Difficulty of sample-average value estimation in nonstationary problems

Charles C Onu, 260663256

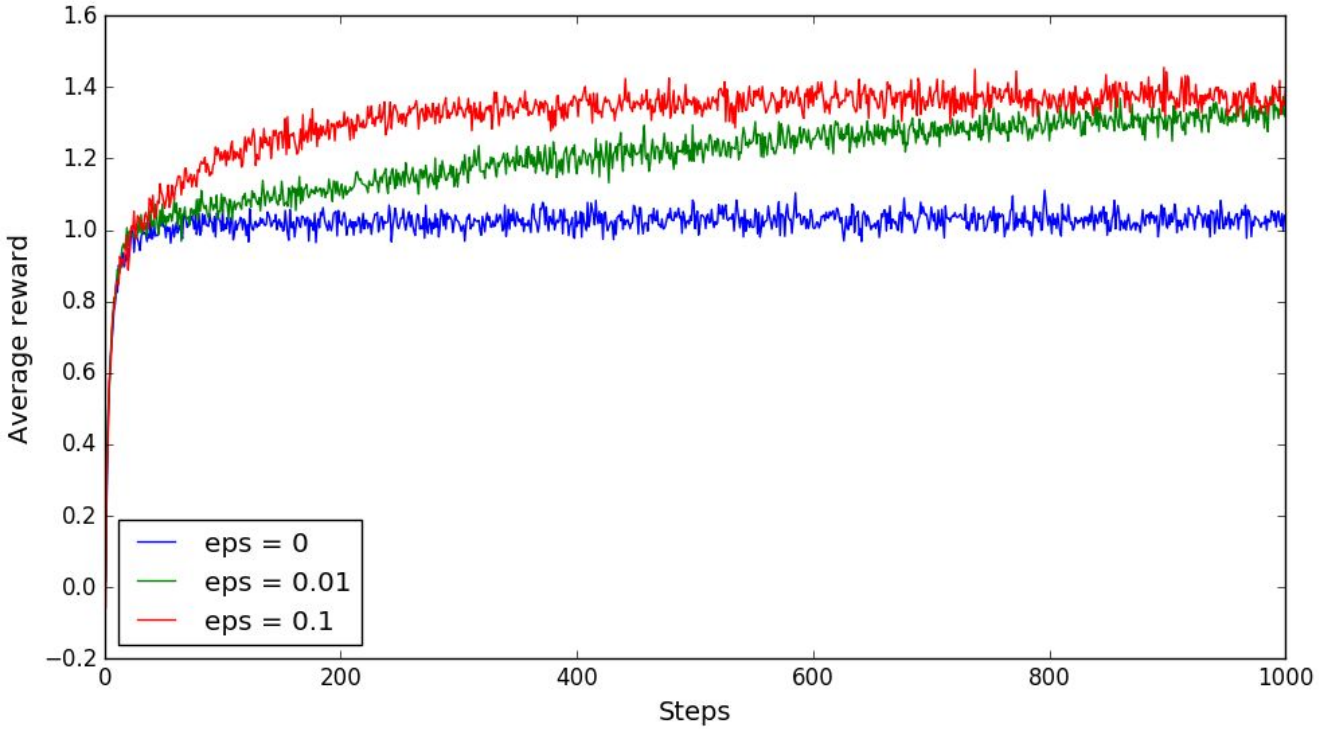
27 Jan, 2017

Stationary environment setup

		initialise	Get reward
1		$q_*(1) = N(0,1)$	$q_t(1) = N(q_*(1),1)$
2		$q_*(2) = N(0,1)$	$q_t(2) = N(q_*(2),1)$
. . .		. . .	. . .
10		$q_*(10) = N(0,1)$	$q_t(10) = N(q_*(10),1)$

