**SYS 411 - Machine Learning**

**Genetic Algorithm Mini-Project**

**Problem Statement**

You are to write a genetic algorithm solution to find the maximum of the function f(x,y) = x- y where x and y are 16 bit unsigned integers.

X is 16 1s and Y is 16 0s

**The purpose of this problem is NOT to exercise your ability to encode a clever representation of the problem, nor to select a correct fitness function for this trivial problem, but to translate your general knowledge of GA’s into actual working code.** For that reason, the focus here is on writing and testing your GA code, and generalizing the solution as much as possible. You may wish to “hard-code” your first pass program and then generalize your code, or design it to be a general solution in the first pass.

You should build a “standard” GA including the following features:

1. User specified population size with random initialization of all individuals
2. User specified mutation rate and crossover rate to be applied to each new generation
3. User specified one-point or two-point crossover of chromosomes
4. Binary tournament selection (to start with) of the mating pool (pull two, put the most fit in the new generation, repeating until a full new gen is created. Binary tournament is a specific version of **n-ary tournament** which you will want in your final solution)
5. Full replacement of each generation (it is possible to create a new generation while still retaining some individuals from the old, but we will not include that capability)

Produce a series of runs of this GA, and discuss the behavior (in particular, # of generations to convergence) for the set of hyper-parameter values you selected to test. Your selection should include a set of different mutation rates and a set of different crossover rates. Keep track of how quickly your system converges to the correct answer with your selected hyper-parameter values.

**Homework Teams**

You may select a partner for this homework

**Submit**

Source code, ready to run

**Brief** written results of your experiment (comments at bottom of code rather than a separate doc)

**Extra Credit**

1) Provide a graphical interface that shows a convergence graph or other graphical representation of population dynamics with respect to fitness (adds up to 25% more credit)

2) Apply your GA solution to a more interesting problem where the encoding and fitness function are not as obvious (adds up to 25% more credit).

3) Run a “full factorial" experiment on the original problem. Produce output to a file to streamline the comparison of experimental parameter combinations. Be sure to apply a valid statistical approach. (adds up to 100% credit)