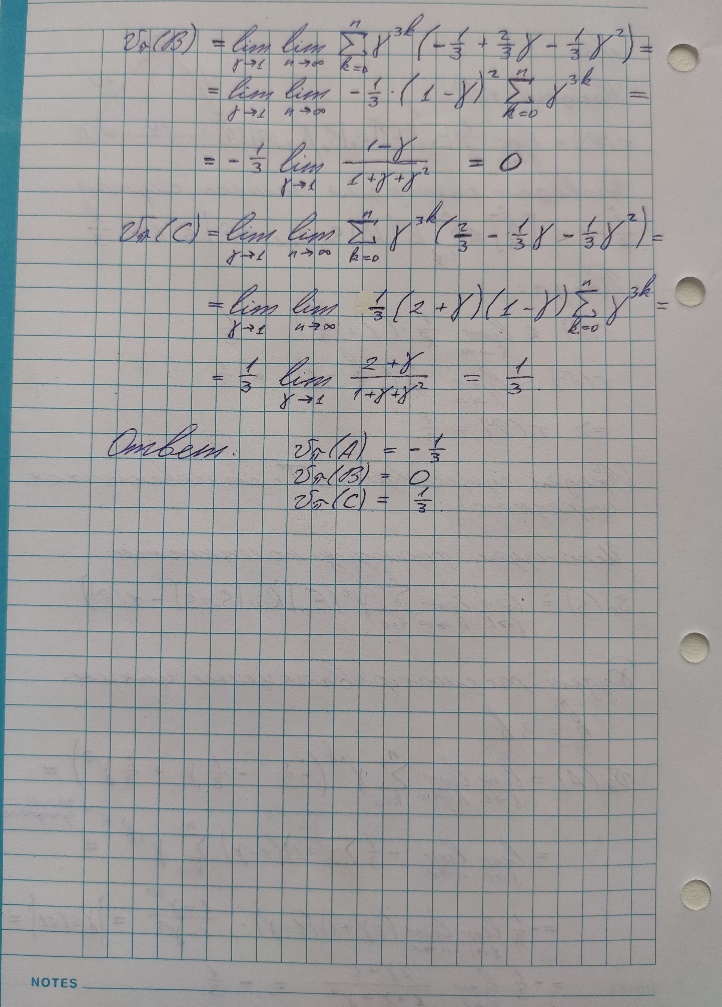
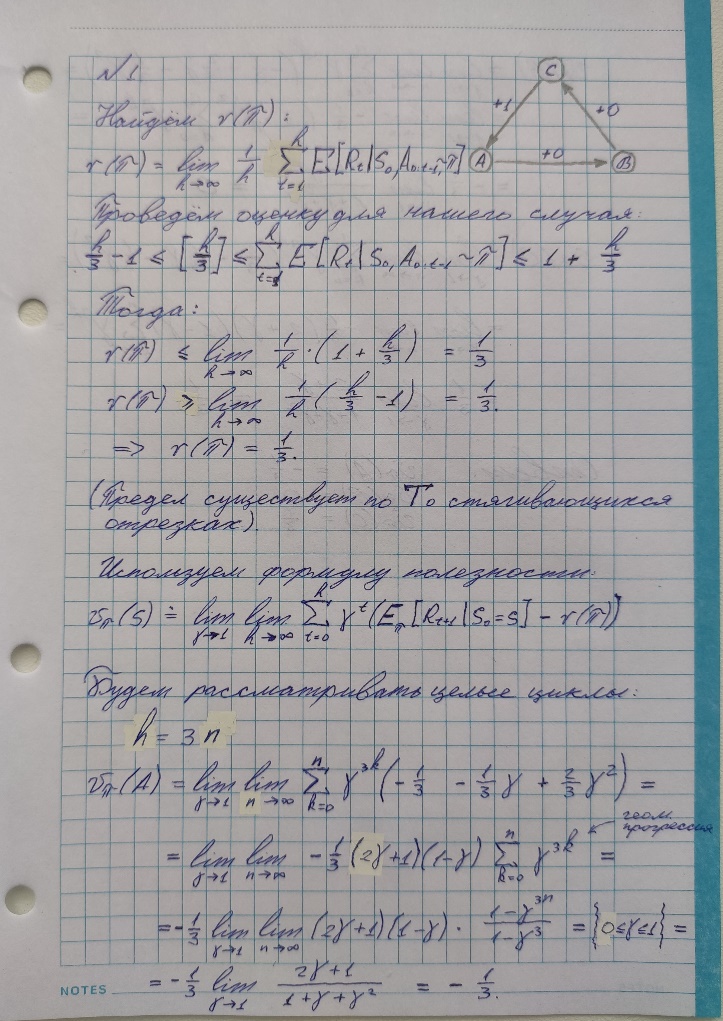
