CS-415: Computational Biology: Project 1a

Taylor Martin
University of Idaho
Moscow, Idaho, USA
mart8517@vandals.uidaho.edu

1 INTRODUCTION

Currently in class, we have developed a simple genetic algorithm framework to:

(1) Generate individuals
(2) Give the randomly generated individuals a genome
(3) Measure the individuals fitness
(4) Mutate the individuals genome
(5) Insert individuals in populations which consist of numerous individuals.
(6) Pick the most fit individuals from the population for reproduction

Within this report, I will propose a couple experiments that I will be performing on this genetic algorithm framework and how I will be interpreting and visualizing the data gathered.

2 EXPERIMENT PROPOSITIONS:

In this report, I propose performing the following experiments on the Genetic Algorithm Framework:

(1) I will increase the rate of mutation of individuals and see how this effects populations of various size magnitudes. For example, an individual mutation rate of 20% with varying population sizes of 50, 100, 200, 500, etc. Observe if increasing the population of individuals with a higher mutation rate would lead to a majority of mutated individuals.
(2) The next experiment that I will perform will be changing the fitness calculation algorithm. I will construct a slightly more biologically accurate fitness algorithm that will calculate the number of triplet "T" occurrences within an individuals genome. Without replacement, so it will not count duplicate triplets. I will then see how this affects the average fitness of populations of various sizes.

3 PROPOSED DATA VISUALIZATION:

My proposed method of data visualization for the outcomes of these experiments is in the form of a scatter plot. I'm going to try and use either a graphing software like Desmos or a python Graphing package like matplotlib. An example of the scatterplot is listed below.
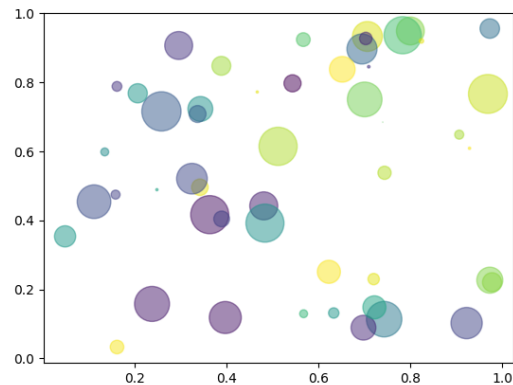

Figure 1: Example Scatter Plot Generated From Experiment Findings