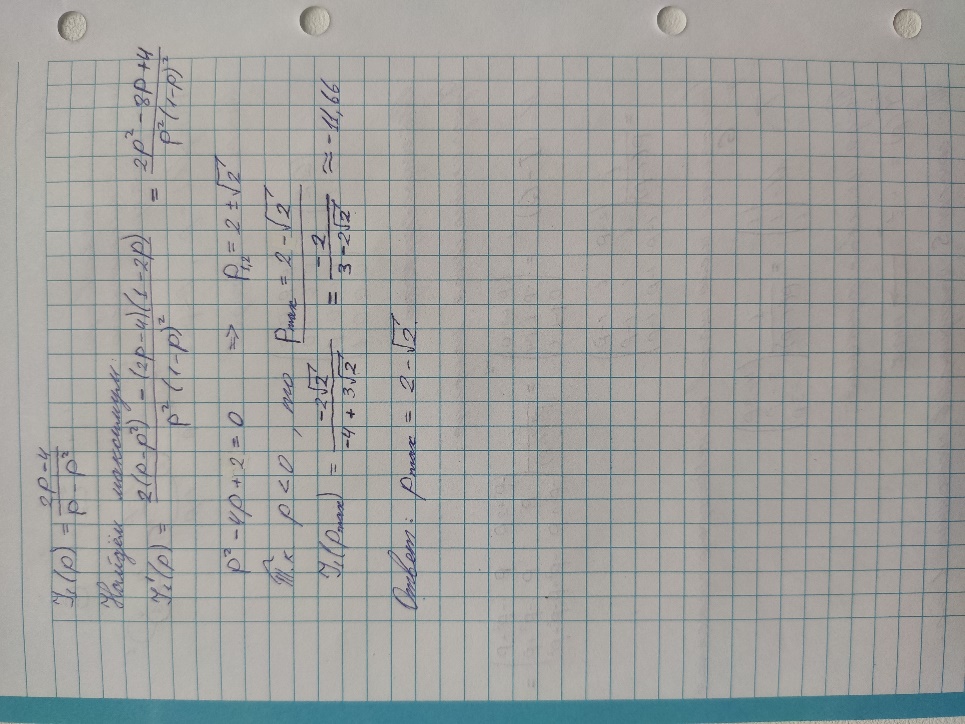
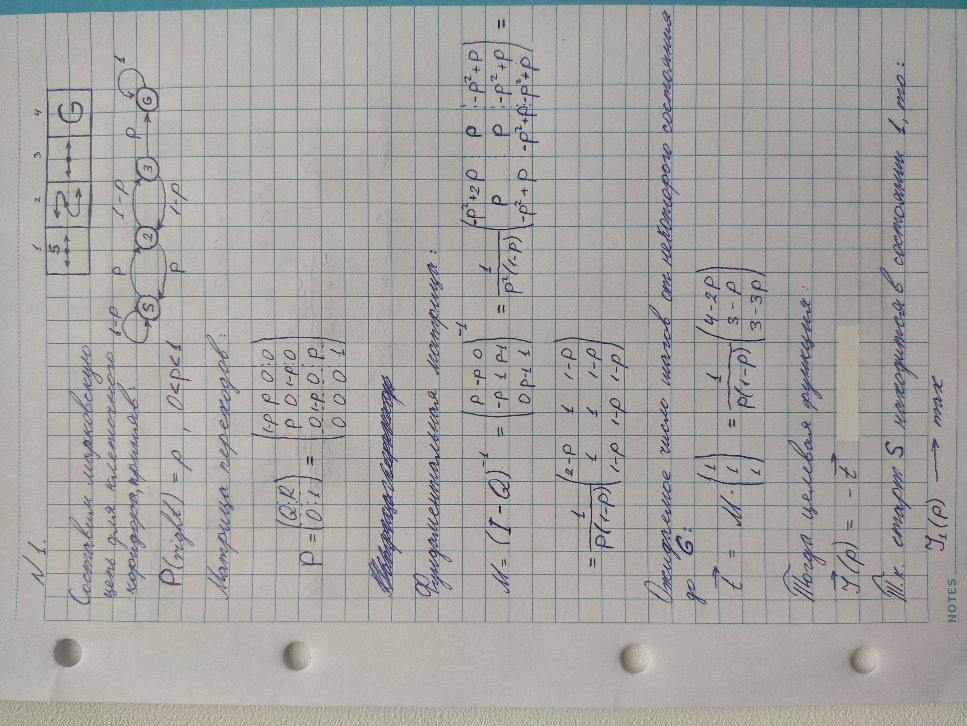
# Теория

## Задача 1



# Практика

## Первые шаги

Заполнил все пропуски “TODO” и преступил к экспериментам.

## Конфиги и Аугментации

Базовый конфиг экспериментов:

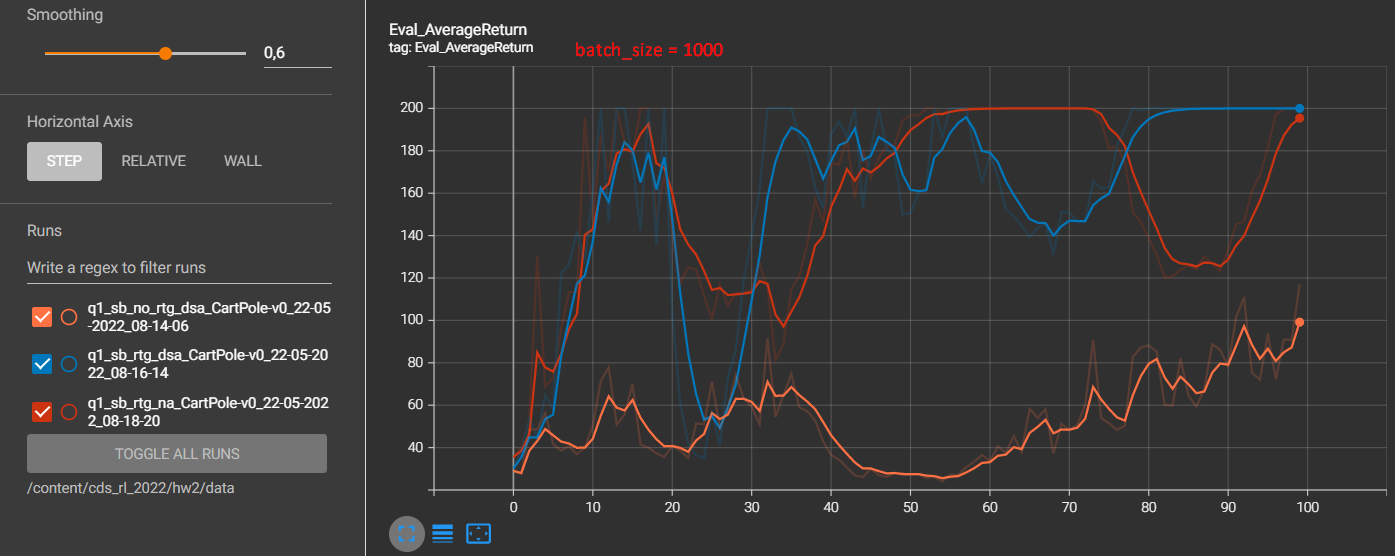


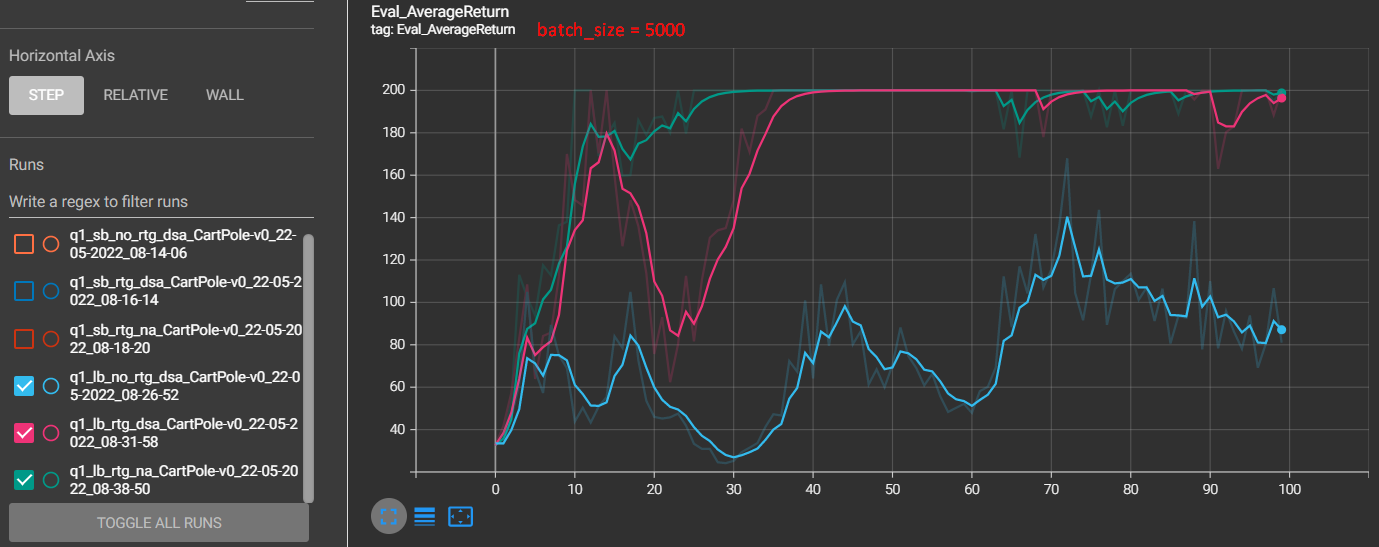


## Важные замечания

## Задание 1

Эксперименты с маленьким размером батча:



Эксперименты с большим размером батча:

1. Какой из вариантов оценки отдачи имеет лучшие результаты без нормализации значения преимущества?