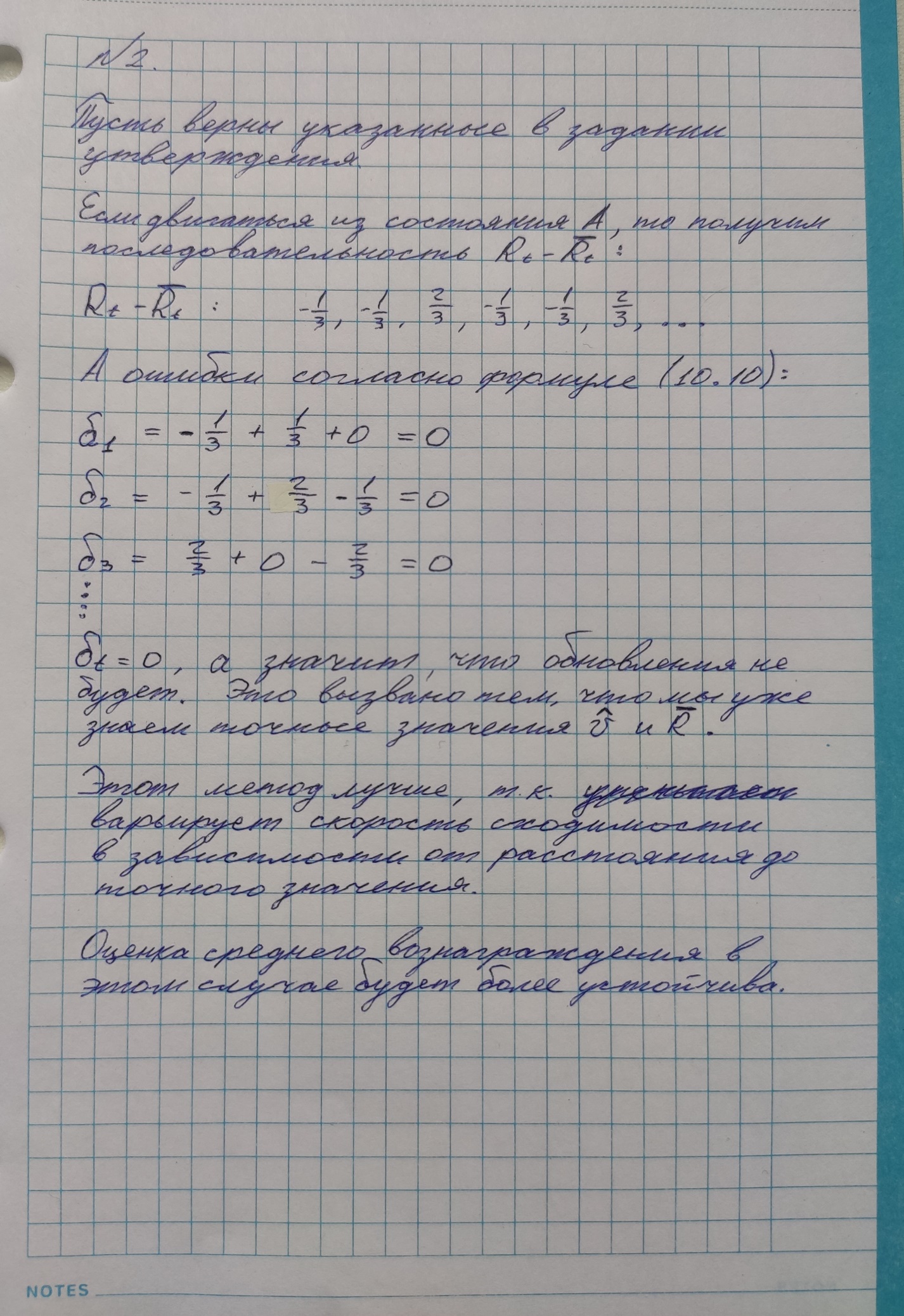
# Теория

