Project Milestone 3:

Full project report

Note that the PPO implementations here follow the convention from [Spinning Up](https://spinningup.openai.com/) rather than [Baselines](https://www.github.com/openai/baselines): they use the early stopping trick, omit observation and reward normalization, and do not use the clipped value loss, among other potential diffs. As a result, while it is easy to fairly compare this PPO to this TRPO, it is not the strongest PPO implementation (in the sense of sample efficiency) and can be improved on substantially.

Demo videos

moving pid video

original

ppo baseline fail

ppo spingup recreate

ppo motor pid crashed

ppo starter agent

ppo implement

Code with documentation

readme