При dsa=True лучший результат дает rtg=True, batch\_size=5000.

1. Помогает ли нормализация значения преимущества?

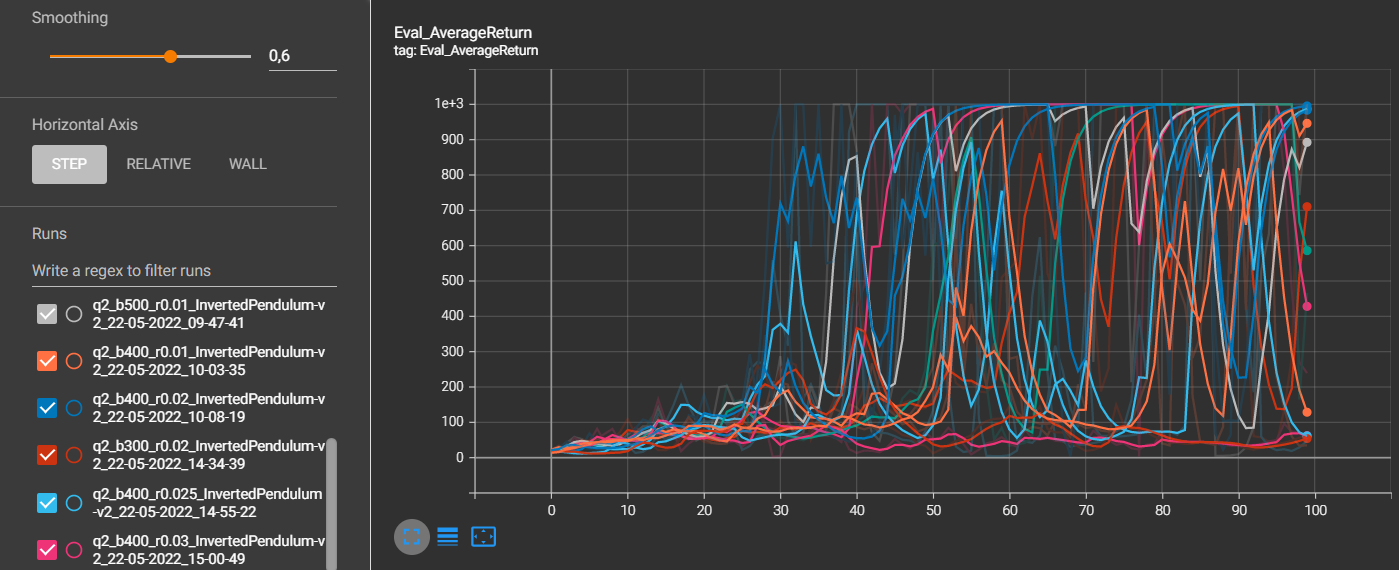
При малом размере батча улучшение незаметно (если вообще есть), а при большом батче – обучение проходит быстрее, нет больших отклонений.

