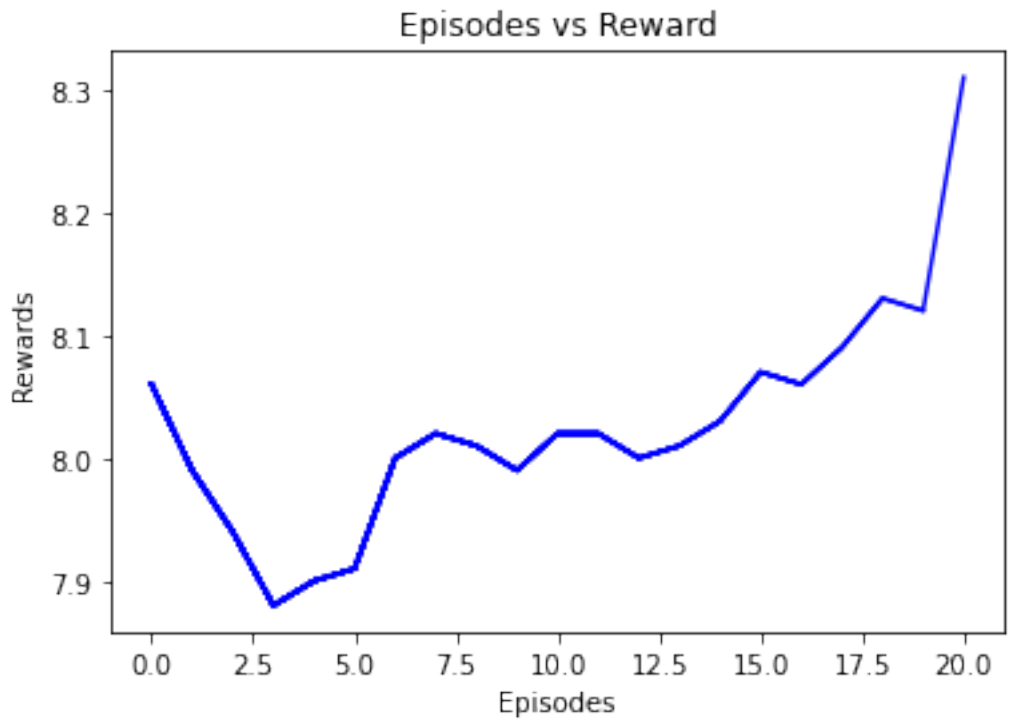


[ ]:
```
rewards, episodes = [], []
best_eval_reward = 0
halt_flag = 0

for e in range(EPISODES):
    done = False
    score = 0

    history = np.zeros([5, 84, 84], dtype=np.uint8)
    step = 0
    state, _ = env.reset()
    next_state = state
    life = number_lives

    get_init_state(history, state, HISTORY_SIZE)

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

        # Perform a fire action if ball is no longer on screen to continue onto␣
 ↪next life
        if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
            action = torch.tensor([[0]]).cuda()
        else:
            action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
#             print(action)
        state = next_state
#         next_state, reward, done, info = env.step(action + 1)
        next_state, reward, terminated, truncated, info = env.step(action + 1)
        done = terminated or truncated

        frame_next_state = get_frame(next_state)
        history[4, :, :] = frame_next_state
        terminal_state = check_live(life, info['lives'])

        life = info['lives']
        r = reward

        # Store the transition in memory
        agent.memory.push(deepcopy(frame_next_state), action.cpu(), r,␣
 ↪terminal_state)
#         agent.memory.push(deepcopy(frame_next_state), action, r,␣
 ↪terminal_state)
        # Start training after random sample generation
        if(frame >= train_frame): # You can set train_frame to a lower value␣
 ↪while testing your starts training earlier
```

```
#           if frame >= 200:
              agent.train_policy_net(frame)
              # Update the target network only for Double DQN only
              if double_dqn and (frame % update_target_network_frequency)== 0:
                  agent.update_target_net()
          score += reward
          history[:4, :, :] = history[1:, :, :]

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

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

              # if the mean of scores of last 100 episode is bigger than 5 save␣
 ↪model
              ### Change this save condition to whatever you prefer ###
              if np.mean(evaluation_reward) > 8.0 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)
                  if np.mean(evaluation_reward) > 8.2:
                      halt_flag = 1
                      break
      if halt_flag == 1:
          break
```

```
episode: 0   score: 10.0   memory length: 508684   epsilon: 0.19080370000889824
steps: 526    lr: 2.560000000000001e-06     evaluation reward: 8.06
episode: 1   score: 5.0   memory length: 509009   epsilon: 0.19016020000889416
steps: 325    lr: 2.560000000000001e-06     evaluation reward: 7.99
episode: 2   score: 6.0   memory length: 509346   epsilon: 0.18949294000888994
steps: 337    lr: 2.560000000000001e-06     evaluation reward: 7.94
episode: 3   score: 5.0   memory length: 509653   epsilon: 0.1888850800088861
steps: 307    lr: 2.560000000000001e-06     evaluation reward: 7.88
```

```
episode: 4   score: 8.0   memory length: 510129   epsilon: 0.18794260000888013
steps: 476    lr: 2.5600000000000001e-06     evaluation reward: 7.9
episode: 5   score: 7.0   memory length: 510504   epsilon: 0.18720010000887544
steps: 375    lr: 2.5600000000000001e-06     evaluation reward: 7.91
episode: 6   score: 12.0   memory length: 510960   epsilon: 0.18629722000886972
steps: 456    lr: 2.5600000000000001e-06     evaluation reward: 8.0
episode: 7   score: 9.0   memory length: 511398   epsilon: 0.18542998000886424
steps: 438    lr: 2.5600000000000001e-06     evaluation reward: 8.02
episode: 8   score: 5.0   memory length: 511707   epsilon: 0.18481816000886037
steps: 309    lr: 2.5600000000000001e-06     evaluation reward: 8.01
episode: 9   score: 6.0   memory length: 512082   epsilon: 0.18407566000885567
steps: 375    lr: 2.5600000000000001e-06     evaluation reward: 7.99
episode: 10   score: 10.0   memory length: 512640   epsilon: 0.18297082000884868
steps: 558    lr: 2.5600000000000001e-06     evaluation reward: 8.02
episode: 11   score: 11.0   memory length: 513209   epsilon: 0.18184420000884155
steps: 569    lr: 2.5600000000000001e-06     evaluation reward: 8.02
episode: 12   score: 9.0   memory length: 513628   epsilon: 0.1810145800088363
steps: 419    lr: 2.5600000000000001e-06     evaluation reward: 8.0
episode: 13   score: 5.0   memory length: 513952   epsilon: 0.18037306000883224
steps: 324    lr: 2.5600000000000001e-06     evaluation reward: 8.01
episode: 14   score: 4.0   memory length: 514229   epsilon: 0.17982460000882877
steps: 277    lr: 2.5600000000000001e-06     evaluation reward: 8.03
episode: 15   score: 10.0   memory length: 514763   epsilon: 0.17876728000882208
steps: 534    lr: 2.5600000000000001e-06     evaluation reward: 8.07
episode: 16   score: 9.0   memory length: 515183   epsilon: 0.17793568000881682
steps: 420    lr: 2.5600000000000001e-06     evaluation reward: 8.06
episode: 17   score: 7.0   memory length: 515583   epsilon: 0.1771436800088118
steps: 400    lr: 2.5600000000000001e-06     evaluation reward: 8.09
episode: 18   score: 13.0   memory length: 516062   epsilon: 0.1761952600088058
steps: 479    lr: 2.5600000000000001e-06     evaluation reward: 8.13
episode: 19   score: 7.0   memory length: 516430   epsilon: 0.1754666200088012
steps: 368    lr: 2.5600000000000001e-06     evaluation reward: 8.12
episode: 20   score: 22.0   memory length: 517214   epsilon: 0.17391430000879138
steps: 784    lr: 2.5600000000000001e-06     evaluation reward: 8.31
```



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

1.3 Double DQN Agent

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

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

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

[ ]:
```
rewards, episodes = [], []
best_eval_reward = 0
halt_flag = 0

for e in range(EPISODES):
    done = False
    score = 0
```

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

    get_init_state(history, state, HISTORY_SIZE)

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

        # Perform a fire action if ball is no longer on screen to continue onto␣
 ↪next life
        if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
            action = torch.tensor([[0]]).cuda()
        else:
            action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
#             print(action)
        state = next_state
#         next_state, reward, done, info = env.step(action + 1)
        next_state, reward, terminated, truncated, info = env.step(action + 1)
        done = terminated or truncated

        frame_next_state = get_frame(next_state)
        history[4, :, :] = frame_next_state
        terminal_state = check_live(life, info['lives'])

        life = info['lives']
        r = reward

        # Store the transition in memory
        agent.memory.push(deepcopy(frame_next_state), action.cpu(), r,␣
 ↪terminal_state)
#         agent.memory.push(deepcopy(frame_next_state), action, r,␣
 ↪terminal_state)
        # Start training after random sample generation
        if(frame >= train_frame): # You can set train_frame to a lower value␣
 ↪while testing your starts training earlier
#         if frame >= 200:
            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_double_dqn.png") # save graph␣
↪for training visualization

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

            # if the mean of scores of last 100 episode is bigger than 5 save␣
↪model
            ### Change this save condition to whatever you prefer ###
            if np.mean(evaluation_reward) > 10.0 and np.mean(evaluation_reward)␣
↪> best_eval_reward:
                torch.save(agent.policy_net, "./save_model/breakout_double_dqn.
↪pth")
                best_eval_reward = np.mean(evaluation_reward)
                if np.mean(evaluation_reward) > 10.1:
                    halt_flag = 1
                    break
    if halt_flag == 1:
        break
```

```
episode: 0   score: 0.0   memory length: 122   epsilon: 1.0    steps: 122    lr:
0.0001     evaluation reward: 0.0
episode: 1   score: 2.0   memory length: 340   epsilon: 1.0    steps: 218    lr:
0.0001     evaluation reward: 1.0
episode: 2   score: 0.0   memory length: 462   epsilon: 1.0    steps: 122    lr:
0.0001     evaluation reward: 0.6666666666666666
episode: 3   score: 1.0   memory length: 630   epsilon: 1.0    steps: 168    lr:
0.0001     evaluation reward: 0.75
episode: 4   score: 1.0   memory length: 802   epsilon: 1.0    steps: 172    lr:
0.0001     evaluation reward: 0.8
episode: 5   score: 1.0   memory length: 971   epsilon: 1.0    steps: 169    lr:
0.0001     evaluation reward: 0.8333333333333334
episode: 6   score: 1.0   memory length: 1140   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 0.8571428571428571
```

```
episode: 7   score: 0.0   memory length: 1263   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 0.75
episode: 8   score: 0.0   memory length: 1385   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 0.6666666666666666
episode: 9   score: 1.0   memory length: 1554   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 0.7
episode: 10   score: 0.0   memory length: 1676   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 0.6363636363636364
episode: 11   score: 1.0   memory length: 1827   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 0.6666666666666666
episode: 12   score: 0.0   memory length: 1950   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 0.6153846153846154
episode: 13   score: 0.0   memory length: 2072   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 0.5714285714285714
episode: 14   score: 3.0   memory length: 2336   epsilon: 1.0    steps: 264
lr: 0.0001     evaluation reward: 0.7333333333333333
episode: 15   score: 1.0   memory length: 2507   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 0.75
episode: 16   score: 3.0   memory length: 2772   epsilon: 1.0    steps: 265
lr: 0.0001     evaluation reward: 0.8823529411764706
episode: 17   score: 2.0   memory length: 2954   epsilon: 1.0    steps: 182
lr: 0.0001     evaluation reward: 0.9444444444444444
episode: 18   score: 1.0   memory length: 3126   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 0.9473684210526315
episode: 19   score: 0.0   memory length: 3249   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 0.9
episode: 20   score: 2.0   memory length: 3464   epsilon: 1.0    steps: 215
lr: 0.0001     evaluation reward: 0.9523809523809523
episode: 21   score: 1.0   memory length: 3633   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 0.9545454545454546
episode: 22   score: 0.0   memory length: 3755   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 0.9130434782608695
episode: 23   score: 4.0   memory length: 4031   epsilon: 1.0    steps: 276
lr: 0.0001     evaluation reward: 1.0416666666666667
episode: 24   score: 0.0   memory length: 4154   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.0
episode: 25   score: 2.0   memory length: 4372   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.0384615384615385
episode: 26   score: 3.0   memory length: 4640   epsilon: 1.0    steps: 268
lr: 0.0001     evaluation reward: 1.1111111111111112
episode: 27   score: 3.0   memory length: 4867   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.1785714285714286
episode: 28   score: 1.0   memory length: 5017   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.1724137931034482
episode: 29   score: 1.0   memory length: 5170   epsilon: 1.0    steps: 153
lr: 0.0001     evaluation reward: 1.1666666666666667
episode: 30   score: 2.0   memory length: 5367   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.1935483870967742
```

```
episode: 31   score: 4.0   memory length: 5644   epsilon: 1.0    steps: 277
lr: 0.0001     evaluation reward: 1.28125
episode: 32   score: 3.0   memory length: 5908   epsilon: 1.0    steps: 264
lr: 0.0001     evaluation reward: 1.3333333333333333
episode: 33   score: 0.0   memory length: 6031   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.2941176470588236
episode: 34   score: 2.0   memory length: 6229   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.3142857142857143
episode: 35   score: 6.0   memory length: 6543   epsilon: 1.0    steps: 314
lr: 0.0001     evaluation reward: 1.4444444444444444
episode: 36   score: 0.0   memory length: 6666   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4054054054054055
episode: 37   score: 2.0   memory length: 6881   epsilon: 1.0    steps: 215
lr: 0.0001     evaluation reward: 1.4210526315789473
episode: 38   score: 0.0   memory length: 7003   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.3846153846153846
episode: 39   score: 3.0   memory length: 7250   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.425
episode: 40   score: 0.0   memory length: 7373   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3902439024390243
episode: 41   score: 1.0   memory length: 7543   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.380952380952381
episode: 42   score: 0.0   memory length: 7665   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.3488372093023255
episode: 43   score: 2.0   memory length: 7863   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.3636363636363635
episode: 44   score: 5.0   memory length: 8200   epsilon: 1.0    steps: 337
lr: 0.0001     evaluation reward: 1.4444444444444444
episode: 45   score: 1.0   memory length: 8368   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.434782608695652
episode: 46   score: 5.0   memory length: 8675   epsilon: 1.0    steps: 307
lr: 0.0001     evaluation reward: 1.5106382978723405
episode: 47   score: 1.0   memory length: 8825   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.5
episode: 48   score: 0.0   memory length: 8948   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.469387755102041
episode: 49   score: 2.0   memory length: 9168   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.48
episode: 50   score: 0.0   memory length: 9291   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4509803921568627
episode: 51   score: 0.0   memory length: 9413   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.4230769230769231
episode: 52   score: 0.0   memory length: 9536   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3962264150943395
episode: 53   score: 0.0   memory length: 9658   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.3703703703703705
episode: 54   score: 4.0   memory length: 9955   epsilon: 1.0    steps: 297
lr: 0.0001     evaluation reward: 1.4181818181818182
```

```
episode: 55   score: 0.0   memory length: 10078   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3928571428571428
episode: 56   score: 1.0   memory length: 10229   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.3859649122807018
episode: 57   score: 7.0   memory length: 10662   epsilon: 1.0    steps: 433
lr: 0.0001     evaluation reward: 1.4827586206896552
episode: 58   score: 1.0   memory length: 10834   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.4745762711864407
episode: 59   score: 5.0   memory length: 11182   epsilon: 1.0    steps: 348
lr: 0.0001     evaluation reward: 1.5333333333333334
episode: 60   score: 0.0   memory length: 11305   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.5081967213114753
episode: 61   score: 1.0   memory length: 11474   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.5
episode: 62   score: 1.0   memory length: 11644   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.492063492063492
episode: 63   score: 1.0   memory length: 11813   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.484375
episode: 64   score: 2.0   memory length: 12011   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.4923076923076923
episode: 65   score: 2.0   memory length: 12227   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.5
episode: 66   score: 4.0   memory length: 12503   epsilon: 1.0    steps: 276
lr: 0.0001     evaluation reward: 1.537313432835821
episode: 67   score: 1.0   memory length: 12672   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.5294117647058822
episode: 68   score: 1.0   memory length: 12840   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.5217391304347827
episode: 69   score: 0.0   memory length: 12962   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.5
episode: 70   score: 1.0   memory length: 13133   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.4929577464788732
episode: 71   score: 2.0   memory length: 13330   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.5
episode: 72   score: 0.0   memory length: 13453   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4794520547945205
episode: 73   score: 2.0   memory length: 13656   epsilon: 1.0    steps: 203
lr: 0.0001     evaluation reward: 1.4864864864864864
episode: 74   score: 0.0   memory length: 13779   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4666666666666666
episode: 75   score: 0.0   memory length: 13901   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.4473684210526316
episode: 76   score: 0.0   memory length: 14024   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4285714285714286
episode: 77   score: 2.0   memory length: 14225   epsilon: 1.0    steps: 201
lr: 0.0001     evaluation reward: 1.435897435897436
episode: 78   score: 1.0   memory length: 14394   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.4303797468354431
```

```
episode: 79   score: 0.0   memory length: 14517   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4125
episode: 80   score: 2.0   memory length: 14735   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.4197530864197532
episode: 81   score: 2.0   memory length: 14935   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.4268292682926829
episode: 82   score: 1.0   memory length: 15105   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.4216867469879517
episode: 83   score: 1.0   memory length: 15277   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.4166666666666667
episode: 84   score: 1.0   memory length: 15447   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.411764705882353
episode: 85   score: 0.0   memory length: 15570   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3953488372093024
episode: 86   score: 2.0   memory length: 15767   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.4022988505747127
episode: 87   score: 4.0   memory length: 16055   epsilon: 1.0    steps: 288
lr: 0.0001     evaluation reward: 1.4318181818181819
episode: 88   score: 2.0   memory length: 16252   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.4382022471910112
episode: 89   score: 0.0   memory length: 16374   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.4222222222222223
episode: 90   score: 4.0   memory length: 16650   epsilon: 1.0    steps: 276
lr: 0.0001     evaluation reward: 1.4505494505494505
episode: 91   score: 1.0   memory length: 16818   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.4456521739130435
episode: 92   score: 3.0   memory length: 17063   epsilon: 1.0    steps: 245
lr: 0.0001     evaluation reward: 1.4623655913978495
episode: 93   score: 1.0   memory length: 17214   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.4574468085106382
episode: 94   score: 0.0   memory length: 17336   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.4421052631578948
episode: 95   score: 1.0   memory length: 17505   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.4375
episode: 96   score: 2.0   memory length: 17702   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.443298969072165
episode: 97   score: 1.0   memory length: 17873   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.4387755102040816
episode: 98   score: 0.0   memory length: 17996   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4242424242424243
episode: 99   score: 1.0   memory length: 18165   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.42
episode: 100   score: 1.0   memory length: 18333   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.43
episode: 101   score: 0.0   memory length: 18455   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.41
episode: 102   score: 0.0   memory length: 18578   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.41
```

```
episode: 103   score: 0.0   memory length: 18701   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4
episode: 104   score: 2.0   memory length: 18919   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.41
episode: 105   score: 1.0   memory length: 19070   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.41
episode: 106   score: 0.0   memory length: 19192   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.4
episode: 107   score: 1.0   memory length: 19343   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.41
episode: 108   score: 2.0   memory length: 19561   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.43
episode: 109   score: 2.0   memory length: 19743   epsilon: 1.0    steps: 182
lr: 0.0001     evaluation reward: 1.44
episode: 110   score: 1.0   memory length: 19911   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.45
episode: 111   score: 0.0   memory length: 20034   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.44
episode: 112   score: 1.0   memory length: 20185   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.45
episode: 113   score: 0.0   memory length: 20308   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.45
episode: 114   score: 2.0   memory length: 20508   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.44
episode: 115   score: 2.0   memory length: 20706   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.45
episode: 116   score: 2.0   memory length: 20926   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.44
episode: 117   score: 1.0   memory length: 21095   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.43
episode: 118   score: 0.0   memory length: 21218   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.42
episode: 119   score: 0.0   memory length: 21340   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.42
episode: 120   score: 2.0   memory length: 21538   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.42
episode: 121   score: 1.0   memory length: 21689   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.42
episode: 122   score: 3.0   memory length: 21918   epsilon: 1.0    steps: 229
lr: 0.0001     evaluation reward: 1.45
episode: 123   score: 0.0   memory length: 22040   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.41
episode: 124   score: 0.0   memory length: 22163   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.41
episode: 125   score: 2.0   memory length: 22385   epsilon: 1.0    steps: 222
lr: 0.0001     evaluation reward: 1.41
episode: 126   score: 0.0   memory length: 22507   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.38
```

```
episode: 127   score: 2.0   memory length: 22725   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.37
episode: 128   score: 0.0   memory length: 22848   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.36
episode: 129   score: 2.0   memory length: 23064   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.37
episode: 130   score: 0.0   memory length: 23186   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.35
episode: 131   score: 0.0   memory length: 23308   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.31
episode: 132   score: 1.0   memory length: 23459   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.29
episode: 133   score: 0.0   memory length: 23582   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.29
episode: 134   score: 1.0   memory length: 23751   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.28
episode: 135   score: 0.0   memory length: 23874   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.22
episode: 136   score: 0.0   memory length: 23997   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.22
episode: 137   score: 1.0   memory length: 24147   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.21
episode: 138   score: 1.0   memory length: 24316   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.22
episode: 139   score: 0.0   memory length: 24439   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.19
episode: 140   score: 0.0   memory length: 24562   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.19
episode: 141   score: 2.0   memory length: 24785   epsilon: 1.0    steps: 223
lr: 0.0001     evaluation reward: 1.2
episode: 142   score: 3.0   memory length: 25030   epsilon: 1.0    steps: 245
lr: 0.0001     evaluation reward: 1.23
episode: 143   score: 4.0   memory length: 25322   epsilon: 1.0    steps: 292
lr: 0.0001     evaluation reward: 1.25
episode: 144   score: 4.0   memory length: 25617   epsilon: 1.0    steps: 295
lr: 0.0001     evaluation reward: 1.24
episode: 145   score: 1.0   memory length: 25786   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.24
episode: 146   score: 2.0   memory length: 26004   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.21
episode: 147   score: 2.0   memory length: 26202   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.22
episode: 148   score: 0.0   memory length: 26325   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.22
episode: 149   score: 0.0   memory length: 26447   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.2
episode: 150   score: 0.0   memory length: 26570   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.2
```

```
episode: 151   score: 1.0   memory length: 26740   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.21
episode: 152   score: 2.0   memory length: 26959   epsilon: 1.0    steps: 219
lr: 0.0001     evaluation reward: 1.23
episode: 153   score: 2.0   memory length: 27157   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.25
episode: 154   score: 0.0   memory length: 27279   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.21
episode: 155   score: 1.0   memory length: 27448   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.22
episode: 156   score: 2.0   memory length: 27666   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.23
episode: 157   score: 3.0   memory length: 27910   epsilon: 1.0    steps: 244
lr: 0.0001     evaluation reward: 1.19
episode: 158   score: 0.0   memory length: 28033   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.18
episode: 159   score: 1.0   memory length: 28202   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.14
episode: 160   score: 0.0   memory length: 28325   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.14
episode: 161   score: 1.0   memory length: 28475   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.14
episode: 162   score: 1.0   memory length: 28644   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.14
episode: 163   score: 0.0   memory length: 28766   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.13
episode: 164   score: 0.0   memory length: 28889   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.11
episode: 165   score: 2.0   memory length: 29108   epsilon: 1.0    steps: 219
lr: 0.0001     evaluation reward: 1.11
episode: 166   score: 3.0   memory length: 29375   epsilon: 1.0    steps: 267
lr: 0.0001     evaluation reward: 1.1
episode: 167   score: 0.0   memory length: 29498   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.09
episode: 168   score: 2.0   memory length: 29696   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.1
episode: 169   score: 2.0   memory length: 29894   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.12
episode: 170   score: 1.0   memory length: 30063   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.12
episode: 171   score: 0.0   memory length: 30186   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.1
episode: 172   score: 1.0   memory length: 30337   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.11
episode: 173   score: 0.0   memory length: 30460   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.09
episode: 174   score: 1.0   memory length: 30612   epsilon: 1.0    steps: 152
lr: 0.0001     evaluation reward: 1.1
```

```
episode: 175   score: 3.0   memory length: 30857   epsilon: 1.0    steps: 245
lr: 0.0001     evaluation reward: 1.13
episode: 176   score: 0.0   memory length: 30980   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.13
episode: 177   score: 1.0   memory length: 31130   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.12
episode: 178   score: 1.0   memory length: 31281   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.12
episode: 179   score: 1.0   memory length: 31451   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.13
episode: 180   score: 0.0   memory length: 31573   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.11
episode: 181   score: 4.0   memory length: 31888   epsilon: 1.0    steps: 315
lr: 0.0001     evaluation reward: 1.13
episode: 182   score: 1.0   memory length: 32039   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.13
episode: 183   score: 0.0   memory length: 32162   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.12
episode: 184   score: 3.0   memory length: 32389   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.14
episode: 185   score: 2.0   memory length: 32569   epsilon: 1.0    steps: 180
lr: 0.0001     evaluation reward: 1.16
episode: 186   score: 2.0   memory length: 32789   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.16
episode: 187   score: 2.0   memory length: 33006   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.14
episode: 188   score: 3.0   memory length: 33232   epsilon: 1.0    steps: 226
lr: 0.0001     evaluation reward: 1.15
episode: 189   score: 1.0   memory length: 33382   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.16
episode: 190   score: 1.0   memory length: 33553   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.13
episode: 191   score: 0.0   memory length: 33675   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.12
episode: 192   score: 6.0   memory length: 34060   epsilon: 1.0    steps: 385
lr: 0.0001     evaluation reward: 1.15
episode: 193   score: 1.0   memory length: 34232   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.15
episode: 194   score: 0.0   memory length: 34355   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.15
episode: 195   score: 4.0   memory length: 34630   epsilon: 1.0    steps: 275
lr: 0.0001     evaluation reward: 1.18
episode: 196   score: 0.0   memory length: 34753   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.16
episode: 197   score: 1.0   memory length: 34922   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.16
episode: 198   score: 2.0   memory length: 35139   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.18
```

```
episode: 199   score: 1.0   memory length: 35308   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.18
episode: 200   score: 1.0   memory length: 35477   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.18
episode: 201   score: 6.0   memory length: 35807   epsilon: 1.0    steps: 330
lr: 0.0001     evaluation reward: 1.24
episode: 202   score: 0.0   memory length: 35930   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.24
episode: 203   score: 0.0   memory length: 36053   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.24
episode: 204   score: 0.0   memory length: 36176   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.22
episode: 205   score: 0.0   memory length: 36299   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.21
episode: 206   score: 3.0   memory length: 36563   epsilon: 1.0    steps: 264
lr: 0.0001     evaluation reward: 1.24
episode: 207   score: 4.0   memory length: 36856   epsilon: 1.0    steps: 293
lr: 0.0001     evaluation reward: 1.27
episode: 208   score: 1.0   memory length: 37025   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.26
episode: 209   score: 4.0   memory length: 37338   epsilon: 1.0    steps: 313
lr: 0.0001     evaluation reward: 1.28
episode: 210   score: 0.0   memory length: 37461   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.27
episode: 211   score: 2.0   memory length: 37682   epsilon: 1.0    steps: 221
lr: 0.0001     evaluation reward: 1.29
episode: 212   score: 2.0   memory length: 37900   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.3
episode: 213   score: 0.0   memory length: 38023   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.3
episode: 214   score: 0.0   memory length: 38146   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.28
episode: 215   score: 1.0   memory length: 38317   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.27
episode: 216   score: 1.0   memory length: 38486   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.26
episode: 217   score: 5.0   memory length: 38830   epsilon: 1.0    steps: 344
lr: 0.0001     evaluation reward: 1.3
episode: 218   score: 2.0   memory length: 39027   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.32
episode: 219   score: 3.0   memory length: 39291   epsilon: 1.0    steps: 264
lr: 0.0001     evaluation reward: 1.35
episode: 220   score: 0.0   memory length: 39414   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.33
episode: 221   score: 0.0   memory length: 39537   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.32
episode: 222   score: 0.0   memory length: 39660   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.29
```

```
episode: 223   score: 3.0   memory length: 39904   epsilon: 1.0    steps: 244
lr: 0.0001     evaluation reward: 1.32
episode: 224   score: 3.0   memory length: 40131   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.35
episode: 225   score: 3.0   memory length: 40376   epsilon: 1.0    steps: 245
lr: 0.0001     evaluation reward: 1.36
episode: 226   score: 0.0   memory length: 40499   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.36
episode: 227   score: 1.0   memory length: 40668   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.35
episode: 228   score: 1.0   memory length: 40818   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.36
episode: 229   score: 2.0   memory length: 41016   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.36
episode: 230   score: 4.0   memory length: 41291   epsilon: 1.0    steps: 275
lr: 0.0001     evaluation reward: 1.4
episode: 231   score: 0.0   memory length: 41414   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4
episode: 232   score: 0.0   memory length: 41537   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.39
episode: 233   score: 1.0   memory length: 41706   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.4
episode: 234   score: 0.0   memory length: 41828   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.39
episode: 235   score: 2.0   memory length: 42045   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.41
episode: 236   score: 2.0   memory length: 42243   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.43
episode: 237   score: 0.0   memory length: 42365   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.42
episode: 238   score: 1.0   memory length: 42533   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.42
episode: 239   score: 3.0   memory length: 42778   epsilon: 1.0    steps: 245
lr: 0.0001     evaluation reward: 1.45
episode: 240   score: 3.0   memory length: 43026   epsilon: 1.0    steps: 248
lr: 0.0001     evaluation reward: 1.48
episode: 241   score: 1.0   memory length: 43195   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.47
episode: 242   score: 2.0   memory length: 43396   epsilon: 1.0    steps: 201
lr: 0.0001     evaluation reward: 1.46
episode: 243   score: 1.0   memory length: 43566   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.43
episode: 244   score: 2.0   memory length: 43764   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.41
episode: 245   score: 3.0   memory length: 43991   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.43
episode: 246   score: 1.0   memory length: 44160   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.42
```

```
episode: 247   score: 1.0   memory length: 44311   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.41
episode: 248   score: 3.0   memory length: 44537   epsilon: 1.0    steps: 226
lr: 0.0001     evaluation reward: 1.44
episode: 249   score: 0.0   memory length: 44660   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.44
episode: 250   score: 2.0   memory length: 44860   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.46
episode: 251   score: 3.0   memory length: 45108   epsilon: 1.0    steps: 248
lr: 0.0001     evaluation reward: 1.48
episode: 252   score: 7.0   memory length: 45515   epsilon: 1.0    steps: 407
lr: 0.0001     evaluation reward: 1.53
episode: 253   score: 3.0   memory length: 45780   epsilon: 1.0    steps: 265
lr: 0.0001     evaluation reward: 1.54
episode: 254   score: 1.0   memory length: 45931   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.55
episode: 255   score: 3.0   memory length: 46179   epsilon: 1.0    steps: 248
lr: 0.0001     evaluation reward: 1.57
episode: 256   score: 2.0   memory length: 46377   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.57
episode: 257   score: 2.0   memory length: 46595   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.56
episode: 258   score: 0.0   memory length: 46718   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.56
episode: 259   score: 4.0   memory length: 46975   epsilon: 1.0    steps: 257
lr: 0.0001     evaluation reward: 1.59
episode: 260   score: 2.0   memory length: 47173   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.61
episode: 261   score: 5.0   memory length: 47495   epsilon: 1.0    steps: 322
lr: 0.0001     evaluation reward: 1.65
episode: 262   score: 0.0   memory length: 47618   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.64
episode: 263   score: 0.0   memory length: 47741   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.64
episode: 264   score: 3.0   memory length: 47969   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.67
episode: 265   score: 1.0   memory length: 48137   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.66
episode: 266   score: 1.0   memory length: 48308   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.64
episode: 267   score: 4.0   memory length: 48583   epsilon: 1.0    steps: 275
lr: 0.0001     evaluation reward: 1.68
episode: 268   score: 4.0   memory length: 48882   epsilon: 1.0    steps: 299
lr: 0.0001     evaluation reward: 1.7
episode: 269   score: 3.0   memory length: 49127   epsilon: 1.0    steps: 245
lr: 0.0001     evaluation reward: 1.71
episode: 270   score: 3.0   memory length: 49392   epsilon: 1.0    steps: 265
lr: 0.0001     evaluation reward: 1.73
```

```
episode: 271   score: 0.0   memory length: 49515   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.73
episode: 272   score: 1.0   memory length: 49665   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.73
episode: 273   score: 0.0   memory length: 49788   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.73
episode: 274   score: 5.0   memory length: 50113   epsilon: 1.0    steps: 325
lr: 0.0001     evaluation reward: 1.77
episode: 275   score: 0.0   memory length: 50235   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.74
episode: 276   score: 0.0   memory length: 50358   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.74
episode: 277   score: 2.0   memory length: 50558   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.75
episode: 278   score: 5.0   memory length: 50906   epsilon: 1.0    steps: 348
lr: 0.0001     evaluation reward: 1.79
episode: 279   score: 3.0   memory length: 51171   epsilon: 1.0    steps: 265
lr: 0.0001     evaluation reward: 1.81
episode: 280   score: 4.0   memory length: 51450   epsilon: 1.0    steps: 279
lr: 0.0001     evaluation reward: 1.85
episode: 281   score: 0.0   memory length: 51573   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.81
episode: 282   score: 1.0   memory length: 51724   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.81
episode: 283   score: 1.0   memory length: 51874   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.82
episode: 284   score: 0.0   memory length: 51997   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.79
episode: 285   score: 4.0   memory length: 52277   epsilon: 1.0    steps: 280
lr: 0.0001     evaluation reward: 1.81
episode: 286   score: 2.0   memory length: 52496   epsilon: 1.0    steps: 219
lr: 0.0001     evaluation reward: 1.81
episode: 287   score: 2.0   memory length: 52693   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.81
episode: 288   score: 3.0   memory length: 52956   epsilon: 1.0    steps: 263
lr: 0.0001     evaluation reward: 1.81
episode: 289   score: 2.0   memory length: 53154   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.82
episode: 290   score: 2.0   memory length: 53352   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.83
episode: 291   score: 0.0   memory length: 53475   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.83
episode: 292   score: 3.0   memory length: 53705   epsilon: 1.0    steps: 230
lr: 0.0001     evaluation reward: 1.8
episode: 293   score: 2.0   memory length: 53903   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.81
episode: 294   score: 3.0   memory length: 54152   epsilon: 1.0    steps: 249
lr: 0.0001     evaluation reward: 1.84
```

```
episode: 295   score: 0.0   memory length: 54274   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.8
episode: 296   score: 3.0   memory length: 54521   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.83
episode: 297   score: 2.0   memory length: 54739   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.84
episode: 298   score: 2.0   memory length: 54957   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.84
episode: 299   score: 1.0   memory length: 55108   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.84
episode: 300   score: 3.0   memory length: 55371   epsilon: 1.0    steps: 263
lr: 0.0001     evaluation reward: 1.86
episode: 301   score: 0.0   memory length: 55494   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.8
episode: 302   score: 2.0   memory length: 55692   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.82
episode: 303   score: 0.0   memory length: 55815   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.82
episode: 304   score: 0.0   memory length: 55938   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.82
episode: 305   score: 2.0   memory length: 56120   epsilon: 1.0    steps: 182
lr: 0.0001     evaluation reward: 1.84
episode: 306   score: 1.0   memory length: 56271   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.82
episode: 307   score: 1.0   memory length: 56422   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.79
episode: 308   score: 3.0   memory length: 56649   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.81
episode: 309   score: 0.0   memory length: 56771   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.77
episode: 310   score: 4.0   memory length: 57046   epsilon: 1.0    steps: 275
lr: 0.0001     evaluation reward: 1.81
episode: 311   score: 0.0   memory length: 57168   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.79
episode: 312   score: 3.0   memory length: 57413   epsilon: 1.0    steps: 245
lr: 0.0001     evaluation reward: 1.8
episode: 313   score: 4.0   memory length: 57713   epsilon: 1.0    steps: 300
lr: 0.0001     evaluation reward: 1.84
episode: 314   score: 0.0   memory length: 57836   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.84
episode: 315   score: 0.0   memory length: 57958   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.83
episode: 316   score: 2.0   memory length: 58156   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.84
episode: 317   score: 6.0   memory length: 58545   epsilon: 1.0    steps: 389
lr: 0.0001     evaluation reward: 1.85
episode: 318   score: 6.0   memory length: 58900   epsilon: 1.0    steps: 355
lr: 0.0001     evaluation reward: 1.89
```

```
episode: 319   score: 0.0   memory length: 59023   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.86
episode: 320   score: 1.0   memory length: 59191   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.87
episode: 321   score: 3.0   memory length: 59417   epsilon: 1.0    steps: 226
lr: 0.0001     evaluation reward: 1.9
episode: 322   score: 1.0   memory length: 59587   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.91
episode: 323   score: 2.0   memory length: 59769   epsilon: 1.0    steps: 182
lr: 0.0001     evaluation reward: 1.9
episode: 324   score: 0.0   memory length: 59892   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.87
episode: 325   score: 0.0   memory length: 60014   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.84
episode: 326   score: 1.0   memory length: 60165   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.85
episode: 327   score: 3.0   memory length: 60411   epsilon: 1.0    steps: 246
lr: 0.0001     evaluation reward: 1.87
episode: 328   score: 0.0   memory length: 60533   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.86
episode: 329   score: 0.0   memory length: 60656   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.84
episode: 330   score: 4.0   memory length: 60973   epsilon: 1.0    steps: 317
lr: 0.0001     evaluation reward: 1.84
episode: 331   score: 2.0   memory length: 61191   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.86
episode: 332   score: 1.0   memory length: 61363   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.87
episode: 333   score: 2.0   memory length: 61581   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.88
episode: 334   score: 2.0   memory length: 61779   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.9
episode: 335   score: 2.0   memory length: 61997   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.9
episode: 336   score: 0.0   memory length: 62120   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.88
episode: 337   score: 2.0   memory length: 62299   epsilon: 1.0    steps: 179
lr: 0.0001     evaluation reward: 1.9
episode: 338   score: 5.0   memory length: 62619   epsilon: 1.0    steps: 320
lr: 0.0001     evaluation reward: 1.94
episode: 339   score: 2.0   memory length: 62817   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.93
episode: 340   score: 3.0   memory length: 63062   epsilon: 1.0    steps: 245
lr: 0.0001     evaluation reward: 1.93
episode: 341   score: 0.0   memory length: 63185   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.92
episode: 342   score: 2.0   memory length: 63403   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.92
```

```
episode: 343   score: 4.0   memory length: 63675   epsilon: 1.0    steps: 272
lr: 0.0001     evaluation reward: 1.95
episode: 344   score: 0.0   memory length: 63798   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.93
episode: 345   score: 1.0   memory length: 63967   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.91
episode: 346   score: 0.0   memory length: 64090   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.9
episode: 347   score: 1.0   memory length: 64241   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.9
episode: 348   score: 0.0   memory length: 64364   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.87
episode: 349   score: 3.0   memory length: 64630   epsilon: 1.0    steps: 266
lr: 0.0001     evaluation reward: 1.9
episode: 350   score: 0.0   memory length: 64753   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.88
episode: 351   score: 2.0   memory length: 64972   epsilon: 1.0    steps: 219
lr: 0.0001     evaluation reward: 1.87
episode: 352   score: 2.0   memory length: 65190   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.82
episode: 353   score: 2.0   memory length: 65387   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.81
episode: 354   score: 3.0   memory length: 65635   epsilon: 1.0    steps: 248
lr: 0.0001     evaluation reward: 1.83
episode: 355   score: 0.0   memory length: 65758   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.8
episode: 356   score: 3.0   memory length: 65986   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.81
episode: 357   score: 1.0   memory length: 66156   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.8
episode: 358   score: 1.0   memory length: 66324   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.81
episode: 359   score: 4.0   memory length: 66600   epsilon: 1.0    steps: 276
lr: 0.0001     evaluation reward: 1.81
episode: 360   score: 0.0   memory length: 66722   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.79
episode: 361   score: 0.0   memory length: 66845   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.74
episode: 362   score: 0.0   memory length: 66968   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.74
episode: 363   score: 2.0   memory length: 67168   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.76
episode: 364   score: 1.0   memory length: 67318   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.74
episode: 365   score: 1.0   memory length: 67469   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.74
episode: 366   score: 2.0   memory length: 67667   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.75
```

```
episode: 367   score: 3.0   memory length: 67896   epsilon: 1.0    steps: 229
lr: 0.0001     evaluation reward: 1.74
episode: 368   score: 2.0   memory length: 68094   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.72
episode: 369   score: 2.0   memory length: 68292   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.71
episode: 370   score: 1.0   memory length: 68463   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.69
episode: 371   score: 3.0   memory length: 68691   epsilon: 1.0    steps: 228
lr: 0.0001     evaluation reward: 1.72
episode: 372   score: 5.0   memory length: 69032   epsilon: 1.0    steps: 341
lr: 0.0001     evaluation reward: 1.76
episode: 373   score: 4.0   memory length: 69347   epsilon: 1.0    steps: 315
lr: 0.0001     evaluation reward: 1.8
episode: 374   score: 1.0   memory length: 69519   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.76
episode: 375   score: 1.0   memory length: 69670   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.77
episode: 376   score: 2.0   memory length: 69867   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.79
episode: 377   score: 0.0   memory length: 69990   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.77
episode: 378   score: 0.0   memory length: 70113   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.72
episode: 379   score: 5.0   memory length: 70431   epsilon: 1.0    steps: 318
lr: 0.0001     evaluation reward: 1.74
episode: 380   score: 1.0   memory length: 70582   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.71
episode: 381   score: 3.0   memory length: 70825   epsilon: 1.0    steps: 243
lr: 0.0001     evaluation reward: 1.74
episode: 382   score: 2.0   memory length: 71022   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.75
episode: 383   score: 2.0   memory length: 71221   epsilon: 1.0    steps: 199
lr: 0.0001     evaluation reward: 1.76
episode: 384   score: 0.0   memory length: 71343   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.76
episode: 385   score: 2.0   memory length: 71541   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.74
episode: 386   score: 0.0   memory length: 71664   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.72
episode: 387   score: 2.0   memory length: 71882   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.72
episode: 388   score: 0.0   memory length: 72004   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.69
episode: 389   score: 0.0   memory length: 72127   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.67
episode: 390   score: 0.0   memory length: 72249   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.65
```

```
episode: 391   score: 0.0   memory length: 72372   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.65
episode: 392   score: 2.0   memory length: 72575   epsilon: 1.0    steps: 203
lr: 0.0001     evaluation reward: 1.64
episode: 393   score: 4.0   memory length: 72866   epsilon: 1.0    steps: 291
lr: 0.0001     evaluation reward: 1.66
episode: 394   score: 1.0   memory length: 73017   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.64
episode: 395   score: 0.0   memory length: 73139   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.64
episode: 396   score: 0.0   memory length: 73262   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.61
episode: 397   score: 0.0   memory length: 73384   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.59
episode: 398   score: 3.0   memory length: 73651   epsilon: 1.0    steps: 267
lr: 0.0001     evaluation reward: 1.6
episode: 399   score: 2.0   memory length: 73849   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.61
episode: 400   score: 0.0   memory length: 73972   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.58
episode: 401   score: 2.0   memory length: 74170   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.6
episode: 402   score: 0.0   memory length: 74293   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.58
episode: 403   score: 0.0   memory length: 74416   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.58
episode: 404   score: 1.0   memory length: 74567   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.59
episode: 405   score: 1.0   memory length: 74717   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.58
episode: 406   score: 1.0   memory length: 74886   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.58
episode: 407   score: 1.0   memory length: 75037   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.58
episode: 408   score: 12.0   memory length: 75374   epsilon: 1.0    steps: 337
lr: 0.0001     evaluation reward: 1.67
episode: 409   score: 1.0   memory length: 75525   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.68
episode: 410   score: 1.0   memory length: 75675   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.65
episode: 411   score: 3.0   memory length: 75923   epsilon: 1.0    steps: 248
lr: 0.0001     evaluation reward: 1.68
episode: 412   score: 1.0   memory length: 76074   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.66
episode: 413   score: 2.0   memory length: 76273   epsilon: 1.0    steps: 199
lr: 0.0001     evaluation reward: 1.64
episode: 414   score: 2.0   memory length: 76471   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.66
```

```
episode: 415   score: 1.0   memory length: 76640   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.67
episode: 416   score: 0.0   memory length: 76762   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.65
episode: 417   score: 3.0   memory length: 76975   epsilon: 1.0    steps: 213
lr: 0.0001     evaluation reward: 1.62
episode: 418   score: 2.0   memory length: 77176   epsilon: 1.0    steps: 201
lr: 0.0001     evaluation reward: 1.58
episode: 419   score: 3.0   memory length: 77423   epsilon: 1.0    steps: 247
lr: 0.0001     evaluation reward: 1.61
episode: 420   score: 0.0   memory length: 77546   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.6
episode: 421   score: 1.0   memory length: 77715   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.58
episode: 422   score: 1.0   memory length: 77866   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.58
episode: 423   score: 1.0   memory length: 78016   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.57
episode: 424   score: 2.0   memory length: 78239   epsilon: 1.0    steps: 223
lr: 0.0001     evaluation reward: 1.59
episode: 425   score: 1.0   memory length: 78389   epsilon: 1.0    steps: 150
lr: 0.0001     evaluation reward: 1.6
episode: 426   score: 1.0   memory length: 78557   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.6
episode: 427   score: 5.0   memory length: 78881   epsilon: 1.0    steps: 324
lr: 0.0001     evaluation reward: 1.62
episode: 428   score: 2.0   memory length: 79097   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.64
episode: 429   score: 2.0   memory length: 79294   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.66
episode: 430   score: 1.0   memory length: 79446   epsilon: 1.0    steps: 152
lr: 0.0001     evaluation reward: 1.63
episode: 431   score: 1.0   memory length: 79618   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.62
episode: 432   score: 0.0   memory length: 79741   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.61
episode: 433   score: 1.0   memory length: 79912   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.6
episode: 434   score: 4.0   memory length: 80190   epsilon: 1.0    steps: 278
lr: 0.0001     evaluation reward: 1.62
episode: 435   score: 2.0   memory length: 80409   epsilon: 1.0    steps: 219
lr: 0.0001     evaluation reward: 1.62
episode: 436   score: 0.0   memory length: 80531   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.62
episode: 437   score: 2.0   memory length: 80728   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.62
episode: 438   score: 0.0   memory length: 80850   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.57
```

```
episode: 439   score: 0.0   memory length: 80972   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.55
episode: 440   score: 1.0   memory length: 81143   epsilon: 1.0    steps: 171
lr: 0.0001     evaluation reward: 1.53
episode: 441   score: 1.0   memory length: 81312   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.54
episode: 442   score: 0.0   memory length: 81434   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.52
episode: 443   score: 0.0   memory length: 81557   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.48
episode: 444   score: 2.0   memory length: 81774   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.5
episode: 445   score: 1.0   memory length: 81925   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.5
episode: 446   score: 1.0   memory length: 82076   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.51
episode: 447   score: 0.0   memory length: 82199   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.5
episode: 448   score: 2.0   memory length: 82399   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.52
episode: 449   score: 1.0   memory length: 82550   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.5
episode: 450   score: 3.0   memory length: 82817   epsilon: 1.0    steps: 267
lr: 0.0001     evaluation reward: 1.53
episode: 451   score: 0.0   memory length: 82940   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.51
episode: 452   score: 3.0   memory length: 83169   epsilon: 1.0    steps: 229
lr: 0.0001     evaluation reward: 1.52
episode: 453   score: 2.0   memory length: 83368   epsilon: 1.0    steps: 199
lr: 0.0001     evaluation reward: 1.52
episode: 454   score: 3.0   memory length: 83594   epsilon: 1.0    steps: 226
lr: 0.0001     evaluation reward: 1.52
episode: 455   score: 3.0   memory length: 83861   epsilon: 1.0    steps: 267
lr: 0.0001     evaluation reward: 1.55
episode: 456   score: 1.0   memory length: 84033   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.53
episode: 457   score: 0.0   memory length: 84156   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.52
episode: 458   score: 1.0   memory length: 84324   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.52
episode: 459   score: 1.0   memory length: 84494   epsilon: 1.0    steps: 170
lr: 0.0001     evaluation reward: 1.49
episode: 460   score: 0.0   memory length: 84617   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.49
episode: 461   score: 2.0   memory length: 84799   epsilon: 1.0    steps: 182
lr: 0.0001     evaluation reward: 1.51
episode: 462   score: 1.0   memory length: 84968   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.52
```

```
episode: 463   score: 0.0   memory length: 85091   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.5
episode: 464   score: 0.0   memory length: 85214   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.49
episode: 465   score: 5.0   memory length: 85539   epsilon: 1.0    steps: 325
lr: 0.0001     evaluation reward: 1.53
episode: 466   score: 0.0   memory length: 85662   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.51
episode: 467   score: 2.0   memory length: 85879   epsilon: 1.0    steps: 217
lr: 0.0001     evaluation reward: 1.5
episode: 468   score: 2.0   memory length: 86097   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.5
episode: 469   score: 1.0   memory length: 86266   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.49
episode: 470   score: 2.0   memory length: 86482   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.5
episode: 471   score: 2.0   memory length: 86680   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.49
episode: 472   score: 2.0   memory length: 86878   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.46
episode: 473   score: 4.0   memory length: 87176   epsilon: 1.0    steps: 298
lr: 0.0001     evaluation reward: 1.46
episode: 474   score: 0.0   memory length: 87299   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.45
episode: 475   score: 2.0   memory length: 87498   epsilon: 1.0    steps: 199
lr: 0.0001     evaluation reward: 1.46
episode: 476   score: 2.0   memory length: 87718   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 1.46
episode: 477   score: 2.0   memory length: 87919   epsilon: 1.0    steps: 201
lr: 0.0001     evaluation reward: 1.48
episode: 478   score: 1.0   memory length: 88088   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.49
episode: 479   score: 2.0   memory length: 88286   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.46
episode: 480   score: 0.0   memory length: 88409   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.45
episode: 481   score: 0.0   memory length: 88531   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.42
episode: 482   score: 4.0   memory length: 88828   epsilon: 1.0    steps: 297
lr: 0.0001     evaluation reward: 1.44
episode: 483   score: 2.0   memory length: 89026   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.44
episode: 484   score: 6.0   memory length: 89419   epsilon: 1.0    steps: 393
lr: 0.0001     evaluation reward: 1.5
episode: 485   score: 4.0   memory length: 89715   epsilon: 1.0    steps: 296
lr: 0.0001     evaluation reward: 1.52
episode: 486   score: 0.0   memory length: 89838   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.52
```

```
episode: 487   score: 2.0   memory length: 90038   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.52
episode: 488   score: 0.0   memory length: 90160   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.52
episode: 489   score: 0.0   memory length: 90283   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.52
episode: 490   score: 1.0   memory length: 90451   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.53
episode: 491   score: 0.0   memory length: 90574   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.53
episode: 492   score: 2.0   memory length: 90772   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.53
episode: 493   score: 0.0   memory length: 90895   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.49
episode: 494   score: 1.0   memory length: 91067   epsilon: 1.0    steps: 172
lr: 0.0001     evaluation reward: 1.49
episode: 495   score: 2.0   memory length: 91264   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.51
episode: 496   score: 0.0   memory length: 91386   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.51
episode: 497   score: 6.0   memory length: 91744   epsilon: 1.0    steps: 358
lr: 0.0001     evaluation reward: 1.57
episode: 498   score: 0.0   memory length: 91867   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.54
episode: 499   score: 2.0   memory length: 92083   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.54
episode: 500   score: 0.0   memory length: 92206   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.54
episode: 501   score: 2.0   memory length: 92407   epsilon: 1.0    steps: 201
lr: 0.0001     evaluation reward: 1.54
episode: 502   score: 3.0   memory length: 92674   epsilon: 1.0    steps: 267
lr: 0.0001     evaluation reward: 1.57
episode: 503   score: 0.0   memory length: 92797   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.57
episode: 504   score: 1.0   memory length: 92948   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.57
episode: 505   score: 0.0   memory length: 93071   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.56
episode: 506   score: 0.0   memory length: 93193   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.55
episode: 507   score: 0.0   memory length: 93316   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.54
episode: 508   score: 0.0   memory length: 93439   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.42
episode: 509   score: 0.0   memory length: 93562   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.41
episode: 510   score: 0.0   memory length: 93685   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4
```

```
episode: 511   score: 1.0   memory length: 93837   epsilon: 1.0    steps: 152
lr: 0.0001     evaluation reward: 1.38
episode: 512   score: 3.0   memory length: 94086   epsilon: 1.0    steps: 249
lr: 0.0001     evaluation reward: 1.4
episode: 513   score: 1.0   memory length: 94237   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.39
episode: 514   score: 0.0   memory length: 94360   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.37
episode: 515   score: 2.0   memory length: 94539   epsilon: 1.0    steps: 179
lr: 0.0001     evaluation reward: 1.38
episode: 516   score: 2.0   memory length: 94737   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.4
episode: 517   score: 3.0   memory length: 94980   epsilon: 1.0    steps: 243
lr: 0.0001     evaluation reward: 1.4
episode: 518   score: 3.0   memory length: 95206   epsilon: 1.0    steps: 226
lr: 0.0001     evaluation reward: 1.41
episode: 519   score: 2.0   memory length: 95408   epsilon: 1.0    steps: 202
lr: 0.0001     evaluation reward: 1.4
episode: 520   score: 0.0   memory length: 95531   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.4
episode: 521   score: 1.0   memory length: 95682   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.4
episode: 522   score: 2.0   memory length: 95880   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.41
episode: 523   score: 1.0   memory length: 96031   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.41
episode: 524   score: 0.0   memory length: 96153   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.39
episode: 525   score: 2.0   memory length: 96350   epsilon: 1.0    steps: 197
lr: 0.0001     evaluation reward: 1.4
episode: 526   score: 0.0   memory length: 96472   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.39
episode: 527   score: 0.0   memory length: 96595   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.34
episode: 528   score: 2.0   memory length: 96813   epsilon: 1.0    steps: 218
lr: 0.0001     evaluation reward: 1.34
episode: 529   score: 2.0   memory length: 97011   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.34
episode: 530   score: 1.0   memory length: 97162   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.34
episode: 531   score: 1.0   memory length: 97331   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.34
episode: 532   score: 2.0   memory length: 97547   epsilon: 1.0    steps: 216
lr: 0.0001     evaluation reward: 1.36
episode: 533   score: 0.0   memory length: 97670   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.35
episode: 534   score: 0.0   memory length: 97793   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.31
```

```
episode: 535   score: 2.0   memory length: 97991   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.31
episode: 536   score: 2.0   memory length: 98189   epsilon: 1.0    steps: 198
lr: 0.0001     evaluation reward: 1.33
episode: 537   score: 0.0   memory length: 98311   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.31
episode: 538   score: 2.0   memory length: 98511   epsilon: 1.0    steps: 200
lr: 0.0001     evaluation reward: 1.33
episode: 539   score: 0.0   memory length: 98634   epsilon: 1.0    steps: 123
lr: 0.0001     evaluation reward: 1.33
episode: 540   score: 1.0   memory length: 98785   epsilon: 1.0    steps: 151
lr: 0.0001     evaluation reward: 1.33
episode: 541   score: 3.0   memory length: 99012   epsilon: 1.0    steps: 227
lr: 0.0001     evaluation reward: 1.35
episode: 542   score: 1.0   memory length: 99181   epsilon: 1.0    steps: 169
lr: 0.0001     evaluation reward: 1.36
episode: 543   score: 1.0   memory length: 99349   epsilon: 1.0    steps: 168
lr: 0.0001     evaluation reward: 1.37
episode: 544   score: 3.0   memory length: 99597   epsilon: 1.0    steps: 248
lr: 0.0001     evaluation reward: 1.38
episode: 545   score: 0.0   memory length: 99719   epsilon: 1.0    steps: 122
lr: 0.0001     evaluation reward: 1.37
episode: 546   score: 4.0   memory length: 99995   epsilon: 1.0    steps: 276
lr: 0.0001     evaluation reward: 1.4
```

```
/u/qilong/QA_generator/CS444/mp5/memory.py:30: FutureWarning: The input object
of type 'Tensor' is an array-like implementing one of the corresponding
protocols (`__array__`, `__array_interface__` or `__array_struct__`); but not a
sequence (or 0-D). In the future, this object will be coerced as if it was first
converted using `np.array(obj)`. To retain the old behaviour, you have to either
modify the type 'Tensor', or assign to an empty array created with
`np.empty(correct_shape, dtype=object)`.
  sample = np.array(sample, dtype=object)
/u/qilong/QA_generator/CS444/mp5/agent_double.py:74: FutureWarning: The input
object of type 'Tensor' is an array-like implementing one of the corresponding
protocols (`__array__`, `__array_interface__` or `__array_struct__`); but not a
sequence (or 0-D). In the future, this object will be coerced as if it was first
converted using `np.array(obj)`. To retain the old behaviour, you have to either
modify the type 'Tensor', or assign to an empty array created with
`np.empty(correct_shape, dtype=object)`.
  mini_batch = np.array(mini_batch, dtype=object).transpose()
```

```
episode: 547   score: 1.0   memory length: 100145   epsilon: 0.9997109200000063
steps: 150    lr: 0.0001     evaluation reward: 1.41
episode: 548   score: 1.0   memory length: 100313   epsilon: 0.9993782800000135
steps: 168    lr: 0.0001     evaluation reward: 1.4
episode: 549   score: 1.0   memory length: 100481   epsilon: 0.9990456400000207
steps: 168    lr: 0.0001     evaluation reward: 1.4
episode: 550   score: 0.0   memory length: 100604   epsilon: 0.998802100000026
```

```
steps: 123    lr: 0.0001     evaluation reward: 1.37
episode: 551   score: 1.0   memory length: 100774   epsilon: 0.9984655000000333
steps: 170    lr: 0.0001     evaluation reward: 1.38
episode: 552   score: 2.0   memory length: 100976   epsilon: 0.998065540000042
steps: 202    lr: 0.0001     evaluation reward: 1.37
episode: 553   score: 0.0   memory length: 101099   epsilon: 0.9978220000000473
steps: 123    lr: 0.0001     evaluation reward: 1.35
episode: 554   score: 0.0   memory length: 101222   epsilon: 0.9975784600000526
steps: 123    lr: 0.0001     evaluation reward: 1.32
episode: 555   score: 2.0   memory length: 101419   epsilon: 0.997188400000061
steps: 197    lr: 0.0001     evaluation reward: 1.31
episode: 556   score: 2.0   memory length: 101616   epsilon: 0.9967983400000695
steps: 197    lr: 0.0001     evaluation reward: 1.32
episode: 557   score: 2.0   memory length: 101834   epsilon: 0.9963667000000789
steps: 218    lr: 0.0001     evaluation reward: 1.34
episode: 558   score: 4.0   memory length: 102121   epsilon: 0.9957984400000912
steps: 287    lr: 0.0001     evaluation reward: 1.37
episode: 559   score: 4.0   memory length: 102396   epsilon: 0.995253940000103
steps: 275    lr: 0.0001     evaluation reward: 1.4
episode: 560   score: 0.0   memory length: 102519   epsilon: 0.9950104000001083
steps: 123    lr: 0.0001     evaluation reward: 1.4
episode: 561   score: 2.0   memory length: 102718   epsilon: 0.9946163800001169
steps: 199    lr: 0.0001     evaluation reward: 1.4
episode: 562   score: 2.0   memory length: 102916   epsilon: 0.9942243400001254
steps: 198    lr: 0.0001     evaluation reward: 1.41
episode: 563   score: 2.0   memory length: 103114   epsilon: 0.9938323000001339
steps: 198    lr: 0.0001     evaluation reward: 1.43
episode: 564   score: 0.0   memory length: 103236   epsilon: 0.9935907400001391
steps: 122    lr: 0.0001     evaluation reward: 1.43
episode: 565   score: 1.0   memory length: 103407   epsilon: 0.9932521600001465
steps: 171    lr: 0.0001     evaluation reward: 1.39
episode: 566   score: 4.0   memory length: 103682   epsilon: 0.9927076600001583
steps: 275    lr: 0.0001     evaluation reward: 1.43
episode: 567   score: 2.0   memory length: 103898   epsilon: 0.9922799800001676
steps: 216    lr: 0.0001     evaluation reward: 1.43
episode: 568   score: 1.0   memory length: 104067   epsilon: 0.9919453600001749
steps: 169    lr: 0.0001     evaluation reward: 1.42
episode: 569   score: 0.0   memory length: 104190   epsilon: 0.9917018200001801
steps: 123    lr: 0.0001     evaluation reward: 1.41
episode: 570   score: 0.0   memory length: 104313   epsilon: 0.9914582800001854
steps: 123    lr: 0.0001     evaluation reward: 1.39
episode: 571   score: 1.0   memory length: 104464   epsilon: 0.9911593000001919
steps: 151    lr: 0.0001     evaluation reward: 1.38
episode: 572   score: 2.0   memory length: 104680   epsilon: 0.9907316200002012
steps: 216    lr: 0.0001     evaluation reward: 1.38
episode: 573   score: 0.0   memory length: 104802   epsilon: 0.9904900600002065
steps: 122    lr: 0.0001     evaluation reward: 1.34
episode: 574   score: 0.0   memory length: 104925   epsilon: 0.9902465200002117
```

```
steps: 123    lr: 0.0001     evaluation reward: 1.34
episode: 575   score: 0.0   memory length: 105047   epsilon: 0.990004960000217
steps: 122    lr: 0.0001     evaluation reward: 1.32
episode: 576   score: 1.0   memory length: 105218   epsilon: 0.9896663800002243
steps: 171    lr: 0.0001     evaluation reward: 1.31
episode: 577   score: 1.0   memory length: 105369   epsilon: 0.9893674000002308
steps: 151    lr: 0.0001     evaluation reward: 1.3
episode: 578   score: 2.0   memory length: 105589   epsilon: 0.9889318000002403
steps: 220    lr: 0.0001     evaluation reward: 1.31
episode: 579   score: 0.0   memory length: 105712   epsilon: 0.9886882600002456
steps: 123    lr: 0.0001     evaluation reward: 1.29
episode: 580   score: 1.0   memory length: 105881   epsilon: 0.9883536400002528
steps: 169    lr: 0.0001     evaluation reward: 1.3
episode: 581   score: 1.0   memory length: 106049   epsilon: 0.98802100000026
steps: 168    lr: 0.0001     evaluation reward: 1.31
episode: 582   score: 2.0   memory length: 106231   epsilon: 0.9876606400002679
steps: 182    lr: 0.0001     evaluation reward: 1.29
episode: 583   score: 6.0   memory length: 106597   epsilon: 0.9869359600002836
steps: 366    lr: 0.0001     evaluation reward: 1.33
episode: 584   score: 2.0   memory length: 106815   epsilon: 0.986504320000293
steps: 218    lr: 0.0001     evaluation reward: 1.29
episode: 585   score: 0.0   memory length: 106938   epsilon: 0.9862607800002983
steps: 123    lr: 0.0001     evaluation reward: 1.25
episode: 586   score: 2.0   memory length: 107138   epsilon: 0.9858647800003069
steps: 200    lr: 0.0001     evaluation reward: 1.27
episode: 587   score: 3.0   memory length: 107402   epsilon: 0.9853420600003182
steps: 264    lr: 0.0001     evaluation reward: 1.28
episode: 588   score: 0.0   memory length: 107524   epsilon: 0.9851005000003235
steps: 122    lr: 0.0001     evaluation reward: 1.28
episode: 589   score: 3.0   memory length: 107750   epsilon: 0.9846530200003332
steps: 226    lr: 0.0001     evaluation reward: 1.31
episode: 590   score: 5.0   memory length: 108039   epsilon: 0.9840808000003456
steps: 289    lr: 0.0001     evaluation reward: 1.35
episode: 591   score: 1.0   memory length: 108191   epsilon: 0.9837798400003521
steps: 152    lr: 0.0001     evaluation reward: 1.36
episode: 592   score: 3.0   memory length: 108417   epsilon: 0.9833323600003618
steps: 226    lr: 0.0001     evaluation reward: 1.37
episode: 593   score: 3.0   memory length: 108667   epsilon: 0.9828373600003726
steps: 250    lr: 0.0001     evaluation reward: 1.4
episode: 594   score: 0.0   memory length: 108790   epsilon: 0.9825938200003779
steps: 123    lr: 0.0001     evaluation reward: 1.39
episode: 595   score: 2.0   memory length: 109008   epsilon: 0.9821621800003872
steps: 218    lr: 0.0001     evaluation reward: 1.39
episode: 596   score: 1.0   memory length: 109177   epsilon: 0.9818275600003945
steps: 169    lr: 0.0001     evaluation reward: 1.4
episode: 597   score: 2.0   memory length: 109375   epsilon: 0.981435520000403
steps: 198    lr: 0.0001     evaluation reward: 1.36
episode: 598   score: 2.0   memory length: 109573   epsilon: 0.9810434800004115
```

```
steps: 198    lr: 0.0001     evaluation reward: 1.38
episode: 599   score: 1.0   memory length: 109743   epsilon: 0.9807068800004188
steps: 170    lr: 0.0001     evaluation reward: 1.37
episode: 600   score: 1.0   memory length: 109912   epsilon: 0.9803722600004261
steps: 169    lr: 0.0001     evaluation reward: 1.38
episode: 601   score: 3.0   memory length: 110163   epsilon: 0.9798752800004369
steps: 251    lr: 0.0001     evaluation reward: 1.39
episode: 602   score: 1.0   memory length: 110333   epsilon: 0.9795386800004442
steps: 170    lr: 0.0001     evaluation reward: 1.37
episode: 603   score: 0.0   memory length: 110456   epsilon: 0.9792951400004495
steps: 123    lr: 0.0001     evaluation reward: 1.37
episode: 604   score: 2.0   memory length: 110635   epsilon: 0.9789407200004572
steps: 179    lr: 0.0001     evaluation reward: 1.38
episode: 605   score: 2.0   memory length: 110834   epsilon: 0.9785467000004657
steps: 199    lr: 0.0001     evaluation reward: 1.4
episode: 606   score: 0.0   memory length: 110957   epsilon: 0.978303160000471
steps: 123    lr: 0.0001     evaluation reward: 1.4
episode: 607   score: 0.0   memory length: 111079   epsilon: 0.9780616000004763
steps: 122    lr: 0.0001     evaluation reward: 1.4
episode: 608   score: 0.0   memory length: 111202   epsilon: 0.9778180600004815
steps: 123    lr: 0.0001     evaluation reward: 1.4
episode: 609   score: 2.0   memory length: 111400   epsilon: 0.9774260200004901
steps: 198    lr: 0.0001     evaluation reward: 1.42
episode: 610   score: 2.0   memory length: 111619   epsilon: 0.9769924000004995
steps: 219    lr: 0.0001     evaluation reward: 1.44
episode: 611   score: 0.0   memory length: 111742   epsilon: 0.9767488600005048
steps: 123    lr: 0.0001     evaluation reward: 1.43
episode: 612   score: 2.0   memory length: 111958   epsilon: 0.976321180000514
steps: 216    lr: 0.0001     evaluation reward: 1.42
episode: 613   score: 3.0   memory length: 112224   epsilon: 0.9757945000005255
steps: 266    lr: 0.0001     evaluation reward: 1.44
episode: 614   score: 0.0   memory length: 112347   epsilon: 0.9755509600005308
steps: 123    lr: 0.0001     evaluation reward: 1.44
episode: 615   score: 0.0   memory length: 112470   epsilon: 0.975307420000536
steps: 123    lr: 0.0001     evaluation reward: 1.42
episode: 616   score: 0.0   memory length: 112592   epsilon: 0.9750658600005413
steps: 122    lr: 0.0001     evaluation reward: 1.4
episode: 617   score: 1.0   memory length: 112760   epsilon: 0.9747332200005485
steps: 168    lr: 0.0001     evaluation reward: 1.38
episode: 618   score: 0.0   memory length: 112883   epsilon: 0.9744896800005538
steps: 123    lr: 0.0001     evaluation reward: 1.35
episode: 619   score: 3.0   memory length: 113127   epsilon: 0.9740065600005643
steps: 244    lr: 0.0001     evaluation reward: 1.36
episode: 620   score: 0.0   memory length: 113249   epsilon: 0.9737650000005695
steps: 122    lr: 0.0001     evaluation reward: 1.36
episode: 621   score: 1.0   memory length: 113418   epsilon: 0.9734303800005768
steps: 169    lr: 0.0001     evaluation reward: 1.36
episode: 622   score: 1.0   memory length: 113568   epsilon: 0.9731333800005832
```

```
steps: 150    lr: 0.0001     evaluation reward: 1.35
episode: 623   score: 4.0   memory length: 113826   epsilon: 0.9726225400005943
steps: 258    lr: 0.0001     evaluation reward: 1.38
episode: 624   score: 3.0   memory length: 114036   epsilon: 0.9722067400006034
steps: 210    lr: 0.0001     evaluation reward: 1.41
episode: 625   score: 2.0   memory length: 114255   epsilon: 0.9717731200006128
steps: 219    lr: 0.0001     evaluation reward: 1.41
episode: 626   score: 0.0   memory length: 114378   epsilon: 0.9715295800006181
steps: 123    lr: 0.0001     evaluation reward: 1.41
episode: 627   score: 2.0   memory length: 114578   epsilon: 0.9711335800006267
steps: 200    lr: 0.0001     evaluation reward: 1.43
episode: 628   score: 2.0   memory length: 114758   epsilon: 0.9707771800006344
steps: 180    lr: 0.0001     evaluation reward: 1.43
episode: 629   score: 0.0   memory length: 114881   epsilon: 0.9705336400006397
steps: 123    lr: 0.0001     evaluation reward: 1.41
episode: 630   score: 3.0   memory length: 115148   epsilon: 0.9700049800006512
steps: 267    lr: 0.0001     evaluation reward: 1.43
episode: 631   score: 1.0   memory length: 115317   epsilon: 0.9696703600006584
steps: 169    lr: 0.0001     evaluation reward: 1.43
episode: 632   score: 0.0   memory length: 115439   epsilon: 0.9694288000006637
steps: 122    lr: 0.0001     evaluation reward: 1.41
episode: 633   score: 4.0   memory length: 115735   epsilon: 0.9688427200006764
steps: 296    lr: 0.0001     evaluation reward: 1.45
episode: 634   score: 0.0   memory length: 115858   epsilon: 0.9685991800006817
steps: 123    lr: 0.0001     evaluation reward: 1.45
episode: 635   score: 0.0   memory length: 115981   epsilon: 0.968355640000687
steps: 123    lr: 0.0001     evaluation reward: 1.43
episode: 636   score: 3.0   memory length: 116245   epsilon: 0.9678329200006983
steps: 264    lr: 0.0001     evaluation reward: 1.44
episode: 637   score: 5.0   memory length: 116572   epsilon: 0.9671854600007124
steps: 327    lr: 0.0001     evaluation reward: 1.49
episode: 638   score: 1.0   memory length: 116741   epsilon: 0.9668508400007196
steps: 169    lr: 0.0001     evaluation reward: 1.48
episode: 639   score: 1.0   memory length: 116893   epsilon: 0.9665498800007262
steps: 152    lr: 0.0001     evaluation reward: 1.49
episode: 640   score: 2.0   memory length: 117109   epsilon: 0.9661222000007355
steps: 216    lr: 0.0001     evaluation reward: 1.5
episode: 641   score: 1.0   memory length: 117262   epsilon: 0.965819260000742
steps: 153    lr: 0.0001     evaluation reward: 1.48
episode: 642   score: 1.0   memory length: 117413   epsilon: 0.9655202800007485
steps: 151    lr: 0.0001     evaluation reward: 1.48
episode: 643   score: 0.0   memory length: 117536   epsilon: 0.9652767400007538
steps: 123    lr: 0.0001     evaluation reward: 1.47
episode: 644   score: 1.0   memory length: 117687   epsilon: 0.9649777600007603
steps: 151    lr: 0.0001     evaluation reward: 1.45
episode: 645   score: 0.0   memory length: 117810   epsilon: 0.9647342200007656
steps: 123    lr: 0.0001     evaluation reward: 1.45
episode: 646   score: 1.0   memory length: 117979   epsilon: 0.9643996000007728
```

```
steps: 169    lr: 0.0001     evaluation reward: 1.42
episode: 647   score: 1.0   memory length: 118130   epsilon: 0.9641006200007793
steps: 151    lr: 0.0001     evaluation reward: 1.42
episode: 648   score: 2.0   memory length: 118327   epsilon: 0.9637105600007878
steps: 197    lr: 0.0001     evaluation reward: 1.43
episode: 649   score: 4.0   memory length: 118644   epsilon: 0.9630829000008014
steps: 317    lr: 0.0001     evaluation reward: 1.46
episode: 650   score: 1.0   memory length: 118815   epsilon: 0.9627443200008088
steps: 171    lr: 0.0001     evaluation reward: 1.47
episode: 651   score: 3.0   memory length: 119063   epsilon: 0.9622532800008194
steps: 248    lr: 0.0001     evaluation reward: 1.49
episode: 652   score: 1.0   memory length: 119232   epsilon: 0.9619186600008267
steps: 169    lr: 0.0001     evaluation reward: 1.48
episode: 653   score: 1.0   memory length: 119401   epsilon: 0.961584040000834
steps: 169    lr: 0.0001     evaluation reward: 1.49
episode: 654   score: 4.0   memory length: 119676   epsilon: 0.9610395400008458
steps: 275    lr: 0.0001     evaluation reward: 1.53
episode: 655   score: 0.0   memory length: 119799   epsilon: 0.9607960000008511
steps: 123    lr: 0.0001     evaluation reward: 1.51
episode: 656   score: 2.0   memory length: 119997   epsilon: 0.9604039600008596
steps: 198    lr: 0.0001     evaluation reward: 1.51
episode: 657   score: 4.0   memory length: 120295   epsilon: 0.9598139200008724
steps: 298    lr: 0.0001     evaluation reward: 1.53
episode: 658   score: 3.0   memory length: 120564   epsilon: 0.959281300000884
steps: 269    lr: 0.0001     evaluation reward: 1.52
episode: 659   score: 0.0   memory length: 120687   epsilon: 0.9590377600008893
steps: 123    lr: 0.0001     evaluation reward: 1.48
episode: 660   score: 2.0   memory length: 120904   epsilon: 0.9586081000008986
steps: 217    lr: 0.0001     evaluation reward: 1.5
episode: 661   score: 3.0   memory length: 121153   epsilon: 0.9581150800009093
steps: 249    lr: 0.0001     evaluation reward: 1.51
episode: 662   score: 1.0   memory length: 121322   epsilon: 0.9577804600009165
steps: 169    lr: 0.0001     evaluation reward: 1.5
episode: 663   score: 4.0   memory length: 121617   epsilon: 0.9571963600009292
steps: 295    lr: 0.0001     evaluation reward: 1.52
episode: 664   score: 2.0   memory length: 121835   epsilon: 0.9567647200009386
steps: 218    lr: 0.0001     evaluation reward: 1.54
episode: 665   score: 2.0   memory length: 122053   epsilon: 0.956333080000948
steps: 218    lr: 0.0001     evaluation reward: 1.55
episode: 666   score: 3.0   memory length: 122265   epsilon: 0.9559133200009571
steps: 212    lr: 0.0001     evaluation reward: 1.54
episode: 667   score: 1.0   memory length: 122437   epsilon: 0.9555727600009645
steps: 172    lr: 0.0001     evaluation reward: 1.53
episode: 668   score: 0.0   memory length: 122560   epsilon: 0.9553292200009698
steps: 123    lr: 0.0001     evaluation reward: 1.52
episode: 669   score: 1.0   memory length: 122710   epsilon: 0.9550322200009762
steps: 150    lr: 0.0001     evaluation reward: 1.53
episode: 670   score: 2.0   memory length: 122928   epsilon: 0.9546005800009856
```

```
steps: 218    lr: 0.0001     evaluation reward: 1.55
episode: 671   score: 2.0   memory length: 123127   epsilon: 0.9542065600009941
steps: 199    lr: 0.0001     evaluation reward: 1.56
episode: 672   score: 4.0   memory length: 123424   epsilon: 0.9536185000010069
steps: 297    lr: 0.0001     evaluation reward: 1.58
episode: 673   score: 0.0   memory length: 123547   epsilon: 0.9533749600010122
steps: 123    lr: 0.0001     evaluation reward: 1.58
episode: 674   score: 0.0   memory length: 123670   epsilon: 0.9531314200010175
steps: 123    lr: 0.0001     evaluation reward: 1.58
episode: 675   score: 2.0   memory length: 123850   epsilon: 0.9527750200010252
steps: 180    lr: 0.0001     evaluation reward: 1.6
episode: 676   score: 0.0   memory length: 123973   epsilon: 0.9525314800010305
steps: 123    lr: 0.0001     evaluation reward: 1.59
episode: 677   score: 2.0   memory length: 124171   epsilon: 0.952139440001039
steps: 198    lr: 0.0001     evaluation reward: 1.6
episode: 678   score: 5.0   memory length: 124514   epsilon: 0.9514603000010537
steps: 343    lr: 0.0001     evaluation reward: 1.63
episode: 679   score: 3.0   memory length: 124739   epsilon: 0.9510148000010634
steps: 225    lr: 0.0001     evaluation reward: 1.66
episode: 680   score: 0.0   memory length: 124862   epsilon: 0.9507712600010687
steps: 123    lr: 0.0001     evaluation reward: 1.65
episode: 681   score: 1.0   memory length: 125013   epsilon: 0.9504722800010752
steps: 151    lr: 0.0001     evaluation reward: 1.65
episode: 682   score: 1.0   memory length: 125182   epsilon: 0.9501376600010825
steps: 169    lr: 0.0001     evaluation reward: 1.64
episode: 683   score: 1.0   memory length: 125353   epsilon: 0.9497990800010898
steps: 171    lr: 0.0001     evaluation reward: 1.59
episode: 684   score: 0.0   memory length: 125475   epsilon: 0.9495575200010951
steps: 122    lr: 0.0001     evaluation reward: 1.57
episode: 685   score: 3.0   memory length: 125701   epsilon: 0.9491100400011048
steps: 226    lr: 0.0001     evaluation reward: 1.6
episode: 686   score: 0.0   memory length: 125823   epsilon: 0.94886848000111
steps: 122    lr: 0.0001     evaluation reward: 1.58
episode: 687   score: 1.0   memory length: 125975   epsilon: 0.9485675200011165
steps: 152    lr: 0.0001     evaluation reward: 1.56
episode: 688   score: 0.0   memory length: 126098   epsilon: 0.9483239800011218
steps: 123    lr: 0.0001     evaluation reward: 1.56
episode: 689   score: 0.0   memory length: 126221   epsilon: 0.9480804400011271
steps: 123    lr: 0.0001     evaluation reward: 1.53
episode: 690   score: 0.0   memory length: 126344   epsilon: 0.9478369000011324
steps: 123    lr: 0.0001     evaluation reward: 1.48
episode: 691   score: 3.0   memory length: 126611   epsilon: 0.9473082400011439
steps: 267    lr: 0.0001     evaluation reward: 1.5
episode: 692   score: 2.0   memory length: 126808   epsilon: 0.9469181800011524
steps: 197    lr: 0.0001     evaluation reward: 1.49
episode: 693   score: 1.0   memory length: 126976   epsilon: 0.9465855400011596
steps: 168    lr: 0.0001     evaluation reward: 1.47
episode: 694   score: 3.0   memory length: 127202   epsilon: 0.9461380600011693
```

```
steps: 226    lr: 0.0001     evaluation reward: 1.5
episode: 695   score: 0.0   memory length: 127325   epsilon: 0.9458945200011746
steps: 123    lr: 0.0001     evaluation reward: 1.48
episode: 696   score: 2.0   memory length: 127523   epsilon: 0.9455024800011831
steps: 198    lr: 0.0001     evaluation reward: 1.49
episode: 697   score: 4.0   memory length: 127813   epsilon: 0.9449282800011956
steps: 290    lr: 0.0001     evaluation reward: 1.51
episode: 698   score: 1.0   memory length: 127963   epsilon: 0.944631280001202
steps: 150    lr: 0.0001     evaluation reward: 1.5
episode: 699   score: 3.0   memory length: 128189   epsilon: 0.9441838000012117
steps: 226    lr: 0.0001     evaluation reward: 1.52
episode: 700   score: 1.0   memory length: 128341   epsilon: 0.9438828400012182
steps: 152    lr: 0.0001     evaluation reward: 1.52
episode: 701   score: 2.0   memory length: 128539   epsilon: 0.9434908000012268
steps: 198    lr: 0.0001     evaluation reward: 1.51
episode: 702   score: 1.0   memory length: 128708   epsilon: 0.943156180001234
steps: 169    lr: 0.0001     evaluation reward: 1.51
episode: 703   score: 1.0   memory length: 128878   epsilon: 0.9428195800012413
steps: 170    lr: 0.0001     evaluation reward: 1.52
episode: 704   score: 3.0   memory length: 129107   epsilon: 0.9423661600012512
steps: 229    lr: 0.0001     evaluation reward: 1.53
episode: 705   score: 0.0   memory length: 129229   epsilon: 0.9421246000012564
steps: 122    lr: 0.0001     evaluation reward: 1.51
episode: 706   score: 0.0   memory length: 129351   epsilon: 0.9418830400012617
steps: 122    lr: 0.0001     evaluation reward: 1.51
episode: 707   score: 0.0   memory length: 129474   epsilon: 0.941639500001267
steps: 123    lr: 0.0001     evaluation reward: 1.51
episode: 708   score: 3.0   memory length: 129718   epsilon: 0.9411563800012774
steps: 244    lr: 0.0001     evaluation reward: 1.54
episode: 709   score: 0.0   memory length: 129841   epsilon: 0.9409128400012827
steps: 123    lr: 0.0001     evaluation reward: 1.52
episode: 710   score: 1.0   memory length: 130011   epsilon: 0.94057624000129
steps: 170    lr: 0.0001     evaluation reward: 1.51
episode: 711   score: 1.0   memory length: 130180   epsilon: 0.9402416200012973
steps: 169    lr: 0.0001     evaluation reward: 1.52
episode: 712   score: 0.0   memory length: 130303   epsilon: 0.9399980800013026
steps: 123    lr: 0.0001     evaluation reward: 1.5
episode: 713   score: 2.0   memory length: 130501   epsilon: 0.9396060400013111
steps: 198    lr: 0.0001     evaluation reward: 1.49
episode: 714   score: 2.0   memory length: 130718   epsilon: 0.9391763800013204
steps: 217    lr: 0.0001     evaluation reward: 1.51
episode: 715   score: 3.0   memory length: 130988   epsilon: 0.938641780001332
steps: 270    lr: 0.0001     evaluation reward: 1.54
episode: 716   score: 1.0   memory length: 131158   epsilon: 0.9383051800013393
steps: 170    lr: 0.0001     evaluation reward: 1.55
episode: 717   score: 0.0   memory length: 131281   epsilon: 0.9380616400013446
steps: 123    lr: 0.0001     evaluation reward: 1.54
episode: 718   score: 4.0   memory length: 131598   epsilon: 0.9374339800013582
```

```
steps: 317    lr: 0.0001     evaluation reward: 1.58
episode: 719   score: 0.0   memory length: 131721   epsilon: 0.9371904400013635
steps: 123    lr: 0.0001     evaluation reward: 1.55
episode: 720   score: 2.0   memory length: 131940   epsilon: 0.936756820001373
steps: 219    lr: 0.0001     evaluation reward: 1.57
episode: 721   score: 0.0   memory length: 132063   epsilon: 0.9365132800013782
steps: 123    lr: 0.0001     evaluation reward: 1.56
episode: 722   score: 0.0   memory length: 132186   epsilon: 0.9362697400013835
steps: 123    lr: 0.0001     evaluation reward: 1.55
episode: 723   score: 4.0   memory length: 132483   epsilon: 0.9356816800013963
steps: 297    lr: 0.0001     evaluation reward: 1.55
episode: 724   score: 0.0   memory length: 132606   epsilon: 0.9354381400014016
steps: 123    lr: 0.0001     evaluation reward: 1.52
episode: 725   score: 4.0   memory length: 132882   epsilon: 0.9348916600014134
steps: 276    lr: 0.0001     evaluation reward: 1.54
episode: 726   score: 3.0   memory length: 133128   epsilon: 0.934404580001424
steps: 246    lr: 0.0001     evaluation reward: 1.57
episode: 727   score: 1.0   memory length: 133279   epsilon: 0.9341056000014305
steps: 151    lr: 0.0001     evaluation reward: 1.56
episode: 728   score: 0.0   memory length: 133402   epsilon: 0.9338620600014358
steps: 123    lr: 0.0001     evaluation reward: 1.54
episode: 729   score: 8.0   memory length: 133754   epsilon: 0.9331651000014509
steps: 352    lr: 0.0001     evaluation reward: 1.62
episode: 730   score: 2.0   memory length: 133951   epsilon: 0.9327750400014594
steps: 197    lr: 0.0001     evaluation reward: 1.61
episode: 731   score: 0.0   memory length: 134074   epsilon: 0.9325315000014647
steps: 123    lr: 0.0001     evaluation reward: 1.6
episode: 732   score: 0.0   memory length: 134197   epsilon: 0.93228796000147
steps: 123    lr: 0.0001     evaluation reward: 1.6
episode: 733   score: 0.0   memory length: 134320   epsilon: 0.9320444200014752
steps: 123    lr: 0.0001     evaluation reward: 1.56
episode: 734   score: 2.0   memory length: 134519   epsilon: 0.9316504000014838
steps: 199    lr: 0.0001     evaluation reward: 1.58
episode: 735   score: 0.0   memory length: 134642   epsilon: 0.9314068600014891
steps: 123    lr: 0.0001     evaluation reward: 1.58
episode: 736   score: 3.0   memory length: 134868   epsilon: 0.9309593800014988
steps: 226    lr: 0.0001     evaluation reward: 1.58
episode: 737   score: 0.0   memory length: 134990   epsilon: 0.930717820001504
steps: 122    lr: 0.0001     evaluation reward: 1.53
episode: 738   score: 4.0   memory length: 135285   epsilon: 0.9301337200015167
steps: 295    lr: 0.0001     evaluation reward: 1.56
episode: 739   score: 5.0   memory length: 135608   epsilon: 0.9294941800015306
steps: 323    lr: 0.0001     evaluation reward: 1.6
episode: 740   score: 0.0   memory length: 135731   epsilon: 0.9292506400015359
steps: 123    lr: 0.0001     evaluation reward: 1.58
episode: 741   score: 3.0   memory length: 135962   epsilon: 0.9287932600015458
steps: 231    lr: 0.0001     evaluation reward: 1.6
episode: 742   score: 0.0   memory length: 136085   epsilon: 0.9285497200015511
```

```
steps: 123    lr: 0.0001     evaluation reward: 1.59
episode: 743   score: 0.0   memory length: 136207   epsilon: 0.9283081600015564
steps: 122    lr: 0.0001     evaluation reward: 1.59
episode: 744   score: 4.0   memory length: 136500   epsilon: 0.927728020001569
steps: 293    lr: 0.0001     evaluation reward: 1.62
episode: 745   score: 2.0   memory length: 136717   epsilon: 0.9272983600015783
steps: 217    lr: 0.0001     evaluation reward: 1.64
episode: 746   score: 0.0   memory length: 136839   epsilon: 0.9270568000015835
steps: 122    lr: 0.0001     evaluation reward: 1.63
episode: 747   score: 2.0   memory length: 137025   epsilon: 0.9266885200015915
steps: 186    lr: 0.0001     evaluation reward: 1.64
episode: 748   score: 4.0   memory length: 137301   epsilon: 0.9261420400016034
steps: 276    lr: 0.0001     evaluation reward: 1.66
episode: 749   score: 3.0   memory length: 137549   epsilon: 0.925651000001614
steps: 248    lr: 0.0001     evaluation reward: 1.65
episode: 750   score: 2.0   memory length: 137731   epsilon: 0.9252906400016219
steps: 182    lr: 0.0001     evaluation reward: 1.66
episode: 751   score: 0.0   memory length: 137853   epsilon: 0.9250490800016271
steps: 122    lr: 0.0001     evaluation reward: 1.63
episode: 752   score: 1.0   memory length: 138021   epsilon: 0.9247164400016343
steps: 168    lr: 0.0001     evaluation reward: 1.63
episode: 753   score: 1.0   memory length: 138191   epsilon: 0.9243798400016416
steps: 170    lr: 0.0001     evaluation reward: 1.63
episode: 754   score: 0.0   memory length: 138314   epsilon: 0.9241363000016469
steps: 123    lr: 0.0001     evaluation reward: 1.59
episode: 755   score: 3.0   memory length: 138579   epsilon: 0.9236116000016583
steps: 265    lr: 0.0001     evaluation reward: 1.62
episode: 756   score: 1.0   memory length: 138729   epsilon: 0.9233146000016648
steps: 150    lr: 0.0001     evaluation reward: 1.61
episode: 757   score: 2.0   memory length: 138930   epsilon: 0.9229166200016734
steps: 201    lr: 0.0001     evaluation reward: 1.59
episode: 758   score: 3.0   memory length: 139203   epsilon: 0.9223760800016851
steps: 273    lr: 0.0001     evaluation reward: 1.59
episode: 759   score: 2.0   memory length: 139401   epsilon: 0.9219840400016936
steps: 198    lr: 0.0001     evaluation reward: 1.61
episode: 760   score: 2.0   memory length: 139599   epsilon: 0.9215920000017022
steps: 198    lr: 0.0001     evaluation reward: 1.61
episode: 761   score: 2.0   memory length: 139799   epsilon: 0.9211960000017108
steps: 200    lr: 0.0001     evaluation reward: 1.6
episode: 762   score: 1.0   memory length: 139971   epsilon: 0.9208554400017182
steps: 172    lr: 0.0001     evaluation reward: 1.6
episode: 763   score: 2.0   memory length: 140187   epsilon: 0.9204277600017274
steps: 216    lr: 0.0001     evaluation reward: 1.58
episode: 764   score: 2.0   memory length: 140385   epsilon: 0.920035720001736
steps: 198    lr: 0.0001     evaluation reward: 1.58
episode: 765   score: 3.0   memory length: 140635   epsilon: 0.9195407200017467
steps: 250    lr: 0.0001     evaluation reward: 1.59
episode: 766   score: 1.0   memory length: 140805   epsilon: 0.919204120001754
```

```
steps: 170    lr: 0.0001     evaluation reward: 1.57
episode: 767   score: 0.0   memory length: 140927   epsilon: 0.9189625600017592
steps: 122    lr: 0.0001     evaluation reward: 1.56
episode: 768   score: 4.0   memory length: 141203   epsilon: 0.9184160800017711
steps: 276    lr: 0.0001     evaluation reward: 1.6
episode: 769   score: 0.0   memory length: 141326   epsilon: 0.9181725400017764
steps: 123    lr: 0.0001     evaluation reward: 1.59
episode: 770   score: 1.0   memory length: 141477   epsilon: 0.9178735600017829
steps: 151    lr: 0.0001     evaluation reward: 1.58
episode: 771   score: 4.0   memory length: 141789   epsilon: 0.9172558000017963
steps: 312    lr: 0.0001     evaluation reward: 1.6
episode: 772   score: 1.0   memory length: 141939   epsilon: 0.9169588000018027
steps: 150    lr: 0.0001     evaluation reward: 1.57
episode: 773   score: 4.0   memory length: 142234   epsilon: 0.9163747000018154
steps: 295    lr: 0.0001     evaluation reward: 1.61
episode: 774   score: 1.0   memory length: 142404   epsilon: 0.9160381000018227
steps: 170    lr: 0.0001     evaluation reward: 1.62
episode: 775   score: 0.0   memory length: 142527   epsilon: 0.915794560001828
steps: 123    lr: 0.0001     evaluation reward: 1.6
episode: 776   score: 2.0   memory length: 142725   epsilon: 0.9154025200018365
steps: 198    lr: 0.0001     evaluation reward: 1.62
episode: 777   score: 3.0   memory length: 142989   epsilon: 0.9148798000018479
steps: 264    lr: 0.0001     evaluation reward: 1.63
episode: 778   score: 2.0   memory length: 143206   epsilon: 0.9144501400018572
steps: 217    lr: 0.0001     evaluation reward: 1.6
episode: 779   score: 5.0   memory length: 143524   epsilon: 0.9138205000018709
steps: 318    lr: 0.0001     evaluation reward: 1.62
episode: 780   score: 0.0   memory length: 143647   epsilon: 0.9135769600018762
steps: 123    lr: 0.0001     evaluation reward: 1.62
episode: 781   score: 0.0   memory length: 143770   epsilon: 0.9133334200018814
steps: 123    lr: 0.0001     evaluation reward: 1.61
episode: 782   score: 2.0   memory length: 143968   epsilon: 0.91294138000189
steps: 198    lr: 0.0001     evaluation reward: 1.62
episode: 783   score: 2.0   memory length: 144166   epsilon: 0.9125493400018985
steps: 198    lr: 0.0001     evaluation reward: 1.63
episode: 784   score: 1.0   memory length: 144317   epsilon: 0.912250360001905
steps: 151    lr: 0.0001     evaluation reward: 1.64
episode: 785   score: 2.0   memory length: 144515   epsilon: 0.9118583200019135
steps: 198    lr: 0.0001     evaluation reward: 1.63
episode: 786   score: 4.0   memory length: 144810   epsilon: 0.9112742200019261
steps: 295    lr: 0.0001     evaluation reward: 1.67
episode: 787   score: 3.0   memory length: 145057   epsilon: 0.9107851600019368
steps: 247    lr: 0.0001     evaluation reward: 1.69
episode: 788   score: 2.0   memory length: 145278   epsilon: 0.9103475800019463
steps: 221    lr: 0.0001     evaluation reward: 1.71
episode: 789   score: 1.0   memory length: 145446   epsilon: 0.9100149400019535
steps: 168    lr: 0.0001     evaluation reward: 1.72
episode: 790   score: 1.0   memory length: 145617   epsilon: 0.9096763600019608
```

```
steps: 171    lr: 0.0001     evaluation reward: 1.73
episode: 791   score: 1.0   memory length: 145788   epsilon: 0.9093377800019682
steps: 171    lr: 0.0001     evaluation reward: 1.71
episode: 792   score: 2.0   memory length: 145986   epsilon: 0.9089457400019767
steps: 198    lr: 0.0001     evaluation reward: 1.71
episode: 793   score: 1.0   memory length: 146156   epsilon: 0.908609140001984
steps: 170    lr: 0.0001     evaluation reward: 1.71
episode: 794   score: 2.0   memory length: 146371   epsilon: 0.9081834400019932
steps: 215    lr: 0.0001     evaluation reward: 1.7
episode: 795   score: 4.0   memory length: 146664   epsilon: 0.9076033000020058
steps: 293    lr: 0.0001     evaluation reward: 1.74
episode: 796   score: 1.0   memory length: 146814   epsilon: 0.9073063000020123
steps: 150    lr: 0.0001     evaluation reward: 1.73
episode: 797   score: 1.0   memory length: 146965   epsilon: 0.9070073200020188
steps: 151    lr: 0.0001     evaluation reward: 1.7
episode: 798   score: 1.0   memory length: 147136   epsilon: 0.9066687400020261
steps: 171    lr: 0.0001     evaluation reward: 1.7
episode: 799   score: 1.0   memory length: 147305   epsilon: 0.9063341200020334
steps: 169    lr: 0.0001     evaluation reward: 1.68
episode: 800   score: 3.0   memory length: 147570   epsilon: 0.9058094200020448
steps: 265    lr: 0.0001     evaluation reward: 1.7
episode: 801   score: 3.0   memory length: 147783   epsilon: 0.9053876800020539
steps: 213    lr: 0.0001     evaluation reward: 1.71
episode: 802   score: 0.0   memory length: 147905   epsilon: 0.9051461200020592
steps: 122    lr: 0.0001     evaluation reward: 1.7
episode: 803   score: 3.0   memory length: 148115   epsilon: 0.9047303200020682
steps: 210    lr: 0.0001     evaluation reward: 1.72
episode: 804   score: 2.0   memory length: 148334   epsilon: 0.9042967000020776
steps: 219    lr: 0.0001     evaluation reward: 1.71
episode: 805   score: 2.0   memory length: 148531   epsilon: 0.9039066400020861
steps: 197    lr: 0.0001     evaluation reward: 1.73
episode: 806   score: 0.0   memory length: 148654   epsilon: 0.9036631000020914
steps: 123    lr: 0.0001     evaluation reward: 1.73
episode: 807   score: 2.0   memory length: 148872   epsilon: 0.9032314600021008
steps: 218    lr: 0.0001     evaluation reward: 1.75
episode: 808   score: 3.0   memory length: 149099   epsilon: 0.9027820000021105
steps: 227    lr: 0.0001     evaluation reward: 1.75
episode: 809   score: 1.0   memory length: 149270   epsilon: 0.9024434200021179
steps: 171    lr: 0.0001     evaluation reward: 1.76
episode: 810   score: 2.0   memory length: 149487   epsilon: 0.9020137600021272
steps: 217    lr: 0.0001     evaluation reward: 1.77
episode: 811   score: 2.0   memory length: 149684   epsilon: 0.9016237000021357
steps: 197    lr: 0.0001     evaluation reward: 1.78
episode: 812   score: 1.0   memory length: 149836   epsilon: 0.9013227400021422
steps: 152    lr: 0.0001     evaluation reward: 1.79
episode: 813   score: 0.0   memory length: 149959   epsilon: 0.9010792000021475
steps: 123    lr: 0.0001     evaluation reward: 1.77
episode: 814   score: 0.0   memory length: 150082   epsilon: 0.9008356600021528
```

```
steps: 123    lr: 0.0001     evaluation reward: 1.75
episode: 815   score: 3.0   memory length: 150330   epsilon: 0.9003446200021634
steps: 248    lr: 0.0001     evaluation reward: 1.75
episode: 816   score: 2.0   memory length: 150527   epsilon: 0.8999545600021719
steps: 197    lr: 0.0001     evaluation reward: 1.76
episode: 817   score: 2.0   memory length: 150745   epsilon: 0.8995229200021813
steps: 218    lr: 0.0001     evaluation reward: 1.78
episode: 818   score: 2.0   memory length: 150948   epsilon: 0.89912098000219
steps: 203    lr: 0.0001     evaluation reward: 1.76
episode: 819   score: 2.0   memory length: 151146   epsilon: 0.8987289400021985
steps: 198    lr: 0.0001     evaluation reward: 1.78
episode: 820   score: 1.0   memory length: 151297   epsilon: 0.898429960002205
steps: 151    lr: 0.0001     evaluation reward: 1.77
episode: 821   score: 2.0   memory length: 151516   epsilon: 0.8979963400022144
steps: 219    lr: 0.0001     evaluation reward: 1.79
episode: 822   score: 5.0   memory length: 151838   epsilon: 0.8973587800022282
steps: 322    lr: 0.0001     evaluation reward: 1.84
episode: 823   score: 4.0   memory length: 152114   epsilon: 0.8968123000022401
steps: 276    lr: 0.0001     evaluation reward: 1.84
episode: 824   score: 1.0   memory length: 152265   epsilon: 0.8965133200022466
steps: 151    lr: 0.0001     evaluation reward: 1.85
episode: 825   score: 4.0   memory length: 152523   epsilon: 0.8960024800022577
steps: 258    lr: 0.0001     evaluation reward: 1.85
episode: 826   score: 2.0   memory length: 152722   epsilon: 0.8956084600022662
steps: 199    lr: 0.0001     evaluation reward: 1.84
episode: 827   score: 6.0   memory length: 153112   epsilon: 0.894836260002283
steps: 390    lr: 0.0001     evaluation reward: 1.89
episode: 828   score: 1.0   memory length: 153283   epsilon: 0.8944976800022904
steps: 171    lr: 0.0001     evaluation reward: 1.9
episode: 829   score: 5.0   memory length: 153610   epsilon: 0.8938502200023044
steps: 327    lr: 0.0001     evaluation reward: 1.87
episode: 830   score: 3.0   memory length: 153874   epsilon: 0.8933275000023158
steps: 264    lr: 0.0001     evaluation reward: 1.88
episode: 831   score: 3.0   memory length: 154138   epsilon: 0.8928047800023271
steps: 264    lr: 0.0001     evaluation reward: 1.91
episode: 832   score: 0.0   memory length: 154260   epsilon: 0.8925632200023323
steps: 122    lr: 0.0001     evaluation reward: 1.91
episode: 833   score: 1.0   memory length: 154429   epsilon: 0.8922286000023396
steps: 169    lr: 0.0001     evaluation reward: 1.92
episode: 834   score: 2.0   memory length: 154647   epsilon: 0.891796960002349
steps: 218    lr: 0.0001     evaluation reward: 1.92
episode: 835   score: 2.0   memory length: 154870   epsilon: 0.8913554200023586
steps: 223    lr: 0.0001     evaluation reward: 1.94
episode: 836   score: 1.0   memory length: 155038   epsilon: 0.8910227800023658
steps: 168    lr: 0.0001     evaluation reward: 1.92
episode: 837   score: 1.0   memory length: 155189   epsilon: 0.8907238000023723
steps: 151    lr: 0.0001     evaluation reward: 1.93
episode: 838   score: 1.0   memory length: 155359   epsilon: 0.8903872000023796
```

```
steps: 170    lr: 0.0001     evaluation reward: 1.9
episode: 839   score: 2.0   memory length: 155557   epsilon: 0.8899951600023881
steps: 198    lr: 0.0001     evaluation reward: 1.87
episode: 840   score: 3.0   memory length: 155805   epsilon: 0.8895041200023988
steps: 248    lr: 0.0001     evaluation reward: 1.9
episode: 841   score: 2.0   memory length: 156023   epsilon: 0.8890724800024081
steps: 218    lr: 0.0001     evaluation reward: 1.89
episode: 842   score: 8.0   memory length: 156487   epsilon: 0.8881537600024281
steps: 464    lr: 0.0001     evaluation reward: 1.97
episode: 843   score: 3.0   memory length: 156755   epsilon: 0.8876231200024396
steps: 268    lr: 0.0001     evaluation reward: 2.0
episode: 844   score: 2.0   memory length: 156953   epsilon: 0.8872310800024481
steps: 198    lr: 0.0001     evaluation reward: 1.98
episode: 845   score: 2.0   memory length: 157171   epsilon: 0.8867994400024575
steps: 218    lr: 0.0001     evaluation reward: 1.98
episode: 846   score: 1.0   memory length: 157322   epsilon: 0.886500460002464
steps: 151    lr: 0.0001     evaluation reward: 1.99
episode: 847   score: 0.0   memory length: 157445   epsilon: 0.8862569200024693
steps: 123    lr: 0.0001     evaluation reward: 1.97
episode: 848   score: 2.0   memory length: 157645   epsilon: 0.8858609200024778
steps: 200    lr: 0.0001     evaluation reward: 1.95
episode: 849   score: 4.0   memory length: 157942   epsilon: 0.8852728600024906
steps: 297    lr: 0.0001     evaluation reward: 1.96
episode: 850   score: 1.0   memory length: 158092   epsilon: 0.8849758600024971
steps: 150    lr: 0.0001     evaluation reward: 1.95
episode: 851   score: 2.0   memory length: 158292   epsilon: 0.8845798600025057
steps: 200    lr: 0.0001     evaluation reward: 1.97
episode: 852   score: 2.0   memory length: 158490   epsilon: 0.8841878200025142
steps: 198    lr: 0.0001     evaluation reward: 1.98
episode: 853   score: 2.0   memory length: 158689   epsilon: 0.8837938000025227
steps: 199    lr: 0.0001     evaluation reward: 1.99
episode: 854   score: 4.0   memory length: 158987   epsilon: 0.8832037600025355
steps: 298    lr: 0.0001     evaluation reward: 2.03
episode: 855   score: 7.0   memory length: 159409   epsilon: 0.8823682000025537
steps: 422    lr: 0.0001     evaluation reward: 2.07
episode: 856   score: 0.0   memory length: 159532   epsilon: 0.882124660002559
steps: 123    lr: 0.0001     evaluation reward: 2.06
episode: 857   score: 5.0   memory length: 159860   epsilon: 0.8814752200025731
steps: 328    lr: 0.0001     evaluation reward: 2.09
episode: 858   score: 5.0   memory length: 160222   epsilon: 0.8807584600025886
steps: 362    lr: 0.0001     evaluation reward: 2.11
episode: 859   score: 1.0   memory length: 160391   epsilon: 0.8804238400025959
steps: 169    lr: 0.0001     evaluation reward: 2.1
episode: 860   score: 3.0   memory length: 160637   epsilon: 0.8799367600026065
steps: 246    lr: 0.0001     evaluation reward: 2.11
episode: 861   score: 0.0   memory length: 160760   epsilon: 0.8796932200026117
steps: 123    lr: 0.0001     evaluation reward: 2.09
episode: 862   score: 1.0   memory length: 160930   epsilon: 0.879356620002619
```

```
steps: 170    lr: 0.0001     evaluation reward: 2.09
episode: 863   score: 3.0   memory length: 161177   epsilon: 0.8788675600026297
steps: 247    lr: 0.0001     evaluation reward: 2.1
episode: 864   score: 5.0   memory length: 161520   epsilon: 0.8781884200026444
steps: 343    lr: 0.0001     evaluation reward: 2.13
episode: 865   score: 3.0   memory length: 161767   epsilon: 0.877699360002655
steps: 247    lr: 0.0001     evaluation reward: 2.13
episode: 866   score: 1.0   memory length: 161937   epsilon: 0.8773627600026623
steps: 170    lr: 0.0001     evaluation reward: 2.13
episode: 867   score: 1.0   memory length: 162087   epsilon: 0.8770657600026688
steps: 150    lr: 0.0001     evaluation reward: 2.14
episode: 868   score: 0.0   memory length: 162210   epsilon: 0.8768222200026741
steps: 123    lr: 0.0001     evaluation reward: 2.1
episode: 869   score: 4.0   memory length: 162466   epsilon: 0.8763153400026851
steps: 256    lr: 0.0001     evaluation reward: 2.14
episode: 870   score: 0.0   memory length: 162589   epsilon: 0.8760718000026904
steps: 123    lr: 0.0001     evaluation reward: 2.13
episode: 871   score: 1.0   memory length: 162740   epsilon: 0.8757728200026969
steps: 151    lr: 0.0001     evaluation reward: 2.1
episode: 872   score: 1.0   memory length: 162890   epsilon: 0.8754758200027033
steps: 150    lr: 0.0001     evaluation reward: 2.1
episode: 873   score: 0.0   memory length: 163013   epsilon: 0.8752322800027086
steps: 123    lr: 0.0001     evaluation reward: 2.06
episode: 874   score: 0.0   memory length: 163136   epsilon: 0.8749887400027139
steps: 123    lr: 0.0001     evaluation reward: 2.05
episode: 875   score: 4.0   memory length: 163434   epsilon: 0.8743987000027267
steps: 298    lr: 0.0001     evaluation reward: 2.09
episode: 876   score: 4.0   memory length: 163707   epsilon: 0.8738581600027384
steps: 273    lr: 0.0001     evaluation reward: 2.11
episode: 877   score: 1.0   memory length: 163858   epsilon: 0.8735591800027449
steps: 151    lr: 0.0001     evaluation reward: 2.09
episode: 878   score: 2.0   memory length: 164077   epsilon: 0.8731255600027543
steps: 219    lr: 0.0001     evaluation reward: 2.09
episode: 879   score: 2.0   memory length: 164293   epsilon: 0.8726978800027636
steps: 216    lr: 0.0001     evaluation reward: 2.06
episode: 880   score: 2.0   memory length: 164493   epsilon: 0.8723018800027722
steps: 200    lr: 0.0001     evaluation reward: 2.08
episode: 881   score: 1.0   memory length: 164644   epsilon: 0.8720029000027787
steps: 151    lr: 0.0001     evaluation reward: 2.09
episode: 882   score: 0.0   memory length: 164767   epsilon: 0.871759360002784
steps: 123    lr: 0.0001     evaluation reward: 2.07
episode: 883   score: 0.0   memory length: 164889   epsilon: 0.8715178000027892
steps: 122    lr: 0.0001     evaluation reward: 2.05
episode: 884   score: 4.0   memory length: 165165   epsilon: 0.8709713200028011
steps: 276    lr: 0.0001     evaluation reward: 2.08
episode: 885   score: 2.0   memory length: 165382   epsilon: 0.8705416600028104
steps: 217    lr: 0.0001     evaluation reward: 2.08
episode: 886   score: 5.0   memory length: 165725   epsilon: 0.8698625200028252
```

```
steps: 343    lr: 0.0001     evaluation reward: 2.09
episode: 887   score: 1.0   memory length: 165875   epsilon: 0.8695655200028316
steps: 150    lr: 0.0001     evaluation reward: 2.07
episode: 888   score: 0.0   memory length: 165998   epsilon: 0.8693219800028369
steps: 123    lr: 0.0001     evaluation reward: 2.05
episode: 889   score: 2.0   memory length: 166217   epsilon: 0.8688883600028463
steps: 219    lr: 0.0001     evaluation reward: 2.06
episode: 890   score: 1.0   memory length: 166368   epsilon: 0.8685893800028528
steps: 151    lr: 0.0001     evaluation reward: 2.06
episode: 891   score: 1.0   memory length: 166537   epsilon: 0.8682547600028601
steps: 169    lr: 0.0001     evaluation reward: 2.06
episode: 892   score: 0.0   memory length: 166660   epsilon: 0.8680112200028653
steps: 123    lr: 0.0001     evaluation reward: 2.04
episode: 893   score: 2.0   memory length: 166877   epsilon: 0.8675815600028747
steps: 217    lr: 0.0001     evaluation reward: 2.05
episode: 894   score: 3.0   memory length: 167089   epsilon: 0.8671618000028838
steps: 212    lr: 0.0001     evaluation reward: 2.06
episode: 895   score: 1.0   memory length: 167240   epsilon: 0.8668628200028903
steps: 151    lr: 0.0001     evaluation reward: 2.03
episode: 896   score: 0.0   memory length: 167363   epsilon: 0.8666192800028956
steps: 123    lr: 0.0001     evaluation reward: 2.02
episode: 897   score: 1.0   memory length: 167531   epsilon: 0.8662866400029028
steps: 168    lr: 0.0001     evaluation reward: 2.02
episode: 898   score: 3.0   memory length: 167795   epsilon: 0.8657639200029141
steps: 264    lr: 0.0001     evaluation reward: 2.04
episode: 899   score: 1.0   memory length: 167946   epsilon: 0.8654649400029206
steps: 151    lr: 0.0001     evaluation reward: 2.04
episode: 900   score: 1.0   memory length: 168117   epsilon: 0.865126360002928
steps: 171    lr: 0.0001     evaluation reward: 2.02
episode: 901   score: 0.0   memory length: 168239   epsilon: 0.8648848000029332
steps: 122    lr: 0.0001     evaluation reward: 1.99
episode: 902   score: 4.0   memory length: 168535   epsilon: 0.8642987200029459
steps: 296    lr: 0.0001     evaluation reward: 2.03
episode: 903   score: 2.0   memory length: 168733   epsilon: 0.8639066800029545
steps: 198    lr: 0.0001     evaluation reward: 2.02
episode: 904   score: 3.0   memory length: 168977   epsilon: 0.8634235600029649
steps: 244    lr: 0.0001     evaluation reward: 2.03
episode: 905   score: 4.0   memory length: 169233   epsilon: 0.8629166800029759
steps: 256    lr: 0.0001     evaluation reward: 2.05
episode: 906   score: 3.0   memory length: 169458   epsilon: 0.8624711800029856
steps: 225    lr: 0.0001     evaluation reward: 2.08
episode: 907   score: 4.0   memory length: 169754   epsilon: 0.8618851000029983
steps: 296    lr: 0.0001     evaluation reward: 2.1
episode: 908   score: 1.0   memory length: 169924   epsilon: 0.8615485000030056
steps: 170    lr: 0.0001     evaluation reward: 2.08
episode: 909   score: 2.0   memory length: 170122   epsilon: 0.8611564600030142
steps: 198    lr: 0.0001     evaluation reward: 2.09
episode: 910   score: 6.0   memory length: 170518   epsilon: 0.8603723800030312
```

```
steps: 396    lr: 0.0001     evaluation reward: 2.13
episode: 911   score: 1.0   memory length: 170688   epsilon: 0.8600357800030385
steps: 170    lr: 0.0001     evaluation reward: 2.12
episode: 912   score: 4.0   memory length: 170983   epsilon: 0.8594516800030512
steps: 295    lr: 0.0001     evaluation reward: 2.15
episode: 913   score: 4.0   memory length: 171258   epsilon: 0.858907180003063
steps: 275    lr: 0.0001     evaluation reward: 2.19
episode: 914   score: 1.0   memory length: 171409   epsilon: 0.8586082000030695
steps: 151    lr: 0.0001     evaluation reward: 2.2
episode: 915   score: 2.0   memory length: 171607   epsilon: 0.858216160003078
steps: 198    lr: 0.0001     evaluation reward: 2.19
episode: 916   score: 4.0   memory length: 171882   epsilon: 0.8576716600030898
steps: 275    lr: 0.0001     evaluation reward: 2.21
episode: 917   score: 0.0   memory length: 172005   epsilon: 0.8574281200030951
steps: 123    lr: 0.0001     evaluation reward: 2.19
episode: 918   score: 3.0   memory length: 172230   epsilon: 0.8569826200031048
steps: 225    lr: 0.0001     evaluation reward: 2.2
episode: 919   score: 4.0   memory length: 172504   epsilon: 0.8564401000031165
steps: 274    lr: 0.0001     evaluation reward: 2.22
episode: 920   score: 5.0   memory length: 172848   epsilon: 0.8557589800031313
steps: 344    lr: 0.0001     evaluation reward: 2.26
episode: 921   score: 0.0   memory length: 172971   epsilon: 0.8555154400031366
steps: 123    lr: 0.0001     evaluation reward: 2.24
episode: 922   score: 1.0   memory length: 173141   epsilon: 0.8551788400031439
steps: 170    lr: 0.0001     evaluation reward: 2.2
episode: 923   score: 0.0   memory length: 173263   epsilon: 0.8549372800031492
steps: 122    lr: 0.0001     evaluation reward: 2.16
episode: 924   score: 2.0   memory length: 173461   epsilon: 0.8545452400031577
steps: 198    lr: 0.0001     evaluation reward: 2.17
episode: 925   score: 3.0   memory length: 173688   epsilon: 0.8540957800031674
steps: 227    lr: 0.0001     evaluation reward: 2.16
episode: 926   score: 5.0   memory length: 174053   epsilon: 0.8533730800031831
steps: 365    lr: 0.0001     evaluation reward: 2.19
episode: 927   score: 3.0   memory length: 174278   epsilon: 0.8529275800031928
steps: 225    lr: 0.0001     evaluation reward: 2.16
episode: 928   score: 3.0   memory length: 174525   epsilon: 0.8524385200032034
steps: 247    lr: 0.0001     evaluation reward: 2.18
episode: 929   score: 5.0   memory length: 174815   epsilon: 0.8518643200032159
steps: 290    lr: 0.0001     evaluation reward: 2.18
episode: 930   score: 2.0   memory length: 175035   epsilon: 0.8514287200032253
steps: 220    lr: 0.0001     evaluation reward: 2.17
episode: 931   score: 2.0   memory length: 175233   epsilon: 0.8510366800032338
steps: 198    lr: 0.0001     evaluation reward: 2.16
episode: 932   score: 6.0   memory length: 175641   epsilon: 0.8502288400032514
steps: 408    lr: 0.0001     evaluation reward: 2.22
episode: 933   score: 2.0   memory length: 175859   epsilon: 0.8497972000032608
steps: 218    lr: 0.0001     evaluation reward: 2.23
episode: 934   score: 1.0   memory length: 176009   epsilon: 0.8495002000032672
```

```
steps: 150    lr: 0.0001     evaluation reward: 2.22
episode: 935   score: 3.0   memory length: 176256   epsilon: 0.8490111400032778
steps: 247    lr: 0.0001     evaluation reward: 2.23
episode: 936   score: 3.0   memory length: 176486   epsilon: 0.8485557400032877
steps: 230    lr: 0.0001     evaluation reward: 2.25
episode: 937   score: 3.0   memory length: 176733   epsilon: 0.8480666800032983
steps: 247    lr: 0.0001     evaluation reward: 2.27
episode: 938   score: 2.0   memory length: 176951   epsilon: 0.8476350400033077
steps: 218    lr: 0.0001     evaluation reward: 2.28
episode: 939   score: 1.0   memory length: 177122   epsilon: 0.847296460003315
steps: 171    lr: 0.0001     evaluation reward: 2.27
episode: 940   score: 2.0   memory length: 177320   epsilon: 0.8469044200033236
steps: 198    lr: 0.0001     evaluation reward: 2.26
episode: 941   score: 2.0   memory length: 177538   epsilon: 0.8464727800033329
steps: 218    lr: 0.0001     evaluation reward: 2.26
episode: 942   score: 3.0   memory length: 177769   epsilon: 0.8460154000033429
steps: 231    lr: 0.0001     evaluation reward: 2.21
episode: 943   score: 0.0   memory length: 177891   epsilon: 0.8457738400033481
steps: 122    lr: 0.0001     evaluation reward: 2.18
episode: 944   score: 0.0   memory length: 178014   epsilon: 0.8455303000033534
steps: 123    lr: 0.0001     evaluation reward: 2.16
episode: 945   score: 4.0   memory length: 178296   epsilon: 0.8449719400033655
steps: 282    lr: 0.0001     evaluation reward: 2.18
episode: 946   score: 5.0   memory length: 178613   epsilon: 0.8443442800033791
steps: 317    lr: 0.0001     evaluation reward: 2.22
episode: 947   score: 1.0   memory length: 178782   epsilon: 0.8440096600033864
steps: 169    lr: 0.0001     evaluation reward: 2.23
episode: 948   score: 3.0   memory length: 179030   epsilon: 0.8435186200033971
steps: 248    lr: 0.0001     evaluation reward: 2.24
episode: 949   score: 3.0   memory length: 179275   epsilon: 0.8430335200034076
steps: 245    lr: 0.0001     evaluation reward: 2.23
episode: 950   score: 2.0   memory length: 179473   epsilon: 0.8426414800034161
steps: 198    lr: 0.0001     evaluation reward: 2.24
episode: 951   score: 4.0   memory length: 179790   epsilon: 0.8420138200034297
steps: 317    lr: 0.0001     evaluation reward: 2.26
episode: 952   score: 0.0   memory length: 179913   epsilon: 0.841770280003435
steps: 123    lr: 0.0001     evaluation reward: 2.24
episode: 953   score: 2.0   memory length: 180114   epsilon: 0.8413723000034437
steps: 201    lr: 0.0001     evaluation reward: 2.24
episode: 954   score: 3.0   memory length: 180342   epsilon: 0.8409208600034535
steps: 228    lr: 0.0001     evaluation reward: 2.23
episode: 955   score: 2.0   memory length: 180541   epsilon: 0.840526840003462
steps: 199    lr: 0.0001     evaluation reward: 2.18
episode: 956   score: 3.0   memory length: 180770   epsilon: 0.8400734200034718
steps: 229    lr: 0.0001     evaluation reward: 2.21
episode: 957   score: 3.0   memory length: 180997   epsilon: 0.8396239600034816
steps: 227    lr: 0.0001     evaluation reward: 2.19
episode: 958   score: 1.0   memory length: 181165   epsilon: 0.8392913200034888
```

```
steps: 168    lr: 0.0001     evaluation reward: 2.15
episode: 959   score: 1.0   memory length: 181316   epsilon: 0.8389923400034953
steps: 151    lr: 0.0001     evaluation reward: 2.15
episode: 960   score: 1.0   memory length: 181485   epsilon: 0.8386577200035026
steps: 169    lr: 0.0001     evaluation reward: 2.13
episode: 961   score: 3.0   memory length: 181711   epsilon: 0.8382102400035123
steps: 226    lr: 0.0001     evaluation reward: 2.16
episode: 962   score: 4.0   memory length: 182009   epsilon: 0.8376202000035251
steps: 298    lr: 0.0001     evaluation reward: 2.19
episode: 963   score: 2.0   memory length: 182206   epsilon: 0.8372301400035336
steps: 197    lr: 0.0001     evaluation reward: 2.18
episode: 964   score: 1.0   memory length: 182357   epsilon: 0.8369311600035401
steps: 151    lr: 0.0001     evaluation reward: 2.14
episode: 965   score: 3.0   memory length: 182628   epsilon: 0.8363945800035517
steps: 271    lr: 0.0001     evaluation reward: 2.14
episode: 966   score: 2.0   memory length: 182826   epsilon: 0.8360025400035602
steps: 198    lr: 0.0001     evaluation reward: 2.15
episode: 967   score: 5.0   memory length: 183150   epsilon: 0.8353610200035742
steps: 324    lr: 0.0001     evaluation reward: 2.19
episode: 968   score: 0.0   memory length: 183273   epsilon: 0.8351174800035794
steps: 123    lr: 0.0001     evaluation reward: 2.19
episode: 969   score: 1.0   memory length: 183442   epsilon: 0.8347828600035867
steps: 169    lr: 0.0001     evaluation reward: 2.16
episode: 970   score: 3.0   memory length: 183692   epsilon: 0.8342878600035974
steps: 250    lr: 0.0001     evaluation reward: 2.19
episode: 971   score: 1.0   memory length: 183861   epsilon: 0.8339532400036047
steps: 169    lr: 0.0001     evaluation reward: 2.19
episode: 972   score: 2.0   memory length: 184042   epsilon: 0.8335948600036125
steps: 181    lr: 0.0001     evaluation reward: 2.2
episode: 973   score: 3.0   memory length: 184288   epsilon: 0.8331077800036231
steps: 246    lr: 0.0001     evaluation reward: 2.23
episode: 974   score: 0.0   memory length: 184411   epsilon: 0.8328642400036284
steps: 123    lr: 0.0001     evaluation reward: 2.23
episode: 975   score: 2.0   memory length: 184609   epsilon: 0.8324722000036369
steps: 198    lr: 0.0001     evaluation reward: 2.21
episode: 976   score: 1.0   memory length: 184760   epsilon: 0.8321732200036434
steps: 151    lr: 0.0001     evaluation reward: 2.18
episode: 977   score: 11.0   memory length: 185230   epsilon: 0.8312426200036636
steps: 470    lr: 0.0001     evaluation reward: 2.28
episode: 978   score: 3.0   memory length: 185477   epsilon: 0.8307535600036742
steps: 247    lr: 0.0001     evaluation reward: 2.29
episode: 979   score: 3.0   memory length: 185687   epsilon: 0.8303377600036832
steps: 210    lr: 0.0001     evaluation reward: 2.3
episode: 980   score: 4.0   memory length: 185965   epsilon: 0.8297873200036952
steps: 278    lr: 0.0001     evaluation reward: 2.32
episode: 981   score: 1.0   memory length: 186116   epsilon: 0.8294883400037016
steps: 151    lr: 0.0001     evaluation reward: 2.32
episode: 982   score: 1.0   memory length: 186286   epsilon: 0.829151740003709
```

```
steps: 170    lr: 0.0001     evaluation reward: 2.33
episode: 983   score: 3.0   memory length: 186550   epsilon: 0.8286290200037203
steps: 264    lr: 0.0001     evaluation reward: 2.36
episode: 984   score: 3.0   memory length: 186776   epsilon: 0.82818154000373
steps: 226    lr: 0.0001     evaluation reward: 2.35
episode: 985   score: 6.0   memory length: 187150   epsilon: 0.8274410200037461
steps: 374    lr: 0.0001     evaluation reward: 2.39
episode: 986   score: 2.0   memory length: 187367   epsilon: 0.8270113600037554
steps: 217    lr: 0.0001     evaluation reward: 2.36
episode: 987   score: 1.0   memory length: 187538   epsilon: 0.8266727800037628
steps: 171    lr: 0.0001     evaluation reward: 2.36
episode: 988   score: 0.0   memory length: 187661   epsilon: 0.826429240003768
steps: 123    lr: 0.0001     evaluation reward: 2.36
episode: 989   score: 5.0   memory length: 188029   epsilon: 0.8257006000037839
steps: 368    lr: 0.0001     evaluation reward: 2.39
episode: 990   score: 3.0   memory length: 188258   epsilon: 0.8252471800037937
steps: 229    lr: 0.0001     evaluation reward: 2.41
episode: 991   score: 2.0   memory length: 188456   epsilon: 0.8248551400038022
steps: 198    lr: 0.0001     evaluation reward: 2.42
episode: 992   score: 3.0   memory length: 188681   epsilon: 0.8244096400038119
steps: 225    lr: 0.0001     evaluation reward: 2.45
episode: 993   score: 1.0   memory length: 188831   epsilon: 0.8241126400038183
steps: 150    lr: 0.0001     evaluation reward: 2.44
episode: 994   score: 2.0   memory length: 189028   epsilon: 0.8237225800038268
steps: 197    lr: 0.0001     evaluation reward: 2.43
episode: 995   score: 5.0   memory length: 189349   epsilon: 0.8230870000038406
steps: 321    lr: 0.0001     evaluation reward: 2.47
episode: 996   score: 1.0   memory length: 189500   epsilon: 0.8227880200038471
steps: 151    lr: 0.0001     evaluation reward: 2.48
episode: 997   score: 3.0   memory length: 189745   epsilon: 0.8223029200038576
steps: 245    lr: 0.0001     evaluation reward: 2.5
episode: 998   score: 0.0   memory length: 189868   epsilon: 0.8220593800038629
steps: 123    lr: 0.0001     evaluation reward: 2.47
episode: 999   score: 4.0   memory length: 190164   epsilon: 0.8214733000038756
steps: 296    lr: 0.0001     evaluation reward: 2.5
episode: 1000   score: 2.0   memory length: 190382   epsilon: 0.821041660003885
steps: 218    lr: 0.0001     evaluation reward: 2.51
episode: 1001   score: 2.0   memory length: 190580   epsilon: 0.8206496200038935
steps: 198    lr: 0.0001     evaluation reward: 2.53
episode: 1002   score: 1.0   memory length: 190731   epsilon: 0.8203506400039
steps: 151    lr: 0.0001     evaluation reward: 2.5
episode: 1003   score: 3.0   memory length: 190975   epsilon: 0.8198675200039105
steps: 244    lr: 0.0001     evaluation reward: 2.51
episode: 1004   score: 1.0   memory length: 191145   epsilon: 0.8195309200039178
steps: 170    lr: 0.0001     evaluation reward: 2.49
episode: 1005   score: 3.0   memory length: 191390   epsilon: 0.8190458200039283
steps: 245    lr: 0.0001     evaluation reward: 2.48
episode: 1006   score: 3.0   memory length: 191638   epsilon: 0.818554780003939
```

```
steps: 248    lr: 0.0001     evaluation reward: 2.48
episode: 1007   score: 2.0   memory length: 191836   epsilon: 0.8181627400039475
steps: 198    lr: 0.0001     evaluation reward: 2.46
episode: 1008   score: 2.0   memory length: 192051   epsilon: 0.8177370400039567
steps: 215    lr: 0.0001     evaluation reward: 2.47
episode: 1009   score: 4.0   memory length: 192331   epsilon: 0.8171826400039688
steps: 280    lr: 0.0001     evaluation reward: 2.49
episode: 1010   score: 5.0   memory length: 192627   epsilon: 0.8165965600039815
steps: 296    lr: 0.0001     evaluation reward: 2.48
episode: 1011   score: 3.0   memory length: 192873   epsilon: 0.8161094800039921
steps: 246    lr: 0.0001     evaluation reward: 2.5
episode: 1012   score: 0.0   memory length: 192995   epsilon: 0.8158679200039973
steps: 122    lr: 0.0001     evaluation reward: 2.46
episode: 1013   score: 1.0   memory length: 193145   epsilon: 0.8155709200040038
steps: 150    lr: 0.0001     evaluation reward: 2.43
episode: 1014   score: 5.0   memory length: 193473   epsilon: 0.8149214800040179
steps: 328    lr: 0.0001     evaluation reward: 2.47
episode: 1015   score: 4.0   memory length: 193760   epsilon: 0.8143532200040302
steps: 287    lr: 0.0001     evaluation reward: 2.49
episode: 1016   score: 4.0   memory length: 194036   epsilon: 0.8138067400040421
steps: 276    lr: 0.0001     evaluation reward: 2.49
episode: 1017   score: 2.0   memory length: 194234   epsilon: 0.8134147000040506
steps: 198    lr: 0.0001     evaluation reward: 2.51
episode: 1018   score: 3.0   memory length: 194502   epsilon: 0.8128840600040621
steps: 268    lr: 0.0001     evaluation reward: 2.51
episode: 1019   score: 2.0   memory length: 194721   epsilon: 0.8124504400040715
steps: 219    lr: 0.0001     evaluation reward: 2.49
episode: 1020   score: 3.0   memory length: 194968   epsilon: 0.8119613800040821
steps: 247    lr: 0.0001     evaluation reward: 2.47
episode: 1021   score: 2.0   memory length: 195185   epsilon: 0.8115317200040915
steps: 217    lr: 0.0001     evaluation reward: 2.49
episode: 1022   score: 2.0   memory length: 195367   epsilon: 0.8111713600040993
steps: 182    lr: 0.0001     evaluation reward: 2.5
episode: 1023   score: 4.0   memory length: 195669   epsilon: 0.8105734000041123
steps: 302    lr: 0.0001     evaluation reward: 2.54
episode: 1024   score: 0.0   memory length: 195791   epsilon: 0.8103318400041175
steps: 122    lr: 0.0001     evaluation reward: 2.52
episode: 1025   score: 1.0   memory length: 195942   epsilon: 0.810032860004124
steps: 151    lr: 0.0001     evaluation reward: 2.5
episode: 1026   score: 3.0   memory length: 196171   epsilon: 0.8095794400041338
steps: 229    lr: 0.0001     evaluation reward: 2.48
episode: 1027   score: 3.0   memory length: 196399   epsilon: 0.8091280000041436
steps: 228    lr: 0.0001     evaluation reward: 2.48
episode: 1028   score: 3.0   memory length: 196648   epsilon: 0.8086349800041543
steps: 249    lr: 0.0001     evaluation reward: 2.48
episode: 1029   score: 1.0   memory length: 196818   epsilon: 0.8082983800041617
steps: 170    lr: 0.0001     evaluation reward: 2.44
episode: 1030   score: 1.0   memory length: 196988   epsilon: 0.807961780004169
```

```
steps: 170    lr: 0.0001     evaluation reward: 2.43
episode: 1031   score: 3.0   memory length: 197199   epsilon: 0.807544000004178
steps: 211    lr: 0.0001     evaluation reward: 2.44
episode: 1032   score: 1.0   memory length: 197368   epsilon: 0.8072093800041853
steps: 169    lr: 0.0001     evaluation reward: 2.39
episode: 1033   score: 2.0   memory length: 197586   epsilon: 0.8067777400041947
steps: 218    lr: 0.0001     evaluation reward: 2.39
episode: 1034   score: 3.0   memory length: 197812   epsilon: 0.8063302600042044
steps: 226    lr: 0.0001     evaluation reward: 2.41
episode: 1035   score: 2.0   memory length: 198029   epsilon: 0.8059006000042137
steps: 217    lr: 0.0001     evaluation reward: 2.4
episode: 1036   score: 1.0   memory length: 198198   epsilon: 0.805565980004221
steps: 169    lr: 0.0001     evaluation reward: 2.38
episode: 1037   score: 2.0   memory length: 198377   epsilon: 0.8052115600042287
steps: 179    lr: 0.0001     evaluation reward: 2.37
episode: 1038   score: 1.0   memory length: 198549   epsilon: 0.8048710000042361
steps: 172    lr: 0.0001     evaluation reward: 2.36
episode: 1039   score: 3.0   memory length: 198777   epsilon: 0.8044195600042459
steps: 228    lr: 0.0001     evaluation reward: 2.38
episode: 1040   score: 2.0   memory length: 198960   epsilon: 0.8040572200042537
steps: 183    lr: 0.0001     evaluation reward: 2.38
episode: 1041   score: 0.0   memory length: 199083   epsilon: 0.803813680004259
steps: 123    lr: 0.0001     evaluation reward: 2.36
episode: 1042   score: 6.0   memory length: 199407   epsilon: 0.8031721600042729
steps: 324    lr: 0.0001     evaluation reward: 2.39
episode: 1043   score: 3.0   memory length: 199653   epsilon: 0.8026850800042835
steps: 246    lr: 0.0001     evaluation reward: 2.42
episode: 1044   score: 3.0   memory length: 199879   epsilon: 0.8022376000042932
steps: 226    lr: 0.0001     evaluation reward: 2.45
episode: 1045   score: 0.0   memory length: 200002   epsilon: 0.8019940600042985
steps: 123    lr: 4e-05     evaluation reward: 2.41
episode: 1046   score: 2.0   memory length: 200219   epsilon: 0.8015644000043078
steps: 217    lr: 4e-05     evaluation reward: 2.38
episode: 1047   score: 5.0   memory length: 200565   epsilon: 0.8008793200043227
steps: 346    lr: 4e-05     evaluation reward: 2.42
episode: 1048   score: 4.0   memory length: 200859   epsilon: 0.8002972000043354
steps: 294    lr: 4e-05     evaluation reward: 2.43
episode: 1049   score: 5.0   memory length: 201184   epsilon: 0.7996537000043493
steps: 325    lr: 4e-05     evaluation reward: 2.45
episode: 1050   score: 3.0   memory length: 201430   epsilon: 0.7991666200043599
steps: 246    lr: 4e-05     evaluation reward: 2.46
episode: 1051   score: 1.0   memory length: 201599   epsilon: 0.7988320000043672
steps: 169    lr: 4e-05     evaluation reward: 2.43
episode: 1052   score: 2.0   memory length: 201798   epsilon: 0.7984379800043757
steps: 199    lr: 4e-05     evaluation reward: 2.45
episode: 1053   score: 5.0   memory length: 202106   epsilon: 0.797828140004389
steps: 308    lr: 4e-05     evaluation reward: 2.48
episode: 1054   score: 7.0   memory length: 202492   epsilon: 0.7970638600044055
```

```
steps: 386    lr: 4e-05     evaluation reward: 2.52
episode: 1055   score: 5.0   memory length: 202806   epsilon: 0.796442140004419
steps: 314    lr: 4e-05     evaluation reward: 2.55
episode: 1056   score: 6.0   memory length: 203180   epsilon: 0.7957016200044351
steps: 374    lr: 4e-05     evaluation reward: 2.58
episode: 1057   score: 4.0   memory length: 203455   epsilon: 0.7951571200044469
steps: 275    lr: 4e-05     evaluation reward: 2.59
episode: 1058   score: 5.0   memory length: 203748   epsilon: 0.7945769800044595
steps: 293    lr: 4e-05     evaluation reward: 2.63
episode: 1059   score: 7.0   memory length: 204152   epsilon: 0.7937770600044769
steps: 404    lr: 4e-05     evaluation reward: 2.69
episode: 1060   score: 2.0   memory length: 204351   epsilon: 0.7933830400044855
steps: 199    lr: 4e-05     evaluation reward: 2.7
episode: 1061   score: 4.0   memory length: 204610   epsilon: 0.7928702200044966
steps: 259    lr: 4e-05     evaluation reward: 2.71
episode: 1062   score: 2.0   memory length: 204807   epsilon: 0.792480160004505
steps: 197    lr: 4e-05     evaluation reward: 2.69
episode: 1063   score: 2.0   memory length: 204988   epsilon: 0.7921217800045128
steps: 181    lr: 4e-05     evaluation reward: 2.69
episode: 1064   score: 2.0   memory length: 205208   epsilon: 0.7916861800045223
steps: 220    lr: 4e-05     evaluation reward: 2.7
episode: 1065   score: 4.0   memory length: 205485   epsilon: 0.7911377200045342
steps: 277    lr: 4e-05     evaluation reward: 2.71
episode: 1066   score: 5.0   memory length: 205846   epsilon: 0.7904229400045497
steps: 361    lr: 4e-05     evaluation reward: 2.74
episode: 1067   score: 4.0   memory length: 206122   epsilon: 0.7898764600045616
steps: 276    lr: 4e-05     evaluation reward: 2.73
episode: 1068   score: 3.0   memory length: 206370   epsilon: 0.7893854200045722
steps: 248    lr: 4e-05     evaluation reward: 2.76
episode: 1069   score: 3.0   memory length: 206601   epsilon: 0.7889280400045822
steps: 231    lr: 4e-05     evaluation reward: 2.78
episode: 1070   score: 8.0   memory length: 207074   epsilon: 0.7879915000046025
steps: 473    lr: 4e-05     evaluation reward: 2.83
episode: 1071   score: 5.0   memory length: 207368   epsilon: 0.7874093800046151
steps: 294    lr: 4e-05     evaluation reward: 2.87
episode: 1072   score: 6.0   memory length: 207720   epsilon: 0.7867124200046303
steps: 352    lr: 4e-05     evaluation reward: 2.91
episode: 1073   score: 6.0   memory length: 208070   epsilon: 0.7860194200046453
steps: 350    lr: 4e-05     evaluation reward: 2.94
episode: 1074   score: 5.0   memory length: 208395   epsilon: 0.7853759200046593
steps: 325    lr: 4e-05     evaluation reward: 2.99
episode: 1075   score: 9.0   memory length: 208896   epsilon: 0.7843839400046808
steps: 501    lr: 4e-05     evaluation reward: 3.06
episode: 1076   score: 4.0   memory length: 209174   epsilon: 0.7838335000046928
steps: 278    lr: 4e-05     evaluation reward: 3.09
episode: 1077   score: 3.0   memory length: 209420   epsilon: 0.7833464200047033
steps: 246    lr: 4e-05     evaluation reward: 3.01
episode: 1078   score: 2.0   memory length: 209618   epsilon: 0.7829543800047118
```

```
steps: 198    lr: 4e-05     evaluation reward: 3.0
episode: 1079   score: 2.0   memory length: 209818   epsilon: 0.7825583800047204
steps: 200    lr: 4e-05     evaluation reward: 2.99
episode: 1080   score: 3.0   memory length: 210064   epsilon: 0.782071300004731
steps: 246    lr: 4e-05     evaluation reward: 2.98
episode: 1081   score: 3.0   memory length: 210309   epsilon: 0.7815862000047415
steps: 245    lr: 4e-05     evaluation reward: 3.0
episode: 1082   score: 3.0   memory length: 210521   epsilon: 0.7811664400047507
steps: 212    lr: 4e-05     evaluation reward: 3.02
episode: 1083   score: 1.0   memory length: 210690   epsilon: 0.7808318200047579
steps: 169    lr: 4e-05     evaluation reward: 3.0
episode: 1084   score: 4.0   memory length: 211011   epsilon: 0.7801962400047717
steps: 321    lr: 4e-05     evaluation reward: 3.01
episode: 1085   score: 3.0   memory length: 211223   epsilon: 0.7797764800047808
steps: 212    lr: 4e-05     evaluation reward: 2.98
episode: 1086   score: 3.0   memory length: 211448   epsilon: 0.7793309800047905
steps: 225    lr: 4e-05     evaluation reward: 2.99
episode: 1087   score: 2.0   memory length: 211646   epsilon: 0.778938940004799
steps: 198    lr: 4e-05     evaluation reward: 3.0
episode: 1088   score: 5.0   memory length: 211970   epsilon: 0.778297420004813
steps: 324    lr: 4e-05     evaluation reward: 3.05
episode: 1089   score: 2.0   memory length: 212169   epsilon: 0.7779034000048215
steps: 199    lr: 4e-05     evaluation reward: 3.02
episode: 1090   score: 2.0   memory length: 212349   epsilon: 0.7775470000048292
steps: 180    lr: 4e-05     evaluation reward: 3.01
episode: 1091   score: 3.0   memory length: 212580   epsilon: 0.7770896200048392
steps: 231    lr: 4e-05     evaluation reward: 3.02
episode: 1092   score: 7.0   memory length: 213002   epsilon: 0.7762540600048573
steps: 422    lr: 4e-05     evaluation reward: 3.06
episode: 1093   score: 4.0   memory length: 213276   epsilon: 0.7757115400048691
steps: 274    lr: 4e-05     evaluation reward: 3.09
episode: 1094   score: 3.0   memory length: 213502   epsilon: 0.7752640600048788
steps: 226    lr: 4e-05     evaluation reward: 3.1
episode: 1095   score: 4.0   memory length: 213800   epsilon: 0.7746740200048916
steps: 298    lr: 4e-05     evaluation reward: 3.09
episode: 1096   score: 1.0   memory length: 213971   epsilon: 0.774335440004899
steps: 171    lr: 4e-05     evaluation reward: 3.09
episode: 1097   score: 1.0   memory length: 214142   epsilon: 0.7739968600049063
steps: 171    lr: 4e-05     evaluation reward: 3.07
episode: 1098   score: 2.0   memory length: 214339   epsilon: 0.7736068000049148
steps: 197    lr: 4e-05     evaluation reward: 3.09
episode: 1099   score: 3.0   memory length: 214565   epsilon: 0.7731593200049245
steps: 226    lr: 4e-05     evaluation reward: 3.08
episode: 1100   score: 3.0   memory length: 214791   epsilon: 0.7727118400049342
steps: 226    lr: 4e-05     evaluation reward: 3.09
episode: 1101   score: 2.0   memory length: 214989   epsilon: 0.7723198000049427
steps: 198    lr: 4e-05     evaluation reward: 3.09
episode: 1102   score: 3.0   memory length: 215216   epsilon: 0.7718703400049525
```

```
steps: 227    lr: 4e-05     evaluation reward: 3.11
episode: 1103   score: 2.0   memory length: 215414   epsilon: 0.771478300004961
steps: 198    lr: 4e-05     evaluation reward: 3.1
episode: 1104   score: 5.0   memory length: 215757   epsilon: 0.7707991600049757
steps: 343    lr: 4e-05     evaluation reward: 3.14
episode: 1105   score: 3.0   memory length: 215969   epsilon: 0.7703794000049848
steps: 212    lr: 4e-05     evaluation reward: 3.14
episode: 1106   score: 4.0   memory length: 216243   epsilon: 0.7698368800049966
steps: 274    lr: 4e-05     evaluation reward: 3.15
episode: 1107   score: 4.0   memory length: 216554   epsilon: 0.76922110000501
steps: 311    lr: 4e-05     evaluation reward: 3.17
episode: 1108   score: 4.0   memory length: 216870   epsilon: 0.7685954200050236
steps: 316    lr: 4e-05     evaluation reward: 3.19
episode: 1109   score: 4.0   memory length: 217166   epsilon: 0.7680093400050363
steps: 296    lr: 4e-05     evaluation reward: 3.19
episode: 1110   score: 0.0   memory length: 217289   epsilon: 0.7677658000050416
steps: 123    lr: 4e-05     evaluation reward: 3.14
episode: 1111   score: 0.0   memory length: 217412   epsilon: 0.7675222600050469
steps: 123    lr: 4e-05     evaluation reward: 3.11
episode: 1112   score: 5.0   memory length: 217736   epsilon: 0.7668807400050608
steps: 324    lr: 4e-05     evaluation reward: 3.16
episode: 1113   score: 5.0   memory length: 218050   epsilon: 0.7662590200050743
steps: 314    lr: 4e-05     evaluation reward: 3.2
episode: 1114   score: 6.0   memory length: 218381   epsilon: 0.7656036400050885
steps: 331    lr: 4e-05     evaluation reward: 3.21
episode: 1115   score: 1.0   memory length: 218532   epsilon: 0.765304660005095
steps: 151    lr: 4e-05     evaluation reward: 3.18
episode: 1116   score: 7.0   memory length: 218917   epsilon: 0.7645423600051116
steps: 385    lr: 4e-05     evaluation reward: 3.21
episode: 1117   score: 0.0   memory length: 219039   epsilon: 0.7643008000051168
steps: 122    lr: 4e-05     evaluation reward: 3.19
episode: 1118   score: 1.0   memory length: 219210   epsilon: 0.7639622200051241
steps: 171    lr: 4e-05     evaluation reward: 3.17
episode: 1119   score: 3.0   memory length: 219419   epsilon: 0.7635484000051331
steps: 209    lr: 4e-05     evaluation reward: 3.18
episode: 1120   score: 6.0   memory length: 219792   epsilon: 0.7628098600051492
steps: 373    lr: 4e-05     evaluation reward: 3.21
episode: 1121   score: 2.0   memory length: 220007   epsilon: 0.7623841600051584
steps: 215    lr: 4e-05     evaluation reward: 3.21
episode: 1122   score: 10.0   memory length: 220513   epsilon:
0.7613822800051802    steps: 506    lr: 4e-05     evaluation reward: 3.29
episode: 1123   score: 5.0   memory length: 220801   epsilon: 0.7608120400051925
steps: 288    lr: 4e-05     evaluation reward: 3.3
episode: 1124   score: 5.0   memory length: 221105   epsilon: 0.7602101200052056
steps: 304    lr: 4e-05     evaluation reward: 3.35
episode: 1125   score: 2.0   memory length: 221302   epsilon: 0.7598200600052141
steps: 197    lr: 4e-05     evaluation reward: 3.36
episode: 1126   score: 4.0   memory length: 221595   epsilon: 0.7592399200052267
```

```
steps: 293    lr: 4e-05     evaluation reward: 3.37
episode: 1127   score: 2.0   memory length: 221775   epsilon: 0.7588835200052344
steps: 180    lr: 4e-05     evaluation reward: 3.36
episode: 1128   score: 4.0   memory length: 222070   epsilon: 0.7582994200052471
steps: 295    lr: 4e-05     evaluation reward: 3.37
episode: 1129   score: 4.0   memory length: 222365   epsilon: 0.7577153200052598
steps: 295    lr: 4e-05     evaluation reward: 3.4
episode: 1130   score: 3.0   memory length: 222577   epsilon: 0.7572955600052689
steps: 212    lr: 4e-05     evaluation reward: 3.42
episode: 1131   score: 4.0   memory length: 222855   epsilon: 0.7567451200052808
steps: 278    lr: 4e-05     evaluation reward: 3.43
episode: 1132   score: 4.0   memory length: 223102   epsilon: 0.7562560600052914
steps: 247    lr: 4e-05     evaluation reward: 3.46
episode: 1133   score: 4.0   memory length: 223400   epsilon: 0.7556660200053043
steps: 298    lr: 4e-05     evaluation reward: 3.48
episode: 1134   score: 1.0   memory length: 223551   epsilon: 0.7553670400053107
steps: 151    lr: 4e-05     evaluation reward: 3.46
episode: 1135   score: 4.0   memory length: 223824   epsilon: 0.7548265000053225
steps: 273    lr: 4e-05     evaluation reward: 3.48
episode: 1136   score: 4.0   memory length: 224067   epsilon: 0.7543453600053329
steps: 243    lr: 4e-05     evaluation reward: 3.51
episode: 1137   score: 2.0   memory length: 224265   epsilon: 0.7539533200053414
steps: 198    lr: 4e-05     evaluation reward: 3.51
episode: 1138   score: 5.0   memory length: 224590   epsilon: 0.7533098200053554
steps: 325    lr: 4e-05     evaluation reward: 3.55
episode: 1139   score: 4.0   memory length: 224870   epsilon: 0.7527554200053674
steps: 280    lr: 4e-05     evaluation reward: 3.56
episode: 1140   score: 1.0   memory length: 225039   epsilon: 0.7524208000053747
steps: 169    lr: 4e-05     evaluation reward: 3.55
episode: 1141   score: 2.0   memory length: 225236   epsilon: 0.7520307400053832
steps: 197    lr: 4e-05     evaluation reward: 3.57
episode: 1142   score: 3.0   memory length: 225461   epsilon: 0.7515852400053928
steps: 225    lr: 4e-05     evaluation reward: 3.54
episode: 1143   score: 4.0   memory length: 225739   epsilon: 0.7510348000054048
steps: 278    lr: 4e-05     evaluation reward: 3.55
episode: 1144   score: 8.0   memory length: 226215   epsilon: 0.7500923200054253
steps: 476    lr: 4e-05     evaluation reward: 3.6
episode: 1145   score: 4.0   memory length: 226494   epsilon: 0.7495399000054372
steps: 279    lr: 4e-05     evaluation reward: 3.64
episode: 1146   score: 10.0   memory length: 227021   epsilon:
0.7484964400054599    steps: 527    lr: 4e-05     evaluation reward: 3.72
episode: 1147   score: 4.0   memory length: 227317   epsilon: 0.7479103600054726
steps: 296    lr: 4e-05     evaluation reward: 3.71
episode: 1148   score: 1.0   memory length: 227468   epsilon: 0.7476113800054791
steps: 151    lr: 4e-05     evaluation reward: 3.68
episode: 1149   score: 4.0   memory length: 227742   epsilon: 0.7470688600054909
steps: 274    lr: 4e-05     evaluation reward: 3.67
episode: 1150   score: 11.0   memory length: 228329   epsilon:
```

0.7459066000055161   steps: 587   lr: 4e-05   evaluation reward: 3.75
episode: 1151   score: 7.0   memory length: 228728   epsilon: 0.7451165800055333
steps: 399   lr: 4e-05   evaluation reward: 3.81
episode: 1152   score: 6.0   memory length: 229119   epsilon: 0.7443424000055501
steps: 391   lr: 4e-05   evaluation reward: 3.85
episode: 1153   score: 1.0   memory length: 229270   epsilon: 0.7440434200055566
steps: 151   lr: 4e-05   evaluation reward: 3.81
episode: 1154   score: 1.0   memory length: 229441   epsilon: 0.7437048400055639
steps: 171   lr: 4e-05   evaluation reward: 3.75
episode: 1155   score: 4.0   memory length: 229719   epsilon: 0.7431544000055759
steps: 278   lr: 4e-05   evaluation reward: 3.74
episode: 1156   score: 3.0   memory length: 229948   epsilon: 0.7427009800055857
steps: 229   lr: 4e-05   evaluation reward: 3.71
episode: 1157   score: 4.0   memory length: 230245   epsilon: 0.7421129200055985
steps: 297   lr: 4e-05   evaluation reward: 3.71
episode: 1158   score: 4.0   memory length: 230540   epsilon: 0.7415288200056112
steps: 295   lr: 4e-05   evaluation reward: 3.7
episode: 1159   score: 2.0   memory length: 230760   epsilon: 0.7410932200056206
steps: 220   lr: 4e-05   evaluation reward: 3.65
episode: 1160   score: 4.0   memory length: 231005   epsilon: 0.7406081200056311
steps: 245   lr: 4e-05   evaluation reward: 3.67
episode: 1161   score: 1.0   memory length: 231176   epsilon: 0.7402695400056385
steps: 171   lr: 4e-05   evaluation reward: 3.64
episode: 1162   score: 3.0   memory length: 231406   epsilon: 0.7398141400056484
steps: 230   lr: 4e-05   evaluation reward: 3.65
episode: 1163   score: 9.0   memory length: 231823   epsilon: 0.7389884800056663
steps: 417   lr: 4e-05   evaluation reward: 3.72
episode: 1164   score: 4.0   memory length: 232098   epsilon: 0.7384439800056781
steps: 275   lr: 4e-05   evaluation reward: 3.74
episode: 1165   score: 3.0   memory length: 232342   epsilon: 0.7379608600056886
steps: 244   lr: 4e-05   evaluation reward: 3.73
episode: 1166   score: 3.0   memory length: 232571   epsilon: 0.7375074400056985
steps: 229   lr: 4e-05   evaluation reward: 3.71
episode: 1167   score: 4.0   memory length: 232866   epsilon: 0.7369233400057111
steps: 295   lr: 4e-05   evaluation reward: 3.71
episode: 1168   score: 3.0   memory length: 233110   epsilon: 0.7364402200057216
steps: 244   lr: 4e-05   evaluation reward: 3.71
episode: 1169   score: 3.0   memory length: 233339   epsilon: 0.7359868000057315
steps: 229   lr: 4e-05   evaluation reward: 3.71
episode: 1170   score: 4.0   memory length: 233596   epsilon: 0.7354779400057425
steps: 257   lr: 4e-05   evaluation reward: 3.67
episode: 1171   score: 9.0   memory length: 234062   epsilon: 0.7345552600057625
steps: 466   lr: 4e-05   evaluation reward: 3.71
episode: 1172   score: 5.0   memory length: 234388   epsilon: 0.7339097800057766
steps: 326   lr: 4e-05   evaluation reward: 3.7
episode: 1173   score: 4.0   memory length: 234663   epsilon: 0.7333652800057884
steps: 275   lr: 4e-05   evaluation reward: 3.68
episode: 1174   score: 1.0   memory length: 234814   epsilon: 0.7330663000057949

```
steps: 151    lr: 4e-05     evaluation reward: 3.64
episode: 1175   score: 7.0   memory length: 235241   epsilon: 0.7322208400058132
steps: 427    lr: 4e-05     evaluation reward: 3.62
episode: 1176   score: 3.0   memory length: 235470   epsilon: 0.7317674200058231
steps: 229    lr: 4e-05     evaluation reward: 3.61
episode: 1177   score: 8.0   memory length: 235909   epsilon: 0.7308982000058419
steps: 439    lr: 4e-05     evaluation reward: 3.66
episode: 1178   score: 2.0   memory length: 236125   epsilon: 0.7304705200058512
steps: 216    lr: 4e-05     evaluation reward: 3.66
episode: 1179   score: 1.0   memory length: 236295   epsilon: 0.7301339200058585
steps: 170    lr: 4e-05     evaluation reward: 3.65
episode: 1180   score: 4.0   memory length: 236556   epsilon: 0.7296171400058697
steps: 261    lr: 4e-05     evaluation reward: 3.66
episode: 1181   score: 3.0   memory length: 236804   epsilon: 0.7291261000058804
steps: 248    lr: 4e-05     evaluation reward: 3.66
episode: 1182   score: 3.0   memory length: 237049   epsilon: 0.7286410000058909
steps: 245    lr: 4e-05     evaluation reward: 3.66
episode: 1183   score: 3.0   memory length: 237278   epsilon: 0.7281875800059008
steps: 229    lr: 4e-05     evaluation reward: 3.68
episode: 1184   score: 4.0   memory length: 237535   epsilon: 0.7276787200059118
steps: 257    lr: 4e-05     evaluation reward: 3.68
episode: 1185   score: 7.0   memory length: 237909   epsilon: 0.7269382000059279
steps: 374    lr: 4e-05     evaluation reward: 3.72
episode: 1186   score: 2.0   memory length: 238107   epsilon: 0.7265461600059364
steps: 198    lr: 4e-05     evaluation reward: 3.71
episode: 1187   score: 9.0   memory length: 238632   epsilon: 0.725506660005959
steps: 525    lr: 4e-05     evaluation reward: 3.78
episode: 1188   score: 3.0   memory length: 238878   epsilon: 0.7250195800059696
steps: 246    lr: 4e-05     evaluation reward: 3.76
episode: 1189   score: 1.0   memory length: 239047   epsilon: 0.7246849600059768
steps: 169    lr: 4e-05     evaluation reward: 3.75
episode: 1190   score: 4.0   memory length: 239325   epsilon: 0.7241345200059888
steps: 278    lr: 4e-05     evaluation reward: 3.77
episode: 1191   score: 3.0   memory length: 239536   epsilon: 0.7237167400059978
steps: 211    lr: 4e-05     evaluation reward: 3.77
episode: 1192   score: 4.0   memory length: 239814   epsilon: 0.7231663000060098
steps: 278    lr: 4e-05     evaluation reward: 3.74
episode: 1193   score: 3.0   memory length: 240043   epsilon: 0.7227128800060196
steps: 229    lr: 4e-05     evaluation reward: 3.73
episode: 1194   score: 2.0   memory length: 240243   epsilon: 0.7223168800060282
steps: 200    lr: 4e-05     evaluation reward: 3.72
episode: 1195   score: 5.0   memory length: 240534   epsilon: 0.7217407000060407
steps: 291    lr: 4e-05     evaluation reward: 3.73
episode: 1196   score: 4.0   memory length: 240827   epsilon: 0.7211605600060533
steps: 293    lr: 4e-05     evaluation reward: 3.76
episode: 1197   score: 2.0   memory length: 241025   epsilon: 0.7207685200060618
steps: 198    lr: 4e-05     evaluation reward: 3.77
episode: 1198   score: 3.0   memory length: 241256   epsilon: 0.7203111400060718
```

