CS 285 Homework 3

Neo Yu Yao Terence

# Experiment 1

\*\*\*\*\*\*\*\*\*\* Iteration 350000 \*\*\*\*\*\*\*\*\*\*\*\*

Training agent...

Beginning logging procedure...

Timestep 350001

mean reward (100 episodes) 97.767258

best mean reward 134.511581

running time 1416.802803

Train\_EnvstepsSoFar : 350001

Train\_AverageReturn : 97.76725769042969

Train\_BestReturn : 134.51158142089844

TimeSinceStart : 1416.8028028011322

Training Loss : 0.2648437023162842

Done logging...