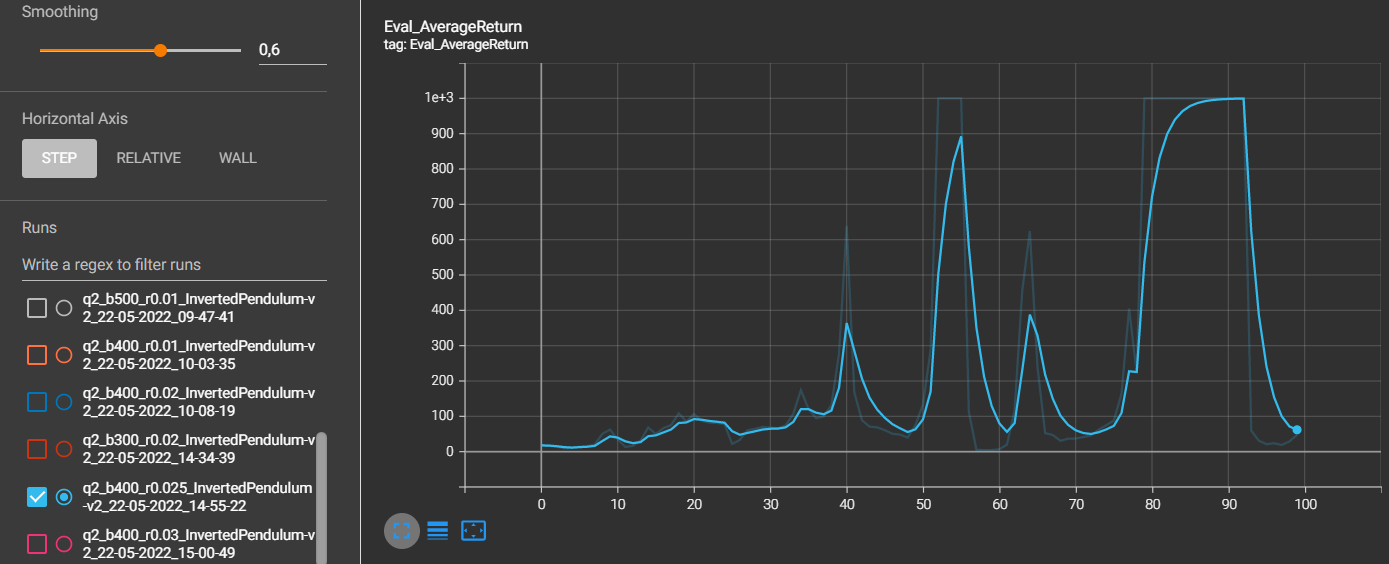
1. Как влияет размер пакета на качество обучения?

Увеличение размера батча привело к повышению отдачи.

