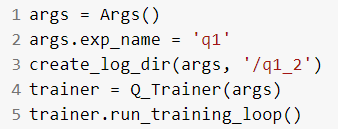
# Практика

Код можно посмотреть в тетрадках.

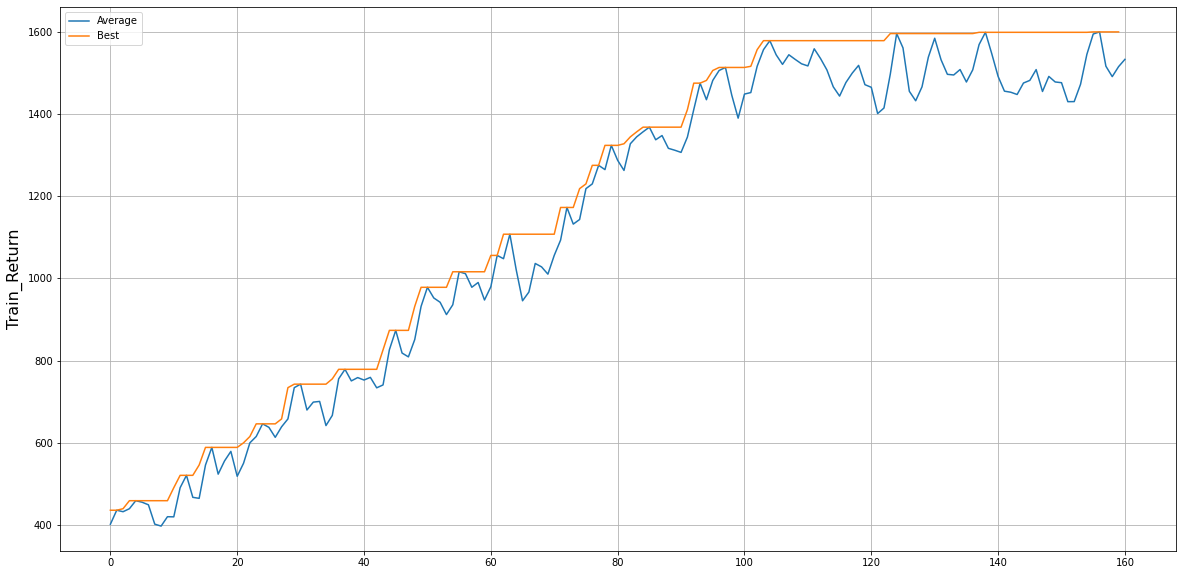
## Задание 1

Качество базовой версии Q-обучения.

Код обучения:



Результаты:

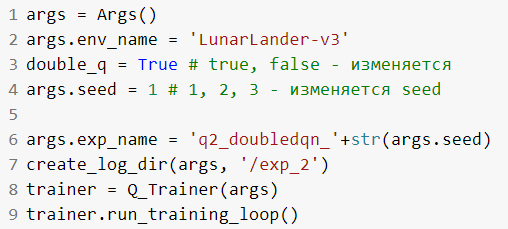


\*10^4 Timestep

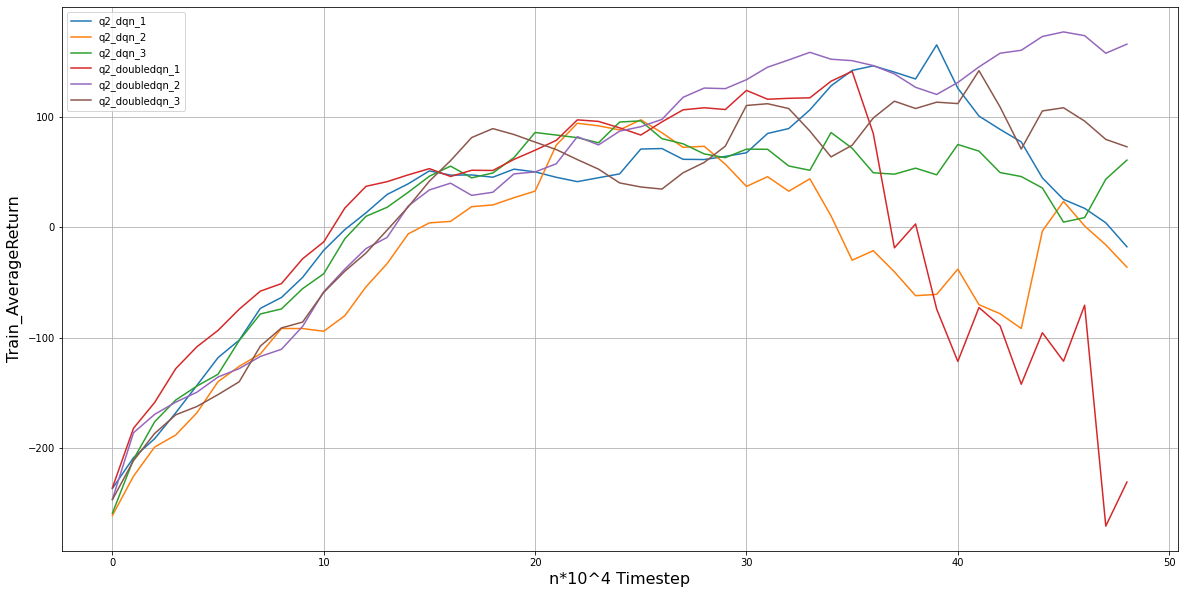
## Задание 2

Качество двойного Q-обучения

Код запуска:

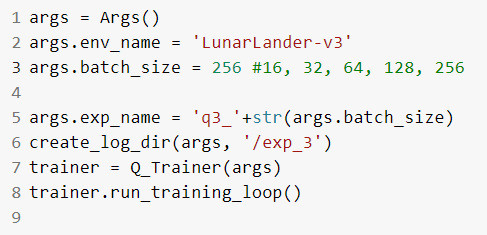


Результаты:

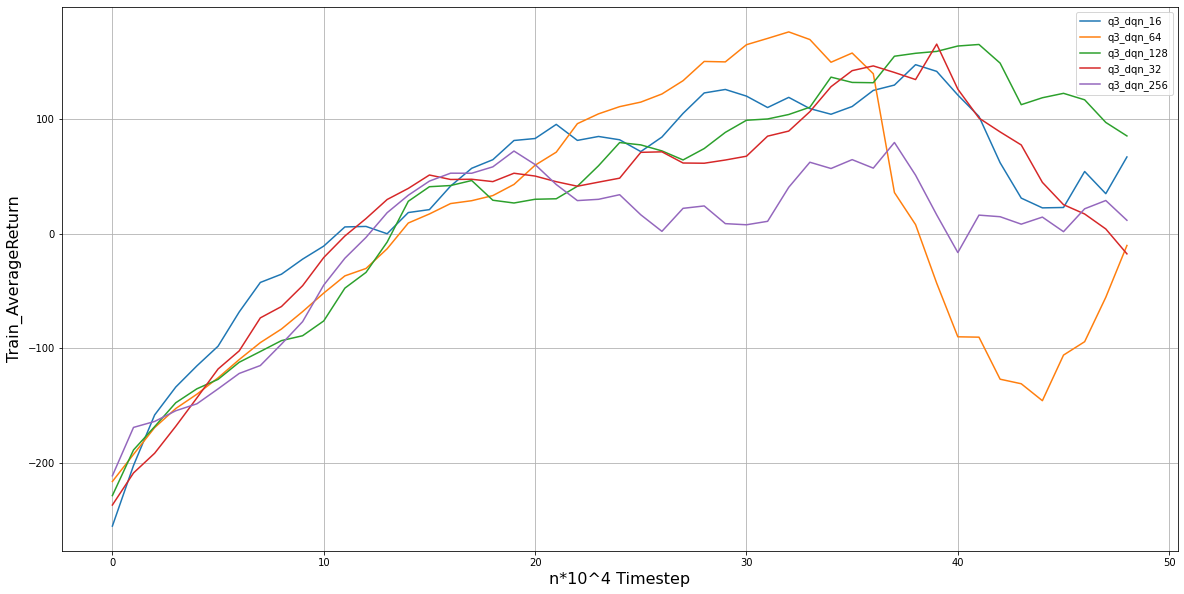


## Задание 3

Эксперимент с гиперпараметрами. Взят самый простой по доступу

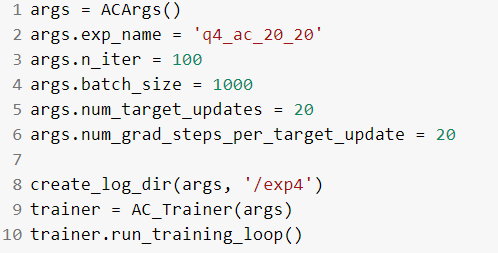


Результатами:

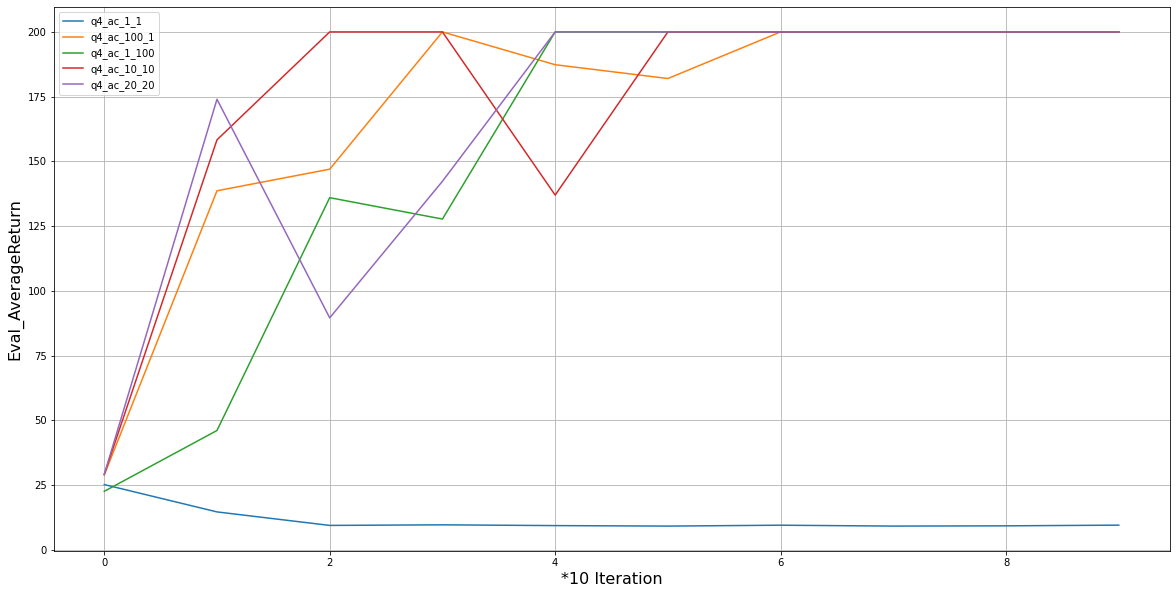


## Задание 4

Код:



Результаты:

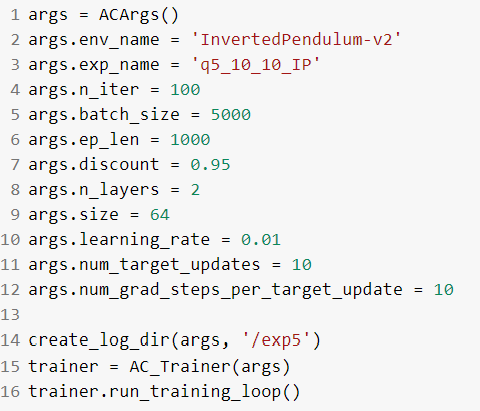


Результат эксперимента показал, что выгодно увеличивать одновременно два параметра num\_target\_updates и num\_grad\_steps\_per\_target\_update. Лучшими параметрами получились 10 и 10. 20 и 20 могли бы быть лучше, но не сильно.

