

Reinforcement Learning results analysis

Marco Marini

June 1, 2019

Contents

1	TDQAgent	1
----------	-----------------	----------

Abstract

Reinforcement Learning results analysis.

Measure

One hundred tests were run a maze environment with two different λ values.

Each test consist of 1000 episodes with a limit of 300 steps per episode.

1 TDQAgent

The returns of TDQAgent slightly increase from about -565 to -555 while the optimal value should be about -4.

The agent was not able to learn the optimal strategy due to probable the impossibility to find the path within limit of 300 steps.

The errors increase from 6 to about 3000. The cause of such increment may be due to bootstrap effect. No optimal and stable policy was found and so the action values decrease incrementally without a lower bound limit. The changes of network parameters do not fit such limits createing incremental errors episode by episodes.

No significant differences of learning rate can be noted changing the hyper parameter λ .

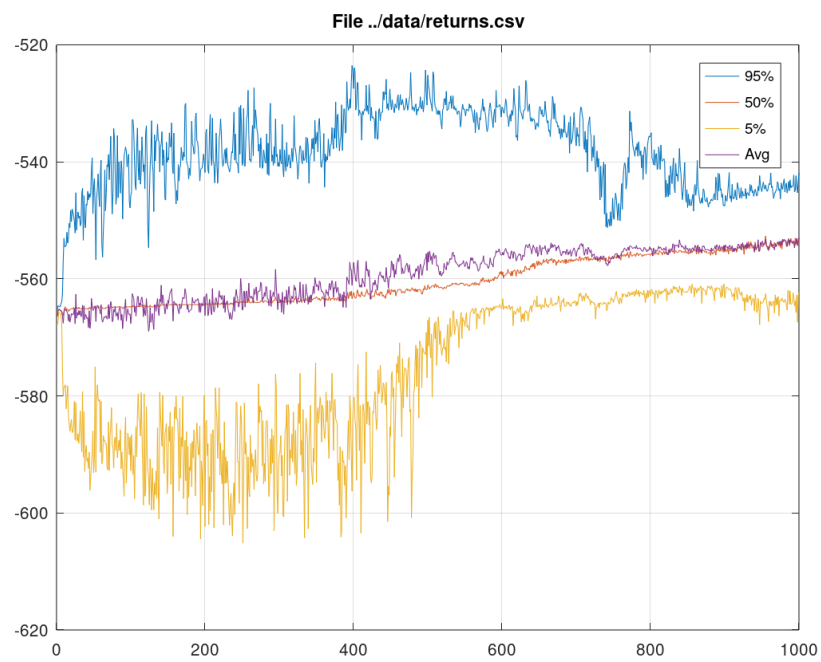


Figure 1: Returns $\lambda = 0$.

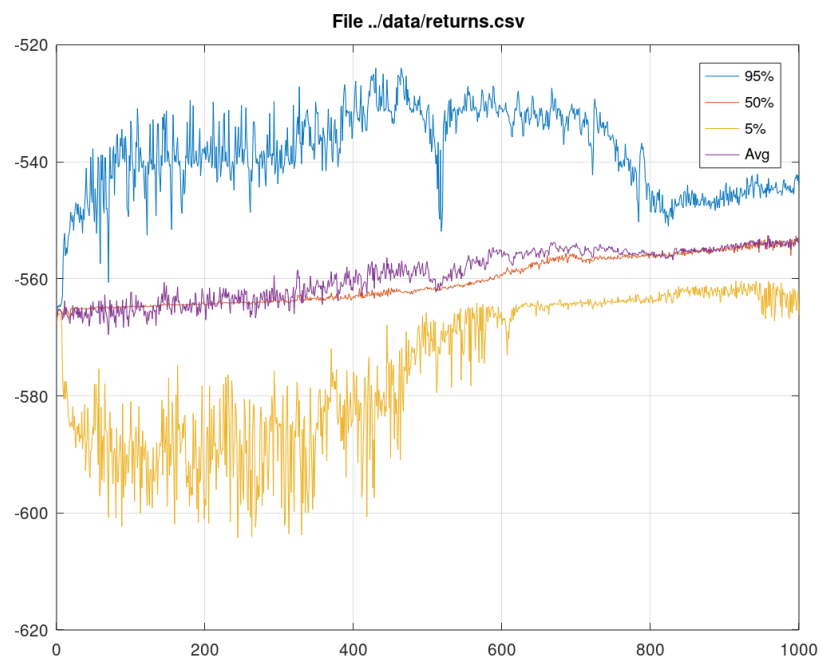


Figure 2: Returns $\lambda = 0.7$.

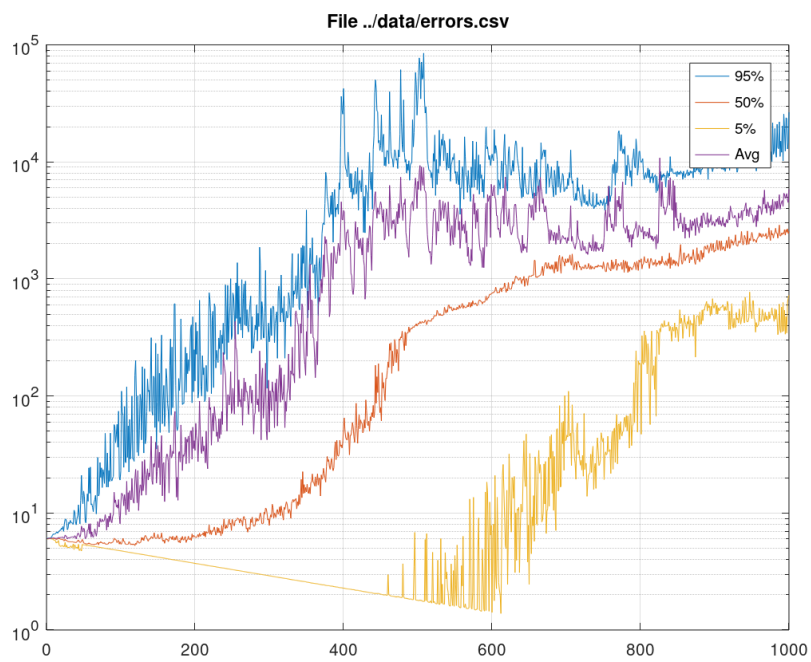


Figure 3: Error $\lambda = 0$.

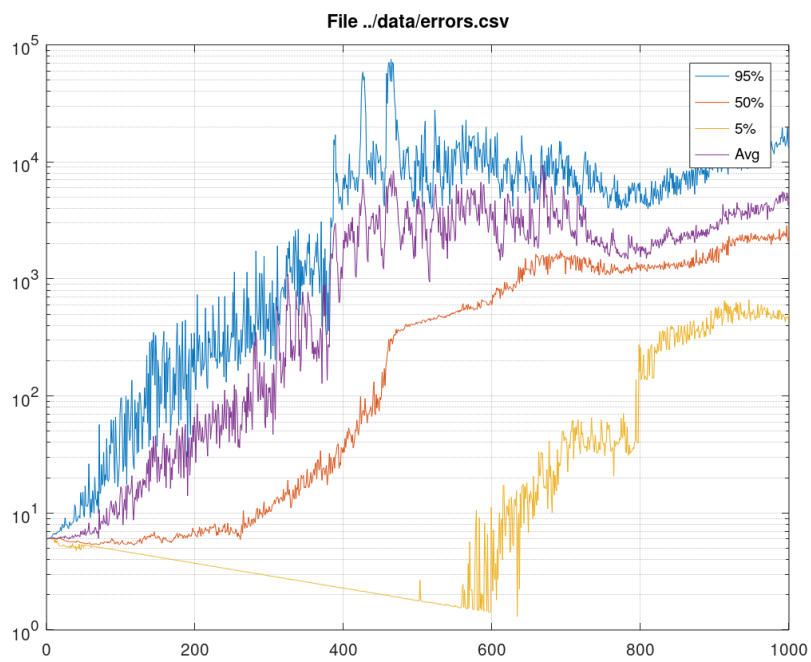


Figure 4: Errorss with $\lambda = 0.7$.