A Distributed Implementation of Reinforcement Learning

Yi-Chun Chen and Yu-Sheng Chen

May 31, 2016

ICME

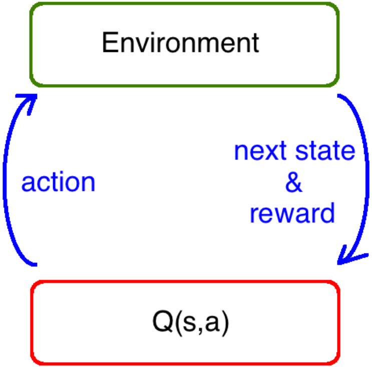


Figure: Vanilla Q-learning

Structure of Distributed Algorithm

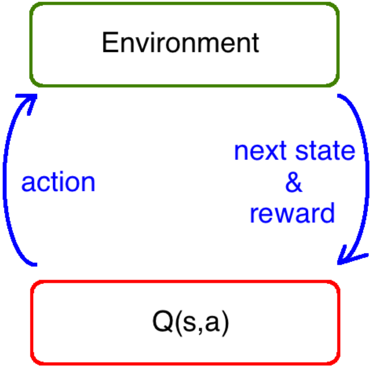


Figure: Vanilla Q-learning

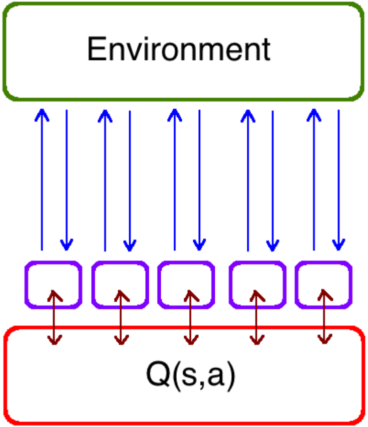


Figure: Distributed Q-learning

Analysis

Vanilla Q-learning v.s. Distributed Q-learning

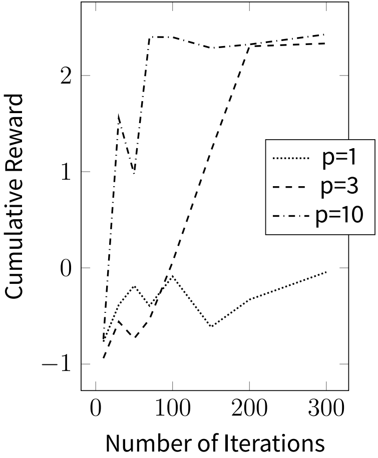
- All-to-one structure
- Samples from interacting with environment are distributed. ($N \to N/p$)
- Convergence to optimal policy can be accelerated. ($> p$ times faster)
  Since samples can cover the environment more uniformly.

Vanilla Q-learning v.s. Distributed Q-learning

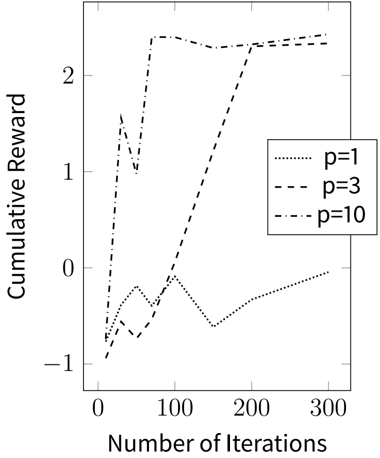
- All-to-one structure
- Samples from interacting with environment are distributed. ($N \to N/p$)
- Convergence to optimal policy can be accelerated. ($> p$ times faster)
  Since samples can cover the environment more uniformly.
- Per iteration,

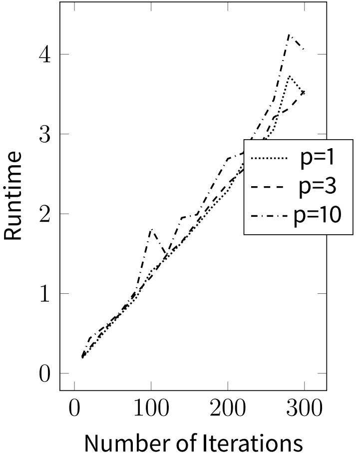
	Work	Depth
Make Decision (on each learner)	$O(\|A\|)$	$O(log(\|A\|))$
Update Q	$O(p)$	$O(log(\|p\|))$
Total Shuffle Size	$O(p\|A\|)$	

Experimental Results

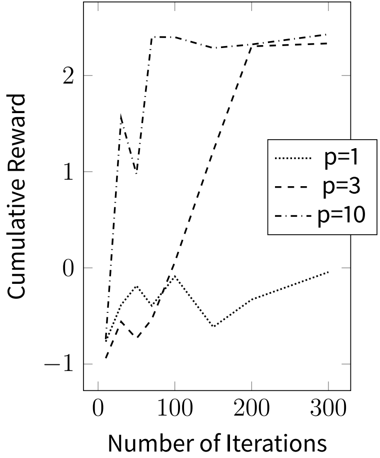


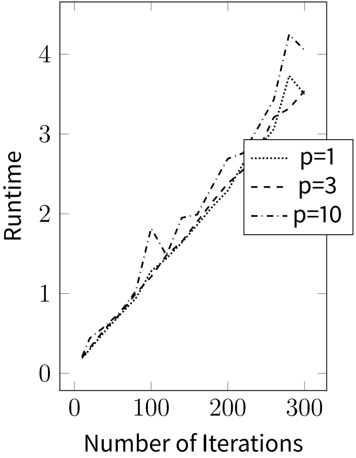
Experimental Results





Experimental Results





< p=10, iter=30 > is better than <p=1, iter=300 > by reward 0.73 > -0.058.