```
steps: 231    lr: 4e-05     evaluation reward: 3.78
episode: 1199   score: 1.0   memory length: 241407   epsilon: 0.7200121600060783
steps: 151    lr: 4e-05     evaluation reward: 3.76
episode: 1200   score: 5.0   memory length: 241712   epsilon: 0.7194082600060914
steps: 305    lr: 4e-05     evaluation reward: 3.78
episode: 1201   score: 3.0   memory length: 241938   epsilon: 0.7189607800061011
steps: 226    lr: 4e-05     evaluation reward: 3.79
episode: 1202   score: 2.0   memory length: 242136   epsilon: 0.7185687400061096
steps: 198    lr: 4e-05     evaluation reward: 3.78
episode: 1203   score: 6.0   memory length: 242481   epsilon: 0.7178856400061244
steps: 345    lr: 4e-05     evaluation reward: 3.82
episode: 1204   score: 3.0   memory length: 242709   epsilon: 0.7174342000061342
steps: 228    lr: 4e-05     evaluation reward: 3.8
episode: 1205   score: 5.0   memory length: 243015   epsilon: 0.7168283200061474
steps: 306    lr: 4e-05     evaluation reward: 3.82
episode: 1206   score: 8.0   memory length: 243456   epsilon: 0.7159551400061663
steps: 441    lr: 4e-05     evaluation reward: 3.86
episode: 1207   score: 5.0   memory length: 243800   epsilon: 0.7152740200061811
steps: 344    lr: 4e-05     evaluation reward: 3.87
episode: 1208   score: 1.0   memory length: 243972   epsilon: 0.7149334600061885
steps: 172    lr: 4e-05     evaluation reward: 3.84
episode: 1209   score: 0.0   memory length: 244095   epsilon: 0.7146899200061938
steps: 123    lr: 4e-05     evaluation reward: 3.8
episode: 1210   score: 2.0   memory length: 244276   epsilon: 0.7143315400062016
steps: 181    lr: 4e-05     evaluation reward: 3.82
episode: 1211   score: 1.0   memory length: 244446   epsilon: 0.7139949400062089
steps: 170    lr: 4e-05     evaluation reward: 3.83
episode: 1212   score: 2.0   memory length: 244645   epsilon: 0.7136009200062174
steps: 199    lr: 4e-05     evaluation reward: 3.8
episode: 1213   score: 4.0   memory length: 244938   epsilon: 0.71302078000623
steps: 293    lr: 4e-05     evaluation reward: 3.79
episode: 1214   score: 6.0   memory length: 245286   epsilon: 0.712331740006245
steps: 348    lr: 4e-05     evaluation reward: 3.79
episode: 1215   score: 2.0   memory length: 245484   epsilon: 0.7119397000062535
steps: 198    lr: 4e-05     evaluation reward: 3.8
episode: 1216   score: 2.0   memory length: 245666   epsilon: 0.7115793400062613
steps: 182    lr: 4e-05     evaluation reward: 3.75
episode: 1217   score: 3.0   memory length: 245897   epsilon: 0.7111219600062713
steps: 231    lr: 4e-05     evaluation reward: 3.78
episode: 1218   score: 5.0   memory length: 246217   epsilon: 0.710488360006285
steps: 320    lr: 4e-05     evaluation reward: 3.82
episode: 1219   score: 6.0   memory length: 246607   epsilon: 0.7097161600063018
steps: 390    lr: 4e-05     evaluation reward: 3.85
episode: 1220   score: 3.0   memory length: 246852   epsilon: 0.7092310600063123
steps: 245    lr: 4e-05     evaluation reward: 3.82
episode: 1221   score: 2.0   memory length: 247052   epsilon: 0.7088350600063209
steps: 200    lr: 4e-05     evaluation reward: 3.82
episode: 1222   score: 6.0   memory length: 247398   epsilon: 0.7081499800063358
```

```
steps: 346    lr: 4e-05     evaluation reward: 3.78
episode: 1223   score: 4.0   memory length: 247676   epsilon: 0.7075995400063477
steps: 278    lr: 4e-05     evaluation reward: 3.77
episode: 1224   score: 3.0   memory length: 247922   epsilon: 0.7071124600063583
steps: 246    lr: 4e-05     evaluation reward: 3.75
episode: 1225   score: 4.0   memory length: 248220   epsilon: 0.7065224200063711
steps: 298    lr: 4e-05     evaluation reward: 3.77
episode: 1226   score: 5.0   memory length: 248525   epsilon: 0.7059185200063842
steps: 305    lr: 4e-05     evaluation reward: 3.78
episode: 1227   score: 7.0   memory length: 248950   epsilon: 0.7050770200064025
steps: 425    lr: 4e-05     evaluation reward: 3.83
episode: 1228   score: 4.0   memory length: 249227   epsilon: 0.7045285600064144
steps: 277    lr: 4e-05     evaluation reward: 3.83
episode: 1229   score: 5.0   memory length: 249555   epsilon: 0.7038791200064285
steps: 328    lr: 4e-05     evaluation reward: 3.84
episode: 1230   score: 3.0   memory length: 249801   epsilon: 0.7033920400064391
steps: 246    lr: 4e-05     evaluation reward: 3.84
episode: 1231   score: 5.0   memory length: 250104   epsilon: 0.7027921000064521
steps: 303    lr: 4e-05     evaluation reward: 3.85
episode: 1232   score: 6.0   memory length: 250422   epsilon: 0.7021624600064658
steps: 318    lr: 4e-05     evaluation reward: 3.87
episode: 1233   score: 4.0   memory length: 250719   epsilon: 0.7015744000064785
steps: 297    lr: 4e-05     evaluation reward: 3.87
episode: 1234   score: 4.0   memory length: 251000   epsilon: 0.7010180200064906
steps: 281    lr: 4e-05     evaluation reward: 3.9
episode: 1235   score: 5.0   memory length: 251327   epsilon: 0.7003705600065047
steps: 327    lr: 4e-05     evaluation reward: 3.91
episode: 1236   score: 3.0   memory length: 251558   epsilon: 0.6999131800065146
steps: 231    lr: 4e-05     evaluation reward: 3.9
episode: 1237   score: 4.0   memory length: 251836   epsilon: 0.6993627400065265
steps: 278    lr: 4e-05     evaluation reward: 3.92
episode: 1238   score: 5.0   memory length: 252181   epsilon: 0.6986796400065414
steps: 345    lr: 4e-05     evaluation reward: 3.92
episode: 1239   score: 3.0   memory length: 252430   epsilon: 0.6981866200065521
steps: 249    lr: 4e-05     evaluation reward: 3.91
episode: 1240   score: 3.0   memory length: 252641   epsilon: 0.6977688400065611
steps: 211    lr: 4e-05     evaluation reward: 3.93
episode: 1241   score: 3.0   memory length: 252907   epsilon: 0.6972421600065726
steps: 266    lr: 4e-05     evaluation reward: 3.94
episode: 1242   score: 6.0   memory length: 253280   epsilon: 0.6965036200065886
steps: 373    lr: 4e-05     evaluation reward: 3.97
episode: 1243   score: 3.0   memory length: 253510   epsilon: 0.6960482200065985
steps: 230    lr: 4e-05     evaluation reward: 3.96
episode: 1244   score: 9.0   memory length: 254030   epsilon: 0.6950186200066208
steps: 520    lr: 4e-05     evaluation reward: 3.97
episode: 1245   score: 5.0   memory length: 254340   epsilon: 0.6944048200066342
steps: 310    lr: 4e-05     evaluation reward: 3.98
episode: 1246   score: 2.0   memory length: 254558   epsilon: 0.6939731800066435
```

```
steps: 218    lr: 4e-05     evaluation reward: 3.9
episode: 1247   score: 4.0   memory length: 254837   epsilon: 0.6934207600066555
steps: 279    lr: 4e-05     evaluation reward: 3.9
episode: 1248   score: 3.0   memory length: 255083   epsilon: 0.6929336800066661
steps: 246    lr: 4e-05     evaluation reward: 3.92
episode: 1249   score: 5.0   memory length: 255400   epsilon: 0.6923060200066797
steps: 317    lr: 4e-05     evaluation reward: 3.93
episode: 1250   score: 4.0   memory length: 255674   epsilon: 0.6917635000066915
steps: 274    lr: 4e-05     evaluation reward: 3.86
episode: 1251   score: 2.0   memory length: 255872   epsilon: 0.6913714600067
steps: 198    lr: 4e-05     evaluation reward: 3.81
episode: 1252   score: 7.0   memory length: 256299   epsilon: 0.6905260000067184
steps: 427    lr: 4e-05     evaluation reward: 3.82
episode: 1253   score: 2.0   memory length: 256497   epsilon: 0.6901339600067269
steps: 198    lr: 4e-05     evaluation reward: 3.83
episode: 1254   score: 5.0   memory length: 256806   epsilon: 0.6895221400067402
steps: 309    lr: 4e-05     evaluation reward: 3.87
episode: 1255   score: 2.0   memory length: 257021   epsilon: 0.6890964400067494
steps: 215    lr: 4e-05     evaluation reward: 3.85
episode: 1256   score: 4.0   memory length: 257280   epsilon: 0.6885836200067605
steps: 259    lr: 4e-05     evaluation reward: 3.86
episode: 1257   score: 1.0   memory length: 257450   epsilon: 0.6882470200067679
steps: 170    lr: 4e-05     evaluation reward: 3.83
episode: 1258   score: 4.0   memory length: 257709   epsilon: 0.687734200006779
steps: 259    lr: 4e-05     evaluation reward: 3.83
episode: 1259   score: 2.0   memory length: 257889   epsilon: 0.6873778000067867
steps: 180    lr: 4e-05     evaluation reward: 3.83
episode: 1260   score: 1.0   memory length: 258040   epsilon: 0.6870788200067932
steps: 151    lr: 4e-05     evaluation reward: 3.8
episode: 1261   score: 4.0   memory length: 258315   epsilon: 0.686534320006805
steps: 275    lr: 4e-05     evaluation reward: 3.83
episode: 1262   score: 3.0   memory length: 258562   epsilon: 0.6860452600068156
steps: 247    lr: 4e-05     evaluation reward: 3.83
episode: 1263   score: 1.0   memory length: 258713   epsilon: 0.6857462800068221
steps: 151    lr: 4e-05     evaluation reward: 3.75
episode: 1264   score: 3.0   memory length: 258961   epsilon: 0.6852552400068328
steps: 248    lr: 4e-05     evaluation reward: 3.74
episode: 1265   score: 4.0   memory length: 259240   epsilon: 0.6847028200068448
steps: 279    lr: 4e-05     evaluation reward: 3.75
episode: 1266   score: 4.0   memory length: 259482   epsilon: 0.6842236600068552
steps: 242    lr: 4e-05     evaluation reward: 3.76
episode: 1267   score: 2.0   memory length: 259683   epsilon: 0.6838256800068638
steps: 201    lr: 4e-05     evaluation reward: 3.74
episode: 1268   score: 4.0   memory length: 259940   epsilon: 0.6833168200068749
steps: 257    lr: 4e-05     evaluation reward: 3.75
episode: 1269   score: 3.0   memory length: 260166   epsilon: 0.6828693400068846
steps: 226    lr: 4e-05     evaluation reward: 3.75
episode: 1270   score: 3.0   memory length: 260379   epsilon: 0.6824476000068938
```

```
steps: 213    lr: 4e-05     evaluation reward: 3.74
episode: 1271   score: 2.0   memory length: 260579   epsilon: 0.6820516000069023
steps: 200    lr: 4e-05     evaluation reward: 3.67
episode: 1272   score: 4.0   memory length: 260854   epsilon: 0.6815071000069142
steps: 275    lr: 4e-05     evaluation reward: 3.66
episode: 1273   score: 4.0   memory length: 261149   epsilon: 0.6809230000069268
steps: 295    lr: 4e-05     evaluation reward: 3.66
episode: 1274   score: 4.0   memory length: 261417   epsilon: 0.6803923600069384
steps: 268    lr: 4e-05     evaluation reward: 3.69
episode: 1275   score: 4.0   memory length: 261693   epsilon: 0.6798458800069502
steps: 276    lr: 4e-05     evaluation reward: 3.66
episode: 1276   score: 8.0   memory length: 262007   epsilon: 0.6792241600069637
steps: 314    lr: 4e-05     evaluation reward: 3.71
episode: 1277   score: 1.0   memory length: 262158   epsilon: 0.6789251800069702
steps: 151    lr: 4e-05     evaluation reward: 3.64
episode: 1278   score: 3.0   memory length: 262368   epsilon: 0.6785093800069792
steps: 210    lr: 4e-05     evaluation reward: 3.65
episode: 1279   score: 1.0   memory length: 262519   epsilon: 0.6782104000069857
steps: 151    lr: 4e-05     evaluation reward: 3.65
episode: 1280   score: 2.0   memory length: 262718   epsilon: 0.6778163800069943
steps: 199    lr: 4e-05     evaluation reward: 3.63
episode: 1281   score: 2.0   memory length: 262918   epsilon: 0.6774203800070029
steps: 200    lr: 4e-05     evaluation reward: 3.62
episode: 1282   score: 7.0   memory length: 263320   epsilon: 0.6766244200070202
steps: 402    lr: 4e-05     evaluation reward: 3.66
episode: 1283   score: 1.0   memory length: 263470   epsilon: 0.6763274200070266
steps: 150    lr: 4e-05     evaluation reward: 3.64
episode: 1284   score: 2.0   memory length: 263668   epsilon: 0.6759353800070351
steps: 198    lr: 4e-05     evaluation reward: 3.62
episode: 1285   score: 6.0   memory length: 264062   epsilon: 0.6751552600070521
steps: 394    lr: 4e-05     evaluation reward: 3.61
episode: 1286   score: 0.0   memory length: 264185   epsilon: 0.6749117200070573
steps: 123    lr: 4e-05     evaluation reward: 3.59
episode: 1287   score: 6.0   memory length: 264548   epsilon: 0.674192980007073
steps: 363    lr: 4e-05     evaluation reward: 3.56
episode: 1288   score: 4.0   memory length: 264823   epsilon: 0.6736484800070848
steps: 275    lr: 4e-05     evaluation reward: 3.57
episode: 1289   score: 4.0   memory length: 265119   epsilon: 0.6730624000070975
steps: 296    lr: 4e-05     evaluation reward: 3.6
episode: 1290   score: 7.0   memory length: 265522   epsilon: 0.6722644600071148
steps: 403    lr: 4e-05     evaluation reward: 3.63
episode: 1291   score: 4.0   memory length: 265820   epsilon: 0.6716744200071276
steps: 298    lr: 4e-05     evaluation reward: 3.64
episode: 1292   score: 6.0   memory length: 266183   epsilon: 0.6709556800071432
steps: 363    lr: 4e-05     evaluation reward: 3.66
episode: 1293   score: 4.0   memory length: 266441   epsilon: 0.6704448400071543
steps: 258    lr: 4e-05     evaluation reward: 3.67
episode: 1294   score: 3.0   memory length: 266686   epsilon: 0.6699597400071649
```

```
steps: 245    lr: 4e-05     evaluation reward: 3.68
episode: 1295   score: 6.0   memory length: 267042   epsilon: 0.6692548600071802
steps: 356    lr: 4e-05     evaluation reward: 3.69
episode: 1296   score: 6.0   memory length: 267405   epsilon: 0.6685361200071958
steps: 363    lr: 4e-05     evaluation reward: 3.71
episode: 1297   score: 7.0   memory length: 267831   epsilon: 0.6676926400072141
steps: 426    lr: 4e-05     evaluation reward: 3.76
episode: 1298   score: 5.0   memory length: 268120   epsilon: 0.6671204200072265
steps: 289    lr: 4e-05     evaluation reward: 3.78
episode: 1299   score: 2.0   memory length: 268338   epsilon: 0.6666887800072359
steps: 218    lr: 4e-05     evaluation reward: 3.79
episode: 1300   score: 4.0   memory length: 268597   epsilon: 0.666175960007247
steps: 259    lr: 4e-05     evaluation reward: 3.78
episode: 1301   score: 4.0   memory length: 268872   epsilon: 0.6656314600072588
steps: 275    lr: 4e-05     evaluation reward: 3.79
episode: 1302   score: 6.0   memory length: 269242   epsilon: 0.6648988600072747
steps: 370    lr: 4e-05     evaluation reward: 3.83
episode: 1303   score: 2.0   memory length: 269440   epsilon: 0.6645068200072832
steps: 198    lr: 4e-05     evaluation reward: 3.79
episode: 1304   score: 5.0   memory length: 269744   epsilon: 0.6639049000072963
steps: 304    lr: 4e-05     evaluation reward: 3.81
episode: 1305   score: 2.0   memory length: 269962   epsilon: 0.6634732600073057
steps: 218    lr: 4e-05     evaluation reward: 3.78
episode: 1306   score: 7.0   memory length: 270364   epsilon: 0.662677300007323
steps: 402    lr: 4e-05     evaluation reward: 3.77
episode: 1307   score: 4.0   memory length: 270643   epsilon: 0.6621248800073349
steps: 279    lr: 4e-05     evaluation reward: 3.76
episode: 1308   score: 3.0   memory length: 270874   epsilon: 0.6616675000073449
steps: 231    lr: 4e-05     evaluation reward: 3.78
episode: 1309   score: 4.0   memory length: 271133   epsilon: 0.661154680007356
steps: 259    lr: 4e-05     evaluation reward: 3.82
episode: 1310   score: 1.0   memory length: 271284   epsilon: 0.6608557000073625
steps: 151    lr: 4e-05     evaluation reward: 3.81
episode: 1311   score: 4.0   memory length: 271604   epsilon: 0.6602221000073762
steps: 320    lr: 4e-05     evaluation reward: 3.84
episode: 1312   score: 7.0   memory length: 271973   epsilon: 0.6594914800073921
steps: 369    lr: 4e-05     evaluation reward: 3.89
episode: 1313   score: 6.0   memory length: 272329   epsilon: 0.6587866000074074
steps: 356    lr: 4e-05     evaluation reward: 3.91
episode: 1314   score: 3.0   memory length: 272575   epsilon: 0.658299520007418
steps: 246    lr: 4e-05     evaluation reward: 3.88
episode: 1315   score: 4.0   memory length: 272815   epsilon: 0.6578243200074283
steps: 240    lr: 4e-05     evaluation reward: 3.9
episode: 1316   score: 3.0   memory length: 273081   epsilon: 0.6572976400074397
steps: 266    lr: 4e-05     evaluation reward: 3.91
episode: 1317   score: 5.0   memory length: 273385   epsilon: 0.6566957200074528
steps: 304    lr: 4e-05     evaluation reward: 3.93
episode: 1318   score: 3.0   memory length: 273636   epsilon: 0.6561987400074636
```

```
steps: 251    lr: 4e-05     evaluation reward: 3.91
episode: 1319   score: 2.0   memory length: 273854   epsilon: 0.655767100007473
steps: 218    lr: 4e-05     evaluation reward: 3.87
episode: 1320   score: 2.0   memory length: 274073   epsilon: 0.6553334800074824
steps: 219    lr: 4e-05     evaluation reward: 3.86
episode: 1321   score: 3.0   memory length: 274304   epsilon: 0.6548761000074923
steps: 231    lr: 4e-05     evaluation reward: 3.87
episode: 1322   score: 4.0   memory length: 274560   epsilon: 0.6543692200075033
steps: 256    lr: 4e-05     evaluation reward: 3.85
episode: 1323   score: 5.0   memory length: 274887   epsilon: 0.6537217600075174
steps: 327    lr: 4e-05     evaluation reward: 3.86
episode: 1324   score: 1.0   memory length: 275038   epsilon: 0.6534227800075239
steps: 151    lr: 4e-05     evaluation reward: 3.84
episode: 1325   score: 8.0   memory length: 275333   epsilon: 0.6528386800075365
steps: 295    lr: 4e-05     evaluation reward: 3.88
episode: 1326   score: 6.0   memory length: 275674   epsilon: 0.6521635000075512
steps: 341    lr: 4e-05     evaluation reward: 3.89
episode: 1327   score: 4.0   memory length: 275947   epsilon: 0.6516229600075629
steps: 273    lr: 4e-05     evaluation reward: 3.86
episode: 1328   score: 2.0   memory length: 276145   epsilon: 0.6512309200075714
steps: 198    lr: 4e-05     evaluation reward: 3.84
episode: 1329   score: 2.0   memory length: 276364   epsilon: 0.6507973000075808
steps: 219    lr: 4e-05     evaluation reward: 3.81
episode: 1330   score: 3.0   memory length: 276575   epsilon: 0.6503795200075899
steps: 211    lr: 4e-05     evaluation reward: 3.81
episode: 1331   score: 8.0   memory length: 277045   epsilon: 0.6494489200076101
steps: 470    lr: 4e-05     evaluation reward: 3.84
episode: 1332   score: 5.0   memory length: 277367   epsilon: 0.648811360007624
steps: 322    lr: 4e-05     evaluation reward: 3.83
episode: 1333   score: 6.0   memory length: 277722   epsilon: 0.6481084600076392
steps: 355    lr: 4e-05     evaluation reward: 3.85
episode: 1334   score: 3.0   memory length: 277949   epsilon: 0.647659000007649
steps: 227    lr: 4e-05     evaluation reward: 3.84
episode: 1335   score: 6.0   memory length: 278305   epsilon: 0.6469541200076643
steps: 356    lr: 4e-05     evaluation reward: 3.85
episode: 1336   score: 5.0   memory length: 278632   epsilon: 0.6463066600076783
steps: 327    lr: 4e-05     evaluation reward: 3.87
episode: 1337   score: 4.0   memory length: 278931   epsilon: 0.6457146400076912
steps: 299    lr: 4e-05     evaluation reward: 3.87
episode: 1338   score: 5.0   memory length: 279239   epsilon: 0.6451048000077044
steps: 308    lr: 4e-05     evaluation reward: 3.87
episode: 1339   score: 1.0   memory length: 279390   epsilon: 0.6448058200077109
steps: 151    lr: 4e-05     evaluation reward: 3.85
episode: 1340   score: 5.0   memory length: 279728   epsilon: 0.6441365800077254
steps: 338    lr: 4e-05     evaluation reward: 3.87
episode: 1341   score: 1.0   memory length: 279878   epsilon: 0.6438395800077319
steps: 150    lr: 4e-05     evaluation reward: 3.85
episode: 1342   score: 3.0   memory length: 280106   epsilon: 0.6433881400077417
```

```
steps: 228    lr: 4e-05     evaluation reward: 3.82
episode: 1343   score: 8.0   memory length: 280559   epsilon: 0.6424912000077612
steps: 453    lr: 4e-05     evaluation reward: 3.87
episode: 1344   score: 8.0   memory length: 280990   epsilon: 0.6416378200077797
steps: 431    lr: 4e-05     evaluation reward: 3.86
episode: 1345   score: 3.0   memory length: 281216   epsilon: 0.6411903400077894
steps: 226    lr: 4e-05     evaluation reward: 3.84
episode: 1346   score: 4.0   memory length: 281476   epsilon: 0.6406755400078006
steps: 260    lr: 4e-05     evaluation reward: 3.86
episode: 1347   score: 2.0   memory length: 281673   epsilon: 0.640285480007809
steps: 197    lr: 4e-05     evaluation reward: 3.84
episode: 1348   score: 4.0   memory length: 281968   epsilon: 0.6397013800078217
steps: 295    lr: 4e-05     evaluation reward: 3.85
episode: 1349   score: 4.0   memory length: 282245   epsilon: 0.6391529200078336
steps: 277    lr: 4e-05     evaluation reward: 3.84
episode: 1350   score: 5.0   memory length: 282590   epsilon: 0.6384698200078485
steps: 345    lr: 4e-05     evaluation reward: 3.85
episode: 1351   score: 5.0   memory length: 282934   epsilon: 0.6377887000078633
steps: 344    lr: 4e-05     evaluation reward: 3.88
episode: 1352   score: 5.0   memory length: 283208   epsilon: 0.637246180007875
steps: 274    lr: 4e-05     evaluation reward: 3.86
episode: 1353   score: 3.0   memory length: 283456   epsilon: 0.6367551400078857
steps: 248    lr: 4e-05     evaluation reward: 3.87
episode: 1354   score: 2.0   memory length: 283654   epsilon: 0.6363631000078942
steps: 198    lr: 4e-05     evaluation reward: 3.84
episode: 1355   score: 2.0   memory length: 283873   epsilon: 0.6359294800079036
steps: 219    lr: 4e-05     evaluation reward: 3.84
episode: 1356   score: 3.0   memory length: 284117   epsilon: 0.6354463600079141
steps: 244    lr: 4e-05     evaluation reward: 3.83
episode: 1357   score: 4.0   memory length: 284374   epsilon: 0.6349375000079251
steps: 257    lr: 4e-05     evaluation reward: 3.86
episode: 1358   score: 4.0   memory length: 284668   epsilon: 0.6343553800079378
steps: 294    lr: 4e-05     evaluation reward: 3.86
episode: 1359   score: 9.0   memory length: 285114   epsilon: 0.633472300007957
steps: 446    lr: 4e-05     evaluation reward: 3.93
episode: 1360   score: 6.0   memory length: 285490   epsilon: 0.6327278200079731
steps: 376    lr: 4e-05     evaluation reward: 3.98
episode: 1361   score: 2.0   memory length: 285687   epsilon: 0.6323377600079816
steps: 197    lr: 4e-05     evaluation reward: 3.96
episode: 1362   score: 4.0   memory length: 285980   epsilon: 0.6317576200079942
steps: 293    lr: 4e-05     evaluation reward: 3.97
episode: 1363   score: 5.0   memory length: 286269   epsilon: 0.6311854000080066
steps: 289    lr: 4e-05     evaluation reward: 4.01
episode: 1364   score: 7.0   memory length: 286659   epsilon: 0.6304132000080234
steps: 390    lr: 4e-05     evaluation reward: 4.05
episode: 1365   score: 0.0   memory length: 286782   epsilon: 0.6301696600080287
steps: 123    lr: 4e-05     evaluation reward: 4.01
episode: 1366   score: 1.0   memory length: 286933   epsilon: 0.6298706800080351
```

```
steps: 151    lr: 4e-05     evaluation reward: 3.98
episode: 1367   score: 3.0   memory length: 287163   epsilon: 0.629415280008045
steps: 230    lr: 4e-05     evaluation reward: 3.99
episode: 1368   score: 3.0   memory length: 287375   epsilon: 0.6289955200080541
steps: 212    lr: 4e-05     evaluation reward: 3.98
episode: 1369   score: 4.0   memory length: 287641   epsilon: 0.6284688400080656
steps: 266    lr: 4e-05     evaluation reward: 3.99
episode: 1370   score: 11.0   memory length: 288149   epsilon:
0.6274630000080874    steps: 508    lr: 4e-05     evaluation reward: 4.07
episode: 1371   score: 2.0   memory length: 288331   epsilon: 0.6271026400080952
steps: 182    lr: 4e-05     evaluation reward: 4.07
episode: 1372   score: 7.0   memory length: 288774   epsilon: 0.6262255000081143
steps: 443    lr: 4e-05     evaluation reward: 4.1
episode: 1373   score: 4.0   memory length: 289055   epsilon: 0.6256691200081264
steps: 281    lr: 4e-05     evaluation reward: 4.1
episode: 1374   score: 4.0   memory length: 289331   epsilon: 0.6251226400081382
steps: 276    lr: 4e-05     evaluation reward: 4.1
episode: 1375   score: 6.0   memory length: 289667   epsilon: 0.6244573600081527
steps: 336    lr: 4e-05     evaluation reward: 4.12
episode: 1376   score: 6.0   memory length: 290044   epsilon: 0.6237109000081689
steps: 377    lr: 4e-05     evaluation reward: 4.1
episode: 1377   score: 6.0   memory length: 290435   epsilon: 0.6229367200081857
steps: 391    lr: 4e-05     evaluation reward: 4.15
episode: 1378   score: 2.0   memory length: 290633   epsilon: 0.6225446800081942
steps: 198    lr: 4e-05     evaluation reward: 4.14
episode: 1379   score: 4.0   memory length: 290891   epsilon: 0.6220338400082053
steps: 258    lr: 4e-05     evaluation reward: 4.17
episode: 1380   score: 4.0   memory length: 291134   epsilon: 0.6215527000082157
steps: 243    lr: 4e-05     evaluation reward: 4.19
episode: 1381   score: 4.0   memory length: 291435   epsilon: 0.6209567200082287
steps: 301    lr: 4e-05     evaluation reward: 4.21
episode: 1382   score: 4.0   memory length: 291711   epsilon: 0.6204102400082405
steps: 276    lr: 4e-05     evaluation reward: 4.18
episode: 1383   score: 6.0   memory length: 292062   epsilon: 0.6197152600082556
steps: 351    lr: 4e-05     evaluation reward: 4.23
episode: 1384   score: 2.0   memory length: 292261   epsilon: 0.6193212400082642
steps: 199    lr: 4e-05     evaluation reward: 4.23
episode: 1385   score: 1.0   memory length: 292412   epsilon: 0.6190222600082707
steps: 151    lr: 4e-05     evaluation reward: 4.18
episode: 1386   score: 3.0   memory length: 292659   epsilon: 0.6185332000082813
steps: 247    lr: 4e-05     evaluation reward: 4.21
episode: 1387   score: 4.0   memory length: 292936   epsilon: 0.6179847400082932
steps: 277    lr: 4e-05     evaluation reward: 4.19
episode: 1388   score: 8.0   memory length: 293374   epsilon: 0.617117500008312
steps: 438    lr: 4e-05     evaluation reward: 4.23
episode: 1389   score: 4.0   memory length: 293654   epsilon: 0.616563100008324
steps: 280    lr: 4e-05     evaluation reward: 4.23
episode: 1390   score: 5.0   memory length: 293943   epsilon: 0.6159908800083365
```

```
steps: 289    lr: 4e-05     evaluation reward: 4.21
episode: 1391   score: 7.0   memory length: 294351   epsilon: 0.615183040008354
steps: 408    lr: 4e-05     evaluation reward: 4.24
episode: 1392   score: 4.0   memory length: 294647   epsilon: 0.6145969600083667
steps: 296    lr: 4e-05     evaluation reward: 4.22
episode: 1393   score: 2.0   memory length: 294847   epsilon: 0.6142009600083753
steps: 200    lr: 4e-05     evaluation reward: 4.2
episode: 1394   score: 4.0   memory length: 295102   epsilon: 0.6136960600083863
steps: 255    lr: 4e-05     evaluation reward: 4.21
episode: 1395   score: 4.0   memory length: 295379   epsilon: 0.6131476000083982
steps: 277    lr: 4e-05     evaluation reward: 4.19
episode: 1396   score: 3.0   memory length: 295590   epsilon: 0.6127298200084073
steps: 211    lr: 4e-05     evaluation reward: 4.16
episode: 1397   score: 5.0   memory length: 295862   epsilon: 0.612191260008419
steps: 272    lr: 4e-05     evaluation reward: 4.14
episode: 1398   score: 2.0   memory length: 296059   epsilon: 0.6118012000084274
steps: 197    lr: 4e-05     evaluation reward: 4.11
episode: 1399   score: 4.0   memory length: 296354   epsilon: 0.6112171000084401
steps: 295    lr: 4e-05     evaluation reward: 4.13
episode: 1400   score: 4.0   memory length: 296652   epsilon: 0.6106270600084529
steps: 298    lr: 4e-05     evaluation reward: 4.13
episode: 1401   score: 3.0   memory length: 296862   epsilon: 0.6102112600084619
steps: 210    lr: 4e-05     evaluation reward: 4.12
episode: 1402   score: 6.0   memory length: 297225   epsilon: 0.6094925200084775
steps: 363    lr: 4e-05     evaluation reward: 4.12
episode: 1403   score: 5.0   memory length: 297552   epsilon: 0.6088450600084916
steps: 327    lr: 4e-05     evaluation reward: 4.15
episode: 1404   score: 6.0   memory length: 297911   epsilon: 0.608134240008507
steps: 359    lr: 4e-05     evaluation reward: 4.16
episode: 1405   score: 4.0   memory length: 298190   epsilon: 0.607581820008519
steps: 279    lr: 4e-05     evaluation reward: 4.18
episode: 1406   score: 3.0   memory length: 298437   epsilon: 0.6070927600085296
steps: 247    lr: 4e-05     evaluation reward: 4.14
episode: 1407   score: 7.0   memory length: 298812   epsilon: 0.6063502600085457
steps: 375    lr: 4e-05     evaluation reward: 4.17
episode: 1408   score: 3.0   memory length: 299057   epsilon: 0.6058651600085563
steps: 245    lr: 4e-05     evaluation reward: 4.17
episode: 1409   score: 5.0   memory length: 299365   epsilon: 0.6052553200085695
steps: 308    lr: 4e-05     evaluation reward: 4.18
episode: 1410   score: 0.0   memory length: 299488   epsilon: 0.6050117800085748
steps: 123    lr: 4e-05     evaluation reward: 4.17
episode: 1411   score: 5.0   memory length: 299832   epsilon: 0.6043306600085896
steps: 344    lr: 4e-05     evaluation reward: 4.18
episode: 1412   score: 6.0   memory length: 300204   epsilon: 0.6035941000086056
steps: 372    lr: 1.6000000000000003e-05     evaluation reward: 4.17
episode: 1413   score: 5.0   memory length: 300549   epsilon: 0.6029110000086204
steps: 345    lr: 1.6000000000000003e-05     evaluation reward: 4.16
episode: 1414   score: 4.0   memory length: 300846   epsilon: 0.6023229400086332
```

steps: 297   lr: 1.6000000000000003e-05   evaluation reward: 4.17
episode: 1415   score: 7.0   memory length: 301269   epsilon: 0.6014854000086514
steps: 423   lr: 1.6000000000000003e-05   evaluation reward: 4.2
episode: 1416   score: 10.0   memory length: 301647   epsilon:
0.6007369600086676   steps: 378   lr: 1.6000000000000003e-05   evaluation
reward: 4.27
episode: 1417   score: 3.0   memory length: 301877   epsilon: 0.6002815600086775
steps: 230   lr: 1.6000000000000003e-05   evaluation reward: 4.25
episode: 1418   score: 1.0   memory length: 302027   epsilon: 0.5999845600086839
steps: 150   lr: 1.6000000000000003e-05   evaluation reward: 4.23
episode: 1419   score: 7.0   memory length: 302391   epsilon: 0.5992638400086996
steps: 364   lr: 1.6000000000000003e-05   evaluation reward: 4.28
episode: 1420   score: 8.0   memory length: 302834   epsilon: 0.5983867000087186
steps: 443   lr: 1.6000000000000003e-05   evaluation reward: 4.34
episode: 1421   score: 4.0   memory length: 303128   epsilon: 0.5978045800087313
steps: 294   lr: 1.6000000000000003e-05   evaluation reward: 4.35
episode: 1422   score: 6.0   memory length: 303492   epsilon: 0.5970838600087469
steps: 364   lr: 1.6000000000000003e-05   evaluation reward: 4.37
episode: 1423   score: 2.0   memory length: 303689   epsilon: 0.5966938000087554
steps: 197   lr: 1.6000000000000003e-05   evaluation reward: 4.34
episode: 1424   score: 2.0   memory length: 303871   epsilon: 0.5963334400087632
steps: 182   lr: 1.6000000000000003e-05   evaluation reward: 4.35
episode: 1425   score: 6.0   memory length: 304220   epsilon: 0.5956424200087782
steps: 349   lr: 1.6000000000000003e-05   evaluation reward: 4.33
episode: 1426   score: 4.0   memory length: 304479   epsilon: 0.5951296000087893
steps: 259   lr: 1.6000000000000003e-05   evaluation reward: 4.31
episode: 1427   score: 9.0   memory length: 304967   epsilon: 0.5941633600088103
steps: 488   lr: 1.6000000000000003e-05   evaluation reward: 4.36
episode: 1428   score: 5.0   memory length: 305283   epsilon: 0.5935376800088239
steps: 316   lr: 1.6000000000000003e-05   evaluation reward: 4.39
episode: 1429   score: 4.0   memory length: 305545   epsilon: 0.5930189200088352
steps: 262   lr: 1.6000000000000003e-05   evaluation reward: 4.41
episode: 1430   score: 3.0   memory length: 305770   epsilon: 0.5925734200088448
steps: 225   lr: 1.6000000000000003e-05   evaluation reward: 4.41
episode: 1431   score: 4.0   memory length: 306013   epsilon: 0.5920922800088553
steps: 243   lr: 1.6000000000000003e-05   evaluation reward: 4.37
episode: 1432   score: 3.0   memory length: 306258   epsilon: 0.5916071800088658
steps: 245   lr: 1.6000000000000003e-05   evaluation reward: 4.35
episode: 1433   score: 6.0   memory length: 306610   epsilon: 0.5909102200088809
steps: 352   lr: 1.6000000000000003e-05   evaluation reward: 4.35
episode: 1434   score: 6.0   memory length: 306976   epsilon: 0.5901855400088967
steps: 366   lr: 1.6000000000000003e-05   evaluation reward: 4.38
episode: 1435   score: 7.0   memory length: 307382   epsilon: 0.5893816600089141
steps: 406   lr: 1.6000000000000003e-05   evaluation reward: 4.39
episode: 1436   score: 2.0   memory length: 307598   epsilon: 0.5889539800089234
steps: 216   lr: 1.6000000000000003e-05   evaluation reward: 4.36
episode: 1437   score: 5.0   memory length: 307889   epsilon: 0.5883778000089359
steps: 291   lr: 1.6000000000000003e-05   evaluation reward: 4.37

```
episode: 1438   score: 3.0   memory length: 308118   epsilon: 0.5879243800089458
steps: 229    lr: 1.6000000000000003e-05     evaluation reward: 4.35
episode: 1439   score: 7.0   memory length: 308531   epsilon: 0.5871066400089635
steps: 413    lr: 1.6000000000000003e-05     evaluation reward: 4.41
episode: 1440   score: 12.0   memory length: 309102   epsilon: 0.585976060008988
steps: 571    lr: 1.6000000000000003e-05     evaluation reward: 4.48
episode: 1441   score: 1.0   memory length: 309253   epsilon: 0.5856770800089945
steps: 151    lr: 1.6000000000000003e-05     evaluation reward: 4.48
episode: 1442   score: 3.0   memory length: 309482   epsilon: 0.5852236600090044
steps: 229    lr: 1.6000000000000003e-05     evaluation reward: 4.48
episode: 1443   score: 6.0   memory length: 309869   epsilon: 0.584457400009021
steps: 387    lr: 1.6000000000000003e-05     evaluation reward: 4.46
episode: 1444   score: 3.0   memory length: 310096   epsilon: 0.5840079400090308
steps: 227    lr: 1.6000000000000003e-05     evaluation reward: 4.41
episode: 1445   score: 5.0   memory length: 310386   epsilon: 0.5834337400090432
steps: 290    lr: 1.6000000000000003e-05     evaluation reward: 4.43
episode: 1446   score: 4.0   memory length: 310645   epsilon: 0.5829209200090544
steps: 259    lr: 1.6000000000000003e-05     evaluation reward: 4.43
episode: 1447   score: 3.0   memory length: 310876   epsilon: 0.5824635400090643
steps: 231    lr: 1.6000000000000003e-05     evaluation reward: 4.44
episode: 1448   score: 5.0   memory length: 311226   epsilon: 0.5817705400090794
steps: 350    lr: 1.6000000000000003e-05     evaluation reward: 4.45
episode: 1449   score: 2.0   memory length: 311426   epsilon: 0.581374540009088
steps: 200    lr: 1.6000000000000003e-05     evaluation reward: 4.43
episode: 1450   score: 4.0   memory length: 311721   epsilon: 0.5807904400091006
steps: 295    lr: 1.6000000000000003e-05     evaluation reward: 4.42
episode: 1451   score: 1.0   memory length: 311873   epsilon: 0.5804894800091072
steps: 152    lr: 1.6000000000000003e-05     evaluation reward: 4.38
episode: 1452   score: 5.0   memory length: 312154   epsilon: 0.5799331000091192
steps: 281    lr: 1.6000000000000003e-05     evaluation reward: 4.38
episode: 1453   score: 3.0   memory length: 312383   epsilon: 0.5794796800091291
steps: 229    lr: 1.6000000000000003e-05     evaluation reward: 4.38
episode: 1454   score: 8.0   memory length: 312781   epsilon: 0.5786916400091462
steps: 398    lr: 1.6000000000000003e-05     evaluation reward: 4.44
episode: 1455   score: 7.0   memory length: 313182   epsilon: 0.5778976600091634
steps: 401    lr: 1.6000000000000003e-05     evaluation reward: 4.49
episode: 1456   score: 5.0   memory length: 313526   epsilon: 0.5772165400091782
steps: 344    lr: 1.6000000000000003e-05     evaluation reward: 4.51
episode: 1457   score: 5.0   memory length: 313869   epsilon: 0.576537400009193
steps: 343    lr: 1.6000000000000003e-05     evaluation reward: 4.52
episode: 1458   score: 5.0   memory length: 314184   epsilon: 0.5759137000092065
steps: 315    lr: 1.6000000000000003e-05     evaluation reward: 4.53
episode: 1459   score: 10.0   memory length: 314725   epsilon:
0.5748425200092298    steps: 541    lr: 1.6000000000000003e-05     evaluation
reward: 4.54
episode: 1460   score: 6.0   memory length: 315063   epsilon: 0.5741732800092443
steps: 338    lr: 1.6000000000000003e-05     evaluation reward: 4.54
episode: 1461   score: 6.0   memory length: 315412   epsilon: 0.5734822600092593
```

steps: 349   lr: 1.6000000000000003e-05   evaluation reward: 4.58
episode: 1462   score: 9.0   memory length: 315850   epsilon: 0.5726150200092781
steps: 438   lr: 1.6000000000000003e-05   evaluation reward: 4.63
episode: 1463   score: 5.0   memory length: 316137   epsilon: 0.5720467600092904
steps: 287   lr: 1.6000000000000003e-05   evaluation reward: 4.63
episode: 1464   score: 3.0   memory length: 316369   epsilon: 0.5715874000093004
steps: 232   lr: 1.6000000000000003e-05   evaluation reward: 4.59
episode: 1465   score: 3.0   memory length: 316599   epsilon: 0.5711320000093103
steps: 230   lr: 1.6000000000000003e-05   evaluation reward: 4.62
episode: 1466   score: 10.0   memory length: 316991   epsilon:
0.5703558400093272   steps: 392   lr: 1.6000000000000003e-05   evaluation
reward: 4.71
episode: 1467   score: 3.0   memory length: 317219   epsilon: 0.569904400009337
steps: 228   lr: 1.6000000000000003e-05   evaluation reward: 4.71
episode: 1468   score: 6.0   memory length: 317554   epsilon: 0.5692411000093514
steps: 335   lr: 1.6000000000000003e-05   evaluation reward: 4.74
episode: 1469   score: 9.0   memory length: 318053   epsilon: 0.5682530800093728
steps: 499   lr: 1.6000000000000003e-05   evaluation reward: 4.79
episode: 1470   score: 6.0   memory length: 318428   epsilon: 0.5675105800093889
steps: 375   lr: 1.6000000000000003e-05   evaluation reward: 4.74
episode: 1471   score: 4.0   memory length: 318691   epsilon: 0.5669898400094002
steps: 263   lr: 1.6000000000000003e-05   evaluation reward: 4.76
episode: 1472   score: 10.0   memory length: 319185   epsilon:
0.5660117200094215   steps: 494   lr: 1.6000000000000003e-05   evaluation
reward: 4.79
episode: 1473   score: 10.0   memory length: 319738   epsilon:
0.5649167800094452   steps: 553   lr: 1.6000000000000003e-05   evaluation
reward: 4.85
episode: 1474   score: 5.0   memory length: 320047   epsilon: 0.5643049600094585
steps: 309   lr: 1.6000000000000003e-05   evaluation reward: 4.86
episode: 1475   score: 1.0   memory length: 320197   epsilon: 0.564007960009465
steps: 150   lr: 1.6000000000000003e-05   evaluation reward: 4.81
episode: 1476   score: 6.0   memory length: 320573   epsilon: 0.5632634800094811
steps: 376   lr: 1.6000000000000003e-05   evaluation reward: 4.81
episode: 1477   score: 4.0   memory length: 320869   epsilon: 0.5626774000094938
steps: 296   lr: 1.6000000000000003e-05   evaluation reward: 4.79
episode: 1478   score: 4.0   memory length: 321148   epsilon: 0.5621249800095058
steps: 279   lr: 1.6000000000000003e-05   evaluation reward: 4.81
episode: 1479   score: 4.0   memory length: 321423   epsilon: 0.5615804800095177
steps: 275   lr: 1.6000000000000003e-05   evaluation reward: 4.81
episode: 1480   score: 3.0   memory length: 321634   epsilon: 0.5611627000095267
steps: 211   lr: 1.6000000000000003e-05   evaluation reward: 4.8
episode: 1481   score: 4.0   memory length: 321912   epsilon: 0.5606122600095387
steps: 278   lr: 1.6000000000000003e-05   evaluation reward: 4.8
episode: 1482   score: 4.0   memory length: 322190   epsilon: 0.5600618200095506
steps: 278   lr: 1.6000000000000003e-05   evaluation reward: 4.8
episode: 1483   score: 3.0   memory length: 322402   epsilon: 0.5596420600095597
steps: 212   lr: 1.6000000000000003e-05   evaluation reward: 4.77

```
episode: 1484   score: 5.0   memory length: 322711   epsilon: 0.559030240009573
steps: 309    lr: 1.6000000000000003e-05     evaluation reward: 4.8
episode: 1485   score: 5.0   memory length: 323013   epsilon: 0.558432280009586
steps: 302    lr: 1.6000000000000003e-05     evaluation reward: 4.84
episode: 1486   score: 8.0   memory length: 323457   epsilon: 0.5575531600096051
steps: 444    lr: 1.6000000000000003e-05     evaluation reward: 4.89
episode: 1487   score: 5.0   memory length: 323799   epsilon: 0.5568760000096198
steps: 342    lr: 1.6000000000000003e-05     evaluation reward: 4.9
episode: 1488   score: 4.0   memory length: 324074   epsilon: 0.5563315000096316
steps: 275    lr: 1.6000000000000003e-05     evaluation reward: 4.86
episode: 1489   score: 8.0   memory length: 324485   epsilon: 0.5555177200096493
steps: 411    lr: 1.6000000000000003e-05     evaluation reward: 4.9
episode: 1490   score: 3.0   memory length: 324715   epsilon: 0.5550623200096592
steps: 230    lr: 1.6000000000000003e-05     evaluation reward: 4.88
episode: 1491   score: 4.0   memory length: 324977   epsilon: 0.5545435600096704
steps: 262    lr: 1.6000000000000003e-05     evaluation reward: 4.85
episode: 1492   score: 0.0   memory length: 325100   epsilon: 0.5543000200096757
steps: 123    lr: 1.6000000000000003e-05     evaluation reward: 4.81
episode: 1493   score: 6.0   memory length: 325478   epsilon: 0.553551580009692
steps: 378    lr: 1.6000000000000003e-05     evaluation reward: 4.85
episode: 1494   score: 6.0   memory length: 325774   epsilon: 0.5529655000097047
steps: 296    lr: 1.6000000000000003e-05     evaluation reward: 4.87
episode: 1495   score: 3.0   memory length: 326021   epsilon: 0.5524764400097153
steps: 247    lr: 1.6000000000000003e-05     evaluation reward: 4.86
episode: 1496   score: 4.0   memory length: 326298   epsilon: 0.5519279800097272
steps: 277    lr: 1.6000000000000003e-05     evaluation reward: 4.87
episode: 1497   score: 3.0   memory length: 326543   epsilon: 0.5514428800097377
steps: 245    lr: 1.6000000000000003e-05     evaluation reward: 4.85
episode: 1498   score: 3.0   memory length: 326758   epsilon: 0.551017180009747
steps: 215    lr: 1.6000000000000003e-05     evaluation reward: 4.86
episode: 1499   score: 5.0   memory length: 327086   epsilon: 0.5503677400097611
steps: 328    lr: 1.6000000000000003e-05     evaluation reward: 4.87
episode: 1500   score: 6.0   memory length: 327460   epsilon: 0.5496272200097772
steps: 374    lr: 1.6000000000000003e-05     evaluation reward: 4.89
episode: 1501   score: 3.0   memory length: 327707   epsilon: 0.5491381600097878
steps: 247    lr: 1.6000000000000003e-05     evaluation reward: 4.89
episode: 1502   score: 2.0   memory length: 327906   epsilon: 0.5487441400097963
steps: 199    lr: 1.6000000000000003e-05     evaluation reward: 4.85
episode: 1503   score: 6.0   memory length: 328274   epsilon: 0.5480155000098121
steps: 368    lr: 1.6000000000000003e-05     evaluation reward: 4.86
episode: 1504   score: 7.0   memory length: 328711   epsilon: 0.5471502400098309
steps: 437    lr: 1.6000000000000003e-05     evaluation reward: 4.87
episode: 1505   score: 9.0   memory length: 329220   epsilon: 0.5461424200098528
steps: 509    lr: 1.6000000000000003e-05     evaluation reward: 4.92
episode: 1506   score: 4.0   memory length: 329495   epsilon: 0.5455979200098646
steps: 275    lr: 1.6000000000000003e-05     evaluation reward: 4.93
episode: 1507   score: 4.0   memory length: 329789   epsilon: 0.5450158000098773
steps: 294    lr: 1.6000000000000003e-05     evaluation reward: 4.9
```

episode: 1508   score: 6.0   memory length: 330126   epsilon: 0.5443485400098917
steps: 337   lr: 1.6000000000000003e-05   evaluation reward: 4.93
episode: 1509   score: 6.0   memory length: 330492   epsilon: 0.5436238600099075
steps: 366   lr: 1.6000000000000003e-05   evaluation reward: 4.94
episode: 1510   score: 5.0   memory length: 330836   epsilon: 0.5429427400099223
steps: 344   lr: 1.6000000000000003e-05   evaluation reward: 4.99
episode: 1511   score: 2.0   memory length: 331054   epsilon: 0.5425111000099316
steps: 218   lr: 1.6000000000000003e-05   evaluation reward: 4.96
episode: 1512   score: 3.0   memory length: 331267   epsilon: 0.5420893600099408
steps: 213   lr: 1.6000000000000003e-05   evaluation reward: 4.93
episode: 1513   score: 12.0   memory length: 331769   epsilon:
0.5410954000099624   steps: 502   lr: 1.6000000000000003e-05   evaluation
reward: 5.0
episode: 1514   score: 5.0   memory length: 332059   epsilon: 0.5405212000099748
steps: 290   lr: 1.6000000000000003e-05   evaluation reward: 5.01
episode: 1515   score: 5.0   memory length: 332349   epsilon: 0.5399470000099873
steps: 290   lr: 1.6000000000000003e-05   evaluation reward: 4.99
episode: 1516   score: 8.0   memory length: 332803   epsilon: 0.5390480800100068
steps: 454   lr: 1.6000000000000003e-05   evaluation reward: 4.97
episode: 1517   score: 7.0   memory length: 333226   epsilon: 0.538210540010025
steps: 423   lr: 1.6000000000000003e-05   evaluation reward: 5.01
episode: 1518   score: 3.0   memory length: 333457   epsilon: 0.5377531600100349
steps: 231   lr: 1.6000000000000003e-05   evaluation reward: 5.03
episode: 1519   score: 4.0   memory length: 333718   epsilon: 0.5372363800100461
steps: 261   lr: 1.6000000000000003e-05   evaluation reward: 5.0
episode: 1520   score: 7.0   memory length: 334100   epsilon: 0.5364800200100626
steps: 382   lr: 1.6000000000000003e-05   evaluation reward: 4.99
episode: 1521   score: 7.0   memory length: 334506   epsilon: 0.53567614001008
steps: 406   lr: 1.6000000000000003e-05   evaluation reward: 5.02
episode: 1522   score: 3.0   memory length: 334719   epsilon: 0.5352544000100892
steps: 213   lr: 1.6000000000000003e-05   evaluation reward: 4.99
episode: 1523   score: 3.0   memory length: 334947   epsilon: 0.534802960010099
steps: 228   lr: 1.6000000000000003e-05   evaluation reward: 5.0
episode: 1524   score: 4.0   memory length: 335223   epsilon: 0.5342564800101108
steps: 276   lr: 1.6000000000000003e-05   evaluation reward: 5.02
episode: 1525   score: 6.0   memory length: 335597   epsilon: 0.5335159600101269
steps: 374   lr: 1.6000000000000003e-05   evaluation reward: 5.02
episode: 1526   score: 4.0   memory length: 335874   epsilon: 0.5329675000101388
steps: 277   lr: 1.6000000000000003e-05   evaluation reward: 5.02
episode: 1527   score: 4.0   memory length: 336173   epsilon: 0.5323754800101517
steps: 299   lr: 1.6000000000000003e-05   evaluation reward: 4.97
episode: 1528   score: 4.0   memory length: 336431   epsilon: 0.5318646400101628
steps: 258   lr: 1.6000000000000003e-05   evaluation reward: 4.96
episode: 1529   score: 9.0   memory length: 336909   epsilon: 0.5309182000101833
steps: 478   lr: 1.6000000000000003e-05   evaluation reward: 5.01
episode: 1530   score: 5.0   memory length: 337213   epsilon: 0.5303162800101964
steps: 304   lr: 1.6000000000000003e-05   evaluation reward: 5.03
episode: 1531   score: 4.0   memory length: 337511   epsilon: 0.5297262400102092

steps: 298   lr: 1.6000000000000003e-05   evaluation reward: 5.03
episode: 1532   score: 4.0   memory length: 337791   epsilon: 0.5291718400102212
steps: 280   lr: 1.6000000000000003e-05   evaluation reward: 5.04
episode: 1533   score: 6.0   memory length: 338132   epsilon: 0.5284966600102359
steps: 341   lr: 1.6000000000000003e-05   evaluation reward: 5.04
episode: 1534   score: 6.0   memory length: 338527   epsilon: 0.5277145600102529
steps: 395   lr: 1.6000000000000003e-05   evaluation reward: 5.04
episode: 1535   score: 7.0   memory length: 338913   epsilon: 0.5269502800102694
steps: 386   lr: 1.6000000000000003e-05   evaluation reward: 5.04
episode: 1536   score: 11.0   memory length: 339317   epsilon:
0.5261503600102868   steps: 404   lr: 1.6000000000000003e-05   evaluation
reward: 5.13
episode: 1537   score: 5.0   memory length: 339641   epsilon: 0.5255088400103007
steps: 324   lr: 1.6000000000000003e-05   evaluation reward: 5.13
episode: 1538   score: 6.0   memory length: 340040   epsilon: 0.5247188200103179
steps: 399   lr: 1.6000000000000003e-05   evaluation reward: 5.16
episode: 1539   score: 1.0   memory length: 340191   epsilon: 0.5244198400103244
steps: 151   lr: 1.6000000000000003e-05   evaluation reward: 5.1
episode: 1540   score: 5.0   memory length: 340503   epsilon: 0.5238020800103378
steps: 312   lr: 1.6000000000000003e-05   evaluation reward: 5.03
episode: 1541   score: 7.0   memory length: 340893   epsilon: 0.5230298800103546
steps: 390   lr: 1.6000000000000003e-05   evaluation reward: 5.09
episode: 1542   score: 1.0   memory length: 341044   epsilon: 0.522730900010361
steps: 151   lr: 1.6000000000000003e-05   evaluation reward: 5.07
episode: 1543   score: 5.0   memory length: 341369   epsilon: 0.522087400010375
steps: 325   lr: 1.6000000000000003e-05   evaluation reward: 5.06
episode: 1544   score: 4.0   memory length: 341625   epsilon: 0.521580520010386
steps: 256   lr: 1.6000000000000003e-05   evaluation reward: 5.07
episode: 1545   score: 7.0   memory length: 342046   epsilon: 0.5207469400104041
steps: 421   lr: 1.6000000000000003e-05   evaluation reward: 5.09
episode: 1546   score: 8.0   memory length: 342502   epsilon: 0.5198440600104237
steps: 456   lr: 1.6000000000000003e-05   evaluation reward: 5.13
episode: 1547   score: 2.0   memory length: 342720   epsilon: 0.5194124200104331
steps: 218   lr: 1.6000000000000003e-05   evaluation reward: 5.12
episode: 1548   score: 2.0   memory length: 342920   epsilon: 0.5190164200104417
steps: 200   lr: 1.6000000000000003e-05   evaluation reward: 5.09
episode: 1549   score: 6.0   memory length: 343276   epsilon: 0.518311540010457
steps: 356   lr: 1.6000000000000003e-05   evaluation reward: 5.13
episode: 1550   score: 4.0   memory length: 343558   epsilon: 0.5177531800104691
steps: 282   lr: 1.6000000000000003e-05   evaluation reward: 5.13
episode: 1551   score: 11.0   memory length: 344063   epsilon:
0.5167532800104908   steps: 505   lr: 1.6000000000000003e-05   evaluation
reward: 5.23
episode: 1552   score: 7.0   memory length: 344501   epsilon: 0.5158860400105096
steps: 438   lr: 1.6000000000000003e-05   evaluation reward: 5.25
episode: 1553   score: 6.0   memory length: 344833   epsilon: 0.5152286800105239
steps: 332   lr: 1.6000000000000003e-05   evaluation reward: 5.28
episode: 1554   score: 10.0   memory length: 345322   epsilon:

0.5142604600105449 steps: 489 lr: 1.6000000000000003e-05 evaluation reward: 5.3
episode: 1555 score: 7.0 memory length: 345749 epsilon: 0.5134150000105633 steps: 427 lr: 1.6000000000000003e-05 evaluation reward: 5.3
episode: 1556 score: 4.0 memory length: 346029 epsilon: 0.5128606000105753 steps: 280 lr: 1.6000000000000003e-05 evaluation reward: 5.29
episode: 1557 score: 4.0 memory length: 346309 epsilon: 0.5123062000105874 steps: 280 lr: 1.6000000000000003e-05 evaluation reward: 5.28
episode: 1558 score: 10.0 memory length: 346841 epsilon: 0.5112528400106102 steps: 532 lr: 1.6000000000000003e-05 evaluation reward: 5.33
episode: 1559 score: 7.0 memory length: 347252 epsilon: 0.5104390600106279 steps: 411 lr: 1.6000000000000003e-05 evaluation reward: 5.3
episode: 1560 score: 8.0 memory length: 347699 epsilon: 0.5095540000106471 steps: 447 lr: 1.6000000000000003e-05 evaluation reward: 5.32
episode: 1561 score: 1.0 memory length: 347850 epsilon: 0.5092550200106536 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 5.27
episode: 1562 score: 4.0 memory length: 348107 epsilon: 0.5087461600106646 steps: 257 lr: 1.6000000000000003e-05 evaluation reward: 5.22
episode: 1563 score: 3.0 memory length: 348317 epsilon: 0.5083303600106737 steps: 210 lr: 1.6000000000000003e-05 evaluation reward: 5.2
episode: 1564 score: 11.0 memory length: 348797 epsilon: 0.5073799600106943 steps: 480 lr: 1.6000000000000003e-05 evaluation reward: 5.28
episode: 1565 score: 8.0 memory length: 349257 epsilon: 0.5064691600107141 steps: 460 lr: 1.6000000000000003e-05 evaluation reward: 5.33
episode: 1566 score: 4.0 memory length: 349534 epsilon: 0.505920700010726 steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 5.27
episode: 1567 score: 7.0 memory length: 349936 epsilon: 0.5051247400107433 steps: 402 lr: 1.6000000000000003e-05 evaluation reward: 5.31
episode: 1568 score: 4.0 memory length: 350195 epsilon: 0.5046119200107544 steps: 259 lr: 1.6000000000000003e-05 evaluation reward: 5.29
episode: 1569 score: 3.0 memory length: 350444 epsilon: 0.5041189000107651 steps: 249 lr: 1.6000000000000003e-05 evaluation reward: 5.23
episode: 1570 score: 6.0 memory length: 350785 epsilon: 0.5034437200107797 steps: 341 lr: 1.6000000000000003e-05 evaluation reward: 5.23
episode: 1571 score: 7.0 memory length: 351174 epsilon: 0.5026735000107965 steps: 389 lr: 1.6000000000000003e-05 evaluation reward: 5.26
episode: 1572 score: 7.0 memory length: 351537 epsilon: 0.5019547600108121 steps: 363 lr: 1.6000000000000003e-05 evaluation reward: 5.23
episode: 1573 score: 7.0 memory length: 351955 epsilon: 0.50112712001083 steps: 418 lr: 1.6000000000000003e-05 evaluation reward: 5.2
episode: 1574 score: 6.0 memory length: 352369 epsilon: 0.5003074000108478 steps: 414 lr: 1.6000000000000003e-05 evaluation reward: 5.21
episode: 1575 score: 7.0 memory length: 352758 epsilon: 0.49953718001085157 steps: 389 lr: 1.6000000000000003e-05 evaluation reward: 5.27
episode: 1576 score: 2.0 memory length: 352957 epsilon: 0.4991431600108491

```
steps: 199    lr: 1.6000000000000003e-05     evaluation reward: 5.23
episode: 1577   score: 8.0   memory length: 353382   epsilon:
0.49830166001084375    steps: 425    lr: 1.6000000000000003e-05     evaluation
reward: 5.27
episode: 1578   score: 5.0   memory length: 353690   epsilon: 0.4976918200108399
steps: 308    lr: 1.6000000000000003e-05     evaluation reward: 5.28
episode: 1579   score: 5.0   memory length: 354016   epsilon: 0.4970463400108358
steps: 326    lr: 1.6000000000000003e-05     evaluation reward: 5.29
episode: 1580   score: 7.0   memory length: 354422   epsilon: 0.4962424600108307
steps: 406    lr: 1.6000000000000003e-05     evaluation reward: 5.33
episode: 1581   score: 4.0   memory length: 354696   epsilon: 0.4956999400108273
steps: 274    lr: 1.6000000000000003e-05     evaluation reward: 5.33
episode: 1582   score: 11.0   memory length: 355270   epsilon:
0.4945634200108201    steps: 574    lr: 1.6000000000000003e-05     evaluation
reward: 5.4
episode: 1583   score: 9.0   memory length: 355595   epsilon: 0.493919920010816
steps: 325    lr: 1.6000000000000003e-05     evaluation reward: 5.46
episode: 1584   score: 5.0   memory length: 355903   epsilon:
0.49331008001081217    steps: 308    lr: 1.6000000000000003e-05     evaluation
reward: 5.46
episode: 1585   score: 8.0   memory length: 356337   epsilon:
0.49245076001080673    steps: 434    lr: 1.6000000000000003e-05     evaluation
reward: 5.49
episode: 1586   score: 6.0   memory length: 356672   epsilon:
0.49178746001080254    steps: 335    lr: 1.6000000000000003e-05     evaluation
reward: 5.47
episode: 1587   score: 8.0   memory length: 357134   epsilon:
0.49087270001079675    steps: 462    lr: 1.6000000000000003e-05     evaluation
reward: 5.5
episode: 1588   score: 4.0   memory length: 357407   epsilon: 0.4903321600107933
steps: 273    lr: 1.6000000000000003e-05     evaluation reward: 5.5
episode: 1589   score: 14.0   memory length: 357929   epsilon:
0.4892986000107868    steps: 522    lr: 1.6000000000000003e-05     evaluation
reward: 5.56
episode: 1590   score: 3.0   memory length: 358160   epsilon: 0.4888412200107839
steps: 231    lr: 1.6000000000000003e-05     evaluation reward: 5.56
episode: 1591   score: 4.0   memory length: 358420   epsilon:
0.48832642001078064    steps: 260    lr: 1.6000000000000003e-05     evaluation
reward: 5.56
episode: 1592   score: 3.0   memory length: 358634   epsilon:
0.48790270001077796    steps: 214    lr: 1.6000000000000003e-05     evaluation
reward: 5.59
episode: 1593   score: 12.0   memory length: 359202   epsilon:
0.48677806001077084    steps: 568    lr: 1.6000000000000003e-05     evaluation
reward: 5.65
episode: 1594   score: 8.0   memory length: 359624   epsilon:
0.48594250001076555    steps: 422    lr: 1.6000000000000003e-05     evaluation
reward: 5.67
```

episode: 1595   score: 3.0   memory length: 359850   epsilon: 0.4854950200107627
steps: 226   lr: 1.6000000000000003e-05   evaluation reward: 5.67
episode: 1596   score: 8.0   memory length: 360299   epsilon: 0.4846060000107571
steps: 449   lr: 1.6000000000000003e-05   evaluation reward: 5.71
episode: 1597   score: 9.0   memory length: 360775   epsilon:
0.48366352001075114   steps: 476   lr: 1.6000000000000003e-05   evaluation
reward: 5.77
episode: 1598   score: 6.0   memory length: 361171   epsilon: 0.4828794400107462
steps: 396   lr: 1.6000000000000003e-05   evaluation reward: 5.8
episode: 1599   score: 4.0   memory length: 361471   epsilon: 0.4822854400107424
steps: 300   lr: 1.6000000000000003e-05   evaluation reward: 5.79
episode: 1600   score: 5.0   memory length: 361744   epsilon: 0.481744900010739
steps: 273   lr: 1.6000000000000003e-05   evaluation reward: 5.78
episode: 1601   score: 5.0   memory length: 362052   epsilon:
0.48113506001073514   steps: 308   lr: 1.6000000000000003e-05   evaluation
reward: 5.8
episode: 1602   score: 12.0   memory length: 362455   epsilon:
0.4803371200107301   steps: 403   lr: 1.6000000000000003e-05   evaluation
reward: 5.9
episode: 1603   score: 6.0   memory length: 362830   epsilon: 0.4795946200107254
steps: 375   lr: 1.6000000000000003e-05   evaluation reward: 5.9
episode: 1604   score: 5.0   memory length: 363134   epsilon: 0.4789927000107216
steps: 304   lr: 1.6000000000000003e-05   evaluation reward: 5.88
episode: 1605   score: 7.0   memory length: 363541   epsilon: 0.4781868400107165
steps: 407   lr: 1.6000000000000003e-05   evaluation reward: 5.86
episode: 1606   score: 2.0   memory length: 363739   epsilon: 0.477794800010714
steps: 198   lr: 1.6000000000000003e-05   evaluation reward: 5.84
episode: 1607   score: 9.0   memory length: 364214   epsilon:
0.47685430001070805   steps: 475   lr: 1.6000000000000003e-05   evaluation
reward: 5.89
episode: 1608   score: 8.0   memory length: 364678   epsilon:
0.47593558001070224   steps: 464   lr: 1.6000000000000003e-05   evaluation
reward: 5.91
episode: 1609   score: 6.0   memory length: 365069   epsilon:
0.47516140001069734   steps: 391   lr: 1.6000000000000003e-05   evaluation
reward: 5.91
episode: 1610   score: 7.0   memory length: 365510   epsilon: 0.4742882200106918
steps: 441   lr: 1.6000000000000003e-05   evaluation reward: 5.93
episode: 1611   score: 7.0   memory length: 365883   epsilon:
0.47354968001068715   steps: 373   lr: 1.6000000000000003e-05   evaluation
reward: 5.98
episode: 1612   score: 8.0   memory length: 366317   epsilon: 0.4726903600106817
steps: 434   lr: 1.6000000000000003e-05   evaluation reward: 6.03
episode: 1613   score: 6.0   memory length: 366649   epsilon:
0.47203300001067755   steps: 332   lr: 1.6000000000000003e-05   evaluation
reward: 5.97
episode: 1614   score: 4.0   memory length: 366907   epsilon: 0.4715221600106743
steps: 258   lr: 1.6000000000000003e-05   evaluation reward: 5.96

episode: 1615   score: 8.0   memory length: 367362   epsilon: 0.4706212600106686
steps: 455   lr: 1.6000000000000003e-05   evaluation reward: 5.99
episode: 1616   score: 7.0   memory length: 367764   epsilon: 0.4698253000106636
steps: 402   lr: 1.6000000000000003e-05   evaluation reward: 5.98
episode: 1617   score: 4.0   memory length: 368040   epsilon: 0.4692788200106601
steps: 276   lr: 1.6000000000000003e-05   evaluation reward: 5.95
episode: 1618   score: 3.0   memory length: 368267   epsilon: 0.4688293600106573
steps: 227   lr: 1.6000000000000003e-05   evaluation reward: 5.95
episode: 1619   score: 14.0   memory length: 368844   epsilon:
0.46768690001065005   steps: 577   lr: 1.6000000000000003e-05   evaluation
reward: 6.05
episode: 1620   score: 1.0   memory length: 368995   epsilon:
0.46738792001064816   steps: 151   lr: 1.6000000000000003e-05   evaluation
reward: 5.99
episode: 1621   score: 9.0   memory length: 369467   epsilon:
0.46645336001064225   steps: 472   lr: 1.6000000000000003e-05   evaluation
reward: 6.01
episode: 1622   score: 6.0   memory length: 369845   epsilon: 0.4657049200106375
steps: 378   lr: 1.6000000000000003e-05   evaluation reward: 6.04
episode: 1623   score: 9.0   memory length: 370303   epsilon: 0.4647980800106318
steps: 458   lr: 1.6000000000000003e-05   evaluation reward: 6.1
episode: 1624   score: 6.0   memory length: 370659   epsilon: 0.4640932000106273
steps: 356   lr: 1.6000000000000003e-05   evaluation reward: 6.12
episode: 1625   score: 9.0   memory length: 371105   epsilon:
0.46321012001062173   steps: 446   lr: 1.6000000000000003e-05   evaluation
reward: 6.15
episode: 1626   score: 10.0   memory length: 371631   epsilon:
0.46216864001061514   steps: 526   lr: 1.6000000000000003e-05   evaluation
reward: 6.21
episode: 1627   score: 6.0   memory length: 371960   epsilon: 0.461517220010611
steps: 329   lr: 1.6000000000000003e-05   evaluation reward: 6.23
episode: 1628   score: 6.0   memory length: 372265   epsilon: 0.4609133200106072
steps: 305   lr: 1.6000000000000003e-05   evaluation reward: 6.25
episode: 1629   score: 7.0   memory length: 372661   epsilon:
0.46012924001060224   steps: 396   lr: 1.6000000000000003e-05   evaluation
reward: 6.23
episode: 1630   score: 4.0   memory length: 372956   epsilon:
0.45954514001059854   steps: 295   lr: 1.6000000000000003e-05   evaluation
reward: 6.22
episode: 1631   score: 7.0   memory length: 373383   epsilon: 0.4586996800105932
steps: 427   lr: 1.6000000000000003e-05   evaluation reward: 6.25
episode: 1632   score: 10.0   memory length: 373832   epsilon:
0.45781066001058757   steps: 449   lr: 1.6000000000000003e-05   evaluation
reward: 6.31
episode: 1633   score: 4.0   memory length: 374106   epsilon:
0.45726814001058413   steps: 274   lr: 1.6000000000000003e-05   evaluation
reward: 6.29
episode: 1634   score: 9.0   memory length: 374627   epsilon: 0.4562365600105776

steps: 521   lr: 1.6000000000000003e-05   evaluation reward: 6.32
episode: 1635   score: 3.0   memory length: 374853   epsilon: 0.4557890800105748
steps: 226   lr: 1.6000000000000003e-05   evaluation reward: 6.28
episode: 1636   score: 3.0   memory length: 375064   epsilon:
0.45537130001057213   steps: 211   lr: 1.6000000000000003e-05   evaluation
reward: 6.2
episode: 1637   score: 7.0   memory length: 375474   epsilon: 0.454559500010567
steps: 410   lr: 1.6000000000000003e-05   evaluation reward: 6.22
episode: 1638   score: 6.0   memory length: 375812   epsilon:
0.45389026001056276   steps: 338   lr: 1.6000000000000003e-05   evaluation
reward: 6.22
episode: 1639   score: 7.0   memory length: 376221   epsilon:
0.45308044001055764   steps: 409   lr: 1.6000000000000003e-05   evaluation
reward: 6.28
episode: 1640   score: 6.0   memory length: 376542   epsilon: 0.4524448600105536
steps: 321   lr: 1.6000000000000003e-05   evaluation reward: 6.29
episode: 1641   score: 6.0   memory length: 376877   epsilon: 0.4517815600105494
steps: 335   lr: 1.6000000000000003e-05   evaluation reward: 6.28
episode: 1642   score: 12.0   memory length: 377333   epsilon:
0.4508786800105437   steps: 456   lr: 1.6000000000000003e-05   evaluation
reward: 6.39
episode: 1643   score: 3.0   memory length: 377563   epsilon: 0.4504232800105408
steps: 230   lr: 1.6000000000000003e-05   evaluation reward: 6.37
episode: 1644   score: 7.0   memory length: 377952   epsilon:
0.44965306001053595   steps: 389   lr: 1.6000000000000003e-05   evaluation
reward: 6.4
episode: 1645   score: 5.0   memory length: 378281   epsilon:
0.44900164001053183   steps: 329   lr: 1.6000000000000003e-05   evaluation
reward: 6.38
episode: 1646   score: 11.0   memory length: 378674   epsilon:
0.4482235000105269   steps: 393   lr: 1.6000000000000003e-05   evaluation
reward: 6.41
episode: 1647   score: 5.0   memory length: 378968   epsilon: 0.4476413800105232
steps: 294   lr: 1.6000000000000003e-05   evaluation reward: 6.44
episode: 1648   score: 13.0   memory length: 379459   epsilon:
0.4466692000105171   steps: 491   lr: 1.6000000000000003e-05   evaluation
reward: 6.55
episode: 1649   score: 7.0   memory length: 379867   epsilon:
0.44586136001051196   steps: 408   lr: 1.6000000000000003e-05   evaluation
reward: 6.56
episode: 1650   score: 6.0   memory length: 380244   epsilon:
0.44511490001050724   steps: 377   lr: 1.6000000000000003e-05   evaluation
reward: 6.58
episode: 1651   score: 9.0   memory length: 380727   epsilon: 0.4441585600105012
steps: 483   lr: 1.6000000000000003e-05   evaluation reward: 6.56
episode: 1652   score: 7.0   memory length: 381100   epsilon: 0.4434200200104965
steps: 373   lr: 1.6000000000000003e-05   evaluation reward: 6.56
episode: 1653   score: 4.0   memory length: 381360   epsilon:

0.44290522001049326   steps: 260   lr: 1.6000000000000003e-05   evaluation
reward: 6.54
episode: 1654   score: 4.0   memory length: 381615   epsilon:
0.44240032001049007   steps: 255   lr: 1.6000000000000003e-05   evaluation
reward: 6.48
episode: 1655   score: 10.0   memory length: 382131   epsilon:
0.4413786400104836   steps: 516   lr: 1.6000000000000003e-05   evaluation
reward: 6.51
episode: 1656   score: 6.0   memory length: 382445   epsilon:
0.44075692001047967   steps: 314   lr: 1.6000000000000003e-05   evaluation
reward: 6.53
episode: 1657   score: 10.0   memory length: 382821   epsilon:
0.44001244001047496   steps: 376   lr: 1.6000000000000003e-05   evaluation
reward: 6.59
episode: 1658   score: 3.0   memory length: 383067   epsilon: 0.4395253600104719
steps: 246   lr: 1.6000000000000003e-05   evaluation reward: 6.52
episode: 1659   score: 10.0   memory length: 383594   epsilon:
0.4384819000104653   steps: 527   lr: 1.6000000000000003e-05   evaluation
reward: 6.55
episode: 1660   score: 10.0   memory length: 384125   epsilon:
0.4374305200104586   steps: 531   lr: 1.6000000000000003e-05   evaluation
reward: 6.57
episode: 1661   score: 4.0   memory length: 384400   epsilon: 0.4368860200104552
steps: 275   lr: 1.6000000000000003e-05   evaluation reward: 6.6
episode: 1662   score: 7.0   memory length: 384824   epsilon:
0.43604650001044987   steps: 424   lr: 1.6000000000000003e-05   evaluation
reward: 6.63
episode: 1663   score: 7.0   memory length: 385186   epsilon:
0.43532974001044533   steps: 362   lr: 1.6000000000000003e-05   evaluation
reward: 6.67
episode: 1664   score: 3.0   memory length: 385436   epsilon: 0.4348347400104422
steps: 250   lr: 1.6000000000000003e-05   evaluation reward: 6.59
episode: 1665   score: 5.0   memory length: 385728   epsilon:
0.43425658001043854   steps: 292   lr: 1.6000000000000003e-05   evaluation
reward: 6.56
episode: 1666   score: 4.0   memory length: 385990   epsilon:
0.43373782001043526   steps: 262   lr: 1.6000000000000003e-05   evaluation
reward: 6.56
episode: 1667   score: 6.0   memory length: 386385   epsilon: 0.4329557200104303
steps: 395   lr: 1.6000000000000003e-05   evaluation reward: 6.55
episode: 1668   score: 7.0   memory length: 386758   epsilon:
0.43221718001042564   steps: 373   lr: 1.6000000000000003e-05   evaluation
reward: 6.58
episode: 1669   score: 9.0   memory length: 387231   epsilon: 0.4312806400104197
steps: 473   lr: 1.6000000000000003e-05   evaluation reward: 6.64
episode: 1670   score: 5.0   memory length: 387575   epsilon: 0.4305995200104154
steps: 344   lr: 1.6000000000000003e-05   evaluation reward: 6.63
episode: 1671   score: 11.0   memory length: 388166   epsilon: 0.429429340010408

steps: 591   lr: 1.6000000000000003e-05   evaluation reward: 6.67
episode: 1672   score: 7.0   memory length: 388572   epsilon: 0.4286254600104029
steps: 406   lr: 1.6000000000000003e-05   evaluation reward: 6.67
episode: 1673   score: 8.0   memory length: 388988   epsilon: 0.4278017800103977
steps: 416   lr: 1.6000000000000003e-05   evaluation reward: 6.68
episode: 1674   score: 6.0   memory length: 389340   epsilon: 0.4271048200103933
steps: 352   lr: 1.6000000000000003e-05   evaluation reward: 6.68
episode: 1675   score: 6.0   memory length: 389680   epsilon:
0.42643162001038903   steps: 340   lr: 1.6000000000000003e-05   evaluation
reward: 6.67
episode: 1676   score: 7.0   memory length: 390079   epsilon:
0.42564160001038404   steps: 399   lr: 1.6000000000000003e-05   evaluation
reward: 6.72
episode: 1677   score: 7.0   memory length: 390504   epsilon: 0.4248001000103787
steps: 425   lr: 1.6000000000000003e-05   evaluation reward: 6.71
episode: 1678   score: 3.0   memory length: 390715   epsilon:
0.42438232001037607   steps: 211   lr: 1.6000000000000003e-05   evaluation
reward: 6.69
episode: 1679   score: 7.0   memory length: 391102   epsilon: 0.4236160600103712
steps: 387   lr: 1.6000000000000003e-05   evaluation reward: 6.71
episode: 1680   score: 4.0   memory length: 391380   epsilon:
0.42306562001036774   steps: 278   lr: 1.6000000000000003e-05   evaluation
reward: 6.68
episode: 1681   score: 7.0   memory length: 391793   epsilon:
0.42224788001036256   steps: 413   lr: 1.6000000000000003e-05   evaluation
reward: 6.71
episode: 1682   score: 7.0   memory length: 392179   epsilon:
0.42148360001035773   steps: 386   lr: 1.6000000000000003e-05   evaluation
reward: 6.67
episode: 1683   score: 3.0   memory length: 392409   epsilon:
0.42102820001035485   steps: 230   lr: 1.6000000000000003e-05   evaluation
reward: 6.61
episode: 1684   score: 4.0   memory length: 392671   epsilon:
0.42050944001035157   steps: 262   lr: 1.6000000000000003e-05   evaluation
reward: 6.6
episode: 1685   score: 7.0   memory length: 393074   epsilon: 0.4197115000103465
steps: 403   lr: 1.6000000000000003e-05   evaluation reward: 6.59
episode: 1686   score: 9.0   memory length: 393445   epsilon:
0.41897692001034187   steps: 371   lr: 1.6000000000000003e-05   evaluation
reward: 6.62
episode: 1687   score: 9.0   memory length: 393948   epsilon:
0.41798098001033557   steps: 503   lr: 1.6000000000000003e-05   evaluation
reward: 6.63
episode: 1688   score: 7.0   memory length: 394353   epsilon: 0.4171790800103305
steps: 405   lr: 1.6000000000000003e-05   evaluation reward: 6.66
episode: 1689   score: 4.0   memory length: 394631   epsilon: 0.416628640010327
steps: 278   lr: 1.6000000000000003e-05   evaluation reward: 6.56
episode: 1690   score: 6.0   memory length: 394988   epsilon:

0.41592178001032254 steps: 357 lr: 1.6000000000000003e-05 evaluation reward: 6.59
episode: 1691 score: 2.0 memory length: 395207 epsilon: 0.4154881600103198 steps: 219 lr: 1.6000000000000003e-05 evaluation reward: 6.57
episode: 1692 score: 5.0 memory length: 395511 epsilon: 0.414886240010316 steps: 304 lr: 1.6000000000000003e-05 evaluation reward: 6.59
episode: 1693 score: 7.0 memory length: 395879 epsilon: 0.4141576000103114 steps: 368 lr: 1.6000000000000003e-05 evaluation reward: 6.54
episode: 1694 score: 4.0 memory length: 396138 epsilon: 0.41364478001030813 steps: 259 lr: 1.6000000000000003e-05 evaluation reward: 6.5
episode: 1695 score: 10.0 memory length: 396653 epsilon: 0.4126250800103017 steps: 515 lr: 1.6000000000000003e-05 evaluation reward: 6.57
episode: 1696 score: 7.0 memory length: 397041 epsilon: 0.4118568400102968 steps: 388 lr: 1.6000000000000003e-05 evaluation reward: 6.56
episode: 1697 score: 3.0 memory length: 397254 epsilon: 0.41143510001029415 steps: 213 lr: 1.6000000000000003e-05 evaluation reward: 6.5
episode: 1698 score: 1.0 memory length: 397405 epsilon: 0.41113612001029226 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 6.45
episode: 1699 score: 5.0 memory length: 397693 epsilon: 0.41056588001028865 steps: 288 lr: 1.6000000000000003e-05 evaluation reward: 6.46
episode: 1700 score: 3.0 memory length: 397919 epsilon: 0.4101184000102858 steps: 226 lr: 1.6000000000000003e-05 evaluation reward: 6.44
episode: 1701 score: 11.0 memory length: 398408 epsilon: 0.4091501800102797 steps: 489 lr: 1.6000000000000003e-05 evaluation reward: 6.5
episode: 1702 score: 6.0 memory length: 398804 epsilon: 0.40836610001027474 steps: 396 lr: 1.6000000000000003e-05 evaluation reward: 6.44
episode: 1703 score: 3.0 memory length: 399017 epsilon: 0.40794436001027207 steps: 213 lr: 1.6000000000000003e-05 evaluation reward: 6.41
episode: 1704 score: 5.0 memory length: 399358 epsilon: 0.4072691800102678 steps: 341 lr: 1.6000000000000003e-05 evaluation reward: 6.41
episode: 1705 score: 3.0 memory length: 399589 epsilon: 0.4068118000102649 steps: 231 lr: 1.6000000000000003e-05 evaluation reward: 6.37
episode: 1706 score: 5.0 memory length: 399878 epsilon: 0.4062395800102613 steps: 289 lr: 1.6000000000000003e-05 evaluation reward: 6.4
episode: 1707 score: 10.0 memory length: 400420 epsilon: 0.4051664200102545 steps: 542 lr: 6.400000000000001e-06 evaluation reward: 6.41
episode: 1708 score: 10.0 memory length: 400964 epsilon: 0.4040893000102477 steps: 544 lr: 6.400000000000001e-06 evaluation reward: 6.43

episode: 1709   score: 7.0   memory length: 401382   epsilon:
0.40326166001024244   steps: 418   lr: 6.400000000000001e-06   evaluation
reward: 6.44
episode: 1710   score: 3.0   memory length: 401613   epsilon:
0.40280428001023955   steps: 231   lr: 6.400000000000001e-06   evaluation
reward: 6.4
episode: 1711   score: 7.0   memory length: 402021   epsilon:
0.40199644001023443   steps: 408   lr: 6.400000000000001e-06   evaluation
reward: 6.4
episode: 1712   score: 6.0   memory length: 402395   epsilon:
0.40125592001022975   steps: 374   lr: 6.400000000000001e-06   evaluation
reward: 6.38
episode: 1713   score: 5.0   memory length: 402684   epsilon:
0.40068370001022613   steps: 289   lr: 6.400000000000001e-06   evaluation
reward: 6.37
episode: 1714   score: 14.0   memory length: 403234   epsilon:
0.39959470001021924   steps: 550   lr: 6.400000000000001e-06   evaluation
reward: 6.47
episode: 1715   score: 9.0   memory length: 403654   epsilon: 0.398763100010214
steps: 420   lr: 6.400000000000001e-06   evaluation reward: 6.48
episode: 1716   score: 9.0   memory length: 404146   epsilon: 0.3977889400102078
steps: 492   lr: 6.400000000000001e-06   evaluation reward: 6.5
episode: 1717   score: 5.0   memory length: 404457   epsilon: 0.3971731600102039
steps: 311   lr: 6.400000000000001e-06   evaluation reward: 6.51
episode: 1718   score: 7.0   memory length: 404823   epsilon:
0.39644848001019933   steps: 366   lr: 6.400000000000001e-06   evaluation
reward: 6.55
episode: 1719   score: 7.0   memory length: 405224   epsilon: 0.3956545000101943
steps: 401   lr: 6.400000000000001e-06   evaluation reward: 6.48
episode: 1720   score: 4.0   memory length: 405486   epsilon:
0.39513574001019103   steps: 262   lr: 6.400000000000001e-06   evaluation
reward: 6.51
episode: 1721   score: 6.0   memory length: 405808   epsilon: 0.394498180010187
steps: 322   lr: 6.400000000000001e-06   evaluation reward: 6.48
episode: 1722   score: 14.0   memory length: 406471   epsilon:
0.3931854400101787   steps: 663   lr: 6.400000000000001e-06   evaluation
reward: 6.56
episode: 1723   score: 7.0   memory length: 406895   epsilon: 0.3923459200101734
steps: 424   lr: 6.400000000000001e-06   evaluation reward: 6.54
episode: 1724   score: 9.0   memory length: 407394   epsilon: 0.3913579000101671
steps: 499   lr: 6.400000000000001e-06   evaluation reward: 6.57
episode: 1725   score: 4.0   memory length: 407649   epsilon:
0.39085300001016393   steps: 255   lr: 6.400000000000001e-06   evaluation
reward: 6.52
episode: 1726   score: 6.0   memory length: 408019   epsilon: 0.3901204000101593
steps: 370   lr: 6.400000000000001e-06   evaluation reward: 6.48
episode: 1727   score: 11.0   memory length: 408399   epsilon:
0.38936800001015454   steps: 380   lr: 6.400000000000001e-06   evaluation

