Reinformcement Learning Theory Learning

Yue Wang

October 30, 2016

Abstract

This is a simple template for an article written in LaTeX.

1 Introduction

RL theory is very long history and these days

- Bullet point one
- Bullet point two

1. Numbered list item one
2. Numbered list item two

2 Notation and Formatting

some notations:

$\mathcal{S}$	the state space
S_t	the state at time t, stochastic
s_t	the state at time t, actual
s	the state, actual
R_t	the reward at time t, stochastic
r_t	the reward at time t, actual
r	the reward, actual

3 RL algorithms

4 RL convergence theory

4.1 counterexample

In [1, chap 11.3] the authors give an intuitive conclusion about when these algorithms will divergence :

> the danger of instability and divergence arises whenever we combine three things:

1. training on a distribution of trainsition other than that naturally generated by the process whose expectation is being estimated(e.g. off-policy learning)
2. scalable function approximation (e.g. li)

4.1.1 counterexample1

In [2]

References

[1] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction," 2011.

[2] D. P. Bertsekas, "A counterexample to temporal differences learning," *Neural Computation*, vol. 7, no. 2, pp. 270–279, 1995.