Agent runs for 1000 time steps ie t = 1,...,1000

e-Greedy methods in stationary problem



NB: rewards are averaged across 2000 runs

Non stationary environment setup

initialise | **Get reward**

1

$q_*(1) = N(0,1)$ | $q_t(1) = N(q_*(1),1)$

2

$q_*(2) = N(0,1)$ | $q_t(2) = N(q_*(2),1)$

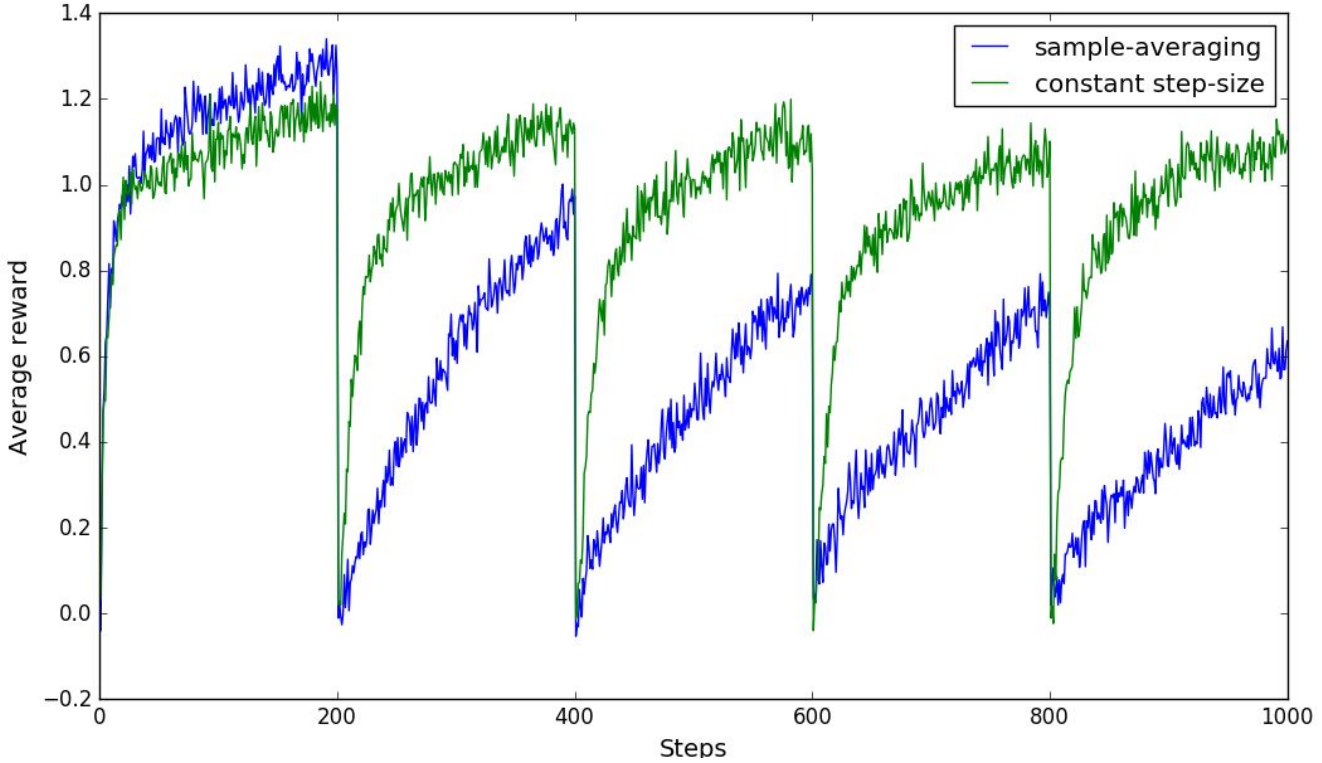
. . . | . . . | . . .

10

$q_*(10) = N(0,1)$ | $q_t(10) = N(q_*(10),1)$

Agent runs for 1000 time steps ie t = 1,...,1000

Sample-averaging vs constant step-size for non stationary problem



Code

https://github.com/rllabmcgill/rlcourse-january27-onucharles