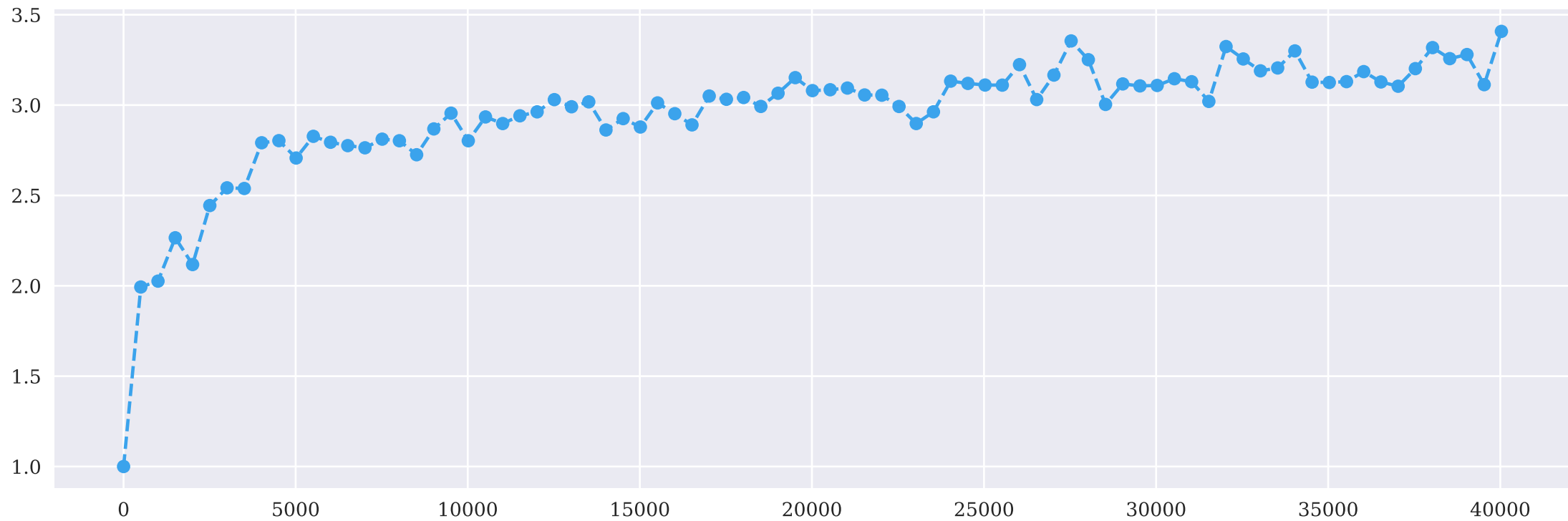


Reward trajectory

Average cumulative episode reward achieved



Episodes approximately trained in each of 23 parallel processes