

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования **САНКТ-
ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

«Использование метода машинного обучения на основе Q-обучения для управления роботом- манипулятором»

Выпускная квалификационная работа

Автор:
Исхаков М.Р., группа Р3435
Руководитель:
к.т.н., доцент Ведяков А.А.

2019

Актуальность

- Интеллектуальные системы принятия решений в реальном времени
- Решение прикладных задач
- Относительно быстрая адаптация к неизвестным средам

Постановка задачи

Цель: Перемещение схвата манипулятора в заданную точку рабочей зоны. Перемещение скользящих по поверхности объектов в заданную точку плоскости.

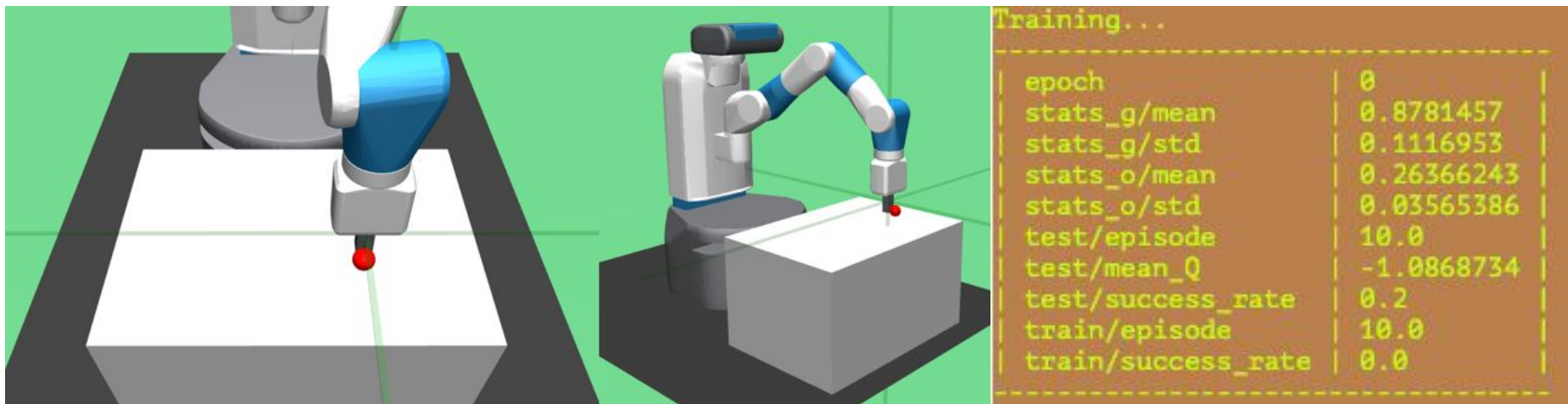
Требования:

- Вывод значений параметров системы по завершению каждой эпохи
- Моделирование физических законов в среде обучения
- Возможность реализации с использованием низкопроизводительной техники
- Визуализация полученных результатов по завершению обучения

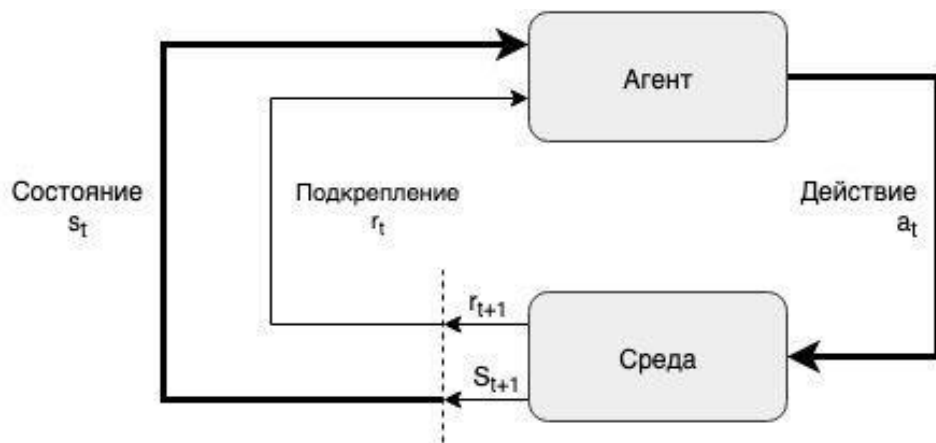
Перечень подлежащих разработке вопросов

1. Выбор инструментария для реализации методов машинного обучения
2. Анализ и выбор алгоритмов обучения, отвечающих поставленной задаче
3. Выбор среды окружения
4. Реализация управления манипулятором посредством алгоритмов обучения с подкреплением
5. Вывод, визуализация и анализ полученных результатов обучения системы

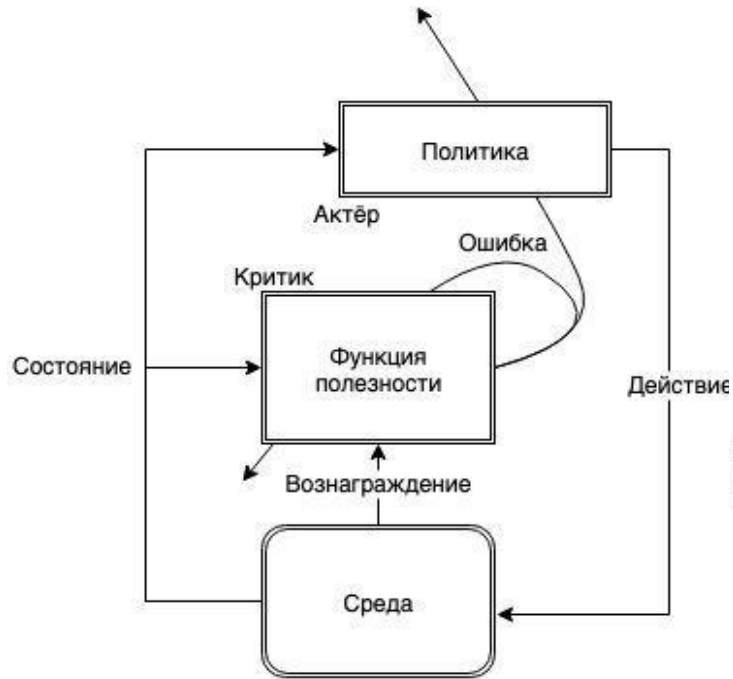
Инструментарий для исследований обучения с подкреплением Gym



Обучение с подкреплением. Q-обучение

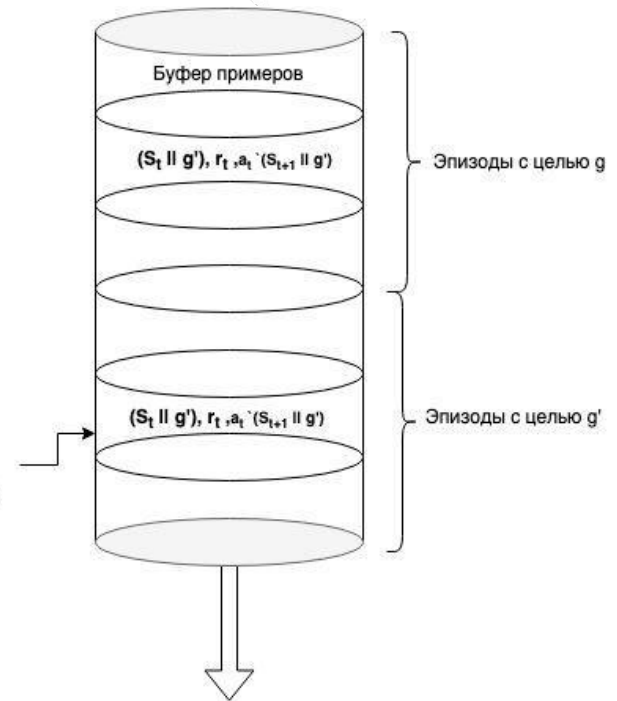


Алгоритмы

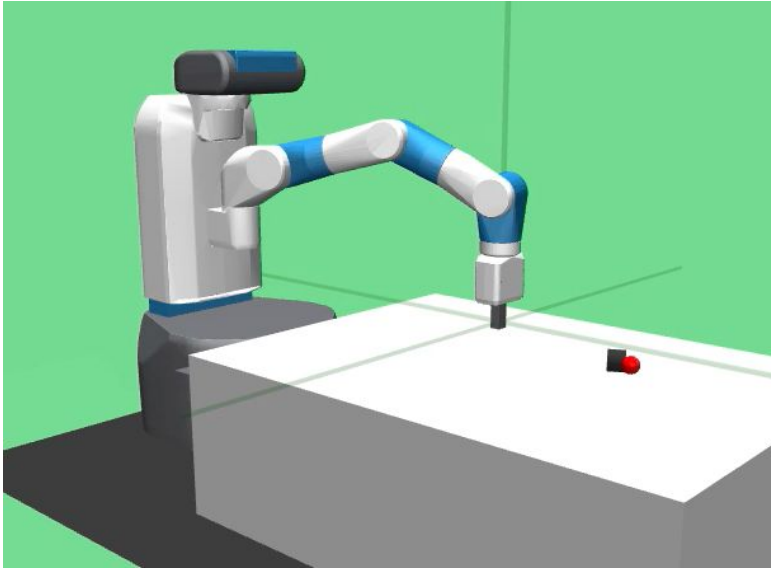


DDPG

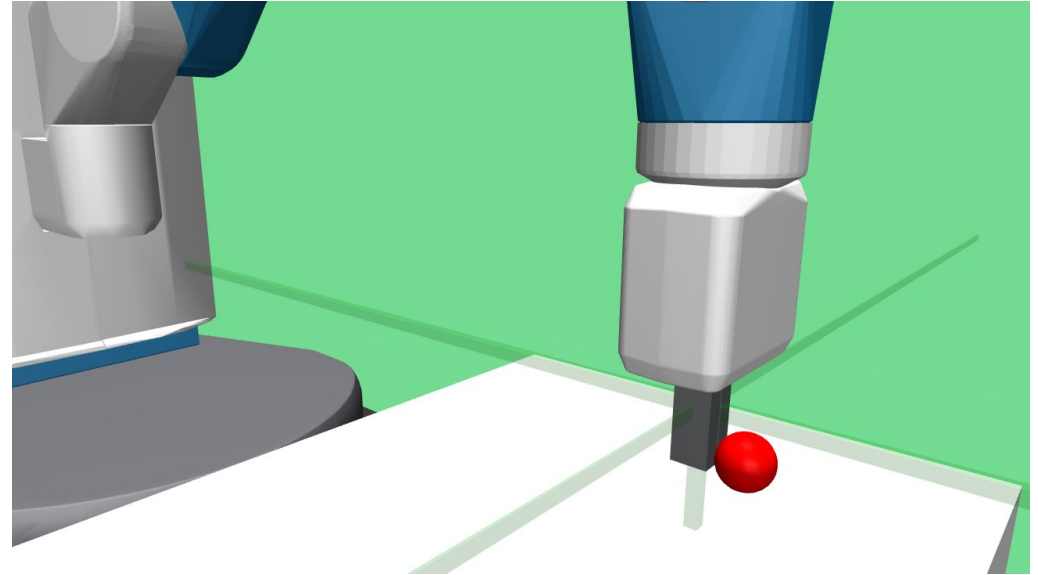
Использование примеров для получения "новой" цели g' и награды $R(s_0, a_0, g)$



HER



FetchSlide-v1



FetchReach-v1

График функции полезности Q от номера эпизода

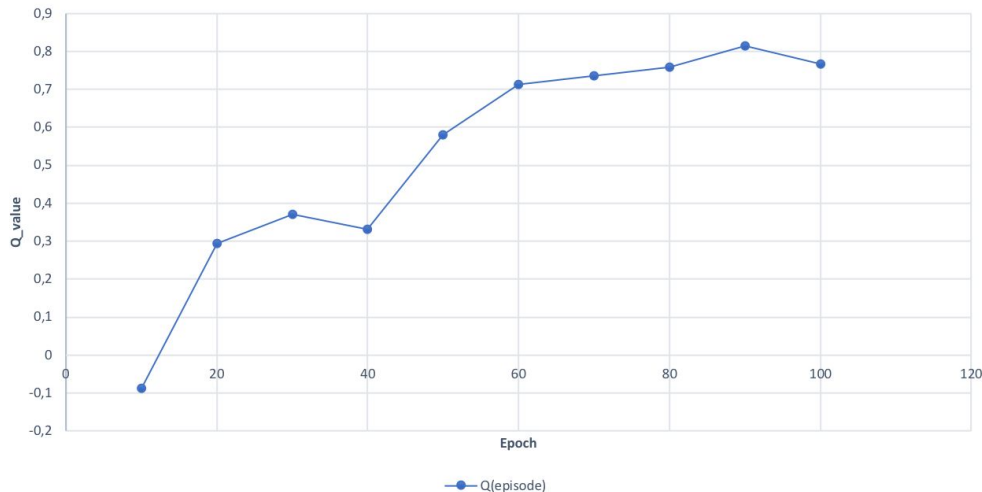
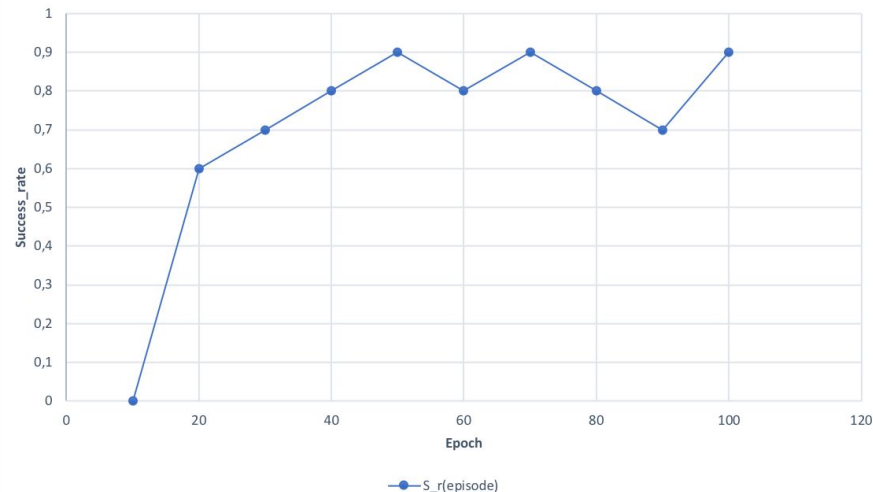


График рейтинга успеха от номера эпизода



HER + DDPG
FetchReach-v1
100 ЭПИЗОДОВ

HER + DDPG
FetchReach-v1
100 ЭПИЗОДОВ



График функции полезности Q

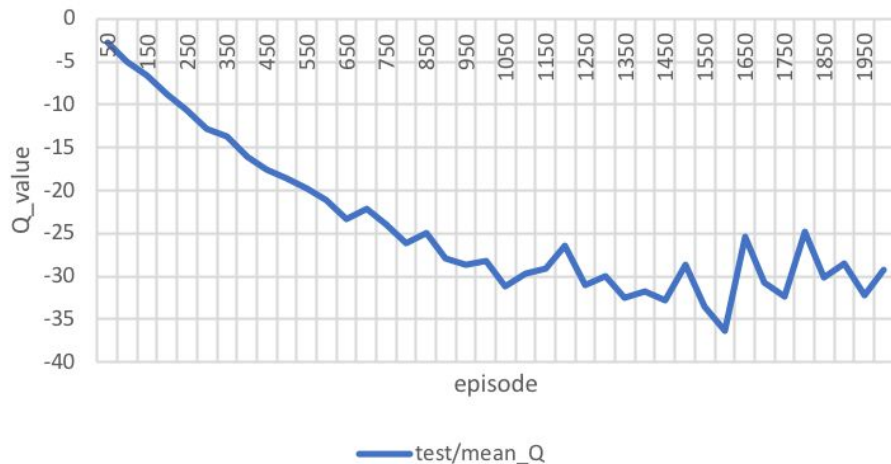
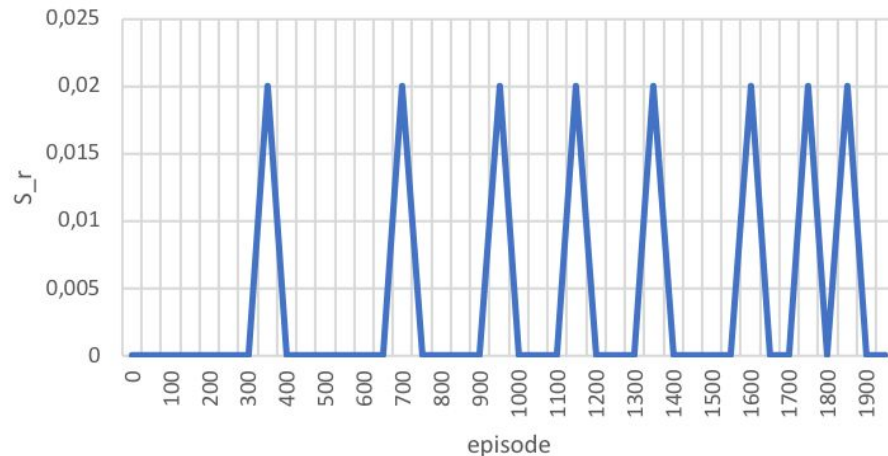


График рейтинга успеха



HER + DDPG
FetchSlide-v1
2000 эпизодов



HER + DDPG
FetchSlide-v1
2000 эпизодов

Заключение

- Получены и проанализированы параметры систем на протяжении обучения
- Реализованный алгоритм решает обратную задачу кинематики с коэффициентом успеха 0,9 к 100 эпизоду обучения
- Согласно полученным результатам реализация исследованных методов возможна с использованием низкопроизводительной техники
- Выполнена визуализация работы обученных систем
- Использован физический симулятор Mujoco