```
reward: 6.53
episode: 1728   score: 6.0   memory length: 408736   epsilon: 0.3887007400101503
steps: 337    lr: 6.400000000000001e-06     evaluation reward: 6.53
episode: 1729   score: 7.0   memory length: 409116   epsilon:
0.38794834001014555    steps: 380    lr: 6.400000000000001e-06     evaluation
reward: 6.53
episode: 1730   score: 4.0   memory length: 409375   epsilon: 0.3874355200101423
steps: 259    lr: 6.400000000000001e-06     evaluation reward: 6.53
episode: 1731   score: 8.0   memory length: 409796   epsilon:
0.38660194001013704    steps: 421    lr: 6.400000000000001e-06     evaluation
reward: 6.54
episode: 1732   score: 6.0   memory length: 410128   epsilon: 0.3859445800101329
steps: 332    lr: 6.400000000000001e-06     evaluation reward: 6.5
episode: 1733   score: 4.0   memory length: 410388   epsilon: 0.3854297800101296
steps: 260    lr: 6.400000000000001e-06     evaluation reward: 6.5
episode: 1734   score: 10.0   memory length: 410909   epsilon:
0.3843982000101231    steps: 521    lr: 6.400000000000001e-06     evaluation
reward: 6.51
episode: 1735   score: 9.0   memory length: 411364   epsilon: 0.3834973000101174
steps: 455    lr: 6.400000000000001e-06     evaluation reward: 6.57
episode: 1736   score: 6.0   memory length: 411705   epsilon: 0.3828221200101131
steps: 341    lr: 6.400000000000001e-06     evaluation reward: 6.6
episode: 1737   score: 8.0   memory length: 412129   epsilon: 0.3819826000101078
steps: 424    lr: 6.400000000000001e-06     evaluation reward: 6.61
episode: 1738   score: 3.0   memory length: 412358   epsilon:
0.38152918001010494    steps: 229    lr: 6.400000000000001e-06     evaluation
reward: 6.58
episode: 1739   score: 13.0   memory length: 412870   epsilon:
0.3805154200100985    steps: 512    lr: 6.400000000000001e-06     evaluation
reward: 6.64
episode: 1740   score: 6.0   memory length: 413208   epsilon: 0.3798461800100943
steps: 338    lr: 6.400000000000001e-06     evaluation reward: 6.64
episode: 1741   score: 5.0   memory length: 413533   epsilon: 0.3792026800100902
steps: 325    lr: 6.400000000000001e-06     evaluation reward: 6.63
episode: 1742   score: 11.0   memory length: 414106   epsilon:
0.37806814001008304    steps: 573    lr: 6.400000000000001e-06     evaluation
reward: 6.62
episode: 1743   score: 8.0   memory length: 414519   epsilon:
0.37725040001007787    steps: 413    lr: 6.400000000000001e-06     evaluation
reward: 6.67
episode: 1744   score: 9.0   memory length: 415023   epsilon:
0.37625248001007155    steps: 504    lr: 6.400000000000001e-06     evaluation
reward: 6.69
episode: 1745   score: 3.0   memory length: 415253   epsilon: 0.3757970800100687
steps: 230    lr: 6.400000000000001e-06     evaluation reward: 6.67
episode: 1746   score: 8.0   memory length: 415674   epsilon: 0.3749635000100634
steps: 421    lr: 6.400000000000001e-06     evaluation reward: 6.64
episode: 1747   score: 7.0   memory length: 416100   epsilon:
```

0.37412002001005806 steps: 426 lr: 6.400000000000001e-06 evaluation reward: 6.66
episode: 1748 score: 8.0 memory length: 416524 epsilon: 0.37328050001005275 steps: 424 lr: 6.400000000000001e-06 evaluation reward: 6.61
episode: 1749 score: 8.0 memory length: 416925 epsilon: 0.37248652001004773 steps: 401 lr: 6.400000000000001e-06 evaluation reward: 6.62
episode: 1750 score: 12.0 memory length: 417552 epsilon: 0.3712450600100399 steps: 627 lr: 6.400000000000001e-06 evaluation reward: 6.68
episode: 1751 score: 5.0 memory length: 417881 epsilon: 0.37059364001003575 steps: 329 lr: 6.400000000000001e-06 evaluation reward: 6.64
episode: 1752 score: 7.0 memory length: 418228 epsilon: 0.3699065800100314 steps: 347 lr: 6.400000000000001e-06 evaluation reward: 6.64
episode: 1753 score: 4.0 memory length: 418525 epsilon: 0.3693185200100277 steps: 297 lr: 6.400000000000001e-06 evaluation reward: 6.64
episode: 1754 score: 8.0 memory length: 418943 epsilon: 0.36849088001002245 steps: 418 lr: 6.400000000000001e-06 evaluation reward: 6.68
episode: 1755 score: 8.0 memory length: 419395 epsilon: 0.3675959200100168 steps: 452 lr: 6.400000000000001e-06 evaluation reward: 6.66
episode: 1756 score: 5.0 memory length: 419707 epsilon: 0.3669781600100129 steps: 312 lr: 6.400000000000001e-06 evaluation reward: 6.65
episode: 1757 score: 6.0 memory length: 420033 epsilon: 0.3663326800100088 steps: 326 lr: 6.400000000000001e-06 evaluation reward: 6.61
episode: 1758 score: 4.0 memory length: 420295 epsilon: 0.3658139200100055 steps: 262 lr: 6.400000000000001e-06 evaluation reward: 6.62
episode: 1759 score: 14.0 memory length: 420805 epsilon: 0.3648041200099991 steps: 510 lr: 6.400000000000001e-06 evaluation reward: 6.66
episode: 1760 score: 9.0 memory length: 421257 epsilon: 0.36390916000999346 steps: 452 lr: 6.400000000000001e-06 evaluation reward: 6.65
episode: 1761 score: 10.0 memory length: 421766 epsilon: 0.3629013400099871 steps: 509 lr: 6.400000000000001e-06 evaluation reward: 6.71
episode: 1762 score: 8.0 memory length: 422206 epsilon: 0.36203014000998157 steps: 440 lr: 6.400000000000001e-06 evaluation reward: 6.72
episode: 1763 score: 8.0 memory length: 422651 epsilon: 0.361149040009976 steps: 445 lr: 6.400000000000001e-06 evaluation reward: 6.73
episode: 1764 score: 4.0 memory length: 422925 epsilon: 0.36060652000997256 steps: 274 lr: 6.400000000000001e-06 evaluation reward: 6.74
episode: 1765 score: 3.0 memory length: 423156 epsilon: 0.36014914000996967 steps: 231 lr: 6.400000000000001e-06 evaluation

reward: 6.72
episode: 1766   score: 5.0   memory length: 423467   epsilon: 0.3595333600099658
steps: 311   lr: 6.400000000000001e-06   evaluation reward: 6.73
episode: 1767   score: 7.0   memory length: 423872   epsilon: 0.3587314600099607
steps: 405   lr: 6.400000000000001e-06   evaluation reward: 6.74
episode: 1768   score: 4.0   memory length: 424134   epsilon: 0.3582127000099574
steps: 262   lr: 6.400000000000001e-06   evaluation reward: 6.71
episode: 1769   score: 9.0   memory length: 424588   epsilon:
0.35731378000995173   steps: 454   lr: 6.400000000000001e-06   evaluation
reward: 6.71
episode: 1770   score: 4.0   memory length: 424850   epsilon:
0.35679502000994845   steps: 262   lr: 6.400000000000001e-06   evaluation
reward: 6.7
episode: 1771   score: 10.0   memory length: 425351   epsilon:
0.3558030400099422   steps: 501   lr: 6.400000000000001e-06   evaluation
reward: 6.69
episode: 1772   score: 14.0   memory length: 425871   epsilon:
0.35477344000993566   steps: 520   lr: 6.400000000000001e-06   evaluation
reward: 6.76
episode: 1773   score: 9.0   memory length: 426329   epsilon: 0.3538666000099299
steps: 458   lr: 6.400000000000001e-06   evaluation reward: 6.77
episode: 1774   score: 9.0   memory length: 426797   epsilon:
0.35293996000992406   steps: 468   lr: 6.400000000000001e-06   evaluation
reward: 6.8
episode: 1775   score: 9.0   memory length: 427228   epsilon:
0.35208658000991866   steps: 431   lr: 6.400000000000001e-06   evaluation
reward: 6.83
episode: 1776   score: 6.0   memory length: 427602   epsilon: 0.351346060009914
steps: 374   lr: 6.400000000000001e-06   evaluation reward: 6.82
episode: 1777   score: 8.0   memory length: 428022   epsilon: 0.3505144600099087
steps: 420   lr: 6.400000000000001e-06   evaluation reward: 6.83
episode: 1778   score: 6.0   memory length: 428374   epsilon: 0.3498175000099043
steps: 352   lr: 6.400000000000001e-06   evaluation reward: 6.86
episode: 1779   score: 5.0   memory length: 428683   epsilon:
0.34920568000990043   steps: 309   lr: 6.400000000000001e-06   evaluation
reward: 6.84
episode: 1780   score: 8.0   memory length: 429157   epsilon: 0.3482671600098945
steps: 474   lr: 6.400000000000001e-06   evaluation reward: 6.88
episode: 1781   score: 10.0   memory length: 429513   epsilon:
0.34756228000989003   steps: 356   lr: 6.400000000000001e-06   evaluation
reward: 6.91
episode: 1782   score: 8.0   memory length: 429895   epsilon:
0.34680592000988525   steps: 382   lr: 6.400000000000001e-06   evaluation
reward: 6.92
episode: 1783   score: 4.0   memory length: 430137   epsilon: 0.3463267600098822
steps: 242   lr: 6.400000000000001e-06   evaluation reward: 6.93
episode: 1784   score: 10.0   memory length: 430707   epsilon:
0.3451981600098751   steps: 570   lr: 6.400000000000001e-06   evaluation

reward: 6.99
episode: 1785   score: 6.0   memory length: 431060   epsilon:
0.34449922000987065   steps: 353   lr: 6.400000000000001e-06   evaluation
reward: 6.98
episode: 1786   score: 12.0   memory length: 431507   epsilon:
0.34361416000986506   steps: 447   lr: 6.400000000000001e-06   evaluation
reward: 7.01
episode: 1787   score: 7.0   memory length: 431934   epsilon: 0.3427687000098597
steps: 427   lr: 6.400000000000001e-06   evaluation reward: 6.99
episode: 1788   score: 11.0   memory length: 432508   epsilon:
0.3416321800098525   steps: 574   lr: 6.400000000000001e-06   evaluation
reward: 7.03
episode: 1789   score: 10.0   memory length: 433032   epsilon:
0.34059466000984595   steps: 524   lr: 6.400000000000001e-06   evaluation
reward: 7.09
episode: 1790   score: 15.0   memory length: 433717   epsilon:
0.33923836000983737   steps: 685   lr: 6.400000000000001e-06   evaluation
reward: 7.18
episode: 1791   score: 4.0   memory length: 433958   epsilon:
0.33876118000983435   steps: 241   lr: 6.400000000000001e-06   evaluation
reward: 7.2
episode: 1792   score: 3.0   memory length: 434168   epsilon: 0.3383453800098317
steps: 210   lr: 6.400000000000001e-06   evaluation reward: 7.18
episode: 1793   score: 3.0   memory length: 434380   epsilon:
0.33792562000982906   steps: 212   lr: 6.400000000000001e-06   evaluation
reward: 7.14
episode: 1794   score: 11.0   memory length: 434868   epsilon:
0.33695938000982295   steps: 488   lr: 6.400000000000001e-06   evaluation
reward: 7.21
episode: 1795   score: 13.0   memory length: 435468   epsilon:
0.33577138000981543   steps: 600   lr: 6.400000000000001e-06   evaluation
reward: 7.24
episode: 1796   score: 8.0   memory length: 435903   epsilon: 0.33491008000981
steps: 435   lr: 6.400000000000001e-06   evaluation reward: 7.25
episode: 1797   score: 3.0   memory length: 436134   epsilon: 0.3344527000098071
steps: 231   lr: 6.400000000000001e-06   evaluation reward: 7.25
episode: 1798   score: 6.0   memory length: 436463   epsilon:
0.33380128000980297   steps: 329   lr: 6.400000000000001e-06   evaluation
reward: 7.3
episode: 1799   score: 11.0   memory length: 436965   epsilon:
0.3328073200097967   steps: 502   lr: 6.400000000000001e-06   evaluation
reward: 7.36
episode: 1800   score: 8.0   memory length: 437446   epsilon:
0.33185494000979066   steps: 481   lr: 6.400000000000001e-06   evaluation
reward: 7.41
episode: 1801   score: 12.0   memory length: 437896   epsilon: 0.330963940009785
steps: 450   lr: 6.400000000000001e-06   evaluation reward: 7.42
episode: 1802   score: 3.0   memory length: 438109   epsilon:

0.33054220000978235   steps: 213   lr: 6.400000000000001e-06   evaluation
reward: 7.39
episode: 1803   score: 3.0   memory length: 438340   epsilon:
0.33008482000977946   steps: 231   lr: 6.400000000000001e-06   evaluation
reward: 7.39
episode: 1804   score: 5.0   memory length: 438630   epsilon: 0.3295106200097758
steps: 290   lr: 6.400000000000001e-06   evaluation reward: 7.39
episode: 1805   score: 14.0   memory length: 439270   epsilon:
0.3282434200097678   steps: 640   lr: 6.400000000000001e-06   evaluation
reward: 7.5
episode: 1806   score: 5.0   memory length: 439562   epsilon:
0.32766526000976415   steps: 292   lr: 6.400000000000001e-06   evaluation
reward: 7.5
episode: 1807   score: 6.0   memory length: 439933   epsilon: 0.3269306800097595
steps: 371   lr: 6.400000000000001e-06   evaluation reward: 7.46
episode: 1808   score: 3.0   memory length: 440146   epsilon:
0.32650894000975683   steps: 213   lr: 6.400000000000001e-06   evaluation
reward: 7.39
episode: 1809   score: 12.0   memory length: 440645   epsilon:
0.3255209200097506   steps: 499   lr: 6.400000000000001e-06   evaluation
reward: 7.44
episode: 1810   score: 6.0   memory length: 440981   epsilon:
0.32485564000974637   steps: 336   lr: 6.400000000000001e-06   evaluation
reward: 7.47
episode: 1811   score: 3.0   memory length: 441211   epsilon: 0.3244002400097435
steps: 230   lr: 6.400000000000001e-06   evaluation reward: 7.43
episode: 1812   score: 7.0   memory length: 441598   epsilon:
0.32363398000973864   steps: 387   lr: 6.400000000000001e-06   evaluation
reward: 7.44
episode: 1813   score: 6.0   memory length: 441971   epsilon:
0.32289544000973397   steps: 373   lr: 6.400000000000001e-06   evaluation
reward: 7.45
episode: 1814   score: 6.0   memory length: 442293   epsilon:
0.32225788000972994   steps: 322   lr: 6.400000000000001e-06   evaluation
reward: 7.37
episode: 1815   score: 4.0   memory length: 442552   epsilon: 0.3217450600097267
steps: 259   lr: 6.400000000000001e-06   evaluation reward: 7.32
episode: 1816   score: 3.0   memory length: 442762   epsilon:
0.32132926000972406   steps: 210   lr: 6.400000000000001e-06   evaluation
reward: 7.26
episode: 1817   score: 4.0   memory length: 443025   epsilon:
0.32080852000972077   steps: 263   lr: 6.400000000000001e-06   evaluation
reward: 7.25
episode: 1818   score: 8.0   memory length: 443440   epsilon:
0.31998682000971557   steps: 415   lr: 6.400000000000001e-06   evaluation
reward: 7.26
episode: 1819   score: 6.0   memory length: 443798   epsilon: 0.3192779800097111
steps: 358   lr: 6.400000000000001e-06   evaluation reward: 7.25

episode: 1820 score: 7.0 memory length: 444170 epsilon: 0.3185414200097064
steps: 372 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1821 score: 6.0 memory length: 444500 epsilon: 0.3178880200097023
steps: 330 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1822 score: 8.0 memory length: 444918 epsilon:
0.31706038000969705 steps: 418 lr: 6.400000000000001e-06 evaluation
reward: 7.22
episode: 1823 score: 8.0 memory length: 445393 epsilon: 0.3161198800096911
steps: 475 lr: 6.400000000000001e-06 evaluation reward: 7.23
episode: 1824 score: 3.0 memory length: 445604 epsilon:
0.31570210000968846 steps: 211 lr: 6.400000000000001e-06 evaluation
reward: 7.17
episode: 1825 score: 9.0 memory length: 446091 epsilon:
0.31473784000968236 steps: 487 lr: 6.400000000000001e-06 evaluation
reward: 7.22
episode: 1826 score: 3.0 memory length: 446304 epsilon: 0.3143161000096797
steps: 213 lr: 6.400000000000001e-06 evaluation reward: 7.19
episode: 1827 score: 8.0 memory length: 446706 epsilon:
0.31352014000967465 steps: 402 lr: 6.400000000000001e-06 evaluation
reward: 7.16
episode: 1828 score: 4.0 memory length: 446968 epsilon:
0.31300138000967137 steps: 262 lr: 6.400000000000001e-06 evaluation
reward: 7.14
episode: 1829 score: 4.0 memory length: 447227 epsilon: 0.3124885600096681
steps: 259 lr: 6.400000000000001e-06 evaluation reward: 7.11
episode: 1830 score: 6.0 memory length: 447602 epsilon:
0.31174606000966343 steps: 375 lr: 6.400000000000001e-06 evaluation
reward: 7.13
episode: 1831 score: 15.0 memory length: 448201 epsilon:
0.3105600400096559 steps: 599 lr: 6.400000000000001e-06 evaluation
reward: 7.2
episode: 1832 score: 5.0 memory length: 448510 epsilon:
0.30994822000965205 steps: 309 lr: 6.400000000000001e-06 evaluation
reward: 7.19
episode: 1833 score: 5.0 memory length: 448817 epsilon: 0.3093403600096482
steps: 307 lr: 6.400000000000001e-06 evaluation reward: 7.2
episode: 1834 score: 7.0 memory length: 449186 epsilon: 0.3086097400096436
steps: 369 lr: 6.400000000000001e-06 evaluation reward: 7.17
episode: 1835 score: 3.0 memory length: 449399 epsilon: 0.3081880000096409
steps: 213 lr: 6.400000000000001e-06 evaluation reward: 7.11
episode: 1836 score: 14.0 memory length: 449928 epsilon:
0.3071405800096343 steps: 529 lr: 6.400000000000001e-06 evaluation
reward: 7.19
episode: 1837 score: 4.0 memory length: 450191 epsilon: 0.306619840009631
steps: 263 lr: 6.400000000000001e-06 evaluation reward: 7.15
episode: 1838 score: 8.0 memory length: 450627 epsilon:
0.30575656000962553 steps: 436 lr: 6.400000000000001e-06 evaluation
reward: 7.2

episode: 1839   score: 4.0   memory length: 450889   epsilon:
0.30523780000962225   steps: 262   lr: 6.400000000000001e-06   evaluation
reward: 7.11
episode: 1840   score: 5.0   memory length: 451178   epsilon:
0.30466558000961863   steps: 289   lr: 6.400000000000001e-06   evaluation
reward: 7.1
episode: 1841   score: 7.0   memory length: 451581   epsilon: 0.3038676400096136
steps: 403   lr: 6.400000000000001e-06   evaluation reward: 7.12
episode: 1842   score: 3.0   memory length: 451794   epsilon: 0.3034459000096109
steps: 213   lr: 6.400000000000001e-06   evaluation reward: 7.04
episode: 1843   score: 7.0   memory length: 452172   epsilon: 0.3026974600096062
steps: 378   lr: 6.400000000000001e-06   evaluation reward: 7.03
episode: 1844   score: 17.0   memory length: 452855   epsilon:
0.3013451200095976   steps: 683   lr: 6.400000000000001e-06   evaluation
reward: 7.11
episode: 1845   score: 13.0   memory length: 453369   epsilon:
0.3003274000095912   steps: 514   lr: 6.400000000000001e-06   evaluation
reward: 7.21
episode: 1846   score: 7.0   memory length: 453796   epsilon:
0.29948194000958583   steps: 427   lr: 6.400000000000001e-06   evaluation
reward: 7.2
episode: 1847   score: 6.0   memory length: 454114   epsilon:
0.29885230000958185   steps: 318   lr: 6.400000000000001e-06   evaluation
reward: 7.19
episode: 1848   score: 3.0   memory length: 454327   epsilon: 0.2984305600095792
steps: 213   lr: 6.400000000000001e-06   evaluation reward: 7.14
episode: 1849   score: 6.0   memory length: 454651   epsilon: 0.2977890400095751
steps: 324   lr: 6.400000000000001e-06   evaluation reward: 7.12
episode: 1850   score: 7.0   memory length: 455068   epsilon: 0.2969633800095699
steps: 417   lr: 6.400000000000001e-06   evaluation reward: 7.07
episode: 1851   score: 6.0   memory length: 455394   epsilon: 0.2963179000095658
steps: 326   lr: 6.400000000000001e-06   evaluation reward: 7.08
episode: 1852   score: 5.0   memory length: 455688   epsilon:
0.29573578000956213   steps: 294   lr: 6.400000000000001e-06   evaluation
reward: 7.06
episode: 1853   score: 8.0   memory length: 456138   epsilon: 0.2948447800095565
steps: 450   lr: 6.400000000000001e-06   evaluation reward: 7.1
episode: 1854   score: 9.0   memory length: 456612   epsilon:
0.29390626000955056   steps: 474   lr: 6.400000000000001e-06   evaluation
reward: 7.11
episode: 1855   score: 5.0   memory length: 456921   epsilon: 0.2932944400095467
steps: 309   lr: 6.400000000000001e-06   evaluation reward: 7.08
episode: 1856   score: 6.0   memory length: 457250   epsilon:
0.29264302000954256   steps: 329   lr: 6.400000000000001e-06   evaluation
reward: 7.09
episode: 1857   score: 5.0   memory length: 457575   epsilon: 0.2919995200095385
steps: 325   lr: 6.400000000000001e-06   evaluation reward: 7.08
episode: 1858   score: 4.0   memory length: 457833   epsilon:

0.29148868000953526   steps: 258   lr: 6.400000000000001e-06   evaluation
reward: 7.08
episode: 1859   score: 19.0   memory length: 458663   epsilon:
0.28984528000952486   steps: 830   lr: 6.400000000000001e-06   evaluation
reward: 7.13
episode: 1860   score: 9.0   memory length: 459146   epsilon: 0.2888889400095188
steps: 483   lr: 6.400000000000001e-06   evaluation reward: 7.13
episode: 1861   score: 7.0   memory length: 459541   epsilon:
0.28810684000951386   steps: 395   lr: 6.400000000000001e-06   evaluation
reward: 7.1
episode: 1862   score: 11.0   memory length: 459971   epsilon:
0.2872554400095085   steps: 430   lr: 6.400000000000001e-06   evaluation
reward: 7.13
episode: 1863   score: 7.0   memory length: 460384   epsilon: 0.2864377000095033
steps: 413   lr: 6.400000000000001e-06   evaluation reward: 7.12
episode: 1864   score: 10.0   memory length: 460876   epsilon:
0.28546354000949714   steps: 492   lr: 6.400000000000001e-06   evaluation
reward: 7.18
episode: 1865   score: 8.0   memory length: 461268   epsilon:
0.28468738000949223   steps: 392   lr: 6.400000000000001e-06   evaluation
reward: 7.23
episode: 1866   score: 7.0   memory length: 461620   epsilon: 0.2839904200094878
steps: 352   lr: 6.400000000000001e-06   evaluation reward: 7.25
episode: 1867   score: 11.0   memory length: 462126   epsilon:
0.2829885400094815   steps: 506   lr: 6.400000000000001e-06   evaluation
reward: 7.29
episode: 1868   score: 6.0   memory length: 462504   epsilon:
0.28224010000947675   steps: 378   lr: 6.400000000000001e-06   evaluation
reward: 7.31
episode: 1869   score: 4.0   memory length: 462806   epsilon:
0.28164214000947296   steps: 302   lr: 6.400000000000001e-06   evaluation
reward: 7.26
episode: 1870   score: 8.0   memory length: 463212   epsilon: 0.2808382600094679
steps: 406   lr: 6.400000000000001e-06   evaluation reward: 7.3
episode: 1871   score: 12.0   memory length: 463774   epsilon:
0.27972550000946084   steps: 562   lr: 6.400000000000001e-06   evaluation
reward: 7.32
episode: 1872   score: 11.0   memory length: 464335   epsilon:
0.2786147200094538   steps: 561   lr: 6.400000000000001e-06   evaluation
reward: 7.29
episode: 1873   score: 5.0   memory length: 464662   epsilon: 0.2779672600094497
steps: 327   lr: 6.400000000000001e-06   evaluation reward: 7.25
episode: 1874   score: 3.0   memory length: 464875   epsilon:
0.27754552000944704   steps: 213   lr: 6.400000000000001e-06   evaluation
reward: 7.19
episode: 1875   score: 7.0   memory length: 465279   epsilon: 0.276745600009442
steps: 404   lr: 6.400000000000001e-06   evaluation reward: 7.17
episode: 1876   score: 10.0   memory length: 465689   epsilon:

0.27593380000943685 steps: 410 lr: 6.400000000000001e-06 evaluation reward: 7.21
episode: 1877 score: 4.0 memory length: 465969 epsilon: 0.27537940000943334 steps: 280 lr: 6.400000000000001e-06 evaluation reward: 7.17
episode: 1878 score: 15.0 memory length: 466526 epsilon: 0.27427654000942636 steps: 557 lr: 6.400000000000001e-06 evaluation reward: 7.26
episode: 1879 score: 10.0 memory length: 467012 epsilon: 0.2733142600094203 steps: 486 lr: 6.400000000000001e-06 evaluation reward: 7.31
episode: 1880 score: 5.0 memory length: 467311 epsilon: 0.27272224000941653 steps: 299 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1881 score: 5.0 memory length: 467618 epsilon: 0.2721143800094127 steps: 307 lr: 6.400000000000001e-06 evaluation reward: 7.23
episode: 1882 score: 7.0 memory length: 468041 epsilon: 0.2712768400094074 steps: 423 lr: 6.400000000000001e-06 evaluation reward: 7.22
episode: 1883 score: 12.0 memory length: 468525 epsilon: 0.2703185200094013 steps: 484 lr: 6.400000000000001e-06 evaluation reward: 7.3
episode: 1884 score: 8.0 memory length: 468945 epsilon: 0.26948692000939606 steps: 420 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1885 score: 3.0 memory length: 469174 epsilon: 0.2690335000093932 steps: 229 lr: 6.400000000000001e-06 evaluation reward: 7.25
episode: 1886 score: 7.0 memory length: 469604 epsilon: 0.2681821000093878 steps: 430 lr: 6.400000000000001e-06 evaluation reward: 7.2
episode: 1887 score: 4.0 memory length: 469866 epsilon: 0.2676633400093845 steps: 262 lr: 6.400000000000001e-06 evaluation reward: 7.17
episode: 1888 score: 11.0 memory length: 470262 epsilon: 0.26687926000937956 steps: 396 lr: 6.400000000000001e-06 evaluation reward: 7.17
episode: 1889 score: 4.0 memory length: 470520 epsilon: 0.2663684200093763 steps: 258 lr: 6.400000000000001e-06 evaluation reward: 7.11
episode: 1890 score: 7.0 memory length: 470903 epsilon: 0.26561008000937153 steps: 383 lr: 6.400000000000001e-06 evaluation reward: 7.03
episode: 1891 score: 8.0 memory length: 471289 epsilon: 0.2648458000093667 steps: 386 lr: 6.400000000000001e-06 evaluation reward: 7.07
episode: 1892 score: 7.0 memory length: 471732 epsilon: 0.26396866000936114 steps: 443 lr: 6.400000000000001e-06 evaluation reward: 7.11
episode: 1893 score: 9.0 memory length: 472194 epsilon: 0.26305390000935536 steps: 462 lr: 6.400000000000001e-06 evaluation reward: 7.17
episode: 1894 score: 4.0 memory length: 472438 epsilon: 0.2625707800093523 steps: 244 lr: 6.400000000000001e-06 evaluation reward: 7.1

episode: 1895   score: 14.0   memory length: 473028   epsilon:
0.2614025800093449   steps: 590   lr: 6.400000000000001e-06   evaluation
reward: 7.11
episode: 1896   score: 6.0   memory length: 473387   epsilon: 0.2606917600093404
steps: 359   lr: 6.400000000000001e-06   evaluation reward: 7.09
episode: 1897   score: 5.0   memory length: 473678   epsilon:
0.26011558000933677   steps: 291   lr: 6.400000000000001e-06   evaluation
reward: 7.11
episode: 1898   score: 7.0   memory length: 474078   epsilon:
0.25932358000933176   steps: 400   lr: 6.400000000000001e-06   evaluation
reward: 7.12
episode: 1899   score: 6.0   memory length: 474443   epsilon: 0.2586008800093272
steps: 365   lr: 6.400000000000001e-06   evaluation reward: 7.07
episode: 1900   score: 10.0   memory length: 474945   epsilon:
0.2576069200093209   steps: 502   lr: 6.400000000000001e-06   evaluation
reward: 7.09
episode: 1901   score: 8.0   memory length: 475260   epsilon:
0.25698322000931695   steps: 315   lr: 6.400000000000001e-06   evaluation
reward: 7.05
episode: 1902   score: 11.0   memory length: 475800   epsilon:
0.2559140200093102   steps: 540   lr: 6.400000000000001e-06   evaluation
reward: 7.13
episode: 1903   score: 11.0   memory length: 476348   epsilon:
0.2548289800093033   steps: 548   lr: 6.400000000000001e-06   evaluation
reward: 7.21
episode: 1904   score: 3.0   memory length: 476561   epsilon:
0.25440724000930065   steps: 213   lr: 6.400000000000001e-06   evaluation
reward: 7.19
episode: 1905   score: 7.0   memory length: 476903   epsilon:
0.25373008000929637   steps: 342   lr: 6.400000000000001e-06   evaluation
reward: 7.12
episode: 1906   score: 7.0   memory length: 477325   epsilon: 0.2528945200092911
steps: 422   lr: 6.400000000000001e-06   evaluation reward: 7.14
episode: 1907   score: 7.0   memory length: 477684   epsilon: 0.2521837000092866
steps: 359   lr: 6.400000000000001e-06   evaluation reward: 7.15
episode: 1908   score: 10.0   memory length: 478167   epsilon:
0.25122736000928053   steps: 483   lr: 6.400000000000001e-06   evaluation
reward: 7.22
episode: 1909   score: 7.0   memory length: 478574   epsilon:
0.25042150000927543   steps: 407   lr: 6.400000000000001e-06   evaluation
reward: 7.17
episode: 1910   score: 14.0   memory length: 479084   epsilon:
0.24941170000926904   steps: 510   lr: 6.400000000000001e-06   evaluation
reward: 7.25
episode: 1911   score: 3.0   memory length: 479313   epsilon:
0.24895828000926618   steps: 229   lr: 6.400000000000001e-06   evaluation
reward: 7.25
episode: 1912   score: 10.0   memory length: 479796   epsilon:

0.24800194000926012 steps: 483 lr: 6.400000000000001e-06 evaluation reward: 7.28
episode: 1913 score: 5.0 memory length: 480085 epsilon: 0.2474297200092565 steps: 289 lr: 6.400000000000001e-06 evaluation reward: 7.27
episode: 1914 score: 5.0 memory length: 480394 epsilon: 0.24681790000925263 steps: 309 lr: 6.400000000000001e-06 evaluation reward: 7.26
episode: 1915 score: 11.0 memory length: 480852 epsilon: 0.2459110600092469 steps: 458 lr: 6.400000000000001e-06 evaluation reward: 7.33
episode: 1916 score: 15.0 memory length: 481436 epsilon: 0.24475474000923958 steps: 584 lr: 6.400000000000001e-06 evaluation reward: 7.45
episode: 1917 score: 6.0 memory length: 481815 epsilon: 0.24400432000923483 steps: 379 lr: 6.400000000000001e-06 evaluation reward: 7.47
episode: 1918 score: 6.0 memory length: 482137 epsilon: 0.2433667600092308 steps: 322 lr: 6.400000000000001e-06 evaluation reward: 7.45
episode: 1919 score: 6.0 memory length: 482505 epsilon: 0.2426381200092262 steps: 368 lr: 6.400000000000001e-06 evaluation reward: 7.45
episode: 1920 score: 10.0 memory length: 483029 epsilon: 0.24160060000921962 steps: 524 lr: 6.400000000000001e-06 evaluation reward: 7.48
episode: 1921 score: 7.0 memory length: 483417 epsilon: 0.24083236000921476 steps: 388 lr: 6.400000000000001e-06 evaluation reward: 7.49
episode: 1922 score: 17.0 memory length: 484136 epsilon: 0.23940874000920576 steps: 719 lr: 6.400000000000001e-06 evaluation reward: 7.58
episode: 1923 score: 11.0 memory length: 484710 epsilon: 0.23827222000919857 steps: 574 lr: 6.400000000000001e-06 evaluation reward: 7.61
episode: 1924 score: 6.0 memory length: 485063 epsilon: 0.23757328000919414 steps: 353 lr: 6.400000000000001e-06 evaluation reward: 7.64
episode: 1925 score: 16.0 memory length: 485653 epsilon: 0.23640508000918675 steps: 590 lr: 6.400000000000001e-06 evaluation reward: 7.71
episode: 1926 score: 4.0 memory length: 485936 epsilon: 0.2358447400091832 steps: 283 lr: 6.400000000000001e-06 evaluation reward: 7.72
episode: 1927 score: 8.0 memory length: 486380 epsilon: 0.23496562000917764 steps: 444 lr: 6.400000000000001e-06 evaluation reward: 7.72
episode: 1928 score: 8.0 memory length: 486756 epsilon: 0.23422114000917293 steps: 376 lr: 6.400000000000001e-06 evaluation reward: 7.76
episode: 1929 score: 7.0 memory length: 487165 epsilon: 0.2334113200091678 steps: 409 lr: 6.400000000000001e-06 evaluation reward: 7.79

```
episode: 1930   score: 3.0   memory length: 487396   epsilon:
0.23295394000916492    steps: 231    lr: 6.400000000000001e-06     evaluation
reward: 7.76
episode: 1931   score: 4.0   memory length: 487655   epsilon:
0.23244112000916167    steps: 259    lr: 6.400000000000001e-06     evaluation
reward: 7.65
episode: 1932   score: 6.0   memory length: 488000   epsilon:
0.23175802000915735    steps: 345    lr: 6.400000000000001e-06     evaluation
reward: 7.66
episode: 1933   score: 7.0   memory length: 488401   epsilon:
0.23096404000915233    steps: 401    lr: 6.400000000000001e-06     evaluation
reward: 7.68
episode: 1934   score: 15.0   memory length: 488999   epsilon:
0.22978000000914484    steps: 598    lr: 6.400000000000001e-06     evaluation
reward: 7.76
episode: 1935   score: 7.0   memory length: 489403   epsilon:
0.22898008000913977    steps: 404    lr: 6.400000000000001e-06     evaluation
reward: 7.8
episode: 1936   score: 4.0   memory length: 489665   epsilon: 0.2284613200091365
steps: 262    lr: 6.400000000000001e-06     evaluation reward: 7.7
episode: 1937   score: 7.0   memory length: 490073   epsilon:
0.22765348000913138    steps: 408    lr: 6.400000000000001e-06     evaluation
reward: 7.73
episode: 1938   score: 9.0   memory length: 490546   epsilon:
0.22671694000912546    steps: 473    lr: 6.400000000000001e-06     evaluation
reward: 7.74
episode: 1939   score: 10.0   memory length: 491031   epsilon:
0.22575664000911938    steps: 485    lr: 6.400000000000001e-06     evaluation
reward: 7.8
episode: 1940   score: 7.0   memory length: 491404   epsilon: 0.2250181000091147
steps: 373    lr: 6.400000000000001e-06     evaluation reward: 7.82
episode: 1941   score: 12.0   memory length: 491975   epsilon:
0.22388752000910755    steps: 571    lr: 6.400000000000001e-06     evaluation
reward: 7.87
episode: 1942   score: 16.0   memory length: 492563   epsilon:
0.2227232800091002    steps: 588    lr: 6.400000000000001e-06     evaluation
reward: 8.0
episode: 1943   score: 10.0   memory length: 493077   epsilon:
0.22170556000909375    steps: 514    lr: 6.400000000000001e-06     evaluation
reward: 8.03
episode: 1944   score: 10.0   memory length: 493595   epsilon:
0.22067992000908726    steps: 518    lr: 6.400000000000001e-06     evaluation
reward: 7.96
episode: 1945   score: 6.0   memory length: 493936   epsilon: 0.220004740009083
steps: 341    lr: 6.400000000000001e-06     evaluation reward: 7.89
episode: 1946   score: 11.0   memory length: 494472   epsilon:
0.21894346000907627    steps: 536    lr: 6.400000000000001e-06     evaluation
reward: 7.93
```

episode: 1947   score: 8.0   memory length: 494918   epsilon: 0.2180603800090707
steps: 446   lr: 6.400000000000001e-06   evaluation reward: 7.95
episode: 1948   score: 3.0   memory length: 495131   epsilon:
0.21763864000906802   steps: 213   lr: 6.400000000000001e-06   evaluation
reward: 7.95
episode: 1949   score: 5.0   memory length: 495443   epsilon: 0.2170208800090641
steps: 312   lr: 6.400000000000001e-06   evaluation reward: 7.94
episode: 1950   score: 3.0   memory length: 495671   epsilon:
0.21656944000906125   steps: 228   lr: 6.400000000000001e-06   evaluation
reward: 7.9
episode: 1951   score: 8.0   memory length: 496091   epsilon: 0.215737840009056
steps: 420   lr: 6.400000000000001e-06   evaluation reward: 7.92
episode: 1952   score: 6.0   memory length: 496470   epsilon:
0.21498742000905124   steps: 379   lr: 6.400000000000001e-06   evaluation
reward: 7.93
episode: 1953   score: 12.0   memory length: 497029   epsilon:
0.21388060000904424   steps: 559   lr: 6.400000000000001e-06   evaluation
reward: 7.97
episode: 1954   score: 7.0   memory length: 497462   epsilon:
0.21302326000903882   steps: 433   lr: 6.400000000000001e-06   evaluation
reward: 7.95
episode: 1955   score: 13.0   memory length: 498079   epsilon:
0.2118016000090311   steps: 617   lr: 6.400000000000001e-06   evaluation
reward: 8.03
episode: 1956   score: 7.0   memory length: 498455   epsilon:
0.21105712000902638   steps: 376   lr: 6.400000000000001e-06   evaluation
reward: 8.04
episode: 1957   score: 8.0   memory length: 498923   epsilon:
0.21013048000902051   steps: 468   lr: 6.400000000000001e-06   evaluation
reward: 8.07
episode: 1958   score: 8.0   memory length: 499381   epsilon:
0.20922364000901478   steps: 458   lr: 6.400000000000001e-06   evaluation
reward: 8.11
episode: 1959   score: 8.0   memory length: 499817   epsilon:
0.20836036000900932   steps: 436   lr: 6.400000000000001e-06   evaluation
reward: 8.0
episode: 1960   score: 9.0   memory length: 500284   epsilon:
0.20743570000900347   steps: 467   lr: 2.560000000000001e-06   evaluation
reward: 8.0
episode: 1961   score: 13.0   memory length: 500907   epsilon:
0.20620216000899566   steps: 623   lr: 2.560000000000001e-06   evaluation
reward: 8.06
episode: 1962   score: 9.0   memory length: 501357   epsilon:
0.20531116000899002   steps: 450   lr: 2.560000000000001e-06   evaluation
reward: 8.04
episode: 1963   score: 4.0   memory length: 501619   epsilon:
0.20479240000898674   steps: 262   lr: 2.560000000000001e-06   evaluation
reward: 8.01

```
episode: 1964   score: 3.0   memory length: 501830   epsilon: 0.2043746200089841
steps: 211    lr: 2.560000000000001e-06     evaluation reward: 7.94
episode: 1965   score: 3.0   memory length: 502043   epsilon:
0.20395288000898143    steps: 213    lr: 2.560000000000001e-06     evaluation
reward: 7.89
episode: 1966   score: 13.0   memory length: 502579   epsilon:
0.20289160000897472    steps: 536    lr: 2.560000000000001e-06     evaluation
reward: 7.95
episode: 1967   score: 10.0   memory length: 502953   epsilon:
0.20215108000897003    steps: 374    lr: 2.560000000000001e-06     evaluation
reward: 7.94
episode: 1968   score: 7.0   memory length: 503358   epsilon:
0.20134918000896496    steps: 405    lr: 2.560000000000001e-06     evaluation
reward: 7.95
episode: 1969   score: 9.0   memory length: 503848   epsilon:
0.20037898000895882    steps: 490    lr: 2.560000000000001e-06     evaluation
reward: 8.0
episode: 1970   score: 3.0   memory length: 504061   epsilon:
0.19995724000895615    steps: 213    lr: 2.560000000000001e-06     evaluation
reward: 7.95
episode: 1971   score: 9.0   memory length: 504549   epsilon:
0.19899100000895004    steps: 488    lr: 2.560000000000001e-06     evaluation
reward: 7.92
episode: 1972   score: 16.0   memory length: 505055   epsilon:
0.1979891200089437    steps: 506    lr: 2.560000000000001e-06     evaluation
reward: 7.97
episode: 1973   score: 9.0   memory length: 505555   epsilon:
0.19699912000893743    steps: 500    lr: 2.560000000000001e-06     evaluation
reward: 8.01
episode: 1974   score: 7.0   memory length: 505930   epsilon:
0.19625662000893274    steps: 375    lr: 2.560000000000001e-06     evaluation
reward: 8.05
episode: 1975   score: 7.0   memory length: 506345   epsilon:
0.19543492000892754    steps: 415    lr: 2.560000000000001e-06     evaluation
reward: 8.05
episode: 1976   score: 4.0   memory length: 506589   epsilon:
0.19495180000892448    steps: 244    lr: 2.560000000000001e-06     evaluation
reward: 7.99
episode: 1977   score: 14.0   memory length: 507243   epsilon:
0.1936568800089163    steps: 654    lr: 2.560000000000001e-06     evaluation
reward: 8.09
episode: 1978   score: 10.0   memory length: 507741   epsilon:
0.19267084000891005    steps: 498    lr: 2.560000000000001e-06     evaluation
reward: 8.04
episode: 1979   score: 8.0   memory length: 508176   epsilon: 0.1918095400089046
steps: 435    lr: 2.560000000000001e-06     evaluation reward: 8.02
episode: 1980   score: 4.0   memory length: 508439   epsilon: 0.1912888000089013
steps: 263    lr: 2.560000000000001e-06     evaluation reward: 8.01
```

```
episode: 1981   score: 6.0   memory length: 508792   epsilon:
0.19058986000889688    steps: 353    lr: 2.560000000000001e-06     evaluation
reward: 8.02
episode: 1982   score: 12.0   memory length: 509273   epsilon:
0.18963748000889086    steps: 481    lr: 2.560000000000001e-06     evaluation
reward: 8.07
episode: 1983   score: 12.0   memory length: 509882   epsilon:
0.18843166000888323    steps: 609    lr: 2.560000000000001e-06     evaluation
reward: 8.07
episode: 1984   score: 6.0   memory length: 510196   epsilon: 0.1878099400088793
steps: 314    lr: 2.560000000000001e-06     evaluation reward: 8.05
episode: 1985   score: 9.0   memory length: 510652   epsilon:
0.18690706000887358    steps: 456    lr: 2.560000000000001e-06     evaluation
reward: 8.11
episode: 1986   score: 10.0   memory length: 511136   epsilon:
0.18594874000886752    steps: 484    lr: 2.560000000000001e-06     evaluation
reward: 8.14
episode: 1987   score: 4.0   memory length: 511398   epsilon:
0.18542998000886424    steps: 262    lr: 2.560000000000001e-06     evaluation
reward: 8.14
episode: 1988   score: 15.0   memory length: 512004   epsilon:
0.18423010000885665    steps: 606    lr: 2.560000000000001e-06     evaluation
reward: 8.18
episode: 1989   score: 5.0   memory length: 512313   epsilon:
0.18361828000885277    steps: 309    lr: 2.560000000000001e-06     evaluation
reward: 8.19
episode: 1990   score: 12.0   memory length: 512846   epsilon:
0.1825629400088461    steps: 533    lr: 2.560000000000001e-06     evaluation
reward: 8.24
episode: 1991   score: 7.0   memory length: 513174   epsilon: 0.181913500008842
steps: 328    lr: 2.560000000000001e-06     evaluation reward: 8.23
episode: 1992   score: 14.0   memory length: 513761   epsilon:
0.18075124000883463    steps: 587    lr: 2.560000000000001e-06     evaluation
reward: 8.3
episode: 1993   score: 8.0   memory length: 514182   epsilon:
0.17991766000882936    steps: 421    lr: 2.560000000000001e-06     evaluation
reward: 8.29
episode: 1994   score: 9.0   memory length: 514691   epsilon:
0.17890984000882298    steps: 509    lr: 2.560000000000001e-06     evaluation
reward: 8.34
episode: 1995   score: 9.0   memory length: 515151   epsilon:
0.17799904000881722    steps: 460    lr: 2.560000000000001e-06     evaluation
reward: 8.29
episode: 1996   score: 11.0   memory length: 515520   epsilon:
0.1772684200088126    steps: 369    lr: 2.560000000000001e-06     evaluation
reward: 8.34
episode: 1997   score: 11.0   memory length: 516081   epsilon:
0.17615764000880557    steps: 561    lr: 2.560000000000001e-06     evaluation
```

reward: 8.4
episode: 1998   score: 11.0   memory length: 516581   epsilon:
0.1751676400087993   steps: 500   lr: 2.560000000000001e-06   evaluation
reward: 8.44
episode: 1999   score: 4.0   memory length: 516841   epsilon:
0.17465284000879605   steps: 260   lr: 2.560000000000001e-06   evaluation
reward: 8.42
episode: 2000   score: 6.0   memory length: 517181   epsilon: 0.1739796400087918
steps: 340   lr: 2.560000000000001e-06   evaluation reward: 8.38
episode: 2001   score: 7.0   memory length: 517552   epsilon:
0.17324506000878714   steps: 371   lr: 2.560000000000001e-06   evaluation
reward: 8.37
episode: 2002   score: 7.0   memory length: 517944   epsilon:
0.17246890000878223   steps: 392   lr: 2.560000000000001e-06   evaluation
reward: 8.33
episode: 2003   score: 7.0   memory length: 518300   epsilon:
0.17176402000877777   steps: 356   lr: 2.560000000000001e-06   evaluation
reward: 8.29
episode: 2004   score: 16.0   memory length: 518952   epsilon:
0.1704730600087696   steps: 652   lr: 2.560000000000001e-06   evaluation
reward: 8.42
episode: 2005   score: 3.0   memory length: 519165   epsilon:
0.17005132000876694   steps: 213   lr: 2.560000000000001e-06   evaluation
reward: 8.38
episode: 2006   score: 3.0   memory length: 519378   epsilon:
0.16962958000876427   steps: 213   lr: 2.560000000000001e-06   evaluation
reward: 8.34
episode: 2007   score: 9.0   memory length: 519879   epsilon: 0.168637600008758
steps: 501   lr: 2.560000000000001e-06   evaluation reward: 8.36
episode: 2008   score: 8.0   memory length: 520335   epsilon:
0.16773472000875228   steps: 456   lr: 2.560000000000001e-06   evaluation
reward: 8.34
episode: 2009   score: 13.0   memory length: 520791   epsilon:
0.16683184000874657   steps: 456   lr: 2.560000000000001e-06   evaluation
reward: 8.4
episode: 2010   score: 5.0   memory length: 521084   epsilon: 0.1662517000087429
steps: 293   lr: 2.560000000000001e-06   evaluation reward: 8.31
episode: 2011   score: 6.0   memory length: 521427   epsilon: 0.1655725600087386
steps: 343   lr: 2.560000000000001e-06   evaluation reward: 8.34
episode: 2012   score: 12.0   memory length: 521888   epsilon:
0.16465978000873283   steps: 461   lr: 2.560000000000001e-06   evaluation
reward: 8.36
episode: 2013   score: 13.0   memory length: 522467   epsilon:
0.16351336000872557   steps: 579   lr: 2.560000000000001e-06   evaluation
reward: 8.44
episode: 2014   score: 9.0   memory length: 523014   epsilon:
0.16243030000871872   steps: 547   lr: 2.560000000000001e-06   evaluation
reward: 8.48

episode: 2015 score: 10.0 memory length: 523567 epsilon: 0.1613353600087118 steps: 553 lr: 2.560000000000001e-06 evaluation reward: 8.47
episode: 2016 score: 14.0 memory length: 524188 epsilon: 0.160105780008704 steps: 621 lr: 2.560000000000001e-06 evaluation reward: 8.46
episode: 2017 score: 9.0 memory length: 524657 epsilon: 0.15917716000869814 steps: 469 lr: 2.560000000000001e-06 evaluation reward: 8.49
episode: 2018 score: 7.0 memory length: 525028 epsilon: 0.1584425800086935 steps: 371 lr: 2.560000000000001e-06 evaluation reward: 8.5
episode: 2019 score: 10.0 memory length: 525554 epsilon: 0.1574011000086869 steps: 526 lr: 2.560000000000001e-06 evaluation reward: 8.54
episode: 2020 score: 15.0 memory length: 526228 epsilon: 0.15606658000867846 steps: 674 lr: 2.560000000000001e-06 evaluation reward: 8.59
episode: 2021 score: 5.0 memory length: 526518 epsilon: 0.15549238000867482 steps: 290 lr: 2.560000000000001e-06 evaluation reward: 8.57
episode: 2022 score: 9.0 memory length: 526863 epsilon: 0.1548092800086705 steps: 345 lr: 2.560000000000001e-06 evaluation reward: 8.49
episode: 2023 score: 8.0 memory length: 527179 epsilon: 0.15418360000866654 steps: 316 lr: 2.560000000000001e-06 evaluation reward: 8.46
episode: 2024 score: 7.0 memory length: 527551 epsilon: 0.15344704000866188 steps: 372 lr: 2.560000000000001e-06 evaluation reward: 8.47
episode: 2025 score: 9.0 memory length: 528009 epsilon: 0.15254020000865615 steps: 458 lr: 2.560000000000001e-06 evaluation reward: 8.4
episode: 2026 score: 9.0 memory length: 528475 epsilon: 0.1516175200086503 steps: 466 lr: 2.560000000000001e-06 evaluation reward: 8.45
episode: 2027 score: 9.0 memory length: 528974 epsilon: 0.15062950000864406 steps: 499 lr: 2.560000000000001e-06 evaluation reward: 8.46
episode: 2028 score: 5.0 memory length: 529267 epsilon: 0.1500493600086404 steps: 293 lr: 2.560000000000001e-06 evaluation reward: 8.43
episode: 2029 score: 4.0 memory length: 529530 epsilon: 0.1495286200086371 steps: 263 lr: 2.560000000000001e-06 evaluation reward: 8.4
episode: 2030 score: 9.0 memory length: 529974 epsilon: 0.14864950000863153 steps: 444 lr: 2.560000000000001e-06 evaluation reward: 8.46
episode: 2031 score: 6.0 memory length: 530315 epsilon: 0.14797432000862726 steps: 341 lr: 2.560000000000001e-06 evaluation reward: 8.48
episode: 2032 score: 4.0 memory length: 530578 epsilon: 0.14745358000862396 steps: 263 lr: 2.560000000000001e-06 evaluation reward: 8.46

episode: 2033   score: 7.0   memory length: 530970   epsilon:
0.14667742000861905   steps: 392   lr: 2.560000000000001e-06   evaluation
reward: 8.46
episode: 2034   score: 12.0   memory length: 531554   epsilon:
0.14552110000861174   steps: 584   lr: 2.560000000000001e-06   evaluation
reward: 8.43
episode: 2035   score: 8.0   memory length: 532025   epsilon:
0.14458852000860584   steps: 471   lr: 2.560000000000001e-06   evaluation
reward: 8.44
episode: 2036   score: 16.0   memory length: 532663   epsilon:
0.14332528000859784   steps: 638   lr: 2.560000000000001e-06   evaluation
reward: 8.56
episode: 2037   score: 8.0   memory length: 533079   epsilon:
0.14250160000859263   steps: 416   lr: 2.560000000000001e-06   evaluation
reward: 8.57
episode: 2038   score: 4.0   memory length: 533342   epsilon:
0.14198086000858934   steps: 263   lr: 2.560000000000001e-06   evaluation
reward: 8.52
episode: 2039   score: 9.0   memory length: 533814   epsilon:
0.14104630000858343   steps: 472   lr: 2.560000000000001e-06   evaluation
reward: 8.51
episode: 2040   score: 3.0   memory length: 534027   epsilon:
0.14062456000858076   steps: 213   lr: 2.560000000000001e-06   evaluation
reward: 8.47
episode: 2041   score: 10.0   memory length: 534586   epsilon:
0.13951774000857375   steps: 559   lr: 2.560000000000001e-06   evaluation
reward: 8.45
episode: 2042   score: 10.0   memory length: 535089   epsilon:
0.13852180000856745   steps: 503   lr: 2.560000000000001e-06   evaluation
reward: 8.39
episode: 2043   score: 11.0   memory length: 535581   epsilon:
0.1375476400085613   steps: 492   lr: 2.560000000000001e-06   evaluation
reward: 8.4
episode: 2044   score: 11.0   memory length: 536134   epsilon:
0.13645270000855436   steps: 553   lr: 2.560000000000001e-06   evaluation
reward: 8.41
episode: 2045   score: 10.0   memory length: 536528   epsilon:
0.13567258000854943   steps: 394   lr: 2.560000000000001e-06   evaluation
reward: 8.45
episode: 2046   score: 13.0   memory length: 536927   epsilon:
0.13488256000854443   steps: 399   lr: 2.560000000000001e-06   evaluation
reward: 8.47
episode: 2047   score: 10.0   memory length: 537319   epsilon:
0.13410640000853952   steps: 392   lr: 2.560000000000001e-06   evaluation
reward: 8.49
episode: 2048   score: 9.0   memory length: 537767   epsilon: 0.1332193600085339
steps: 448   lr: 2.560000000000001e-06   evaluation reward: 8.55
episode: 2049   score: 13.0   memory length: 538416   epsilon:

0.13193434000852577 steps: 649 lr: 2.560000000000001e-06 evaluation reward: 8.63
episode: 2050 score: 6.0 memory length: 538777 epsilon: 0.13121956000852125 steps: 361 lr: 2.560000000000001e-06 evaluation reward: 8.66
episode: 2051 score: 8.0 memory length: 539211 epsilon: 0.13036024000851582 steps: 434 lr: 2.560000000000001e-06 evaluation reward: 8.66
episode: 2052 score: 4.0 memory length: 539473 epsilon: 0.12984148000851253 steps: 262 lr: 2.560000000000001e-06 evaluation reward: 8.64
episode: 2053 score: 9.0 memory length: 539976 epsilon: 0.12884554000850623 steps: 503 lr: 2.560000000000001e-06 evaluation reward: 8.61
episode: 2054 score: 16.0 memory length: 540535 epsilon: 0.12773872000849923 steps: 559 lr: 2.560000000000001e-06 evaluation reward: 8.7
episode: 2055 score: 10.0 memory length: 540909 epsilon: 0.12699820000849454 steps: 374 lr: 2.560000000000001e-06 evaluation reward: 8.67
episode: 2056 score: 7.0 memory length: 541296 epsilon: 0.1262319400084897 steps: 387 lr: 2.560000000000001e-06 evaluation reward: 8.67
episode: 2057 score: 11.0 memory length: 541846 epsilon: 0.1251429400084828 steps: 550 lr: 2.560000000000001e-06 evaluation reward: 8.7
episode: 2058 score: 7.0 memory length: 542218 epsilon: 0.12440638000848231 steps: 372 lr: 2.560000000000001e-06 evaluation reward: 8.69
episode: 2059 score: 11.0 memory length: 542794 epsilon: 0.12326590000848309 steps: 576 lr: 2.560000000000001e-06 evaluation reward: 8.72
episode: 2060 score: 13.0 memory length: 543403 epsilon: 0.12206008000848391 steps: 609 lr: 2.560000000000001e-06 evaluation reward: 8.76
episode: 2061 score: 8.0 memory length: 543880 epsilon: 0.12111562000848455 steps: 477 lr: 2.560000000000001e-06 evaluation reward: 8.71
episode: 2062 score: 5.0 memory length: 544188 epsilon: 0.12050578000848497 steps: 308 lr: 2.560000000000001e-06 evaluation reward: 8.67
episode: 2063 score: 12.0 memory length: 544681 epsilon: 0.11952964000848564 steps: 493 lr: 2.560000000000001e-06 evaluation reward: 8.75
episode: 2064 score: 10.0 memory length: 545163 epsilon: 0.11857528000848629 steps: 482 lr: 2.560000000000001e-06 evaluation reward: 8.82
episode: 2065 score: 10.0 memory length: 545648 epsilon: 0.11761498000848694 steps: 485 lr: 2.560000000000001e-06 evaluation

reward: 8.89
episode: 2066   score: 11.0   memory length: 546220   epsilon:
0.11648242000848771   steps: 572   lr: 2.560000000000001e-06   evaluation
reward: 8.87
episode: 2067   score: 3.0   memory length: 546433   epsilon: 0.116060680008488
steps: 213   lr: 2.560000000000001e-06   evaluation reward: 8.8
episode: 2068   score: 11.0   memory length: 546952   epsilon:
0.1150330600084887   steps: 519   lr: 2.560000000000001e-06   evaluation
reward: 8.84
episode: 2069   score: 12.0   memory length: 547558   epsilon:
0.11383318000848952   steps: 606   lr: 2.560000000000001e-06   evaluation
reward: 8.87
episode: 2070   score: 10.0   memory length: 548071   epsilon:
0.11281744000849021   steps: 513   lr: 2.560000000000001e-06   evaluation
reward: 8.94
episode: 2071   score: 9.0   memory length: 548517   epsilon:
0.11193436000849082   steps: 446   lr: 2.560000000000001e-06   evaluation
reward: 8.94
episode: 2072   score: 7.0   memory length: 548881   epsilon:
0.11121364000849131   steps: 364   lr: 2.560000000000001e-06   evaluation
reward: 8.85
episode: 2073   score: 9.0   memory length: 549385   epsilon:
0.11021572000849199   steps: 504   lr: 2.560000000000001e-06   evaluation
reward: 8.85
episode: 2074   score: 13.0   memory length: 549948   epsilon:
0.10910098000849275   steps: 563   lr: 2.560000000000001e-06   evaluation
reward: 8.91
episode: 2075   score: 6.0   memory length: 550301   epsilon:
0.10840204000849323   steps: 353   lr: 2.560000000000001e-06   evaluation
reward: 8.9
episode: 2076   score: 10.0   memory length: 550855   epsilon:
0.10730512000849397   steps: 554   lr: 2.560000000000001e-06   evaluation
reward: 8.96
episode: 2077   score: 14.0   memory length: 551534   epsilon:
0.10596070000849489   steps: 679   lr: 2.560000000000001e-06   evaluation
reward: 8.96
episode: 2078   score: 5.0   memory length: 551843   epsilon:
0.10534888000849531   steps: 309   lr: 2.560000000000001e-06   evaluation
reward: 8.91
episode: 2079   score: 11.0   memory length: 552224   epsilon:
0.10459450000849582   steps: 381   lr: 2.560000000000001e-06   evaluation
reward: 8.94
episode: 2080   score: 3.0   memory length: 552437   epsilon:
0.10417276000849611   steps: 213   lr: 2.560000000000001e-06   evaluation
reward: 8.93
episode: 2081   score: 9.0   memory length: 552876   epsilon: 0.1033035400084967
steps: 439   lr: 2.560000000000001e-06   evaluation reward: 8.96
episode: 2082   score: 17.0   memory length: 553516   epsilon:

0.10203634000849757   steps: 640   lr: 2.560000000000001e-06   evaluation
reward: 9.01
episode: 2083   score: 3.0   memory length: 553729   epsilon:
0.10161460000849785   steps: 213   lr: 2.560000000000001e-06   evaluation
reward: 8.92
episode: 2084   score: 8.0   memory length: 554187   epsilon:
0.10070776000849847   steps: 458   lr: 2.560000000000001e-06   evaluation
reward: 8.94
episode: 2085   score: 7.0   memory length: 554579   epsilon: 0.099931600008499
steps: 392   lr: 2.560000000000001e-06   evaluation reward: 8.92
episode: 2086   score: 8.0   memory length: 555017   epsilon: 0.0990643600084996
steps: 438   lr: 2.560000000000001e-06   evaluation reward: 8.9
episode: 2087   score: 6.0   memory length: 555357   epsilon:
0.09839116000850005   steps: 340   lr: 2.560000000000001e-06   evaluation
reward: 8.92
episode: 2088   score: 6.0   memory length: 555717   epsilon:
0.09767836000850054   steps: 360   lr: 2.560000000000001e-06   evaluation
reward: 8.83
episode: 2089   score: 10.0   memory length: 556224   epsilon:
0.09667450000850122   steps: 507   lr: 2.560000000000001e-06   evaluation
reward: 8.88
episode: 2090   score: 7.0   memory length: 556614   epsilon:
0.09590230000850175   steps: 390   lr: 2.560000000000001e-06   evaluation
reward: 8.83
episode: 2091   score: 12.0   memory length: 557177   epsilon:
0.09478756000850251   steps: 563   lr: 2.560000000000001e-06   evaluation
reward: 8.88
episode: 2092   score: 6.0   memory length: 557499   epsilon:
0.09415000000850295   steps: 322   lr: 2.560000000000001e-06   evaluation
reward: 8.8
episode: 2093   score: 11.0   memory length: 557916   epsilon:
0.09332434000850351   steps: 417   lr: 2.560000000000001e-06   evaluation
reward: 8.83
episode: 2094   score: 7.0   memory length: 558288   epsilon:
0.09258778000850401   steps: 372   lr: 2.560000000000001e-06   evaluation
reward: 8.81
episode: 2095   score: 10.0   memory length: 558755   epsilon:
0.09166312000850464   steps: 467   lr: 2.560000000000001e-06   evaluation
reward: 8.82
episode: 2096   score: 5.0   memory length: 559080   epsilon:
0.09101962000850508   steps: 325   lr: 2.560000000000001e-06   evaluation
reward: 8.76
episode: 2097   score: 11.0   memory length: 559614   epsilon:
0.0899623000085058   steps: 534   lr: 2.560000000000001e-06   evaluation
reward: 8.76
episode: 2098   score: 8.0   memory length: 560017   epsilon:
0.08916436000850635   steps: 403   lr: 2.560000000000001e-06   evaluation
reward: 8.73

```
episode: 2099   score: 7.0   memory length: 560366   epsilon:
0.08847334000850682    steps: 349    lr: 2.560000000000001e-06     evaluation
reward: 8.76
episode: 2100   score: 8.0   memory length: 560804   epsilon:
0.08760610000850741    steps: 438    lr: 2.560000000000001e-06     evaluation
reward: 8.78
episode: 2101   score: 8.0   memory length: 561207   epsilon:
0.08680816000850795    steps: 403    lr: 2.560000000000001e-06     evaluation
reward: 8.79
episode: 2102   score: 10.0   memory length: 561742   epsilon:
0.08574886000850868    steps: 535    lr: 2.560000000000001e-06     evaluation
reward: 8.82
episode: 2103   score: 11.0   memory length: 562258   epsilon:
0.08472718000850937    steps: 516    lr: 2.560000000000001e-06     evaluation
reward: 8.86
episode: 2104   score: 7.0   memory length: 562631   epsilon:
0.08398864000850988    steps: 373    lr: 2.560000000000001e-06     evaluation
reward: 8.77
episode: 2105   score: 9.0   memory length: 563088   epsilon: 0.0830837800085105
steps: 457    lr: 2.560000000000001e-06     evaluation reward: 8.83
episode: 2106   score: 15.0   memory length: 563653   epsilon:
0.08196508000851126    steps: 565    lr: 2.560000000000001e-06     evaluation
reward: 8.95
episode: 2107   score: 8.0   memory length: 564085   epsilon:
0.08110972000851184    steps: 432    lr: 2.560000000000001e-06     evaluation
reward: 8.94
episode: 2108   score: 9.0   memory length: 564558   epsilon:
0.08017318000851248    steps: 473    lr: 2.560000000000001e-06     evaluation
reward: 8.95
episode: 2109   score: 5.0   memory length: 564880   epsilon:
0.07953562000851291    steps: 322    lr: 2.560000000000001e-06     evaluation
reward: 8.87
episode: 2110   score: 7.0   memory length: 565326   epsilon:
0.07865254000851352    steps: 446    lr: 2.560000000000001e-06     evaluation
reward: 8.89
episode: 2111   score: 5.0   memory length: 565675   epsilon:
0.07796152000851399    steps: 349    lr: 2.560000000000001e-06     evaluation
reward: 8.88
episode: 2112   score: 7.0   memory length: 566048   epsilon:
0.07722298000851449    steps: 373    lr: 2.560000000000001e-06     evaluation
reward: 8.83
episode: 2113   score: 19.0   memory length: 566675   epsilon:
0.07598152000851534    steps: 627    lr: 2.560000000000001e-06     evaluation
reward: 8.89
episode: 2114   score: 8.0   memory length: 567129   epsilon:
0.07508260000851595    steps: 454    lr: 2.560000000000001e-06     evaluation
reward: 8.88
episode: 2115   score: 17.0   memory length: 567788   epsilon:
```

0.07377778000851684   steps: 659   lr: 2.560000000000001e-06   evaluation
reward: 8.95
episode: 2116   score: 5.0   memory length: 568082   epsilon:
0.07319566000851724   steps: 294   lr: 2.560000000000001e-06   evaluation
reward: 8.86
episode: 2117   score: 9.0   memory length: 568554   epsilon:
0.07226110000851788   steps: 472   lr: 2.560000000000001e-06   evaluation
reward: 8.86
episode: 2118   score: 8.0   memory length: 568997   epsilon:
0.07138396000851847   steps: 443   lr: 2.560000000000001e-06   evaluation
reward: 8.87
episode: 2119   score: 10.0   memory length: 569462   epsilon:
0.0704632600085191   steps: 465   lr: 2.560000000000001e-06   evaluation
reward: 8.87
episode: 2120   score: 15.0   memory length: 570083   epsilon:
0.06923368000851994   steps: 621   lr: 2.560000000000001e-06   evaluation
reward: 8.87
episode: 2121   score: 13.0   memory length: 570499   epsilon:
0.0684100000085205   steps: 416   lr: 2.560000000000001e-06   evaluation
reward: 8.95
episode: 2122   score: 22.0   memory length: 571180   epsilon:
0.06706162000852142   steps: 681   lr: 2.560000000000001e-06   evaluation
reward: 9.08
episode: 2123   score: 8.0   memory length: 571583   epsilon:
0.06626368000852197   steps: 403   lr: 2.560000000000001e-06   evaluation
reward: 9.08
episode: 2124   score: 5.0   memory length: 571877   epsilon:
0.06568156000852236   steps: 294   lr: 2.560000000000001e-06   evaluation
reward: 9.06
episode: 2125   score: 12.0   memory length: 572330   epsilon:
0.06478462000852298   steps: 453   lr: 2.560000000000001e-06   evaluation
reward: 9.09
episode: 2126   score: 10.0   memory length: 572812   epsilon:
0.06383026000852363   steps: 482   lr: 2.560000000000001e-06   evaluation
reward: 9.1
episode: 2127   score: 7.0   memory length: 573165   epsilon: 0.0631313200085241
steps: 353   lr: 2.560000000000001e-06   evaluation reward: 9.08
episode: 2128   score: 10.0   memory length: 573636   epsilon:
0.06219874000852474   steps: 471   lr: 2.560000000000001e-06   evaluation
reward: 9.13
episode: 2129   score: 7.0   memory length: 573989   epsilon:
0.06149980000852521   steps: 353   lr: 2.560000000000001e-06   evaluation
reward: 9.16
episode: 2130   score: 11.0   memory length: 574428   epsilon:
0.06063058000852581   steps: 439   lr: 2.560000000000001e-06   evaluation
reward: 9.18
episode: 2131   score: 8.0   memory length: 574872   epsilon:
0.05975146000852641   steps: 444   lr: 2.560000000000001e-06   evaluation

```
reward: 9.2
episode: 2132   score: 10.0   memory length: 575264   epsilon:
0.05897530000852694    steps: 392    lr: 2.560000000000001e-06     evaluation
reward: 9.26
episode: 2133   score: 5.0   memory length: 575539   epsilon:
0.05843080000852731    steps: 275    lr: 2.560000000000001e-06     evaluation
reward: 9.24
episode: 2134   score: 8.0   memory length: 575946   epsilon:
0.05762494000852786    steps: 407    lr: 2.560000000000001e-06     evaluation
reward: 9.2
episode: 2135   score: 12.0   memory length: 576422   epsilon:
0.0566824600085285    steps: 476    lr: 2.560000000000001e-06     evaluation
reward: 9.24
episode: 2136   score: 14.0   memory length: 576995   epsilon:
0.055547920008529275    steps: 573    lr: 2.560000000000001e-06     evaluation
reward: 9.22
episode: 2137   score: 9.0   memory length: 577321   epsilon:
0.054902440008529715    steps: 326    lr: 2.560000000000001e-06     evaluation
reward: 9.23
episode: 2138   score: 9.0   memory length: 577822   epsilon:
0.05391046000853039    steps: 501    lr: 2.560000000000001e-06     evaluation
reward: 9.28
episode: 2139   score: 7.0   memory length: 578175   epsilon:
0.05321152000853087    steps: 353    lr: 2.560000000000001e-06     evaluation
reward: 9.26
episode: 2140   score: 9.0   memory length: 578631   epsilon:
0.052308640008531485    steps: 456    lr: 2.560000000000001e-06     evaluation
reward: 9.32
episode: 2141   score: 7.0   memory length: 579018   epsilon:
0.05154238000853201    steps: 387    lr: 2.560000000000001e-06     evaluation
reward: 9.29
episode: 2142   score: 7.0   memory length: 579393   epsilon:
0.050799880008532514    steps: 375    lr: 2.560000000000001e-06     evaluation
reward: 9.26
episode: 2143   score: 17.0   memory length: 579970   epsilon:
0.04965742000853329    steps: 577    lr: 2.560000000000001e-06     evaluation
reward: 9.32
episode: 2144   score: 10.0   memory length: 580380   epsilon:
0.04884562000853385    steps: 410    lr: 2.560000000000001e-06     evaluation
reward: 9.31
episode: 2145   score: 7.0   memory length: 580772   epsilon:
0.048069460008534376    steps: 392    lr: 2.560000000000001e-06     evaluation
reward: 9.28
episode: 2146   score: 10.0   memory length: 581284   epsilon:
0.04705570000853507    steps: 512    lr: 2.560000000000001e-06     evaluation
reward: 9.25
episode: 2147   score: 9.0   memory length: 581740   epsilon:
0.04615282000853568    steps: 456    lr: 2.560000000000001e-06     evaluation
```

reward: 9.24
episode: 2148   score: 10.0   memory length: 582247   epsilon:
0.04514896000853637   steps: 507   lr: 2.560000000000001e-06   evaluation
reward: 9.25
episode: 2149   score: 9.0   memory length: 582734   epsilon:
0.044184700008537026   steps: 487   lr: 2.560000000000001e-06   evaluation
reward: 9.21
episode: 2150   score: 4.0   memory length: 582997   epsilon:
0.04366396000853738   steps: 263   lr: 2.560000000000001e-06   evaluation
reward: 9.19
episode: 2151   score: 9.0   memory length: 583479   epsilon:
0.04270960000853803   steps: 482   lr: 2.560000000000001e-06   evaluation
reward: 9.2
episode: 2152   score: 6.0   memory length: 583822   epsilon:
0.042030460008538495   steps: 343   lr: 2.560000000000001e-06   evaluation
reward: 9.22
episode: 2153   score: 8.0   memory length: 584225   epsilon:
0.04123252000853904   steps: 403   lr: 2.560000000000001e-06   evaluation
reward: 9.21
episode: 2154   score: 8.0   memory length: 584656   epsilon:
0.04037914000853962   steps: 431   lr: 2.560000000000001e-06   evaluation
reward: 9.13
episode: 2155   score: 9.0   memory length: 585112   epsilon:
0.03947626000854024   steps: 456   lr: 2.560000000000001e-06   evaluation
reward: 9.12
episode: 2156   score: 10.0   memory length: 585617   epsilon:
0.03847636000854092   steps: 505   lr: 2.560000000000001e-06   evaluation
reward: 9.15
episode: 2157   score: 9.0   memory length: 586074   epsilon:
0.037571500008541536   steps: 457   lr: 2.560000000000001e-06   evaluation
reward: 9.13
episode: 2158   score: 7.0   memory length: 586427   epsilon:
0.03687256000854201   steps: 353   lr: 2.560000000000001e-06   evaluation
reward: 9.13
episode: 2159   score: 9.0   memory length: 586753   epsilon:
0.03622708000854245   steps: 326   lr: 2.560000000000001e-06   evaluation
reward: 9.11
episode: 2160   score: 18.0   memory length: 587414   epsilon:
0.034918300008543346   steps: 661   lr: 2.560000000000001e-06   evaluation
reward: 9.16
episode: 2161   score: 11.0   memory length: 587955   epsilon:
0.03384712000854408   steps: 541   lr: 2.560000000000001e-06   evaluation
reward: 9.19
episode: 2162   score: 14.0   memory length: 588465   epsilon:
0.032837320008544765   steps: 510   lr: 2.560000000000001e-06   evaluation
reward: 9.28
episode: 2163   score: 9.0   memory length: 588908   epsilon:
0.031960180008545364   steps: 443   lr: 2.560000000000001e-06   evaluation

reward: 9.25
episode: 2164   score: 10.0   memory length: 589382   epsilon:
0.031021660008546004   steps: 474   lr: 2.560000000000001e-06   evaluation
reward: 9.25
episode: 2165   score: 7.0   memory length: 589735   epsilon:
0.03032272000854648   steps: 353   lr: 2.560000000000001e-06   evaluation
reward: 9.22
episode: 2166   score: 13.0   memory length: 590361   epsilon:
0.029083240008547326   steps: 626   lr: 2.560000000000001e-06   evaluation
reward: 9.24
episode: 2167   score: 10.0   memory length: 590874   epsilon:
0.02806750000854802   steps: 513   lr: 2.560000000000001e-06   evaluation
reward: 9.31
episode: 2168   score: 5.0   memory length: 591168   epsilon:
0.027485380008548416   steps: 294   lr: 2.560000000000001e-06   evaluation
reward: 9.25
episode: 2169   score: 11.0   memory length: 591607   epsilon:
0.02661616000854901   steps: 439   lr: 2.560000000000001e-06   evaluation
reward: 9.24
episode: 2170   score: 10.0   memory length: 592153   epsilon:
0.025535080008549746   steps: 546   lr: 2.560000000000001e-06   evaluation
reward: 9.24
episode: 2171   score: 7.0   memory length: 592560   epsilon:
0.024729220008550296   steps: 407   lr: 2.560000000000001e-06   evaluation
reward: 9.22
episode: 2172   score: 10.0   memory length: 593033   epsilon:
0.023792680008550934   steps: 473   lr: 2.560000000000001e-06   evaluation
reward: 9.25
episode: 2173   score: 7.0   memory length: 593386   epsilon:
0.02309374000855141   steps: 353   lr: 2.560000000000001e-06   evaluation
reward: 9.23
episode: 2174   score: 7.0   memory length: 593739   epsilon:
0.022394800008551888   steps: 353   lr: 2.560000000000001e-06   evaluation
reward: 9.17
episode: 2175   score: 10.0   memory length: 594115   epsilon:
0.021650320008552396   steps: 376   lr: 2.560000000000001e-06   evaluation
reward: 9.21
episode: 2176   score: 7.0   memory length: 594468   epsilon:
0.020951380008552872   steps: 353   lr: 2.560000000000001e-06   evaluation
reward: 9.18
episode: 2177   score: 7.0   memory length: 594821   epsilon:
0.02025244000855335   steps: 353   lr: 2.560000000000001e-06   evaluation
reward: 9.11
episode: 2178   score: 7.0   memory length: 595174   epsilon:
0.019553500008553826   steps: 353   lr: 2.560000000000001e-06   evaluation
reward: 9.13
episode: 2179   score: 12.0   memory length: 595754   epsilon:
0.01840510000855461   steps: 580   lr: 2.560000000000001e-06   evaluation

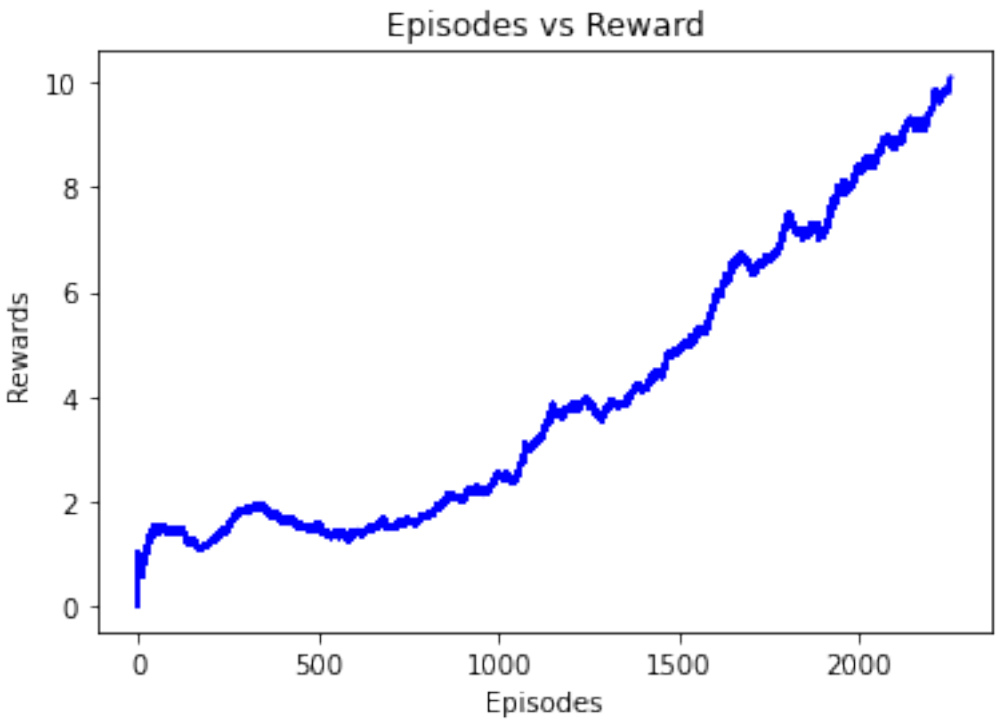
reward: 9.14
episode: 2180 score: 9.0 memory length: 596173 epsilon:
0.017575480008555175 steps: 419 lr: 2.560000000000001e-06 evaluation
reward: 9.2
episode: 2181 score: 7.0 memory length: 596526 epsilon:
0.01687654000855565 steps: 353 lr: 2.560000000000001e-06 evaluation
reward: 9.18
episode: 2182 score: 9.0 memory length: 597010 epsilon:
0.015918220008556305 steps: 484 lr: 2.560000000000001e-06 evaluation
reward: 9.1
episode: 2183 score: 7.0 memory length: 597363 epsilon:
0.015219280008556426 steps: 353 lr: 2.560000000000001e-06 evaluation
reward: 9.14
episode: 2184 score: 7.0 memory length: 597716 epsilon:
0.01452034000855629 steps: 353 lr: 2.560000000000001e-06 evaluation
reward: 9.13
episode: 2185 score: 7.0 memory length: 598088 epsilon:
0.013783780008556148 steps: 372 lr: 2.560000000000001e-06 evaluation
reward: 9.13
episode: 2186 score: 11.0 memory length: 598621 epsilon:
0.012728440008555943 steps: 533 lr: 2.560000000000001e-06 evaluation
reward: 9.16
episode: 2187 score: 13.0 memory length: 599162 epsilon:
0.011657260008555735 steps: 541 lr: 2.560000000000001e-06 evaluation
reward: 9.23
episode: 2188 score: 10.0 memory length: 599675 epsilon:
0.010641520008555538 steps: 513 lr: 2.560000000000001e-06 evaluation
reward: 9.27
episode: 2189 score: 13.0 memory length: 600154 epsilon:
0.009998020008555413 steps: 479 lr: 1.0240000000000005e-06 evaluation
reward: 9.3
episode: 2190 score: 15.0 memory length: 600754 epsilon:
0.009998020008555413 steps: 600 lr: 1.0240000000000005e-06 evaluation
reward: 9.38
episode: 2191 score: 11.0 memory length: 601318 epsilon:
0.009998020008555413 steps: 564 lr: 1.0240000000000005e-06 evaluation
reward: 9.37
episode: 2192 score: 9.0 memory length: 601805 epsilon:
0.009998020008555413 steps: 487 lr: 1.0240000000000005e-06 evaluation
reward: 9.4
episode: 2193 score: 10.0 memory length: 602318 epsilon:
0.009998020008555413 steps: 513 lr: 1.0240000000000005e-06 evaluation
reward: 9.39
episode: 2194 score: 8.0 memory length: 602750 epsilon:
0.009998020008555413 steps: 432 lr: 1.0240000000000005e-06 evaluation
reward: 9.4
episode: 2195 score: 8.0 memory length: 603182 epsilon:
0.009998020008555413 steps: 432 lr: 1.0240000000000005e-06 evaluation

```
reward: 9.38
episode: 2196   score: 9.0   memory length: 603645   epsilon:
0.009998020008555413    steps: 463    lr: 1.0240000000000005e-06     evaluation
reward: 9.42
episode: 2197   score: 10.0   memory length: 604158   epsilon:
0.009998020008555413    steps: 513    lr: 1.0240000000000005e-06     evaluation
reward: 9.41
episode: 2198   score: 9.0   memory length: 604579   epsilon:
0.009998020008555413    steps: 421    lr: 1.0240000000000005e-06     evaluation
reward: 9.42
episode: 2199   score: 8.0   memory length: 604983   epsilon:
0.009998020008555413    steps: 404    lr: 1.0240000000000005e-06     evaluation
reward: 9.43
episode: 2200   score: 7.0   memory length: 605336   epsilon:
0.009998020008555413    steps: 353    lr: 1.0240000000000005e-06     evaluation
reward: 9.42
episode: 2201   score: 14.0   memory length: 605846   epsilon:
0.009998020008555413    steps: 510    lr: 1.0240000000000005e-06     evaluation
reward: 9.48
episode: 2202   score: 11.0   memory length: 606376   epsilon:
0.009998020008555413    steps: 530    lr: 1.0240000000000005e-06     evaluation
reward: 9.49
episode: 2203   score: 10.0   memory length: 606889   epsilon:
0.009998020008555413    steps: 513    lr: 1.0240000000000005e-06     evaluation
reward: 9.48
episode: 2204   score: 11.0   memory length: 607315   epsilon:
0.009998020008555413    steps: 426    lr: 1.0240000000000005e-06     evaluation
reward: 9.52
episode: 2205   score: 10.0   memory length: 607828   epsilon:
0.009998020008555413    steps: 513    lr: 1.0240000000000005e-06     evaluation
reward: 9.53
episode: 2206   score: 21.0   memory length: 608503   epsilon:
0.009998020008555413    steps: 675    lr: 1.0240000000000005e-06     evaluation
reward: 9.59
episode: 2207   score: 10.0   memory length: 609016   epsilon:
0.009998020008555413    steps: 513    lr: 1.0240000000000005e-06     evaluation
reward: 9.61
episode: 2208   score: 10.0   memory length: 609529   epsilon:
0.009998020008555413    steps: 513    lr: 1.0240000000000005e-06     evaluation
reward: 9.62
episode: 2209   score: 19.0   memory length: 610156   epsilon:
0.009998020008555413    steps: 627    lr: 1.0240000000000005e-06     evaluation
reward: 9.76
episode: 2210   score: 7.0   memory length: 610509   epsilon:
0.009998020008555413    steps: 353    lr: 1.0240000000000005e-06     evaluation
reward: 9.76
episode: 2211   score: 13.0   memory length: 610988   epsilon:
0.009998020008555413    steps: 479    lr: 1.0240000000000005e-06     evaluation
```

```
reward: 9.84
episode: 2212   score: 10.0   memory length: 611501   epsilon:
0.009998020008555413    steps: 513    lr: 1.0240000000000005e-06     evaluation
reward: 9.87
episode: 2213   score: 12.0   memory length: 612008   epsilon:
0.009998020008555413    steps: 507    lr: 1.0240000000000005e-06     evaluation
reward: 9.8
episode: 2214   score: 11.0   memory length: 612524   epsilon:
0.009998020008555413    steps: 516    lr: 1.0240000000000005e-06     evaluation
reward: 9.83
episode: 2215   score: 7.0   memory length: 612877   epsilon:
0.009998020008555413    steps: 353    lr: 1.0240000000000005e-06     evaluation
reward: 9.73
episode: 2216   score: 7.0   memory length: 613230   epsilon:
0.009998020008555413    steps: 353    lr: 1.0240000000000005e-06     evaluation
reward: 9.75
episode: 2217   score: 7.0   memory length: 613583   epsilon:
0.009998020008555413    steps: 353    lr: 1.0240000000000005e-06     evaluation
reward: 9.73
episode: 2218   score: 12.0   memory length: 614106   epsilon:
0.009998020008555413    steps: 523    lr: 1.0240000000000005e-06     evaluation
reward: 9.77
episode: 2219   score: 13.0   memory length: 614630   epsilon:
0.009998020008555413    steps: 524    lr: 1.0240000000000005e-06     evaluation
reward: 9.8
episode: 2220   score: 11.0   memory length: 615163   epsilon:
0.009998020008555413    steps: 533    lr: 1.0240000000000005e-06     evaluation
reward: 9.76
episode: 2221   score: 16.0   memory length: 615805   epsilon:
0.009998020008555413    steps: 642    lr: 1.0240000000000005e-06     evaluation
reward: 9.79
episode: 2222   score: 8.0   memory length: 616209   epsilon:
0.009998020008555413    steps: 404    lr: 1.0240000000000005e-06     evaluation
reward: 9.65
episode: 2223   score: 7.0   memory length: 616562   epsilon:
0.009998020008555413    steps: 353    lr: 1.0240000000000005e-06     evaluation
reward: 9.64
episode: 2224   score: 7.0   memory length: 616915   epsilon:
0.009998020008555413    steps: 353    lr: 1.0240000000000005e-06     evaluation
reward: 9.66
episode: 2225   score: 11.0   memory length: 617354   epsilon:
0.009998020008555413    steps: 439    lr: 1.0240000000000005e-06     evaluation
reward: 9.65
episode: 2226   score: 11.0   memory length: 617780   epsilon:
0.009998020008555413    steps: 426    lr: 1.0240000000000005e-06     evaluation
reward: 9.66
episode: 2227   score: 13.0   memory length: 618314   epsilon:
0.009998020008555413    steps: 534    lr: 1.0240000000000005e-06     evaluation
```

reward: 9.72
episode: 2228   score: 19.0   memory length: 618942   epsilon:
0.009998020008555413   steps: 628   lr: 1.0240000000000005e-06   evaluation
reward: 9.81
episode: 2229   score: 8.0   memory length: 619346   epsilon:
0.009998020008555413   steps: 404   lr: 1.0240000000000005e-06   evaluation
reward: 9.82
episode: 2230   score: 7.0   memory length: 619699   epsilon:
0.009998020008555413   steps: 353   lr: 1.0240000000000005e-06   evaluation
reward: 9.78
episode: 2231   score: 5.0   memory length: 619993   epsilon:
0.009998020008555413   steps: 294   lr: 1.0240000000000005e-06   evaluation
reward: 9.75
episode: 2232   score: 11.0   memory length: 620432   epsilon:
0.009998020008555413   steps: 439   lr: 1.0240000000000005e-06   evaluation
reward: 9.76
episode: 2233   score: 11.0   memory length: 620871   epsilon:
0.009998020008555413   steps: 439   lr: 1.0240000000000005e-06   evaluation
reward: 9.82
episode: 2234   score: 11.0   memory length: 621310   epsilon:
0.009998020008555413   steps: 439   lr: 1.0240000000000005e-06   evaluation
reward: 9.85
episode: 2235   score: 11.0   memory length: 621736   epsilon:
0.009998020008555413   steps: 426   lr: 1.0240000000000005e-06   evaluation
reward: 9.84
episode: 2236   score: 12.0   memory length: 622316   epsilon:
0.009998020008555413   steps: 580   lr: 1.0240000000000005e-06   evaluation
reward: 9.82
episode: 2237   score: 10.0   memory length: 622809   epsilon:
0.009998020008555413   steps: 493   lr: 1.0240000000000005e-06   evaluation
reward: 9.83
episode: 2238   score: 8.0   memory length: 623213   epsilon:
0.009998020008555413   steps: 404   lr: 1.0240000000000005e-06   evaluation
reward: 9.82
episode: 2239   score: 7.0   memory length: 623566   epsilon:
0.009998020008555413   steps: 353   lr: 1.0240000000000005e-06   evaluation
reward: 9.82
episode: 2240   score: 7.0   memory length: 623919   epsilon:
0.009998020008555413   steps: 353   lr: 1.0240000000000005e-06   evaluation
reward: 9.8
episode: 2241   score: 10.0   memory length: 624336   epsilon:
0.009998020008555413   steps: 417   lr: 1.0240000000000005e-06   evaluation
reward: 9.83
episode: 2242   score: 13.0   memory length: 624853   epsilon:
0.009998020008555413   steps: 517   lr: 1.0240000000000005e-06   evaluation
reward: 9.89
episode: 2243   score: 12.0   memory length: 625407   epsilon:
0.009998020008555413   steps: 554   lr: 1.0240000000000005e-06   evaluation

```
reward: 9.84
episode: 2244   score: 7.0   memory length: 625760   epsilon:
0.009998020008555413    steps: 353    lr: 1.0240000000000005e-06     evaluation
reward: 9.81
episode: 2245   score: 8.0   memory length: 626164   epsilon:
0.009998020008555413    steps: 404    lr: 1.0240000000000005e-06     evaluation
reward: 9.82
episode: 2246   score: 10.0   memory length: 626540   epsilon:
0.009998020008555413    steps: 376    lr: 1.0240000000000005e-06     evaluation
reward: 9.82
episode: 2247   score: 9.0   memory length: 626996   epsilon:
0.009998020008555413    steps: 456    lr: 1.0240000000000005e-06     evaluation
reward: 9.82
episode: 2248   score: 11.0   memory length: 627529   epsilon:
0.009998020008555413    steps: 533    lr: 1.0240000000000005e-06     evaluation
reward: 9.83
episode: 2249   score: 18.0   memory length: 628076   epsilon:
0.009998020008555413    steps: 547    lr: 1.0240000000000005e-06     evaluation
reward: 9.92
episode: 2250   score: 12.0   memory length: 628631   epsilon:
0.009998020008555413    steps: 555    lr: 1.0240000000000005e-06     evaluation
reward: 10.0
episode: 2251   score: 8.0   memory length: 629079   epsilon:
0.009998020008555413    steps: 448    lr: 1.0240000000000005e-06     evaluation
reward: 9.99
episode: 2252   score: 11.0   memory length: 629518   epsilon:
0.009998020008555413    steps: 439    lr: 1.0240000000000005e-06     evaluation
reward: 10.04
episode: 2253   score: 11.0   memory length: 630075   epsilon:
0.009998020008555413    steps: 557    lr: 1.0240000000000005e-06     evaluation
reward: 10.07
episode: 2254   score: 10.0   memory length: 630610   epsilon:
0.009998020008555413    steps: 535    lr: 1.0240000000000005e-06     evaluation
reward: 10.09
episode: 2255   score: 11.0   memory length: 631049   epsilon:
0.009998020008555413    steps: 439    lr: 1.0240000000000005e-06     evaluation
reward: 10.11
```



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

2 Visualize Agent Performance

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

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

[ ]:
```
!pip install moviepy
```

```
Defaulting to user installation because normal site-packages is not writeable
Collecting moviepy
  Downloading moviepy-1.0.3.tar.gz (388 kB)
     |                | 388 kB 4.9 MB/s eta 0:00:01
Collecting decorator<5.0,>=4.0.2
  Downloading decorator-4.4.2-py2.py3-none-any.whl (9.2 kB)
Requirement already satisfied: tqdm<5.0,>=4.11.2 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from moviepy)
(4.64.0)
Requirement already satisfied: requests<3.0,>=2.8.1 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from moviepy)
(2.27.1)
Collecting proglog<=1.0.0
```

```
  Downloading proglog-0.1.10-py3-none-any.whl (6.1 kB)
Requirement already satisfied: numpy>=1.17.3 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from moviepy)
(1.21.5)
Requirement already satisfied: imageio<3.0,>=2.5 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from moviepy)
(2.9.0)
Collecting imageio_ffmpeg>=0.2.0
  Downloading imageio_ffmpeg-0.4.9-py3-none-manylinux2010_x86_64.whl (26.9 MB)
     |                 | 26.9 MB 123.3 MB/s eta
0:00:01               | 20.2 MB 123.3 MB/s eta 0:00:01
Requirement already satisfied: pillow in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
imageio<3.0,>=2.5->moviepy) (9.0.1)
Requirement already satisfied: setuptools in
/u/qilong/.local/lib/python3.9/site-packages (from
imageio_ffmpeg>=0.2.0->moviepy) (69.5.1)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
requests<3.0,>=2.8.1->moviepy) (1.26.9)
Requirement already satisfied: idna<4,>=2.5 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
requests<3.0,>=2.8.1->moviepy) (3.3)
Requirement already satisfied: certifi>=2017.4.17 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
requests<3.0,>=2.8.1->moviepy) (2021.10.8)
Requirement already satisfied: charset-normalizer~=2.0.0 in
/sw/external/python/anaconda3_gpu/lib/python3.9/site-packages (from
requests<3.0,>=2.8.1->moviepy) (2.0.4)
Building wheels for collected packages: moviepy
  Building wheel for moviepy (setup.py) … done
  Created wheel for moviepy: filename=moviepy-1.0.3-py3-none-any.whl
size=110728
sha256=3a23de15db3feafec79be85e1ff4f8366abf146c22d452069fe60e8cfa16afc6
  Stored in directory: /u/qilong/.cache/pip/wheels/29/15/e4/4f790bec6acd51a00b67
e8ee1394f0bc6e0135c315f8ff399a
Successfully built moviepy
Installing collected packages: proglog, imageio-ffmpeg, decorator, moviepy
Successfully installed decorator-4.4.2 imageio-ffmpeg-0.4.9 moviepy-1.0.3
proglog-0.1.10
```

2.1 DQN

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

```
import base64

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

from pyvirtualdisplay import Display

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

# Displaying the game live
def show_state(env, step=0, info=""):
    plt.figure(3)
    plt.clf()
    plt.imshow(env.render(mode='rgb_array'))
    plt.title("%s | Step: %d %s" % ("Agent Playing",step, info))
    plt.axis('off')

    ipythondisplay.clear_output(wait=True)
    ipythondisplay.display(plt.gcf())

# Recording the game and replaying the game afterwards
def show_video():
    mp4list = glob.glob('video/*.mp4')
    if len(mp4list) >= 0:
        mp4 = mp4list[0]
        video = io.open(mp4, 'r+b').read()
        encoded = base64.b64encode(video)
        ipythondisplay.display(HTML(data='''<video style="height: 400px;"␣
 ↪controls="" loop="" autoplay="" alt="test">
                <source type="video/mp4" src="data:video/mp4; base64,{0}">
             </video>'''.format(encoded.decode('ascii'))))
    else:
        print("Could not find video")


def wrap_env(env):
    env = RecordVideo(env, './video')
    return env

from agent import Agent
action_size = 3
```

```
# Load agent
agent = Agent(action_size)
agent.load_policy_net("./save_model/breakout_dqn.pth")
agent.epsilon = 0.0 # Set agent to only exploit the best action

env = wrap_env(env)

done = False
score = 0
step = 0
state, _ = env.reset()
next_state = state
life = number_lives
history = np.zeros([5, 84, 84], dtype=np.uint8)
get_init_state(history, state, HISTORY_SIZE)
frame = 0
while not done:
#     show_state(env,step) # uncommenting this provides another way to␣
 ↪visualize the game
    step += 1
    frame += 1

    # Perform a fire action if ball is no longer on screen
    if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
        action = 0
    else:
        action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
    state = next_state

    next_state, reward, done, _, info = env.step(action + 1)

    frame_next_state = get_frame(next_state)
    history[4, :, :] = frame_next_state
    terminal_state = check_live(life, info['lives'])

    life = info['lives']
    r = np.clip(reward, -1, 1)
    r = reward

    # Store the transition in memory
    agent.memory.push(deepcopy(frame_next_state), action, r, terminal_state)
    # Start training after random sample generation
    score += reward

    history[:4, :, :] = history[1:, :, :]
env.close()
show_video()
```

```
/u/qilong/.local/lib/python3.9/site-packages/gym/wrappers/record_video.py:75:
UserWarning: WARN: Overwriting existing videos at
/u/qilong/QA_generator/CS444/mp5/video folder (try specifying a different
`video_folder` for the `RecordVideo` wrapper if this is not desired)
  logger.warn(
/u/qilong/.local/lib/python3.9/site-
packages/gym/utils/passive_env_checker.py:289: UserWarning: WARN: No render
fps was declared in the environment (env.metadata['render_fps'] is None or not
defined), rendering may occur at inconsistent fps.
  logger.warn(
Moviepy - Building video /u/qilong/QA_generator/CS444/mp5/video/rl-video-
episode-0.mp4.
Moviepy - Writing video /u/qilong/QA_generator/CS444/mp5/video/rl-video-
episode-0.mp4



Moviepy - Done !
Moviepy - video ready /u/qilong/QA_generator/CS444/mp5/video/rl-video-
episode-0.mp4


<IPython.core.display.HTML object>
```

2.2 Double DQN

[ ]:
```python
from gym.wrappers import RecordVideo # If importing monitor raises issues, try␣
 ↪using `from gym.wrappers import RecordVideo`
import glob
import io
import base64

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

from pyvirtualdisplay import Display

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

# Displaying the game live
def show_state(env, step=0, info=""):
```

```
    plt.figure(3)
    plt.clf()
    plt.imshow(env.render(mode='rgb_array'))
    plt.title("%s | Step: %d %s" % ("Agent Playing",step, info))
    plt.axis('off')

    ipythondisplay.clear_output(wait=True)
    ipythondisplay.display(plt.gcf())

# Recording the game and replaying the game afterwards
def show_video():
    mp4list = glob.glob('video/*.mp4')
    if len(mp4list) >= 0:
        mp4 = mp4list[0]
        video = io.open(mp4, 'r+b').read()
        encoded = base64.b64encode(video)
        ipythondisplay.display(HTML(data='''<video style="height: 400px;"␣
 ↪controls="" loop="" autoplay="" alt="test">
              <source type="video/mp4" src="data:video/mp4; base64,{0}">
            </video>'''.format(encoded.decode('ascii'))))
    else:
        print("Could not find video")


def wrap_env(env):
    env = RecordVideo(env, './video')
    return env

from agent import Agent
action_size = 3


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

env = wrap_env(env)

done = False
score = 0
step = 0
state, _ = env.reset()
next_state = state
life = number_lives
history = np.zeros([5, 84, 84], dtype=np.uint8)
get_init_state(history, state, HISTORY_SIZE)
```

```
frame = 0
while not done:
#     show_state(env,step) # uncommenting this provides another way to␣
 ↪visualize the game
    step += 1
    frame += 1

    # Perform a fire action if ball is no longer on screen
    if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
        action = 0
    else:
        action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
    state = next_state

    next_state, reward, done, _, info = env.step(action + 1)

    frame_next_state = get_frame(next_state)
    history[4, :, :] = frame_next_state
    terminal_state = check_live(life, info['lives'])

    life = info['lives']
    r = np.clip(reward, -1, 1)
    r = reward

    # Store the transition in memory
    agent.memory.push(deepcopy(frame_next_state), action, r, terminal_state)
    # Start training after random sample generation
    score += reward

    history[:4, :, :] = history[1:, :, :]
env.close()
show_video()
```

```
/u/qilong/.local/lib/python3.9/site-packages/gym/wrappers/record_video.py:75:
UserWarning: WARN: Overwriting existing videos at
/u/qilong/QA_generator/CS444/mp5/video folder (try specifying a different
`video_folder` for the `RecordVideo` wrapper if this is not desired)
  logger.warn(
/u/qilong/.local/lib/python3.9/site-
packages/gym/utils/passive_env_checker.py:289: UserWarning: WARN: No render
fps was declared in the environment (env.metadata['render_fps'] is None or not
defined), rendering may occur at inconsistent fps.
  logger.warn(
Moviepy - Building video /u/qilong/QA_generator/CS444/mp5/video/rl-video-
episode-0.mp4.
```

```
Moviepy - Writing video /u/qilong/QA_generator/CS444/mp5/video/rl-video-
episode-0.mp4



Moviepy - Done !
Moviepy - video ready /u/qilong/QA_generator/CS444/mp5/video/rl-video-
episode-0.mp4


<IPython.core.display.HTML object>
```

3 Extra Credit 2: Demon Attack

[ ]:
```
env = gym.make("ALE/DemonAttack-v5")
state = env.reset()
print(state[0].shape)
```

```
A.L.E: Arcade Learning Environment (version 0.8.1+53f58b7)
[Powered by Stella]
(210, 160, 3)
```

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

```
4
(210, 160, 3)
3
```

3.1 DQN

[ ]:
```
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
```

[ ]:
```
rewards, episodes = [], []
best_eval_reward = 0
halt_flag = 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

        # Take action
        action = agent.get_action(np.float32(history[:4, :, :]) / 255.)

        state = next_state
        next_state, reward, terminated, truncated, info = env.step(action + 1)
        done = terminated or truncated

        frame_next_state = get_frame(next_state)
        history[4, :, :] = frame_next_state
        terminal_state = check_live(life, info['lives'])

        life = info['lives']
        r = reward

        # Store the transition in memory
        agent.memory.push(deepcopy(frame_next_state), action.cpu(), r,␣
 ↪terminal_state)
#         agent.memory.push(deepcopy(frame_next_state), action, r,␣
 ↪terminal_state)
        # Start training after random sample generation
        if(frame >= train_frame): # You can set train_frame to a lower value␣
 ↪while testing your starts training earlier
#         if frame >= 100:
            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/demonattack_dqn.png") # save graph for␣
↪training visualization

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

```
episode: 0   score: 140.0   memory length: 647   epsilon: 1.0    steps: 647
lr: 0.0001     evaluation reward: 140.0
episode: 1   score: 100.0   memory length: 1245   epsilon: 1.0    steps: 598
lr: 0.0001     evaluation reward: 120.0
episode: 2   score: 140.0   memory length: 1970   epsilon: 1.0    steps: 725
lr: 0.0001     evaluation reward: 126.66666666666667
episode: 3   score: 70.0   memory length: 2480   epsilon: 1.0    steps: 510
lr: 0.0001     evaluation reward: 112.5
episode: 4   score: 70.0   memory length: 2883   epsilon: 1.0    steps: 403
lr: 0.0001     evaluation reward: 104.0
episode: 5   score: 175.0   memory length: 4149   epsilon: 1.0    steps: 1266
lr: 0.0001     evaluation reward: 115.83333333333333
episode: 6   score: 100.0   memory length: 4717   epsilon: 1.0    steps: 568
lr: 0.0001     evaluation reward: 113.57142857142857
episode: 7   score: 110.0   memory length: 5451   epsilon: 1.0    steps: 734
lr: 0.0001     evaluation reward: 113.125
episode: 8   score: 60.0   memory length: 6012   epsilon: 1.0    steps: 561
lr: 0.0001     evaluation reward: 107.22222222222223
episode: 9   score: 140.0   memory length: 6785   epsilon: 1.0    steps: 773
lr: 0.0001     evaluation reward: 110.5
episode: 10   score: 100.0   memory length: 7518   epsilon: 1.0    steps: 733
lr: 0.0001     evaluation reward: 109.54545454545455
episode: 11   score: 120.0   memory length: 8363   epsilon: 1.0    steps: 845
lr: 0.0001     evaluation reward: 110.41666666666667
episode: 12   score: 70.0   memory length: 9163   epsilon: 1.0    steps: 800
lr: 0.0001     evaluation reward: 107.3076923076923
```

```
episode: 13   score: 355.0   memory length: 10783   epsilon: 1.0    steps: 1620
lr: 0.0001     evaluation reward: 125.0
episode: 14   score: 80.0   memory length: 11261   epsilon: 1.0    steps: 478
lr: 0.0001     evaluation reward: 122.0
episode: 15   score: 130.0   memory length: 12048   epsilon: 1.0    steps: 787
lr: 0.0001     evaluation reward: 122.5
episode: 16   score: 250.0   memory length: 13194   epsilon: 1.0    steps: 1146
lr: 0.0001     evaluation reward: 130.0
episode: 17   score: 150.0   memory length: 14130   epsilon: 1.0    steps: 936
lr: 0.0001     evaluation reward: 131.11111111111111
episode: 18   score: 130.0   memory length: 15225   epsilon: 1.0    steps: 1095
lr: 0.0001     evaluation reward: 131.05263157894737
episode: 19   score: 190.0   memory length: 16555   epsilon: 1.0    steps: 1330
lr: 0.0001     evaluation reward: 134.0
episode: 20   score: 100.0   memory length: 17224   epsilon: 1.0    steps: 669
lr: 0.0001     evaluation reward: 132.38095238095238
episode: 21   score: 90.0   memory length: 18037   epsilon: 1.0    steps: 813
lr: 0.0001     evaluation reward: 130.45454545454547
episode: 22   score: 130.0   memory length: 18861   epsilon: 1.0    steps: 824
lr: 0.0001     evaluation reward: 130.43478260869566
episode: 23   score: 60.0   memory length: 19167   epsilon: 1.0    steps: 306
lr: 0.0001     evaluation reward: 127.5
episode: 24   score: 120.0   memory length: 20038   epsilon: 1.0    steps: 871
lr: 0.0001     evaluation reward: 127.2
episode: 25   score: 250.0   memory length: 21554   epsilon: 1.0    steps: 1516
lr: 0.0001     evaluation reward: 131.92307692307693
episode: 26   score: 100.0   memory length: 22045   epsilon: 1.0    steps: 491
lr: 0.0001     evaluation reward: 130.74074074074073
episode: 27   score: 130.0   memory length: 23608   epsilon: 1.0    steps: 1563
lr: 0.0001     evaluation reward: 130.71428571428572
episode: 28   score: 205.0   memory length: 24813   epsilon: 1.0    steps: 1205
lr: 0.0001     evaluation reward: 133.27586206896552
episode: 29   score: 220.0   memory length: 25555   epsilon: 1.0    steps: 742
lr: 0.0001     evaluation reward: 136.16666666666666
episode: 30   score: 235.0   memory length: 27029   epsilon: 1.0    steps: 1474
lr: 0.0001     evaluation reward: 139.3548387096774
episode: 31   score: 220.0   memory length: 27988   epsilon: 1.0    steps: 959
lr: 0.0001     evaluation reward: 141.875
episode: 32   score: 70.0   memory length: 28847   epsilon: 1.0    steps: 859
lr: 0.0001     evaluation reward: 139.6969696969697
episode: 33   score: 90.0   memory length: 29487   epsilon: 1.0    steps: 640
lr: 0.0001     evaluation reward: 138.23529411764707
episode: 34   score: 310.0   memory length: 31183   epsilon: 1.0    steps: 1696
lr: 0.0001     evaluation reward: 143.14285714285714
episode: 35   score: 140.0   memory length: 32063   epsilon: 1.0    steps: 880
lr: 0.0001     evaluation reward: 143.05555555555554
episode: 36   score: 340.0   memory length: 33452   epsilon: 1.0    steps: 1389
lr: 0.0001     evaluation reward: 148.3783783783784
```

```
episode: 37   score: 220.0   memory length: 34757   epsilon: 1.0    steps: 1305
lr: 0.0001     evaluation reward: 150.26315789473685
episode: 38   score: 280.0   memory length: 36238   epsilon: 1.0    steps: 1481
lr: 0.0001     evaluation reward: 153.5897435897436
episode: 39   score: 220.0   memory length: 37208   epsilon: 1.0    steps: 970
lr: 0.0001     evaluation reward: 155.25
episode: 40   score: 220.0   memory length: 38095   epsilon: 1.0    steps: 887
lr: 0.0001     evaluation reward: 156.82926829268294
episode: 41   score: 120.0   memory length: 39361   epsilon: 1.0    steps: 1266
lr: 0.0001     evaluation reward: 155.95238095238096
episode: 42   score: 140.0   memory length: 40158   epsilon: 1.0    steps: 797
lr: 0.0001     evaluation reward: 155.58139534883722
episode: 43   score: 90.0   memory length: 41039   epsilon: 1.0    steps: 881
lr: 0.0001     evaluation reward: 154.0909090909091
episode: 44   score: 80.0   memory length: 41666   epsilon: 1.0    steps: 627
lr: 0.0001     evaluation reward: 152.44444444444446
episode: 45   score: 80.0   memory length: 42407   epsilon: 1.0    steps: 741
lr: 0.0001     evaluation reward: 150.8695652173913
episode: 46   score: 70.0   memory length: 42859   epsilon: 1.0    steps: 452
lr: 0.0001     evaluation reward: 149.14893617021278
episode: 47   score: 120.0   memory length: 43762   epsilon: 1.0    steps: 903
lr: 0.0001     evaluation reward: 148.54166666666666
episode: 48   score: 90.0   memory length: 44526   epsilon: 1.0    steps: 764
lr: 0.0001     evaluation reward: 147.3469387755102
episode: 49   score: 310.0   memory length: 47428   epsilon: 1.0    steps: 2902
lr: 0.0001     evaluation reward: 150.6
episode: 50   score: 60.0   memory length: 47845   epsilon: 1.0    steps: 417
lr: 0.0001     evaluation reward: 148.8235294117647
episode: 51   score: 130.0   memory length: 48673   epsilon: 1.0    steps: 828
lr: 0.0001     evaluation reward: 148.46153846153845
episode: 52   score: 160.0   memory length: 49935   epsilon: 1.0    steps: 1262
lr: 0.0001     evaluation reward: 148.67924528301887
episode: 53   score: 140.0   memory length: 50680   epsilon: 1.0    steps: 745
lr: 0.0001     evaluation reward: 148.5185185185185
episode: 54   score: 60.0   memory length: 51209   epsilon: 1.0    steps: 529
lr: 0.0001     evaluation reward: 146.9090909090909
episode: 55   score: 110.0   memory length: 51919   epsilon: 1.0    steps: 710
lr: 0.0001     evaluation reward: 146.25
episode: 56   score: 50.0   memory length: 52246   epsilon: 1.0    steps: 327
lr: 0.0001     evaluation reward: 144.56140350877192
episode: 57   score: 295.0   memory length: 53429   epsilon: 1.0    steps: 1183
lr: 0.0001     evaluation reward: 147.1551724137931
episode: 58   score: 20.0   memory length: 53666   epsilon: 1.0    steps: 237
lr: 0.0001     evaluation reward: 145.0
episode: 59   score: 120.0   memory length: 54218   epsilon: 1.0    steps: 552
lr: 0.0001     evaluation reward: 144.58333333333334
episode: 60   score: 100.0   memory length: 54705   epsilon: 1.0    steps: 487
lr: 0.0001     evaluation reward: 143.85245901639345
```

```
episode: 61   score: 130.0   memory length: 55186   epsilon: 1.0    steps: 481
lr: 0.0001     evaluation reward: 143.6290322580645
episode: 62   score: 100.0   memory length: 55941   epsilon: 1.0    steps: 755
lr: 0.0001     evaluation reward: 142.93650793650792
episode: 63   score: 90.0   memory length: 56617   epsilon: 1.0    steps: 676
lr: 0.0001     evaluation reward: 142.109375
episode: 64   score: 80.0   memory length: 57216   epsilon: 1.0    steps: 599
lr: 0.0001     evaluation reward: 141.15384615384616
episode: 65   score: 80.0   memory length: 57695   epsilon: 1.0    steps: 479
lr: 0.0001     evaluation reward: 140.22727272727272
episode: 66   score: 130.0   memory length: 58652   epsilon: 1.0    steps: 957
lr: 0.0001     evaluation reward: 140.07462686567163
episode: 67   score: 100.0   memory length: 59050   epsilon: 1.0    steps: 398
lr: 0.0001     evaluation reward: 139.48529411764707
episode: 68   score: 90.0   memory length: 59604   epsilon: 1.0    steps: 554
lr: 0.0001     evaluation reward: 138.768115942029
episode: 69   score: 325.0   memory length: 61285   epsilon: 1.0    steps: 1681
lr: 0.0001     evaluation reward: 141.42857142857142
episode: 70   score: 205.0   memory length: 62810   epsilon: 1.0    steps: 1525
lr: 0.0001     evaluation reward: 142.32394366197184
episode: 71   score: 70.0   memory length: 63286   epsilon: 1.0    steps: 476
lr: 0.0001     evaluation reward: 141.31944444444446
episode: 72   score: 100.0   memory length: 64012   epsilon: 1.0    steps: 726
lr: 0.0001     evaluation reward: 140.75342465753425
episode: 73   score: 120.0   memory length: 65476   epsilon: 1.0    steps: 1464
lr: 0.0001     evaluation reward: 140.47297297297297
episode: 74   score: 190.0   memory length: 66334   epsilon: 1.0    steps: 858
lr: 0.0001     evaluation reward: 141.13333333333333
episode: 75   score: 90.0   memory length: 66809   epsilon: 1.0    steps: 475
lr: 0.0001     evaluation reward: 140.46052631578948
episode: 76   score: 100.0   memory length: 67457   epsilon: 1.0    steps: 648
lr: 0.0001     evaluation reward: 139.93506493506493
episode: 77   score: 70.0   memory length: 68013   epsilon: 1.0    steps: 556
lr: 0.0001     evaluation reward: 139.03846153846155
episode: 78   score: 140.0   memory length: 68862   epsilon: 1.0    steps: 849
lr: 0.0001     evaluation reward: 139.0506329113924
episode: 79   score: 90.0   memory length: 69318   epsilon: 1.0    steps: 456
lr: 0.0001     evaluation reward: 138.4375
episode: 80   score: 130.0   memory length: 70087   epsilon: 1.0    steps: 769
lr: 0.0001     evaluation reward: 138.33333333333334
episode: 81   score: 220.0   memory length: 70953   epsilon: 1.0    steps: 866
lr: 0.0001     evaluation reward: 139.32926829268294
episode: 82   score: 190.0   memory length: 72405   epsilon: 1.0    steps: 1452
lr: 0.0001     evaluation reward: 139.93975903614458
episode: 83   score: 110.0   memory length: 73153   epsilon: 1.0    steps: 748
lr: 0.0001     evaluation reward: 139.58333333333334
episode: 84   score: 50.0   memory length: 73373   epsilon: 1.0    steps: 220
lr: 0.0001     evaluation reward: 138.52941176470588
```

```
episode: 85   score: 80.0   memory length: 73950   epsilon: 1.0    steps: 577
lr: 0.0001     evaluation reward: 137.84883720930233
episode: 86   score: 325.0   memory length: 75520   epsilon: 1.0    steps: 1570
lr: 0.0001     evaluation reward: 140.0
episode: 87   score: 150.0   memory length: 76295   epsilon: 1.0    steps: 775
lr: 0.0001     evaluation reward: 140.11363636363637
episode: 88   score: 110.0   memory length: 76931   epsilon: 1.0    steps: 636
lr: 0.0001     evaluation reward: 139.77528089887642
episode: 89   score: 80.0   memory length: 77687   epsilon: 1.0    steps: 756
lr: 0.0001     evaluation reward: 139.11111111111111
episode: 90   score: 295.0   memory length: 79293   epsilon: 1.0    steps: 1606
lr: 0.0001     evaluation reward: 140.82417582417582
episode: 91   score: 140.0   memory length: 79921   epsilon: 1.0    steps: 628
lr: 0.0001     evaluation reward: 140.81521739130434
episode: 92   score: 60.0   memory length: 80239   epsilon: 1.0    steps: 318
lr: 0.0001     evaluation reward: 139.94623655913978
episode: 93   score: 60.0   memory length: 80652   epsilon: 1.0    steps: 413
lr: 0.0001     evaluation reward: 139.09574468085106
episode: 94   score: 90.0   memory length: 82454   epsilon: 1.0    steps: 1802
lr: 0.0001     evaluation reward: 138.57894736842104
episode: 95   score: 140.0   memory length: 83115   epsilon: 1.0    steps: 661
lr: 0.0001     evaluation reward: 138.59375
episode: 96   score: 90.0   memory length: 83680   epsilon: 1.0    steps: 565
lr: 0.0001     evaluation reward: 138.09278350515464
episode: 97   score: 80.0   memory length: 84196   epsilon: 1.0    steps: 516
lr: 0.0001     evaluation reward: 137.5
episode: 98   score: 50.0   memory length: 84452   epsilon: 1.0    steps: 256
lr: 0.0001     evaluation reward: 136.6161616161616
episode: 99   score: 90.0   memory length: 85072   epsilon: 1.0    steps: 620
lr: 0.0001     evaluation reward: 136.15
episode: 100   score: 110.0   memory length: 85743   epsilon: 1.0    steps: 671
lr: 0.0001     evaluation reward: 135.85
episode: 101   score: 20.0   memory length: 85944   epsilon: 1.0    steps: 201
lr: 0.0001     evaluation reward: 135.05
episode: 102   score: 190.0   memory length: 86933   epsilon: 1.0    steps: 989
lr: 0.0001     evaluation reward: 135.55
episode: 103   score: 150.0   memory length: 88279   epsilon: 1.0    steps: 1346
lr: 0.0001     evaluation reward: 136.35
episode: 104   score: 110.0   memory length: 88996   epsilon: 1.0    steps: 717
lr: 0.0001     evaluation reward: 136.75
episode: 105   score: 40.0   memory length: 89241   epsilon: 1.0    steps: 245
lr: 0.0001     evaluation reward: 135.4
episode: 106   score: 130.0   memory length: 90187   epsilon: 1.0    steps: 946
lr: 0.0001     evaluation reward: 135.7
episode: 107   score: 130.0   memory length: 91005   epsilon: 1.0    steps: 818
lr: 0.0001     evaluation reward: 135.9
episode: 108   score: 205.0   memory length: 92416   epsilon: 1.0    steps: 1411
lr: 0.0001     evaluation reward: 137.35
```

```
episode: 109   score: 140.0   memory length: 93224   epsilon: 1.0    steps: 808
lr: 0.0001     evaluation reward: 137.35
episode: 110   score: 140.0   memory length: 94479   epsilon: 1.0    steps: 1255
lr: 0.0001     evaluation reward: 137.75
episode: 111   score: 130.0   memory length: 95401   epsilon: 1.0    steps: 922
lr: 0.0001     evaluation reward: 137.85
episode: 112   score: 150.0   memory length: 96028   epsilon: 1.0    steps: 627
lr: 0.0001     evaluation reward: 138.65
episode: 113   score: 140.0   memory length: 96564   epsilon: 1.0    steps: 536
lr: 0.0001     evaluation reward: 136.5
episode: 114   score: 190.0   memory length: 97552   epsilon: 1.0    steps: 988
lr: 0.0001     evaluation reward: 137.6
episode: 115   score: 80.0   memory length: 98391   epsilon: 1.0    steps: 839
lr: 0.0001     evaluation reward: 137.1
episode: 116   score: 250.0   memory length: 99402   epsilon: 1.0    steps: 1011
lr: 0.0001     evaluation reward: 137.1
```

```
/u/qilong/QA_generator/CS444/mp5/memory.py:30: FutureWarning: The input object
of type 'Tensor' is an array-like implementing one of the corresponding
protocols (`__array__`, `__array_interface__` or `__array_struct__`); but not a
sequence (or 0-D). In the future, this object will be coerced as if it was first
converted using `np.array(obj)`. To retain the old behaviour, you have to either
modify the type 'Tensor', or assign to an empty array created with
`np.empty(correct_shape, dtype=object)`.
  sample = np.array(sample, dtype=object)
/u/qilong/QA_generator/CS444/mp5/agent.py:63: FutureWarning: The input object of
type 'Tensor' is an array-like implementing one of the corresponding protocols
(`__array__`, `__array_interface__` or `__array_struct__`); but not a sequence
(or 0-D). In the future, this object will be coerced as if it was first
converted using `np.array(obj)`. To retain the old behaviour, you have to either
modify the type 'Tensor', or assign to an empty array created with
`np.empty(correct_shape, dtype=object)`.
  mini_batch = np.array(mini_batch, dtype=object).transpose()
```

```
episode: 117   score: 140.0   memory length: 100111   epsilon:
0.9997782400000048    steps: 709    lr: 0.0001     evaluation reward: 137.0
episode: 118   score: 50.0   memory length: 100429   epsilon: 0.9991486000000185
steps: 318    lr: 0.0001     evaluation reward: 136.2
episode: 119   score: 110.0   memory length: 101428   epsilon:
0.9971705800000614    steps: 999    lr: 0.0001     evaluation reward: 135.4
episode: 120   score: 80.0   memory length: 101858   epsilon: 0.9963191800000799
steps: 430    lr: 0.0001     evaluation reward: 135.2
episode: 121   score: 70.0   memory length: 102293   epsilon: 0.9954578800000986
steps: 435    lr: 0.0001     evaluation reward: 135.0
episode: 122   score: 160.0   memory length: 103781   epsilon:
0.9925116400001626    steps: 1488    lr: 0.0001     evaluation reward: 135.3
episode: 123   score: 190.0   memory length: 105192   epsilon:
0.9897178600002232    steps: 1411    lr: 0.0001     evaluation reward: 136.6
episode: 124   score: 295.0   memory length: 106559   epsilon: 0.987011200000282
```

```
steps: 1367    lr: 0.0001     evaluation reward: 138.35
episode: 125   score: 150.0   memory length: 107434   epsilon:
0.9852787000003196    steps: 875    lr: 0.0001     evaluation reward: 137.35
episode: 126   score: 60.0   memory length: 107921   epsilon: 0.9843144400003405
steps: 487    lr: 0.0001     evaluation reward: 136.95
episode: 127   score: 265.0   memory length: 110047   epsilon:
0.9801049600004319    steps: 2126    lr: 0.0001     evaluation reward: 138.3
episode: 128   score: 80.0   memory length: 110795   epsilon: 0.978623920000464
steps: 748    lr: 0.0001     evaluation reward: 137.05
episode: 129   score: 140.0   memory length: 111346   epsilon:
0.9775329400004877    steps: 551    lr: 0.0001     evaluation reward: 136.25
episode: 130   score: 160.0   memory length: 112694   epsilon:
0.9748639000005457    steps: 1348    lr: 0.0001     evaluation reward: 135.5
episode: 131   score: 70.0   memory length: 113271   epsilon: 0.9737214400005705
steps: 577    lr: 0.0001     evaluation reward: 134.0
episode: 132   score: 120.0   memory length: 113790   epsilon:
0.9726938200005928    steps: 519    lr: 0.0001     evaluation reward: 134.5
episode: 133   score: 130.0   memory length: 114493   epsilon: 0.971301880000623
steps: 703    lr: 0.0001     evaluation reward: 134.9
episode: 134   score: 175.0   memory length: 115794   epsilon:
0.9687259000006789    steps: 1301    lr: 0.0001     evaluation reward: 133.55
episode: 135   score: 205.0   memory length: 117165   epsilon:
0.9660113200007379    steps: 1371    lr: 0.0001     evaluation reward: 134.2
episode: 136   score: 110.0   memory length: 117748   epsilon:
0.9648569800007629    steps: 583    lr: 0.0001     evaluation reward: 131.9
episode: 137   score: 265.0   memory length: 119096   epsilon:
0.9621879400008209    steps: 1348    lr: 0.0001     evaluation reward: 132.35
episode: 138   score: 50.0   memory length: 119428   epsilon: 0.9615305800008351
steps: 332    lr: 0.0001     evaluation reward: 130.05
episode: 139   score: 220.0   memory length: 121102   epsilon:
0.9582160600009071    steps: 1674    lr: 0.0001     evaluation reward: 130.05
episode: 140   score: 220.0   memory length: 122955   epsilon:
0.9545471200009867    steps: 1853    lr: 0.0001     evaluation reward: 130.05
episode: 141   score: 130.0   memory length: 123823   epsilon: 0.952828480001024
steps: 868    lr: 0.0001     evaluation reward: 130.15
episode: 142   score: 190.0   memory length: 124868   epsilon: 0.950759380001069
steps: 1045    lr: 0.0001     evaluation reward: 130.65
episode: 143   score: 40.0   memory length: 125088   epsilon: 0.9503237800010784
steps: 220    lr: 0.0001     evaluation reward: 130.15
episode: 144   score: 190.0   memory length: 126450   epsilon: 0.947627020001137
steps: 1362    lr: 0.0001     evaluation reward: 131.25
episode: 145   score: 100.0   memory length: 127224   epsilon:
0.9460945000011702    steps: 774    lr: 0.0001     evaluation reward: 131.45
episode: 146   score: 90.0   memory length: 127967   epsilon: 0.9446233600012022
steps: 743    lr: 0.0001     evaluation reward: 131.65
episode: 147   score: 250.0   memory length: 129012   epsilon:
0.9425542600012471    steps: 1045    lr: 0.0001     evaluation reward: 132.95
episode: 148   score: 70.0   memory length: 129572   epsilon: 0.9414454600012712
```

```
steps: 560    lr: 0.0001     evaluation reward: 132.75
episode: 149   score: 325.0   memory length: 131120   epsilon:
0.9383804200013377    steps: 1548    lr: 0.0001     evaluation reward: 132.9
episode: 150   score: 205.0   memory length: 132592   epsilon: 0.935465860001401
steps: 1472    lr: 0.0001     evaluation reward: 134.35
episode: 151   score: 120.0   memory length: 133466   epsilon:
0.9337353400014385    steps: 874    lr: 0.0001     evaluation reward: 134.25
episode: 152   score: 130.0   memory length: 134265   epsilon:
0.9321533200014729    steps: 799    lr: 0.0001     evaluation reward: 133.95
episode: 153   score: 140.0   memory length: 134782   epsilon:
0.9311296600014951    steps: 517    lr: 0.0001     evaluation reward: 133.95
episode: 154   score: 90.0   memory length: 135409   epsilon: 0.9298882000015221
steps: 627    lr: 0.0001     evaluation reward: 134.25
episode: 155   score: 60.0   memory length: 135938   epsilon: 0.9288407800015448
steps: 529    lr: 0.0001     evaluation reward: 133.75
episode: 156   score: 90.0   memory length: 136748   epsilon: 0.9272369800015796
steps: 810    lr: 0.0001     evaluation reward: 134.15
episode: 157   score: 110.0   memory length: 138350   epsilon:
0.9240650200016485    steps: 1602    lr: 0.0001     evaluation reward: 132.3
episode: 158   score: 100.0   memory length: 138769   epsilon:
0.9232354000016665    steps: 419    lr: 0.0001     evaluation reward: 133.1
episode: 159   score: 80.0   memory length: 139199   epsilon: 0.922384000001685
steps: 430    lr: 0.0001     evaluation reward: 132.7
episode: 160   score: 60.0   memory length: 139644   epsilon: 0.9215029000017041
steps: 445    lr: 0.0001     evaluation reward: 132.3
episode: 161   score: 140.0   memory length: 140559   epsilon:
0.9196912000017434    steps: 915    lr: 0.0001     evaluation reward: 132.4
episode: 162   score: 100.0   memory length: 142436   epsilon:
0.9159747400018241    steps: 1877    lr: 0.0001     evaluation reward: 132.4
episode: 163   score: 205.0   memory length: 143543   epsilon:
0.9137828800018717    steps: 1107    lr: 0.0001     evaluation reward: 133.55
episode: 164   score: 150.0   memory length: 144341   epsilon: 0.912202840001906
steps: 798    lr: 0.0001     evaluation reward: 134.25
episode: 165   score: 130.0   memory length: 145147   epsilon:
0.9106069600019406    steps: 806    lr: 0.0001     evaluation reward: 134.75
episode: 166   score: 340.0   memory length: 146812   epsilon:
0.9073102600020122    steps: 1665    lr: 0.0001     evaluation reward: 136.85
episode: 167   score: 120.0   memory length: 147546   epsilon:
0.9058569400020438    steps: 734    lr: 0.0001     evaluation reward: 137.05
episode: 168   score: 140.0   memory length: 148124   epsilon:
0.9047125000020686    steps: 578    lr: 0.0001     evaluation reward: 137.55
episode: 169   score: 385.0   memory length: 149622   epsilon: 0.901746460002133
steps: 1498    lr: 0.0001     evaluation reward: 138.15
episode: 170   score: 235.0   memory length: 150762   epsilon: 0.899489260002182
steps: 1140    lr: 0.0001     evaluation reward: 138.45
episode: 171   score: 295.0   memory length: 152078   epsilon:
0.8968835800022386    steps: 1316    lr: 0.0001     evaluation reward: 140.7
episode: 172   score: 235.0   memory length: 153654   epsilon:
```

0.8937631000023063 steps: 1576 lr: 0.0001 evaluation reward: 142.05
episode: 173 score: 295.0 memory length: 154842 epsilon:
0.8914108600023574 steps: 1188 lr: 0.0001 evaluation reward: 143.8
episode: 174 score: 140.0 memory length: 155479 epsilon:
0.8901496000023847 steps: 637 lr: 0.0001 evaluation reward: 143.3
episode: 175 score: 190.0 memory length: 156696 epsilon:
0.8877399400024371 steps: 1217 lr: 0.0001 evaluation reward: 144.3
episode: 176 score: 100.0 memory length: 157430 epsilon:
0.8862866200024686 steps: 734 lr: 0.0001 evaluation reward: 144.3
episode: 177 score: 150.0 memory length: 158425 epsilon:
0.8843165200025114 steps: 995 lr: 0.0001 evaluation reward: 145.1
episode: 178 score: 70.0 memory length: 158762 epsilon: 0.8836492600025259
steps: 337 lr: 0.0001 evaluation reward: 144.4
episode: 179 score: 160.0 memory length: 160314 epsilon:
0.8805763000025926 steps: 1552 lr: 0.0001 evaluation reward: 145.1
episode: 180 score: 100.0 memory length: 160960 epsilon:
0.8792972200026203 steps: 646 lr: 0.0001 evaluation reward: 144.8
episode: 181 score: 130.0 memory length: 161716 epsilon:
0.8778003400026528 steps: 756 lr: 0.0001 evaluation reward: 143.9
episode: 182 score: 190.0 memory length: 162484 epsilon:
0.8762797000026858 steps: 768 lr: 0.0001 evaluation reward: 143.9
episode: 183 score: 120.0 memory length: 163143 epsilon:
0.8749748800027142 steps: 659 lr: 0.0001 evaluation reward: 144.0
episode: 184 score: 175.0 memory length: 164100 epsilon:
0.8730800200027553 steps: 957 lr: 0.0001 evaluation reward: 145.25
episode: 185 score: 50.0 memory length: 164418 epsilon: 0.872450380002769
steps: 318 lr: 0.0001 evaluation reward: 144.95
episode: 186 score: 235.0 memory length: 165284 epsilon:
0.8707357000028062 steps: 866 lr: 0.0001 evaluation reward: 144.05
episode: 187 score: 100.0 memory length: 165956 epsilon:
0.8694051400028351 steps: 672 lr: 0.0001 evaluation reward: 143.55
episode: 188 score: 250.0 memory length: 166864 epsilon:
0.8676073000028741 steps: 908 lr: 0.0001 evaluation reward: 144.95
episode: 189 score: 100.0 memory length: 167680 epsilon:
0.8659916200029092 steps: 816 lr: 0.0001 evaluation reward: 145.15
episode: 190 score: 160.0 memory length: 168753 epsilon:
0.8638670800029553 steps: 1073 lr: 0.0001 evaluation reward: 143.8
episode: 191 score: 220.0 memory length: 170037 epsilon:
0.8613247600030105 steps: 1284 lr: 0.0001 evaluation reward: 144.6
episode: 192 score: 110.0 memory length: 170485 epsilon:
0.8604377200030298 steps: 448 lr: 0.0001 evaluation reward: 145.1
episode: 193 score: 220.0 memory length: 171468 epsilon: 0.858491380003072
steps: 983 lr: 0.0001 evaluation reward: 146.7
episode: 194 score: 150.0 memory length: 172698 epsilon:
0.8560559800031249 steps: 1230 lr: 0.0001 evaluation reward: 147.3
episode: 195 score: 340.0 memory length: 174375 epsilon: 0.852735520003197
steps: 1677 lr: 0.0001 evaluation reward: 149.3
episode: 196 score: 250.0 memory length: 176938 epsilon:

```
0.8476607800033071    steps: 2563    lr: 0.0001     evaluation reward: 150.9
episode: 197   score: 120.0   memory length: 178050   epsilon:
0.8454590200033549    steps: 1112    lr: 0.0001     evaluation reward: 151.3
episode: 198   score: 265.0   memory length: 179089   epsilon:
0.8434018000033996    steps: 1039    lr: 0.0001     evaluation reward: 153.45
episode: 199   score: 310.0   memory length: 180521   epsilon:
0.8405664400034611    steps: 1432    lr: 0.0001     evaluation reward: 155.65
episode: 200   score: 190.0   memory length: 181473   epsilon:
0.8386814800035021    steps: 952    lr: 0.0001     evaluation reward: 156.45
episode: 201   score: 110.0   memory length: 182009   epsilon:
0.8376202000035251    steps: 536    lr: 0.0001     evaluation reward: 157.35
episode: 202   score: 295.0   memory length: 183722   epsilon:
0.8342284600035987    steps: 1713    lr: 0.0001     evaluation reward: 158.4
episode: 203   score: 205.0   memory length: 184798   epsilon: 0.832097980003645
steps: 1076    lr: 0.0001     evaluation reward: 158.95
episode: 204   score: 100.0   memory length: 185467   epsilon:
0.8307733600036737    steps: 669    lr: 0.0001     evaluation reward: 158.85
episode: 205   score: 70.0   memory length: 185914   epsilon: 0.829888300003693
steps: 447    lr: 0.0001     evaluation reward: 159.15
episode: 206   score: 140.0   memory length: 186573   epsilon:
0.8285834800037213    steps: 659    lr: 0.0001     evaluation reward: 159.25
episode: 207   score: 175.0   memory length: 187864   epsilon:
0.8260273000037768    steps: 1291    lr: 0.0001     evaluation reward: 159.7
episode: 208   score: 60.0   memory length: 188314   epsilon: 0.8251363000037961
steps: 450    lr: 0.0001     evaluation reward: 158.25
episode: 209   score: 100.0   memory length: 189121   epsilon:
0.8235384400038308    steps: 807    lr: 0.0001     evaluation reward: 157.85
episode: 210   score: 310.0   memory length: 190458   epsilon:
0.8208911800038883    steps: 1337    lr: 0.0001     evaluation reward: 159.55
episode: 211   score: 120.0   memory length: 191412   epsilon:
0.8190022600039293    steps: 954    lr: 0.0001     evaluation reward: 159.45
episode: 212   score: 160.0   memory length: 192625   epsilon:
0.8166005200039814    steps: 1213    lr: 0.0001     evaluation reward: 159.55
episode: 213   score: 190.0   memory length: 193377   epsilon:
0.8151115600040137    steps: 752    lr: 0.0001     evaluation reward: 160.05
episode: 214   score: 70.0   memory length: 193743   epsilon: 0.8143868800040295
steps: 366    lr: 0.0001     evaluation reward: 158.85
episode: 215   score: 420.0   memory length: 195624   epsilon:
0.8106625000041103    steps: 1881    lr: 0.0001     evaluation reward: 162.25
episode: 216   score: 100.0   memory length: 196169   epsilon:
0.8095834000041338    steps: 545    lr: 0.0001     evaluation reward: 160.75
episode: 217   score: 120.0   memory length: 197095   epsilon:
0.8077499200041736    steps: 926    lr: 0.0001     evaluation reward: 160.55
episode: 218   score: 175.0   memory length: 198011   epsilon:
0.8059362400042129    steps: 916    lr: 0.0001     evaluation reward: 161.8
episode: 219   score: 160.0   memory length: 199085   epsilon:
0.8038097200042591    steps: 1074    lr: 0.0001     evaluation reward: 162.3
episode: 220   score: 220.0   memory length: 200876   epsilon:
```

0.8002635400043361   steps: 1791   lr: 4e-05   evaluation reward: 163.7
episode: 221   score: 90.0   memory length: 201649   epsilon: 0.7987330000043693
steps: 773   lr: 4e-05   evaluation reward: 163.9
episode: 222   score: 140.0   memory length: 202587   epsilon:
0.7968757600044096   steps: 938   lr: 4e-05   evaluation reward: 163.7
episode: 223   score: 140.0   memory length: 203331   epsilon:
0.7954026400044416   steps: 744   lr: 4e-05   evaluation reward: 163.2
episode: 224   score: 100.0   memory length: 204109   epsilon: 0.793862200004475
steps: 778   lr: 4e-05   evaluation reward: 161.25
episode: 225   score: 120.0   memory length: 205141   epsilon:
0.7918188400045194   steps: 1032   lr: 4e-05   evaluation reward: 160.95
episode: 226   score: 235.0   memory length: 206128   epsilon:
0.7898645800045618   steps: 987   lr: 4e-05   evaluation reward: 162.7
episode: 227   score: 100.0   memory length: 206711   epsilon:
0.7887102400045869   steps: 583   lr: 4e-05   evaluation reward: 161.05
episode: 228   score: 120.0   memory length: 207478   epsilon:
0.7871915800046199   steps: 767   lr: 4e-05   evaluation reward: 161.45
episode: 229   score: 160.0   memory length: 208169   epsilon:
0.7858234000046496   steps: 691   lr: 4e-05   evaluation reward: 161.65
episode: 230   score: 160.0   memory length: 208926   epsilon:
0.7843245400046821   steps: 757   lr: 4e-05   evaluation reward: 161.65
episode: 231   score: 160.0   memory length: 209936   epsilon:
0.7823247400047255   steps: 1010   lr: 4e-05   evaluation reward: 162.55
episode: 232   score: 120.0   memory length: 210702   epsilon:
0.7808080600047584   steps: 766   lr: 4e-05   evaluation reward: 162.55
episode: 233   score: 190.0   memory length: 211597   epsilon:
0.7790359600047969   steps: 895   lr: 4e-05   evaluation reward: 163.15
episode: 234   score: 120.0   memory length: 212133   epsilon: 0.77797468000482
steps: 536   lr: 4e-05   evaluation reward: 162.6
episode: 235   score: 220.0   memory length: 213723   epsilon:
0.7748264800048883   steps: 1590   lr: 4e-05   evaluation reward: 162.75
episode: 236   score: 50.0   memory length: 214125   epsilon: 0.7740305200049056
steps: 402   lr: 4e-05   evaluation reward: 162.15
episode: 237   score: 280.0   memory length: 215698   epsilon:
0.7709159800049732   steps: 1573   lr: 4e-05   evaluation reward: 162.3
episode: 238   score: 70.0   memory length: 216302   epsilon: 0.7697200600049992
steps: 604   lr: 4e-05   evaluation reward: 162.5
episode: 239   score: 250.0   memory length: 217383   epsilon:
0.7675796800050456   steps: 1081   lr: 4e-05   evaluation reward: 162.8
episode: 240   score: 325.0   memory length: 220165   epsilon:
0.7620713200051652   steps: 2782   lr: 4e-05   evaluation reward: 163.85
episode: 241   score: 90.0   memory length: 221068   epsilon: 0.760283380005204
steps: 903   lr: 4e-05   evaluation reward: 163.45
episode: 242   score: 370.0   memory length: 222437   epsilon:
0.7575727600052629   steps: 1369   lr: 4e-05   evaluation reward: 165.25
episode: 243   score: 130.0   memory length: 223077   epsilon:
0.7563055600052904   steps: 640   lr: 4e-05   evaluation reward: 166.15
episode: 244   score: 120.0   memory length: 223717   epsilon:

```
0.7550383600053179    steps: 640    lr: 4e-05     evaluation reward: 165.45
episode: 245   score: 140.0   memory length: 224534   epsilon: 0.753420700005353
steps: 817    lr: 4e-05     evaluation reward: 165.85
episode: 246   score: 190.0   memory length: 225828   epsilon:
0.7508585800054086    steps: 1294    lr: 4e-05     evaluation reward: 166.85
episode: 247   score: 140.0   memory length: 226852   epsilon:
0.7488310600054526    steps: 1024    lr: 4e-05     evaluation reward: 165.75
episode: 248   score: 100.0   memory length: 227383   epsilon:
0.7477796800054755    steps: 531    lr: 4e-05     evaluation reward: 166.05
episode: 249   score: 140.0   memory length: 228181   epsilon:
0.7461996400055098    steps: 798    lr: 4e-05     evaluation reward: 164.2
episode: 250   score: 80.0   memory length: 228726   epsilon: 0.7451205400055332
steps: 545    lr: 4e-05     evaluation reward: 162.95
episode: 251   score: 385.0   memory length: 230298   epsilon:
0.7420079800056008    steps: 1572    lr: 4e-05     evaluation reward: 165.6
episode: 252   score: 110.0   memory length: 231082   epsilon:
0.7404556600056345    steps: 784    lr: 4e-05     evaluation reward: 165.4
episode: 253   score: 120.0   memory length: 231726   epsilon:
0.7391805400056621    steps: 644    lr: 4e-05     evaluation reward: 165.2
episode: 254   score: 60.0   memory length: 232116   epsilon: 0.7384083400056789
steps: 390    lr: 4e-05     evaluation reward: 164.9
episode: 255   score: 60.0   memory length: 232518   epsilon: 0.7376123800056962
steps: 402    lr: 4e-05     evaluation reward: 164.9
episode: 256   score: 60.0   memory length: 232921   epsilon: 0.7368144400057135
steps: 403    lr: 4e-05     evaluation reward: 164.6
episode: 257   score: 120.0   memory length: 233602   epsilon:
0.7354660600057428    steps: 681    lr: 4e-05     evaluation reward: 164.7
episode: 258   score: 40.0   memory length: 233903   epsilon: 0.7348700800057557
steps: 301    lr: 4e-05     evaluation reward: 164.1
episode: 259   score: 40.0   memory length: 234191   epsilon: 0.7342998400057681
steps: 288    lr: 4e-05     evaluation reward: 163.7
episode: 260   score: 460.0   memory length: 236390   epsilon:
0.7299458200058626    steps: 2199    lr: 4e-05     evaluation reward: 167.7
episode: 261   score: 90.0   memory length: 237044   epsilon: 0.7286509000058907
steps: 654    lr: 4e-05     evaluation reward: 167.2
episode: 262   score: 160.0   memory length: 237907   epsilon:
0.7269421600059278    steps: 863    lr: 4e-05     evaluation reward: 167.8
episode: 263   score: 130.0   memory length: 238530   epsilon:
0.7257086200059546    steps: 623    lr: 4e-05     evaluation reward: 167.05
episode: 264   score: 205.0   memory length: 239307   epsilon: 0.724170160005988
steps: 777    lr: 4e-05     evaluation reward: 167.6
episode: 265   score: 205.0   memory length: 240171   epsilon:
0.7224594400060251    steps: 864    lr: 4e-05     evaluation reward: 168.35
episode: 266   score: 190.0   memory length: 241029   epsilon: 0.720760600006062
steps: 858    lr: 4e-05     evaluation reward: 166.85
episode: 267   score: 235.0   memory length: 242099   epsilon: 0.718642000006108
steps: 1070    lr: 4e-05     evaluation reward: 168.0
episode: 268   score: 100.0   memory length: 242589   epsilon:
```

0.7176718000061291 steps: 490 lr: 4e-05 evaluation reward: 167.6
episode: 269 score: 100.0 memory length: 243089 epsilon:
0.7166818000061506 steps: 500 lr: 4e-05 evaluation reward: 164.75
episode: 270 score: 140.0 memory length: 244033 epsilon:
0.7148126800061911 steps: 944 lr: 4e-05 evaluation reward: 163.8
episode: 271 score: 130.0 memory length: 244710 epsilon:
0.7134722200062202 steps: 677 lr: 4e-05 evaluation reward: 162.15
episode: 272 score: 110.0 memory length: 245254 epsilon:
0.7123951000062436 steps: 544 lr: 4e-05 evaluation reward: 160.9
episode: 273 score: 160.0 memory length: 245998 epsilon:
0.7109219800062756 steps: 744 lr: 4e-05 evaluation reward: 159.55
episode: 274 score: 60.0 memory length: 246316 epsilon: 0.7102923400062893
steps: 318 lr: 4e-05 evaluation reward: 158.75
episode: 275 score: 235.0 memory length: 247279 epsilon:
0.7083856000063307 steps: 963 lr: 4e-05 evaluation reward: 159.2
episode: 276 score: 70.0 memory length: 247759 epsilon: 0.7074352000063513
steps: 480 lr: 4e-05 evaluation reward: 158.9
episode: 277 score: 265.0 memory length: 248924 epsilon:
0.7051285000064014 steps: 1165 lr: 4e-05 evaluation reward: 160.05
episode: 278 score: 280.0 memory length: 250171 epsilon: 0.702659440006455
steps: 1247 lr: 4e-05 evaluation reward: 162.15
episode: 279 score: 310.0 memory length: 251371 epsilon:
0.7002834400065066 steps: 1200 lr: 4e-05 evaluation reward: 163.65
episode: 280 score: 80.0 memory length: 251928 epsilon: 0.6991805800065305
steps: 557 lr: 4e-05 evaluation reward: 163.45
episode: 281 score: 110.0 memory length: 252603 epsilon:
0.6978440800065595 steps: 675 lr: 4e-05 evaluation reward: 163.25
episode: 282 score: 80.0 memory length: 253063 epsilon: 0.6969332800065793
steps: 460 lr: 4e-05 evaluation reward: 162.15
episode: 283 score: 120.0 memory length: 253806 epsilon:
0.6954621400066112 steps: 743 lr: 4e-05 evaluation reward: 162.15
episode: 284 score: 265.0 memory length: 254746 epsilon:
0.6936009400066516 steps: 940 lr: 4e-05 evaluation reward: 163.05
episode: 285 score: 150.0 memory length: 255578 epsilon:
0.6919535800066874 steps: 832 lr: 4e-05 evaluation reward: 164.05
episode: 286 score: 80.0 memory length: 256397 epsilon: 0.6903319600067226
steps: 819 lr: 4e-05 evaluation reward: 162.5
episode: 287 score: 90.0 memory length: 256898 epsilon: 0.6893399800067441
steps: 501 lr: 4e-05 evaluation reward: 162.4
episode: 288 score: 325.0 memory length: 258190 epsilon:
0.6867818200067997 steps: 1292 lr: 4e-05 evaluation reward: 163.15
episode: 289 score: 70.0 memory length: 258572 epsilon: 0.6860254600068161
steps: 382 lr: 4e-05 evaluation reward: 162.85
episode: 290 score: 190.0 memory length: 259613 epsilon:
0.6839642800068608 steps: 1041 lr: 4e-05 evaluation reward: 163.15
episode: 291 score: 150.0 memory length: 260276 epsilon:
0.6826515400068893 steps: 663 lr: 4e-05 evaluation reward: 162.45
episode: 292 score: 50.0 memory length: 260535 epsilon: 0.6821387200069005

steps: 259   lr: 4e-05   evaluation reward: 161.85
episode: 293   score: 190.0   memory length: 261350   epsilon:
0.6805250200069355   steps: 815   lr: 4e-05   evaluation reward: 161.55
episode: 294   score: 80.0   memory length: 261698   epsilon: 0.6798359800069504
steps: 348   lr: 4e-05   evaluation reward: 160.85
episode: 295   score: 150.0   memory length: 262529   epsilon:
0.6781906000069862   steps: 831   lr: 4e-05   evaluation reward: 158.95
episode: 296   score: 140.0   memory length: 263271   epsilon:
0.6767214400070181   steps: 742   lr: 4e-05   evaluation reward: 157.85
episode: 297   score: 190.0   memory length: 263973   epsilon:
0.6753314800070482   steps: 702   lr: 4e-05   evaluation reward: 158.55
episode: 298   score: 265.0   memory length: 264969   epsilon: 0.673359400007091
steps: 996   lr: 4e-05   evaluation reward: 158.55
episode: 299   score: 140.0   memory length: 265905   epsilon:
0.6715061200071313   steps: 936   lr: 4e-05   evaluation reward: 156.85
episode: 300   score: 250.0   memory length: 266933   epsilon:
0.6694706800071755   steps: 1028   lr: 4e-05   evaluation reward: 157.45
episode: 301   score: 140.0   memory length: 267782   epsilon: 0.667789660007212
steps: 849   lr: 4e-05   evaluation reward: 157.75
episode: 302   score: 80.0   memory length: 268285   epsilon: 0.6667937200072336
steps: 503   lr: 4e-05   evaluation reward: 155.6
episode: 303   score: 370.0   memory length: 269635   epsilon:
0.6641207200072916   steps: 1350   lr: 4e-05   evaluation reward: 157.25
episode: 304   score: 60.0   memory length: 270096   epsilon: 0.6632079400073114
steps: 461   lr: 4e-05   evaluation reward: 156.85
episode: 305   score: 60.0   memory length: 270426   epsilon: 0.6625545400073256
steps: 330   lr: 4e-05   evaluation reward: 156.75
episode: 306   score: 80.0   memory length: 270756   epsilon: 0.6619011400073398
steps: 330   lr: 4e-05   evaluation reward: 156.15
episode: 307   score: 90.0   memory length: 271243   epsilon: 0.6609368800073607
steps: 487   lr: 4e-05   evaluation reward: 155.3
episode: 308   score: 120.0   memory length: 271937   epsilon:
0.6595627600073906   steps: 694   lr: 4e-05   evaluation reward: 155.9
episode: 309   score: 90.0   memory length: 272551   epsilon: 0.658347040007417
steps: 614   lr: 4e-05   evaluation reward: 155.8
episode: 310   score: 70.0   memory length: 272932   epsilon: 0.6575926600074333
steps: 381   lr: 4e-05   evaluation reward: 153.4
episode: 311   score: 160.0   memory length: 273529   epsilon: 0.656410600007459
steps: 597   lr: 4e-05   evaluation reward: 153.8
episode: 312   score: 110.0   memory length: 274207   epsilon:
0.6550681600074881   steps: 678   lr: 4e-05   evaluation reward: 153.3
episode: 313   score: 120.0   memory length: 274821   epsilon:
0.6538524400075145   steps: 614   lr: 4e-05   evaluation reward: 152.6
episode: 314   score: 110.0   memory length: 275521   epsilon:
0.6524664400075446   steps: 700   lr: 4e-05   evaluation reward: 153.0
episode: 315   score: 140.0   memory length: 276369   epsilon:
0.6507874000075811   steps: 848   lr: 4e-05   evaluation reward: 150.2
episode: 316   score: 205.0   memory length: 277443   epsilon:

0.6486608800076272   steps: 1074   lr: 4e-05   evaluation reward: 151.25
episode: 317   score: 110.0   memory length: 277957   epsilon:
0.6476431600076493   steps: 514   lr: 4e-05   evaluation reward: 151.15
episode: 318   score: 100.0   memory length: 278620   epsilon:
0.6463304200076778   steps: 663   lr: 4e-05   evaluation reward: 150.4
episode: 319   score: 60.0   memory length: 278973   epsilon: 0.645631480007693
steps: 353   lr: 4e-05   evaluation reward: 149.4
episode: 320   score: 205.0   memory length: 280013   epsilon:
0.6435722800077377   steps: 1040   lr: 4e-05   evaluation reward: 149.25
episode: 321   score: 160.0   memory length: 281002   epsilon:
0.6416140600077802   steps: 989   lr: 4e-05   evaluation reward: 149.95
episode: 322   score: 60.0   memory length: 281334   epsilon: 0.6409567000077945
steps: 332   lr: 4e-05   evaluation reward: 149.15
episode: 323   score: 110.0   memory length: 282049   epsilon:
0.6395410000078252   steps: 715   lr: 4e-05   evaluation reward: 148.85
episode: 324   score: 110.0   memory length: 283081   epsilon:
0.6374976400078696   steps: 1032   lr: 4e-05   evaluation reward: 148.95
episode: 325   score: 190.0   memory length: 284158   epsilon:
0.6353651800079159   steps: 1077   lr: 4e-05   evaluation reward: 149.65
episode: 326   score: 220.0   memory length: 284988   epsilon:
0.6337217800079515   steps: 830   lr: 4e-05   evaluation reward: 149.5
episode: 327   score: 110.0   memory length: 285539   epsilon:
0.6326308000079752   steps: 551   lr: 4e-05   evaluation reward: 149.6
episode: 328   score: 140.0   memory length: 286452   epsilon:
0.6308230600080145   steps: 913   lr: 4e-05   evaluation reward: 149.8
episode: 329   score: 110.0   memory length: 287064   epsilon:
0.6296113000080408   steps: 612   lr: 4e-05   evaluation reward: 149.3
episode: 330   score: 120.0   memory length: 287725   epsilon:
0.6283025200080692   steps: 661   lr: 4e-05   evaluation reward: 148.9
episode: 331   score: 205.0   memory length: 288671   epsilon:
0.6264294400081099   steps: 946   lr: 4e-05   evaluation reward: 149.35
episode: 332   score: 50.0   memory length: 289054   epsilon: 0.6256711000081263
steps: 383   lr: 4e-05   evaluation reward: 148.65
episode: 333   score: 90.0   memory length: 289671   epsilon: 0.6244494400081528
steps: 617   lr: 4e-05   evaluation reward: 147.65
episode: 334   score: 80.0   memory length: 290261   epsilon: 0.6232812400081782
steps: 590   lr: 4e-05   evaluation reward: 147.25
episode: 335   score: 120.0   memory length: 290728   epsilon:
0.6223565800081983   steps: 467   lr: 4e-05   evaluation reward: 146.25
episode: 336   score: 90.0   memory length: 291052   epsilon: 0.6217150600082122
steps: 324   lr: 4e-05   evaluation reward: 146.65
episode: 337   score: 140.0   memory length: 291762   epsilon:
0.6203092600082427   steps: 710   lr: 4e-05   evaluation reward: 145.25
episode: 338   score: 80.0   memory length: 292442   epsilon: 0.6189628600082719
steps: 680   lr: 4e-05   evaluation reward: 145.35
episode: 339   score: 130.0   memory length: 293271   epsilon:
0.6173214400083076   steps: 829   lr: 4e-05   evaluation reward: 144.15
episode: 340   score: 265.0   memory length: 294221   epsilon:

0.6154404400083484   steps: 950   lr: 4e-05   evaluation reward: 143.55
episode: 341   score: 90.0   memory length: 294746   epsilon: 0.614400940008371
steps: 525   lr: 4e-05   evaluation reward: 143.55
episode: 342   score: 90.0   memory length: 295246   epsilon: 0.6134109400083925
steps: 500   lr: 4e-05   evaluation reward: 140.75
episode: 343   score: 130.0   memory length: 295946   epsilon:
0.6120249400084226   steps: 700   lr: 4e-05   evaluation reward: 140.75
episode: 344   score: 220.0   memory length: 296989   epsilon:
0.6099598000084674   steps: 1043   lr: 4e-05   evaluation reward: 141.75
episode: 345   score: 150.0   memory length: 297967   epsilon:
0.6080233600085094   steps: 978   lr: 4e-05   evaluation reward: 141.85
episode: 346   score: 130.0   memory length: 298596   epsilon:
0.6067779400085365   steps: 629   lr: 4e-05   evaluation reward: 141.25
episode: 347   score: 160.0   memory length: 299512   epsilon:
0.6049642600085758   steps: 916   lr: 4e-05   evaluation reward: 141.45
episode: 348   score: 235.0   memory length: 300546   epsilon:
0.6029169400086203   steps: 1034   lr: 1.6000000000000003e-05   evaluation
reward: 142.8
episode: 349   score: 100.0   memory length: 301152   epsilon:
0.6017170600086463   steps: 606   lr: 1.6000000000000003e-05   evaluation
reward: 142.4
episode: 350   score: 235.0   memory length: 302353   epsilon: 0.599339080008698
steps: 1201   lr: 1.6000000000000003e-05   evaluation reward: 143.95
episode: 351   score: 90.0   memory length: 302943   epsilon: 0.5981708800087233
steps: 590   lr: 1.6000000000000003e-05   evaluation reward: 141.0
episode: 352   score: 130.0   memory length: 303560   epsilon:
0.5969492200087498   steps: 617   lr: 1.6000000000000003e-05   evaluation
reward: 141.2
episode: 353   score: 100.0   memory length: 304238   epsilon: 0.595606780008779
steps: 678   lr: 1.6000000000000003e-05   evaluation reward: 141.0
episode: 354   score: 140.0   memory length: 304997   epsilon:
0.5941039600088116   steps: 759   lr: 1.6000000000000003e-05   evaluation
reward: 141.8
episode: 355   score: 295.0   memory length: 306031   epsilon: 0.592056640008856
steps: 1034   lr: 1.6000000000000003e-05   evaluation reward: 144.15
episode: 356   score: 295.0   memory length: 307162   epsilon:
0.5898172600089047   steps: 1131   lr: 1.6000000000000003e-05   evaluation
reward: 146.5
episode: 357   score: 120.0   memory length: 307822   epsilon: 0.588510460008933
steps: 660   lr: 1.6000000000000003e-05   evaluation reward: 146.5
episode: 358   score: 100.0   memory length: 308556   epsilon:
0.5870571400089646   steps: 734   lr: 1.6000000000000003e-05   evaluation
reward: 147.1
episode: 359   score: 30.0   memory length: 308816   epsilon: 0.5865423400089758
steps: 260   lr: 1.6000000000000003e-05   evaluation reward: 147.0
episode: 360   score: 140.0   memory length: 309474   epsilon: 0.585239500009004
steps: 658   lr: 1.6000000000000003e-05   evaluation reward: 143.8
episode: 361   score: 140.0   memory length: 310190   epsilon:

0.5838218200090348 steps: 716 lr: 1.6000000000000003e-05 evaluation reward: 144.3
episode: 362 score: 100.0 memory length: 310649 epsilon: 0.5829130000090545 steps: 459 lr: 1.6000000000000003e-05 evaluation reward: 143.7
episode: 363 score: 160.0 memory length: 311532 epsilon: 0.5811646600090925 steps: 883 lr: 1.6000000000000003e-05 evaluation reward: 144.0
episode: 364 score: 40.0 memory length: 311774 epsilon: 0.5806855000091029 steps: 242 lr: 1.6000000000000003e-05 evaluation reward: 142.35
episode: 365 score: 160.0 memory length: 312627 epsilon: 0.5789965600091396 steps: 853 lr: 1.6000000000000003e-05 evaluation reward: 141.9
episode: 366 score: 130.0 memory length: 313326 epsilon: 0.5776125400091696 steps: 699 lr: 1.6000000000000003e-05 evaluation reward: 141.3
episode: 367 score: 110.0 memory length: 314004 epsilon: 0.5762701000091988 steps: 678 lr: 1.6000000000000003e-05 evaluation reward: 140.05
episode: 368 score: 160.0 memory length: 314793 epsilon: 0.5747078800092327 steps: 789 lr: 1.6000000000000003e-05 evaluation reward: 140.65
episode: 369 score: 280.0 memory length: 315979 epsilon: 0.5723596000092837 steps: 1186 lr: 1.6000000000000003e-05 evaluation reward: 142.45
episode: 370 score: 100.0 memory length: 316577 epsilon: 0.5711755600093094 steps: 598 lr: 1.6000000000000003e-05 evaluation reward: 142.05
episode: 371 score: 280.0 memory length: 317824 epsilon: 0.568706500009363 steps: 1247 lr: 1.6000000000000003e-05 evaluation reward: 143.55
episode: 372 score: 175.0 memory length: 318801 epsilon: 0.566772040009405 steps: 977 lr: 1.6000000000000003e-05 evaluation reward: 144.2
episode: 373 score: 340.0 memory length: 319979 epsilon: 0.5644396000094556 steps: 1178 lr: 1.6000000000000003e-05 evaluation reward: 146.0
episode: 374 score: 100.0 memory length: 320399 epsilon: 0.5636080000094736 steps: 420 lr: 1.6000000000000003e-05 evaluation reward: 146.4
episode: 375 score: 90.0 memory length: 320850 epsilon: 0.562715020009493 steps: 451 lr: 1.6000000000000003e-05 evaluation reward: 144.95
episode: 376 score: 110.0 memory length: 321445 epsilon: 0.5615369200095186 steps: 595 lr: 1.6000000000000003e-05 evaluation reward: 145.35
episode: 377 score: 220.0 memory length: 322461 epsilon: 0.5595252400095623 steps: 1016 lr: 1.6000000000000003e-05 evaluation reward: 144.9
episode: 378 score: 220.0 memory length: 323309 epsilon: 0.5578462000095987 steps: 848 lr: 1.6000000000000003e-05 evaluation

reward: 144.3
episode: 379   score: 280.0   memory length: 324406   epsilon:
0.5556741400096459   steps: 1097   lr: 1.6000000000000003e-05   evaluation
reward: 144.0
episode: 380   score: 100.0   memory length: 324984   epsilon:
0.5545297000096707   steps: 578   lr: 1.6000000000000003e-05   evaluation
reward: 144.2
episode: 381   score: 250.0   memory length: 325839   epsilon:
0.5528368000097075   steps: 855   lr: 1.6000000000000003e-05   evaluation
reward: 145.6
episode: 382   score: 110.0   memory length: 326521   epsilon:
0.5514864400097368   steps: 682   lr: 1.6000000000000003e-05   evaluation
reward: 145.9
episode: 383   score: 80.0   memory length: 327065   epsilon: 0.5504093200097602
steps: 544   lr: 1.6000000000000003e-05   evaluation reward: 145.5
episode: 384   score: 90.0   memory length: 327684   epsilon: 0.5491837000097868
steps: 619   lr: 1.6000000000000003e-05   evaluation reward: 143.75
episode: 385   score: 100.0   memory length: 328262   epsilon:
0.5480392600098116   steps: 578   lr: 1.6000000000000003e-05   evaluation
reward: 143.25
episode: 386   score: 250.0   memory length: 329160   epsilon:
0.5462612200098502   steps: 898   lr: 1.6000000000000003e-05   evaluation
reward: 144.95
episode: 387   score: 100.0   memory length: 329638   epsilon:
0.5453147800098708   steps: 478   lr: 1.6000000000000003e-05   evaluation
reward: 145.05
episode: 388   score: 110.0   memory length: 330348   epsilon:
0.5439089800099013   steps: 710   lr: 1.6000000000000003e-05   evaluation
reward: 142.9
episode: 389   score: 220.0   memory length: 331257   epsilon:
0.5421091600099404   steps: 909   lr: 1.6000000000000003e-05   evaluation
reward: 144.4
episode: 390   score: 220.0   memory length: 332442   epsilon:
0.5397628600099913   steps: 1185   lr: 1.6000000000000003e-05   evaluation
reward: 144.7
episode: 391   score: 160.0   memory length: 333388   epsilon: 0.537889780010032
steps: 946   lr: 1.6000000000000003e-05   evaluation reward: 144.8
episode: 392   score: 175.0   memory length: 334293   epsilon:
0.5360978800100709   steps: 905   lr: 1.6000000000000003e-05   evaluation
reward: 146.05
episode: 393   score: 140.0   memory length: 334847   epsilon:
0.5350009600100947   steps: 554   lr: 1.6000000000000003e-05   evaluation
reward: 145.55
episode: 394   score: 160.0   memory length: 335797   epsilon:
0.5331199600101355   steps: 950   lr: 1.6000000000000003e-05   evaluation
reward: 146.35
episode: 395   score: 110.0   memory length: 336313   epsilon:
0.5320982800101577   steps: 516   lr: 1.6000000000000003e-05   evaluation

```
reward: 145.95
episode: 396   score: 80.0   memory length: 336999   epsilon: 0.5307400000101872
steps: 686    lr: 1.6000000000000003e-05     evaluation reward: 145.35
episode: 397   score: 80.0   memory length: 337294   epsilon: 0.5301559000101999
steps: 295    lr: 1.6000000000000003e-05     evaluation reward: 144.25
episode: 398   score: 160.0   memory length: 337927   epsilon:
0.5289025600102271    steps: 633    lr: 1.6000000000000003e-05     evaluation
reward: 143.2
episode: 399   score: 160.0   memory length: 338742   epsilon:
0.5272888600102621    steps: 815    lr: 1.6000000000000003e-05     evaluation
reward: 143.4
episode: 400   score: 60.0   memory length: 339255   epsilon: 0.5262731200102841
steps: 513    lr: 1.6000000000000003e-05     evaluation reward: 141.5
episode: 401   score: 140.0   memory length: 340106   epsilon:
0.5245881400103207    steps: 851    lr: 1.6000000000000003e-05     evaluation
reward: 141.5
episode: 402   score: 205.0   memory length: 340995   epsilon:
0.5228279200103589    steps: 889    lr: 1.6000000000000003e-05     evaluation
reward: 142.75
episode: 403   score: 110.0   memory length: 341687   epsilon:
0.5214577600103887    steps: 692    lr: 1.6000000000000003e-05     evaluation
reward: 140.15
episode: 404   score: 140.0   memory length: 342344   epsilon:
0.5201569000104169    steps: 657    lr: 1.6000000000000003e-05     evaluation
reward: 140.95
episode: 405   score: 120.0   memory length: 342911   epsilon:
0.5190342400104413    steps: 567    lr: 1.6000000000000003e-05     evaluation
reward: 141.55
episode: 406   score: 60.0   memory length: 343468   epsilon: 0.5179313800104652
steps: 557    lr: 1.6000000000000003e-05     evaluation reward: 141.35
episode: 407   score: 175.0   memory length: 344381   epsilon:
0.5161236400105045    steps: 913    lr: 1.6000000000000003e-05     evaluation
reward: 142.2
episode: 408   score: 60.0   memory length: 344925   epsilon: 0.5150465200105279
steps: 544    lr: 1.6000000000000003e-05     evaluation reward: 141.6
episode: 409   score: 100.0   memory length: 345575   epsilon:
0.5137595200105558    steps: 650    lr: 1.6000000000000003e-05     evaluation
reward: 141.7
episode: 410   score: 175.0   memory length: 346553   epsilon:
0.5118230800105978    steps: 978    lr: 1.6000000000000003e-05     evaluation
reward: 142.75
episode: 411   score: 90.0   memory length: 346970   epsilon: 0.5109974200106158
steps: 417    lr: 1.6000000000000003e-05     evaluation reward: 142.05
episode: 412   score: 190.0   memory length: 347867   epsilon:
0.5092213600106543    steps: 897    lr: 1.6000000000000003e-05     evaluation
reward: 142.85
episode: 413   score: 235.0   memory length: 348736   epsilon:
0.5075007400106917    steps: 869    lr: 1.6000000000000003e-05     evaluation
```

```
reward: 144.0
episode: 414   score: 150.0   memory length: 349763   epsilon:
0.5054672800107358    steps: 1027    lr: 1.6000000000000003e-05     evaluation
reward: 144.4
episode: 415   score: 110.0   memory length: 350437   epsilon:
0.5041327600107648    steps: 674    lr: 1.6000000000000003e-05     evaluation
reward: 144.1
episode: 416   score: 120.0   memory length: 351087   epsilon:
0.5028457600107927    steps: 650    lr: 1.6000000000000003e-05     evaluation
reward: 143.25
episode: 417   score: 175.0   memory length: 351976   epsilon:
0.5010855400108309    steps: 889    lr: 1.6000000000000003e-05     evaluation
reward: 143.9
episode: 418   score: 235.0   memory length: 352831   epsilon:
0.49939264001085065    steps: 855    lr: 1.6000000000000003e-05     evaluation
reward: 145.25
episode: 419   score: 110.0   memory length: 353490   epsilon:
0.4980878200108424    steps: 659    lr: 1.6000000000000003e-05     evaluation
reward: 145.75
episode: 420   score: 370.0   memory length: 354816   epsilon:
0.4954623400108258    steps: 1326    lr: 1.6000000000000003e-05     evaluation
reward: 147.4
episode: 421   score: 110.0   memory length: 355333   epsilon:
0.4944386800108193    steps: 517    lr: 1.6000000000000003e-05     evaluation
reward: 146.9
episode: 422   score: 130.0   memory length: 356082   epsilon:
0.4929556600108099    steps: 749    lr: 1.6000000000000003e-05     evaluation
reward: 147.6
episode: 423   score: 150.0   memory length: 356943   epsilon:
0.49125088001079914    steps: 861    lr: 1.6000000000000003e-05     evaluation
reward: 148.0
episode: 424   score: 370.0   memory length: 358126   epsilon:
0.4889085400107843    steps: 1183    lr: 1.6000000000000003e-05     evaluation
reward: 150.6
episode: 425   score: 160.0   memory length: 358842   epsilon:
0.48749086001077535    steps: 716    lr: 1.6000000000000003e-05     evaluation
reward: 150.3
episode: 426   score: 150.0   memory length: 359616   epsilon:
0.48595834001076565    steps: 774    lr: 1.6000000000000003e-05     evaluation
reward: 149.6
episode: 427   score: 80.0   memory length: 360157   epsilon: 0.4848871600107589
steps: 541    lr: 1.6000000000000003e-05     evaluation reward: 149.3
episode: 428   score: 220.0   memory length: 361311   epsilon:
0.4826022400107444    steps: 1154    lr: 1.6000000000000003e-05     evaluation
reward: 150.1
episode: 429   score: 310.0   memory length: 362290   epsilon:
0.48066382001073216    steps: 979    lr: 1.6000000000000003e-05     evaluation
reward: 152.1
```

episode: 430   score: 150.0   memory length: 363027   epsilon:
0.4792045600107229   steps: 737   lr: 1.6000000000000003e-05   evaluation
reward: 152.4
episode: 431   score: 280.0   memory length: 364390   epsilon:
0.47650582001070585   steps: 1363   lr: 1.6000000000000003e-05   evaluation
reward: 153.15
episode: 432   score: 60.0   memory length: 364662   epsilon:
0.47596726001070244   steps: 272   lr: 1.6000000000000003e-05   evaluation
reward: 153.25
episode: 433   score: 235.0   memory length: 365380   epsilon:
0.47454562001069345   steps: 718   lr: 1.6000000000000003e-05   evaluation
reward: 154.7
episode: 434   score: 60.0   memory length: 365902   epsilon: 0.4735120600106869
steps: 522   lr: 1.6000000000000003e-05   evaluation reward: 154.5
episode: 435   score: 110.0   memory length: 366676   epsilon:
0.4719795400106772   steps: 774   lr: 1.6000000000000003e-05   evaluation
reward: 154.4
episode: 436   score: 140.0   memory length: 367480   epsilon:
0.47038762001066714   steps: 804   lr: 1.6000000000000003e-05   evaluation
reward: 154.9
episode: 437   score: 40.0   memory length: 367750   epsilon:
0.46985302001066376   steps: 270   lr: 1.6000000000000003e-05   evaluation
reward: 153.9
episode: 438   score: 60.0   memory length: 368012   epsilon: 0.4693342600106605
steps: 262   lr: 1.6000000000000003e-05   evaluation reward: 153.7
episode: 439   score: 265.0   memory length: 369042   epsilon:
0.4672948600106476   steps: 1030   lr: 1.6000000000000003e-05   evaluation
reward: 155.05
episode: 440   score: 175.0   memory length: 369810   epsilon:
0.46577422001063795   steps: 768   lr: 1.6000000000000003e-05   evaluation
reward: 154.15
episode: 441   score: 250.0   memory length: 370880   epsilon:
0.46365562001062455   steps: 1070   lr: 1.6000000000000003e-05   evaluation
reward: 155.75
episode: 442   score: 265.0   memory length: 371860   epsilon:
0.46171522001061227   steps: 980   lr: 1.6000000000000003e-05   evaluation
reward: 157.5
episode: 443   score: 250.0   memory length: 373066   epsilon:
0.45932734001059716   steps: 1206   lr: 1.6000000000000003e-05   evaluation
reward: 158.7
episode: 444   score: 265.0   memory length: 374085   epsilon:
0.4573097200105844   steps: 1019   lr: 1.6000000000000003e-05   evaluation
reward: 159.15
episode: 445   score: 100.0   memory length: 374480   epsilon:
0.45652762001057945   steps: 395   lr: 1.6000000000000003e-05   evaluation
reward: 158.65
episode: 446   score: 205.0   memory length: 375425   epsilon:
0.4546565200105676   steps: 945   lr: 1.6000000000000003e-05   evaluation

```
reward: 159.4
episode: 447   score: 150.0   memory length: 376289   epsilon:
0.4529458000105568    steps: 864    lr: 1.6000000000000003e-05     evaluation
reward: 159.3
episode: 448   score: 205.0   memory length: 377206   epsilon:
0.4511301400105453    steps: 917    lr: 1.6000000000000003e-05     evaluation
reward: 159.0
episode: 449   score: 20.0   memory length: 377493   epsilon: 0.4505618800105417
steps: 287    lr: 1.6000000000000003e-05     evaluation reward: 158.2
episode: 450   score: 100.0   memory length: 378087   epsilon:
0.44938576001053426    steps: 594    lr: 1.6000000000000003e-05     evaluation
reward: 156.85
episode: 451   score: 110.0   memory length: 378733   epsilon:
0.44810668001052617    steps: 646    lr: 1.6000000000000003e-05     evaluation
reward: 157.05
episode: 452   score: 130.0   memory length: 379295   epsilon:
0.44699392001051913    steps: 562    lr: 1.6000000000000003e-05     evaluation
reward: 157.05
episode: 453   score: 205.0   memory length: 380384   epsilon:
0.4448377000105055    steps: 1089    lr: 1.6000000000000003e-05     evaluation
reward: 158.1
episode: 454   score: 70.0   memory length: 380864   epsilon: 0.4438873000104995
steps: 480    lr: 1.6000000000000003e-05     evaluation reward: 157.4
episode: 455   score: 80.0   memory length: 381556   epsilon: 0.4425171400104908
steps: 692    lr: 1.6000000000000003e-05     evaluation reward: 155.25
episode: 456   score: 190.0   memory length: 382609   epsilon:
0.4404322000104776    steps: 1053    lr: 1.6000000000000003e-05     evaluation
reward: 154.2
episode: 457   score: 30.0   memory length: 382868   epsilon:
0.43991938001047437    steps: 259    lr: 1.6000000000000003e-05     evaluation
reward: 153.3
episode: 458   score: 295.0   memory length: 383902   epsilon:
0.4378720600104614    steps: 1034    lr: 1.6000000000000003e-05     evaluation
reward: 155.25
episode: 459   score: 205.0   memory length: 385243   epsilon:
0.4352168800104446    steps: 1341    lr: 1.6000000000000003e-05     evaluation
reward: 157.0
episode: 460   score: 190.0   memory length: 386355   epsilon:
0.4330151200104307    steps: 1112    lr: 1.6000000000000003e-05     evaluation
reward: 157.5
episode: 461   score: 205.0   memory length: 387146   epsilon:
0.4314489400104208    steps: 791    lr: 1.6000000000000003e-05     evaluation
reward: 158.15
episode: 462   score: 220.0   memory length: 387927   epsilon: 0.429902560010411
steps: 781    lr: 1.6000000000000003e-05     evaluation reward: 159.35
episode: 463   score: 80.0   memory length: 388513   epsilon:
0.42874228001040365    steps: 586    lr: 1.6000000000000003e-05     evaluation
reward: 158.55
```

```
episode: 464   score: 280.0   memory length: 389687   epsilon:
0.42641776001038895    steps: 1174    lr: 1.6000000000000003e-05     evaluation
reward: 160.95
episode: 465   score: 220.0   memory length: 390837   epsilon:
0.42414076001037454    steps: 1150    lr: 1.6000000000000003e-05     evaluation
reward: 161.55
episode: 466   score: 110.0   memory length: 391573   epsilon:
0.4226834800103653    steps: 736    lr: 1.6000000000000003e-05     evaluation
reward: 161.35
episode: 467   score: 310.0   memory length: 392892   epsilon:
0.4200718600103488    steps: 1319    lr: 1.6000000000000003e-05     evaluation
reward: 163.35
episode: 468   score: 100.0   memory length: 393489   epsilon:
0.4188898000103413    steps: 597    lr: 1.6000000000000003e-05     evaluation
reward: 162.75
episode: 469   score: 70.0   memory length: 394031   epsilon:
0.41781664001033453    steps: 542    lr: 1.6000000000000003e-05     evaluation
reward: 160.65
episode: 470   score: 130.0   memory length: 394899   epsilon:
0.41609800001032365    steps: 868    lr: 1.6000000000000003e-05     evaluation
reward: 160.95
episode: 471   score: 60.0   memory length: 395401   epsilon:
0.41510404001031737    steps: 502    lr: 1.6000000000000003e-05     evaluation
reward: 158.75
episode: 472   score: 50.0   memory length: 395687   epsilon: 0.4145377600103138
steps: 286    lr: 1.6000000000000003e-05     evaluation reward: 157.5
episode: 473   score: 80.0   memory length: 396146   epsilon:
0.41362894001030803    steps: 459    lr: 1.6000000000000003e-05     evaluation
reward: 154.9
episode: 474   score: 190.0   memory length: 396986   epsilon:
0.4119657400102975    steps: 840    lr: 1.6000000000000003e-05     evaluation
reward: 155.8
episode: 475   score: 265.0   memory length: 398032   epsilon:
0.4098946600102844    steps: 1046    lr: 1.6000000000000003e-05     evaluation
reward: 157.55
episode: 476   score: 160.0   memory length: 398904   epsilon:
0.4081681000102735    steps: 872    lr: 1.6000000000000003e-05     evaluation
reward: 158.05
episode: 477   score: 110.0   memory length: 399257   epsilon:
0.40746916001026906    steps: 353    lr: 1.6000000000000003e-05     evaluation
reward: 156.95
episode: 478   score: 325.0   memory length: 400289   epsilon:
0.40542580001025613    steps: 1032    lr: 6.400000000000001e-06     evaluation
reward: 158.0
episode: 479   score: 110.0   memory length: 401201   epsilon:
0.4036200400102447    steps: 912    lr: 6.400000000000001e-06     evaluation
reward: 156.3
episode: 480   score: 150.0   memory length: 401747   epsilon:
```

0.40253896001023787   steps: 546   lr: 6.400000000000001e-06   evaluation
reward: 156.8
episode: 481   score: 175.0   memory length: 402346   epsilon:
0.40135294001023036   steps: 599   lr: 6.400000000000001e-06   evaluation
reward: 156.05
episode: 482   score: 265.0   memory length: 403546   epsilon:
0.39897694001021533   steps: 1200   lr: 6.400000000000001e-06   evaluation
reward: 157.6
episode: 483   score: 325.0   memory length: 404999   epsilon:
0.39610000001019713   steps: 1453   lr: 6.400000000000001e-06   evaluation
reward: 160.05
episode: 484   score: 40.0   memory length: 405347   epsilon:
0.39541096001019277   steps: 348   lr: 6.400000000000001e-06   evaluation
reward: 159.55
episode: 485   score: 120.0   memory length: 406010   epsilon:
0.39409822001018446   steps: 663   lr: 6.400000000000001e-06   evaluation
reward: 159.75
episode: 486   score: 355.0   memory length: 407437   epsilon:
0.3912727600101666   steps: 1427   lr: 6.400000000000001e-06   evaluation
reward: 160.8
episode: 487   score: 175.0   memory length: 408190   epsilon:
0.38978182001015715   steps: 753   lr: 6.400000000000001e-06   evaluation
reward: 161.55
episode: 488   score: 310.0   memory length: 409621   epsilon:
0.3869484400101392   steps: 1431   lr: 6.400000000000001e-06   evaluation
reward: 163.55
episode: 489   score: 90.0   memory length: 410107   epsilon:
0.38598616001013314   steps: 486   lr: 6.400000000000001e-06   evaluation
reward: 162.25
episode: 490   score: 150.0   memory length: 411019   epsilon:
0.3841804000101217   steps: 912   lr: 6.400000000000001e-06   evaluation
reward: 161.55
episode: 491   score: 100.0   memory length: 411576   epsilon:
0.38307754001011474   steps: 557   lr: 6.400000000000001e-06   evaluation
reward: 160.95
episode: 492   score: 160.0   memory length: 412540   epsilon:
0.38116882001010266   steps: 964   lr: 6.400000000000001e-06   evaluation
reward: 160.8
episode: 493   score: 130.0   memory length: 413517   epsilon:
0.3792343600100904   steps: 977   lr: 6.400000000000001e-06   evaluation
reward: 160.7
episode: 494   score: 190.0   memory length: 414512   epsilon:
0.37726426001007796   steps: 995   lr: 6.400000000000001e-06   evaluation
reward: 161.0
episode: 495   score: 150.0   memory length: 415169   epsilon:
0.3759634000100697   steps: 657   lr: 6.400000000000001e-06   evaluation
reward: 161.4
episode: 496   score: 110.0   memory length: 415561   epsilon:

0.3751872400100648   steps: 392   lr: 6.400000000000001e-06   evaluation
reward: 161.7
episode: 497   score: 160.0   memory length: 416221   epsilon:
0.37388044001005655   steps: 660   lr: 6.400000000000001e-06   evaluation
reward: 162.5
episode: 498   score: 140.0   memory length: 417114   epsilon:
0.37211230001004536   steps: 893   lr: 6.400000000000001e-06   evaluation
reward: 162.3
episode: 499   score: 90.0   memory length: 417651   epsilon:
0.37104904001003863   steps: 537   lr: 6.400000000000001e-06   evaluation
reward: 161.6
episode: 500   score: 190.0   memory length: 418589   epsilon:
0.3691918000100269   steps: 938   lr: 6.400000000000001e-06   evaluation
reward: 162.9
episode: 501   score: 60.0   memory length: 419041   epsilon: 0.3682968400100212
steps: 452   lr: 6.400000000000001e-06   evaluation reward: 162.1
episode: 502   score: 220.0   memory length: 420088   epsilon:
0.3662237800100081   steps: 1047   lr: 6.400000000000001e-06   evaluation
reward: 162.25
episode: 503   score: 220.0   memory length: 421115   epsilon:
0.36419032000999524   steps: 1027   lr: 6.400000000000001e-06   evaluation
reward: 163.35
episode: 504   score: 280.0   memory length: 422259   epsilon:
0.3619252000099809   steps: 1144   lr: 6.400000000000001e-06   evaluation
reward: 164.75
episode: 505   score: 100.0   memory length: 422867   epsilon:
0.3607213600099733   steps: 608   lr: 6.400000000000001e-06   evaluation
reward: 164.55
episode: 506   score: 265.0   memory length: 423796   epsilon:
0.35888194000996165   steps: 929   lr: 6.400000000000001e-06   evaluation
reward: 166.6
episode: 507   score: 60.0   memory length: 424160   epsilon: 0.3581612200099571
steps: 364   lr: 6.400000000000001e-06   evaluation reward: 165.45
episode: 508   score: 250.0   memory length: 425201   epsilon:
0.35610004000994405   steps: 1041   lr: 6.400000000000001e-06   evaluation
reward: 167.35
episode: 509   score: 175.0   memory length: 426120   epsilon:
0.35428042000993254   steps: 919   lr: 6.400000000000001e-06   evaluation
reward: 168.1
episode: 510   score: 60.0   memory length: 426491   epsilon: 0.3535458400099279
steps: 371   lr: 6.400000000000001e-06   evaluation reward: 166.95
episode: 511   score: 220.0   memory length: 427827   epsilon:
0.35090056000991116   steps: 1336   lr: 6.400000000000001e-06   evaluation
reward: 168.25
episode: 512   score: 385.0   memory length: 429286   epsilon:
0.3480117400098929   steps: 1459   lr: 6.400000000000001e-06   evaluation
reward: 170.2
episode: 513   score: 235.0   memory length: 430304   epsilon:

0.3459961000098801 steps: 1018 lr: 6.400000000000001e-06 evaluation reward: 170.2
episode: 514 score: 110.0 memory length: 430728 epsilon: 0.3451565800098748 steps: 424 lr: 6.400000000000001e-06 evaluation reward: 169.8
episode: 515 score: 310.0 memory length: 431821 epsilon: 0.3429924400098611 steps: 1093 lr: 6.400000000000001e-06 evaluation reward: 171.8
episode: 516 score: 310.0 memory length: 432807 epsilon: 0.34104016000984877 steps: 986 lr: 6.400000000000001e-06 evaluation reward: 173.7
episode: 517 score: 220.0 memory length: 433566 epsilon: 0.33953734000983926 steps: 759 lr: 6.400000000000001e-06 evaluation reward: 174.15
episode: 518 score: 460.0 memory length: 435130 epsilon: 0.33644062000981967 steps: 1564 lr: 6.400000000000001e-06 evaluation reward: 176.4
episode: 519 score: 295.0 memory length: 436096 epsilon: 0.33452794000980757 steps: 966 lr: 6.400000000000001e-06 evaluation reward: 178.25
episode: 520 score: 60.0 memory length: 436484 epsilon: 0.3337597000098027 steps: 388 lr: 6.400000000000001e-06 evaluation reward: 175.15
episode: 521 score: 50.0 memory length: 436880 epsilon: 0.33297562000979775 steps: 396 lr: 6.400000000000001e-06 evaluation reward: 174.55
episode: 522 score: 175.0 memory length: 437519 epsilon: 0.33171040000978974 steps: 639 lr: 6.400000000000001e-06 evaluation reward: 175.0
episode: 523 score: 400.0 memory length: 438501 epsilon: 0.32976604000977744 steps: 982 lr: 6.400000000000001e-06 evaluation reward: 177.5
episode: 524 score: 90.0 memory length: 439007 epsilon: 0.3287641600097711 steps: 506 lr: 6.400000000000001e-06 evaluation reward: 174.7
episode: 525 score: 150.0 memory length: 439635 epsilon: 0.32752072000976323 steps: 628 lr: 6.400000000000001e-06 evaluation reward: 174.6
episode: 526 score: 90.0 memory length: 440104 epsilon: 0.32659210000975736 steps: 469 lr: 6.400000000000001e-06 evaluation reward: 174.0
episode: 527 score: 80.0 memory length: 440612 epsilon: 0.325586260009751 steps: 508 lr: 6.400000000000001e-06 evaluation reward: 174.0
episode: 528 score: 150.0 memory length: 441399 epsilon: 0.32402800000974114 steps: 787 lr: 6.400000000000001e-06 evaluation reward: 173.3
episode: 529 score: 120.0 memory length: 442299 epsilon: 0.32224600000972986 steps: 900 lr: 6.400000000000001e-06 evaluation reward: 171.4
episode: 530 score: 30.0 memory length: 442562 epsilon:

0.32172526000972657   steps: 263   lr: 6.400000000000001e-06   evaluation
reward: 170.2
episode: 531   score: 235.0   memory length: 443410   epsilon:
0.32004622000971594   steps: 848   lr: 6.400000000000001e-06   evaluation
reward: 169.75
episode: 532   score: 110.0   memory length: 444082   epsilon:
0.3187156600097075   steps: 672   lr: 6.400000000000001e-06   evaluation
reward: 170.25
episode: 533   score: 265.0   memory length: 444852   epsilon:
0.3171910600096979   steps: 770   lr: 6.400000000000001e-06   evaluation
reward: 170.55
episode: 534   score: 130.0   memory length: 445471   epsilon:
0.3159654400096901   steps: 619   lr: 6.400000000000001e-06   evaluation
reward: 171.25
episode: 535   score: 90.0   memory length: 445863   epsilon: 0.3151892800096852
steps: 392   lr: 6.400000000000001e-06   evaluation reward: 171.05
episode: 536   score: 120.0   memory length: 446398   epsilon:
0.3141299800096785   steps: 535   lr: 6.400000000000001e-06   evaluation
reward: 170.85
episode: 537   score: 80.0   memory length: 446816   epsilon: 0.3133023400096733
steps: 418   lr: 6.400000000000001e-06   evaluation reward: 171.25
episode: 538   score: 190.0   memory length: 447600   epsilon:
0.31175002000966345   steps: 784   lr: 6.400000000000001e-06   evaluation
reward: 172.55
episode: 539   score: 220.0   memory length: 448551   epsilon:
0.30986704000965154   steps: 951   lr: 6.400000000000001e-06   evaluation
reward: 172.1
episode: 540   score: 30.0   memory length: 448790   epsilon:
0.30939382000964855   steps: 239   lr: 6.400000000000001e-06   evaluation
reward: 170.65
episode: 541   score: 175.0   memory length: 449644   epsilon:
0.30770290000963785   steps: 854   lr: 6.400000000000001e-06   evaluation
reward: 169.9
episode: 542   score: 295.0   memory length: 450707   epsilon:
0.30559816000962453   steps: 1063   lr: 6.400000000000001e-06   evaluation
reward: 170.2
episode: 543   score: 960.0   memory length: 452528   epsilon:
0.3019925800096017   steps: 1821   lr: 6.400000000000001e-06   evaluation
reward: 177.3
episode: 544   score: 205.0   memory length: 453382   epsilon: 0.300301660009591
steps: 854   lr: 6.400000000000001e-06   evaluation reward: 176.7
episode: 545   score: 100.0   memory length: 453902   epsilon:
0.2992720600095845   steps: 520   lr: 6.400000000000001e-06   evaluation
reward: 176.7
episode: 546   score: 235.0   memory length: 455047   epsilon:
0.29700496000957016   steps: 1145   lr: 6.400000000000001e-06   evaluation
reward: 177.0
episode: 547   score: 220.0   memory length: 456030   epsilon:

0.29505862000955785 steps: 983 lr: 6.400000000000001e-06 evaluation reward: 177.7
episode: 548 score: 100.0 memory length: 456764 epsilon: 0.29360530000954865 steps: 734 lr: 6.400000000000001e-06 evaluation reward: 176.65
episode: 549 score: 325.0 memory length: 458261 epsilon: 0.2906412400095299 steps: 1497 lr: 6.400000000000001e-06 evaluation reward: 179.7
episode: 550 score: 190.0 memory length: 459088 epsilon: 0.28900378000951954 steps: 827 lr: 6.400000000000001e-06 evaluation reward: 180.6
episode: 551 score: 140.0 memory length: 459764 epsilon: 0.28766530000951107 steps: 676 lr: 6.400000000000001e-06 evaluation reward: 180.9
episode: 552 score: 175.0 memory length: 460677 epsilon: 0.28585756000949963 steps: 913 lr: 6.400000000000001e-06 evaluation reward: 181.35
episode: 553 score: 235.0 memory length: 461798 epsilon: 0.2836379800094856 steps: 1121 lr: 6.400000000000001e-06 evaluation reward: 181.65
episode: 554 score: 60.0 memory length: 462370 epsilon: 0.2825054200094784 steps: 572 lr: 6.400000000000001e-06 evaluation reward: 181.55
episode: 555 score: 110.0 memory length: 462881 epsilon: 0.281493640009472 steps: 511 lr: 6.400000000000001e-06 evaluation reward: 181.85
episode: 556 score: 130.0 memory length: 463446 epsilon: 0.28037494000946495 steps: 565 lr: 6.400000000000001e-06 evaluation reward: 181.25
episode: 557 score: 295.0 memory length: 464807 epsilon: 0.2776801600094479 steps: 1361 lr: 6.400000000000001e-06 evaluation reward: 183.9
episode: 558 score: 190.0 memory length: 465656 epsilon: 0.27599914000943726 steps: 849 lr: 6.400000000000001e-06 evaluation reward: 182.85
episode: 559 score: 140.0 memory length: 466781 epsilon: 0.27377164000942317 steps: 1125 lr: 6.400000000000001e-06 evaluation reward: 182.2
episode: 560 score: 70.0 memory length: 467147 epsilon: 0.2730469600094186 steps: 366 lr: 6.400000000000001e-06 evaluation reward: 181.0
episode: 561 score: 120.0 memory length: 467755 epsilon: 0.27184312000941097 steps: 608 lr: 6.400000000000001e-06 evaluation reward: 180.15
episode: 562 score: 160.0 memory length: 468416 epsilon: 0.2705343400094027 steps: 661 lr: 6.400000000000001e-06 evaluation reward: 179.55
episode: 563 score: 110.0 memory length: 468948 epsilon: 0.269480980009396 steps: 532 lr: 6.400000000000001e-06 evaluation reward: 179.85
episode: 564 score: 235.0 memory length: 469783 epsilon: 0.26782768000938556 steps: 835 lr: 6.400000000000001e-06 evaluation

reward: 179.4
episode: 565   score: 385.0   memory length: 471184   epsilon: 0.265053700009368
steps: 1401   lr: 6.400000000000001e-06   evaluation reward: 181.05
episode: 566   score: 120.0   memory length: 472216   epsilon:
0.2630103400093551   steps: 1032   lr: 6.400000000000001e-06   evaluation
reward: 181.15
episode: 567   score: 175.0   memory length: 473232   epsilon:
0.26099866000934235   steps: 1016   lr: 6.400000000000001e-06   evaluation
reward: 179.8
episode: 568   score: 140.0   memory length: 473889   epsilon:
0.2596978000093341   steps: 657   lr: 6.400000000000001e-06   evaluation
reward: 180.2
episode: 569   score: 130.0   memory length: 474453   epsilon:
0.25858108000932706   steps: 564   lr: 6.400000000000001e-06   evaluation
reward: 180.8
episode: 570   score: 265.0   memory length: 475446   epsilon:
0.2566149400093146   steps: 993   lr: 6.400000000000001e-06   evaluation
reward: 182.15
episode: 571   score: 370.0   memory length: 476591   epsilon:
0.2543478400093003   steps: 1145   lr: 6.400000000000001e-06   evaluation
reward: 185.25
episode: 572   score: 160.0   memory length: 477319   epsilon:
0.25290640000929115   steps: 728   lr: 6.400000000000001e-06   evaluation
reward: 186.35
episode: 573   score: 310.0   memory length: 478546   epsilon:
0.2504769400092758   steps: 1227   lr: 6.400000000000001e-06   evaluation
reward: 188.65
episode: 574   score: 120.0   memory length: 479225   epsilon:
0.24913252000926728   steps: 679   lr: 6.400000000000001e-06   evaluation
reward: 187.95
episode: 575   score: 190.0   memory length: 480078   epsilon:
0.2474435800092566   steps: 853   lr: 6.400000000000001e-06   evaluation
reward: 187.2
episode: 576   score: 90.0   memory length: 480514   epsilon:
0.24658030000925113   steps: 436   lr: 6.400000000000001e-06   evaluation
reward: 186.5
episode: 577   score: 60.0   memory length: 480852   epsilon: 0.2459110600092469
steps: 338   lr: 6.400000000000001e-06   evaluation reward: 186.0
episode: 578   score: 325.0   memory length: 482360   epsilon: 0.242925220009228
steps: 1508   lr: 6.400000000000001e-06   evaluation reward: 186.0
episode: 579   score: 130.0   memory length: 482960   epsilon:
0.2417372200092205   steps: 600   lr: 6.400000000000001e-06   evaluation
reward: 186.2
episode: 580   score: 40.0   memory length: 483341   epsilon:
0.24098284000921572   steps: 381   lr: 6.400000000000001e-06   evaluation
reward: 185.1
episode: 581   score: 160.0   memory length: 484159   epsilon:
0.23936320000920547   steps: 818   lr: 6.400000000000001e-06   evaluation

```
reward: 184.95
episode: 582   score: 60.0   memory length: 484684   epsilon: 0.2383237000091989
steps: 525    lr: 6.400000000000001e-06     evaluation reward: 182.9
episode: 583   score: 295.0   memory length: 485947   epsilon:
0.23582296000918307    steps: 1263    lr: 6.400000000000001e-06     evaluation
reward: 182.6
episode: 584   score: 100.0   memory length: 486389   epsilon:
0.23494780000917753    steps: 442    lr: 6.400000000000001e-06     evaluation
reward: 183.2
episode: 585   score: 235.0   memory length: 487433   epsilon:
0.23288068000916445    steps: 1044    lr: 6.400000000000001e-06     evaluation
reward: 184.35
episode: 586   score: 130.0   memory length: 488155   epsilon:
0.2314511200091554    steps: 722    lr: 6.400000000000001e-06     evaluation
reward: 182.1
episode: 587   score: 60.0   memory length: 488616   epsilon:
0.23053834000914963    steps: 461    lr: 6.400000000000001e-06     evaluation
reward: 180.95
episode: 588   score: 150.0   memory length: 489342   epsilon:
0.22910086000914054    steps: 726    lr: 6.400000000000001e-06     evaluation
reward: 179.35
episode: 589   score: 120.0   memory length: 489903   epsilon:
0.2279900800091335    steps: 561    lr: 6.400000000000001e-06     evaluation
reward: 179.65
episode: 590   score: 120.0   memory length: 490388   epsilon:
0.22702978000912744    steps: 485    lr: 6.400000000000001e-06     evaluation
reward: 179.35
episode: 591   score: 175.0   memory length: 491085   epsilon:
0.2256497200091187    steps: 697    lr: 6.400000000000001e-06     evaluation
reward: 180.1
episode: 592   score: 140.0   memory length: 492199   epsilon:
0.22344400000910475    steps: 1114    lr: 6.400000000000001e-06     evaluation
reward: 179.9
episode: 593   score: 175.0   memory length: 493094   epsilon:
0.22167190000909354    steps: 895    lr: 6.400000000000001e-06     evaluation
reward: 180.35
episode: 594   score: 80.0   memory length: 493550   epsilon:
0.22076902000908782    steps: 456    lr: 6.400000000000001e-06     evaluation
reward: 179.25
episode: 595   score: 20.0   memory length: 493804   epsilon:
0.22026610000908464    steps: 254    lr: 6.400000000000001e-06     evaluation
reward: 177.95
episode: 596   score: 130.0   memory length: 494305   epsilon:
0.21927412000907837    steps: 501    lr: 6.400000000000001e-06     evaluation
reward: 178.15
episode: 597   score: 100.0   memory length: 495027   epsilon:
0.21784456000906932    steps: 722    lr: 6.400000000000001e-06     evaluation
reward: 177.55
```

episode: 598   score: 60.0   memory length: 495451   epsilon: 0.217005040009064
steps: 424   lr: 6.400000000000001e-06   evaluation reward: 176.75
episode: 599   score: 120.0   memory length: 496101   epsilon:
0.21571804000905587   steps: 650   lr: 6.400000000000001e-06   evaluation
reward: 177.05
episode: 600   score: 340.0   memory length: 497482   epsilon:
0.21298366000903857   steps: 1381   lr: 6.400000000000001e-06   evaluation
reward: 178.55
episode: 601   score: 400.0   memory length: 498664   epsilon:
0.21064330000902376   steps: 1182   lr: 6.400000000000001e-06   evaluation
reward: 181.95
episode: 602   score: 40.0   memory length: 498951   epsilon:
0.21007504000902016   steps: 287   lr: 6.400000000000001e-06   evaluation
reward: 180.15
episode: 603   score: 100.0   memory length: 499737   epsilon:
0.20851876000901032   steps: 786   lr: 6.400000000000001e-06   evaluation
reward: 178.95
episode: 604   score: 310.0   memory length: 501358   epsilon: 0.20530918000899
steps: 1621   lr: 2.560000000000001e-06   evaluation reward: 179.25
episode: 605   score: 120.0   memory length: 502098   epsilon:
0.20384398000898074   steps: 740   lr: 2.560000000000001e-06   evaluation
reward: 179.45
episode: 606   score: 60.0   memory length: 502370   epsilon:
0.20330542000897733   steps: 272   lr: 2.560000000000001e-06   evaluation
reward: 177.4
episode: 607   score: 80.0   memory length: 502731   epsilon: 0.2025906400089728
steps: 361   lr: 2.560000000000001e-06   evaluation reward: 177.6
episode: 608   score: 310.0   memory length: 504004   epsilon:
0.20007010000895686   steps: 1273   lr: 2.560000000000001e-06   evaluation
reward: 178.2
episode: 609   score: 175.0   memory length: 505084   epsilon:
0.19793170000894333   steps: 1080   lr: 2.560000000000001e-06   evaluation
reward: 178.2
episode: 610   score: 140.0   memory length: 505680   epsilon:
0.19675162000893587   steps: 596   lr: 2.560000000000001e-06   evaluation
reward: 179.0
episode: 611   score: 190.0   memory length: 506275   epsilon:
0.19557352000892841   steps: 595   lr: 2.560000000000001e-06   evaluation
reward: 178.7
episode: 612   score: 370.0   memory length: 507505   epsilon: 0.193138120008913
steps: 1230   lr: 2.560000000000001e-06   evaluation reward: 178.55
episode: 613   score: 235.0   memory length: 508552   epsilon:
0.1910650600088999   steps: 1047   lr: 2.560000000000001e-06   evaluation
reward: 178.55
episode: 614   score: 175.0   memory length: 509197   epsilon:
0.1897879600088918   steps: 645   lr: 2.560000000000001e-06   evaluation
reward: 179.2
episode: 615   score: 175.0   memory length: 510037   epsilon:

0.1881247600088813 steps: 840 lr: 2.560000000000001e-06 evaluation reward: 177.85
episode: 616 score: 175.0 memory length: 510885 epsilon: 0.18644572000887066 steps: 848 lr: 2.560000000000001e-06 evaluation reward: 176.5
episode: 617 score: 160.0 memory length: 511549 epsilon: 0.18513100000886235 steps: 664 lr: 2.560000000000001e-06 evaluation reward: 175.9
episode: 618 score: 190.0 memory length: 512473 epsilon: 0.18330148000885077 steps: 924 lr: 2.560000000000001e-06 evaluation reward: 173.2
episode: 619 score: 160.0 memory length: 513302 epsilon: 0.18166006000884038 steps: 829 lr: 2.560000000000001e-06 evaluation reward: 171.85
episode: 620 score: 30.0 memory length: 513642 epsilon: 0.18098686000883613 steps: 340 lr: 2.560000000000001e-06 evaluation reward: 171.55
episode: 621 score: 355.0 memory length: 514999 epsilon: 0.17830000000881913 steps: 1357 lr: 2.560000000000001e-06 evaluation reward: 174.6
episode: 622 score: 90.0 memory length: 515511 epsilon: 0.1772862400088127 steps: 512 lr: 2.560000000000001e-06 evaluation reward: 173.75
episode: 623 score: 40.0 memory length: 515815 epsilon: 0.1766843200088089 steps: 304 lr: 2.560000000000001e-06 evaluation reward: 170.15
episode: 624 score: 190.0 memory length: 516654 epsilon: 0.1750231000087984 steps: 839 lr: 2.560000000000001e-06 evaluation reward: 171.15
episode: 625 score: 140.0 memory length: 517300 epsilon: 0.1737440200087903 steps: 646 lr: 2.560000000000001e-06 evaluation reward: 171.05
episode: 626 score: 120.0 memory length: 517945 epsilon: 0.17246692000878222 steps: 645 lr: 2.560000000000001e-06 evaluation reward: 171.35
episode: 627 score: 265.0 memory length: 518921 epsilon: 0.17053444000877 steps: 976 lr: 2.560000000000001e-06 evaluation reward: 173.2
episode: 628 score: 160.0 memory length: 519630 epsilon: 0.1691306200087611 steps: 709 lr: 2.560000000000001e-06 evaluation reward: 173.3
episode: 629 score: 175.0 memory length: 520553 epsilon: 0.16730308000874955 steps: 923 lr: 2.560000000000001e-06 evaluation reward: 173.85
episode: 630 score: 175.0 memory length: 521656 epsilon: 0.16511914000873573 steps: 1103 lr: 2.560000000000001e-06 evaluation reward: 175.3
episode: 631 score: 130.0 memory length: 522213 epsilon: 0.16401628000872875 steps: 557 lr: 2.560000000000001e-06 evaluation reward: 174.25
episode: 632 score: 220.0 memory length: 523222 epsilon:

0.16201846000871611   steps: 1009   lr: 2.560000000000001e-06   evaluation
reward: 175.35
episode: 633   score: 70.0   memory length: 523651   epsilon:
0.16116904000871074   steps: 429   lr: 2.560000000000001e-06   evaluation
reward: 173.4
episode: 634   score: 250.0   memory length: 524852   epsilon:
0.1587910600086957   steps: 1201   lr: 2.560000000000001e-06   evaluation
reward: 174.6
episode: 635   score: 340.0   memory length: 525936   epsilon:
0.15664474000868212   steps: 1084   lr: 2.560000000000001e-06   evaluation
reward: 177.1
episode: 636   score: 160.0   memory length: 526707   epsilon:
0.15511816000867246   steps: 771   lr: 2.560000000000001e-06   evaluation
reward: 177.5
episode: 637   score: 280.0   memory length: 527684   epsilon:
0.15318370000866022   steps: 977   lr: 2.560000000000001e-06   evaluation
reward: 179.5
episode: 638   score: 140.0   memory length: 528544   epsilon:
0.15148090000864944   steps: 860   lr: 2.560000000000001e-06   evaluation
reward: 179.0
episode: 639   score: 90.0   memory length: 529003   epsilon: 0.1505720800086437
steps: 459   lr: 2.560000000000001e-06   evaluation reward: 177.7
episode: 640   score: 70.0   memory length: 529455   epsilon:
0.14967712000863803   steps: 452   lr: 2.560000000000001e-06   evaluation
reward: 178.1
episode: 641   score: 130.0   memory length: 530110   epsilon:
0.14838022000862983   steps: 655   lr: 2.560000000000001e-06   evaluation
reward: 177.65
episode: 642   score: 110.0   memory length: 530686   epsilon:
0.1472397400086226   steps: 576   lr: 2.560000000000001e-06   evaluation
reward: 175.8
episode: 643   score: 120.0   memory length: 531098   epsilon:
0.14642398000861745   steps: 412   lr: 2.560000000000001e-06   evaluation
reward: 167.4
episode: 644   score: 205.0   memory length: 532042   epsilon:
0.14455486000860562   steps: 944   lr: 2.560000000000001e-06   evaluation
reward: 167.4
episode: 645   score: 265.0   memory length: 532767   epsilon:
0.14311936000859654   steps: 725   lr: 2.560000000000001e-06   evaluation
reward: 169.05
episode: 646   score: 325.0   memory length: 533938   epsilon:
0.14080078000858187   steps: 1171   lr: 2.560000000000001e-06   evaluation
reward: 169.95
episode: 647   score: 160.0   memory length: 534808   epsilon:
0.13907818000857097   steps: 870   lr: 2.560000000000001e-06   evaluation
reward: 169.35
episode: 648   score: 340.0   memory length: 535820   epsilon:
0.1370744200085583   steps: 1012   lr: 2.560000000000001e-06   evaluation

```
reward: 171.75
episode: 649   score: 50.0   memory length: 536248   epsilon:
0.13622698000855293    steps: 428    lr: 2.560000000000001e-06     evaluation
reward: 169.0
episode: 650   score: 235.0   memory length: 537240   epsilon:
0.1342628200085405    steps: 992    lr: 2.560000000000001e-06     evaluation
reward: 169.45
episode: 651   score: 280.0   memory length: 538705   epsilon:
0.13136212000852215    steps: 1465    lr: 2.560000000000001e-06     evaluation
reward: 170.85
episode: 652   score: 250.0   memory length: 539716   epsilon:
0.1293603400085095    steps: 1011    lr: 2.560000000000001e-06     evaluation
reward: 171.6
episode: 653   score: 40.0   memory length: 540007   epsilon:
0.12878416000850584    steps: 291    lr: 2.560000000000001e-06     evaluation
reward: 169.65
episode: 654   score: 60.0   memory length: 540325   epsilon:
0.12815452000850186    steps: 318    lr: 2.560000000000001e-06     evaluation
reward: 169.65
episode: 655   score: 295.0   memory length: 541486   epsilon:
0.12585574000848732    steps: 1161    lr: 2.560000000000001e-06     evaluation
reward: 171.5
episode: 656   score: 80.0   memory length: 541944   epsilon:
0.12494890000848194    steps: 458    lr: 2.560000000000001e-06     evaluation
reward: 171.0
episode: 657   score: 150.0   memory length: 542532   epsilon:
0.12378466000848273    steps: 588    lr: 2.560000000000001e-06     evaluation
reward: 169.55
episode: 658   score: 190.0   memory length: 543368   epsilon:
0.12212938000848386    steps: 836    lr: 2.560000000000001e-06     evaluation
reward: 169.55
episode: 659   score: 250.0   memory length: 544535   epsilon:
0.11981872000848544    steps: 1167    lr: 2.560000000000001e-06     evaluation
reward: 170.65
episode: 660   score: 235.0   memory length: 545579   epsilon:
0.11775160000848685    steps: 1044    lr: 2.560000000000001e-06     evaluation
reward: 172.3
episode: 661   score: 190.0   memory length: 546258   epsilon:
0.11640718000848777    steps: 679    lr: 2.560000000000001e-06     evaluation
reward: 173.0
episode: 662   score: 205.0   memory length: 547053   epsilon:
0.11483308000848884    steps: 795    lr: 2.560000000000001e-06     evaluation
reward: 173.45
episode: 663   score: 140.0   memory length: 547731   epsilon:
0.11349064000848975    steps: 678    lr: 2.560000000000001e-06     evaluation
reward: 173.75
episode: 664   score: 120.0   memory length: 548375   epsilon:
0.11221552000849062    steps: 644    lr: 2.560000000000001e-06     evaluation
```

```
reward: 172.6
episode: 665   score: 110.0   memory length: 548935   epsilon:
0.11110672000849138    steps: 560    lr: 2.560000000000001e-06     evaluation
reward: 169.85
episode: 666   score: 90.0   memory length: 549634   epsilon:
0.10972270000849232    steps: 699    lr: 2.560000000000001e-06     evaluation
reward: 169.55
episode: 667   score: 40.0   memory length: 549854   epsilon:
0.10928710000849262    steps: 220    lr: 2.560000000000001e-06     evaluation
reward: 168.2
episode: 668   score: 235.0   memory length: 550663   epsilon:
0.10768528000849371    steps: 809    lr: 2.560000000000001e-06     evaluation
reward: 169.15
episode: 669   score: 70.0   memory length: 551059   epsilon:
0.10690120000849425    steps: 396    lr: 2.560000000000001e-06     evaluation
reward: 168.55
episode: 670   score: 140.0   memory length: 551576   epsilon:
0.10587754000849495    steps: 517    lr: 2.560000000000001e-06     evaluation
reward: 167.3
episode: 671   score: 325.0   memory length: 552955   epsilon:
0.10314712000849681    steps: 1379    lr: 2.560000000000001e-06     evaluation
reward: 166.85
episode: 672   score: 100.0   memory length: 553505   epsilon:
0.10205812000849755    steps: 550    lr: 2.560000000000001e-06     evaluation
reward: 166.25
episode: 673   score: 175.0   memory length: 554263   epsilon:
0.10055728000849858    steps: 758    lr: 2.560000000000001e-06     evaluation
reward: 164.9
episode: 674   score: 190.0   memory length: 555176   epsilon:
0.09874954000849981    steps: 913    lr: 2.560000000000001e-06     evaluation
reward: 165.6
episode: 675   score: 190.0   memory length: 555999   epsilon:
0.09712000000850092    steps: 823    lr: 2.560000000000001e-06     evaluation
reward: 165.6
episode: 676   score: 250.0   memory length: 557219   epsilon:
0.09470440000850257    steps: 1220    lr: 2.560000000000001e-06     evaluation
reward: 167.2
episode: 677   score: 130.0   memory length: 557913   epsilon:
0.0933302800085035    steps: 694    lr: 2.560000000000001e-06     evaluation
reward: 167.9
episode: 678   score: 325.0   memory length: 559119   epsilon:
0.09094240000850513    steps: 1206    lr: 2.560000000000001e-06     evaluation
reward: 167.9
episode: 679   score: 175.0   memory length: 559968   epsilon:
0.08926138000850628    steps: 849    lr: 2.560000000000001e-06     evaluation
reward: 168.35
episode: 680   score: 175.0   memory length: 560832   epsilon:
0.08755066000850745    steps: 864    lr: 2.560000000000001e-06     evaluation
```

reward: 169.7
episode: 681   score: 265.0   memory length: 562021   epsilon:
0.08519644000850905   steps: 1189   lr: 2.560000000000001e-06   evaluation
reward: 170.75
episode: 682   score: 130.0   memory length: 562531   epsilon:
0.08418664000850974   steps: 510   lr: 2.560000000000001e-06   evaluation
reward: 171.45
episode: 683   score: 60.0   memory length: 562841   epsilon:
0.08357284000851016   steps: 310   lr: 2.560000000000001e-06   evaluation
reward: 169.1
episode: 684   score: 110.0   memory length: 563360   epsilon:
0.08254522000851086   steps: 519   lr: 2.560000000000001e-06   evaluation
reward: 169.2
episode: 685   score: 355.0   memory length: 564350   epsilon:
0.0805850200085122   steps: 990   lr: 2.560000000000001e-06   evaluation
reward: 170.4
episode: 686   score: 295.0   memory length: 565471   epsilon:
0.07836544000851371   steps: 1121   lr: 2.560000000000001e-06   evaluation
reward: 172.05
episode: 687   score: 110.0   memory length: 566080   epsilon:
0.07715962000851453   steps: 609   lr: 2.560000000000001e-06   evaluation
reward: 172.55
episode: 688   score: 400.0   memory length: 567271   epsilon:
0.07480144000851614   steps: 1191   lr: 2.560000000000001e-06   evaluation
reward: 175.05
episode: 689   score: 120.0   memory length: 567928   epsilon:
0.07350058000851703   steps: 657   lr: 2.560000000000001e-06   evaluation
reward: 175.05
episode: 690   score: 60.0   memory length: 568249   epsilon:
0.07286500000851746   steps: 321   lr: 2.560000000000001e-06   evaluation
reward: 174.45
episode: 691   score: 340.0   memory length: 569487   epsilon:
0.07041376000851914   steps: 1238   lr: 2.560000000000001e-06   evaluation
reward: 176.1
episode: 692   score: 100.0   memory length: 570071   epsilon:
0.06925744000851992   steps: 584   lr: 2.560000000000001e-06   evaluation
reward: 175.7
episode: 693   score: 385.0   memory length: 571380   epsilon:
0.06666562000852169   steps: 1309   lr: 2.560000000000001e-06   evaluation
reward: 177.8
episode: 694   score: 100.0   memory length: 571796   epsilon:
0.06584194000852225   steps: 416   lr: 2.560000000000001e-06   evaluation
reward: 178.0
episode: 695   score: 130.0   memory length: 572358   epsilon:
0.06472918000852301   steps: 562   lr: 2.560000000000001e-06   evaluation
reward: 179.1
episode: 696   score: 130.0   memory length: 572826   epsilon:
0.06380254000852364   steps: 468   lr: 2.560000000000001e-06   evaluation

```
reward: 179.1
episode: 697   score: 355.0   memory length: 573773   epsilon:
0.061927480008524924    steps: 947    lr: 2.560000000000001e-06     evaluation
reward: 181.65
episode: 698   score: 120.0   memory length: 574339   epsilon:
0.06080680000852569    steps: 566    lr: 2.560000000000001e-06     evaluation
reward: 182.25
episode: 699   score: 110.0   memory length: 574839   epsilon:
0.059816800008526363    steps: 500    lr: 2.560000000000001e-06     evaluation
reward: 182.15
episode: 700   score: 140.0   memory length: 575363   epsilon:
0.05877928000852707    steps: 524    lr: 2.560000000000001e-06     evaluation
reward: 180.15
episode: 701   score: 150.0   memory length: 575880   epsilon:
0.05775562000852777    steps: 517    lr: 2.560000000000001e-06     evaluation
reward: 177.65
episode: 702   score: 250.0   memory length: 576703   epsilon:
0.05612608000852888    steps: 823    lr: 2.560000000000001e-06     evaluation
reward: 179.75
episode: 703   score: 90.0   memory length: 577191   epsilon:
0.05515984000852954    steps: 488    lr: 2.560000000000001e-06     evaluation
reward: 179.65
episode: 704   score: 355.0   memory length: 578349   epsilon:
0.052867000008531104    steps: 1158    lr: 2.560000000000001e-06     evaluation
reward: 180.1
episode: 705   score: 190.0   memory length: 579200   epsilon:
0.05118202000853225    steps: 851    lr: 2.560000000000001e-06     evaluation
reward: 180.8
episode: 706   score: 325.0   memory length: 580477   epsilon:
0.04865356000853398    steps: 1277    lr: 2.560000000000001e-06     evaluation
reward: 183.45
episode: 707   score: 140.0   memory length: 581127   epsilon:
0.047366560008534855    steps: 650    lr: 2.560000000000001e-06     evaluation
reward: 184.05
episode: 708   score: 310.0   memory length: 582303   epsilon:
0.045038080008536444    steps: 1176    lr: 2.560000000000001e-06     evaluation
reward: 184.05
episode: 709   score: 205.0   memory length: 583100   epsilon:
0.04346002000853752    steps: 797    lr: 2.560000000000001e-06     evaluation
reward: 184.35
episode: 710   score: 140.0   memory length: 583727   epsilon:
0.04221856000853837    steps: 627    lr: 2.560000000000001e-06     evaluation
reward: 184.35
episode: 711   score: 265.0   memory length: 584670   epsilon:
0.04035142000853964    steps: 943    lr: 2.560000000000001e-06     evaluation
reward: 185.1
episode: 712   score: 190.0   memory length: 585295   epsilon:
0.039113920008540484    steps: 625    lr: 2.560000000000001e-06     evaluation
```

```
reward: 183.3
episode: 713   score: 190.0   memory length: 586004   epsilon:
0.03771010000854144    steps: 709    lr: 2.560000000000001e-06     evaluation
reward: 182.85
episode: 714   score: 205.0   memory length: 586775   epsilon:
0.03618352000854248    steps: 771    lr: 2.560000000000001e-06     evaluation
reward: 183.15
episode: 715   score: 220.0   memory length: 587592   epsilon:
0.034565860008543586    steps: 817    lr: 2.560000000000001e-06     evaluation
reward: 183.6
episode: 716   score: 160.0   memory length: 588256   epsilon:
0.03325114000854448    steps: 664    lr: 2.560000000000001e-06     evaluation
reward: 183.45
episode: 717   score: 100.0   memory length: 588807   epsilon:
0.03216016000854523    steps: 551    lr: 2.560000000000001e-06     evaluation
reward: 182.85
episode: 718   score: 130.0   memory length: 589385   epsilon:
0.031015720008546008    steps: 578    lr: 2.560000000000001e-06     evaluation
reward: 182.25
episode: 719   score: 150.0   memory length: 589885   epsilon:
0.030025720008546683    steps: 500    lr: 2.560000000000001e-06     evaluation
reward: 182.15
episode: 720   score: 265.0   memory length: 590906   epsilon:
0.028004140008548062    steps: 1021    lr: 2.560000000000001e-06     evaluation
reward: 184.5
episode: 721   score: 40.0   memory length: 591107   epsilon:
0.027606160008548333    steps: 201    lr: 2.560000000000001e-06     evaluation
reward: 181.35
episode: 722   score: 90.0   memory length: 591822   epsilon: 0.0261904600085493
steps: 715    lr: 2.560000000000001e-06     evaluation reward: 181.35
episode: 723   score: 70.0   memory length: 592172   epsilon:
0.02549746000854977    steps: 350    lr: 2.560000000000001e-06     evaluation
reward: 181.65
episode: 724   score: 190.0   memory length: 593120   epsilon:
0.023620420008551052    steps: 948    lr: 2.560000000000001e-06     evaluation
reward: 181.65
episode: 725   score: 60.0   memory length: 593508   epsilon:
0.022852180008551576    steps: 388    lr: 2.560000000000001e-06     evaluation
reward: 180.85
episode: 726   score: 50.0   memory length: 593814   epsilon:
0.02224630000855199    steps: 306    lr: 2.560000000000001e-06     evaluation
reward: 180.15
episode: 727   score: 140.0   memory length: 594510   epsilon:
0.02086822000855293    steps: 696    lr: 2.560000000000001e-06     evaluation
reward: 178.9
episode: 728   score: 190.0   memory length: 595464   epsilon:
0.018979300008554217    steps: 954    lr: 2.560000000000001e-06     evaluation
reward: 179.2
```

episode: 729   score: 120.0   memory length: 595982   epsilon:
0.017953660008554917   steps: 518   lr: 2.560000000000001e-06   evaluation
reward: 178.65
episode: 730   score: 310.0   memory length: 597598   epsilon:
0.014753980008556336   steps: 1616   lr: 2.560000000000001e-06   evaluation
reward: 180.0
episode: 731   score: 110.0   memory length: 598085   epsilon:
0.013789720008556149   steps: 487   lr: 2.560000000000001e-06   evaluation
reward: 179.8
episode: 732   score: 150.0   memory length: 598902   epsilon:
0.012172060008555835   steps: 817   lr: 2.560000000000001e-06   evaluation
reward: 179.1
episode: 733   score: 60.0   memory length: 599199   epsilon:
0.011584000008555721   steps: 297   lr: 2.560000000000001e-06   evaluation
reward: 179.0
episode: 734   score: 90.0   memory length: 599639   epsilon:
0.010712800008555552   steps: 440   lr: 2.560000000000001e-06   evaluation
reward: 177.4
episode: 735   score: 110.0   memory length: 600301   epsilon:
0.009998020008555413   steps: 662   lr: 1.0240000000000005e-06   evaluation
reward: 175.1
episode: 736   score: 310.0   memory length: 601359   epsilon:
0.009998020008555413   steps: 1058   lr: 1.0240000000000005e-06   evaluation
reward: 176.6
episode: 737   score: 50.0   memory length: 601668   epsilon:
0.009998020008555413   steps: 309   lr: 1.0240000000000005e-06   evaluation
reward: 174.3
episode: 738   score: 280.0   memory length: 602738   epsilon:
0.009998020008555413   steps: 1070   lr: 1.0240000000000005e-06   evaluation
reward: 175.7
episode: 739   score: 265.0   memory length: 603583   epsilon:
0.009998020008555413   steps: 845   lr: 1.0240000000000005e-06   evaluation
reward: 177.45
episode: 740   score: 190.0   memory length: 604501   epsilon:
0.009998020008555413   steps: 918   lr: 1.0240000000000005e-06   evaluation
reward: 178.65
episode: 741   score: 80.0   memory length: 605063   epsilon:
0.009998020008555413   steps: 562   lr: 1.0240000000000005e-06   evaluation
reward: 178.15
episode: 742   score: 310.0   memory length: 605943   epsilon:
0.009998020008555413   steps: 880   lr: 1.0240000000000005e-06   evaluation
reward: 180.15
episode: 743   score: 190.0   memory length: 607097   epsilon:
0.009998020008555413   steps: 1154   lr: 1.0240000000000005e-06   evaluation
reward: 180.85
episode: 744   score: 250.0   memory length: 607927   epsilon:
0.009998020008555413   steps: 830   lr: 1.0240000000000005e-06   evaluation
reward: 181.3

```
episode: 745   score: 235.0   memory length: 608824   epsilon:
0.009998020008555413    steps: 897    lr: 1.0240000000000005e-06     evaluation
reward: 181.0
episode: 746   score: 190.0   memory length: 609569   epsilon:
0.009998020008555413    steps: 745    lr: 1.0240000000000005e-06     evaluation
reward: 179.65
episode: 747   score: 235.0   memory length: 610338   epsilon:
0.009998020008555413    steps: 769    lr: 1.0240000000000005e-06     evaluation
reward: 180.4
episode: 748   score: 190.0   memory length: 611199   epsilon:
0.009998020008555413    steps: 861    lr: 1.0240000000000005e-06     evaluation
reward: 178.9
episode: 749   score: 175.0   memory length: 612099   epsilon:
0.009998020008555413    steps: 900    lr: 1.0240000000000005e-06     evaluation
reward: 180.15
episode: 750   score: 80.0   memory length: 612557   epsilon:
0.009998020008555413    steps: 458    lr: 1.0240000000000005e-06     evaluation
reward: 178.6
episode: 751   score: 340.0   memory length: 613791   epsilon:
0.009998020008555413    steps: 1234    lr: 1.0240000000000005e-06     evaluation
reward: 179.2
episode: 752   score: 205.0   memory length: 614610   epsilon:
0.009998020008555413    steps: 819    lr: 1.0240000000000005e-06     evaluation
reward: 178.75
episode: 753   score: 40.0   memory length: 614913   epsilon:
0.009998020008555413    steps: 303    lr: 1.0240000000000005e-06     evaluation
reward: 178.75
episode: 754   score: 325.0   memory length: 616156   epsilon:
0.009998020008555413    steps: 1243    lr: 1.0240000000000005e-06     evaluation
reward: 181.4
episode: 755   score: 110.0   memory length: 616639   epsilon:
0.009998020008555413    steps: 483    lr: 1.0240000000000005e-06     evaluation
reward: 179.55
episode: 756   score: 295.0   memory length: 617491   epsilon:
0.009998020008555413    steps: 852    lr: 1.0240000000000005e-06     evaluation
reward: 181.7
episode: 757   score: 130.0   memory length: 618073   epsilon:
0.009998020008555413    steps: 582    lr: 1.0240000000000005e-06     evaluation
reward: 181.5
episode: 758   score: 130.0   memory length: 618618   epsilon:
0.009998020008555413    steps: 545    lr: 1.0240000000000005e-06     evaluation
reward: 180.9
episode: 759   score: 190.0   memory length: 619574   epsilon:
0.009998020008555413    steps: 956    lr: 1.0240000000000005e-06     evaluation
reward: 180.3
episode: 760   score: 235.0   memory length: 620339   epsilon:
0.009998020008555413    steps: 765    lr: 1.0240000000000005e-06     evaluation
reward: 180.3
```

```
episode: 761   score: 250.0   memory length: 621307   epsilon:
0.009998020008555413    steps: 968    lr: 1.0240000000000005e-06     evaluation
reward: 180.9
episode: 762   score: 250.0   memory length: 622235   epsilon:
0.009998020008555413    steps: 928    lr: 1.0240000000000005e-06     evaluation
reward: 181.35
episode: 763   score: 140.0   memory length: 622870   epsilon:
0.009998020008555413    steps: 635    lr: 1.0240000000000005e-06     evaluation
reward: 181.35
episode: 764   score: 90.0   memory length: 623233   epsilon:
0.009998020008555413    steps: 363    lr: 1.0240000000000005e-06     evaluation
reward: 181.05
episode: 765   score: 250.0   memory length: 624265   epsilon:
0.009998020008555413    steps: 1032    lr: 1.0240000000000005e-06     evaluation
reward: 182.45
episode: 766   score: 90.0   memory length: 624645   epsilon:
0.009998020008555413    steps: 380    lr: 1.0240000000000005e-06     evaluation
reward: 182.45
episode: 767   score: 80.0   memory length: 625025   epsilon:
0.009998020008555413    steps: 380    lr: 1.0240000000000005e-06     evaluation
reward: 182.85
episode: 768   score: 280.0   memory length: 626043   epsilon:
0.009998020008555413    steps: 1018    lr: 1.0240000000000005e-06     evaluation
reward: 183.3
episode: 769   score: 190.0   memory length: 626565   epsilon:
0.009998020008555413    steps: 522    lr: 1.0240000000000005e-06     evaluation
reward: 184.5
episode: 770   score: 70.0   memory length: 626941   epsilon:
0.009998020008555413    steps: 376    lr: 1.0240000000000005e-06     evaluation
reward: 183.8
episode: 771   score: 150.0   memory length: 627613   epsilon:
0.009998020008555413    steps: 672    lr: 1.0240000000000005e-06     evaluation
reward: 182.05
episode: 772   score: 90.0   memory length: 628064   epsilon:
0.009998020008555413    steps: 451    lr: 1.0240000000000005e-06     evaluation
reward: 181.95
episode: 773   score: 30.0   memory length: 628284   epsilon:
0.009998020008555413    steps: 220    lr: 1.0240000000000005e-06     evaluation
reward: 180.5
episode: 774   score: 150.0   memory length: 628974   epsilon:
0.009998020008555413    steps: 690    lr: 1.0240000000000005e-06     evaluation
reward: 180.1
episode: 775   score: 110.0   memory length: 629421   epsilon:
0.009998020008555413    steps: 447    lr: 1.0240000000000005e-06     evaluation
reward: 179.3
episode: 776   score: 110.0   memory length: 629842   epsilon:
0.009998020008555413    steps: 421    lr: 1.0240000000000005e-06     evaluation
reward: 177.9
```

```
episode: 777   score: 205.0   memory length: 630694   epsilon:
0.009998020008555413    steps: 852    lr: 1.0240000000000005e-06     evaluation
reward: 178.65
episode: 778   score: 250.0   memory length: 631455   epsilon:
0.009998020008555413    steps: 761    lr: 1.0240000000000005e-06     evaluation
reward: 177.9
episode: 779   score: 220.0   memory length: 632368   epsilon:
0.009998020008555413    steps: 913    lr: 1.0240000000000005e-06     evaluation
reward: 178.35
episode: 780   score: 100.0   memory length: 632880   epsilon:
0.009998020008555413    steps: 512    lr: 1.0240000000000005e-06     evaluation
reward: 177.6
episode: 781   score: 100.0   memory length: 633454   epsilon:
0.009998020008555413    steps: 574    lr: 1.0240000000000005e-06     evaluation
reward: 175.95
episode: 782   score: 130.0   memory length: 633971   epsilon:
0.009998020008555413    steps: 517    lr: 1.0240000000000005e-06     evaluation
reward: 175.95
episode: 783   score: 160.0   memory length: 634746   epsilon:
0.009998020008555413    steps: 775    lr: 1.0240000000000005e-06     evaluation
reward: 176.95
episode: 784   score: 220.0   memory length: 635494   epsilon:
0.009998020008555413    steps: 748    lr: 1.0240000000000005e-06     evaluation
reward: 178.05
episode: 785   score: 150.0   memory length: 636117   epsilon:
0.009998020008555413    steps: 623    lr: 1.0240000000000005e-06     evaluation
reward: 176.0
episode: 786   score: 120.0   memory length: 636666   epsilon:
0.009998020008555413    steps: 549    lr: 1.0240000000000005e-06     evaluation
reward: 174.25
episode: 787   score: 160.0   memory length: 637425   epsilon:
0.009998020008555413    steps: 759    lr: 1.0240000000000005e-06     evaluation
reward: 174.75
episode: 788   score: 560.0   memory length: 638664   epsilon:
0.009998020008555413    steps: 1239    lr: 1.0240000000000005e-06     evaluation
reward: 176.35
episode: 789   score: 110.0   memory length: 639110   epsilon:
0.009998020008555413    steps: 446    lr: 1.0240000000000005e-06     evaluation
reward: 176.25
episode: 790   score: 40.0   memory length: 639333   epsilon:
0.009998020008555413    steps: 223    lr: 1.0240000000000005e-06     evaluation
reward: 176.05
episode: 791   score: 150.0   memory length: 639973   epsilon:
0.009998020008555413    steps: 640    lr: 1.0240000000000005e-06     evaluation
reward: 174.15
episode: 792   score: 120.0   memory length: 640674   epsilon:
0.009998020008555413    steps: 701    lr: 1.0240000000000005e-06     evaluation
reward: 174.35
```

```
episode: 793   score: 140.0   memory length: 641323   epsilon:
0.009998020008555413    steps: 649    lr: 1.0240000000000005e-06     evaluation
reward: 171.9
episode: 794   score: 265.0   memory length: 642337   epsilon:
0.009998020008555413    steps: 1014    lr: 1.0240000000000005e-06     evaluation
reward: 173.55
episode: 795   score: 325.0   memory length: 643490   epsilon:
0.009998020008555413    steps: 1153    lr: 1.0240000000000005e-06     evaluation
reward: 175.5
episode: 796   score: 150.0   memory length: 644121   epsilon:
0.009998020008555413    steps: 631    lr: 1.0240000000000005e-06     evaluation
reward: 175.7
episode: 797   score: 100.0   memory length: 644638   epsilon:
0.009998020008555413    steps: 517    lr: 1.0240000000000005e-06     evaluation
reward: 173.15
episode: 798   score: 295.0   memory length: 646028   epsilon:
0.009998020008555413    steps: 1390    lr: 1.0240000000000005e-06     evaluation
reward: 174.9
episode: 799   score: 160.0   memory length: 646933   epsilon:
0.009998020008555413    steps: 905    lr: 1.0240000000000005e-06     evaluation
reward: 175.4
episode: 800   score: 190.0   memory length: 647932   epsilon:
0.009998020008555413    steps: 999    lr: 1.0240000000000005e-06     evaluation
reward: 175.9
episode: 801   score: 250.0   memory length: 648786   epsilon:
0.009998020008555413    steps: 854    lr: 1.0240000000000005e-06     evaluation
reward: 176.9
episode: 802   score: 90.0   memory length: 649362   epsilon:
0.009998020008555413    steps: 576    lr: 1.0240000000000005e-06     evaluation
reward: 175.3
episode: 803   score: 110.0   memory length: 649797   epsilon:
0.009998020008555413    steps: 435    lr: 1.0240000000000005e-06     evaluation
reward: 175.5
episode: 804   score: 150.0   memory length: 650361   epsilon:
0.009998020008555413    steps: 564    lr: 1.0240000000000005e-06     evaluation
reward: 173.45
episode: 805   score: 130.0   memory length: 650805   epsilon:
0.009998020008555413    steps: 444    lr: 1.0240000000000005e-06     evaluation
reward: 172.85
episode: 806   score: 235.0   memory length: 651658   epsilon:
0.009998020008555413    steps: 853    lr: 1.0240000000000005e-06     evaluation
reward: 171.95
episode: 807   score: 175.0   memory length: 652338   epsilon:
0.009998020008555413    steps: 680    lr: 1.0240000000000005e-06     evaluation
reward: 172.3
episode: 808   score: 130.0   memory length: 652816   epsilon:
0.009998020008555413    steps: 478    lr: 1.0240000000000005e-06     evaluation
reward: 170.5
```

episode: 809   score: 190.0   memory length: 653546   epsilon:
0.009998020008555413   steps: 730   lr: 1.0240000000000005e-06   evaluation
reward: 170.35
episode: 810   score: 175.0   memory length: 654282   epsilon:
0.009998020008555413   steps: 736   lr: 1.0240000000000005e-06   evaluation
reward: 170.7
episode: 811   score: 190.0   memory length: 655036   epsilon:
0.009998020008555413   steps: 754   lr: 1.0240000000000005e-06   evaluation
reward: 169.95
episode: 812   score: 355.0   memory length: 656047   epsilon:
0.009998020008555413   steps: 1011   lr: 1.0240000000000005e-06   evaluation
reward: 171.6
episode: 813   score: 60.0   memory length: 656342   epsilon:
0.009998020008555413   steps: 295   lr: 1.0240000000000005e-06   evaluation
reward: 170.3
episode: 814   score: 265.0   memory length: 657212   epsilon:
0.009998020008555413   steps: 870   lr: 1.0240000000000005e-06   evaluation
reward: 170.9
episode: 815   score: 190.0   memory length: 657972   epsilon:
0.009998020008555413   steps: 760   lr: 1.0240000000000005e-06   evaluation
reward: 170.6
episode: 816   score: 175.0   memory length: 658880   epsilon:
0.009998020008555413   steps: 908   lr: 1.0240000000000005e-06   evaluation
reward: 170.75
episode: 817   score: 160.0   memory length: 659462   epsilon:
0.009998020008555413   steps: 582   lr: 1.0240000000000005e-06   evaluation
reward: 171.35
episode: 818   score: 130.0   memory length: 659920   epsilon:
0.009998020008555413   steps: 458   lr: 1.0240000000000005e-06   evaluation
reward: 171.35
episode: 819   score: 340.0   memory length: 661052   epsilon:
0.009998020008555413   steps: 1132   lr: 1.0240000000000005e-06   evaluation
reward: 173.25
episode: 820   score: 120.0   memory length: 661537   epsilon:
0.009998020008555413   steps: 485   lr: 1.0240000000000005e-06   evaluation
reward: 171.8
episode: 821   score: 355.0   memory length: 662488   epsilon:
0.009998020008555413   steps: 951   lr: 1.0240000000000005e-06   evaluation
reward: 174.95
episode: 822   score: 120.0   memory length: 663148   epsilon:
0.009998020008555413   steps: 660   lr: 1.0240000000000005e-06   evaluation
reward: 175.25
episode: 823   score: 80.0   memory length: 663528   epsilon:
0.009998020008555413   steps: 380   lr: 1.0240000000000005e-06   evaluation
reward: 175.35
episode: 824   score: 140.0   memory length: 664253   epsilon:
0.009998020008555413   steps: 725   lr: 1.0240000000000005e-06   evaluation
reward: 174.85

episode: 825   score: 190.0   memory length: 665006   epsilon:
0.009998020008555413   steps: 753   lr: 1.0240000000000005e-06   evaluation
reward: 176.15
episode: 826   score: 325.0   memory length: 665988   epsilon:
0.009998020008555413   steps: 982   lr: 1.0240000000000005e-06   evaluation
reward: 178.9
episode: 827   score: 280.0   memory length: 667047   epsilon:
0.009998020008555413   steps: 1059   lr: 1.0240000000000005e-06   evaluation
reward: 180.3
episode: 828   score: 140.0   memory length: 667749   epsilon:
0.009998020008555413   steps: 702   lr: 1.0240000000000005e-06   evaluation
reward: 179.8
episode: 829   score: 150.0   memory length: 668269   epsilon:
0.009998020008555413   steps: 520   lr: 1.0240000000000005e-06   evaluation
reward: 180.1
episode: 830   score: 325.0   memory length: 669190   epsilon:
0.009998020008555413   steps: 921   lr: 1.0240000000000005e-06   evaluation
reward: 180.25
episode: 831   score: 310.0   memory length: 670130   epsilon:
0.009998020008555413   steps: 940   lr: 1.0240000000000005e-06   evaluation
reward: 182.25
episode: 832   score: 280.0   memory length: 671462   epsilon:
0.009998020008555413   steps: 1332   lr: 1.0240000000000005e-06   evaluation
reward: 183.55
episode: 833   score: 110.0   memory length: 671953   epsilon:
0.009998020008555413   steps: 491   lr: 1.0240000000000005e-06   evaluation
reward: 184.05
episode: 834   score: 175.0   memory length: 672762   epsilon:
0.009998020008555413   steps: 809   lr: 1.0240000000000005e-06   evaluation
reward: 184.9
episode: 835   score: 70.0   memory length: 673227   epsilon:
0.009998020008555413   steps: 465   lr: 1.0240000000000005e-06   evaluation
reward: 184.5
episode: 836   score: 90.0   memory length: 673621   epsilon:
0.009998020008555413   steps: 394   lr: 1.0240000000000005e-06   evaluation
reward: 182.3
episode: 837   score: 310.0   memory length: 674633   epsilon:
0.009998020008555413   steps: 1012   lr: 1.0240000000000005e-06   evaluation
reward: 184.9
episode: 838   score: 295.0   memory length: 675755   epsilon:
0.009998020008555413   steps: 1122   lr: 1.0240000000000005e-06   evaluation
reward: 185.05
episode: 839   score: 220.0   memory length: 676769   epsilon:
0.009998020008555413   steps: 1014   lr: 1.0240000000000005e-06   evaluation
reward: 184.6
episode: 840   score: 130.0   memory length: 677191   epsilon:
0.009998020008555413   steps: 422   lr: 1.0240000000000005e-06   evaluation
reward: 184.0

episode: 841   score: 140.0   memory length: 677913   epsilon:
0.009998020008555413   steps: 722   lr: 1.0240000000000005e-06   evaluation
reward: 184.6
episode: 842   score: 235.0   memory length: 678921   epsilon:
0.009998020008555413   steps: 1008   lr: 1.0240000000000005e-06   evaluation
reward: 183.85
episode: 843   score: 150.0   memory length: 679835   epsilon:
0.009998020008555413   steps: 914   lr: 1.0240000000000005e-06   evaluation
reward: 183.45
episode: 844   score: 205.0   memory length: 680596   epsilon:
0.009998020008555413   steps: 761   lr: 1.0240000000000005e-06   evaluation
reward: 183.0
episode: 845   score: 340.0   memory length: 681865   epsilon:
0.009998020008555413   steps: 1269   lr: 1.0240000000000005e-06   evaluation
reward: 184.05
episode: 846   score: 90.0   memory length: 682358   epsilon:
0.009998020008555413   steps: 493   lr: 1.0240000000000005e-06   evaluation
reward: 183.05
episode: 847   score: 110.0   memory length: 682699   epsilon:
0.009998020008555413   steps: 341   lr: 1.0240000000000005e-06   evaluation
reward: 181.8
episode: 848   score: 110.0   memory length: 683199   epsilon:
0.009998020008555413   steps: 500   lr: 1.0240000000000005e-06   evaluation
reward: 181.0
episode: 849   score: 340.0   memory length: 684457   epsilon:
0.009998020008555413   steps: 1258   lr: 1.0240000000000005e-06   evaluation
reward: 182.65
episode: 850   score: 205.0   memory length: 685287   epsilon:
0.009998020008555413   steps: 830   lr: 1.0240000000000005e-06   evaluation
reward: 183.9
episode: 851   score: 50.0   memory length: 685571   epsilon:
0.009998020008555413   steps: 284   lr: 1.0240000000000005e-06   evaluation
reward: 181.0
episode: 852   score: 190.0   memory length: 686611   epsilon:
0.009998020008555413   steps: 1040   lr: 1.0240000000000005e-06   evaluation
reward: 180.85
episode: 853   score: 235.0   memory length: 687565   epsilon:
0.009998020008555413   steps: 954   lr: 1.0240000000000005e-06   evaluation
reward: 182.8
episode: 854   score: 30.0   memory length: 687823   epsilon:
0.009998020008555413   steps: 258   lr: 1.0240000000000005e-06   evaluation
reward: 179.85
episode: 855   score: 80.0   memory length: 688225   epsilon:
0.009998020008555413   steps: 402   lr: 1.0240000000000005e-06   evaluation
reward: 179.55
episode: 856   score: 120.0   memory length: 688882   epsilon:
0.009998020008555413   steps: 657   lr: 1.0240000000000005e-06   evaluation
reward: 177.8

episode: 857   score: 150.0   memory length: 689626   epsilon:
0.009998020008555413   steps: 744   lr: 1.0240000000000005e-06   evaluation
reward: 178.0
episode: 858   score: 295.0   memory length: 690813   epsilon:
0.009998020008555413   steps: 1187   lr: 1.0240000000000005e-06   evaluation
reward: 179.65
episode: 859   score: 175.0   memory length: 691664   epsilon:
0.009998020008555413   steps: 851   lr: 1.0240000000000005e-06   evaluation
reward: 179.5
episode: 860   score: 120.0   memory length: 692220   epsilon:
0.009998020008555413   steps: 556   lr: 1.0240000000000005e-06   evaluation
reward: 178.35
episode: 861   score: 130.0   memory length: 692873   epsilon:
0.009998020008555413   steps: 653   lr: 1.0240000000000005e-06   evaluation
reward: 177.15
episode: 862   score: 250.0   memory length: 693913   epsilon:
0.009998020008555413   steps: 1040   lr: 1.0240000000000005e-06   evaluation
reward: 177.15
episode: 863   score: 250.0   memory length: 694629   epsilon:
0.009998020008555413   steps: 716   lr: 1.0240000000000005e-06   evaluation
reward: 178.25
episode: 864   score: 220.0   memory length: 695402   epsilon:
0.009998020008555413   steps: 773   lr: 1.0240000000000005e-06   evaluation
reward: 179.55
episode: 865   score: 190.0   memory length: 696126   epsilon:
0.009998020008555413   steps: 724   lr: 1.0240000000000005e-06   evaluation
reward: 178.95
episode: 866   score: 370.0   memory length: 697126   epsilon:
0.009998020008555413   steps: 1000   lr: 1.0240000000000005e-06   evaluation
reward: 181.75
episode: 867   score: 250.0   memory length: 697983   epsilon:
0.009998020008555413   steps: 857   lr: 1.0240000000000005e-06   evaluation
reward: 183.45
episode: 868   score: 370.0   memory length: 699150   epsilon:
0.009998020008555413   steps: 1167   lr: 1.0240000000000005e-06   evaluation
reward: 184.35
episode: 869   score: 175.0   memory length: 699842   epsilon:
0.009998020008555413   steps: 692   lr: 1.0240000000000005e-06   evaluation
reward: 184.2
episode: 870   score: 205.0   memory length: 700480   epsilon:
0.009998020008555413   steps: 638   lr: 4.0960000000000023e-07   evaluation
reward: 185.55
episode: 871   score: 110.0   memory length: 700944   epsilon:
0.009998020008555413   steps: 464   lr: 4.0960000000000023e-07   evaluation
reward: 185.15
episode: 872   score: 175.0   memory length: 701578   epsilon:
0.009998020008555413   steps: 634   lr: 4.0960000000000023e-07   evaluation
reward: 186.0

episode: 873   score: 160.0   memory length: 702338   epsilon:
0.009998020008555413   steps: 760   lr: 4.0960000000000023e-07   evaluation
reward: 187.3
episode: 874   score: 110.0   memory length: 703144   epsilon:
0.009998020008555413   steps: 806   lr: 4.0960000000000023e-07   evaluation
reward: 186.9
episode: 875   score: 120.0   memory length: 703600   epsilon:
0.009998020008555413   steps: 456   lr: 4.0960000000000023e-07   evaluation
reward: 187.0
episode: 876   score: 325.0   memory length: 704723   epsilon:
0.009998020008555413   steps: 1123   lr: 4.0960000000000023e-07   evaluation
reward: 189.15
episode: 877   score: 60.0   memory length: 705029   epsilon:
0.009998020008555413   steps: 306   lr: 4.0960000000000023e-07   evaluation
reward: 187.7
episode: 878   score: 280.0   memory length: 706167   epsilon:
0.009998020008555413   steps: 1138   lr: 4.0960000000000023e-07   evaluation
reward: 188.0
episode: 879   score: 140.0   memory length: 706727   epsilon:
0.009998020008555413   steps: 560   lr: 4.0960000000000023e-07   evaluation
reward: 187.2
episode: 880   score: 190.0   memory length: 707744   epsilon:
0.009998020008555413   steps: 1017   lr: 4.0960000000000023e-07   evaluation
reward: 188.1
episode: 881   score: 110.0   memory length: 708242   epsilon:
0.009998020008555413   steps: 498   lr: 4.0960000000000023e-07   evaluation
reward: 188.2
episode: 882   score: 130.0   memory length: 708950   epsilon:
0.009998020008555413   steps: 708   lr: 4.0960000000000023e-07   evaluation
reward: 188.2
episode: 883   score: 160.0   memory length: 709739   epsilon:
0.009998020008555413   steps: 789   lr: 4.0960000000000023e-07   evaluation
reward: 188.2
episode: 884   score: 160.0   memory length: 710407   epsilon:
0.009998020008555413   steps: 668   lr: 4.0960000000000023e-07   evaluation
reward: 187.6
episode: 885   score: 70.0   memory length: 710769   epsilon:
0.009998020008555413   steps: 362   lr: 4.0960000000000023e-07   evaluation
reward: 186.8
episode: 886   score: 175.0   memory length: 711351   epsilon:
0.009998020008555413   steps: 582   lr: 4.0960000000000023e-07   evaluation
reward: 187.35
episode: 887   score: 140.0   memory length: 711791   epsilon:
0.009998020008555413   steps: 440   lr: 4.0960000000000023e-07   evaluation
reward: 187.15
episode: 888   score: 220.0   memory length: 712460   epsilon:
0.009998020008555413   steps: 669   lr: 4.0960000000000023e-07   evaluation
reward: 183.75

```
episode: 889   score: 120.0   memory length: 712887   epsilon:
0.009998020008555413    steps: 427    lr: 4.0960000000000023e-07     evaluation
reward: 183.85
episode: 890   score: 370.0   memory length: 713819   epsilon:
0.009998020008555413    steps: 932    lr: 4.0960000000000023e-07     evaluation
reward: 187.15
episode: 891   score: 205.0   memory length: 714832   epsilon:
0.009998020008555413    steps: 1013    lr: 4.0960000000000023e-07     evaluation
reward: 187.7
episode: 892   score: 80.0   memory length: 715370   epsilon:
0.009998020008555413    steps: 538    lr: 4.0960000000000023e-07     evaluation
reward: 187.3
episode: 893   score: 235.0   memory length: 716019   epsilon:
0.009998020008555413    steps: 649    lr: 4.0960000000000023e-07     evaluation
reward: 188.25
episode: 894   score: 325.0   memory length: 717032   epsilon:
0.009998020008555413    steps: 1013    lr: 4.0960000000000023e-07     evaluation
reward: 188.85
episode: 895   score: 100.0   memory length: 717552   epsilon:
0.009998020008555413    steps: 520    lr: 4.0960000000000023e-07     evaluation
reward: 186.6
episode: 896   score: 150.0   memory length: 718286   epsilon:
0.009998020008555413    steps: 734    lr: 4.0960000000000023e-07     evaluation
reward: 186.6
episode: 897   score: 295.0   memory length: 719202   epsilon:
0.009998020008555413    steps: 916    lr: 4.0960000000000023e-07     evaluation
reward: 188.55
episode: 898   score: 160.0   memory length: 719669   epsilon:
0.009998020008555413    steps: 467    lr: 4.0960000000000023e-07     evaluation
reward: 187.2
episode: 899   score: 220.0   memory length: 720264   epsilon:
0.009998020008555413    steps: 595    lr: 4.0960000000000023e-07     evaluation
reward: 187.8
episode: 900   score: 440.0   memory length: 721808   epsilon:
0.009998020008555413    steps: 1544    lr: 4.0960000000000023e-07     evaluation
reward: 190.3
episode: 901   score: 100.0   memory length: 722389   epsilon:
0.009998020008555413    steps: 581    lr: 4.0960000000000023e-07     evaluation
reward: 188.8
episode: 902   score: 175.0   memory length: 723149   epsilon:
0.009998020008555413    steps: 760    lr: 4.0960000000000023e-07     evaluation
reward: 189.65
episode: 903   score: 175.0   memory length: 723766   epsilon:
0.009998020008555413    steps: 617    lr: 4.0960000000000023e-07     evaluation
reward: 190.3
episode: 904   score: 250.0   memory length: 724440   epsilon:
0.009998020008555413    steps: 674    lr: 4.0960000000000023e-07     evaluation
reward: 191.3
```

```
episode: 905   score: 265.0   memory length: 725405   epsilon:
0.009998020008555413    steps: 965    lr: 4.0960000000000023e-07     evaluation
reward: 192.65
episode: 906   score: 250.0   memory length: 726224   epsilon:
0.009998020008555413    steps: 819    lr: 4.0960000000000023e-07     evaluation
reward: 192.8
episode: 907   score: 205.0   memory length: 727049   epsilon:
0.009998020008555413    steps: 825    lr: 4.0960000000000023e-07     evaluation
reward: 193.1
episode: 908   score: 150.0   memory length: 727789   epsilon:
0.009998020008555413    steps: 740    lr: 4.0960000000000023e-07     evaluation
reward: 193.3
episode: 909   score: 325.0   memory length: 728830   epsilon:
0.009998020008555413    steps: 1041    lr: 4.0960000000000023e-07     evaluation
reward: 194.65
episode: 910   score: 150.0   memory length: 729427   epsilon:
0.009998020008555413    steps: 597    lr: 4.0960000000000023e-07     evaluation
reward: 194.4
episode: 911   score: 100.0   memory length: 730008   epsilon:
0.009998020008555413    steps: 581    lr: 4.0960000000000023e-07     evaluation
reward: 193.5
episode: 912   score: 120.0   memory length: 730431   epsilon:
0.009998020008555413    steps: 423    lr: 4.0960000000000023e-07     evaluation
reward: 191.15
episode: 913   score: 355.0   memory length: 731587   epsilon:
0.009998020008555413    steps: 1156    lr: 4.0960000000000023e-07     evaluation
reward: 194.1
episode: 914   score: 250.0   memory length: 732503   epsilon:
0.009998020008555413    steps: 916    lr: 4.0960000000000023e-07     evaluation
reward: 193.95
episode: 915   score: 80.0   memory length: 732921   epsilon:
0.009998020008555413    steps: 418    lr: 4.0960000000000023e-07     evaluation
reward: 192.85
episode: 916   score: 110.0   memory length: 733292   epsilon:
0.009998020008555413    steps: 371    lr: 4.0960000000000023e-07     evaluation
reward: 192.2
episode: 917   score: 220.0   memory length: 734176   epsilon:
0.009998020008555413    steps: 884    lr: 4.0960000000000023e-07     evaluation
reward: 192.8
episode: 918   score: 90.0   memory length: 734701   epsilon:
0.009998020008555413    steps: 525    lr: 4.0960000000000023e-07     evaluation
reward: 192.4
episode: 919   score: 140.0   memory length: 735440   epsilon:
0.009998020008555413    steps: 739    lr: 4.0960000000000023e-07     evaluation
reward: 190.4
episode: 920   score: 280.0   memory length: 736131   epsilon:
0.009998020008555413    steps: 691    lr: 4.0960000000000023e-07     evaluation
reward: 192.0
```

episode: 921   score: 325.0   memory length: 737213   epsilon:
0.009998020008555413   steps: 1082   lr: 4.0960000000000023e-07   evaluation
reward: 191.7
episode: 922   score: 110.0   memory length: 737777   epsilon:
0.009998020008555413   steps: 564   lr: 4.0960000000000023e-07   evaluation
reward: 191.6
episode: 923   score: 190.0   memory length: 738730   epsilon:
0.009998020008555413   steps: 953   lr: 4.0960000000000023e-07   evaluation
reward: 192.7
episode: 924   score: 205.0   memory length: 739500   epsilon:
0.009998020008555413   steps: 770   lr: 4.0960000000000023e-07   evaluation
reward: 193.35
episode: 925   score: 140.0   memory length: 740008   epsilon:
0.009998020008555413   steps: 508   lr: 4.0960000000000023e-07   evaluation
reward: 192.85
episode: 926   score: 70.0   memory length: 740343   epsilon:
0.009998020008555413   steps: 335   lr: 4.0960000000000023e-07   evaluation
reward: 190.3
episode: 927   score: 370.0   memory length: 741454   epsilon:
0.009998020008555413   steps: 1111   lr: 4.0960000000000023e-07   evaluation
reward: 191.2
episode: 928   score: 190.0   memory length: 742229   epsilon:
0.009998020008555413   steps: 775   lr: 4.0960000000000023e-07   evaluation
reward: 191.7
episode: 929   score: 325.0   memory length: 743407   epsilon:
0.009998020008555413   steps: 1178   lr: 4.0960000000000023e-07   evaluation
reward: 193.45
episode: 930   score: 340.0   memory length: 744541   epsilon:
0.009998020008555413   steps: 1134   lr: 4.0960000000000023e-07   evaluation
reward: 193.6
episode: 931   score: 325.0   memory length: 745474   epsilon:
0.009998020008555413   steps: 933   lr: 4.0960000000000023e-07   evaluation
reward: 193.75
episode: 932   score: 175.0   memory length: 746406   epsilon:
0.009998020008555413   steps: 932   lr: 4.0960000000000023e-07   evaluation
reward: 192.7
episode: 933   score: 175.0   memory length: 746916   epsilon:
0.009998020008555413   steps: 510   lr: 4.0960000000000023e-07   evaluation
reward: 193.35
episode: 934   score: 100.0   memory length: 747439   epsilon:
0.009998020008555413   steps: 523   lr: 4.0960000000000023e-07   evaluation
reward: 192.6
episode: 935   score: 355.0   memory length: 748528   epsilon:
0.009998020008555413   steps: 1089   lr: 4.0960000000000023e-07   evaluation
reward: 195.45
episode: 936   score: 80.0   memory length: 749041   epsilon:
0.009998020008555413   steps: 513   lr: 4.0960000000000023e-07   evaluation
reward: 195.35

episode: 937   score: 355.0   memory length: 750332   epsilon:
0.009998020008555413   steps: 1291   lr: 4.0960000000000023e-07   evaluation
reward: 195.8
episode: 938   score: 90.0   memory length: 750800   epsilon:
0.009998020008555413   steps: 468   lr: 4.0960000000000023e-07   evaluation
reward: 193.75
episode: 939   score: 280.0   memory length: 751602   epsilon:
0.009998020008555413   steps: 802   lr: 4.0960000000000023e-07   evaluation
reward: 194.35
episode: 940   score: 370.0   memory length: 752860   epsilon:
0.009998020008555413   steps: 1258   lr: 4.0960000000000023e-07   evaluation
reward: 196.75
episode: 941   score: 120.0   memory length: 753644   epsilon:
0.009998020008555413   steps: 784   lr: 4.0960000000000023e-07   evaluation
reward: 196.55
episode: 942   score: 160.0   memory length: 754404   epsilon:
0.009998020008555413   steps: 760   lr: 4.0960000000000023e-07   evaluation
reward: 195.8
episode: 943   score: 90.0   memory length: 754840   epsilon:
0.009998020008555413   steps: 436   lr: 4.0960000000000023e-07   evaluation
reward: 195.2
episode: 944   score: 340.0   memory length: 755957   epsilon:
0.009998020008555413   steps: 1117   lr: 4.0960000000000023e-07   evaluation
reward: 196.55
episode: 945   score: 150.0   memory length: 756647   epsilon:
0.009998020008555413   steps: 690   lr: 4.0960000000000023e-07   evaluation
reward: 194.65
episode: 946   score: 150.0   memory length: 757306   epsilon:
0.009998020008555413   steps: 659   lr: 4.0960000000000023e-07   evaluation
reward: 195.25
episode: 947   score: 250.0   memory length: 758117   epsilon:
0.009998020008555413   steps: 811   lr: 4.0960000000000023e-07   evaluation
reward: 196.65
episode: 948   score: 175.0   memory length: 759027   epsilon:
0.009998020008555413   steps: 910   lr: 4.0960000000000023e-07   evaluation
reward: 197.3
episode: 949   score: 160.0   memory length: 759743   epsilon:
0.009998020008555413   steps: 716   lr: 4.0960000000000023e-07   evaluation
reward: 195.5
episode: 950   score: 280.0   memory length: 761046   epsilon:
0.009998020008555413   steps: 1303   lr: 4.0960000000000023e-07   evaluation
reward: 196.25
episode: 951   score: 190.0   memory length: 761579   epsilon:
0.009998020008555413   steps: 533   lr: 4.0960000000000023e-07   evaluation
reward: 197.65
episode: 952   score: 80.0   memory length: 761971   epsilon:
0.009998020008555413   steps: 392   lr: 4.0960000000000023e-07   evaluation
reward: 196.55

episode: 953   score: 620.0   memory length: 763421   epsilon:
0.009998020008555413   steps: 1450   lr: 4.0960000000000023e-07   evaluation
reward: 200.4
episode: 954   score: 355.0   memory length: 764390   epsilon:
0.009998020008555413   steps: 969   lr: 4.0960000000000023e-07   evaluation
reward: 203.65
episode: 955   score: 175.0   memory length: 765025   epsilon:
0.009998020008555413   steps: 635   lr: 4.0960000000000023e-07   evaluation
reward: 204.6
episode: 956   score: 295.0   memory length: 766108   epsilon:
0.009998020008555413   steps: 1083   lr: 4.0960000000000023e-07   evaluation
reward: 206.35
episode: 957   score: 310.0   memory length: 766987   epsilon:
0.009998020008555413   steps: 879   lr: 4.0960000000000023e-07   evaluation
reward: 207.95
episode: 958   score: 80.0   memory length: 767363   epsilon:
0.009998020008555413   steps: 376   lr: 4.0960000000000023e-07   evaluation
reward: 205.8
episode: 959   score: 100.0   memory length: 767799   epsilon:
0.009998020008555413   steps: 436   lr: 4.0960000000000023e-07   evaluation
reward: 205.05
episode: 960   score: 310.0   memory length: 768999   epsilon:
0.009998020008555413   steps: 1200   lr: 4.0960000000000023e-07   evaluation
reward: 206.95
episode: 961   score: 190.0   memory length: 769666   epsilon:
0.009998020008555413   steps: 667   lr: 4.0960000000000023e-07   evaluation
reward: 207.55
episode: 962   score: 235.0   memory length: 770428   epsilon:
0.009998020008555413   steps: 762   lr: 4.0960000000000023e-07   evaluation
reward: 207.4
episode: 963   score: 110.0   memory length: 770855   epsilon:
0.009998020008555413   steps: 427   lr: 4.0960000000000023e-07   evaluation
reward: 206.0
episode: 964   score: 130.0   memory length: 771507   epsilon:
0.009998020008555413   steps: 652   lr: 4.0960000000000023e-07   evaluation
reward: 205.1
episode: 965   score: 130.0   memory length: 772067   epsilon:
0.009998020008555413   steps: 560   lr: 4.0960000000000023e-07   evaluation
reward: 204.5
episode: 966   score: 160.0   memory length: 772968   epsilon:
0.009998020008555413   steps: 901   lr: 4.0960000000000023e-07   evaluation
reward: 202.4
episode: 967   score: 1460.0   memory length: 774687   epsilon:
0.009998020008555413   steps: 1719   lr: 4.0960000000000023e-07   evaluation
reward: 214.5
episode: 968   score: 150.0   memory length: 775294   epsilon:
0.009998020008555413   steps: 607   lr: 4.0960000000000023e-07   evaluation
reward: 212.3

episode: 969   score: 80.0   memory length: 775860   epsilon:
0.009998020008555413   steps: 566   lr: 4.0960000000000023e-07   evaluation
reward: 211.35
episode: 970   score: 295.0   memory length: 777013   epsilon:
0.009998020008555413   steps: 1153   lr: 4.0960000000000023e-07   evaluation
reward: 212.25
episode: 971   score: 355.0   memory length: 777957   epsilon:
0.009998020008555413   steps: 944   lr: 4.0960000000000023e-07   evaluation
reward: 214.7
episode: 972   score: 310.0   memory length: 778915   epsilon:
0.009998020008555413   steps: 958   lr: 4.0960000000000023e-07   evaluation
reward: 216.05
episode: 973   score: 235.0   memory length: 779732   epsilon:
0.009998020008555413   steps: 817   lr: 4.0960000000000023e-07   evaluation
reward: 216.8
episode: 974   score: 250.0   memory length: 780495   epsilon:
0.009998020008555413   steps: 763   lr: 4.0960000000000023e-07   evaluation
reward: 218.2
episode: 975   score: 340.0   memory length: 781441   epsilon:
0.009998020008555413   steps: 946   lr: 4.0960000000000023e-07   evaluation
reward: 220.4
episode: 976   score: 325.0   memory length: 782302   epsilon:
0.009998020008555413   steps: 861   lr: 4.0960000000000023e-07   evaluation
reward: 220.4
episode: 977   score: 40.0   memory length: 782536   epsilon:
0.009998020008555413   steps: 234   lr: 4.0960000000000023e-07   evaluation
reward: 220.2
episode: 978   score: 175.0   memory length: 783180   epsilon:
0.009998020008555413   steps: 644   lr: 4.0960000000000023e-07   evaluation
reward: 219.15
episode: 979   score: 500.0   memory length: 784635   epsilon:
0.009998020008555413   steps: 1455   lr: 4.0960000000000023e-07   evaluation
reward: 222.75
episode: 980   score: 130.0   memory length: 785218   epsilon:
0.009998020008555413   steps: 583   lr: 4.0960000000000023e-07   evaluation
reward: 222.15
episode: 981   score: 235.0   memory length: 785955   epsilon:
0.009998020008555413   steps: 737   lr: 4.0960000000000023e-07   evaluation
reward: 223.4
episode: 982   score: 110.0   memory length: 786610   epsilon:
0.009998020008555413   steps: 655   lr: 4.0960000000000023e-07   evaluation
reward: 223.2
episode: 983   score: 250.0   memory length: 787331   epsilon:
0.009998020008555413   steps: 721   lr: 4.0960000000000023e-07   evaluation
reward: 224.1
episode: 984   score: 265.0   memory length: 788650   epsilon:
0.009998020008555413   steps: 1319   lr: 4.0960000000000023e-07   evaluation
reward: 225.15

episode: 985   score: 295.0   memory length: 789477   epsilon:
0.009998020008555413   steps: 827   lr: 4.0960000000000023e-07   evaluation
reward: 227.4
episode: 986   score: 250.0   memory length: 790288   epsilon:
0.009998020008555413   steps: 811   lr: 4.0960000000000023e-07   evaluation
reward: 228.15
episode: 987   score: 400.0   memory length: 791536   epsilon:
0.009998020008555413   steps: 1248   lr: 4.0960000000000023e-07   evaluation
reward: 230.75
episode: 988   score: 175.0   memory length: 792173   epsilon:
0.009998020008555413   steps: 637   lr: 4.0960000000000023e-07   evaluation
reward: 230.3
episode: 989   score: 325.0   memory length: 793361   epsilon:
0.009998020008555413   steps: 1188   lr: 4.0960000000000023e-07   evaluation
reward: 232.35
episode: 990   score: 220.0   memory length: 794324   epsilon:
0.009998020008555413   steps: 963   lr: 4.0960000000000023e-07   evaluation
reward: 230.85
episode: 991   score: 150.0   memory length: 794920   epsilon:
0.009998020008555413   steps: 596   lr: 4.0960000000000023e-07   evaluation
reward: 230.3
episode: 992   score: 295.0   memory length: 795717   epsilon:
0.009998020008555413   steps: 797   lr: 4.0960000000000023e-07   evaluation
reward: 232.45
episode: 993   score: 120.0   memory length: 796134   epsilon:
0.009998020008555413   steps: 417   lr: 4.0960000000000023e-07   evaluation
reward: 231.3
episode: 994   score: 90.0   memory length: 796464   epsilon:
0.009998020008555413   steps: 330   lr: 4.0960000000000023e-07   evaluation
reward: 228.95
episode: 995   score: 80.0   memory length: 796799   epsilon:
0.009998020008555413   steps: 335   lr: 4.0960000000000023e-07   evaluation
reward: 228.75
episode: 996   score: 120.0   memory length: 797344   epsilon:
0.009998020008555413   steps: 545   lr: 4.0960000000000023e-07   evaluation
reward: 228.45
episode: 997   score: 220.0   memory length: 798105   epsilon:
0.009998020008555413   steps: 761   lr: 4.0960000000000023e-07   evaluation
reward: 227.7
episode: 998   score: 295.0   memory length: 799123   epsilon:
0.009998020008555413   steps: 1018   lr: 4.0960000000000023e-07   evaluation
reward: 229.05
episode: 999   score: 175.0   memory length: 799676   epsilon:
0.009998020008555413   steps: 553   lr: 4.0960000000000023e-07   evaluation
reward: 228.6
episode: 1000   score: 190.0   memory length: 800361   epsilon:
0.009998020008555413   steps: 685   lr: 1.638400000000001e-07   evaluation
reward: 226.1

episode: 1001   score: 130.0   memory length: 800988   epsilon:
0.009998020008555413   steps: 627   lr: 1.6384000000000001e-07   evaluation
reward: 226.4
episode: 1002   score: 70.0   memory length: 801306   epsilon:
0.009998020008555413   steps: 318   lr: 1.6384000000000001e-07   evaluation
reward: 225.35
episode: 1003   score: 280.0   memory length: 802168   epsilon:
0.009998020008555413   steps: 862   lr: 1.6384000000000001e-07   evaluation
reward: 226.4
episode: 1004   score: 295.0   memory length: 803407   epsilon:
0.009998020008555413   steps: 1239   lr: 1.6384000000000001e-07   evaluation
reward: 226.85
episode: 1005   score: 325.0   memory length: 804392   epsilon:
0.009998020008555413   steps: 985   lr: 1.6384000000000001e-07   evaluation
reward: 227.45
episode: 1006   score: 310.0   memory length: 805474   epsilon:
0.009998020008555413   steps: 1082   lr: 1.6384000000000001e-07   evaluation
reward: 228.05
episode: 1007   score: 370.0   memory length: 806599   epsilon:
0.009998020008555413   steps: 1125   lr: 1.6384000000000001e-07   evaluation
reward: 229.7
episode: 1008   score: 420.0   memory length: 807631   epsilon:
0.009998020008555413   steps: 1032   lr: 1.6384000000000001e-07   evaluation
reward: 232.4
episode: 1009   score: 205.0   memory length: 808377   epsilon:
0.009998020008555413   steps: 746   lr: 1.6384000000000001e-07   evaluation
reward: 231.2
episode: 1010   score: 265.0   memory length: 809359   epsilon:
0.009998020008555413   steps: 982   lr: 1.6384000000000001e-07   evaluation
reward: 232.35
episode: 1011   score: 100.0   memory length: 809808   epsilon:
0.009998020008555413   steps: 449   lr: 1.6384000000000001e-07   evaluation
reward: 232.35
episode: 1012   score: 265.0   memory length: 810585   epsilon:
0.009998020008555413   steps: 777   lr: 1.6384000000000001e-07   evaluation
reward: 233.8
episode: 1013   score: 120.0   memory length: 811212   epsilon:
0.009998020008555413   steps: 627   lr: 1.6384000000000001e-07   evaluation
reward: 231.45
episode: 1014   score: 140.0   memory length: 811796   epsilon:
0.009998020008555413   steps: 584   lr: 1.6384000000000001e-07   evaluation
reward: 230.35
episode: 1015   score: 310.0   memory length: 812877   epsilon:
0.009998020008555413   steps: 1081   lr: 1.6384000000000001e-07   evaluation
reward: 232.65
episode: 1016   score: 265.0   memory length: 813765   epsilon:
0.009998020008555413   steps: 888   lr: 1.6384000000000001e-07   evaluation
reward: 234.2

```
episode: 1017   score: 150.0   memory length: 814345   epsilon:
0.009998020008555413    steps: 580    lr: 1.6384000000000001e-07     evaluation
reward: 233.5
episode: 1018   score: 220.0   memory length: 815018   epsilon:
0.009998020008555413    steps: 673    lr: 1.6384000000000001e-07     evaluation
reward: 234.8
episode: 1019   score: 235.0   memory length: 815730   epsilon:
0.009998020008555413    steps: 712    lr: 1.6384000000000001e-07     evaluation
reward: 235.75
episode: 1020   score: 295.0   memory length: 816761   epsilon:
0.009998020008555413    steps: 1031    lr: 1.6384000000000001e-07     evaluation
reward: 235.9
episode: 1021   score: 120.0   memory length: 817390   epsilon:
0.009998020008555413    steps: 629    lr: 1.6384000000000001e-07     evaluation
reward: 233.85
episode: 1022   score: 175.0   memory length: 818407   epsilon:
0.009998020008555413    steps: 1017    lr: 1.6384000000000001e-07     evaluation
reward: 234.5
episode: 1023   score: 120.0   memory length: 818936   epsilon:
0.009998020008555413    steps: 529    lr: 1.6384000000000001e-07     evaluation
reward: 233.8
episode: 1024   score: 420.0   memory length: 820568   epsilon:
0.009998020008555413    steps: 1632    lr: 1.6384000000000001e-07     evaluation
reward: 235.95
episode: 1025   score: 310.0   memory length: 821640   epsilon:
0.009998020008555413    steps: 1072    lr: 1.6384000000000001e-07     evaluation
reward: 237.65
episode: 1026   score: 235.0   memory length: 822302   epsilon:
0.009998020008555413    steps: 662    lr: 1.6384000000000001e-07     evaluation
reward: 239.3
episode: 1027   score: 190.0   memory length: 822967   epsilon:
0.009998020008555413    steps: 665    lr: 1.6384000000000001e-07     evaluation
reward: 237.5
episode: 1028   score: 120.0   memory length: 823667   epsilon:
0.009998020008555413    steps: 700    lr: 1.6384000000000001e-07     evaluation
reward: 236.8
episode: 1029   score: 190.0   memory length: 824380   epsilon:
0.009998020008555413    steps: 713    lr: 1.6384000000000001e-07     evaluation
reward: 235.45
episode: 1030   score: 160.0   memory length: 825101   epsilon:
0.009998020008555413    steps: 721    lr: 1.6384000000000001e-07     evaluation
reward: 233.65
episode: 1031   score: 295.0   memory length: 826024   epsilon:
0.009998020008555413    steps: 923    lr: 1.6384000000000001e-07     evaluation
reward: 233.35
episode: 1032   score: 175.0   memory length: 826767   epsilon:
0.009998020008555413    steps: 743    lr: 1.6384000000000001e-07     evaluation
reward: 233.35
```

```
episode: 1033   score: 220.0   memory length: 827514   epsilon:
0.009998020008555413    steps: 747    lr: 1.6384000000000001e-07     evaluation
reward: 233.8
episode: 1034   score: 110.0   memory length: 827992   epsilon:
0.009998020008555413    steps: 478    lr: 1.6384000000000001e-07     evaluation
reward: 233.9
episode: 1035   score: 160.0   memory length: 828664   epsilon:
0.009998020008555413    steps: 672    lr: 1.6384000000000001e-07     evaluation
reward: 231.95
episode: 1036   score: 175.0   memory length: 829326   epsilon:
0.009998020008555413    steps: 662    lr: 1.6384000000000001e-07     evaluation
reward: 232.9
episode: 1037   score: 150.0   memory length: 829940   epsilon:
0.009998020008555413    steps: 614    lr: 1.6384000000000001e-07     evaluation
reward: 230.85
episode: 1038   score: 1160.0   memory length: 831617   epsilon:
0.009998020008555413    steps: 1677    lr: 1.6384000000000001e-07     evaluation
reward: 241.55
episode: 1039   score: 740.0   memory length: 833574   epsilon:
0.009998020008555413    steps: 1957    lr: 1.6384000000000001e-07     evaluation
reward: 246.15
episode: 1040   score: 205.0   memory length: 834381   epsilon:
0.009998020008555413    steps: 807    lr: 1.6384000000000001e-07     evaluation
reward: 244.5
episode: 1041   score: 150.0   memory length: 834937   epsilon:
0.009998020008555413    steps: 556    lr: 1.6384000000000001e-07     evaluation
reward: 244.8
episode: 1042   score: 190.0   memory length: 835658   epsilon:
0.009998020008555413    steps: 721    lr: 1.6384000000000001e-07     evaluation
reward: 245.1
episode: 1043   score: 70.0   memory length: 835914   epsilon:
0.009998020008555413    steps: 256    lr: 1.6384000000000001e-07     evaluation
reward: 244.9
episode: 1044   score: 370.0   memory length: 836728   epsilon:
0.009998020008555413    steps: 814    lr: 1.6384000000000001e-07     evaluation
reward: 245.2
episode: 1045   score: 175.0   memory length: 837515   epsilon:
0.009998020008555413    steps: 787    lr: 1.6384000000000001e-07     evaluation
reward: 245.45
episode: 1046   score: 205.0   memory length: 838291   epsilon:
0.009998020008555413    steps: 776    lr: 1.6384000000000001e-07     evaluation
reward: 246.0
episode: 1047   score: 340.0   memory length: 839472   epsilon:
0.009998020008555413    steps: 1181    lr: 1.6384000000000001e-07     evaluation
reward: 246.9
episode: 1048   score: 205.0   memory length: 840336   epsilon:
0.009998020008555413    steps: 864    lr: 1.6384000000000001e-07     evaluation
reward: 247.2
```

episode: 1049   score: 325.0   memory length: 841400   epsilon:
0.009998020008555413   steps: 1064   lr: 1.6384000000000001e-07   evaluation
reward: 248.85
episode: 1050   score: 90.0   memory length: 841888   epsilon:
0.009998020008555413   steps: 488   lr: 1.6384000000000001e-07   evaluation
reward: 246.95
episode: 1051   score: 580.0   memory length: 843292   epsilon:
0.009998020008555413   steps: 1404   lr: 1.6384000000000001e-07   evaluation
reward: 250.85
episode: 1052   score: 310.0   memory length: 844587   epsilon:
0.009998020008555413   steps: 1295   lr: 1.6384000000000001e-07   evaluation
reward: 253.15
episode: 1053   score: 235.0   memory length: 845583   epsilon:
0.009998020008555413   steps: 996   lr: 1.6384000000000001e-07   evaluation
reward: 249.3
episode: 1054   score: 205.0   memory length: 846372   epsilon:
0.009998020008555413   steps: 789   lr: 1.6384000000000001e-07   evaluation
reward: 247.8
episode: 1055   score: 250.0   memory length: 847142   epsilon:
0.009998020008555413   steps: 770   lr: 1.6384000000000001e-07   evaluation
reward: 248.55
episode: 1056   score: 780.0   memory length: 848649   epsilon:
0.009998020008555413   steps: 1507   lr: 1.6384000000000001e-07   evaluation
reward: 253.4
episode: 1057   score: 160.0   memory length: 849491   epsilon:
0.009998020008555413   steps: 842   lr: 1.6384000000000001e-07   evaluation
reward: 251.9
episode: 1058   score: 205.0   memory length: 850280   epsilon:
0.009998020008555413   steps: 789   lr: 1.6384000000000001e-07   evaluation
reward: 253.15
episode: 1059   score: 175.0   memory length: 850926   epsilon:
0.009998020008555413   steps: 646   lr: 1.6384000000000001e-07   evaluation
reward: 253.9
episode: 1060   score: 205.0   memory length: 851883   epsilon:
0.009998020008555413   steps: 957   lr: 1.6384000000000001e-07   evaluation
reward: 252.85
episode: 1061   score: 120.0   memory length: 852353   epsilon:
0.009998020008555413   steps: 470   lr: 1.6384000000000001e-07   evaluation
reward: 252.15
episode: 1062   score: 70.0   memory length: 852693   epsilon:
0.009998020008555413   steps: 340   lr: 1.6384000000000001e-07   evaluation
reward: 250.5
episode: 1063   score: 150.0   memory length: 853212   epsilon:
0.009998020008555413   steps: 519   lr: 1.6384000000000001e-07   evaluation
reward: 250.9
episode: 1064   score: 140.0   memory length: 853850   epsilon:
0.009998020008555413   steps: 638   lr: 1.6384000000000001e-07   evaluation
reward: 251.0

episode: 1065   score: 160.0   memory length: 854734   epsilon:
0.009998020008555413   steps: 884   lr: 1.6384000000000001e-07   evaluation
reward: 251.3
episode: 1066   score: 140.0   memory length: 855188   epsilon:
0.009998020008555413   steps: 454   lr: 1.6384000000000001e-07   evaluation
reward: 251.1
episode: 1067   score: 235.0   memory length: 855922   epsilon:
0.009998020008555413   steps: 734   lr: 1.6384000000000001e-07   evaluation
reward: 238.85
episode: 1068   score: 250.0   memory length: 856868   epsilon:
0.009998020008555413   steps: 946   lr: 1.6384000000000001e-07   evaluation
reward: 239.85
episode: 1069   score: 140.0   memory length: 857346   epsilon:
0.009998020008555413   steps: 478   lr: 1.6384000000000001e-07   evaluation
reward: 240.45
episode: 1070   score: 325.0   memory length: 858293   epsilon:
0.009998020008555413   steps: 947   lr: 1.6384000000000001e-07   evaluation
reward: 240.75
episode: 1071   score: 280.0   memory length: 859338   epsilon:
0.009998020008555413   steps: 1045   lr: 1.6384000000000001e-07   evaluation
reward: 240.0
episode: 1072   score: 140.0   memory length: 860042   epsilon:
0.009998020008555413   steps: 704   lr: 1.6384000000000001e-07   evaluation
reward: 238.3
episode: 1073   score: 280.0   memory length: 860883   epsilon:
0.009998020008555413   steps: 841   lr: 1.6384000000000001e-07   evaluation
reward: 238.75
episode: 1074   score: 250.0   memory length: 861770   epsilon:
0.009998020008555413   steps: 887   lr: 1.6384000000000001e-07   evaluation
reward: 238.75
episode: 1075   score: 325.0   memory length: 862748   epsilon:
0.009998020008555413   steps: 978   lr: 1.6384000000000001e-07   evaluation
reward: 238.6
episode: 1076   score: 160.0   memory length: 863587   epsilon:
0.009998020008555413   steps: 839   lr: 1.6384000000000001e-07   evaluation
reward: 236.95
episode: 1077   score: 205.0   memory length: 864299   epsilon:
0.009998020008555413   steps: 712   lr: 1.6384000000000001e-07   evaluation
reward: 238.6
episode: 1078   score: 190.0   memory length: 865100   epsilon:
0.009998020008555413   steps: 801   lr: 1.6384000000000001e-07   evaluation
reward: 238.75
episode: 1079   score: 100.0   memory length: 865766   epsilon:
0.009998020008555413   steps: 666   lr: 1.6384000000000001e-07   evaluation
reward: 234.75
episode: 1080   score: 140.0   memory length: 866402   epsilon:
0.009998020008555413   steps: 636   lr: 1.6384000000000001e-07   evaluation
reward: 234.85

episode: 1081   score: 280.0   memory length: 867669   epsilon:
0.009998020008555413   steps: 1267   lr: 1.638400000000001e-07   evaluation
reward: 235.3
episode: 1082   score: 370.0   memory length: 868844   epsilon:
0.009998020008555413   steps: 1175   lr: 1.638400000000001e-07   evaluation
reward: 237.9
episode: 1083   score: 250.0   memory length: 870073   epsilon:
0.009998020008555413   steps: 1229   lr: 1.638400000000001e-07   evaluation
reward: 237.9
episode: 1084   score: 325.0   memory length: 871207   epsilon:
0.009998020008555413   steps: 1134   lr: 1.638400000000001e-07   evaluation
reward: 238.5
episode: 1085   score: 190.0   memory length: 871978   epsilon:
0.009998020008555413   steps: 771   lr: 1.638400000000001e-07   evaluation
reward: 237.45
episode: 1086   score: 20.0   memory length: 872242   epsilon:
0.009998020008555413   steps: 264   lr: 1.638400000000001e-07   evaluation
reward: 235.15
episode: 1087   score: 250.0   memory length: 873183   epsilon:
0.009998020008555413   steps: 941   lr: 1.638400000000001e-07   evaluation
reward: 233.65
episode: 1088   score: 520.0   memory length: 874476   epsilon:
0.009998020008555413   steps: 1293   lr: 1.638400000000001e-07   evaluation
reward: 237.1
episode: 1089   score: 295.0   memory length: 875449   epsilon:
0.009998020008555413   steps: 973   lr: 1.638400000000001e-07   evaluation
reward: 236.8
episode: 1090   score: 355.0   memory length: 876733   epsilon:
0.009998020008555413   steps: 1284   lr: 1.638400000000001e-07   evaluation
reward: 238.15
episode: 1091   score: 175.0   memory length: 877588   epsilon:
0.009998020008555413   steps: 855   lr: 1.638400000000001e-07   evaluation
reward: 238.4
episode: 1092   score: 265.0   memory length: 878459   epsilon:
0.009998020008555413   steps: 871   lr: 1.638400000000001e-07   evaluation
reward: 238.1
episode: 1093   score: 80.0   memory length: 878832   epsilon:
0.009998020008555413   steps: 373   lr: 1.638400000000001e-07   evaluation
reward: 237.7
episode: 1094   score: 160.0   memory length: 879472   epsilon:
0.009998020008555413   steps: 640   lr: 1.638400000000001e-07   evaluation
reward: 238.4
episode: 1095   score: 150.0   memory length: 880131   epsilon:
0.009998020008555413   steps: 659   lr: 1.638400000000001e-07   evaluation
reward: 239.1
episode: 1096   score: 325.0   memory length: 881241   epsilon:
0.009998020008555413   steps: 1110   lr: 1.638400000000001e-07   evaluation
reward: 241.15

```
episode: 1097   score: 60.0   memory length: 881592   epsilon:
0.009998020008555413    steps: 351    lr: 1.6384000000000001e-07     evaluation
reward: 239.55
episode: 1098   score: 90.0   memory length: 882454   epsilon:
0.009998020008555413    steps: 862    lr: 1.6384000000000001e-07     evaluation
reward: 237.5
episode: 1099   score: 175.0   memory length: 883096   epsilon:
0.009998020008555413    steps: 642    lr: 1.6384000000000001e-07     evaluation
reward: 237.5
episode: 1100   score: 110.0   memory length: 883613   epsilon:
0.009998020008555413    steps: 517    lr: 1.6384000000000001e-07     evaluation
reward: 236.7
episode: 1101   score: 480.0   memory length: 884902   epsilon:
0.009998020008555413    steps: 1289    lr: 1.6384000000000001e-07     evaluation
reward: 240.2
episode: 1102   score: 140.0   memory length: 885414   epsilon:
0.009998020008555413    steps: 512    lr: 1.6384000000000001e-07     evaluation
reward: 240.9
episode: 1103   score: 680.0   memory length: 887026   epsilon:
0.009998020008555413    steps: 1612    lr: 1.6384000000000001e-07     evaluation
reward: 244.9
episode: 1104   score: 265.0   memory length: 888006   epsilon:
0.009998020008555413    steps: 980    lr: 1.6384000000000001e-07     evaluation
reward: 244.6
episode: 1105   score: 310.0   memory length: 889054   epsilon:
0.009998020008555413    steps: 1048    lr: 1.6384000000000001e-07     evaluation
reward: 244.45
episode: 1106   score: 110.0   memory length: 889826   epsilon:
0.009998020008555413    steps: 772    lr: 1.6384000000000001e-07     evaluation
reward: 242.45
episode: 1107   score: 160.0   memory length: 890474   epsilon:
0.009998020008555413    steps: 648    lr: 1.6384000000000001e-07     evaluation
reward: 240.35
episode: 1108   score: 160.0   memory length: 891084   epsilon:
0.009998020008555413    steps: 610    lr: 1.6384000000000001e-07     evaluation
reward: 237.75
episode: 1109   score: 130.0   memory length: 891747   epsilon:
0.009998020008555413    steps: 663    lr: 1.6384000000000001e-07     evaluation
reward: 237.0
episode: 1110   score: 220.0   memory length: 892494   epsilon:
0.009998020008555413    steps: 747    lr: 1.6384000000000001e-07     evaluation
reward: 236.55
episode: 1111   score: 60.0   memory length: 892827   epsilon:
0.009998020008555413    steps: 333    lr: 1.6384000000000001e-07     evaluation
reward: 236.15
episode: 1112   score: 310.0   memory length: 893755   epsilon:
0.009998020008555413    steps: 928    lr: 1.6384000000000001e-07     evaluation
reward: 236.6
```

episode: 1113   score: 355.0   memory length: 895017   epsilon:
0.009998020008555413   steps: 1262   lr: 1.638400000000001e-07   evaluation
reward: 238.95
episode: 1114   score: 370.0   memory length: 896236   epsilon:
0.009998020008555413   steps: 1219   lr: 1.638400000000001e-07   evaluation
reward: 241.25
episode: 1115   score: 190.0   memory length: 896942   epsilon:
0.009998020008555413   steps: 706   lr: 1.638400000000001e-07   evaluation
reward: 240.05
episode: 1116   score: 280.0   memory length: 897933   epsilon:
0.009998020008555413   steps: 991   lr: 1.638400000000001e-07   evaluation
reward: 240.2
episode: 1117   score: 90.0   memory length: 898388   epsilon:
0.009998020008555413   steps: 455   lr: 1.638400000000001e-07   evaluation
reward: 239.6
episode: 1118   score: 90.0   memory length: 898905   epsilon:
0.009998020008555413   steps: 517   lr: 1.638400000000001e-07   evaluation
reward: 238.3
episode: 1119   score: 370.0   memory length: 899978   epsilon:
0.009998020008555413   steps: 1073   lr: 1.638400000000001e-07   evaluation
reward: 239.65
episode: 1120   score: 140.0   memory length: 900800   epsilon:
0.009998020008555413   steps: 822   lr: 6.553600000000004e-08   evaluation
reward: 238.1
episode: 1121   score: 355.0   memory length: 901909   epsilon:
0.009998020008555413   steps: 1109   lr: 6.553600000000004e-08   evaluation
reward: 240.45
episode: 1122   score: 175.0   memory length: 902750   epsilon:
0.009998020008555413   steps: 841   lr: 6.553600000000004e-08   evaluation
reward: 240.45
episode: 1123   score: 540.0   memory length: 904060   epsilon:
0.009998020008555413   steps: 1310   lr: 6.553600000000004e-08   evaluation
reward: 244.65
episode: 1124   score: 265.0   memory length: 904877   epsilon:
0.009998020008555413   steps: 817   lr: 6.553600000000004e-08   evaluation
reward: 243.1
episode: 1125   score: 130.0   memory length: 905331   epsilon:
0.009998020008555413   steps: 454   lr: 6.553600000000004e-08   evaluation
reward: 241.3
episode: 1126   score: 175.0   memory length: 906094   epsilon:
0.009998020008555413   steps: 763   lr: 6.553600000000004e-08   evaluation
reward: 240.7
episode: 1127   score: 355.0   memory length: 907778   epsilon:
0.009998020008555413   steps: 1684   lr: 6.553600000000004e-08   evaluation
reward: 242.35
episode: 1128   score: 235.0   memory length: 908641   epsilon:
0.009998020008555413   steps: 863   lr: 6.553600000000004e-08   evaluation
reward: 243.5

```
episode: 1129   score: 90.0   memory length: 909031   epsilon:
0.009998020008555413    steps: 390    lr: 6.553600000000004e-08     evaluation
reward: 242.5
episode: 1130   score: 110.0   memory length: 909472   epsilon:
0.009998020008555413    steps: 441    lr: 6.553600000000004e-08     evaluation
reward: 242.0
episode: 1131   score: 235.0   memory length: 910279   epsilon:
0.009998020008555413    steps: 807    lr: 6.553600000000004e-08     evaluation
reward: 241.4
episode: 1132   score: 205.0   memory length: 911106   epsilon:
0.009998020008555413    steps: 827    lr: 6.553600000000004e-08     evaluation
reward: 241.7
episode: 1133   score: 220.0   memory length: 911836   epsilon:
0.009998020008555413    steps: 730    lr: 6.553600000000004e-08     evaluation
reward: 241.7
episode: 1134   score: 90.0   memory length: 912288   epsilon:
0.009998020008555413    steps: 452    lr: 6.553600000000004e-08     evaluation
reward: 241.5
episode: 1135   score: 175.0   memory length: 912925   epsilon:
0.009998020008555413    steps: 637    lr: 6.553600000000004e-08     evaluation
reward: 241.65
episode: 1136   score: 250.0   memory length: 913868   epsilon:
0.009998020008555413    steps: 943    lr: 6.553600000000004e-08     evaluation
reward: 242.4
episode: 1137   score: 235.0   memory length: 914688   epsilon:
0.009998020008555413    steps: 820    lr: 6.553600000000004e-08     evaluation
reward: 243.25
episode: 1138   score: 370.0   memory length: 915951   epsilon:
0.009998020008555413    steps: 1263    lr: 6.553600000000004e-08     evaluation
reward: 235.35
episode: 1139   score: 100.0   memory length: 916444   epsilon:
0.009998020008555413    steps: 493    lr: 6.553600000000004e-08     evaluation
reward: 228.95
episode: 1140   score: 440.0   memory length: 917690   epsilon:
0.009998020008555413    steps: 1246    lr: 6.553600000000004e-08     evaluation
reward: 231.3
episode: 1141   score: 220.0   memory length: 918402   epsilon:
0.009998020008555413    steps: 712    lr: 6.553600000000004e-08     evaluation
reward: 232.0
episode: 1142   score: 280.0   memory length: 919403   epsilon:
0.009998020008555413    steps: 1001    lr: 6.553600000000004e-08     evaluation
reward: 232.9
episode: 1143   score: 295.0   memory length: 920279   epsilon:
0.009998020008555413    steps: 876    lr: 6.553600000000004e-08     evaluation
reward: 235.15
episode: 1144   score: 480.0   memory length: 921698   epsilon:
0.009998020008555413    steps: 1419    lr: 6.553600000000004e-08     evaluation
reward: 236.25
```

```
episode: 1145   score: 295.0   memory length: 922782   epsilon:
0.009998020008555413    steps: 1084    lr: 6.553600000000004e-08     evaluation
reward: 237.45
episode: 1146   score: 295.0   memory length: 923764   epsilon:
0.009998020008555413    steps: 982    lr: 6.553600000000004e-08     evaluation
reward: 238.35
episode: 1147   score: 190.0   memory length: 924546   epsilon:
0.009998020008555413    steps: 782    lr: 6.553600000000004e-08     evaluation
reward: 236.85
episode: 1148   score: 40.0   memory length: 924812   epsilon:
0.009998020008555413    steps: 266    lr: 6.553600000000004e-08     evaluation
reward: 235.2
episode: 1149   score: 160.0   memory length: 925609   epsilon:
0.009998020008555413    steps: 797    lr: 6.553600000000004e-08     evaluation
reward: 233.55
episode: 1150   score: 250.0   memory length: 926479   epsilon:
0.009998020008555413    steps: 870    lr: 6.553600000000004e-08     evaluation
reward: 235.15
episode: 1151   score: 150.0   memory length: 927128   epsilon:
0.009998020008555413    steps: 649    lr: 6.553600000000004e-08     evaluation
reward: 230.85
episode: 1152   score: 520.0   memory length: 928302   epsilon:
0.009998020008555413    steps: 1174    lr: 6.553600000000004e-08     evaluation
reward: 232.95
episode: 1153   score: 90.0   memory length: 928716   epsilon:
0.009998020008555413    steps: 414    lr: 6.553600000000004e-08     evaluation
reward: 231.5
episode: 1154   score: 130.0   memory length: 929300   epsilon:
0.009998020008555413    steps: 584    lr: 6.553600000000004e-08     evaluation
reward: 230.75
episode: 1155   score: 310.0   memory length: 930313   epsilon:
0.009998020008555413    steps: 1013    lr: 6.553600000000004e-08     evaluation
reward: 231.35
episode: 1156   score: 310.0   memory length: 931389   epsilon:
0.009998020008555413    steps: 1076    lr: 6.553600000000004e-08     evaluation
reward: 226.65
episode: 1157   score: 520.0   memory length: 932792   epsilon:
0.009998020008555413    steps: 1403    lr: 6.553600000000004e-08     evaluation
reward: 230.25
episode: 1158   score: 295.0   memory length: 933682   epsilon:
0.009998020008555413    steps: 890    lr: 6.553600000000004e-08     evaluation
reward: 231.15
episode: 1159   score: 220.0   memory length: 934672   epsilon:
0.009998020008555413    steps: 990    lr: 6.553600000000004e-08     evaluation
reward: 231.6
episode: 1160   score: 235.0   memory length: 935397   epsilon:
0.009998020008555413    steps: 725    lr: 6.553600000000004e-08     evaluation
reward: 231.9
```

```
episode: 1161   score: 130.0   memory length: 935969   epsilon:
0.009998020008555413    steps: 572    lr: 6.553600000000004e-08     evaluation
reward: 232.0
episode: 1162   score: 100.0   memory length: 936377   epsilon:
0.009998020008555413    steps: 408    lr: 6.553600000000004e-08     evaluation
reward: 232.3
episode: 1163   score: 130.0   memory length: 936844   epsilon:
0.009998020008555413    steps: 467    lr: 6.553600000000004e-08     evaluation
reward: 232.1
episode: 1164   score: 420.0   memory length: 938045   epsilon:
0.009998020008555413    steps: 1201    lr: 6.553600000000004e-08     evaluation
reward: 234.9
episode: 1165   score: 205.0   memory length: 938835   epsilon:
0.009998020008555413    steps: 790    lr: 6.553600000000004e-08     evaluation
reward: 235.35
episode: 1166   score: 205.0   memory length: 939595   epsilon:
0.009998020008555413    steps: 760    lr: 6.553600000000004e-08     evaluation
reward: 236.0
episode: 1167   score: 310.0   memory length: 940777   epsilon:
0.009998020008555413    steps: 1182    lr: 6.553600000000004e-08     evaluation
reward: 236.75
episode: 1168   score: 120.0   memory length: 941279   epsilon:
0.009998020008555413    steps: 502    lr: 6.553600000000004e-08     evaluation
reward: 235.45
episode: 1169   score: 235.0   memory length: 942100   epsilon:
0.009998020008555413    steps: 821    lr: 6.553600000000004e-08     evaluation
reward: 236.4
episode: 1170   score: 265.0   memory length: 943023   epsilon:
0.009998020008555413    steps: 923    lr: 6.553600000000004e-08     evaluation
reward: 235.8
episode: 1171   score: 175.0   memory length: 944034   epsilon:
0.009998020008555413    steps: 1011    lr: 6.553600000000004e-08     evaluation
reward: 234.75
episode: 1172   score: 90.0   memory length: 944506   epsilon:
0.009998020008555413    steps: 472    lr: 6.553600000000004e-08     evaluation
reward: 234.25
episode: 1173   score: 130.0   memory length: 945006   epsilon:
0.009998020008555413    steps: 500    lr: 6.553600000000004e-08     evaluation
reward: 232.75
episode: 1174   score: 205.0   memory length: 945806   epsilon:
0.009998020008555413    steps: 800    lr: 6.553600000000004e-08     evaluation
reward: 232.3
episode: 1175   score: 280.0   memory length: 946686   epsilon:
0.009998020008555413    steps: 880    lr: 6.553600000000004e-08     evaluation
reward: 231.85
episode: 1176   score: 190.0   memory length: 947508   epsilon:
0.009998020008555413    steps: 822    lr: 6.553600000000004e-08     evaluation
reward: 232.15
```

episode: 1177   score: 250.0   memory length: 948304   epsilon:
0.009998020008555413   steps: 796   lr: 6.553600000000004e-08   evaluation
reward: 232.6
episode: 1178   score: 190.0   memory length: 949320   epsilon:
0.009998020008555413   steps: 1016   lr: 6.553600000000004e-08   evaluation
reward: 232.6
episode: 1179   score: 100.0   memory length: 949824   epsilon:
0.009998020008555413   steps: 504   lr: 6.553600000000004e-08   evaluation
reward: 232.6
episode: 1180   score: 60.0   memory length: 950221   epsilon:
0.009998020008555413   steps: 397   lr: 6.553600000000004e-08   evaluation
reward: 231.8
episode: 1181   score: 50.0   memory length: 950551   epsilon:
0.009998020008555413   steps: 330   lr: 6.553600000000004e-08   evaluation
reward: 229.5
episode: 1182   score: 30.0   memory length: 950791   epsilon:
0.009998020008555413   steps: 240   lr: 6.553600000000004e-08   evaluation
reward: 226.1
episode: 1183   score: 110.0   memory length: 951463   epsilon:
0.009998020008555413   steps: 672   lr: 6.553600000000004e-08   evaluation
reward: 224.7
episode: 1184   score: 295.0   memory length: 952306   epsilon:
0.009998020008555413   steps: 843   lr: 6.553600000000004e-08   evaluation
reward: 224.4
episode: 1185   score: 110.0   memory length: 952704   epsilon:
0.009998020008555413   steps: 398   lr: 6.553600000000004e-08   evaluation
reward: 223.6
episode: 1186   score: 340.0   memory length: 953665   epsilon:
0.009998020008555413   steps: 961   lr: 6.553600000000004e-08   evaluation
reward: 226.8
episode: 1187   score: 250.0   memory length: 954349   epsilon:
0.009998020008555413   steps: 684   lr: 6.553600000000004e-08   evaluation
reward: 226.8
episode: 1188   score: 190.0   memory length: 954920   epsilon:
0.009998020008555413   steps: 571   lr: 6.553600000000004e-08   evaluation
reward: 223.5
episode: 1189   score: 130.0   memory length: 955501   epsilon:
0.009998020008555413   steps: 581   lr: 6.553600000000004e-08   evaluation
reward: 221.85
episode: 1190   score: 140.0   memory length: 955875   epsilon:
0.009998020008555413   steps: 374   lr: 6.553600000000004e-08   evaluation
reward: 219.7
episode: 1191   score: 280.0   memory length: 956466   epsilon:
0.009998020008555413   steps: 591   lr: 6.553600000000004e-08   evaluation
reward: 220.75
episode: 1192   score: 175.0   memory length: 957388   epsilon:
0.009998020008555413   steps: 922   lr: 6.553600000000004e-08   evaluation
reward: 219.85

```
episode: 1193   score: 175.0   memory length: 957951   epsilon:
0.009998020008555413    steps: 563    lr: 6.553600000000004e-08     evaluation
reward: 220.8
episode: 1194   score: 325.0   memory length: 959126   epsilon:
0.009998020008555413    steps: 1175    lr: 6.553600000000004e-08     evaluation
reward: 222.45
episode: 1195   score: 160.0   memory length: 959945   epsilon:
0.009998020008555413    steps: 819    lr: 6.553600000000004e-08     evaluation
reward: 222.55
episode: 1196   score: 205.0   memory length: 960533   epsilon:
0.009998020008555413    steps: 588    lr: 6.553600000000004e-08     evaluation
reward: 221.35
episode: 1197   score: 175.0   memory length: 961205   epsilon:
0.009998020008555413    steps: 672    lr: 6.553600000000004e-08     evaluation
reward: 222.5
episode: 1198   score: 100.0   memory length: 961787   epsilon:
0.009998020008555413    steps: 582    lr: 6.553600000000004e-08     evaluation
reward: 222.6
episode: 1199   score: 250.0   memory length: 962678   epsilon:
0.009998020008555413    steps: 891    lr: 6.553600000000004e-08     evaluation
reward: 223.35
episode: 1200   score: 295.0   memory length: 963645   epsilon:
0.009998020008555413    steps: 967    lr: 6.553600000000004e-08     evaluation
reward: 225.2
episode: 1201   score: 280.0   memory length: 964706   epsilon:
0.009998020008555413    steps: 1061    lr: 6.553600000000004e-08     evaluation
reward: 223.2
episode: 1202   score: 840.0   memory length: 965904   epsilon:
0.009998020008555413    steps: 1198    lr: 6.553600000000004e-08     evaluation
reward: 230.2
episode: 1203   score: 175.0   memory length: 966646   epsilon:
0.009998020008555413    steps: 742    lr: 6.553600000000004e-08     evaluation
reward: 225.15
episode: 1204   score: 140.0   memory length: 967241   epsilon:
0.009998020008555413    steps: 595    lr: 6.553600000000004e-08     evaluation
reward: 223.9
episode: 1205   score: 250.0   memory length: 968167   epsilon:
0.009998020008555413    steps: 926    lr: 6.553600000000004e-08     evaluation
reward: 223.3
episode: 1206   score: 250.0   memory length: 969015   epsilon:
0.009998020008555413    steps: 848    lr: 6.553600000000004e-08     evaluation
reward: 224.7
episode: 1207   score: 160.0   memory length: 969831   epsilon:
0.009998020008555413    steps: 816    lr: 6.553600000000004e-08     evaluation
reward: 224.7
episode: 1208   score: 110.0   memory length: 970197   epsilon:
0.009998020008555413    steps: 366    lr: 6.553600000000004e-08     evaluation
reward: 224.2
```

episode: 1209   score: 150.0   memory length: 970699   epsilon:
0.009998020008555413   steps: 502   lr: 6.553600000000004e-08   evaluation
reward: 224.4
episode: 1210   score: 160.0   memory length: 971701   epsilon:
0.009998020008555413   steps: 1002   lr: 6.553600000000004e-08   evaluation
reward: 223.8
episode: 1211   score: 190.0   memory length: 972414   epsilon:
0.009998020008555413   steps: 713   lr: 6.553600000000004e-08   evaluation
reward: 225.1
episode: 1212   score: 130.0   memory length: 973003   epsilon:
0.009998020008555413   steps: 589   lr: 6.553600000000004e-08   evaluation
reward: 223.3
episode: 1213   score: 220.0   memory length: 973887   epsilon:
0.009998020008555413   steps: 884   lr: 6.553600000000004e-08   evaluation
reward: 221.95
episode: 1214   score: 175.0   memory length: 974467   epsilon:
0.009998020008555413   steps: 580   lr: 6.553600000000004e-08   evaluation
reward: 220.0
episode: 1215   score: 110.0   memory length: 975124   epsilon:
0.009998020008555413   steps: 657   lr: 6.553600000000004e-08   evaluation
reward: 219.2
episode: 1216   score: 220.0   memory length: 975985   epsilon:
0.009998020008555413   steps: 861   lr: 6.553600000000004e-08   evaluation
reward: 218.6
episode: 1217   score: 205.0   memory length: 976693   epsilon:
0.009998020008555413   steps: 708   lr: 6.553600000000004e-08   evaluation
reward: 219.75
episode: 1218   score: 220.0   memory length: 977442   epsilon:
0.009998020008555413   steps: 749   lr: 6.553600000000004e-08   evaluation
reward: 221.05
episode: 1219   score: 140.0   memory length: 978102   epsilon:
0.009998020008555413   steps: 660   lr: 6.553600000000004e-08   evaluation
reward: 218.75
episode: 1220   score: 250.0   memory length: 979053   epsilon:
0.009998020008555413   steps: 951   lr: 6.553600000000004e-08   evaluation
reward: 219.85
episode: 1221   score: 190.0   memory length: 979885   epsilon:
0.009998020008555413   steps: 832   lr: 6.553600000000004e-08   evaluation
reward: 218.2
episode: 1222   score: 235.0   memory length: 980592   epsilon:
0.009998020008555413   steps: 707   lr: 6.553600000000004e-08   evaluation
reward: 218.8
episode: 1223   score: 385.0   memory length: 981673   epsilon:
0.009998020008555413   steps: 1081   lr: 6.553600000000004e-08   evaluation
reward: 217.25
episode: 1224   score: 280.0   memory length: 982709   epsilon:
0.009998020008555413   steps: 1036   lr: 6.553600000000004e-08   evaluation
reward: 217.4

episode: 1225   score: 370.0   memory length: 983835   epsilon:
0.009998020008555413   steps: 1126   lr: 6.553600000000004e-08   evaluation
reward: 219.8
episode: 1226   score: 540.0   memory length: 985144   epsilon:
0.009998020008555413   steps: 1309   lr: 6.553600000000004e-08   evaluation
reward: 223.45
episode: 1227   score: 120.0   memory length: 985686   epsilon:
0.009998020008555413   steps: 542   lr: 6.553600000000004e-08   evaluation
reward: 221.1
episode: 1228   score: 325.0   memory length: 986731   epsilon:
0.009998020008555413   steps: 1045   lr: 6.553600000000004e-08   evaluation
reward: 222.0
episode: 1229   score: 325.0   memory length: 987874   epsilon:
0.009998020008555413   steps: 1143   lr: 6.553600000000004e-08   evaluation
reward: 224.35
episode: 1230   score: 325.0   memory length: 988877   epsilon:
0.009998020008555413   steps: 1003   lr: 6.553600000000004e-08   evaluation
reward: 226.5
episode: 1231   score: 265.0   memory length: 989816   epsilon:
0.009998020008555413   steps: 939   lr: 6.553600000000004e-08   evaluation
reward: 226.8
episode: 1232   score: 295.0   memory length: 990983   epsilon:
0.009998020008555413   steps: 1167   lr: 6.553600000000004e-08   evaluation
reward: 227.7
episode: 1233   score: 150.0   memory length: 991496   epsilon:
0.009998020008555413   steps: 513   lr: 6.553600000000004e-08   evaluation
reward: 227.0
episode: 1234   score: 235.0   memory length: 992362   epsilon:
0.009998020008555413   steps: 866   lr: 6.553600000000004e-08   evaluation
reward: 228.45
episode: 1235   score: 250.0   memory length: 993298   epsilon:
0.009998020008555413   steps: 936   lr: 6.553600000000004e-08   evaluation
reward: 229.2
episode: 1236   score: 90.0   memory length: 993681   epsilon:
0.009998020008555413   steps: 383   lr: 6.553600000000004e-08   evaluation
reward: 227.6
episode: 1237   score: 120.0   memory length: 993984   epsilon:
0.009998020008555413   steps: 303   lr: 6.553600000000004e-08   evaluation
reward: 226.45
episode: 1238   score: 150.0   memory length: 994690   epsilon:
0.009998020008555413   steps: 706   lr: 6.553600000000004e-08   evaluation
reward: 224.25
episode: 1239   score: 310.0   memory length: 995646   epsilon:
0.009998020008555413   steps: 956   lr: 6.553600000000004e-08   evaluation
reward: 226.35
episode: 1240   score: 110.0   memory length: 996211   epsilon:
0.009998020008555413   steps: 565   lr: 6.553600000000004e-08   evaluation
reward: 223.05

```
episode: 1241   score: 100.0   memory length: 996598   epsilon:
0.009998020008555413    steps: 387    lr: 6.553600000000004e-08     evaluation
reward: 221.85
episode: 1242   score: 250.0   memory length: 997459   epsilon:
0.009998020008555413    steps: 861    lr: 6.553600000000004e-08     evaluation
reward: 221.55
episode: 1243   score: 110.0   memory length: 997886   epsilon:
0.009998020008555413    steps: 427    lr: 6.553600000000004e-08     evaluation
reward: 219.7
episode: 1244   score: 110.0   memory length: 998306   epsilon:
0.009998020008555413    steps: 420    lr: 6.553600000000004e-08     evaluation
reward: 216.0
episode: 1245   score: 190.0   memory length: 999110   epsilon:
0.009998020008555413    steps: 804    lr: 6.553600000000004e-08     evaluation
reward: 214.95
episode: 1246   score: 205.0   memory length: 999773   epsilon:
0.009998020008555413    steps: 663    lr: 6.553600000000004e-08     evaluation
reward: 214.05
episode: 1247   score: 385.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1030    lr: 2.6214400000000017e-08     evaluation
reward: 216.0
episode: 1248   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 584    lr: 2.6214400000000017e-08     evaluation
reward: 216.9
episode: 1249   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 702    lr: 2.6214400000000017e-08     evaluation
reward: 217.65
episode: 1250   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 678    lr: 2.6214400000000017e-08     evaluation
reward: 217.05
episode: 1251   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 855    lr: 2.6214400000000017e-08     evaluation
reward: 217.75
episode: 1252   score: 760.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1646    lr: 2.6214400000000017e-08     evaluation
reward: 220.15
episode: 1253   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 845    lr: 2.6214400000000017e-08     evaluation
reward: 222.05
episode: 1254   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 556    lr: 2.6214400000000017e-08     evaluation
reward: 222.35
episode: 1255   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 925    lr: 2.6214400000000017e-08     evaluation
reward: 222.2
episode: 1256   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 883    lr: 2.6214400000000017e-08     evaluation
reward: 222.05
```

```
episode: 1257   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 903    lr: 2.6214400000000017e-08     evaluation
reward: 219.35
episode: 1258   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1050    lr: 2.6214400000000017e-08     evaluation
reward: 219.5
episode: 1259   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 426    lr: 2.6214400000000017e-08     evaluation
reward: 218.8
episode: 1260   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 554    lr: 2.6214400000000017e-08     evaluation
reward: 217.95
episode: 1261   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 657    lr: 2.6214400000000017e-08     evaluation
reward: 218.05
episode: 1262   score: 520.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1469    lr: 2.6214400000000017e-08     evaluation
reward: 222.25
episode: 1263   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 508    lr: 2.6214400000000017e-08     evaluation
reward: 222.25
episode: 1264   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 619    lr: 2.6214400000000017e-08     evaluation
reward: 219.65
episode: 1265   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 810    lr: 2.6214400000000017e-08     evaluation
reward: 219.2
episode: 1266   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 862    lr: 2.6214400000000017e-08     evaluation
reward: 219.2
episode: 1267   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 702    lr: 2.6214400000000017e-08     evaluation
reward: 217.4
episode: 1268   score: 10.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 208    lr: 2.6214400000000017e-08     evaluation
reward: 216.3
episode: 1269   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 564    lr: 2.6214400000000017e-08     evaluation
reward: 215.15
episode: 1270   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 947    lr: 2.6214400000000017e-08     evaluation
reward: 214.85
episode: 1271   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1326    lr: 2.6214400000000017e-08     evaluation
reward: 216.35
episode: 1272   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 889    lr: 2.6214400000000017e-08     evaluation
reward: 217.8
```

episode: 1273   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1180   lr: 2.6214400000000017e-08   evaluation
reward: 219.75
episode: 1274   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 862   lr: 2.6214400000000017e-08   evaluation
reward: 220.2
episode: 1275   score: 10.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 210   lr: 2.6214400000000017e-08   evaluation
reward: 217.5
episode: 1276   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 991   lr: 2.6214400000000017e-08   evaluation
reward: 218.7
episode: 1277   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 660   lr: 2.6214400000000017e-08   evaluation
reward: 217.7
episode: 1278   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 788   lr: 2.6214400000000017e-08   evaluation
reward: 218.45
episode: 1279   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 471   lr: 2.6214400000000017e-08   evaluation
reward: 218.65
episode: 1280   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 981   lr: 2.6214400000000017e-08   evaluation
reward: 220.4
episode: 1281   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 753   lr: 2.6214400000000017e-08   evaluation
reward: 221.95
episode: 1282   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 526   lr: 2.6214400000000017e-08   evaluation
reward: 223.05
episode: 1283   score: 340.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1438   lr: 2.6214400000000017e-08   evaluation
reward: 225.35
episode: 1284   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 758   lr: 2.6214400000000017e-08   evaluation
reward: 224.75
episode: 1285   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1046   lr: 2.6214400000000017e-08   evaluation
reward: 226.3
episode: 1286   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 430   lr: 2.6214400000000017e-08   evaluation
reward: 223.8
episode: 1287   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 819   lr: 2.6214400000000017e-08   evaluation
reward: 223.95
episode: 1288   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 530   lr: 2.6214400000000017e-08   evaluation
reward: 223.35

episode: 1289   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 634   lr: 2.6214400000000017e-08   evaluation
reward: 223.8
episode: 1290   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 718   lr: 2.6214400000000017e-08   evaluation
reward: 224.45
episode: 1291   score: 50.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 302   lr: 2.6214400000000017e-08   evaluation
reward: 222.15
episode: 1292   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 602   lr: 2.6214400000000017e-08   evaluation
reward: 222.15
episode: 1293   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 967   lr: 2.6214400000000017e-08   evaluation
reward: 223.05
episode: 1294   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 649   lr: 2.6214400000000017e-08   evaluation
reward: 221.85
episode: 1295   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 672   lr: 2.6214400000000017e-08   evaluation
reward: 221.65
episode: 1296   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 525   lr: 2.6214400000000017e-08   evaluation
reward: 220.4
episode: 1297   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 736   lr: 2.6214400000000017e-08   evaluation
reward: 221.0
episode: 1298   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 394   lr: 2.6214400000000017e-08   evaluation
reward: 220.9
episode: 1299   score: 460.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1076   lr: 2.6214400000000017e-08   evaluation
reward: 223.0
episode: 1300   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 949   lr: 2.6214400000000017e-08   evaluation
reward: 223.0
episode: 1301   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 807   lr: 2.6214400000000017e-08   evaluation
reward: 223.0
episode: 1302   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 500   lr: 2.6214400000000017e-08   evaluation
reward: 215.9
episode: 1303   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 822   lr: 2.6214400000000017e-08   evaluation
reward: 216.05
episode: 1304   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 486   lr: 2.6214400000000017e-08   evaluation
reward: 215.45

```
episode: 1305   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 865    lr: 2.6214400000000017e-08     evaluation
reward: 215.3
episode: 1306   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 491    lr: 2.6214400000000017e-08     evaluation
reward: 213.7
episode: 1307   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 784    lr: 2.6214400000000017e-08     evaluation
reward: 213.2
episode: 1308   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 755    lr: 2.6214400000000017e-08     evaluation
reward: 213.4
episode: 1309   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 648    lr: 2.6214400000000017e-08     evaluation
reward: 213.1
episode: 1310   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1190    lr: 2.6214400000000017e-08     evaluation
reward: 214.75
episode: 1311   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 364    lr: 2.6214400000000017e-08     evaluation
reward: 214.05
episode: 1312   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 576    lr: 2.6214400000000017e-08     evaluation
reward: 213.75
episode: 1313   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 936    lr: 2.6214400000000017e-08     evaluation
reward: 213.3
episode: 1314   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 653    lr: 2.6214400000000017e-08     evaluation
reward: 213.3
episode: 1315   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 913    lr: 2.6214400000000017e-08     evaluation
reward: 213.7
episode: 1316   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 991    lr: 2.6214400000000017e-08     evaluation
reward: 214.3
episode: 1317   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 730    lr: 2.6214400000000017e-08     evaluation
reward: 214.0
episode: 1318   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 379    lr: 2.6214400000000017e-08     evaluation
reward: 212.7
episode: 1319   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 457    lr: 2.6214400000000017e-08     evaluation
reward: 212.5
episode: 1320   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 500    lr: 2.6214400000000017e-08     evaluation
reward: 211.6
```

episode: 1321   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1152   lr: 2.6214400000000017e-08   evaluation
reward: 212.5
episode: 1322   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1176   lr: 2.6214400000000017e-08   evaluation
reward: 213.4
episode: 1323   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 975   lr: 2.6214400000000017e-08   evaluation
reward: 212.65
episode: 1324   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 517   lr: 2.6214400000000017e-08   evaluation
reward: 210.75
episode: 1325   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 721   lr: 2.6214400000000017e-08   evaluation
reward: 209.4
episode: 1326   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 664   lr: 2.6214400000000017e-08   evaluation
reward: 205.75
episode: 1327   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 522   lr: 2.6214400000000017e-08   evaluation
reward: 205.55
episode: 1328   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 979   lr: 2.6214400000000017e-08   evaluation
reward: 204.65
episode: 1329   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 510   lr: 2.6214400000000017e-08   evaluation
reward: 202.4
episode: 1330   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 263   lr: 2.6214400000000017e-08   evaluation
reward: 199.85
episode: 1331   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 811   lr: 2.6214400000000017e-08   evaluation
reward: 199.1
episode: 1332   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 969   lr: 2.6214400000000017e-08   evaluation
reward: 199.25
episode: 1333   score: 370.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1220   lr: 2.6214400000000017e-08   evaluation
reward: 201.45
episode: 1334   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 723   lr: 2.6214400000000017e-08   evaluation
reward: 200.85
episode: 1335   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 487   lr: 2.6214400000000017e-08   evaluation
reward: 199.85
episode: 1336   score: 20.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 226   lr: 2.6214400000000017e-08   evaluation
reward: 199.15

episode: 1337   score: 420.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1011   lr: 2.6214400000000017e-08   evaluation
reward: 202.15
episode: 1338   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 511   lr: 2.6214400000000017e-08   evaluation
reward: 202.15
episode: 1339   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 733   lr: 2.6214400000000017e-08   evaluation
reward: 200.35
episode: 1340   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 392   lr: 2.6214400000000017e-08   evaluation
reward: 200.05
episode: 1341   score: 60.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 328   lr: 2.6214400000000017e-08   evaluation
reward: 199.65
episode: 1342   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 781   lr: 2.6214400000000017e-08   evaluation
reward: 199.35
episode: 1343   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 650   lr: 2.6214400000000017e-08   evaluation
reward: 200.0
episode: 1344   score: 1300.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 2076   lr: 2.6214400000000017e-08   evaluation
reward: 211.9
episode: 1345   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1032   lr: 2.6214400000000017e-08   evaluation
reward: 212.05
episode: 1346   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 840   lr: 2.6214400000000017e-08   evaluation
reward: 212.5
episode: 1347   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 742   lr: 2.6214400000000017e-08   evaluation
reward: 211.15
episode: 1348   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 517   lr: 2.6214400000000017e-08   evaluation
reward: 211.15
episode: 1349   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 943   lr: 2.6214400000000017e-08   evaluation
reward: 211.15
episode: 1350   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 762   lr: 2.6214400000000017e-08   evaluation
reward: 210.25
episode: 1351   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 606   lr: 2.6214400000000017e-08   evaluation
reward: 209.35
episode: 1352   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1326   lr: 2.6214400000000017e-08   evaluation
reward: 204.7

episode: 1353 score: 1400.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 1872 lr: 2.6214400000000017e-08 evaluation reward: 215.9
episode: 1354 score: 190.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 806 lr: 2.6214400000000017e-08 evaluation reward: 216.2
episode: 1355 score: 175.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 808 lr: 2.6214400000000017e-08 evaluation reward: 215.0
episode: 1356 score: 160.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 617 lr: 2.6214400000000017e-08 evaluation reward: 213.65
episode: 1357 score: 130.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 671 lr: 2.6214400000000017e-08 evaluation reward: 212.45
episode: 1358 score: 140.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 596 lr: 2.6214400000000017e-08 evaluation reward: 210.75
episode: 1359 score: 250.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 651 lr: 2.6214400000000017e-08 evaluation reward: 211.75
episode: 1360 score: 310.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 1334 lr: 2.6214400000000017e-08 evaluation reward: 213.35
episode: 1361 score: 150.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 521 lr: 2.6214400000000017e-08 evaluation reward: 213.45
episode: 1362 score: 160.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 588 lr: 2.6214400000000017e-08 evaluation reward: 209.85
episode: 1363 score: 400.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 1033 lr: 2.6214400000000017e-08 evaluation reward: 212.55
episode: 1364 score: 150.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 582 lr: 2.6214400000000017e-08 evaluation reward: 212.45
episode: 1365 score: 160.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 692 lr: 2.6214400000000017e-08 evaluation reward: 212.45
episode: 1366 score: 265.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 957 lr: 2.6214400000000017e-08 evaluation reward: 213.05
episode: 1367 score: 310.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 835 lr: 2.6214400000000017e-08 evaluation reward: 214.85
episode: 1368 score: 175.0 memory length: 1000000 epsilon: 0.009998020008555413 steps: 695 lr: 2.6214400000000017e-08 evaluation reward: 216.5

episode: 1369   score: 460.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1357   lr: 2.6214400000000017e-08   evaluation
reward: 219.9
episode: 1370   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 633   lr: 2.6214400000000017e-08   evaluation
reward: 220.05
episode: 1371   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 716   lr: 2.6214400000000017e-08   evaluation
reward: 219.15
episode: 1372   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 522   lr: 2.6214400000000017e-08   evaluation
reward: 218.2
episode: 1373   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 723   lr: 2.6214400000000017e-08   evaluation
reward: 216.45
episode: 1374   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 546   lr: 2.6214400000000017e-08   evaluation
reward: 215.15
episode: 1375   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 435   lr: 2.6214400000000017e-08   evaluation
reward: 216.15
episode: 1376   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 902   lr: 2.6214400000000017e-08   evaluation
reward: 215.7
episode: 1377   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 448   lr: 1.0485760000000008e-08   evaluation
reward: 215.1
episode: 1378   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 582   lr: 1.0485760000000008e-08   evaluation
reward: 214.35
episode: 1379   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 793   lr: 1.0485760000000008e-08   evaluation
reward: 215.05
episode: 1380   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 432   lr: 1.0485760000000008e-08   evaluation
reward: 213.4
episode: 1381   score: 40.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 267   lr: 1.0485760000000008e-08   evaluation
reward: 211.75
episode: 1382   score: 385.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1069   lr: 1.0485760000000008e-08   evaluation
reward: 214.2
episode: 1383   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 669   lr: 1.0485760000000008e-08   evaluation
reward: 212.85
episode: 1384   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 532   lr: 1.0485760000000008e-08   evaluation
reward: 211.9

episode: 1385   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 949   lr: 1.0485760000000008e-08   evaluation
reward: 212.05
episode: 1386   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 456   lr: 1.0485760000000008e-08   evaluation
reward: 212.45
episode: 1387   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 937   lr: 1.0485760000000008e-08   evaluation
reward: 212.9
episode: 1388   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1251   lr: 1.0485760000000008e-08   evaluation
reward: 214.7
episode: 1389   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 695   lr: 1.0485760000000008e-08   evaluation
reward: 214.85
episode: 1390   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 563   lr: 1.0485760000000008e-08   evaluation
reward: 213.8
episode: 1391   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 811   lr: 1.0485760000000008e-08   evaluation
reward: 215.5
episode: 1392   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 457   lr: 1.0485760000000008e-08   evaluation
reward: 214.65
episode: 1393   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 948   lr: 1.0485760000000008e-08   evaluation
reward: 214.35
episode: 1394   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 891   lr: 1.0485760000000008e-08   evaluation
reward: 214.8
episode: 1395   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 643   lr: 1.0485760000000008e-08   evaluation
reward: 215.15
episode: 1396   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 601   lr: 1.0485760000000008e-08   evaluation
reward: 215.75
episode: 1397   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 951   lr: 1.0485760000000008e-08   evaluation
reward: 215.75
episode: 1398   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 809   lr: 1.0485760000000008e-08   evaluation
reward: 216.75
episode: 1399   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 389   lr: 1.0485760000000008e-08   evaluation
reward: 213.05
episode: 1400   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 974   lr: 1.0485760000000008e-08   evaluation
reward: 213.2

episode: 1401   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1165   lr: 1.0485760000000008e-08   evaluation
reward: 213.35
episode: 1402   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 442   lr: 1.0485760000000008e-08   evaluation
reward: 212.95
episode: 1403   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 857   lr: 1.0485760000000008e-08   evaluation
reward: 213.25
episode: 1404   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 911   lr: 1.0485760000000008e-08   evaluation
reward: 214.65
episode: 1405   score: 50.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 263   lr: 1.0485760000000008e-08   evaluation
reward: 212.8
episode: 1406   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1141   lr: 1.0485760000000008e-08   evaluation
reward: 214.85
episode: 1407   score: 370.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1055   lr: 1.0485760000000008e-08   evaluation
reward: 217.45
episode: 1408   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 791   lr: 1.0485760000000008e-08   evaluation
reward: 217.75
episode: 1409   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1382   lr: 1.0485760000000008e-08   evaluation
reward: 219.8
episode: 1410   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 817   lr: 1.0485760000000008e-08   evaluation
reward: 219.05
episode: 1411   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 859   lr: 1.0485760000000008e-08   evaluation
reward: 220.05
episode: 1412   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 594   lr: 1.0485760000000008e-08   evaluation
reward: 220.15
episode: 1413   score: 370.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 992   lr: 1.0485760000000008e-08   evaluation
reward: 222.1
episode: 1414   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 513   lr: 1.0485760000000008e-08   evaluation
reward: 221.25
episode: 1415   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 956   lr: 1.0485760000000008e-08   evaluation
reward: 222.85
episode: 1416   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 552   lr: 1.0485760000000008e-08   evaluation
reward: 221.15

```
episode: 1417   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1027    lr: 1.0485760000000008e-08     evaluation
reward: 221.3
episode: 1418   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 826    lr: 1.0485760000000008e-08     evaluation
reward: 223.35
episode: 1419   score: 440.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1182    lr: 1.0485760000000008e-08     evaluation
reward: 226.55
episode: 1420   score: 40.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 241    lr: 1.0485760000000008e-08     evaluation
reward: 225.35
episode: 1421   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 554    lr: 1.0485760000000008e-08     evaluation
reward: 223.65
episode: 1422   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 661    lr: 1.0485760000000008e-08     evaluation
reward: 222.3
episode: 1423   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 545    lr: 1.0485760000000008e-08     evaluation
reward: 219.9
episode: 1424   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 428    lr: 1.0485760000000008e-08     evaluation
reward: 219.8
episode: 1425   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 574    lr: 1.0485760000000008e-08     evaluation
reward: 218.75
episode: 1426   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 522    lr: 1.0485760000000008e-08     evaluation
reward: 218.2
episode: 1427   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 893    lr: 1.0485760000000008e-08     evaluation
reward: 219.7
episode: 1428   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 341    lr: 1.0485760000000008e-08     evaluation
reward: 218.25
episode: 1429   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 643    lr: 1.0485760000000008e-08     evaluation
reward: 219.15
episode: 1430   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 550    lr: 1.0485760000000008e-08     evaluation
reward: 219.75
episode: 1431   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 591    lr: 1.0485760000000008e-08     evaluation
reward: 219.75
episode: 1432   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1108    lr: 1.0485760000000008e-08     evaluation
reward: 219.9
```

episode: 1433 score: 235.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 758 lr: 1.0485760000000008e-08 evaluation
reward: 218.55
episode: 1434 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1099 lr: 1.0485760000000008e-08 evaluation
reward: 219.75
episode: 1435 score: 90.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 415 lr: 1.0485760000000008e-08 evaluation
reward: 219.15
episode: 1436 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 982 lr: 1.0485760000000008e-08 evaluation
reward: 221.9
episode: 1437 score: 265.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1008 lr: 1.0485760000000008e-08 evaluation
reward: 220.35
episode: 1438 score: 120.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 468 lr: 1.0485760000000008e-08 evaluation
reward: 220.05
episode: 1439 score: 160.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 839 lr: 1.0485760000000008e-08 evaluation
reward: 220.35
episode: 1440 score: 370.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1144 lr: 1.0485760000000008e-08 evaluation
reward: 223.25
episode: 1441 score: 100.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 542 lr: 1.0485760000000008e-08 evaluation
reward: 223.65
episode: 1442 score: 100.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 493 lr: 1.0485760000000008e-08 evaluation
reward: 222.45
episode: 1443 score: 160.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 678 lr: 1.0485760000000008e-08 evaluation
reward: 222.3
episode: 1444 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 854 lr: 1.0485760000000008e-08 evaluation
reward: 211.05
episode: 1445 score: 70.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 323 lr: 1.0485760000000008e-08 evaluation
reward: 209.7
episode: 1446 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 885 lr: 1.0485760000000008e-08 evaluation
reward: 208.95
episode: 1447 score: 130.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 646 lr: 1.0485760000000008e-08 evaluation
reward: 207.75
episode: 1448 score: 325.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1036 lr: 1.0485760000000008e-08 evaluation
reward: 209.7

episode: 1449   score: 355.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 936   lr: 1.0485760000000008e-08   evaluation
reward: 210.9
episode: 1450   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 432   lr: 1.0485760000000008e-08   evaluation
reward: 210.9
episode: 1451   score: 355.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1029   lr: 1.0485760000000008e-08   evaluation
reward: 213.15
episode: 1452   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 992   lr: 1.0485760000000008e-08   evaluation
reward: 213.3
episode: 1453   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 727   lr: 1.0485760000000008e-08   evaluation
reward: 201.05
episode: 1454   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 638   lr: 1.0485760000000008e-08   evaluation
reward: 200.65
episode: 1455   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 985   lr: 1.0485760000000008e-08   evaluation
reward: 201.55
episode: 1456   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 902   lr: 1.0485760000000008e-08   evaluation
reward: 203.2
episode: 1457   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 854   lr: 1.0485760000000008e-08   evaluation
reward: 204.55
episode: 1458   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 802   lr: 1.0485760000000008e-08   evaluation
reward: 204.9
episode: 1459   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 929   lr: 1.0485760000000008e-08   evaluation
reward: 204.45
episode: 1460   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 784   lr: 1.0485760000000008e-08   evaluation
reward: 203.25
episode: 1461   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 574   lr: 1.0485760000000008e-08   evaluation
reward: 203.65
episode: 1462   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1096   lr: 1.0485760000000008e-08   evaluation
reward: 205.0
episode: 1463   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 493   lr: 1.0485760000000008e-08   evaluation
reward: 202.5
episode: 1464   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 519   lr: 1.0485760000000008e-08   evaluation
reward: 201.9

episode: 1465   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1405   lr: 1.0485760000000008e-08   evaluation
reward: 203.4
episode: 1466   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 698   lr: 1.0485760000000008e-08   evaluation
reward: 203.1
episode: 1467   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 711   lr: 1.0485760000000008e-08   evaluation
reward: 201.9
episode: 1468   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 785   lr: 1.0485760000000008e-08   evaluation
reward: 203.1
episode: 1469   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1241   lr: 1.0485760000000008e-08   evaluation
reward: 201.45
episode: 1470   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 896   lr: 1.0485760000000008e-08   evaluation
reward: 201.0
episode: 1471   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 351   lr: 1.0485760000000008e-08   evaluation
reward: 199.35
episode: 1472   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 791   lr: 1.0485760000000008e-08   evaluation
reward: 199.55
episode: 1473   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1303   lr: 1.0485760000000008e-08   evaluation
reward: 201.3
episode: 1474   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 910   lr: 1.0485760000000008e-08   evaluation
reward: 202.15
episode: 1475   score: 385.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1051   lr: 1.0485760000000008e-08   evaluation
reward: 204.9
episode: 1476   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 424   lr: 1.0485760000000008e-08   evaluation
reward: 202.95
episode: 1477   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 613   lr: 1.0485760000000008e-08   evaluation
reward: 203.05
episode: 1478   score: 60.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 298   lr: 1.0485760000000008e-08   evaluation
reward: 201.75
episode: 1479   score: 420.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1215   lr: 1.0485760000000008e-08   evaluation
reward: 204.05
episode: 1480   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 530   lr: 1.0485760000000008e-08   evaluation
reward: 204.85

episode: 1481   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 472   lr: 1.0485760000000008e-08   evaluation
reward: 205.95
episode: 1482   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 460   lr: 1.0485760000000008e-08   evaluation
reward: 203.2
episode: 1483   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 930   lr: 1.0485760000000008e-08   evaluation
reward: 203.65
episode: 1484   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 991   lr: 1.0485760000000008e-08   evaluation
reward: 205.5
episode: 1485   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1041   lr: 1.0485760000000008e-08   evaluation
reward: 205.5
episode: 1486   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 881   lr: 1.0485760000000008e-08   evaluation
reward: 206.55
episode: 1487   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 883   lr: 1.0485760000000008e-08   evaluation
reward: 205.95
episode: 1488   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 398   lr: 1.0485760000000008e-08   evaluation
reward: 203.55
episode: 1489   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 428   lr: 1.0485760000000008e-08   evaluation
reward: 202.45
episode: 1490   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 962   lr: 1.0485760000000008e-08   evaluation
reward: 203.5
episode: 1491   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 788   lr: 1.0485760000000008e-08   evaluation
reward: 203.2
episode: 1492   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 587   lr: 1.0485760000000008e-08   evaluation
reward: 203.1
episode: 1493   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 810   lr: 1.0485760000000008e-08   evaluation
reward: 203.25
episode: 1494   score: 370.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1350   lr: 1.0485760000000008e-08   evaluation
reward: 204.45
episode: 1495   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 799   lr: 1.0485760000000008e-08   evaluation
reward: 205.2
episode: 1496   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 861   lr: 1.0485760000000008e-08   evaluation
reward: 205.55

episode: 1497   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 352   lr: 1.0485760000000008e-08   evaluation
reward: 204.2
episode: 1498   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 311   lr: 1.0485760000000008e-08   evaluation
reward: 203.4
episode: 1499   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 774   lr: 1.0485760000000008e-08   evaluation
reward: 204.1
episode: 1500   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1057   lr: 1.0485760000000008e-08   evaluation
reward: 203.2
episode: 1501   score: 1200.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1508   lr: 1.0485760000000008e-08   evaluation
reward: 212.25
episode: 1502   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 739   lr: 1.0485760000000008e-08   evaluation
reward: 212.55
episode: 1503   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 561   lr: 1.0485760000000008e-08   evaluation
reward: 211.25
episode: 1504   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 974   lr: 1.0485760000000008e-08   evaluation
reward: 212.15
episode: 1505   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 608   lr: 1.0485760000000008e-08   evaluation
reward: 213.05
episode: 1506   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 808   lr: 1.0485760000000008e-08   evaluation
reward: 212.45
episode: 1507   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1213   lr: 4.194304000000004e-09   evaluation
reward: 212.0
episode: 1508   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 461   lr: 4.194304000000004e-09   evaluation
reward: 211.9
episode: 1509   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 594   lr: 4.194304000000004e-09   evaluation
reward: 210.15
episode: 1510   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1091   lr: 4.194304000000004e-09   evaluation
reward: 210.15
episode: 1511   score: 50.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 300   lr: 4.194304000000004e-09   evaluation
reward: 208.45
episode: 1512   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 410   lr: 4.194304000000004e-09   evaluation
reward: 208.25

```
episode: 1513   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1227    lr: 4.194304000000004e-09     evaluation
reward: 207.5
episode: 1514   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 599    lr: 4.194304000000004e-09     evaluation
reward: 207.9
episode: 1515   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 859    lr: 4.194304000000004e-09     evaluation
reward: 207.45
episode: 1516   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 597    lr: 4.194304000000004e-09     evaluation
reward: 207.65
episode: 1517   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 794    lr: 4.194304000000004e-09     evaluation
reward: 208.25
episode: 1518   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 648    lr: 4.194304000000004e-09     evaluation
reward: 206.6
episode: 1519   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 398    lr: 4.194304000000004e-09     evaluation
reward: 203.2
episode: 1520   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 839    lr: 4.194304000000004e-09     evaluation
reward: 204.7
episode: 1521   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1029    lr: 4.194304000000004e-09     evaluation
reward: 205.95
episode: 1522   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 565    lr: 4.194304000000004e-09     evaluation
reward: 205.35
episode: 1523   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 864    lr: 4.194304000000004e-09     evaluation
reward: 206.4
episode: 1524   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 457    lr: 4.194304000000004e-09     evaluation
reward: 206.9
episode: 1525   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 622    lr: 4.194304000000004e-09     evaluation
reward: 207.1
episode: 1526   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 457    lr: 4.194304000000004e-09     evaluation
reward: 206.8
episode: 1527   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 465    lr: 4.194304000000004e-09     evaluation
reward: 205.3
episode: 1528   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 408    lr: 4.194304000000004e-09     evaluation
reward: 205.3
```

episode: 1529 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 847 lr: 4.194304000000004e-09 evaluation
reward: 206.35
episode: 1530 score: 265.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 870 lr: 4.194304000000004e-09 evaluation
reward: 207.7
episode: 1531 score: 110.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 642 lr: 4.194304000000004e-09 evaluation
reward: 206.9
episode: 1532 score: 190.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 722 lr: 4.194304000000004e-09 evaluation
reward: 205.55
episode: 1533 score: 265.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 850 lr: 4.194304000000004e-09 evaluation
reward: 205.85
episode: 1534 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 731 lr: 4.194304000000004e-09 evaluation
reward: 204.65
episode: 1535 score: 325.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1223 lr: 4.194304000000004e-09 evaluation
reward: 207.0
episode: 1536 score: 140.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 681 lr: 4.194304000000004e-09 evaluation
reward: 205.45
episode: 1537 score: 280.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1031 lr: 4.194304000000004e-09 evaluation
reward: 205.6
episode: 1538 score: 140.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 554 lr: 4.194304000000004e-09 evaluation
reward: 205.8
episode: 1539 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 693 lr: 4.194304000000004e-09 evaluation
reward: 205.95
episode: 1540 score: 30.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 256 lr: 4.194304000000004e-09 evaluation
reward: 202.55
episode: 1541 score: 370.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1074 lr: 4.194304000000004e-09 evaluation
reward: 205.25
episode: 1542 score: 355.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 923 lr: 4.194304000000004e-09 evaluation
reward: 207.8
episode: 1543 score: 580.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1189 lr: 4.194304000000004e-09 evaluation
reward: 212.0
episode: 1544 score: 325.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1019 lr: 4.194304000000004e-09 evaluation
reward: 213.5

episode: 1545 score: 325.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 982 lr: 4.194304000000004e-09 evaluation
reward: 216.05
episode: 1546 score: 190.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 731 lr: 4.194304000000004e-09 evaluation
reward: 216.2
episode: 1547 score: 420.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1356 lr: 4.194304000000004e-09 evaluation
reward: 219.1
episode: 1548 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 563 lr: 4.194304000000004e-09 evaluation
reward: 217.6
episode: 1549 score: 280.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 969 lr: 4.194304000000004e-09 evaluation
reward: 216.85
episode: 1550 score: 190.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 808 lr: 4.194304000000004e-09 evaluation
reward: 217.75
episode: 1551 score: 80.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 462 lr: 4.194304000000004e-09 evaluation
reward: 215.0
episode: 1552 score: 520.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1390 lr: 4.194304000000004e-09 evaluation
reward: 217.1
episode: 1553 score: 100.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 640 lr: 4.194304000000004e-09 evaluation
reward: 216.35
episode: 1554 score: 90.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 335 lr: 4.194304000000004e-09 evaluation
reward: 215.75
episode: 1555 score: 150.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 524 lr: 4.194304000000004e-09 evaluation
reward: 214.6
episode: 1556 score: 370.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1150 lr: 4.194304000000004e-09 evaluation
reward: 215.05
episode: 1557 score: 120.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 583 lr: 4.194304000000004e-09 evaluation
reward: 213.6
episode: 1558 score: 150.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 532 lr: 4.194304000000004e-09 evaluation
reward: 213.35
episode: 1559 score: 370.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1137 lr: 4.194304000000004e-09 evaluation
reward: 215.0
episode: 1560 score: 70.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 323 lr: 4.194304000000004e-09 evaluation
reward: 213.8

```
episode: 1561   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 568    lr: 4.194304000000004e-09     evaluation
reward: 213.5
episode: 1562   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 775    lr: 4.194304000000004e-09     evaluation
reward: 211.95
episode: 1563   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 392    lr: 4.194304000000004e-09     evaluation
reward: 211.35
episode: 1564   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 721    lr: 4.194304000000004e-09     evaluation
reward: 212.2
episode: 1565   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 651    lr: 4.194304000000004e-09     evaluation
reward: 210.5
episode: 1566   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 474    lr: 4.194304000000004e-09     evaluation
reward: 209.9
episode: 1567   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 706    lr: 4.194304000000004e-09     evaluation
reward: 210.35
episode: 1568   score: 740.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1201    lr: 4.194304000000004e-09     evaluation
reward: 214.8
episode: 1569   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 610    lr: 4.194304000000004e-09     evaluation
reward: 213.35
episode: 1570   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 558    lr: 4.194304000000004e-09     evaluation
reward: 212.1
episode: 1571   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 919    lr: 4.194304000000004e-09     evaluation
reward: 213.3
episode: 1572   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1078    lr: 4.194304000000004e-09     evaluation
reward: 214.2
episode: 1573   score: 60.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 286    lr: 4.194304000000004e-09     evaluation
reward: 211.55
episode: 1574   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 463    lr: 4.194304000000004e-09     evaluation
reward: 210.8
episode: 1575   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 431    lr: 4.194304000000004e-09     evaluation
reward: 207.95
episode: 1576   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 448    lr: 4.194304000000004e-09     evaluation
reward: 208.15
```

```
episode: 1577   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 532    lr: 4.194304000000004e-09     evaluation
reward: 208.25
episode: 1578   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 979    lr: 4.194304000000004e-09     evaluation
reward: 210.3
episode: 1579   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 448    lr: 4.194304000000004e-09     evaluation
reward: 207.2
episode: 1580   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 866    lr: 4.194304000000004e-09     evaluation
reward: 207.45
episode: 1581   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 861    lr: 4.194304000000004e-09     evaluation
reward: 208.0
episode: 1582   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 624    lr: 4.194304000000004e-09     evaluation
reward: 208.65
episode: 1583   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 901    lr: 4.194304000000004e-09     evaluation
reward: 209.25
episode: 1584   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 817    lr: 4.194304000000004e-09     evaluation
reward: 207.9
episode: 1585   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1045    lr: 4.194304000000004e-09     evaluation
reward: 208.05
episode: 1586   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 735    lr: 4.194304000000004e-09     evaluation
reward: 207.2
episode: 1587   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 566    lr: 4.194304000000004e-09     evaluation
reward: 205.8
episode: 1588   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 572    lr: 4.194304000000004e-09     evaluation
reward: 207.0
episode: 1589   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1123    lr: 4.194304000000004e-09     evaluation
reward: 209.3
episode: 1590   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 600    lr: 4.194304000000004e-09     evaluation
reward: 208.75
episode: 1591   score: 780.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1618    lr: 4.194304000000004e-09     evaluation
reward: 214.65
episode: 1592   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 390    lr: 4.194304000000004e-09     evaluation
reward: 214.85
```

```
episode: 1593   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 552    lr: 4.194304000000004e-09     evaluation
reward: 213.65
episode: 1594   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 882    lr: 4.194304000000004e-09     evaluation
reward: 212.45
episode: 1595   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 785    lr: 4.194304000000004e-09     evaluation
reward: 212.6
episode: 1596   score: 60.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 381    lr: 4.194304000000004e-09     evaluation
reward: 211.45
episode: 1597   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 459    lr: 4.194304000000004e-09     evaluation
reward: 211.65
episode: 1598   score: 385.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1208    lr: 4.194304000000004e-09     evaluation
reward: 214.4
episode: 1599   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 777    lr: 4.194304000000004e-09     evaluation
reward: 214.55
episode: 1600   score: 20.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 201    lr: 4.194304000000004e-09     evaluation
reward: 212.55
episode: 1601   score: 540.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1503    lr: 4.194304000000004e-09     evaluation
reward: 205.95
episode: 1602   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 555    lr: 4.194304000000004e-09     evaluation
reward: 206.15
episode: 1603   score: 340.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 998    lr: 4.194304000000004e-09     evaluation
reward: 208.65
episode: 1604   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 926    lr: 4.194304000000004e-09     evaluation
reward: 207.45
episode: 1605   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 910    lr: 4.194304000000004e-09     evaluation
reward: 208.55
episode: 1606   score: 50.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 230    lr: 4.194304000000004e-09     evaluation
reward: 206.7
episode: 1607   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 557    lr: 4.194304000000004e-09     evaluation
reward: 205.05
episode: 1608   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 595    lr: 4.194304000000004e-09     evaluation
reward: 204.75
```

episode: 1609   score: 340.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1099   lr: 4.194304000000004e-09   evaluation
reward: 206.65
episode: 1610   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 711   lr: 4.194304000000004e-09   evaluation
reward: 206.05
episode: 1611   score: 50.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 261   lr: 4.194304000000004e-09   evaluation
reward: 206.05
episode: 1612   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 476   lr: 4.194304000000004e-09   evaluation
reward: 206.55
episode: 1613   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 684   lr: 4.194304000000004e-09   evaluation
reward: 205.2
episode: 1614   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1191   lr: 4.194304000000004e-09   evaluation
reward: 206.7
episode: 1615   score: 340.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1251   lr: 4.194304000000004e-09   evaluation
reward: 207.45
episode: 1616   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1040   lr: 4.194304000000004e-09   evaluation
reward: 209.1
episode: 1617   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 810   lr: 4.194304000000004e-09   evaluation
reward: 208.2
episode: 1618   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1066   lr: 4.194304000000004e-09   evaluation
reward: 209.4
episode: 1619   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 679   lr: 4.194304000000004e-09   evaluation
reward: 209.6
episode: 1620   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 828   lr: 4.194304000000004e-09   evaluation
reward: 209.6
episode: 1621   score: 60.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 364   lr: 4.194304000000004e-09   evaluation
reward: 207.85
episode: 1622   score: 355.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1176   lr: 4.194304000000004e-09   evaluation
reward: 210.1
episode: 1623   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 878   lr: 4.194304000000004e-09   evaluation
reward: 210.1
episode: 1624   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 773   lr: 4.194304000000004e-09   evaluation
reward: 211.15

```
episode: 1625   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 637    lr: 4.194304000000004e-09     evaluation
reward: 211.85
episode: 1626   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 384    lr: 4.194304000000004e-09     evaluation
reward: 211.85
episode: 1627   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 993    lr: 4.194304000000004e-09     evaluation
reward: 213.95
episode: 1628   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 558    lr: 4.194304000000004e-09     evaluation
reward: 214.55
episode: 1629   score: 500.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1393    lr: 4.194304000000004e-09     evaluation
reward: 216.6
episode: 1630   score: 980.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1138    lr: 4.194304000000004e-09     evaluation
reward: 223.75
episode: 1631   score: 385.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1178    lr: 4.194304000000004e-09     evaluation
reward: 226.5
episode: 1632   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 626    lr: 4.194304000000004e-09     evaluation
reward: 225.6
episode: 1633   score: 720.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1390    lr: 4.194304000000004e-09     evaluation
reward: 230.15
episode: 1634   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 454    lr: 4.194304000000004e-09     evaluation
reward: 229.6
episode: 1635   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 631    lr: 4.194304000000004e-09     evaluation
reward: 227.85
episode: 1636   score: 440.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1623    lr: 4.194304000000004e-09     evaluation
reward: 230.85
episode: 1637   score: 50.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 300    lr: 4.194304000000004e-09     evaluation
reward: 228.55
episode: 1638   score: 60.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 316    lr: 4.194304000000004e-09     evaluation
reward: 227.75
episode: 1639   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 604    lr: 4.194304000000004e-09     evaluation
reward: 226.9
episode: 1640   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 566    lr: 1.6777216000000016e-09     evaluation
reward: 228.1
```

episode: 1641 score: 80.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 519 lr: 1.6777216000000016e-09 evaluation
reward: 225.2
episode: 1642 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 775 lr: 1.6777216000000016e-09 evaluation
reward: 223.4
episode: 1643 score: 90.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 469 lr: 1.6777216000000016e-09 evaluation
reward: 218.5
episode: 1644 score: 190.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 693 lr: 1.6777216000000016e-09 evaluation
reward: 217.15
episode: 1645 score: 100.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 392 lr: 1.6777216000000016e-09 evaluation
reward: 214.9
episode: 1646 score: 205.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 865 lr: 1.6777216000000016e-09 evaluation
reward: 215.05
episode: 1647 score: 325.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1114 lr: 1.6777216000000016e-09 evaluation
reward: 214.1
episode: 1648 score: 190.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 625 lr: 1.6777216000000016e-09 evaluation
reward: 214.25
episode: 1649 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1073 lr: 1.6777216000000016e-09 evaluation
reward: 214.4
episode: 1650 score: 70.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 352 lr: 1.6777216000000016e-09 evaluation
reward: 213.2
episode: 1651 score: 110.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 354 lr: 1.6777216000000016e-09 evaluation
reward: 213.5
episode: 1652 score: 160.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 627 lr: 1.6777216000000016e-09 evaluation
reward: 209.9
episode: 1653 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1233 lr: 1.6777216000000016e-09 evaluation
reward: 211.85
episode: 1654 score: 400.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1317 lr: 1.6777216000000016e-09 evaluation
reward: 214.95
episode: 1655 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 734 lr: 1.6777216000000016e-09 evaluation
reward: 215.2
episode: 1656 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 702 lr: 1.6777216000000016e-09 evaluation
reward: 213.7

episode: 1657 score: 150.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 820 lr: 1.6777216000000016e-09 evaluation
reward: 214.0
episode: 1658 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1082 lr: 1.6777216000000016e-09 evaluation
reward: 215.45
episode: 1659 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 707 lr: 1.6777216000000016e-09 evaluation
reward: 213.5
episode: 1660 score: 190.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 662 lr: 1.6777216000000016e-09 evaluation
reward: 214.7
episode: 1661 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 594 lr: 1.6777216000000016e-09 evaluation
reward: 215.3
episode: 1662 score: 120.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 327 lr: 1.6777216000000016e-09 evaluation
reward: 215.1
episode: 1663 score: 80.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 620 lr: 1.6777216000000016e-09 evaluation
reward: 215.0
episode: 1664 score: 100.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 513 lr: 1.6777216000000016e-09 evaluation
reward: 214.25
episode: 1665 score: 280.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 924 lr: 1.6777216000000016e-09 evaluation
reward: 215.65
episode: 1666 score: 340.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1147 lr: 1.6777216000000016e-09 evaluation
reward: 217.3
episode: 1667 score: 190.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 691 lr: 1.6777216000000016e-09 evaluation
reward: 216.85
episode: 1668 score: 160.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 603 lr: 1.6777216000000016e-09 evaluation
reward: 211.05
episode: 1669 score: 310.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 927 lr: 1.6777216000000016e-09 evaluation
reward: 212.65
episode: 1670 score: 190.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 824 lr: 1.6777216000000016e-09 evaluation
reward: 213.75
episode: 1671 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1058 lr: 1.6777216000000016e-09 evaluation
reward: 214.05
episode: 1672 score: 110.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 470 lr: 1.6777216000000016e-09 evaluation
reward: 212.65

episode: 1673 score: 100.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 455 lr: 1.6777216000000016e-09 evaluation
reward: 213.05
episode: 1674 score: 130.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 616 lr: 1.6777216000000016e-09 evaluation
reward: 213.05
episode: 1675 score: 140.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 657 lr: 1.6777216000000016e-09 evaluation
reward: 213.45
episode: 1676 score: 120.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 464 lr: 1.6777216000000016e-09 evaluation
reward: 213.75
episode: 1677 score: 140.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 530 lr: 1.6777216000000016e-09 evaluation
reward: 214.05
episode: 1678 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 577 lr: 1.6777216000000016e-09 evaluation
reward: 213.15
episode: 1679 score: 205.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 921 lr: 1.6777216000000016e-09 evaluation
reward: 214.1
episode: 1680 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 865 lr: 1.6777216000000016e-09 evaluation
reward: 215.3
episode: 1681 score: 90.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 545 lr: 1.6777216000000016e-09 evaluation
reward: 214.15
episode: 1682 score: 90.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 460 lr: 1.6777216000000016e-09 evaluation
reward: 213.3
episode: 1683 score: 150.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 586 lr: 1.6777216000000016e-09 evaluation
reward: 211.7
episode: 1684 score: 150.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 418 lr: 1.6777216000000016e-09 evaluation
reward: 211.3
episode: 1685 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 636 lr: 1.6777216000000016e-09 evaluation
reward: 210.1
episode: 1686 score: 205.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 752 lr: 1.6777216000000016e-09 evaluation
reward: 210.65
episode: 1687 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 759 lr: 1.6777216000000016e-09 evaluation
reward: 211.75
episode: 1688 score: 235.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 667 lr: 1.6777216000000016e-09 evaluation
reward: 212.2

```
episode: 1689   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 696    lr: 1.6777216000000016e-09     evaluation
reward: 210.7
episode: 1690   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 489    lr: 1.6777216000000016e-09     evaluation
reward: 210.5
episode: 1691   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 628    lr: 1.6777216000000016e-09     evaluation
reward: 203.6
episode: 1692   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1008    lr: 1.6777216000000016e-09     evaluation
reward: 205.85
episode: 1693   score: 520.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1160    lr: 1.6777216000000016e-09     evaluation
reward: 209.75
episode: 1694   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 817    lr: 1.6777216000000016e-09     evaluation
reward: 209.6
episode: 1695   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 594    lr: 1.6777216000000016e-09     evaluation
reward: 208.7
episode: 1696   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 373    lr: 1.6777216000000016e-09     evaluation
reward: 208.9
episode: 1697   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 650    lr: 1.6777216000000016e-09     evaluation
reward: 209.45
episode: 1698   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 994    lr: 1.6777216000000016e-09     evaluation
reward: 207.95
episode: 1699   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 428    lr: 1.6777216000000016e-09     evaluation
reward: 207.1
episode: 1700   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 786    lr: 1.6777216000000016e-09     evaluation
reward: 208.95
episode: 1701   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 935    lr: 1.6777216000000016e-09     evaluation
reward: 206.5
episode: 1702   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 851    lr: 1.6777216000000016e-09     evaluation
reward: 206.85
episode: 1703   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 590    lr: 1.6777216000000016e-09     evaluation
reward: 205.05
episode: 1704   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 713    lr: 1.6777216000000016e-09     evaluation
reward: 205.2
```

```
episode: 1705   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 754    lr: 1.6777216000000016e-09     evaluation
reward: 205.35
episode: 1706   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 674    lr: 1.6777216000000016e-09     evaluation
reward: 206.9
episode: 1707   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 652    lr: 1.6777216000000016e-09     evaluation
reward: 206.8
episode: 1708   score: 370.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1104    lr: 1.6777216000000016e-09     evaluation
reward: 209.3
episode: 1709   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 461    lr: 1.6777216000000016e-09     evaluation
reward: 207.0
episode: 1710   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 917    lr: 1.6777216000000016e-09     evaluation
reward: 207.9
episode: 1711   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 436    lr: 1.6777216000000016e-09     evaluation
reward: 208.6
episode: 1712   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1173    lr: 1.6777216000000016e-09     evaluation
reward: 210.45
episode: 1713   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1015    lr: 1.6777216000000016e-09     evaluation
reward: 210.6
episode: 1714   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 727    lr: 1.6777216000000016e-09     evaluation
reward: 209.7
episode: 1715   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 970    lr: 1.6777216000000016e-09     evaluation
reward: 209.1
episode: 1716   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 966    lr: 1.6777216000000016e-09     evaluation
reward: 208.65
episode: 1717   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 582    lr: 1.6777216000000016e-09     evaluation
reward: 208.95
episode: 1718   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1033    lr: 1.6777216000000016e-09     evaluation
reward: 209.25
episode: 1719   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 935    lr: 1.6777216000000016e-09     evaluation
reward: 210.25
episode: 1720   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1343    lr: 1.6777216000000016e-09     evaluation
reward: 211.6
```

```
episode: 1721   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 679    lr: 1.6777216000000016e-09     evaluation
reward: 212.5
episode: 1722   score: 340.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1080    lr: 1.6777216000000016e-09     evaluation
reward: 212.35
episode: 1723   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 540    lr: 1.6777216000000016e-09     evaluation
reward: 212.5
episode: 1724   score: 355.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 932    lr: 1.6777216000000016e-09     evaluation
reward: 213.7
episode: 1725   score: 385.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1015    lr: 1.6777216000000016e-09     evaluation
reward: 215.35
episode: 1726   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 696    lr: 1.6777216000000016e-09     evaluation
reward: 216.05
episode: 1727   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 873    lr: 1.6777216000000016e-09     evaluation
reward: 215.6
episode: 1728   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 810    lr: 1.6777216000000016e-09     evaluation
reward: 215.85
episode: 1729   score: 370.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1288    lr: 1.6777216000000016e-09     evaluation
reward: 214.55
episode: 1730   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1018    lr: 1.6777216000000016e-09     evaluation
reward: 207.25
episode: 1731   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1096    lr: 1.6777216000000016e-09     evaluation
reward: 206.65
episode: 1732   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 861    lr: 1.6777216000000016e-09     evaluation
reward: 207.55
episode: 1733   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 703    lr: 1.6777216000000016e-09     evaluation
reward: 202.1
episode: 1734   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 735    lr: 1.6777216000000016e-09     evaluation
reward: 202.65
episode: 1735   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 513    lr: 1.6777216000000016e-09     evaluation
reward: 202.35
episode: 1736   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 999    lr: 1.6777216000000016e-09     evaluation
reward: 200.75
```

episode: 1737   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 782   lr: 1.6777216000000016e-09   evaluation
reward: 202.75
episode: 1738   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 372   lr: 1.6777216000000016e-09   evaluation
reward: 202.85
episode: 1739   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 436   lr: 1.6777216000000016e-09   evaluation
reward: 203.15
episode: 1740   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 563   lr: 1.6777216000000016e-09   evaluation
reward: 203.15
episode: 1741   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 982   lr: 1.6777216000000016e-09   evaluation
reward: 205.15
episode: 1742   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 508   lr: 1.6777216000000016e-09   evaluation
reward: 204.5
episode: 1743   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 599   lr: 1.6777216000000016e-09   evaluation
reward: 205.65
episode: 1744   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 879   lr: 1.6777216000000016e-09   evaluation
reward: 206.4
episode: 1745   score: 400.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1038   lr: 1.6777216000000016e-09   evaluation
reward: 209.4
episode: 1746   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 917   lr: 1.6777216000000016e-09   evaluation
reward: 209.1
episode: 1747   score: 460.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1416   lr: 1.6777216000000016e-09   evaluation
reward: 210.45
episode: 1748   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 774   lr: 1.6777216000000016e-09   evaluation
reward: 211.8
episode: 1749   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 726   lr: 1.6777216000000016e-09   evaluation
reward: 210.9
episode: 1750   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 558   lr: 1.6777216000000016e-09   evaluation
reward: 211.6
episode: 1751   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1156   lr: 1.6777216000000016e-09   evaluation
reward: 213.45
episode: 1752   score: 355.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1564   lr: 1.6777216000000016e-09   evaluation
reward: 215.4

```
episode: 1753   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 549    lr: 1.6777216000000016e-09     evaluation
reward: 214.2
episode: 1754   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 968    lr: 1.6777216000000016e-09     evaluation
reward: 212.7
episode: 1755   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1007    lr: 1.6777216000000016e-09     evaluation
reward: 213.15
episode: 1756   score: 50.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 293    lr: 1.6777216000000016e-09     evaluation
reward: 211.45
episode: 1757   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 699    lr: 1.6777216000000016e-09     evaluation
reward: 211.7
episode: 1758   score: 1100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1378    lr: 1.6777216000000016e-09     evaluation
reward: 219.75
episode: 1759   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 550    lr: 1.6777216000000016e-09     evaluation
reward: 219.1
episode: 1760   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 997    lr: 1.6777216000000016e-09     evaluation
reward: 220.15
episode: 1761   score: 355.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1245    lr: 1.6777216000000016e-09     evaluation
reward: 221.5
episode: 1762   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 929    lr: 1.6777216000000016e-09     evaluation
reward: 222.35
episode: 1763   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 622    lr: 1.6777216000000016e-09     evaluation
reward: 222.65
episode: 1764   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 583    lr: 1.6777216000000016e-09     evaluation
reward: 223.4
episode: 1765   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 455    lr: 1.6777216000000016e-09     evaluation
reward: 221.5
episode: 1766   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 556    lr: 1.6777216000000016e-09     evaluation
reward: 219.3
episode: 1767   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 922    lr: 1.6777216000000016e-09     evaluation
reward: 220.65
episode: 1768   score: 370.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1212    lr: 1.6777216000000016e-09     evaluation
reward: 222.75
```

episode: 1769   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 939   lr: 6.710886400000007e-10   evaluation
reward: 220.45
episode: 1770   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 638   lr: 6.710886400000007e-10   evaluation
reward: 219.75
episode: 1771   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 702   lr: 6.710886400000007e-10   evaluation
reward: 219.05
episode: 1772   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 822   lr: 6.710886400000007e-10   evaluation
reward: 220.75
episode: 1773   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1070   lr: 6.710886400000007e-10   evaluation
reward: 222.7
episode: 1774   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1254   lr: 6.710886400000007e-10   evaluation
reward: 224.35
episode: 1775   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 863   lr: 6.710886400000007e-10   evaluation
reward: 225.15
episode: 1776   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 955   lr: 6.710886400000007e-10   evaluation
reward: 227.2
episode: 1777   score: 520.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 964   lr: 6.710886400000007e-10   evaluation
reward: 231.0
episode: 1778   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 535   lr: 6.710886400000007e-10   evaluation
reward: 230.65
episode: 1779   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 820   lr: 6.710886400000007e-10   evaluation
reward: 230.65
episode: 1780   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 714   lr: 6.710886400000007e-10   evaluation
reward: 229.3
episode: 1781   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1068   lr: 6.710886400000007e-10   evaluation
reward: 231.35
episode: 1782   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 669   lr: 6.710886400000007e-10   evaluation
reward: 232.5
episode: 1783   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1227   lr: 6.710886400000007e-10   evaluation
reward: 233.95
episode: 1784   score: 860.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1146   lr: 6.710886400000007e-10   evaluation
reward: 241.05

episode: 1785 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1101 lr: 6.710886400000007e-10 evaluation
reward: 241.5
episode: 1786 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 536 lr: 6.710886400000007e-10 evaluation
reward: 241.2
episode: 1787 score: 280.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1051 lr: 6.710886400000007e-10 evaluation
reward: 241.8
episode: 1788 score: 120.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 519 lr: 6.710886400000007e-10 evaluation
reward: 240.65
episode: 1789 score: 90.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 508 lr: 6.710886400000007e-10 evaluation
reward: 239.95
episode: 1790 score: 100.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 550 lr: 6.710886400000007e-10 evaluation
reward: 239.65
episode: 1791 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 776 lr: 6.710886400000007e-10 evaluation
reward: 241.7
episode: 1792 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 837 lr: 6.710886400000007e-10 evaluation
reward: 240.65
episode: 1793 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 922 lr: 6.710886400000007e-10 evaluation
reward: 237.2
episode: 1794 score: 160.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 691 lr: 6.710886400000007e-10 evaluation
reward: 236.45
episode: 1795 score: 130.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 525 lr: 6.710886400000007e-10 evaluation
reward: 236.0
episode: 1796 score: 80.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 436 lr: 6.710886400000007e-10 evaluation
reward: 236.0
episode: 1797 score: 110.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 371 lr: 6.710886400000007e-10 evaluation
reward: 235.35
episode: 1798 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 659 lr: 6.710886400000007e-10 evaluation
reward: 235.2
episode: 1799 score: 325.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1099 lr: 6.710886400000007e-10 evaluation
reward: 237.55
episode: 1800 score: 120.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 589 lr: 6.710886400000007e-10 evaluation
reward: 236.7

episode: 1801 score: 420.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1168 lr: 6.710886400000007e-10 evaluation
reward: 237.95
episode: 1802 score: 340.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1016 lr: 6.710886400000007e-10 evaluation
reward: 239.6
episode: 1803 score: 140.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 613 lr: 6.710886400000007e-10 evaluation
reward: 239.4
episode: 1804 score: 580.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1344 lr: 6.710886400000007e-10 evaluation
reward: 243.15
episode: 1805 score: 90.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 506 lr: 6.710886400000007e-10 evaluation
reward: 241.4
episode: 1806 score: 355.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1168 lr: 6.710886400000007e-10 evaluation
reward: 242.9
episode: 1807 score: 265.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 919 lr: 6.710886400000007e-10 evaluation
reward: 244.05
episode: 1808 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1001 lr: 6.710886400000007e-10 evaluation
reward: 243.3
episode: 1809 score: 140.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 577 lr: 6.710886400000007e-10 evaluation
reward: 243.6
episode: 1810 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 816 lr: 6.710886400000007e-10 evaluation
reward: 243.0
episode: 1811 score: 385.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1236 lr: 6.710886400000007e-10 evaluation
reward: 245.65
episode: 1812 score: 190.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 774 lr: 6.710886400000007e-10 evaluation
reward: 244.3
episode: 1813 score: 265.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1128 lr: 6.710886400000007e-10 evaluation
reward: 245.2
episode: 1814 score: 110.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 565 lr: 6.710886400000007e-10 evaluation
reward: 244.4
episode: 1815 score: 140.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 543 lr: 6.710886400000007e-10 evaluation
reward: 243.0
episode: 1816 score: 265.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 638 lr: 6.710886400000007e-10 evaluation
reward: 243.15

episode: 1817   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1127   lr: 6.710886400000007e-10   evaluation
reward: 244.05
episode: 1818   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 975   lr: 6.710886400000007e-10   evaluation
reward: 243.15
episode: 1819   score: 385.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1329   lr: 6.710886400000007e-10   evaluation
reward: 244.8
episode: 1820   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 359   lr: 6.710886400000007e-10   evaluation
reward: 242.55
episode: 1821   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 893   lr: 6.710886400000007e-10   evaluation
reward: 242.55
episode: 1822   score: 370.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1114   lr: 6.710886400000007e-10   evaluation
reward: 242.85
episode: 1823   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 460   lr: 6.710886400000007e-10   evaluation
reward: 241.75
episode: 1824   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 755   lr: 6.710886400000007e-10   evaluation
reward: 240.85
episode: 1825   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 628   lr: 6.710886400000007e-10   evaluation
reward: 238.1
episode: 1826   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 841   lr: 6.710886400000007e-10   evaluation
reward: 239.15
episode: 1827   score: 540.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 1453   lr: 6.710886400000007e-10   evaluation
reward: 241.9
episode: 1828   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 975   lr: 6.710886400000007e-10   evaluation
reward: 242.05
episode: 1829   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 689   lr: 6.710886400000007e-10   evaluation
reward: 240.4
episode: 1830   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 435   lr: 6.710886400000007e-10   evaluation
reward: 239.0
episode: 1831   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 895   lr: 6.710886400000007e-10   evaluation
reward: 238.4
episode: 1832   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 446   lr: 6.710886400000007e-10   evaluation
reward: 237.5

episode: 1833   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 509   lr: 6.710886400000007e-10   evaluation
reward: 236.65
episode: 1834   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 729   lr: 6.710886400000007e-10   evaluation
reward: 236.65
episode: 1835   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 841   lr: 6.710886400000007e-10   evaluation
reward: 237.8
episode: 1836   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 523   lr: 6.710886400000007e-10   evaluation
reward: 236.4
episode: 1837   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 853   lr: 6.710886400000007e-10   evaluation
reward: 235.5
episode: 1838   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 429   lr: 6.710886400000007e-10   evaluation
reward: 235.6
episode: 1839   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 803   lr: 6.710886400000007e-10   evaluation
reward: 236.0
episode: 1840   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 703   lr: 6.710886400000007e-10   evaluation
reward: 236.4
episode: 1841   score: 60.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 316   lr: 6.710886400000007e-10   evaluation
reward: 234.2
episode: 1842   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 520   lr: 6.710886400000007e-10   evaluation
reward: 234.3
episode: 1843   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 776   lr: 6.710886400000007e-10   evaluation
reward: 234.15
episode: 1844   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 597   lr: 6.710886400000007e-10   evaluation
reward: 232.7
episode: 1845   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 551   lr: 6.710886400000007e-10   evaluation
reward: 229.7
episode: 1846   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 806   lr: 6.710886400000007e-10   evaluation
reward: 230.15
episode: 1847   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 793   lr: 6.710886400000007e-10   evaluation
reward: 227.6
episode: 1848   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413   steps: 598   lr: 6.710886400000007e-10   evaluation
reward: 225.75

episode: 1849 score: 130.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 487 lr: 6.710886400000007e-10 evaluation
reward: 225.0
episode: 1850 score: 110.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 449 lr: 6.710886400000007e-10 evaluation
reward: 224.7
episode: 1851 score: 295.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 873 lr: 6.710886400000007e-10 evaluation
reward: 224.7
episode: 1852 score: 130.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 468 lr: 6.710886400000007e-10 evaluation
reward: 222.45
episode: 1853 score: 80.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 474 lr: 6.710886400000007e-10 evaluation
reward: 221.5
episode: 1854 score: 80.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 379 lr: 6.710886400000007e-10 evaluation
reward: 219.8
episode: 1855 score: 50.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 273 lr: 6.710886400000007e-10 evaluation
reward: 218.1
episode: 1856 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 729 lr: 6.710886400000007e-10 evaluation
reward: 219.8
episode: 1857 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 690 lr: 6.710886400000007e-10 evaluation
reward: 219.8
episode: 1858 score: 540.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1741 lr: 6.710886400000007e-10 evaluation
reward: 214.2
episode: 1859 score: 280.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 1113 lr: 6.710886400000007e-10 evaluation
reward: 215.9
episode: 1860 score: 120.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 584 lr: 6.710886400000007e-10 evaluation
reward: 214.15
episode: 1861 score: 175.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 807 lr: 6.710886400000007e-10 evaluation
reward: 212.35
episode: 1862 score: 80.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 460 lr: 6.710886400000007e-10 evaluation
reward: 211.1
episode: 1863 score: 220.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 767 lr: 6.710886400000007e-10 evaluation
reward: 212.2
episode: 1864 score: 130.0 memory length: 1000000 epsilon:
0.009998020008555413 steps: 425 lr: 6.710886400000007e-10 evaluation
reward: 211.75

```
episode: 1865   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 965    lr: 6.710886400000007e-10     evaluation
reward: 213.95
episode: 1866   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 534    lr: 6.710886400000007e-10     evaluation
reward: 213.85
episode: 1867   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 765    lr: 6.710886400000007e-10     evaluation
reward: 213.25
episode: 1868   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 917    lr: 6.710886400000007e-10     evaluation
reward: 211.9
episode: 1869   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 820    lr: 6.710886400000007e-10     evaluation
reward: 212.85
episode: 1870   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 564    lr: 6.710886400000007e-10     evaluation
reward: 213.15
episode: 1871   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 886    lr: 6.710886400000007e-10     evaluation
reward: 214.6
episode: 1872   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 725    lr: 6.710886400000007e-10     evaluation
reward: 213.85
episode: 1873   score: 370.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1081    lr: 6.710886400000007e-10     evaluation
reward: 214.6
episode: 1874   score: 355.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1025    lr: 6.710886400000007e-10     evaluation
reward: 215.2
episode: 1875   score: 1200.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1933    lr: 6.710886400000007e-10     evaluation
reward: 225.0
episode: 1876   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 445    lr: 6.710886400000007e-10     evaluation
reward: 222.95
episode: 1877   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 759    lr: 6.710886400000007e-10     evaluation
reward: 219.8
episode: 1878   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 961    lr: 6.710886400000007e-10     evaluation
reward: 220.15
episode: 1879   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 520    lr: 6.710886400000007e-10     evaluation
reward: 219.3
episode: 1880   score: 30.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 199    lr: 6.710886400000007e-10     evaluation
reward: 218.0
```

```
episode: 1881   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 528    lr: 6.710886400000007e-10     evaluation
reward: 216.35
episode: 1882   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 740    lr: 6.710886400000007e-10     evaluation
reward: 215.7
episode: 1883   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 562    lr: 6.710886400000007e-10     evaluation
reward: 214.15
episode: 1884   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 707    lr: 6.710886400000007e-10     evaluation
reward: 207.3
episode: 1885   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1006    lr: 6.710886400000007e-10     evaluation
reward: 208.35
episode: 1886   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 838    lr: 6.710886400000007e-10     evaluation
reward: 209.4
episode: 1887   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 769    lr: 6.710886400000007e-10     evaluation
reward: 208.95
episode: 1888   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 439    lr: 6.710886400000007e-10     evaluation
reward: 208.85
episode: 1889   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 614    lr: 6.710886400000007e-10     evaluation
reward: 209.25
episode: 1890   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 503    lr: 6.710886400000007e-10     evaluation
reward: 209.35
episode: 1891   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 396    lr: 6.710886400000007e-10     evaluation
reward: 207.3
episode: 1892   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1028    lr: 6.710886400000007e-10     evaluation
reward: 208.2
episode: 1893   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 743    lr: 6.710886400000007e-10     evaluation
reward: 208.35
episode: 1894   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 384    lr: 6.710886400000007e-10     evaluation
reward: 207.65
episode: 1902   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1083    lr: 2.684354560000003e-10     evaluation
reward: 204.95
episode: 1903   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 821    lr: 2.684354560000003e-10     evaluation
reward: 205.05
```

```
episode: 1904   score: 700.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1442    lr: 2.684354560000003e-10     evaluation
reward: 206.25
episode: 1905   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 890    lr: 2.684354560000003e-10     evaluation
reward: 207.7
episode: 1906   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 762    lr: 2.684354560000003e-10     evaluation
reward: 206.35
episode: 1907   score: 110.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 429    lr: 2.684354560000003e-10     evaluation
reward: 204.8
episode: 1908   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 470    lr: 2.684354560000003e-10     evaluation
reward: 203.15
episode: 1909   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 523    lr: 2.684354560000003e-10     evaluation
reward: 202.95
episode: 1910   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 325    lr: 2.684354560000003e-10     evaluation
reward: 201.45
episode: 1911   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 728    lr: 2.684354560000003e-10     evaluation
reward: 199.2
episode: 1912   score: 30.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 228    lr: 2.684354560000003e-10     evaluation
reward: 197.6
episode: 1913   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 797    lr: 2.684354560000003e-10     evaluation
reward: 197.3
episode: 1914   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 676    lr: 2.684354560000003e-10     evaluation
reward: 197.95
episode: 1915   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1197    lr: 2.684354560000003e-10     evaluation
reward: 199.35
episode: 1916   score: 355.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1101    lr: 2.684354560000003e-10     evaluation
reward: 200.25
episode: 1917   score: 40.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 262    lr: 2.684354560000003e-10     evaluation
reward: 197.85
episode: 1918   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1018    lr: 2.684354560000003e-10     evaluation
reward: 198.75
episode: 1919   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 875    lr: 2.684354560000003e-10     evaluation
reward: 197.4
```

```
episode: 1920   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 599    lr: 2.684354560000003e-10     evaluation
reward: 197.6
episode: 1921   score: 340.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1209    lr: 2.684354560000003e-10     evaluation
reward: 199.5
episode: 1922   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 676    lr: 2.684354560000003e-10     evaluation
reward: 198.0
episode: 1923   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 515    lr: 2.684354560000003e-10     evaluation
reward: 198.8
episode: 1924   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 804    lr: 2.684354560000003e-10     evaluation
reward: 197.75
episode: 1925   score: 175.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 618    lr: 2.684354560000003e-10     evaluation
reward: 198.4
episode: 1926   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 533    lr: 2.684354560000003e-10     evaluation
reward: 197.35
episode: 1927   score: 385.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1172    lr: 2.684354560000003e-10     evaluation
reward: 195.8
episode: 1928   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 721    lr: 2.684354560000003e-10     evaluation
reward: 196.4
episode: 1929   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 565    lr: 2.684354560000003e-10     evaluation
reward: 195.95
episode: 1930   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 472    lr: 2.684354560000003e-10     evaluation
reward: 195.85
episode: 1931   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 886    lr: 2.684354560000003e-10     evaluation
reward: 195.55
episode: 1932   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 935    lr: 2.684354560000003e-10     evaluation
reward: 197.05
episode: 1933   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 936    lr: 2.684354560000003e-10     evaluation
reward: 198.5
episode: 1934   score: 60.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 390    lr: 2.684354560000003e-10     evaluation
reward: 197.35
episode: 1935   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 775    lr: 2.684354560000003e-10     evaluation
reward: 197.2
```

```
episode: 1936   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 904    lr: 2.6843545600000003e-10     evaluation
reward: 198.6
episode: 1937   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 465    lr: 2.6843545600000003e-10     evaluation
reward: 198.2
episode: 1938   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 614    lr: 2.6843545600000003e-10     evaluation
reward: 198.9
episode: 1939   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 745    lr: 2.6843545600000003e-10     evaluation
reward: 199.2
episode: 1940   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 353    lr: 2.6843545600000003e-10     evaluation
reward: 198.0
episode: 1941   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 943    lr: 2.6843545600000003e-10     evaluation
reward: 200.2
episode: 1942   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 427    lr: 2.6843545600000003e-10     evaluation
reward: 200.0
episode: 1943   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 249    lr: 2.6843545600000003e-10     evaluation
reward: 198.8
episode: 1944   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 505    lr: 2.6843545600000003e-10     evaluation
reward: 199.1
episode: 1945   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1149    lr: 2.6843545600000003e-10     evaluation
reward: 201.2
episode: 1946   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 445    lr: 2.6843545600000003e-10     evaluation
reward: 199.9
episode: 1947   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 757    lr: 2.6843545600000003e-10     evaluation
reward: 199.9
episode: 1948   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1019    lr: 2.6843545600000003e-10     evaluation
reward: 200.55
episode: 1949   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 841    lr: 2.6843545600000003e-10     evaluation
reward: 201.75
episode: 1950   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 922    lr: 2.6843545600000003e-10     evaluation
reward: 202.55
episode: 1951   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1199    lr: 2.6843545600000003e-10     evaluation
reward: 202.85
```

```
episode: 1952   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 576    lr: 2.6843545600000003e-10     evaluation
reward: 203.05
episode: 1953   score: 120.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 452    lr: 2.6843545600000003e-10     evaluation
reward: 203.45
episode: 1954   score: 325.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1152    lr: 2.6843545600000003e-10     evaluation
reward: 205.9
episode: 1955   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 458    lr: 2.6843545600000003e-10     evaluation
reward: 206.7
episode: 1956   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 925    lr: 2.6843545600000003e-10     evaluation
reward: 207.3
episode: 1957   score: 1160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 2058    lr: 2.6843545600000003e-10     evaluation
reward: 217.15
episode: 1958   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 963    lr: 2.6843545600000003e-10     evaluation
reward: 214.25
episode: 1959   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 977    lr: 2.6843545600000003e-10     evaluation
reward: 214.1
episode: 1960   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 363    lr: 2.6843545600000003e-10     evaluation
reward: 213.9
episode: 1961   score: 640.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1422    lr: 2.6843545600000003e-10     evaluation
reward: 218.55
episode: 1962   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 873    lr: 2.6843545600000003e-10     evaluation
reward: 219.8
episode: 1963   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 926    lr: 2.6843545600000003e-10     evaluation
reward: 220.4
episode: 1964   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 684    lr: 2.6843545600000003e-10     evaluation
reward: 220.7
episode: 1965   score: 190.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 686    lr: 2.6843545600000003e-10     evaluation
reward: 219.5
episode: 1966   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 627    lr: 2.6843545600000003e-10     evaluation
reward: 220.0
episode: 1967   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1297    lr: 2.6843545600000003e-10     evaluation
reward: 220.15
```

```
episode: 1968   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 889    lr: 2.6843545600000003e-10     evaluation
reward: 219.85
episode: 1969   score: 280.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 740    lr: 2.6843545600000003e-10     evaluation
reward: 220.9
episode: 1970   score: 265.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 925    lr: 2.6843545600000003e-10     evaluation
reward: 222.05
episode: 1971   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 363    lr: 2.6843545600000003e-10     evaluation
reward: 220.1
episode: 1972   score: 420.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1163    lr: 2.6843545600000003e-10     evaluation
reward: 222.25
episode: 1973   score: 70.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 272    lr: 2.6843545600000003e-10     evaluation
reward: 219.25
episode: 1974   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 795    lr: 2.6843545600000003e-10     evaluation
reward: 217.9
episode: 1975   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 816    lr: 2.6843545600000003e-10     evaluation
reward: 207.95
episode: 1976   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 817    lr: 2.6843545600000003e-10     evaluation
reward: 209.7
episode: 1977   score: 340.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1082    lr: 2.6843545600000003e-10     evaluation
reward: 211.05
episode: 1978   score: 460.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1534    lr: 2.6843545600000003e-10     evaluation
reward: 213.9
episode: 1979   score: 295.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 844    lr: 2.6843545600000003e-10     evaluation
reward: 215.65
episode: 1980   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 372    lr: 2.6843545600000003e-10     evaluation
reward: 216.15
episode: 1981   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 518    lr: 2.6843545600000003e-10     evaluation
reward: 216.35
episode: 1982   score: 960.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1690    lr: 2.6843545600000003e-10     evaluation
reward: 224.55
episode: 1983   score: 235.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 850    lr: 2.6843545600000003e-10     evaluation
reward: 225.5
```

```
episode: 1984   score: 420.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1215    lr: 2.684354560000003e-10     evaluation
reward: 227.95
episode: 1985   score: 220.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1006    lr: 2.684354560000003e-10     evaluation
reward: 226.9
episode: 1986   score: 80.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 393    lr: 2.684354560000003e-10     evaluation
reward: 224.9
episode: 1987   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 727    lr: 2.684354560000003e-10     evaluation
reward: 225.05
episode: 1988   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 567    lr: 2.684354560000003e-10     evaluation
reward: 226.0
episode: 1989   score: 205.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 752    lr: 2.684354560000003e-10     evaluation
reward: 226.75
episode: 1990   score: 100.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 478    lr: 2.684354560000003e-10     evaluation
reward: 226.65
episode: 1991   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 713    lr: 2.684354560000003e-10     evaluation
reward: 227.35
episode: 1992   score: 250.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 798    lr: 2.684354560000003e-10     evaluation
reward: 226.75
episode: 1993   score: 310.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1181    lr: 2.684354560000003e-10     evaluation
reward: 227.95
episode: 1994   score: 140.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 461    lr: 2.684354560000003e-10     evaluation
reward: 228.45
episode: 1995   score: 130.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 594    lr: 2.684354560000003e-10     evaluation
reward: 228.15
episode: 1996   score: 160.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 562    lr: 2.684354560000003e-10     evaluation
reward: 227.25
episode: 1997   score: 520.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 1515    lr: 2.684354560000003e-10     evaluation
reward: 230.55
episode: 1998   score: 90.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 480    lr: 2.684354560000003e-10     evaluation
reward: 229.4
episode: 1999   score: 150.0   memory length: 1000000   epsilon:
0.009998020008555413    steps: 644    lr: 2.684354560000003e-10     evaluation
reward: 230.4
```

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

[ ]:
```
from gym.wrappers import RecordVideo # If importing monitor raises issues, try␣
 ↪using `from gym.wrappers import RecordVideo`
import glob
import io
import base64

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

from pyvirtualdisplay import Display

env = gym.make('ALE/DemonAttack-v5', render_mode='rgb_array')
state = env.reset()
number_lives = find_max_lives(env)
state_size = env.observation_space.shape
action_size = 3 #fire, left, and right

# Displaying the game live
def show_state(env, step=0, info=""):
    plt.figure(3)
    plt.clf()
```

```
    plt.imshow(env.render(mode='rgb_array'))
    plt.title("%s | Step: %d %s" % ("Agent Playing",step, info))
    plt.axis('off')

    ipythondisplay.clear_output(wait=True)
    ipythondisplay.display(plt.gcf())

# Recording the game and replaying the game afterwards
def show_video():
    mp4list = glob.glob('video/*.mp4')
    if len(mp4list) >= 0:
        mp4 = mp4list[0]
        video = io.open(mp4, 'r+b').read()
        encoded = base64.b64encode(video)
        ipythondisplay.display(HTML(data='''<video style="height: 400px;"␣
 ↪controls="" loop="" autoplay="" alt="test">
               <source type="video/mp4" src="data:video/mp4; base64,{0}">
            </video>'''.format(encoded.decode('ascii'))))
    else:
        print("Could not find video")


def wrap_env(env):
    env = RecordVideo(env, './video')
    return env

from agent import Agent
action_size = 3


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

env = wrap_env(env)

done = False
score = 0
step = 0
state, _ = env.reset()
next_state = state
life = number_lives
history = np.zeros([5, 84, 84], dtype=np.uint8)
get_init_state(history, state, HISTORY_SIZE)
frame = 0
while not done:
```

```
#     show_state(env,step) # uncommenting this provides another way to␣
 ↪visualize the game
    step += 1
    frame += 1

    action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
    state = next_state

    next_state, reward, done, _, info = env.step(action + 1)

    frame_next_state = get_frame(next_state)
    history[4, :, :] = frame_next_state
    terminal_state = check_live(life, info['lives'])

    life = info['lives']
    r = np.clip(reward, -1, 1)
    r = reward

    # Store the transition in memory
    agent.memory.push(deepcopy(frame_next_state), action, r, terminal_state)
    # Start training after random sample generation
    score += reward

    history[:4, :, :] = history[1:, :, :]
env.close()
show_video()
```

```
/u/qilong/.local/lib/python3.9/site-packages/gym/wrappers/record_video.py:75:
UserWarning: WARN: Overwriting existing videos at
/u/qilong/QA_generator/CS444/mp5/video folder (try specifying a different
`video_folder` for the `RecordVideo` wrapper if this is not desired)
  logger.warn(
/u/qilong/.local/lib/python3.9/site-
packages/gym/utils/passive_env_checker.py:289: UserWarning: WARN: No render
fps was declared in the environment (env.metadata['render_fps'] is None or not
defined), rendering may occur at inconsistent fps.
  logger.warn(
Moviepy - Building video /u/qilong/QA_generator/CS444/mp5/video/rl-video-
episode-0.mp4.
Moviepy - Writing video /u/qilong/QA_generator/CS444/mp5/video/rl-video-
episode-0.mp4



Moviepy - Done !
```

```
Moviepy - video ready /u/qilong/QA_generator/CS444/mp5/video/rl-video-
episode-0.mp4


<IPython.core.display.HTML object>
```

4 Save as PDF

[2]:
```
%%capture

from google.colab import drive
drive.mount('/content/drive')
# install tex; first run may take several minutes
! apt-get install texlive-xetex
# file path and save location below are default; please change if they do not␣
 ↪match yours
! jupyter nbconvert --output-dir='/content/drive/MyDrive/' '/content/drive/
 ↪MyDrive/CS444/assignment5/MP5.ipynb' --to pdf
```