

PPO: Performance with tuned hyperparameters and linear learning rate annealing