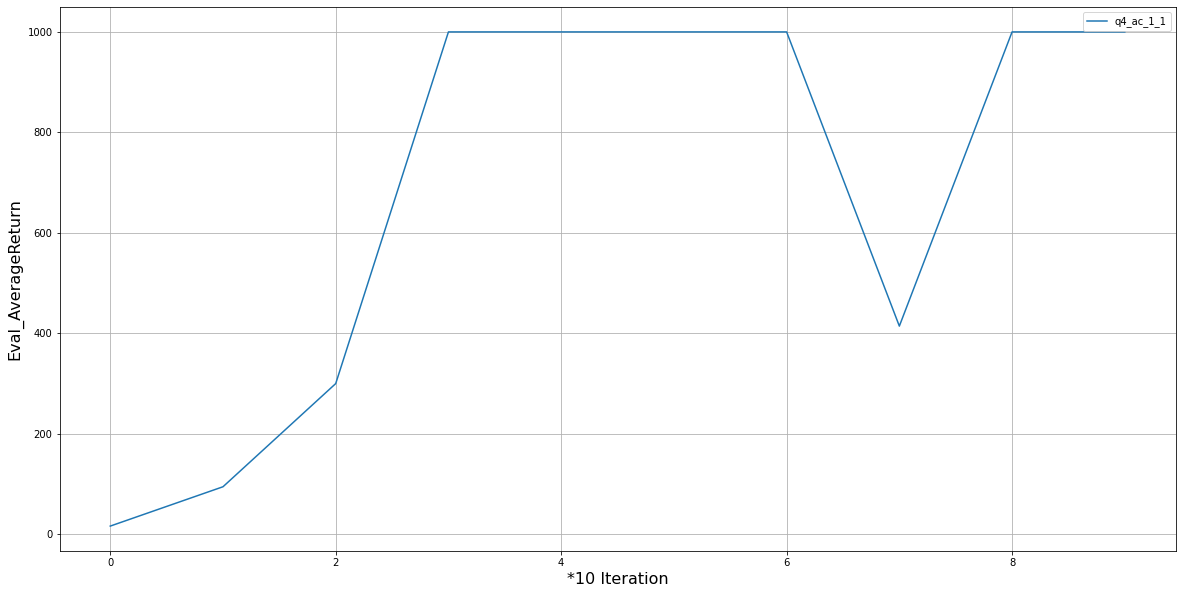
## Задание 5

## InvertedPendulum-v2

Код:

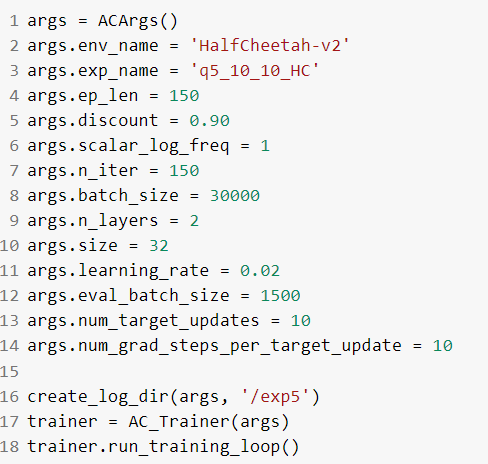


Результаты:



## HalfCheetah-v2

Код:



Результаты:

