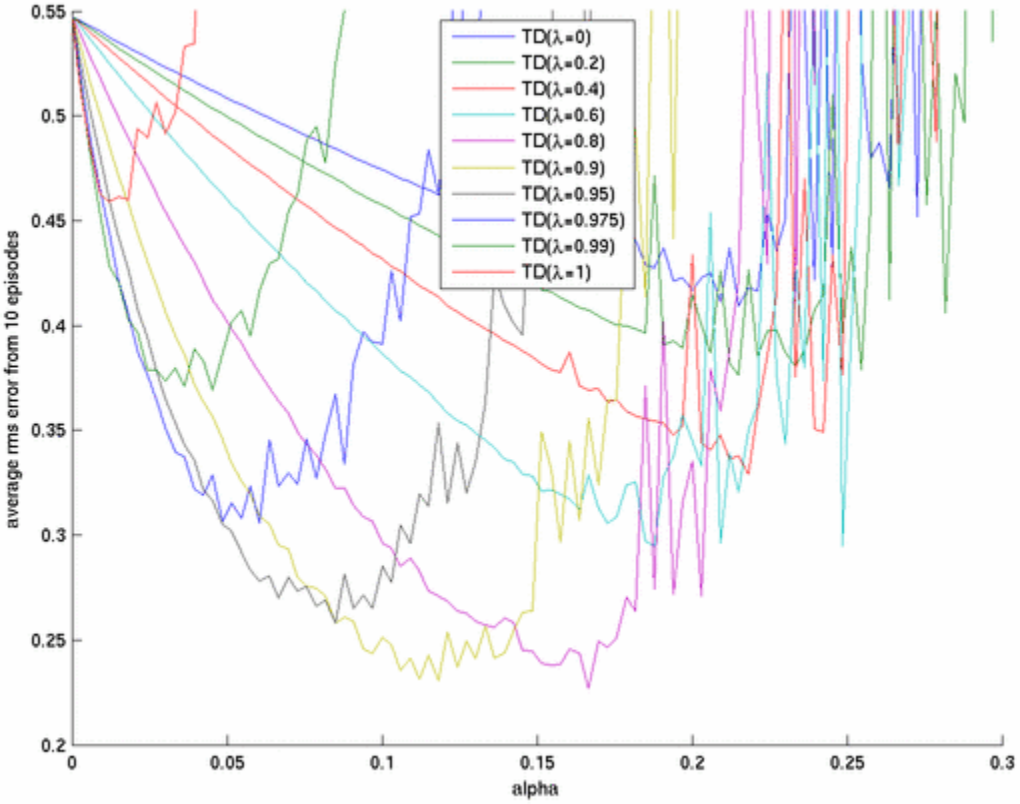
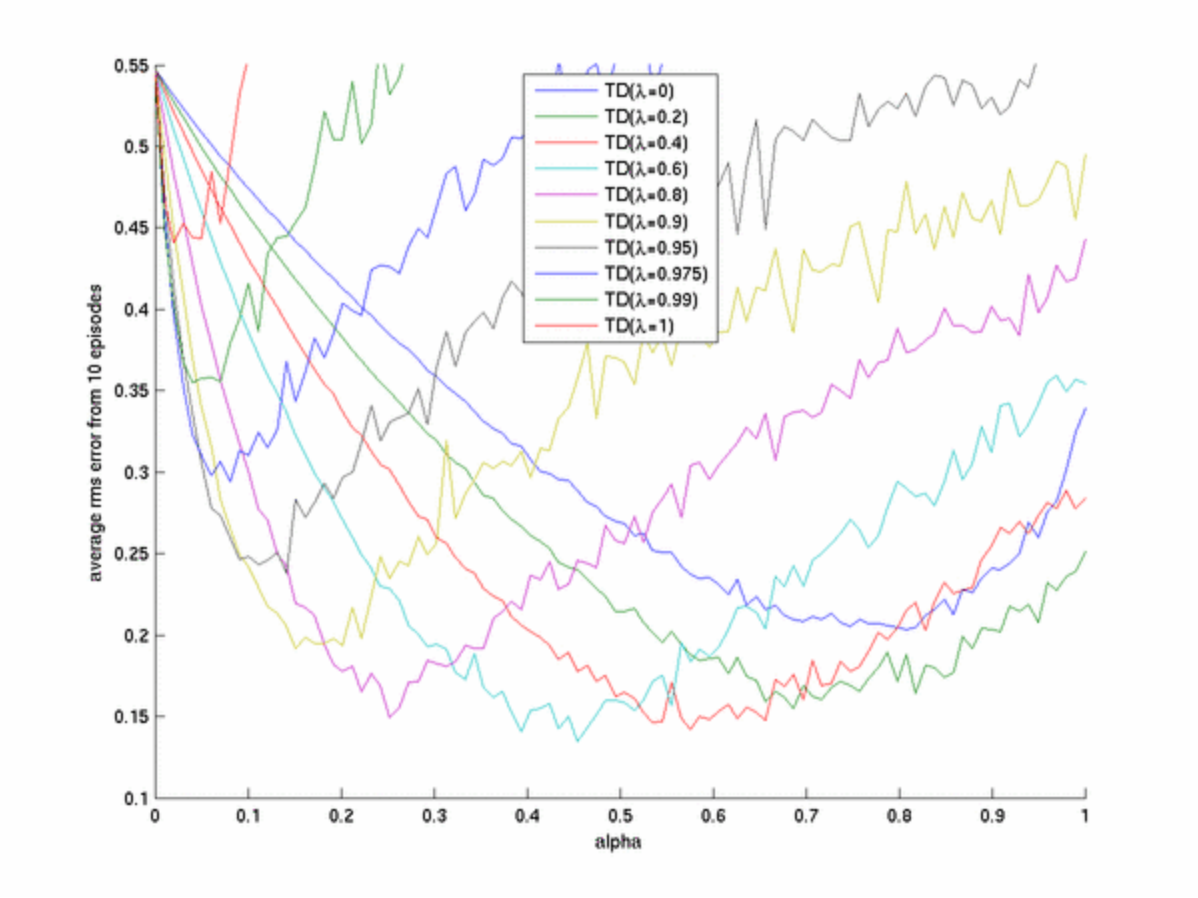
Here you will find experiments and results obtained when performing TD(lambda) learning on the random walk examples from this chapter. I ran these experiments for one hundred alphas uniformly spaced between the limits presented on the plots the book supplied. Since the book discusses the *offline* version of TD(lambda) we present these results first.



This result looks quite similar to that presented in the book. While the book does not talk about this in this section, the offline algorithm can be easily turned into an online algorithm by performing the update of the state value function immediatly after each state change rather than accumulating the updates during the eposode and applying them all at once. This version is coded up in the Matlab file `rw_online_tdl_learn.m` and the results from running this presented below.



I believe that some very long eposodes contributed to the osscilations seen in the learning curves. These osscilations would be smoothed out by taking more experiments/eposodes in these experiments.

John Weatherwax
Last modified: Sun May 15 08:46:34 EDT 2005