## Задача 1



## Задача 2



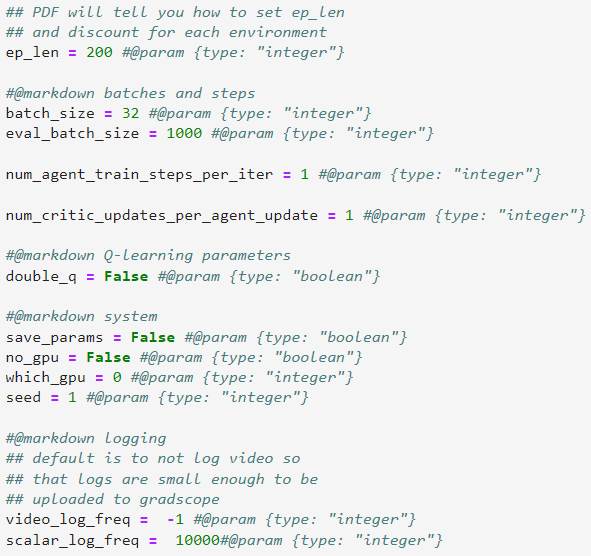
# Практика

## Первые шаги

Заполнил все пропуски “TODO” и преступил к экспериментам.

## Конфиги и Аугментации

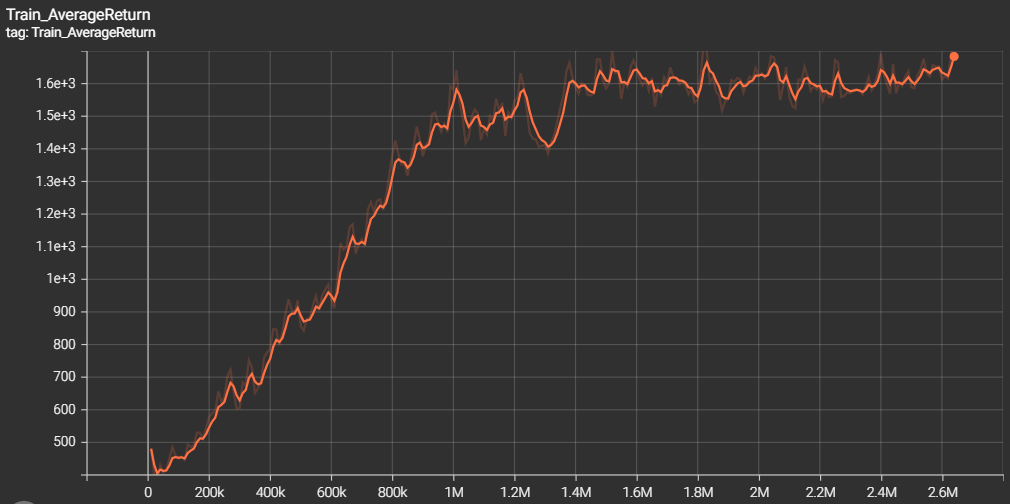
Базовый конфиг экспериментов для DQN:

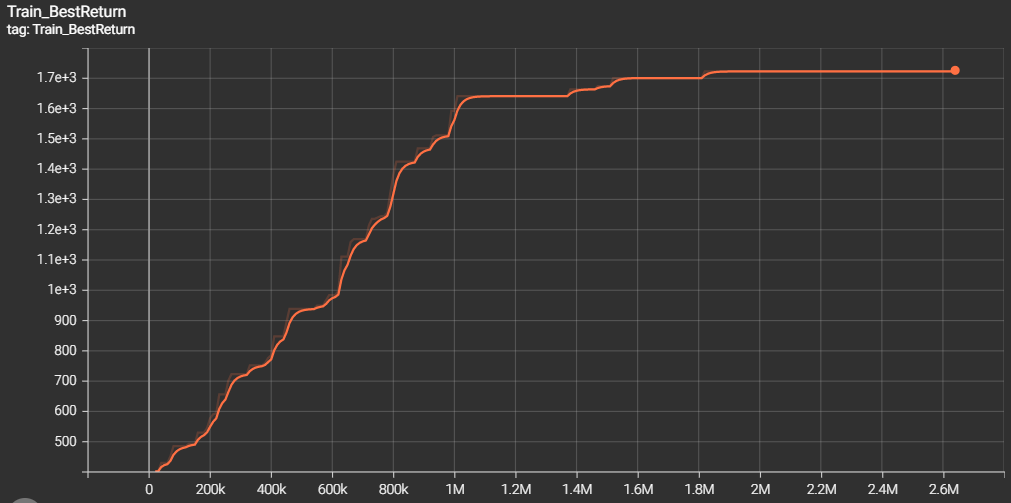


Базовый конфиг экспериментов для Actor-Critic:

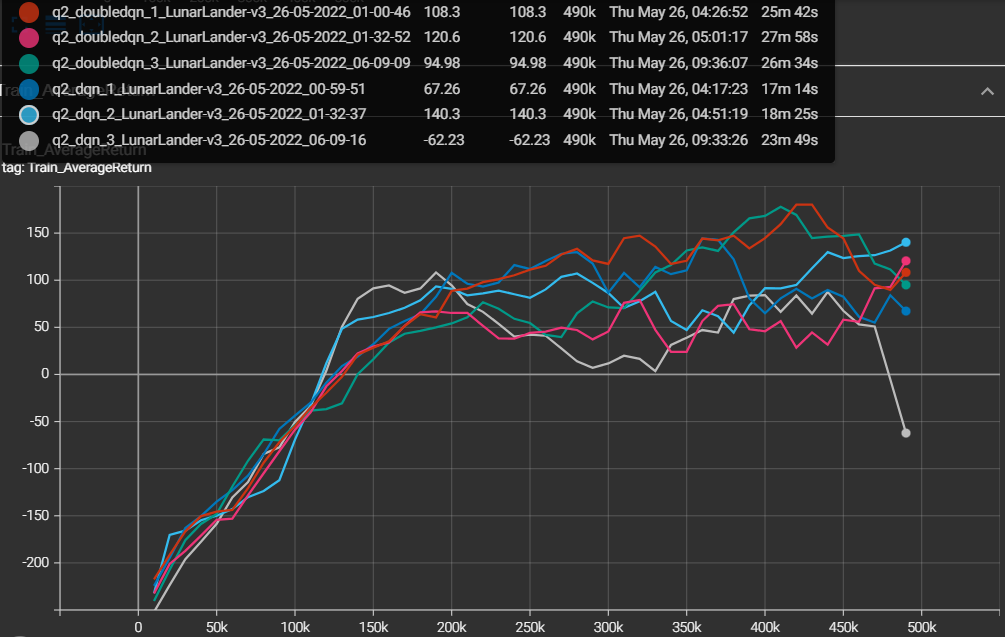
## Важные замечания

## Задание 1

Ниже приведены графики средней и максимальной отдачи при обучении в среде MsPacman-v0:

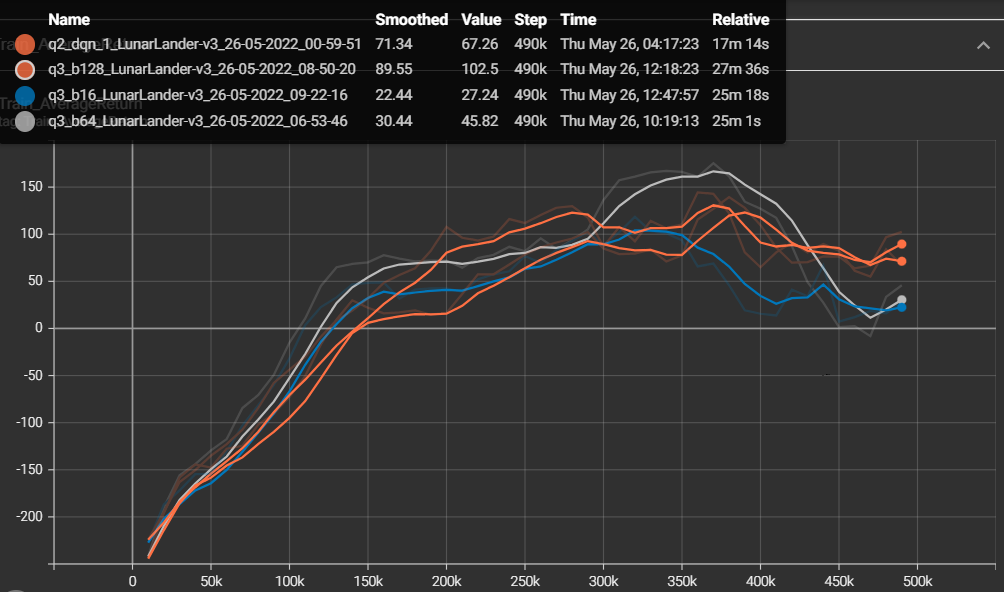


## Задание 2

Сравнение DQN и DDQN с разным seed. Видно, что DDQN в среднем показывает себя лучше:

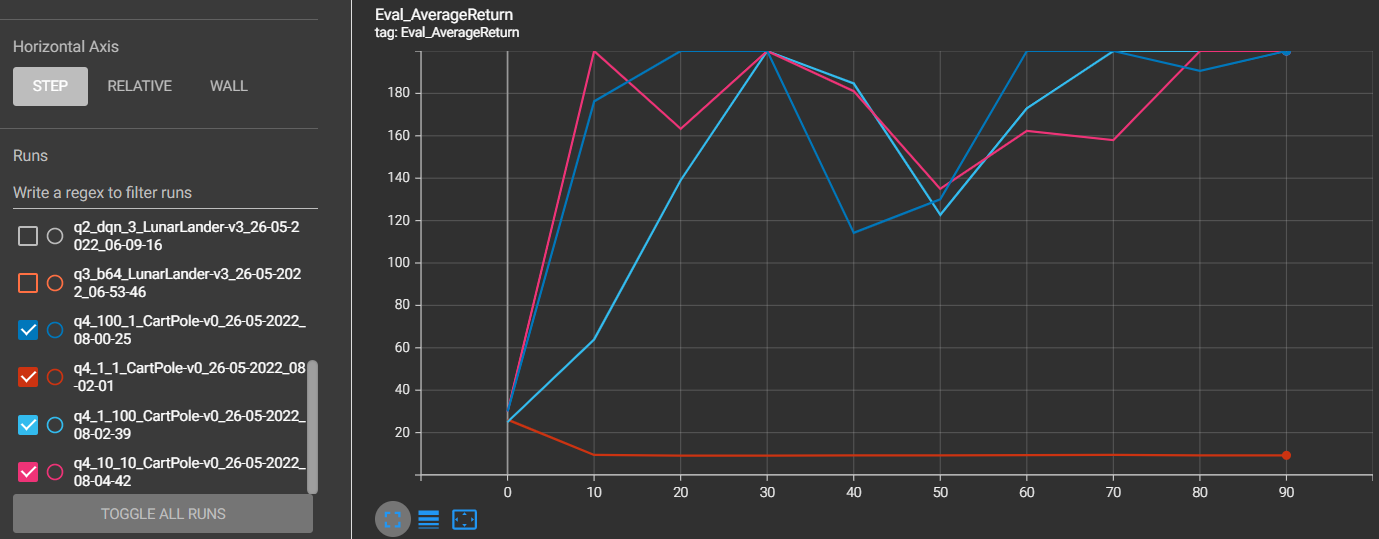
## Задание 3

Была исследована зависимость качества от размера батча. b [16, 32, 64, 128]:



Лучший результат в итоге показал b=128

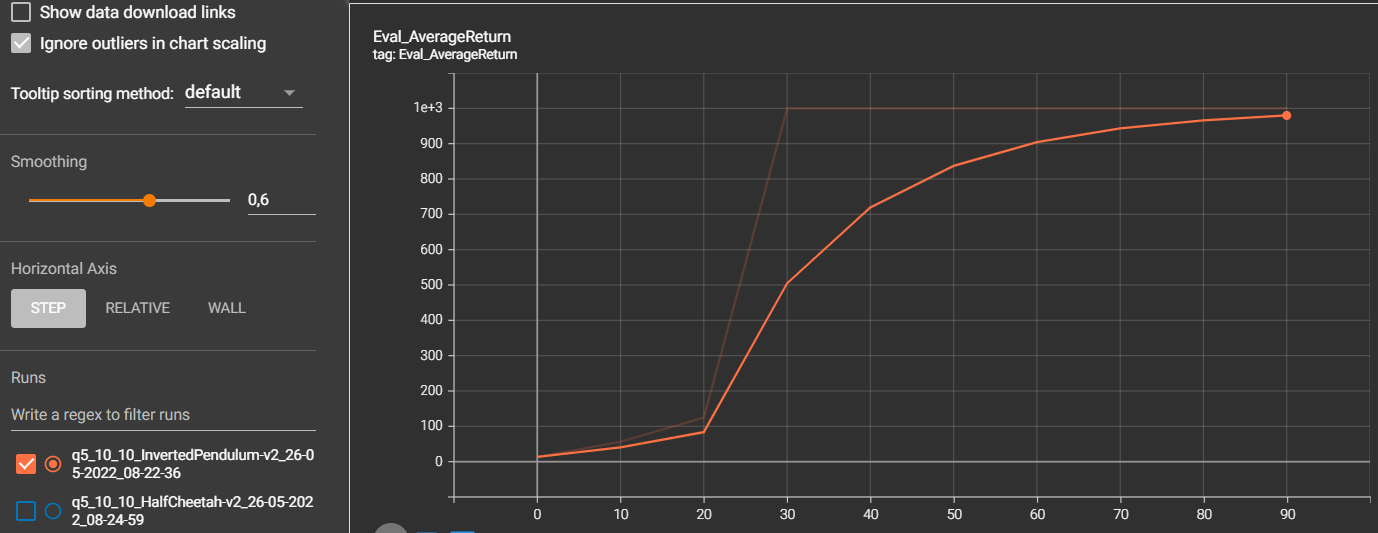
## Задание 4

Сравнение запусков метода автор-критик:

Сложно сказать, какой из экспериментов показал себя лучше, т.к. они очень схожи. Выбрал q4\_10\_10, т.к. он первым принимает максимальное значение.

## Задание 5

Средняя отдача в среде InvertedPendulum-v2, отдача в 1000 получается уже на 30-ой итерации:



Средняя отдача в среде HalfCheetah-v2, итоговая отдача получается несколько выше, чем в hw2:

