Assignment 2 – CS4300

A\* Search in a Wumpus World

Braden Scothern and Kyle Heaton

# Introduction

Summary of problem, basic approach, describe experimental framework  
Clearly state questions answered by the work

* What is the is mean number of search nodes produced by A\* for options 1 and 2?
* Is option 1 10% better than option 2?

The rand function in Matlab is touted to sample the uniform probability distribution. It uses a pseudo-random number generation algorithm, and the issues to be addressed are:

• Is the mean of a set of samples 0.5?

• Is the variance of a set of sample 1/12?

# Method

Description of algorithms  
Specific features of method, parameters in play, what data sets are used

Are the new children nodes added to the tree immediately when a node is expanded from the frontier?  
When is a node checked if it is a solution?  
Produce results even with no solution to goal (include stats of search trees that fail)  
Other issues you deem important to understand the data

The method used here is simply to generate a large number of samples and compute the mean, variance and confidence of the result. An alternative would be to run a large number of trials where each trial would get a fixed number of samples from rand, and then compute the mean and variance of each trial, and then compute the mean and variance over all those trials. This latter approach was not implemented.

# Verification of Program

Show that your code works

Pick 3 example boards and show hand developed solutions for each of the two options and compare results to Matlab  
Determine min(1) and max tree sizes for each method and report if actual minima and maxima from the experiments are in range

The method was applied to 1 sample and the results are found to be:

• the mean is the one sample

• the variance is 0.

>> a = rand

a = 0.8431

>> mean(a)

ans = 0.8431

>> var(a)

ans = 0

>>

Also, a large number of samples was generated and each was checked to make sure it was in the interval [0,1].

>> find(samps1)

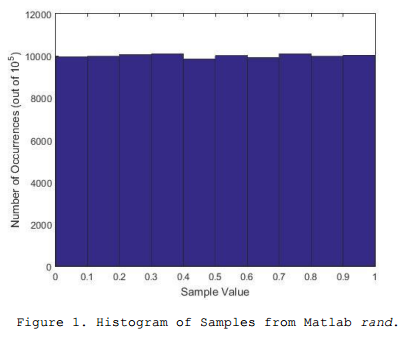
ans = Empty matrix: 0-by-1

# Data and Analysis

Description of data and how it is used/organized for training and testing  
Results from experiments (graphs, tables, means, variance, confidence intervals)  
Describe relationships found or qualitative description of results  
Discuss implications of statistical results

Give 4x4 tables with the number of nodes in the search tree for each of these options when the gold is at (x.y)  
Plot actual size of individual trial results and a histogram

Figure 1 shows the data collected from Matlab rand.



# Interpretation

Answer questions posed in into, use analysis to support conclusions  
Discuss future work and extensions

The results produced by rand are very close to the theoretical values of the mean and variance and the confidence interval at the 95% level is very short. Thus, the mean and variance results are close enough to the theoretical values to be acceptable in most applications.

# Critique

Demonstrate concepts understood, generate result of experiment with respect to raised questions and context  
Specific issues with technique (parameters, constraints)  
Problems and improvements

A better way to set up this experiment would be to pose a (null) hypothesis and then apply standard statistical methods to reject the hypothesis (or not). In addition, the number of samples could be varied from a small number to some very large number to examine the change in mean and variance as a function of number of samples; in theory, due to the Law of Large Numbers, this could show convergence properties.

# Log

A total of 27 minutes was spent performing the experiment in Matlab and writing the report.

# Appendix

List of own code, name, date, etc…