

# PR1: Reinforcement Learning - CS 7642

GID: mmendiola3

Presentation: <http://youtube.com/magahet/blah>

## 1. Overview

We will review and replicate the experiments presented in Sutton's 1988 paper, "Learning to Predict by the Methods of Temporal Differences". In the process, we will highlight required assumptions for parameters not explicitly stated, and compare and contrast the methodology and results of Sutton's work to this one. Finally, we will review the challenges faced in recreating Sutton's original experiments and how these were handled.

## Random Walk

### Batch TD( $\lambda$ )

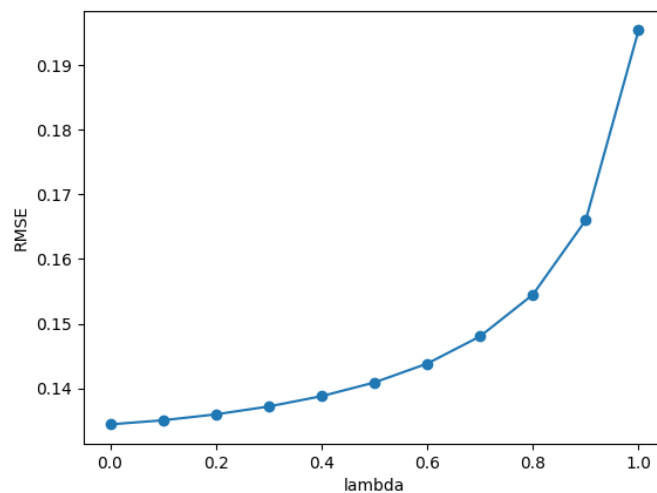


Figure 1: fig3

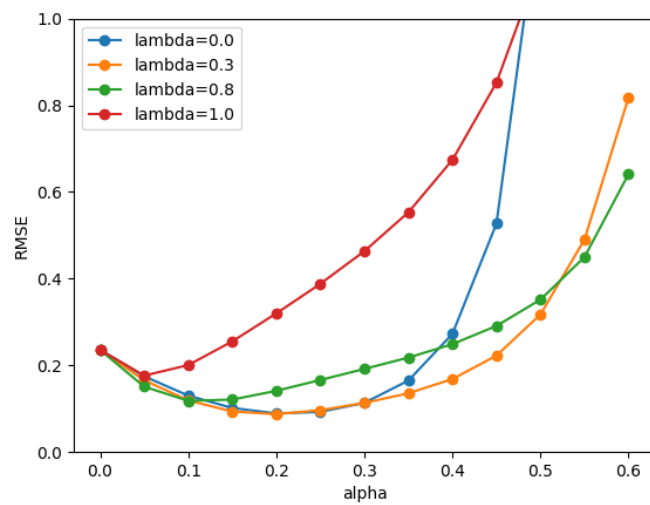


Figure 2: fig4

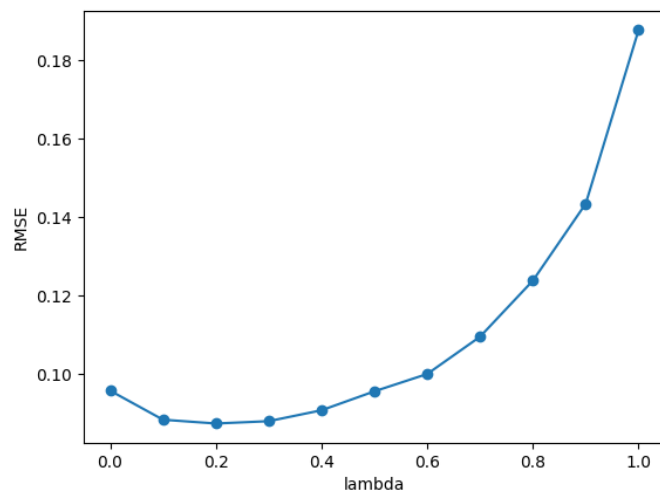


Figure 3: fig5

## Online TD( $\lambda$ )

- description of the experiment replicated
- how the experiment was implemented
- outcome of the experiment
- how well the results match the results given in the paper
- significant differences
- pitfalls you ran into while trying to replicate the experiment from the paper (e.g. unclear parameters, contradictory descriptions of the procedure to follow, results that differ wildly from the published results)
- steps to overcome those pitfalls?
- assumptions made. Why are these assumptions justified?

Figure 1 shows blah.

Observations:

- The median value is fairly close between states; between 3% and 4% for each state.
- OH has the lowest median at ~3%.
- IN has the highest median at ~4%.kkkkkkkkkkkkkkkk