



Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное  
учреждение высшего образования «Московский государственный  
технический университет имени Н.Э. Баумана (национальный ис-  
следовательский университет)» (МГТУ им. Н.Э. Баумана)

---

ФАКУЛЬТЕТ Информатика и системы управления и искусственный интеллект

КАФЕДРА Системы обработки информации и управления

**Лабораторная работа №7**  
**По курсу**  
**«Методы машинного обучения в АСОИУ»**  
**«Алгоритмы Actor-Critic»**

Выполнил:

ИУ5-22М Киричков Е. Е.

26.05.2024

Проверил:

Балашов А.М.

Москва, 2024

## Задание:

Реализуйте любой алгоритм семейства Actor-Critic для произвольной среды.

Программа реализует алгоритм обучения с подкреплением Advantage Actor-Critic (A2C) для решения задачи управления маятником в окружении "CartPole-v1" из библиотеки OpenAI Gym. Основная цель состоит в обучении агента (модели) так, чтобы он мог удерживать маятник в вертикальном положении как можно дольше.

```
Ввод [1]: import sys
import torch
import gym
import numpy as np
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F
from torch.autograd import Variable
import matplotlib.pyplot as plt
import pandas as pd
from IPython.display import Image
```

```
Ввод [2]: # Гиперпараметры
hidden_size = 256
learning_rate = 3e-4
entropy_coef = 0.01

# Константы
GAMMA = 0.99
num_steps = 250
max_episodes = 2500 # Увеличенное количество эпизодов
CONST_ENV_NAME = "CartPole-v1"
CONST_DEVICE = torch.device("cuda" if torch.cuda.is_available() else "cpu")
```

```
Ввод [3]: # Класс, реализующий модель Actor-Critic
class ActorCritic(nn.Module):
    def __init__(self, num_inputs, num_actions, hidden_size, learning_rate=3e-4):
        super(ActorCritic, self).__init__()

        self.num_actions = num_actions
        self.critic_linear1 = nn.Linear(num_inputs, hidden_size)
        self.critic_linear2 = nn.Linear(hidden_size, 1)

        self.actor_linear1 = nn.Linear(num_inputs, hidden_size)
        self.actor_linear2 = nn.Linear(hidden_size, num_actions)

    def forward(self, state):
        value = F.relu(self.critic_linear1(state))
        value = self.critic_linear2(value)

        policy_dist = F.relu(self.actor_linear1(state))
        policy_dist = F.softmax(self.actor_linear2(policy_dist), dim=1)

        return value, policy_dist

    def select_action(self, state):
        _, policy_dist = self.forward(state)
        action = policy_dist.multinomial(num_samples=1).detach()
        return action[0]
```

Ввод [4]: *# Функция для реализации алгоритма Advantage Actor-Critic (A2C)*

```
def a2c(env):
    num_inputs = env.observation_space.shape[0]
    num_outputs = env.action_space.n

    actor_critic = ActorCritic(num_inputs, num_outputs, hidden_size)
    ac_optimizer = optim.Adam(actor_critic.parameters(), lr=learning_rate)

    all_lengths = []
    average_lengths = []
    all_rewards = []
    entropy_term = 0

    for episode in range(max_episodes):
        log_probs = []
        values = []
        rewards = []

        state = env.reset()
        if isinstance(state, tuple):
            state = state[0]
        state = torch.tensor(state, dtype=torch.float32).unsqueeze(0)
        for steps in range(num_steps):
            value, policy_dist = actor_critic.forward(state)
            value = value.detach().numpy()[0, 0]
            dist = policy_dist.detach().numpy()

            action = np.random.choice(num_outputs, p=np.squeeze(dist))
            log_prob = torch.log(policy_dist.squeeze(0)[action])
            entropy = -np.sum(np.mean(dist) * np.log(dist))
            result = env.step(action)
            new_state = result[0]
            reward = result[1]
            done = result[2]

            if isinstance(new_state, tuple):
                new_state = new_state[0]
            new_state = torch.tensor(new_state, dtype=torch.float32).unsqueeze(0)

            rewards.append(reward)
            values.append(value)
            log_probs.append(log_prob)
            entropy_term += entropy
            state = new_state

            if done or steps == num_steps - 1:
                Qval, _ = actor_critic.forward(new_state)
                Qval = Qval.detach().numpy()[0, 0]
                all_rewards.append(np.sum(rewards))
                all_lengths.append(steps)
                average_lengths.append(np.mean(all_lengths[-10:]))
                if episode % 10 == 0:
                    sys.stdout.write("episode: {}, reward: {}, total length: {}, average\n".format(episode, np.sum(all_rewards), sum(all_lengths), np.mean(average_lengths)))
                    sys.stdout.flush()
                break

        # Вычисление значений Q
        Qvals = np.zeros_like(values)
        for t in reversed(range(len(rewards))):
            Qval = rewards[t] + GAMMA * Qval
            Qvals[t] = Qval

        # Обновление модели actor-critic
        values = torch.FloatTensor(values)
        Qvals = torch.FloatTensor(Qvals)
        log_probs = torch.stack(log_probs)

        advantage = Qvals - values
        actor_loss = (-log_probs * advantage).mean()
        critic_loss = 0.5 * advantage.pow(2).mean()
        ac_loss = actor_loss + critic_loss + entropy_coef * entropy_term
```

```

        ac_optimizer.zero_grad()
        ac_loss.backward()
        ac_optimizer.step()

# Построение графиков результатов
smoothed_rewards = pd.Series(all_rewards).rolling(10).mean()
plt.plot(all_rewards, label='Награды')
plt.plot(smoothed_rewards, label='Сглаженные награды')
plt.xlabel('Эпизод')
plt.ylabel('Награда')
plt.legend()
plt.show()

plt.plot(all_lengths, label='Длины эпизодов')
plt.plot(average_lengths, label='Средние длины')
plt.xlabel('Эпизод')
plt.ylabel('Длина эпизода')
plt.legend()
plt.show()

return actor_critic

```

Ввод [5]:

```

# Функция для запуска агента в среде с использованием обученной модели
def play_agent(actor_critic):
    """
    Игра с обученным агентом
    """
    env2 = gym.make(CONST_ENV_NAME, render_mode='human')
    state = env2.reset()
    if isinstance(state, tuple):
        state = state[0]
    state = torch.tensor(state, dtype=torch.float32, device=CONST_DEVICE).unsqueeze(0)
    done = False
    res = []
    while not done:
        action = actor_critic.select_action(state)
        action = action.item()
        result = env2.step(action)
        observation = result[0]
        reward = result[1]
        done = result[2]
        truncated = result[3]
        env2.render()

        res.append((action, reward))

    if done or truncated:
        done = True
    else:
        if isinstance(observation, tuple):
            observation = observation[0]
        state = torch.tensor(observation, dtype=torch.float32, device=CONST_DEVICE)

    print('Данные эпизода: ', res)

```

```
Ввод [6]: if __name__ == "__main__":  
    env = gym.make(CONST_ENV_NAME)  
    trained_actor_critic = a2c(env)  
  
    # Запуск сессии с обученным агентом  
    play_agent(trained_actor_critic)
```

```
/Users/evseykirichkov/anaconda3/lib/python3.11/site-packages/gym/utils/passive_env_checker.py:233: DeprecationWarning: `np.bool8` is a deprecated alias for `np.bool_`.  
(Deprecated NumPy 1.24)  
    if not isinstance(terminated, (bool, np.bool8)):
```

episode: 0, reward: 14.0, total length: 13, average length: 13.0  
episode: 10, reward: 30.0, total length: 29, average length: 16.9  
episode: 20, reward: 32.0, total length: 31, average length: 21.2  
episode: 30, reward: 43.0, total length: 42, average length: 24.6  
episode: 40, reward: 10.0, total length: 9, average length: 13.7  
episode: 50, reward: 17.0, total length: 16, average length: 17.8  
episode: 60, reward: 32.0, total length: 31, average length: 18.7  
episode: 70, reward: 13.0, total length: 12, average length: 21.7  
episode: 80, reward: 18.0, total length: 17, average length: 22.6  
episode: 90, reward: 12.0, total length: 11, average length: 36.0  
episode: 100, reward: 29.0, total length: 28, average length: 22.1  
episode: 110, reward: 17.0, total length: 16, average length: 31.5  
episode: 120, reward: 26.0, total length: 25, average length: 23.6  
episode: 130, reward: 38.0, total length: 37, average length: 19.7  
episode: 140, reward: 16.0, total length: 15, average length: 20.3  
episode: 150, reward: 51.0, total length: 50, average length: 32.0  
episode: 160, reward: 25.0, total length: 24, average length: 31.9  
episode: 170, reward: 57.0, total length: 56, average length: 31.7  
episode: 180, reward: 74.0, total length: 73, average length: 30.0  
episode: 190, reward: 35.0, total length: 34, average length: 30.4  
episode: 200, reward: 13.0, total length: 12, average length: 27.6  
episode: 210, reward: 29.0, total length: 28, average length: 26.6  
episode: 220, reward: 14.0, total length: 13, average length: 25.7  
episode: 230, reward: 18.0, total length: 17, average length: 31.6  
episode: 240, reward: 31.0, total length: 30, average length: 33.1  
episode: 250, reward: 26.0, total length: 25, average length: 37.9  
episode: 260, reward: 17.0, total length: 16, average length: 22.6  
episode: 270, reward: 38.0, total length: 37, average length: 36.9  
episode: 280, reward: 13.0, total length: 12, average length: 37.8  
episode: 290, reward: 13.0, total length: 12, average length: 29.1  
episode: 300, reward: 15.0, total length: 14, average length: 33.7  
episode: 310, reward: 83.0, total length: 82, average length: 47.7  
episode: 320, reward: 45.0, total length: 44, average length: 34.1  
episode: 330, reward: 37.0, total length: 36, average length: 38.4  
episode: 340, reward: 11.0, total length: 10, average length: 31.3  
episode: 350, reward: 22.0, total length: 21, average length: 33.9  
episode: 360, reward: 27.0, total length: 26, average length: 24.6  
episode: 370, reward: 23.0, total length: 22, average length: 26.0  
episode: 380, reward: 23.0, total length: 22, average length: 37.7  
episode: 390, reward: 44.0, total length: 43, average length: 26.7  
episode: 400, reward: 49.0, total length: 48, average length: 28.0  
episode: 410, reward: 38.0, total length: 37, average length: 34.7  
episode: 420, reward: 38.0, total length: 37, average length: 27.3  
episode: 430, reward: 25.0, total length: 24, average length: 31.7  
episode: 440, reward: 71.0, total length: 70, average length: 48.9  
episode: 450, reward: 20.0, total length: 19, average length: 31.7  
episode: 460, reward: 27.0, total length: 26, average length: 35.9  
episode: 470, reward: 18.0, total length: 17, average length: 43.4  
episode: 480, reward: 29.0, total length: 28, average length: 41.4  
episode: 490, reward: 42.0, total length: 41, average length: 62.5  
episode: 500, reward: 42.0, total length: 41, average length: 45.0  
episode: 510, reward: 43.0, total length: 42, average length: 47.1  
episode: 520, reward: 63.0, total length: 62, average length: 43.8  
episode: 530, reward: 35.0, total length: 34, average length: 47.9  
episode: 540, reward: 84.0, total length: 83, average length: 45.5  
episode: 550, reward: 82.0, total length: 81, average length: 69.1  
episode: 560, reward: 131.0, total length: 130, average length: 74.8  
episode: 570, reward: 24.0, total length: 23, average length: 57.5  
episode: 580, reward: 47.0, total length: 46, average length: 67.5  
episode: 590, reward: 99.0, total length: 98, average length: 53.8  
episode: 600, reward: 64.0, total length: 63, average length: 72.6  
episode: 610, reward: 47.0, total length: 46, average length: 68.0  
episode: 620, reward: 36.0, total length: 35, average length: 68.0  
episode: 630, reward: 23.0, total length: 22, average length: 49.0  
episode: 640, reward: 98.0, total length: 97, average length: 81.4  
episode: 650, reward: 132.0, total length: 131, average length: 87.6  
episode: 660, reward: 22.0, total length: 21, average length: 86.8  
episode: 670, reward: 91.0, total length: 90, average length: 99.3  
episode: 680, reward: 117.0, total length: 116, average length: 86.4  
episode: 690, reward: 172.0, total length: 171, average length: 77.6

episode: 700, reward: 55.0, total length: 54, average length: 88.4  
episode: 710, reward: 138.0, total length: 137, average length: 81.3  
episode: 720, reward: 183.0, total length: 182, average length: 99.1  
episode: 730, reward: 132.0, total length: 131, average length: 138.1  
episode: 740, reward: 152.0, total length: 151, average length: 100.3  
episode: 750, reward: 57.0, total length: 56, average length: 81.8  
episode: 760, reward: 163.0, total length: 162, average length: 137.0  
episode: 770, reward: 207.0, total length: 206, average length: 137.4  
episode: 780, reward: 139.0, total length: 138, average length: 90.6  
episode: 790, reward: 71.0, total length: 70, average length: 134.7  
episode: 800, reward: 119.0, total length: 118, average length: 130.2  
episode: 810, reward: 175.0, total length: 174, average length: 145.7  
episode: 820, reward: 160.0, total length: 159, average length: 110.1  
episode: 830, reward: 28.0, total length: 27, average length: 100.2  
episode: 840, reward: 61.0, total length: 60, average length: 99.1  
episode: 850, reward: 109.0, total length: 108, average length: 146.3  
episode: 860, reward: 37.0, total length: 36, average length: 163.5  
episode: 870, reward: 232.0, total length: 231, average length: 155.2  
episode: 880, reward: 201.0, total length: 200, average length: 144.5  
episode: 890, reward: 162.0, total length: 161, average length: 176.8  
episode: 900, reward: 84.0, total length: 83, average length: 149.9  
episode: 910, reward: 78.0, total length: 77, average length: 116.9  
episode: 920, reward: 98.0, total length: 97, average length: 111.8  
episode: 930, reward: 166.0, total length: 165, average length: 175.9  
episode: 940, reward: 250.0, total length: 249, average length: 190.8  
episode: 950, reward: 208.0, total length: 207, average length: 144.6  
episode: 960, reward: 116.0, total length: 115, average length: 146.1  
episode: 970, reward: 229.0, total length: 228, average length: 182.7  
episode: 980, reward: 128.0, total length: 127, average length: 139.1  
episode: 990, reward: 195.0, total length: 194, average length: 168.2  
episode: 1000, reward: 164.0, total length: 163, average length: 149.8  
episode: 1010, reward: 143.0, total length: 142, average length: 138.5  
episode: 1020, reward: 183.0, total length: 182, average length: 166.3  
episode: 1030, reward: 250.0, total length: 249, average length: 186.4  
episode: 1040, reward: 250.0, total length: 249, average length: 165.1  
episode: 1050, reward: 189.0, total length: 188, average length: 197.3  
episode: 1060, reward: 176.0, total length: 175, average length: 183.6  
episode: 1070, reward: 180.0, total length: 179, average length: 163.9  
episode: 1080, reward: 248.0, total length: 247, average length: 201.1  
episode: 1090, reward: 118.0, total length: 117, average length: 185.4  
episode: 1100, reward: 194.0, total length: 193, average length: 187.5  
episode: 1110, reward: 180.0, total length: 179, average length: 202.3  
episode: 1120, reward: 203.0, total length: 202, average length: 217.0  
episode: 1130, reward: 147.0, total length: 146, average length: 170.4  
episode: 1140, reward: 250.0, total length: 249, average length: 191.6  
episode: 1150, reward: 250.0, total length: 249, average length: 207.7  
episode: 1160, reward: 162.0, total length: 161, average length: 146.6  
episode: 1170, reward: 132.0, total length: 131, average length: 156.6  
episode: 1180, reward: 202.0, total length: 201, average length: 198.7

episode: 1190, reward: 250.0, total length: 249, average length: 197.0  
episode: 1200, reward: 91.0, total length: 90, average length: 170.8  
episode: 1210, reward: 250.0, total length: 249, average length: 181.2  
episode: 1220, reward: 167.0, total length: 166, average length: 173.5  
episode: 1230, reward: 145.0, total length: 144, average length: 155.0  
episode: 1240, reward: 178.0, total length: 177, average length: 174.4  
episode: 1250, reward: 155.0, total length: 154, average length: 197.5  
episode: 1260, reward: 250.0, total length: 249, average length: 202.1  
episode: 1270, reward: 250.0, total length: 249, average length: 179.4  
episode: 1280, reward: 244.0, total length: 243, average length: 193.5  
episode: 1290, reward: 151.0, total length: 150, average length: 123.6  
episode: 1300, reward: 146.0, total length: 145, average length: 169.4  
episode: 1310, reward: 250.0, total length: 249, average length: 186.6  
episode: 1320, reward: 62.0, total length: 61, average length: 164.4  
episode: 1330, reward: 250.0, total length: 249, average length: 155.9  
episode: 1340, reward: 188.0, total length: 187, average length: 129.2  
episode: 1350, reward: 106.0, total length: 105, average length: 184.4  
episode: 1360, reward: 37.0, total length: 36, average length: 167.3  
episode: 1370, reward: 248.0, total length: 247, average length: 232.6  
episode: 1380, reward: 202.0, total length: 201, average length: 200.5  
episode: 1390, reward: 204.0, total length: 203, average length: 212.7  
episode: 1400, reward: 150.0, total length: 149, average length: 184.1  
episode: 1410, reward: 250.0, total length: 249, average length: 160.8  
episode: 1420, reward: 250.0, total length: 249, average length: 200.6  
episode: 1430, reward: 250.0, total length: 249, average length: 165.7  
episode: 1440, reward: 250.0, total length: 249, average length: 195.8  
episode: 1450, reward: 206.0, total length: 205, average length: 183.9  
episode: 1460, reward: 250.0, total length: 249, average length: 202.9  
episode: 1470, reward: 235.0, total length: 234, average length: 202.4  
episode: 1480, reward: 250.0, total length: 249, average length: 190.1  
episode: 1490, reward: 250.0, total length: 249, average length: 197.6  
episode: 1500, reward: 105.0, total length: 104, average length: 137.4  
episode: 1510, reward: 207.0, total length: 206, average length: 194.8  
episode: 1520, reward: 247.0, total length: 246, average length: 141.8  
episode: 1530, reward: 220.0, total length: 219, average length: 182.0  
episode: 1540, reward: 250.0, total length: 249, average length: 193.3  
episode: 1550, reward: 250.0, total length: 249, average length: 173.7  
episode: 1560, reward: 250.0, total length: 249, average length: 168.9  
episode: 1570, reward: 119.0, total length: 118, average length: 186.1  
episode: 1580, reward: 250.0, total length: 249, average length: 183.5  
episode: 1590, reward: 250.0, total length: 249, average length: 172.7  
episode: 1600, reward: 170.0, total length: 169, average length: 208.0  
episode: 1610, reward: 210.0, total length: 209, average length: 208.3  
episode: 1620, reward: 250.0, total length: 249, average length: 185.9  
episode: 1630, reward: 79.0, total length: 78, average length: 139.2  
episode: 1640, reward: 58.0, total length: 57, average length: 148.8  
episode: 1650, reward: 133.0, total length: 132, average length: 161.5  
episode: 1660, reward: 221.0, total length: 220, average length: 154.8  
episode: 1670, reward: 139.0, total length: 138, average length: 171.8  
episode: 1680, reward: 184.0, total length: 183, average length: 146.8  
episode: 1690, reward: 250.0, total length: 249, average length: 194.4  
episode: 1700, reward: 227.0, total length: 226, average length: 201.2  
episode: 1710, reward: 138.0, total length: 137, average length: 221.3  
episode: 1720, reward: 250.0, total length: 249, average length: 175.5  
episode: 1730, reward: 152.0, total length: 151, average length: 207.9  
episode: 1740, reward: 98.0, total length: 97, average length: 165.1  
episode: 1750, reward: 250.0, total length: 249, average length: 204.9  
episode: 1760, reward: 250.0, total length: 249, average length: 166.1  
episode: 1770, reward: 123.0, total length: 122, average length: 224.7  
episode: 1780, reward: 151.0, total length: 150, average length: 197.8  
episode: 1790, reward: 248.0, total length: 247, average length: 209.2  
episode: 1800, reward: 234.0, total length: 233, average length: 221.0  
episode: 1810, reward: 250.0, total length: 249, average length: 199.7  
episode: 1820, reward: 250.0, total length: 249, average length: 219.3  
episode: 1830, reward: 250.0, total length: 249, average length: 209.8  
episode: 1840, reward: 204.0, total length: 203, average length: 213.8  
episode: 1850, reward: 250.0, total length: 249, average length: 244.3  
episode: 1860, reward: 145.0, total length: 144, average length: 226.6  
episode: 1870, reward: 250.0, total length: 249, average length: 207.2  
episode: 1880, reward: 250.0, total length: 249, average length: 216.0



episode: 1890, reward: 250.0, total length: 249, average length: 227.1  
episode: 1900, reward: 95.0, total length: 94, average length: 217.4  
episode: 1910, reward: 250.0, total length: 249, average length: 231.3  
episode: 1920, reward: 199.0, total length: 198, average length: 172.1  
episode: 1930, reward: 180.0, total length: 179, average length: 214.8  
episode: 1940, reward: 250.0, total length: 249, average length: 170.7  
episode: 1950, reward: 62.0, total length: 61, average length: 159.2  
episode: 1960, reward: 250.0, total length: 249, average length: 191.1  
episode: 1970, reward: 219.0, total length: 218, average length: 215.3  
episode: 1980, reward: 214.0, total length: 213, average length: 174.1  
episode: 1990, reward: 184.0, total length: 183, average length: 156.5  
episode: 2000, reward: 52.0, total length: 51, average length: 132.9  
episode: 2010, reward: 117.0, total length: 116, average length: 107.2  
episode: 2020, reward: 119.0, total length: 118, average length: 134.1  
episode: 2030, reward: 126.0, total length: 125, average length: 110.9  
episode: 2040, reward: 127.0, total length: 126, average length: 133.3  
episode: 2050, reward: 163.0, total length: 162, average length: 137.2  
episode: 2060, reward: 205.0, total length: 204, average length: 150.0  
episode: 2070, reward: 195.0, total length: 194, average length: 159.3  
episode: 2080, reward: 35.0, total length: 34, average length: 182.8  
episode: 2090, reward: 250.0, total length: 249, average length: 175.2  
episode: 2100, reward: 53.0, total length: 52, average length: 195.8  
episode: 2110, reward: 250.0, total length: 249, average length: 208.5  
episode: 2120, reward: 250.0, total length: 249, average length: 223.8  
episode: 2130, reward: 158.0, total length: 157, average length: 194.6  
episode: 2140, reward: 108.0, total length: 107, average length: 191.5  
episode: 2150, reward: 206.0, total length: 205, average length: 184.2  
episode: 2160, reward: 183.0, total length: 182, average length: 200.2  
episode: 2170, reward: 223.0, total length: 222, average length: 203.4  
episode: 2180, reward: 218.0, total length: 217, average length: 211.0  
episode: 2190, reward: 223.0, total length: 222, average length: 210.4  
episode: 2200, reward: 67.0, total length: 66, average length: 179.0  
episode: 2210, reward: 173.0, total length: 172, average length: 198.7  
episode: 2220, reward: 250.0, total length: 249, average length: 198.5  
episode: 2230, reward: 128.0, total length: 127, average length: 199.5  
episode: 2240, reward: 211.0, total length: 210, average length: 185.5  
episode: 2250, reward: 232.0, total length: 231, average length: 209.8  
episode: 2260, reward: 250.0, total length: 249, average length: 220.4  
episode: 2270, reward: 250.0, total length: 249, average length: 206.4  
episode: 2280, reward: 250.0, total length: 249, average length: 218.9  
episode: 2290, reward: 250.0, total length: 249, average length: 234.0  
episode: 2300, reward: 250.0, total length: 249, average length: 205.7  
episode: 2310, reward: 250.0, total length: 249, average length: 200.5  
episode: 2320, reward: 186.0, total length: 185, average length: 214.5  
episode: 2330, reward: 250.0, total length: 249, average length: 226.1  
episode: 2340, reward: 250.0, total length: 249, average length: 195.3  
  
episode: 2350, reward: 250.0, total length: 249, average length: 216.2  
episode: 2360, reward: 207.0, total length: 206, average length: 209.7  
episode: 2370, reward: 248.0, total length: 247, average length: 185.7  
episode: 2380, reward: 242.0, total length: 241, average length: 172.8  
episode: 2390, reward: 202.0, total length: 201, average length: 182.0  
episode: 2400, reward: 250.0, total length: 249, average length: 171.5  
episode: 2410, reward: 250.0, total length: 249, average length: 188.8  
episode: 2420, reward: 250.0, total length: 249, average length: 208.0  
episode: 2430, reward: 249.0, total length: 248, average length: 210.1  
episode: 2440, reward: 104.0, total length: 103, average length: 192.2  
episode: 2450, reward: 250.0, total length: 249, average length: 217.1  
episode: 2460, reward: 250.0, total length: 249, average length: 203.2  
episode: 2470, reward: 250.0, total length: 249, average length: 224.6  
episode: 2480, reward: 178.0, total length: 177, average length: 208.3  
episode: 2490, reward: 250.0, total length: 249, average length: 212.2



