



The University of Vermont

CS253A QR: Reinforcement Learning
Assignment №8: n-step SARSA

Ayat Ospanov

November 7, 2018

The experiment was held for $n = \{1, 2, 3, 4, 5, 6\}$. The results are given in Figure 1 and Figure 2. From the original graph it is barely visible which is better and which is not. But we can see that there exponential curves, e.g. $n=2$. So lets approximate this lines with polynomials of 2nd degree. Doing this we get Figure 2. It is now easily distinguishable which n is best for Pacman. At least for the scope we have, i.e. for 10000 iterations.

Having $n=\{3, 4\}$ has high constant convergence rate, while $n=\{5, 6\}$ is slightly worse. $n=1$ is the worst case and converges very slow. For $n=2$ we can see exponential convergence (it is a bowl shape because we approximated it with the polynomial, not exponential function) and in long term it will converge the fastest. So $n = 2$ is the best.

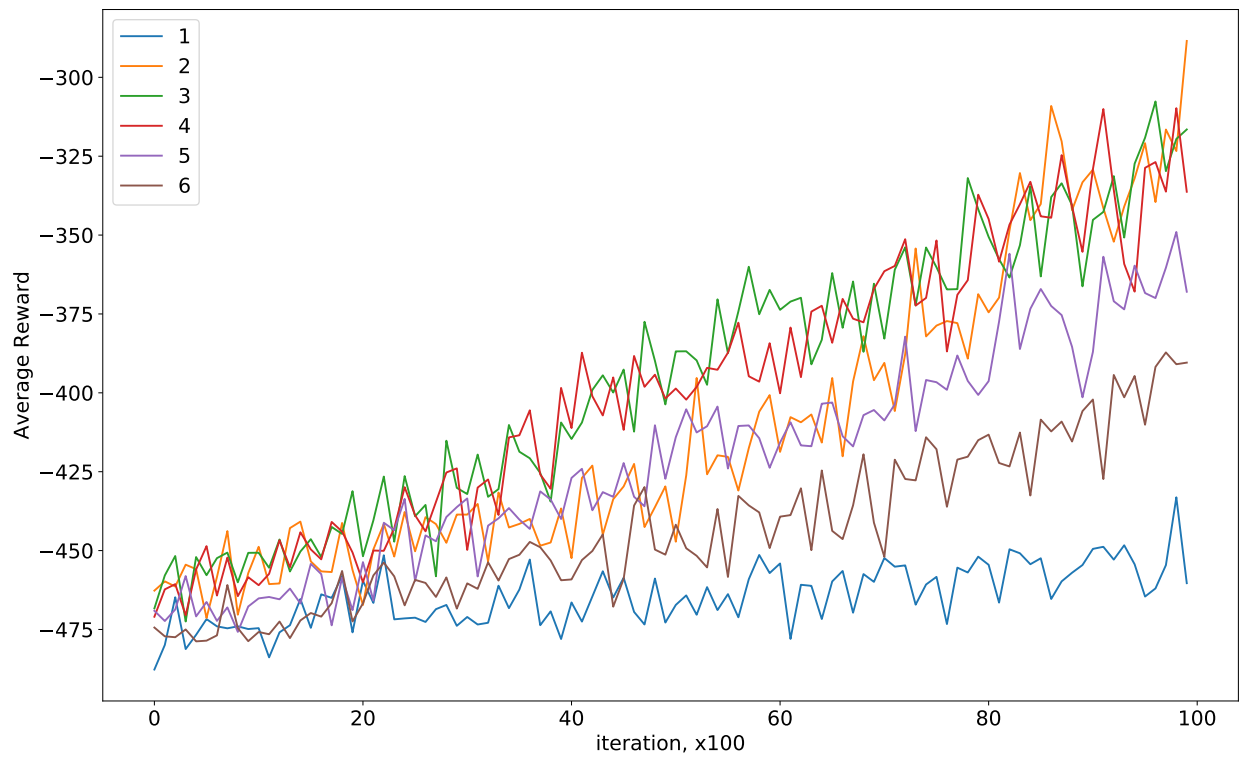


Figure 1: Average reward for 10000 iterations over 10 runs

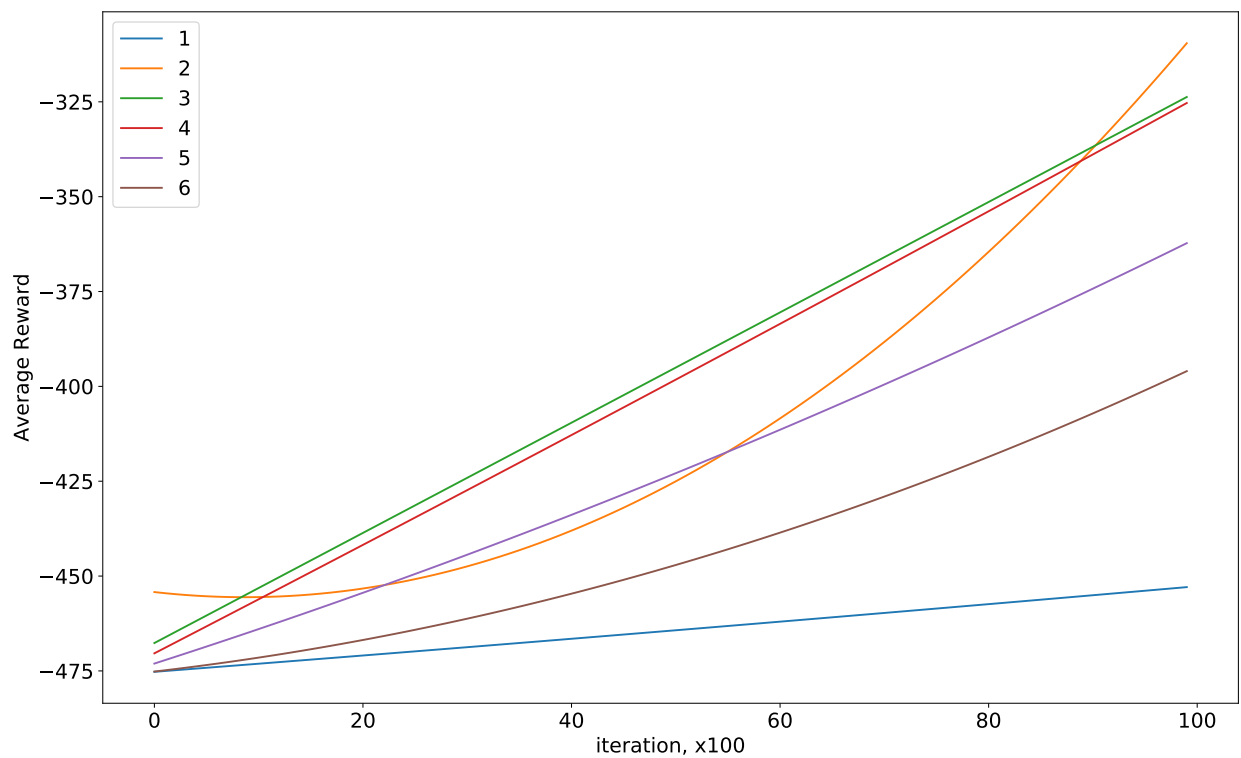


Figure 2: Smoothed average reward for 10000 iterations over 10 runs