## Задание 2

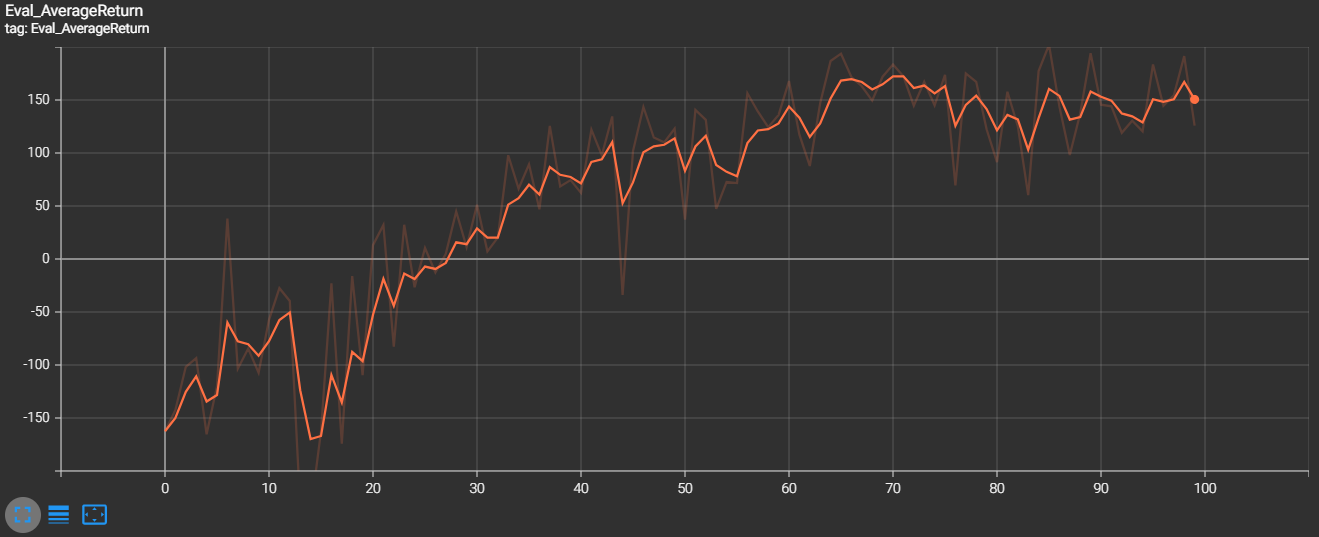
Была проведена целая куча экспериментов с batch\_size от 300 до 1000 и lr от 0.005 до 0.03:



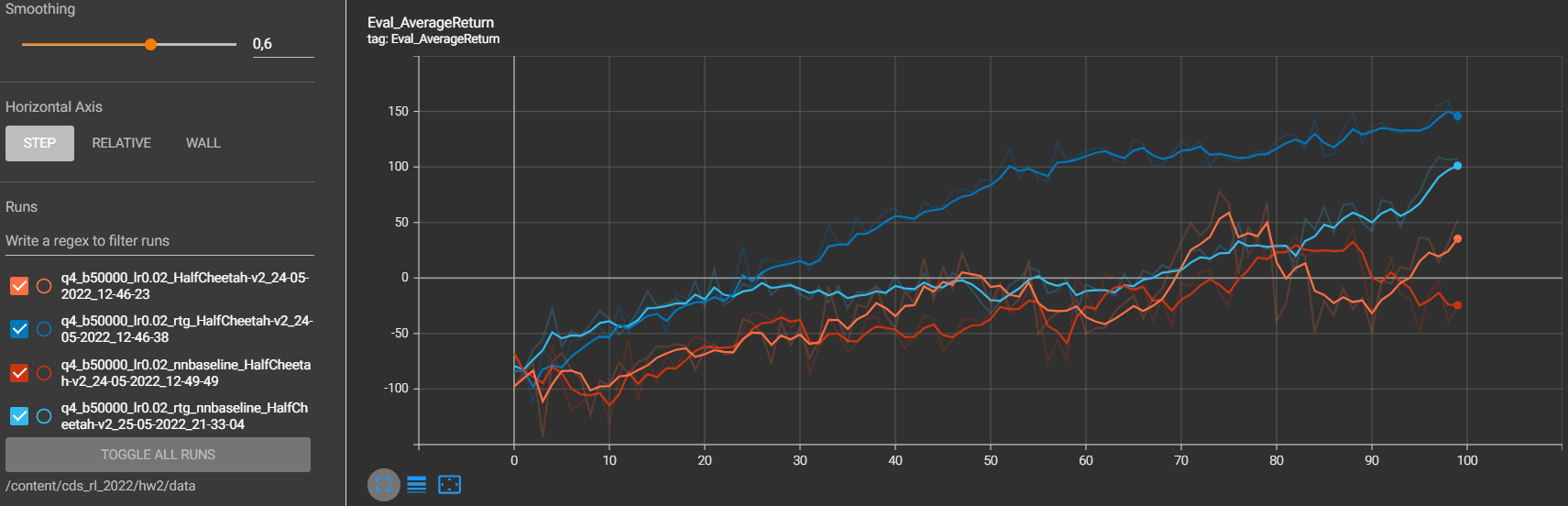
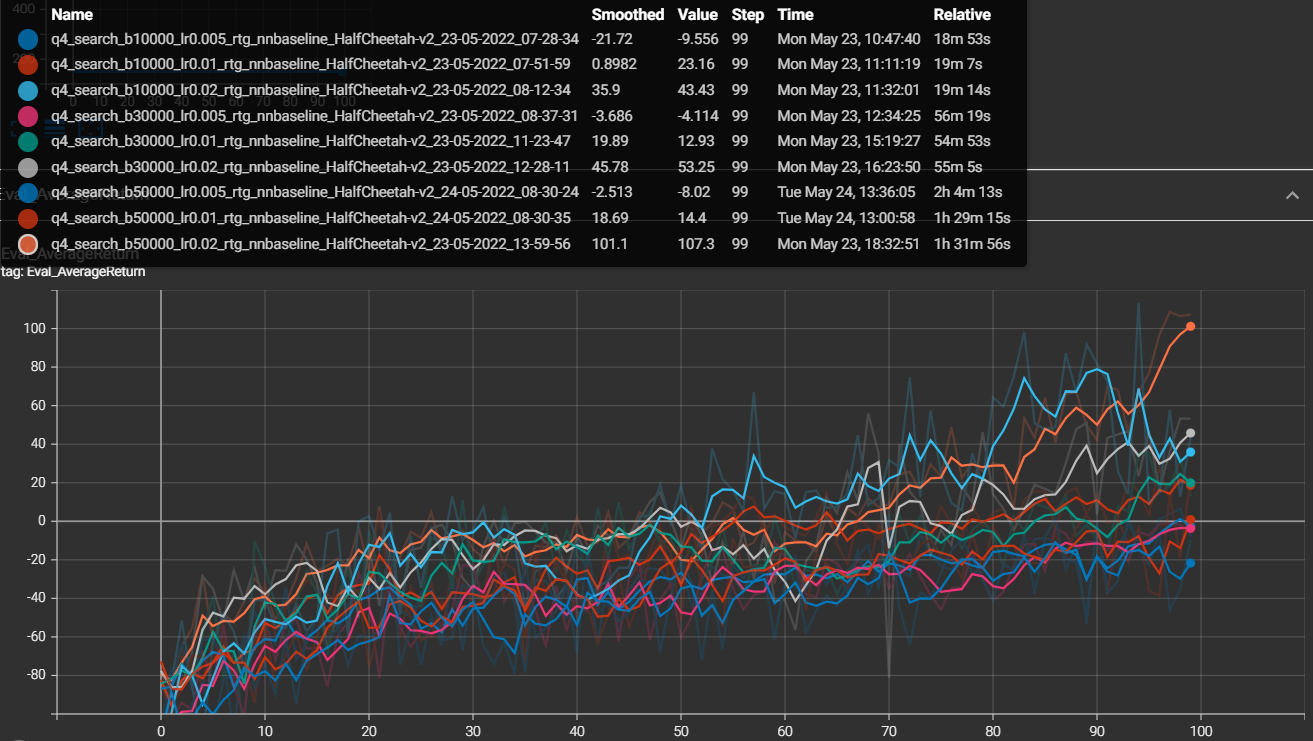
Удалось найти batch\_size=400 и lr=0.025, при которых алгоритм достигает достичь 1000 менее, чем за 100 итераций:

## Задание 3

Средняя отдача для LunarLanderContinuous-v2 с использованием градиента стратегии. Итоговый уровень немного ниже ожидаемого (150 < 180):



## Задание 4

По результатам поиска, наилучшую отдачу показал эксперимент с batch\_size=50000 и lr=0.02:

## Задание 5

Увеличение λ приводит к увеличению скорости обучения, средняя отдача выше: