

Supplementary Material: Shaping Model-Free
Reinforcement Learning with Model-Based
Pseudorewards

Omniscient Dyna (maze task, mountain car to come; figure captions to come)

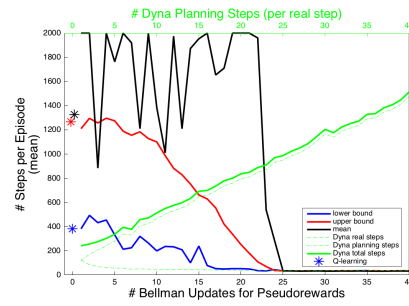
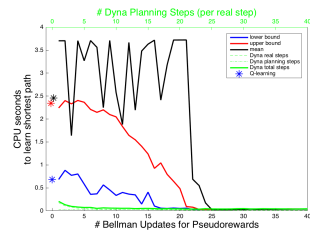
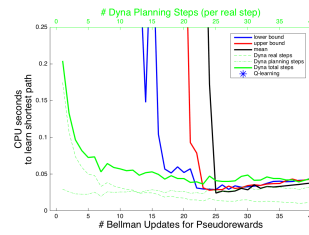


Figure 1: omniscient Dyna



(a) CPU time



(b) close-up

Figure 2: omniscient Dyna

Non-omniscient pseudoreward estimation (maze task, mountain car to come)

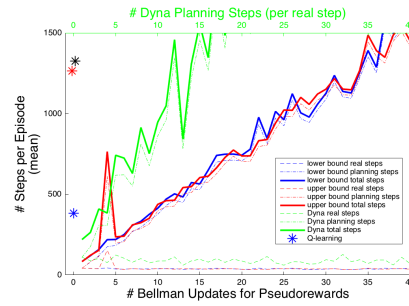
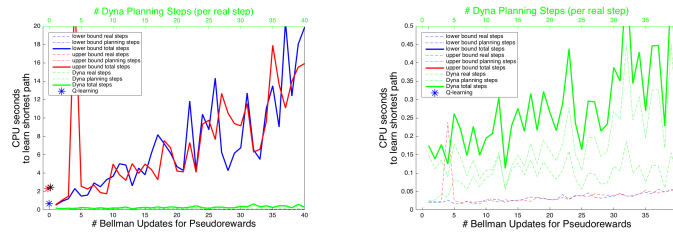


Figure 3: Non-omniscient pseudoreward estimation



(a) CPU time

(b) close-up

Figure 4: Non-omniscient pseudoreward estimation

Prioritized sweeping (maze task, mountain car to come)

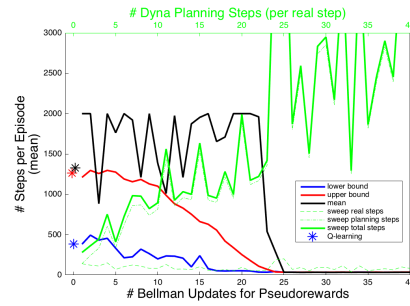


Figure 5: Prioritized sweeping

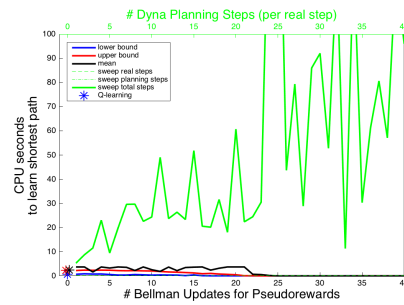


Figure 6: CPU time