**Results for the pendulums in case of a curriculum learning**

Some notes:  
quad and square root, also exp could be better than the others; experiments are usually done with PPO, since A2C doesn’t work properly (if another – it’ll be a note).

1. Изображение выглядит как текст, снимок экрана, Шрифт

   Автоматически созданное описание

TP\_PPO\_r1\_curriculum\_Model1\_PPO\_290816\_78\_tests – best model

1. Изображение выглядит как текст, снимок экрана, Шрифт

   Автоматически созданное описание

0 successful tests in all 3 models with random seeds during 300k training

PPO\_quad\_v0\_rpf09\_Model0\_PPO\_fail

1. Изображение выглядит как текст, снимок экрана, Шрифт

   Автоматически созданное описание

0 successful tests

PPO\_sqrt\_v0\_rpf09\_Model0\_PPO\_fail

1. Изображение выглядит как текст, снимок экрана, Шрифт

   Автоматически созданное описание

PPO\_exp\_v0\_rpf09\_Model1\_PPO\_253952\_91\_tests – 91 tests, Carl!!! The best model for now. Is it because of the ‘exp’ or the control values?

1. Изображение выглядит как текст, снимок экрана, Шрифт

   Автоматически созданное описание

PPO\_exp\_v1\_rpf09\_Model0\_PPO\_348160\_71\_tests – not bad, but late. Could be tested with more time since total t was 350k and the increase was at the very end. But are we aiming for the speed or quality?

1. Изображение выглядит как текст, снимок экрана, Шрифт

   Автоматически созданное описание

PPO\_exp\_v2\_rpf09\_Model2\_PPO\_303104\_70\_tests – average

1. Изображение выглядит как текст, снимок экрана, Шрифт

   Автоматически созданное описание

PPO\_discrete\_v0\_rpf09\_Model2\_PPO\_399360\_73\_tests – not bad, but again, could be needed more testing time for that.

1. Изображение выглядит как текст, снимок экрана, Шрифт

   Автоматически созданное описание

PPO\_x5\_v0\_rpf09\_Model2\_PPO\_415744\_72\_tests

PPO\_x5\_v0\_rpf09\_Model0\_PPO\_462848\_70\_tests – an interesting thing

PPO\_x5\_v0\_rpf09\_Model1\_PPO\_464896\_66\_tests

The best results on average of 3 cases. But started to succeed very lately. Could be needed more time to test more (?)

1. Изображение выглядит как текст, снимок экрана, Шрифт, число

   Автоматически созданное описание

The case with the same controlValues as the exp with 91 tests, but for sqrt.

PPO\_sqrt\_v1\_rpf09\_Model2\_PPO\_495616\_83\_tests – second best. Other two models have 59 and 0 max successful tests. For the ‘exp’ the best was for the model1, while this one is the best for the model2.