|  |  |
| --- | --- |
| C:\Users\USER\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\616878FB.tmp | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

|  |  |
| --- | --- |
| ФАКУЛЬТЕТ | Информатика, искусственный интеллект и системы управления |
| КАФЕДРА | Системы обработки информации и управления |

**Отчет по лабораторной работе №5**

**«Обучение на основе временных различий»**

Студент группы ИУ5-25М

Денисов Алексей Сергеевич

Москва, 2023 г.

**Цель работы**

Ознакомление с базовыми методами обучения с подкреплением на основе временных различий.

**Задание**

На основе рассмотренного на лекции примера реализуйте следующие алгоритмы:

* SARSA
* Q-обучение
* Двойное Q-обучение

для любой среды обучения с подкреплением из библиотеки Gym.

**Описание выполнения**

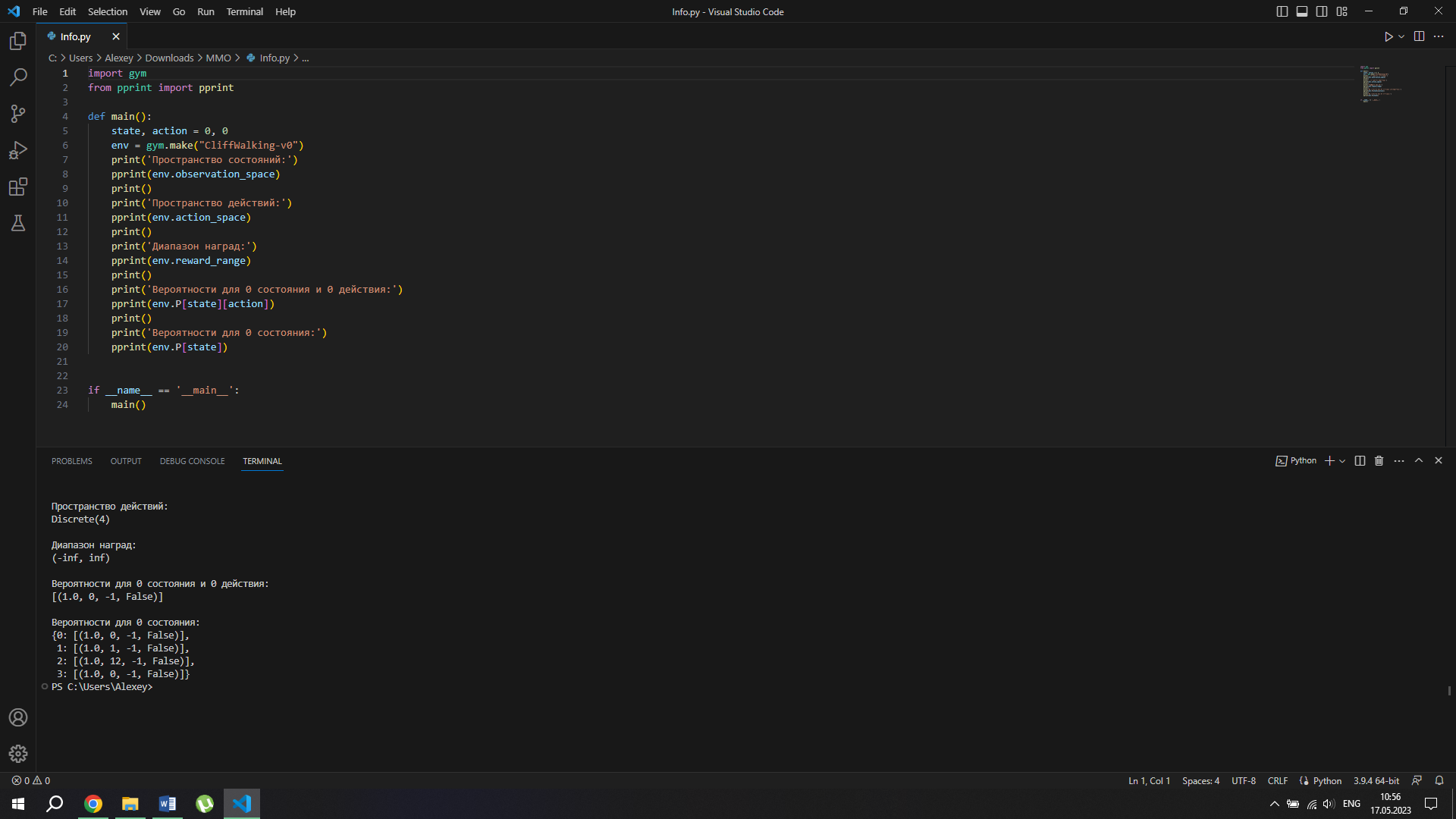


Рисунок 1 – Информация о среде

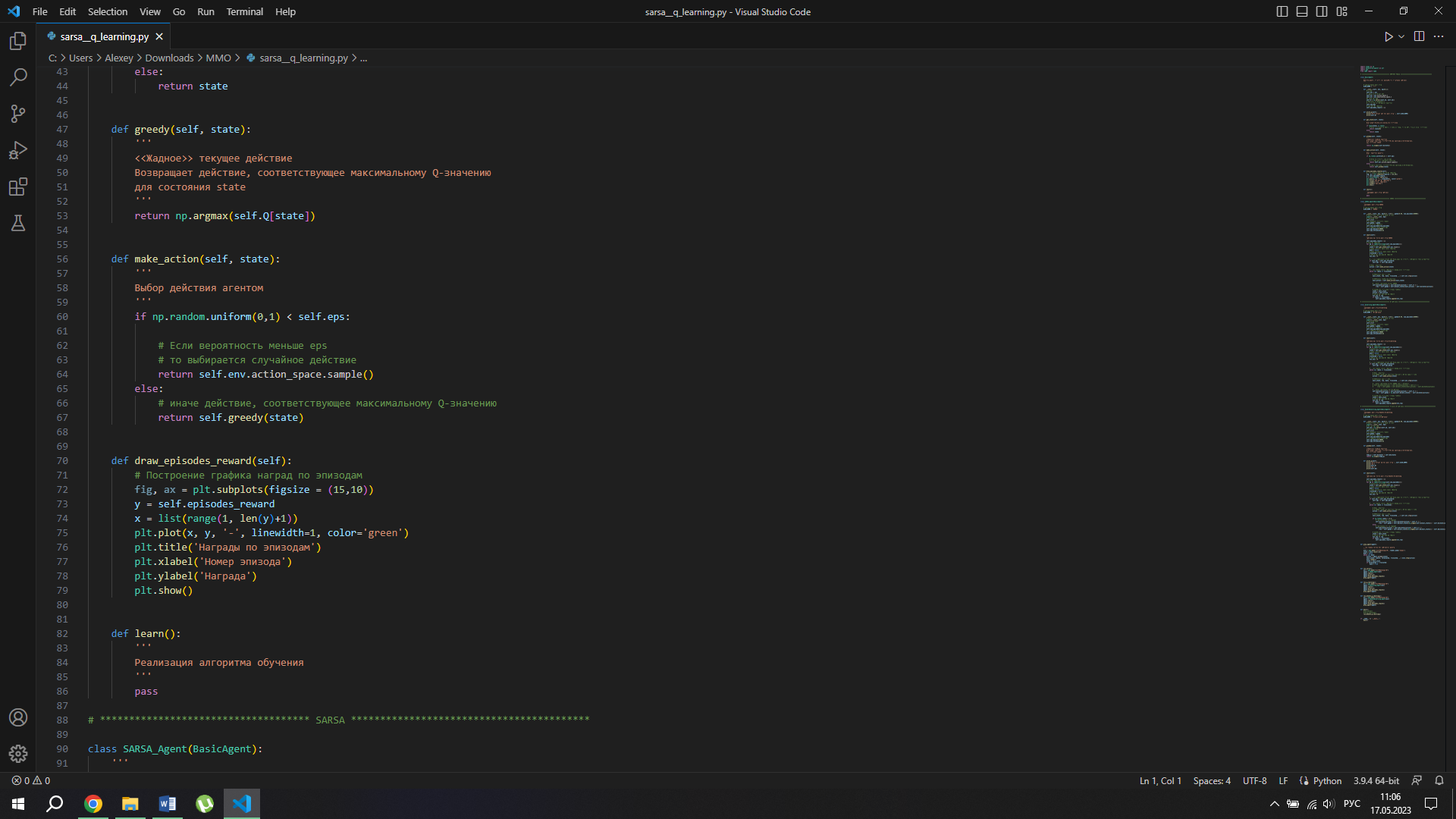
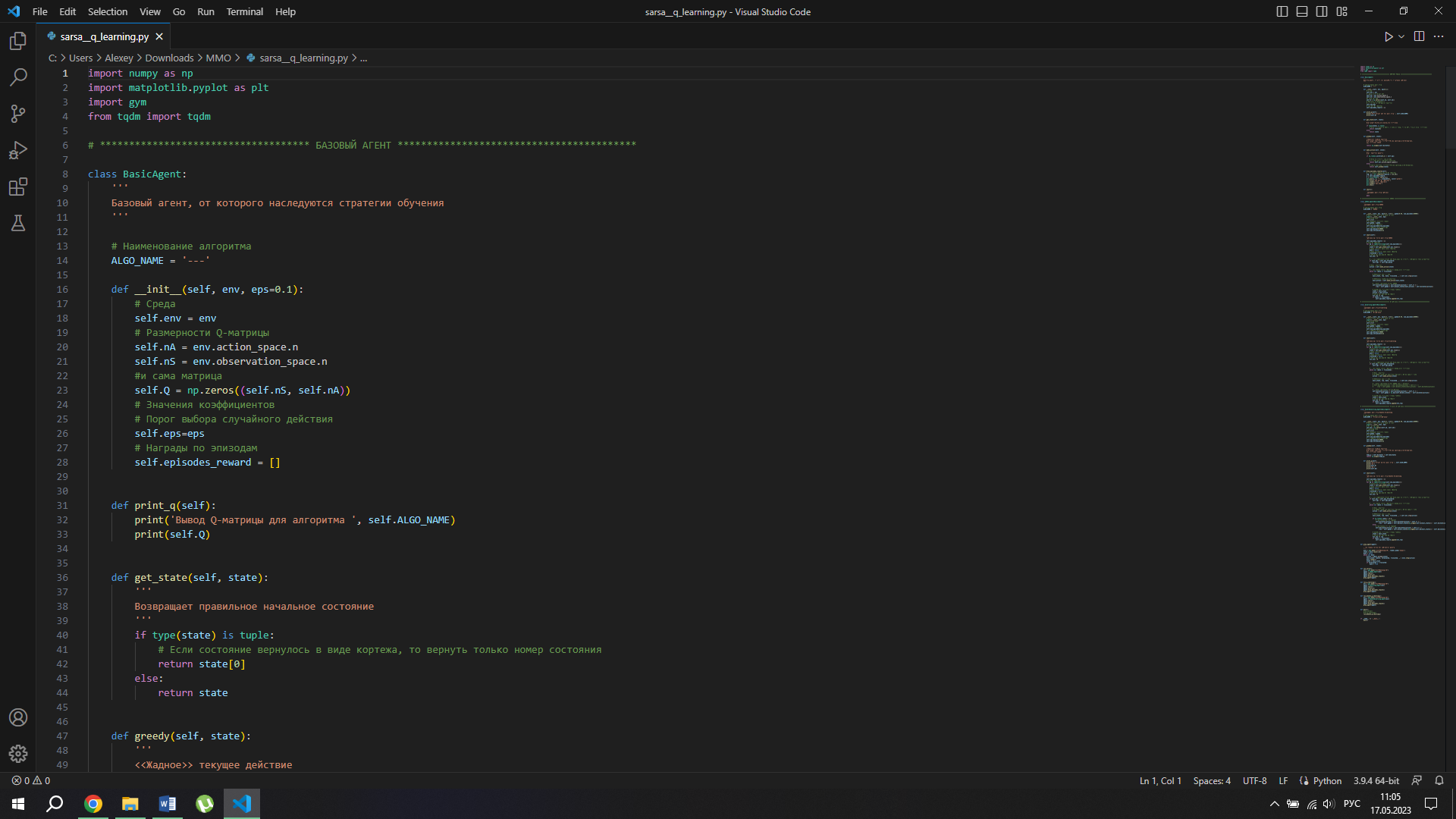


Рисунок 2-3 – Установка базовых параметров

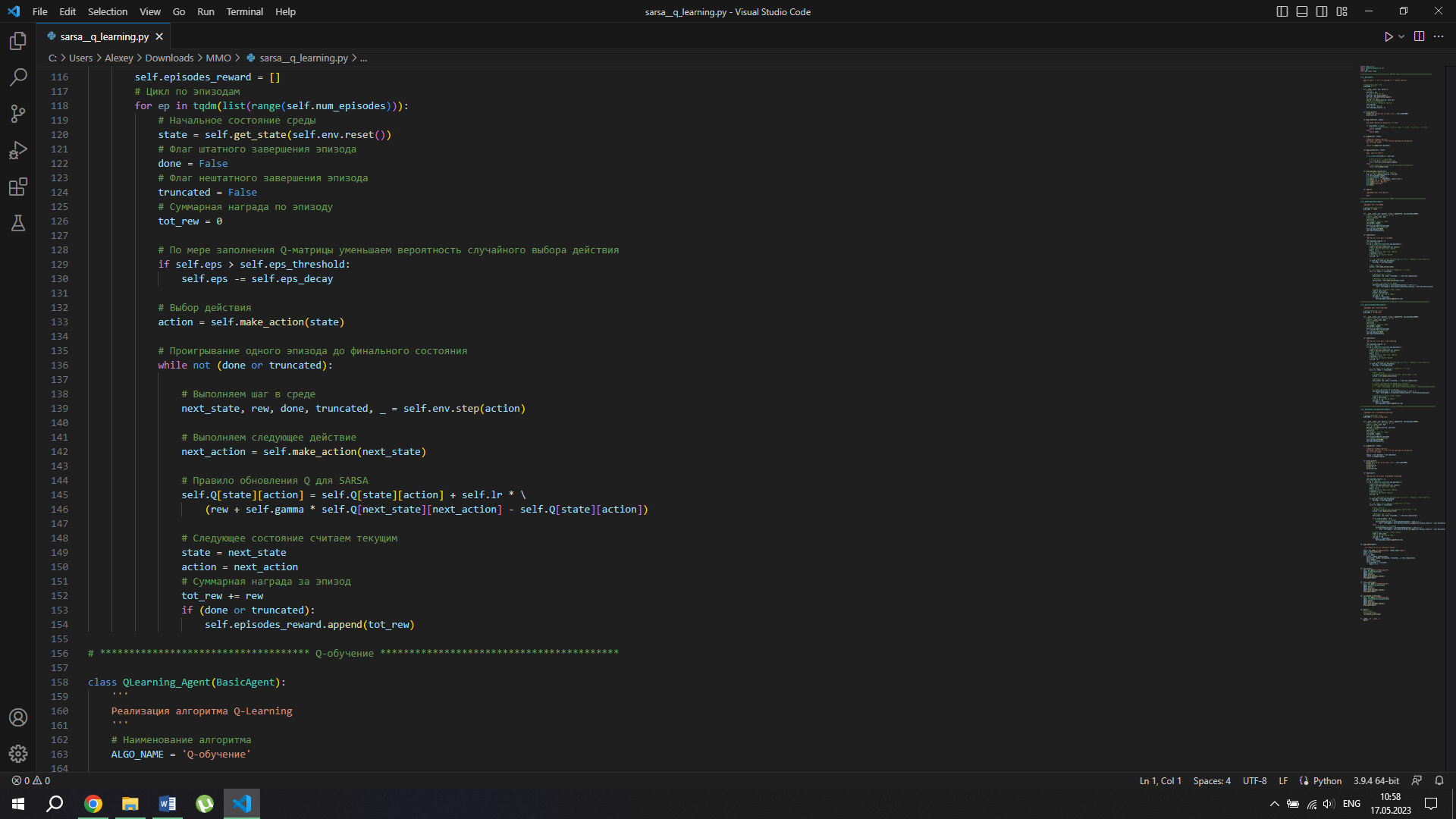
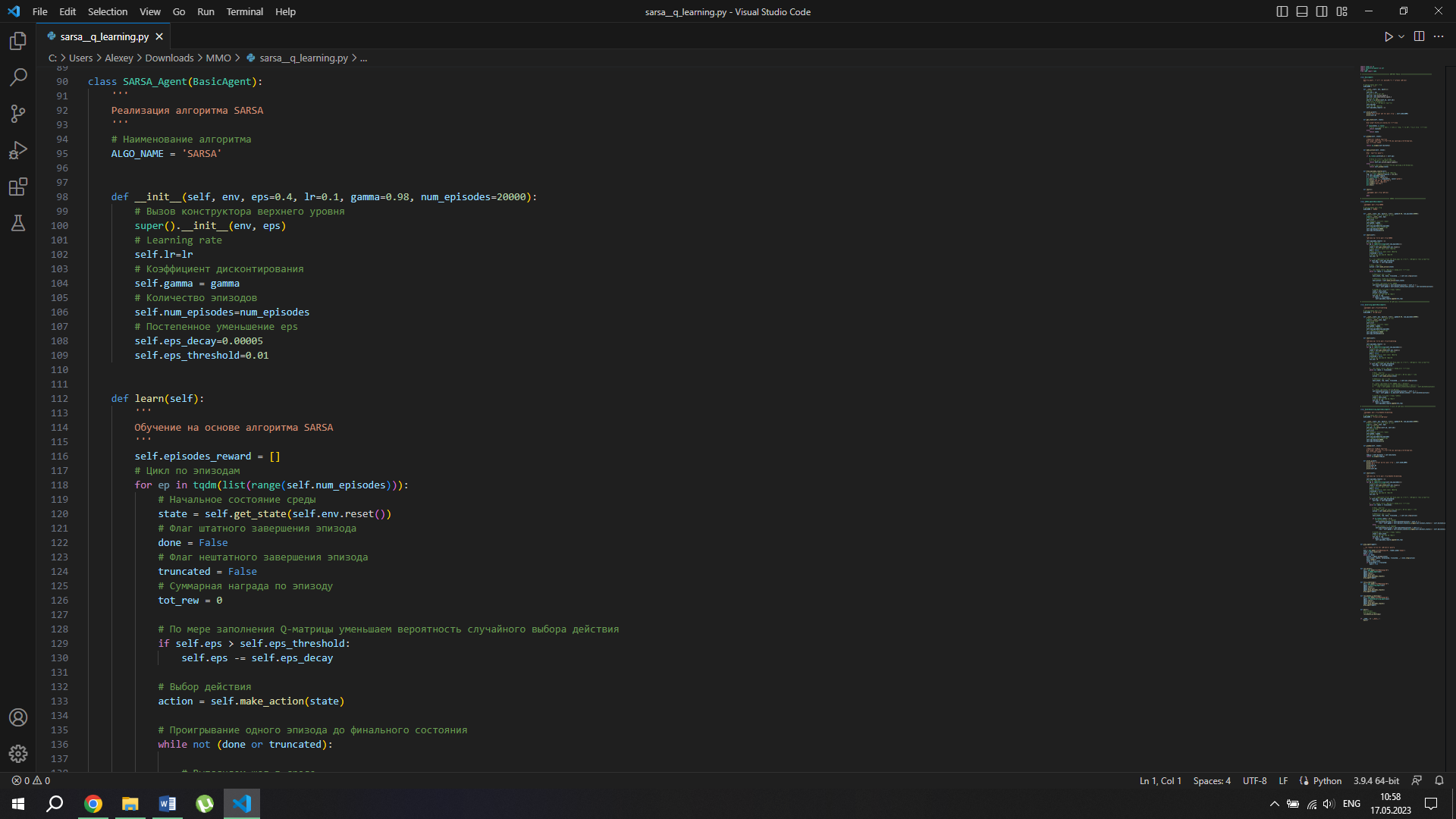


Рисунок 3-4 – Реализация алгоритма SARSA

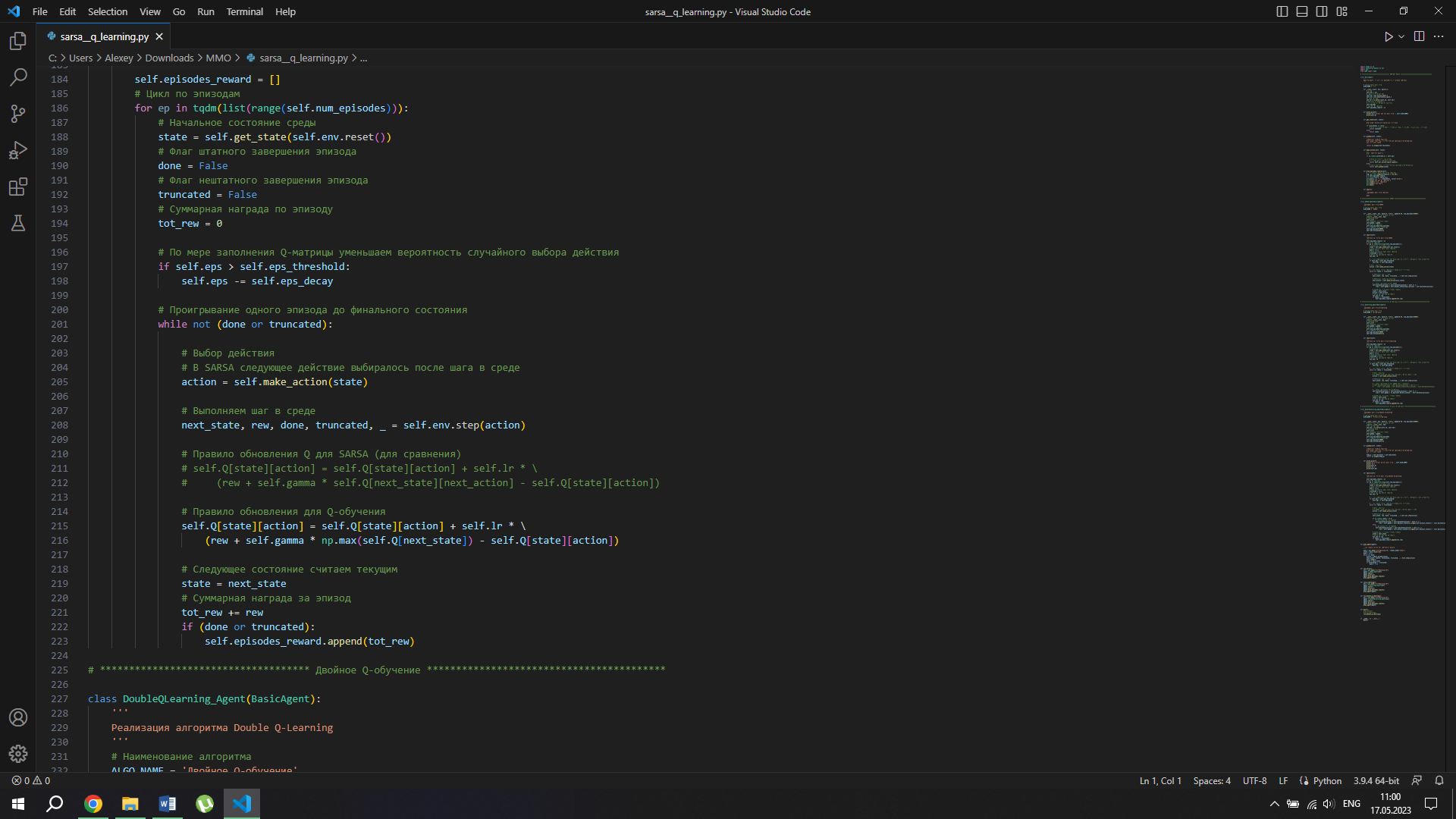
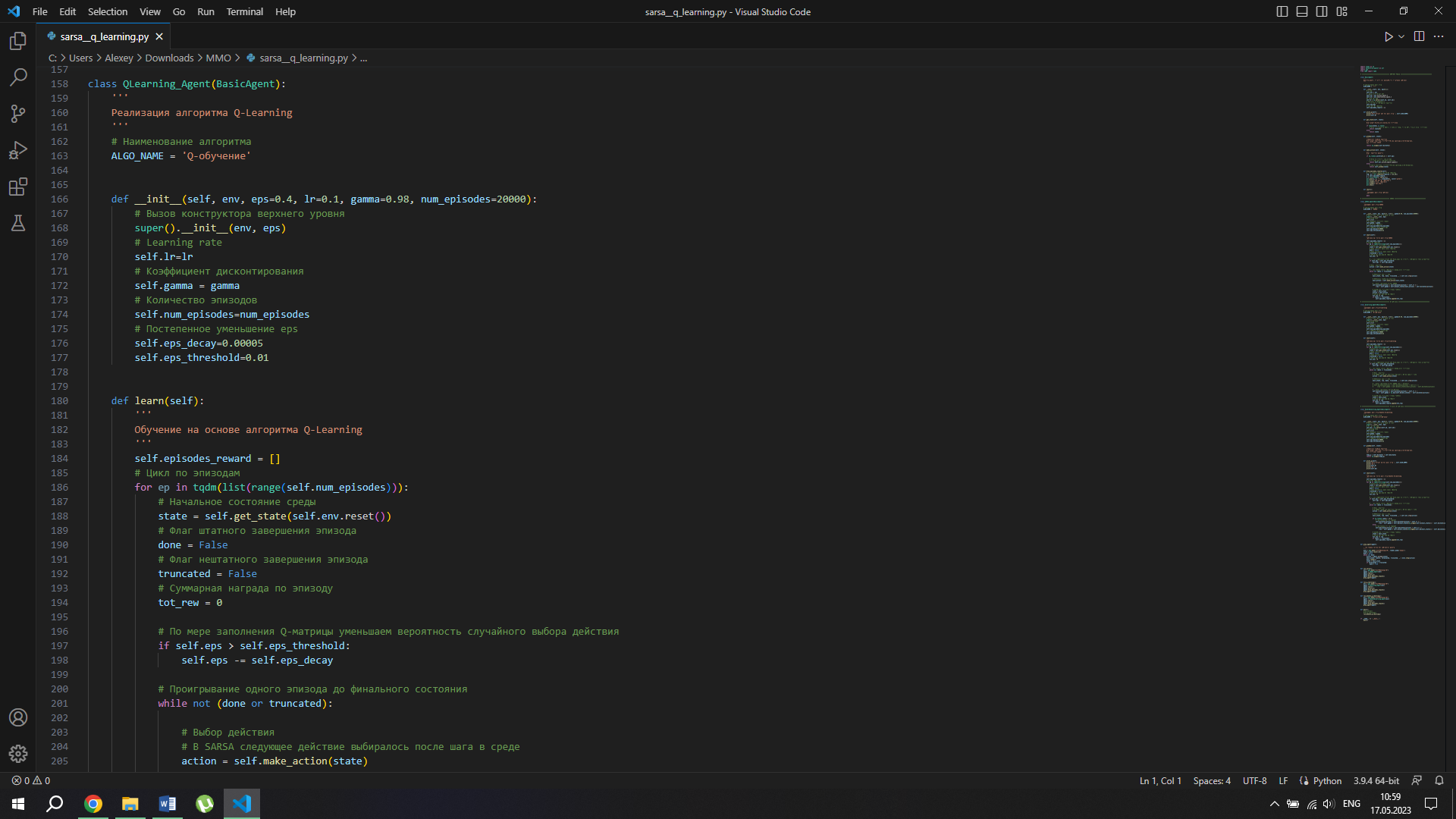


Рисунок 5-6 – Реализация Q-обучения

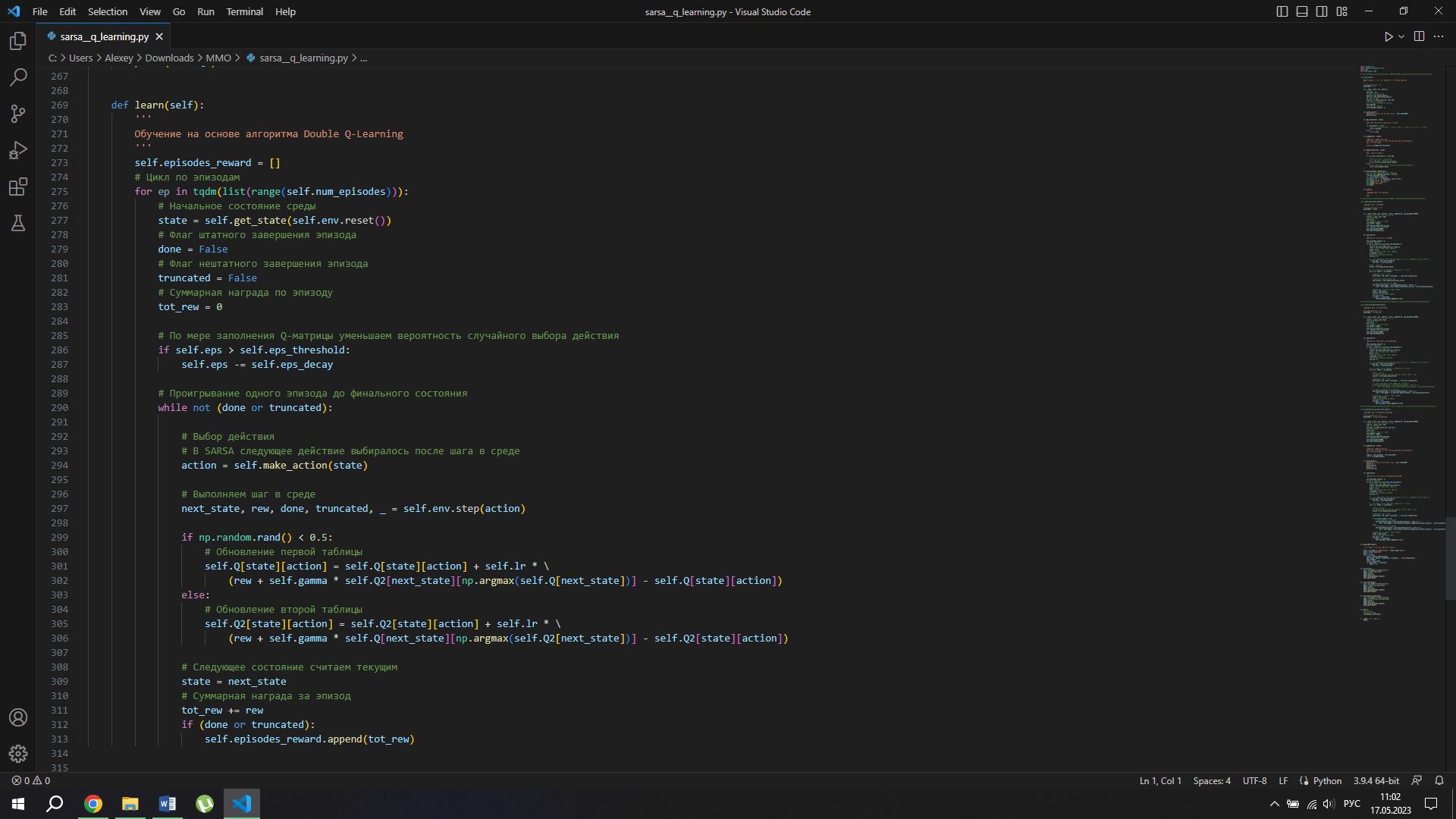
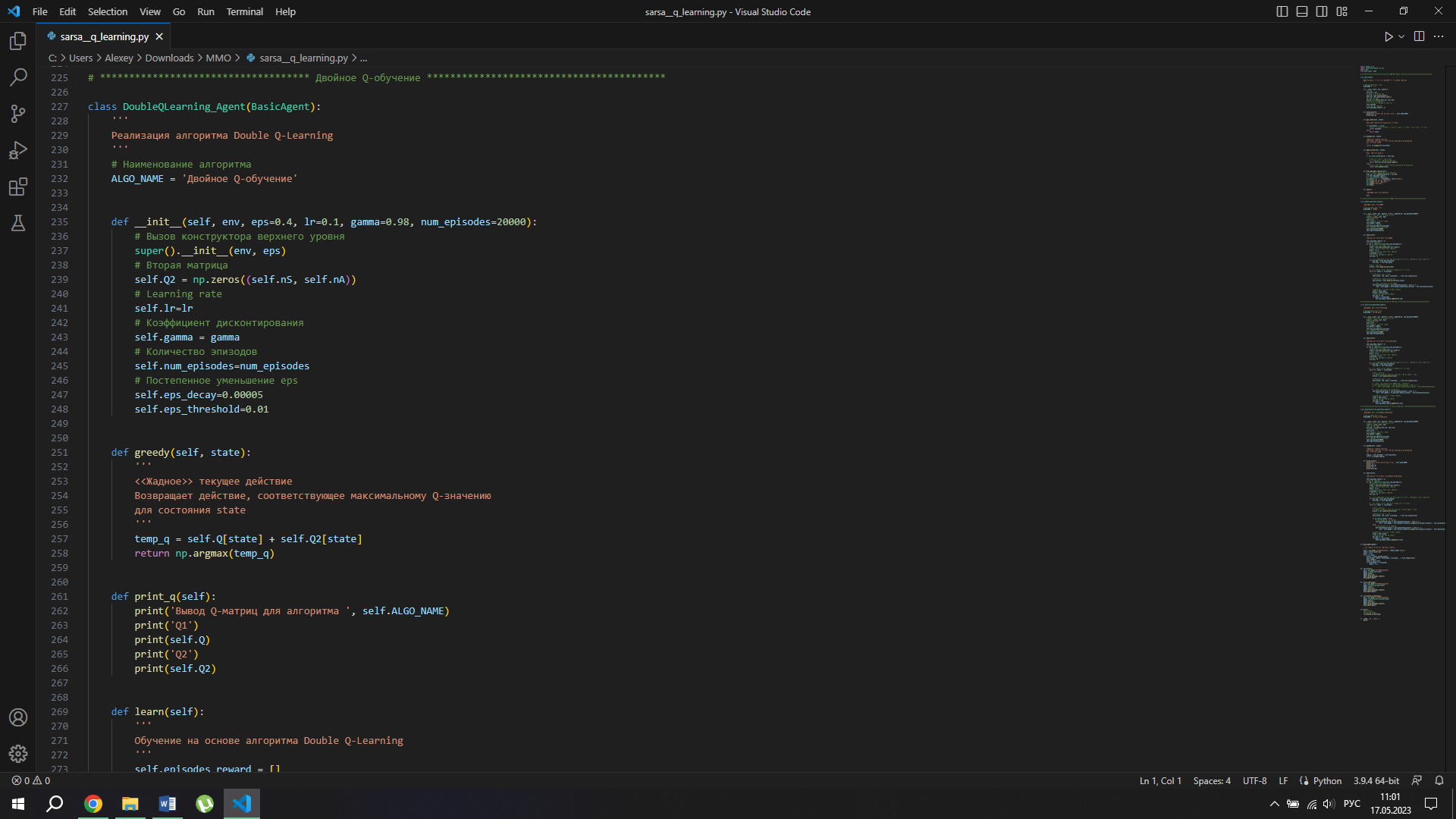


Рисунок 7-8 – Реализация двойного Q-обучения

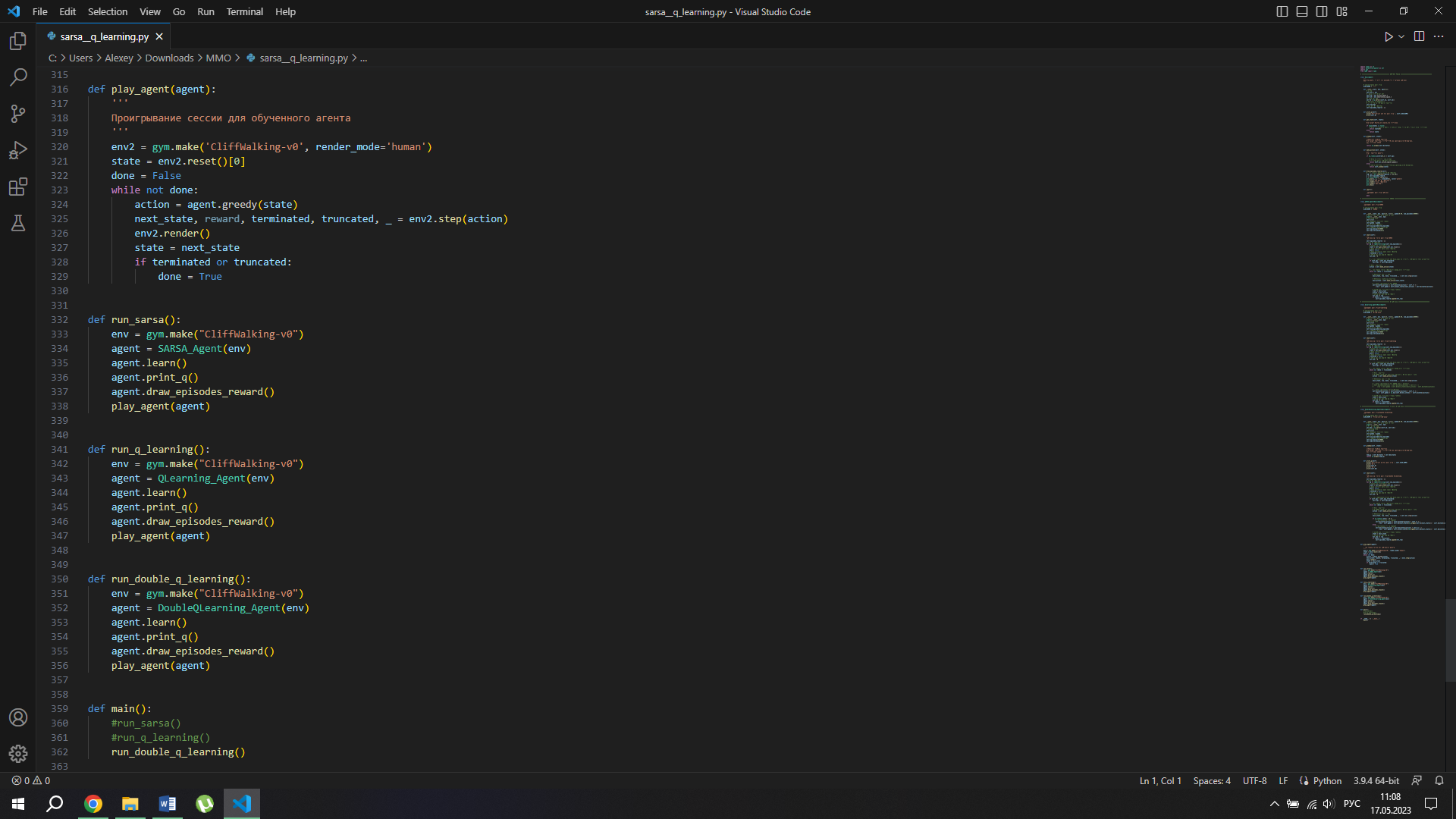


Рисунок 9 – Инициализация алгоритмов

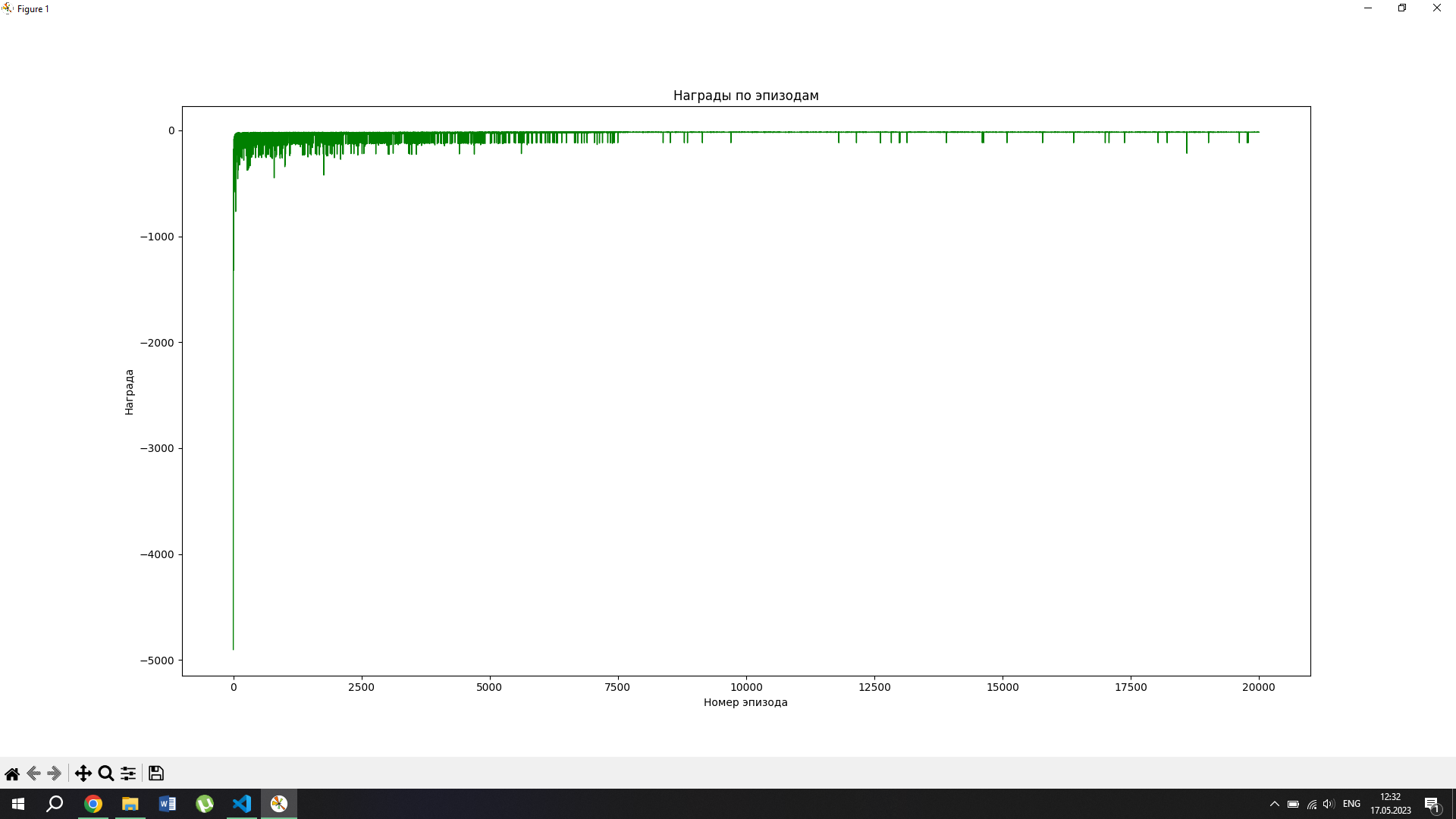


Рисунок 10 – Награды по эпизодам (SARSA)

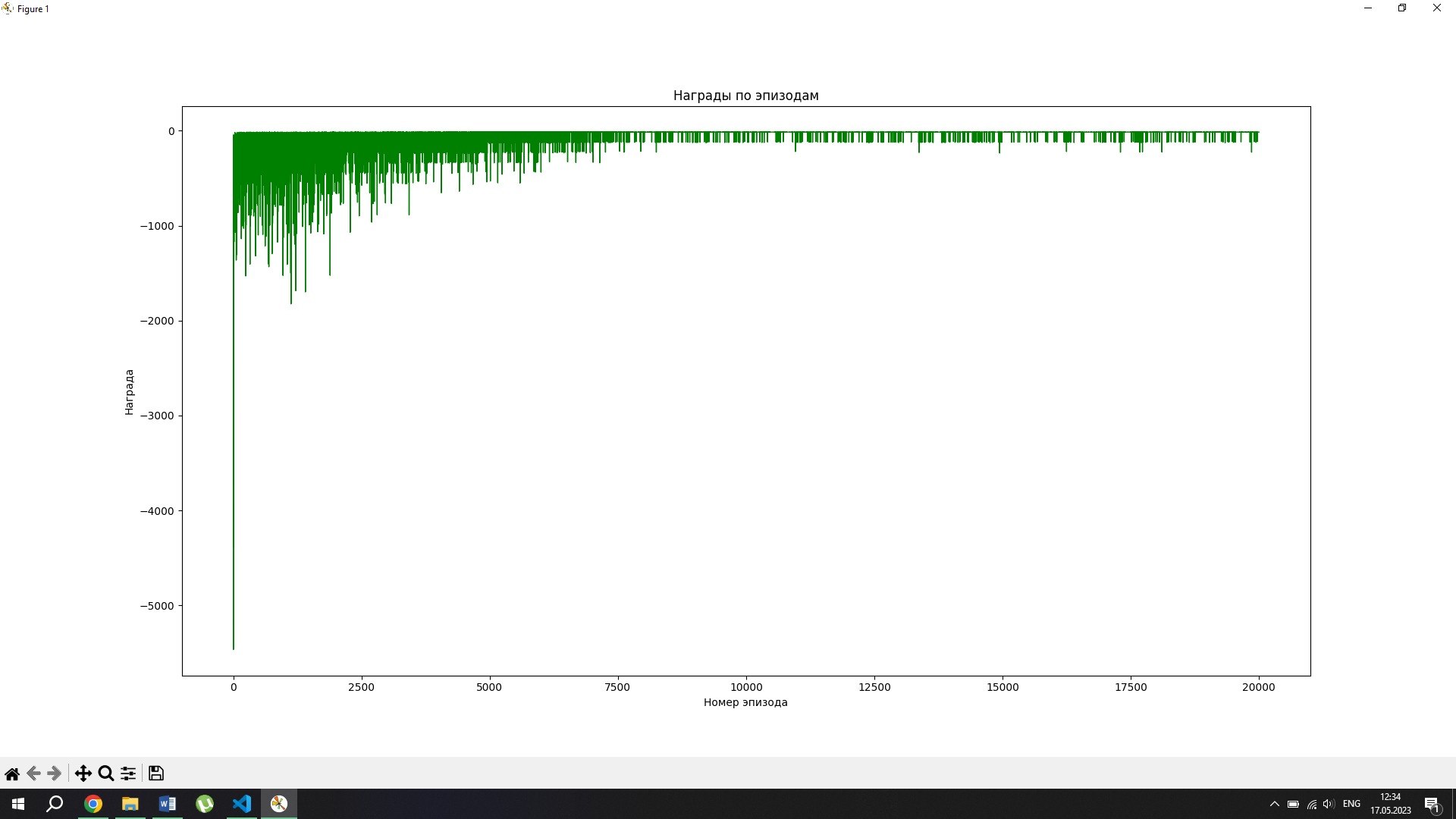


Рисунок 11 – Награды по эпизодам (Q-обучение)

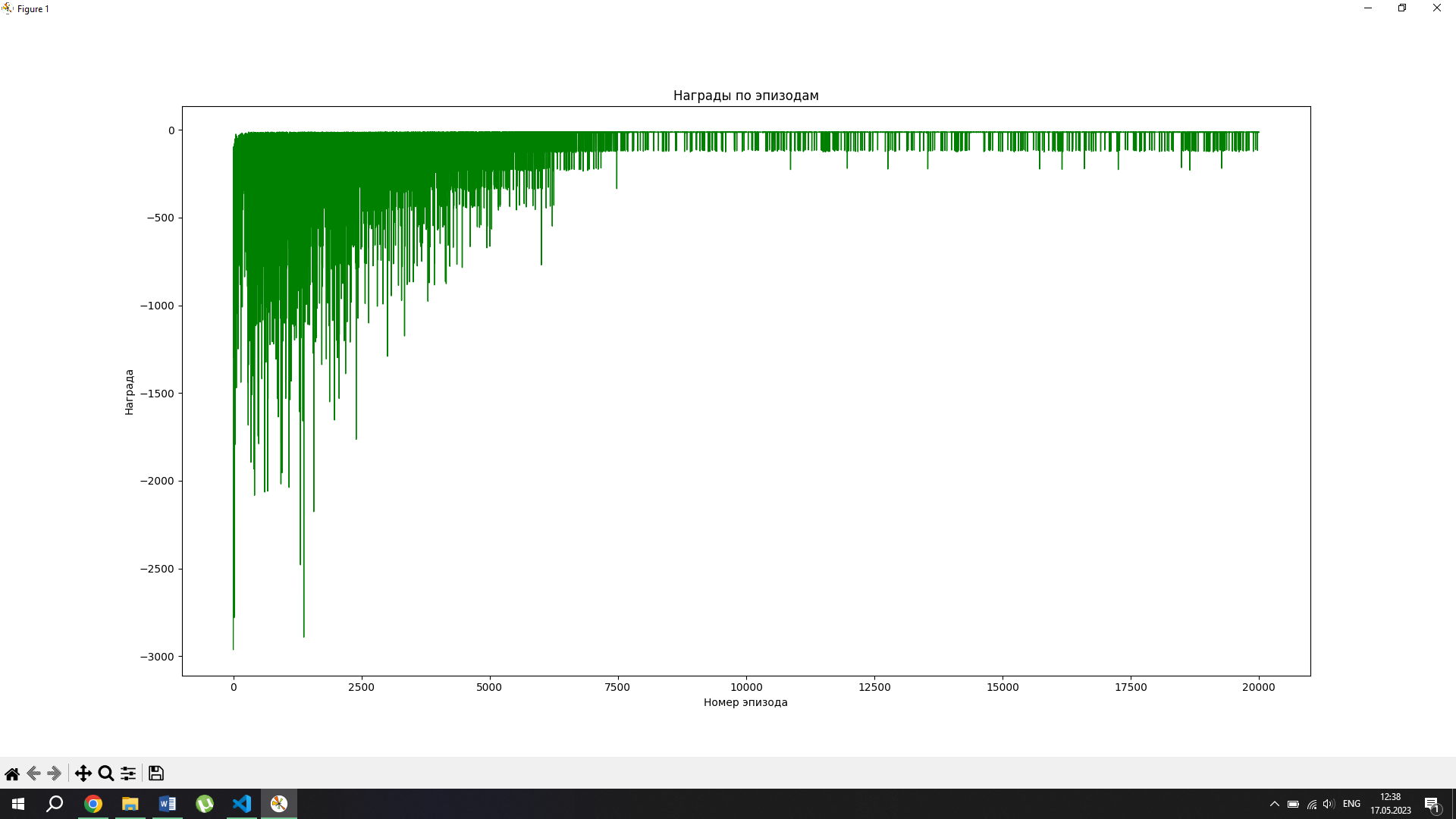


Рисунок 12 – Награды по эпизодам (Двойное Q-обучение)

Все алгоритмы смогли справиться с задачей нахождения маршрута в игре Cliff Walking, SARSA при этом нашла менее оптимальный путь, в то время как Q-алгоритмы нашли кратчайший.

**Вывод**

Таким образом, удалось реализовать алгоритмы SARSA, Q-обучения и двойного Q-обучения для среды обучения с подкреплением, таким образом ознакомившись с базовыми методами обучения с подкреплением на основе временных различий.