|  |  |
| --- | --- |
| lu135925on3bu_tmp_3360867a00ce4d37 | **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 и системы управления и искусственный интеллект

КАФЕДРА                  Системы обработки информации и управления

**Лабораторная работа №4** **по курсу «Методы машинного обучения в автоматизированных системах обработки информации и управления»**

Подготовили:

**У Жун**

**ИУ5И-25М**

25.05.2024

Проверил:

**Гапанюк Ю. Е.**

*2024 г*.

**Цель лабораторной работы:**

ознакомление с базовыми методами обучения с подкреплением.

**Задание:**

На основе рассмотренного на лекции примера реализуйте алгоритм Policy Iteration для любой среды обучения с подкреплением (кроме рассмотренной на лекции среды Toy Text / Frozen Lake) из библиотеки Gym (или аналогичной библиотеки).

Алгоритм итерации стратегии состоит из двух основных частей: оценка политики (policy evaluation) и улучшение политики (policy improvement).

Выбор окружения: Я выбрала окружение CartPole из библиотеки Gym.

Дискретизация состояний: я дискретизировала пространство состояний среды CartPole, чтобы использовать табличный метод для хранения функций значений и политик.

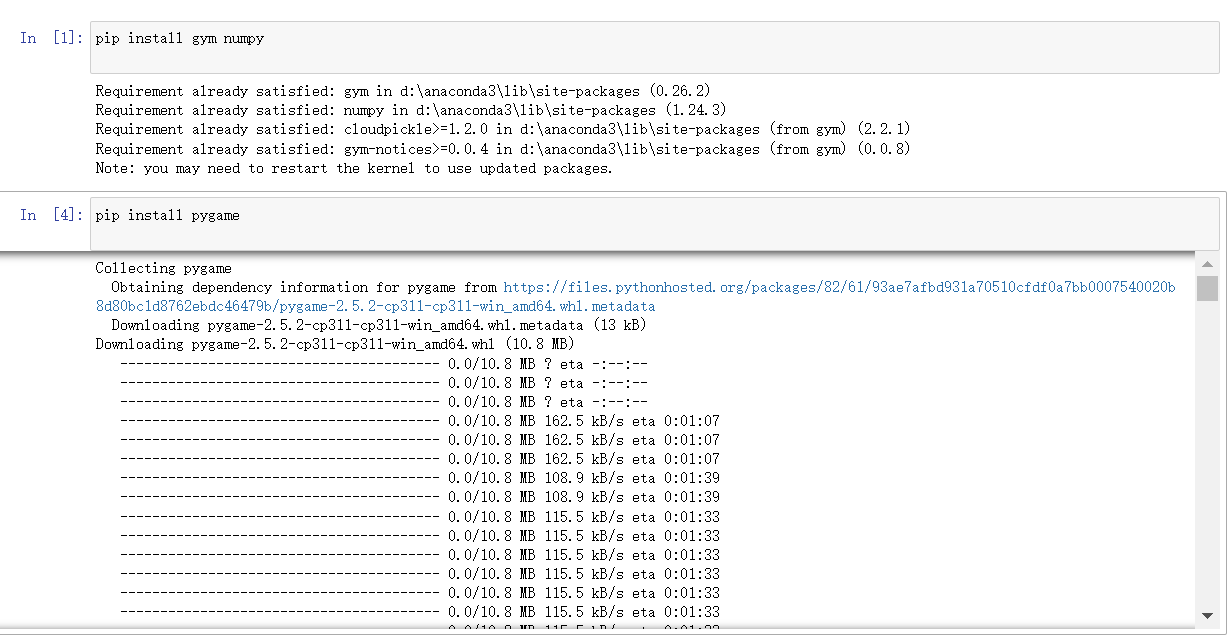
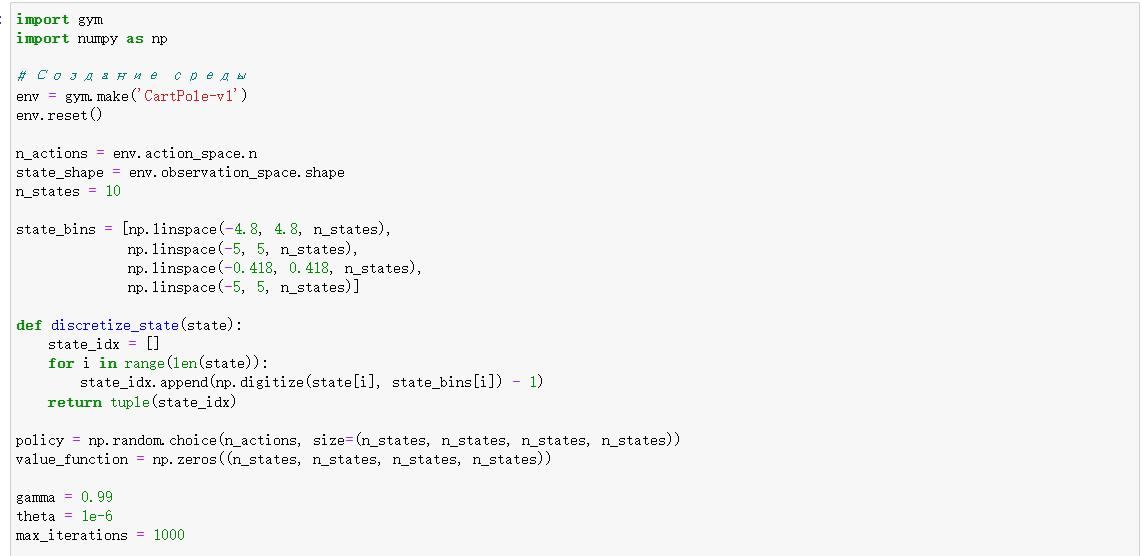
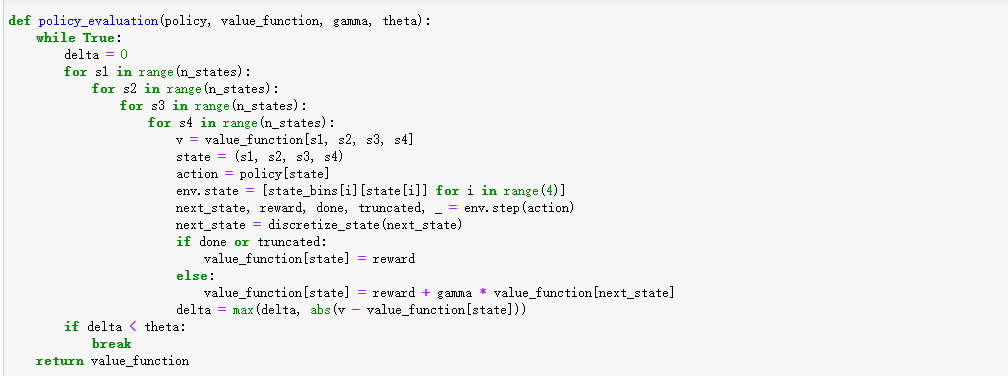
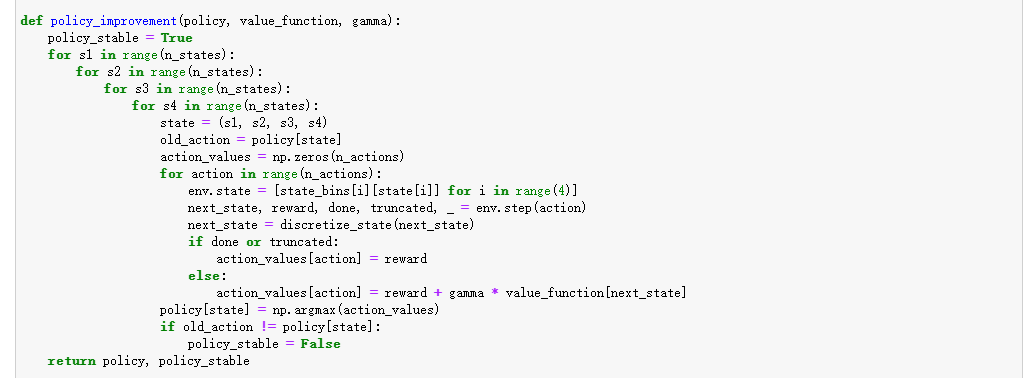
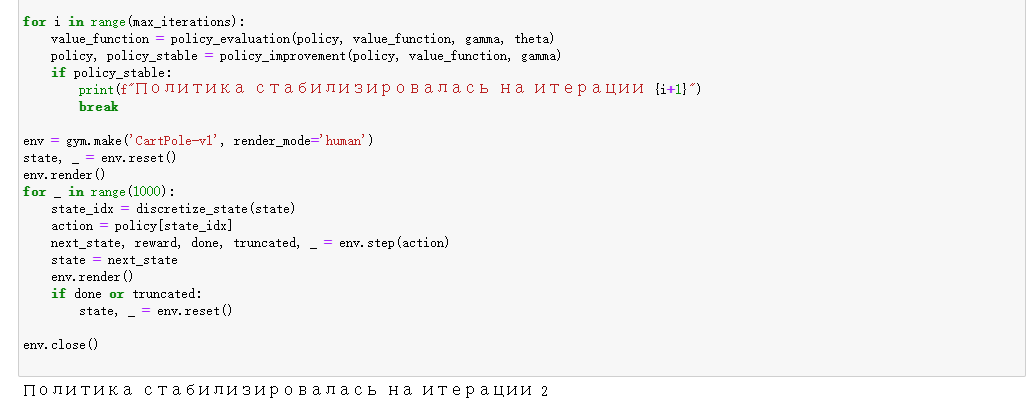
    

Рис.1-агент был обучен с помощью алгоритма итерации политик

появление окна визуализации среды CartPole и стабильное выполнение агента в этой среде означает, что задание успешно выполнено.



Рис. 2 – Окно визуализации среды CartPole