|  |
| --- |
| CS 8803 |
| Midterm Project |
| Replicating published results |

|  |
| --- |
| Jacob Kilver  2-21-2016 |

# Introduction

Reproducing results published in scientific journals is an important part of taking part in the research community. This project attempts to reproduce the results seen in Figures 3, 4, and 5 from Sutton (1988).

# Description of experiments

Figures 3-5 from Sutton (1988) deal with the bounded random walk problem. Figure 1 shows the state diagram for this Markov Decision Process (MDP). Beginning in the center state D, each state has a single action, which 50% of the time results in a transition to the left and 50% of the time results in a transition to the right. Thus, states A and G are terminal states. The reward function is zero everywhere except for the transition from state F to G, which has a reward of 1.0.

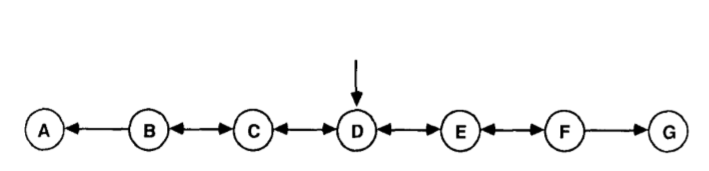


Figure : Bounded Random Walk State diagram

The TD Lambda algorithm was used to find the expected value of states B through F in this MDP. Two experiments were conducted, each with trainings sets of size 100, with each training set containing 10 sequences (10 walks through the MDP). The first experiment dealt with the “repeated presentation” paradigm – each training set was presented to the algorithm repeatedly until the state values converged. The second experiment dealt with how the learning rate effects error when each training set is presented only once.

# Implementation details

BURLAP (<http://burlap.cs.brown.edu/>) was used to replicate these results. The TD Lambda algorithm is readily available in this software package, along with tools to implement MDPs represented as graphs. However, constructing static training sets proved difficult and were not able to be used for these experiments, which likely affected the results presented below.

# Results

# Comparison to published results

# Conclusion