

Multi-step RL: Unifying Algorithm

Kirill Bobyrev

November 5, 2017

Plan

- 1 Introduction
- 2 From MC and one-step TD to multi-step Bootstrapping
- 3 Algorithm
- 4 Conclusion

Results

Monte Carlo methods

One-step TD methods

n -step TD methods

Overview

Algorithm description

Initialize $S_0 \neq \text{terminal}$

Select A_0 according to $\pi(.|S_0)$

Store $S_0, A_0, Q(S_0, A_0)$

for $t = 0, \dots, T + n - 1$ **do**

if $t < T$ **then**

 Take Action A_t , observe R and store S_{t+1}

end if

end for

Intuition and Examples

Choosing σ

Synopsis

- n -step methods are derived from both MC and TD(λ)
- $Q(\sigma)$ unifies n -step Sarsa and Tree-backup
- $Q(\sigma)|_{\sigma=0}$ is Tree-backup
- $Q(\sigma)|_{\sigma=1}$ is n -step 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.



Richard S. Sutton, Andrew G. Barto.

Reinforcement Learning: An Introduction.
MIT Press, Cambridge, MA, 19 Jun 2017 Draft.

The End