Multi-step RL: Unifying Algorithm

Kirill Bobyrev

October 28, 2017

Overview

Important results

Algorithm description

Initialize $S_0 \neq \textit{terminal}$
Select A_0 according to $\pi(.|S_0)$
Store S_0, A_0, $Q(S_0, A_0)$
for $t = 0, \ldots, T + n - 1$ **do**
  if $t < T$ **then**
    Take Action A_t, observe R and store S_{t+1}
  end if
end for

Synopsis

- $Q(\sigma)$ unifies Sarsa and Tree-backup
- $Q(\sigma)|_{\sigma=0}$ is Tree-backup
- $Q(\sigma)|_{\sigma=1}$ is Sarsa

References

Kristopher De Asis, J. Fernando Hernandez-Garcia, G. Zacharias Holland, Richard S. Sutton.
Multi-step Reinforcement Learning: A Unifying Algorithm.
arXiv, 3 Mar 2017.

The End