Genetic Programming:

Sensitivity to Parameters

By: Sarah Bailey

MATH 3220-01

December 15, 2010

Outline

- Genetic Programming
- Lil-GP System
- Ant and Regression Algorithms
- GP Parameters

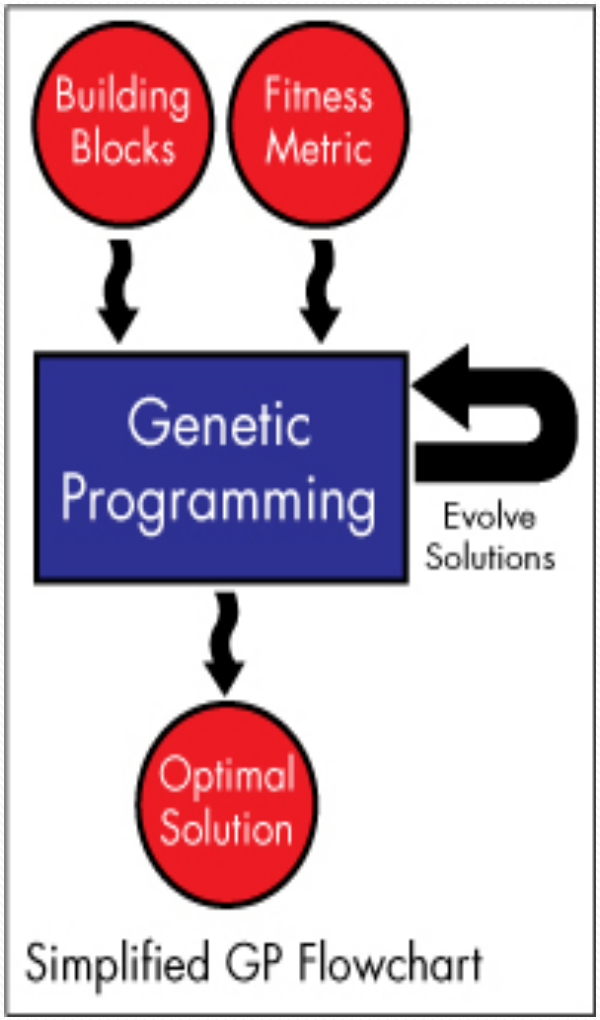
What is Genetic Programming?

- An **evolutionary algorithm**-based methodology inspired by **biological evolution** to find computer programs that perform a user-defined task.

http://en.wikipedia.org/wiki/Genetic_programming

GP Continued

- Challenge in Computer Science
  - Goal of GP: To achieve **automatic programming**
- Domain- Independent method
  - The user provides: building blocks & parameters
  - GP provides: optimal solution



http://geneticprogramming.us/What_is_Genetic_Programming>

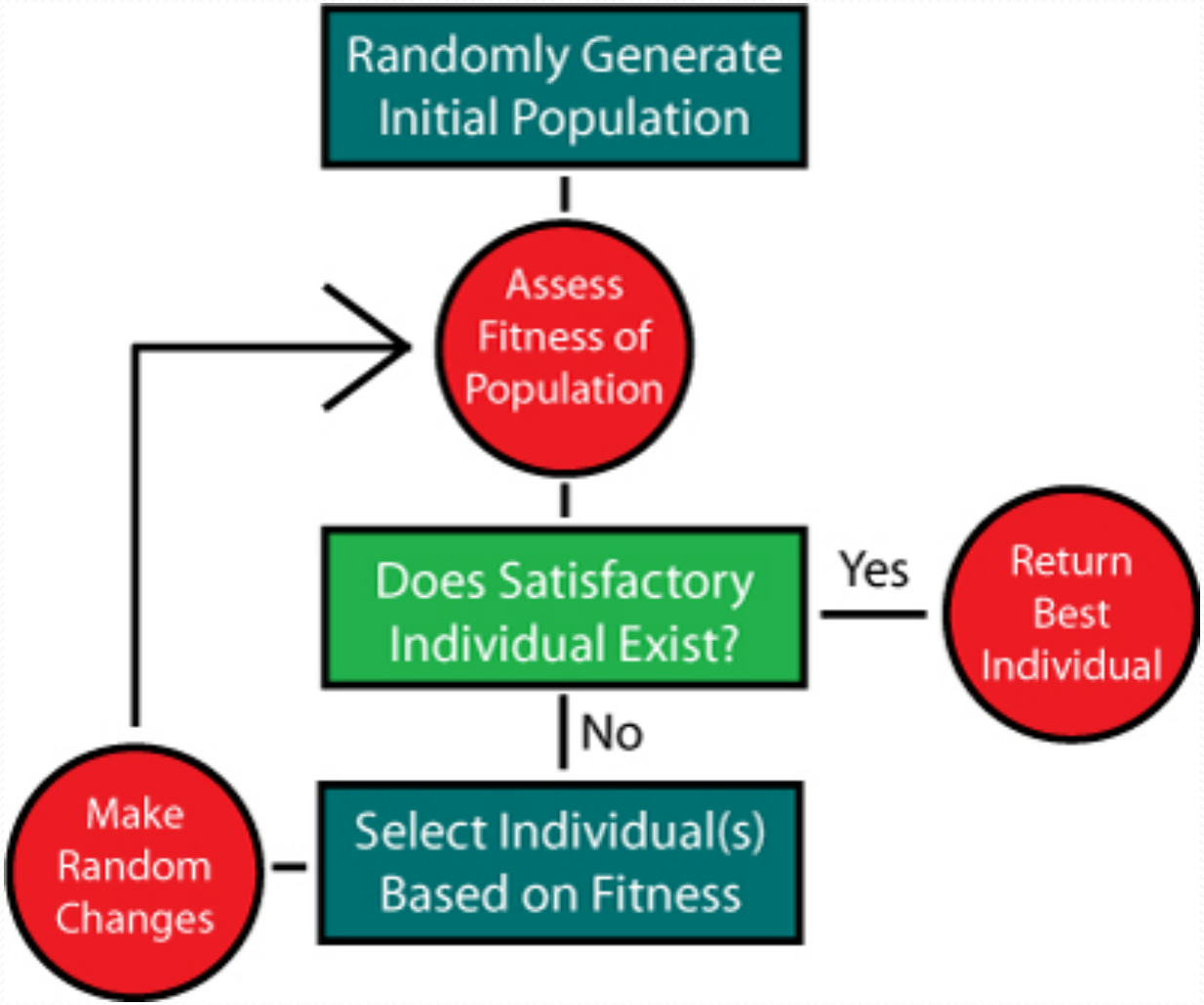
Preparatory Steps of GP

- The user needs to specify:
  - The set of terminals
  - The set primitive functions
  - The fitness measure
  - Parameters used in the run
  - Termination criterion and method for designating the results of the run

How GP Works

- Starts with a random initial population (generation 0).
- Execute each program and assign a fitness value.
- Select *n* number of individuals from the population
- Create new individuals by applying genetic operations
  - Reproduction, Crossover and Mutation.
- Keep the best-so-far solutions for future generations
- Repeat steps 2 – 5 until an optimal solution has been reached or until the algorithm has reached it's max iteration amount.

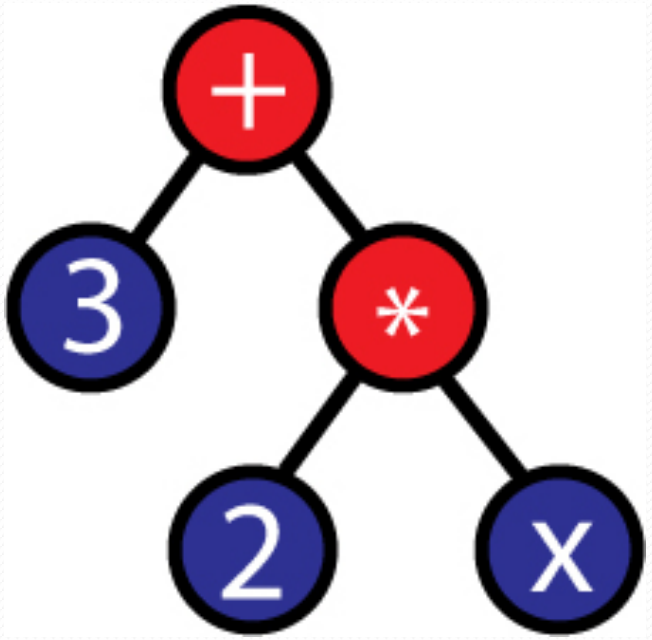
http://www.genetic-programming.com/gpanimatedtutorial.html



Picture: http://geneticprogramming.us/What_is_Genetic_Programming>

GP Representation

- Tree structure
  - Functional Nodes
    - Operators
  - Terminal Nodes
    - Operands
  - In-Order Traversal
    - Left. Root. Right



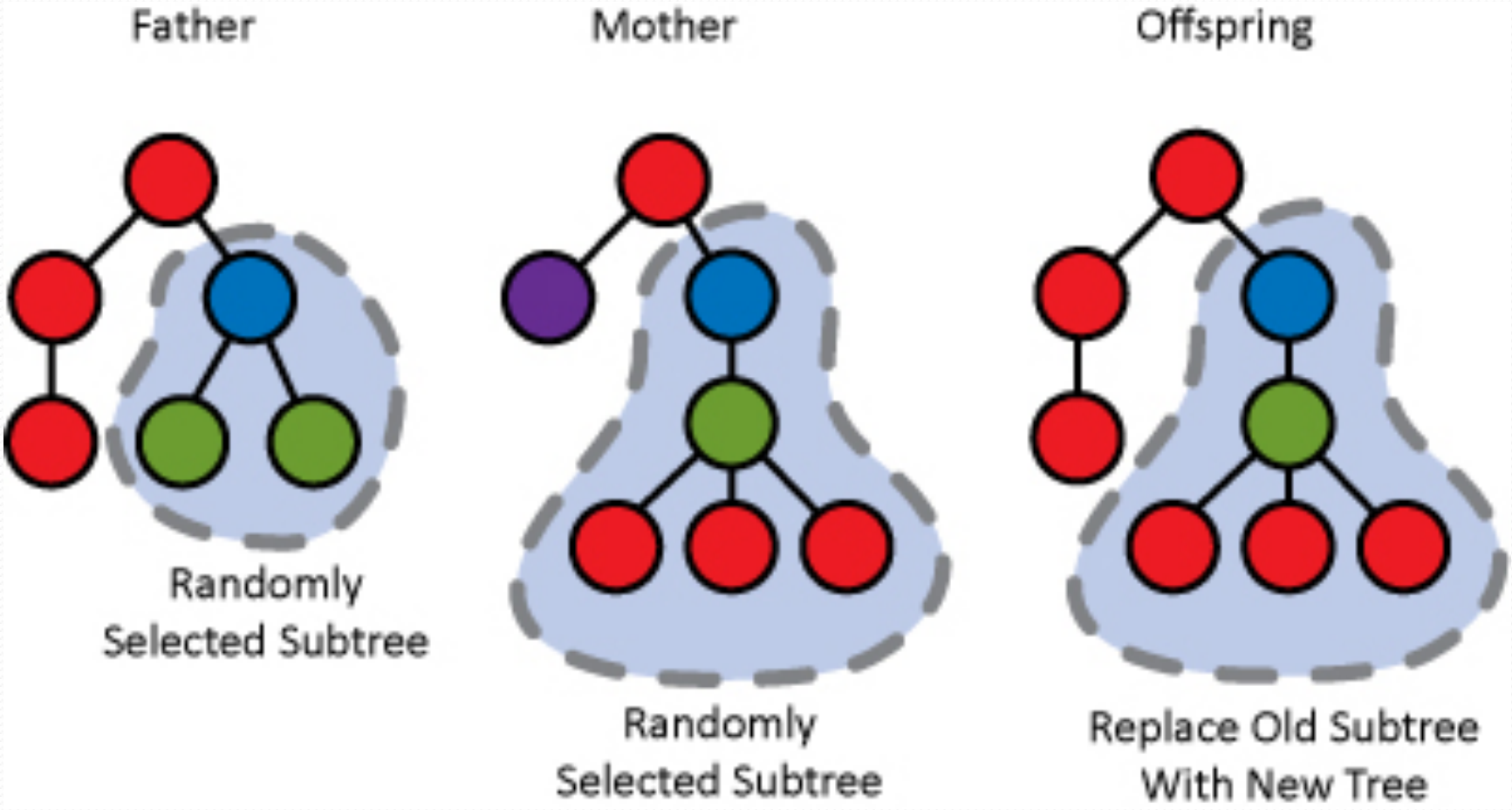
3 + 2 * x

Picture: http://geneticprogramming.us/What_is_Genetic_Programming>

Genetic Operations

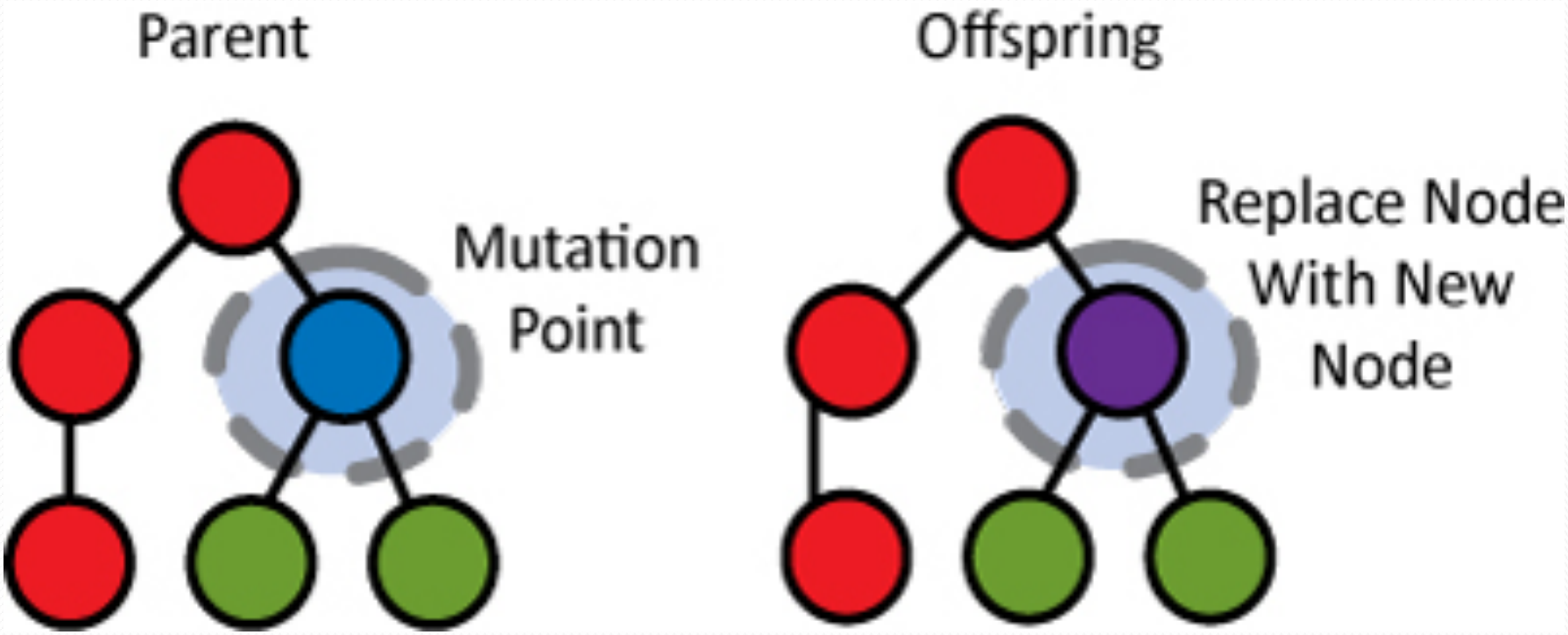
- Crossover
  - Combines genetic material from two parents by swapping a part of one parent with a part from the other
- Mutation
  - Operates on one individual and randomly chooses a point in it's tree and replaces it with an randomly generated subtree at that point
- Reproduction
  - Makes an exact copy of an individual and replaces it into the population

Crossover in Action



http://geneticprogramming.us/Genetic_Operations.html

Mutation In Action



http://geneticprogramming.us/Genetic_Operations.html

Lil-Gp: Genetic Programming System

- Written in C-language
- Three genetic Operators
- Supports random constants
- Output files are easily useable in programs like Excel
- Parameter files specify how individuals are selected
- Three Problems
  - Boolean 11- Multiplexer, Symbolic Regression, and Artificial Ant

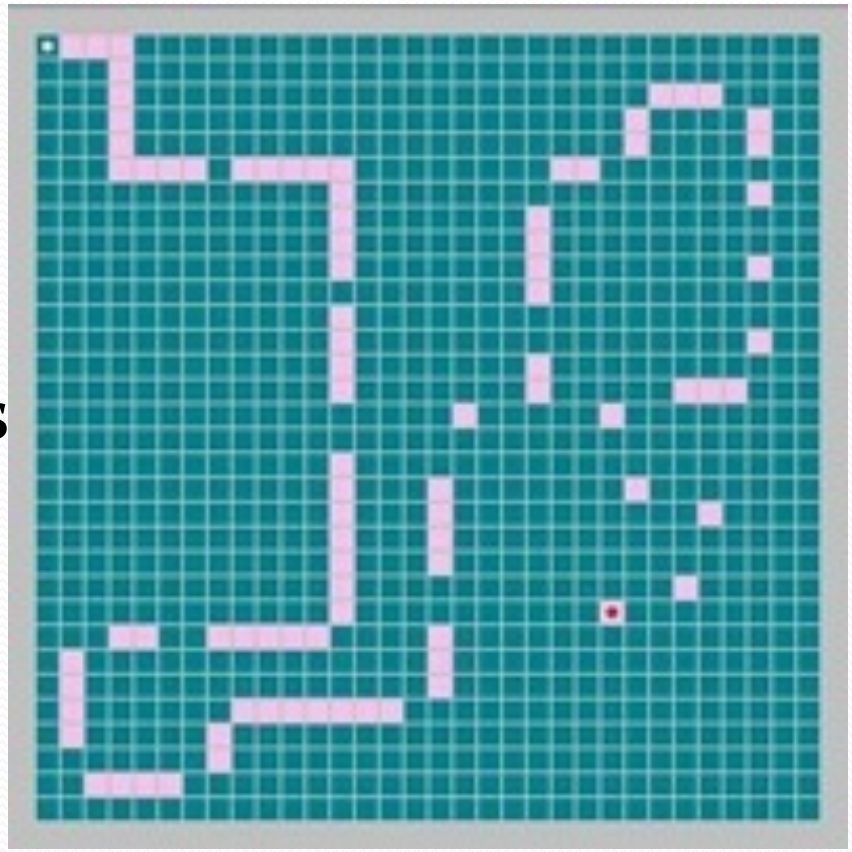
lil-gp 1.01 User's Manual

Artificial Ant Algorithm

- Represented by a grid provided with a trail of food pellets distributed over the grid.
- Santa Fe Trail (32 X 32 grid; 89 food pellets)
- The GP creates a path by walking through the map.
- It's goal is to eat as much food in the given amount of time

Santa Fe Ant Trail

- It will run for some *n* steps.
- Fitness will be measured by the number of food pellets the ant encounters
- Has 7 operators:
  - If-food-ahead; progn2; progn3; progn4
  - Move; left; right



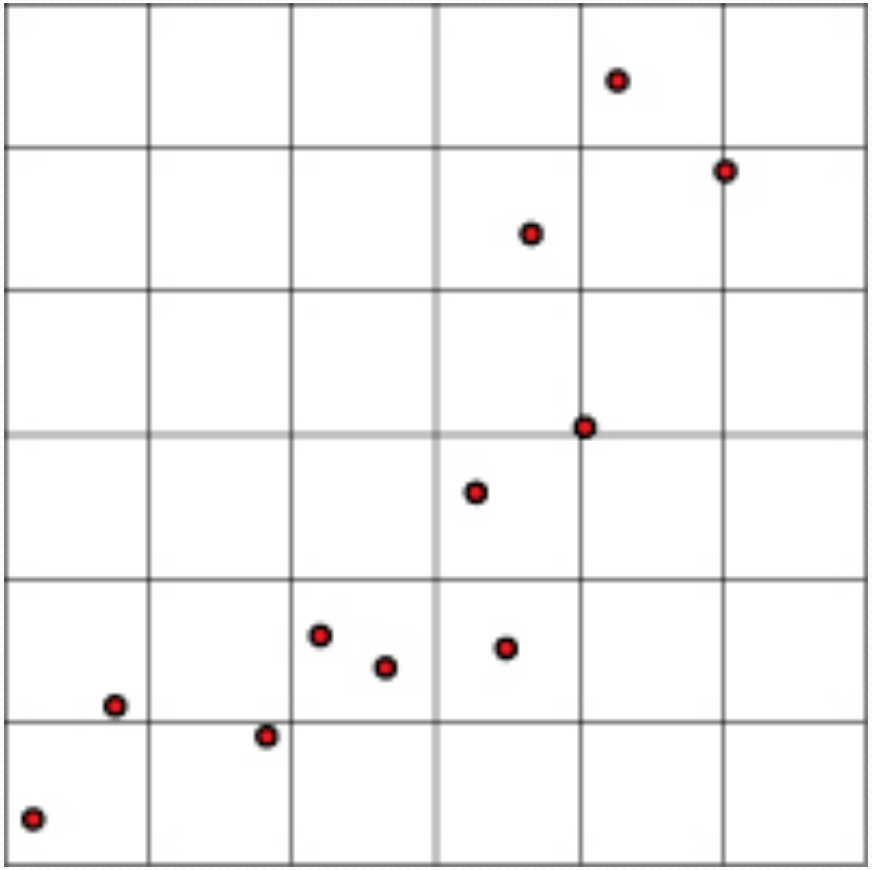
http://www.geneticprogramming.us/Three_Simple_Problems.html

Regression Algorithm

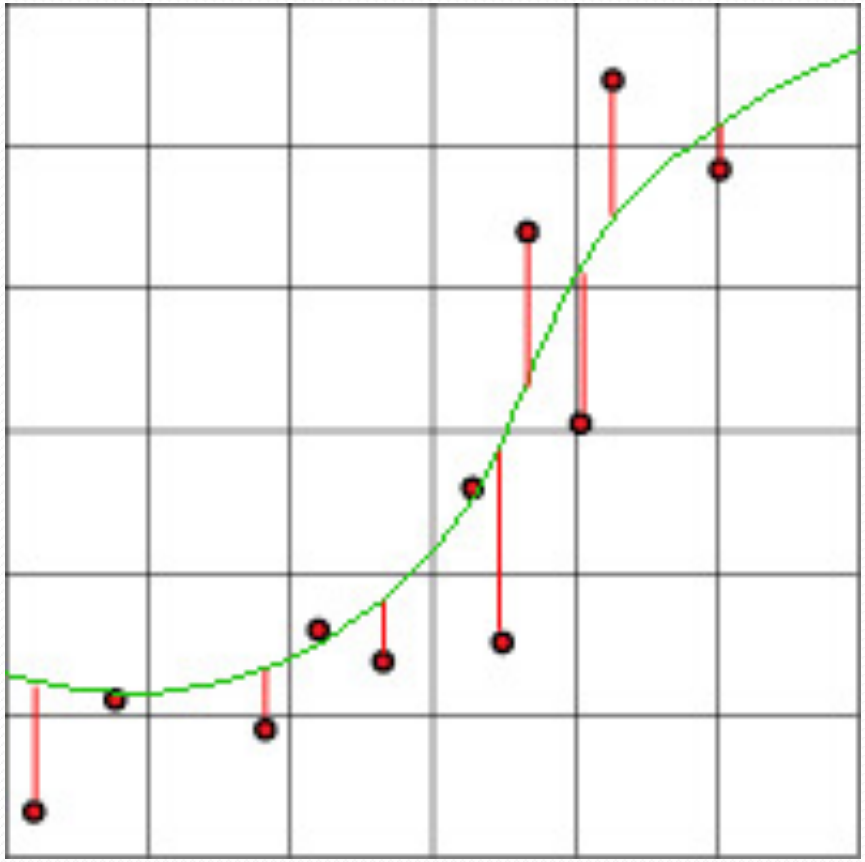
- The programs are designed to generate a function which fits a targeted curve.
- The user inputs a set of data points
- GP outputs a function that fits the data
- The terminal set consists of 1 -2 members:
  - The input value x, and (optionally) an ephemeral random constant
- The function set has 8 members:
  - Multiply, protected-divide, add, subtract, sine, cosine, exponentiation and protected-log

Regression Example

Given data points



GP results



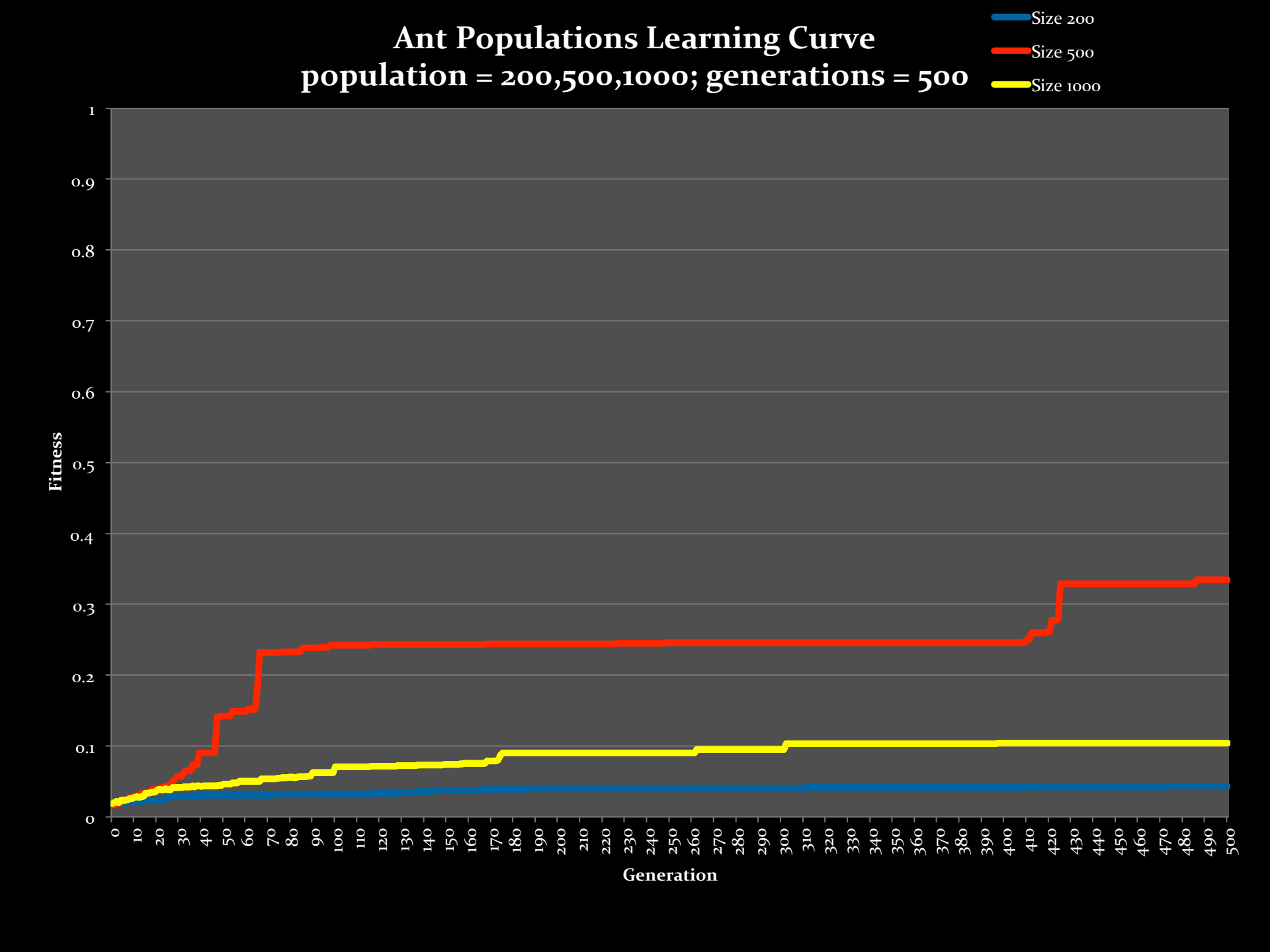
http://www.geneticprogramming.us/Fitness.html

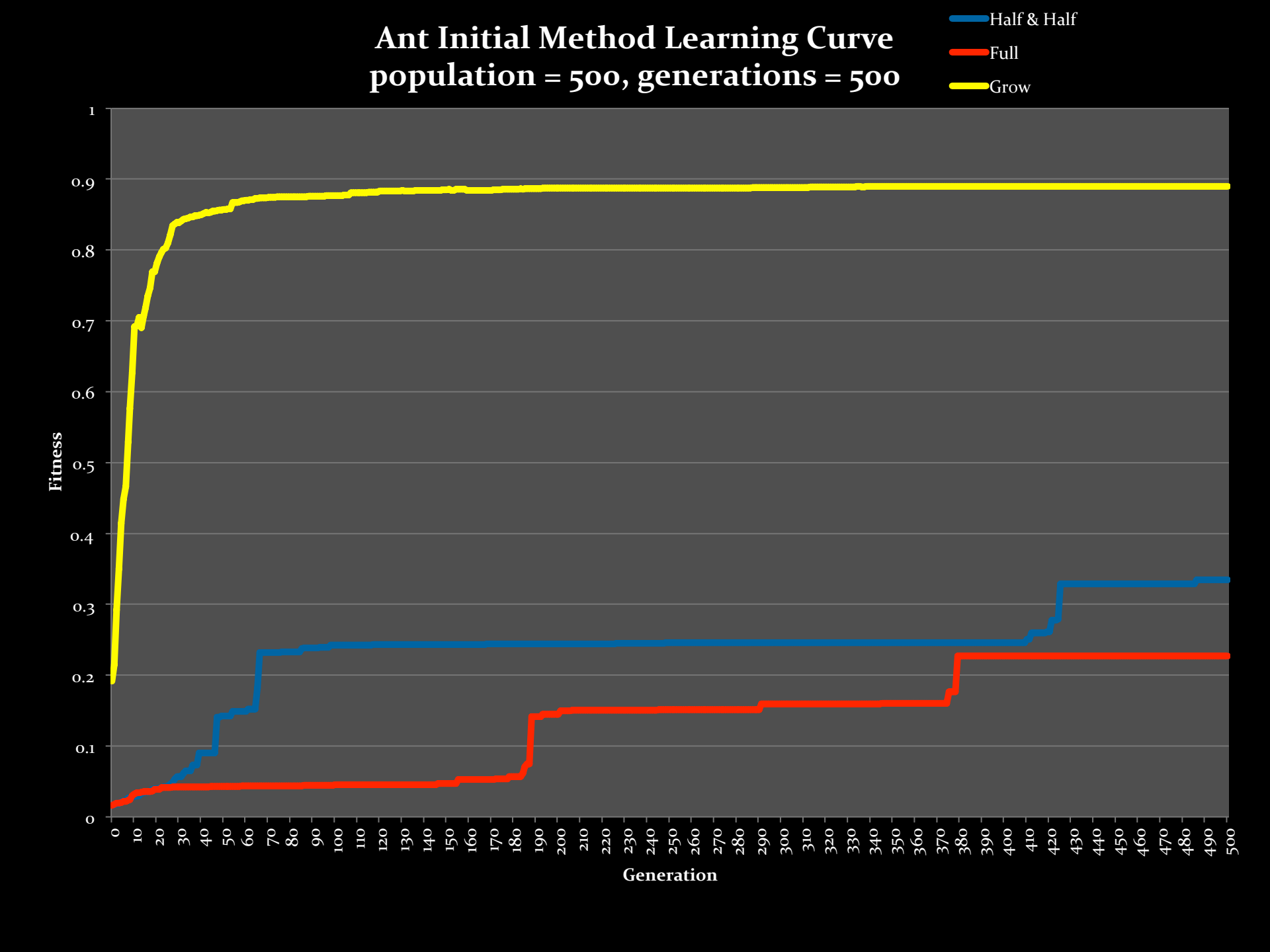
GP Parameters

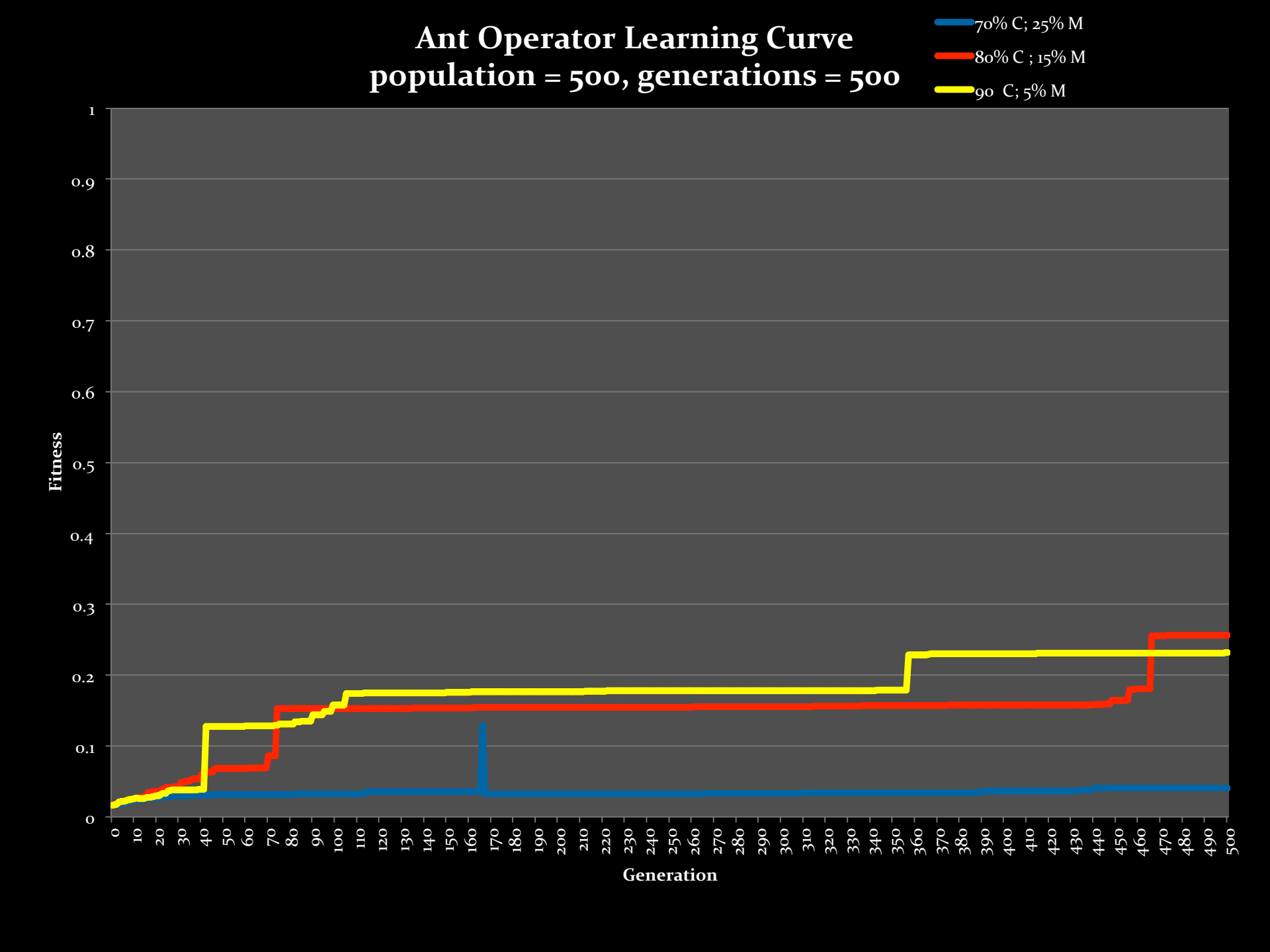
- Tree size limitation
  - Maximum number of nodes/depth of a tree
- Fitness
- Selection Techniques
  - Tournament (*)
  - Fitness Proportional
  - Ranking Selection

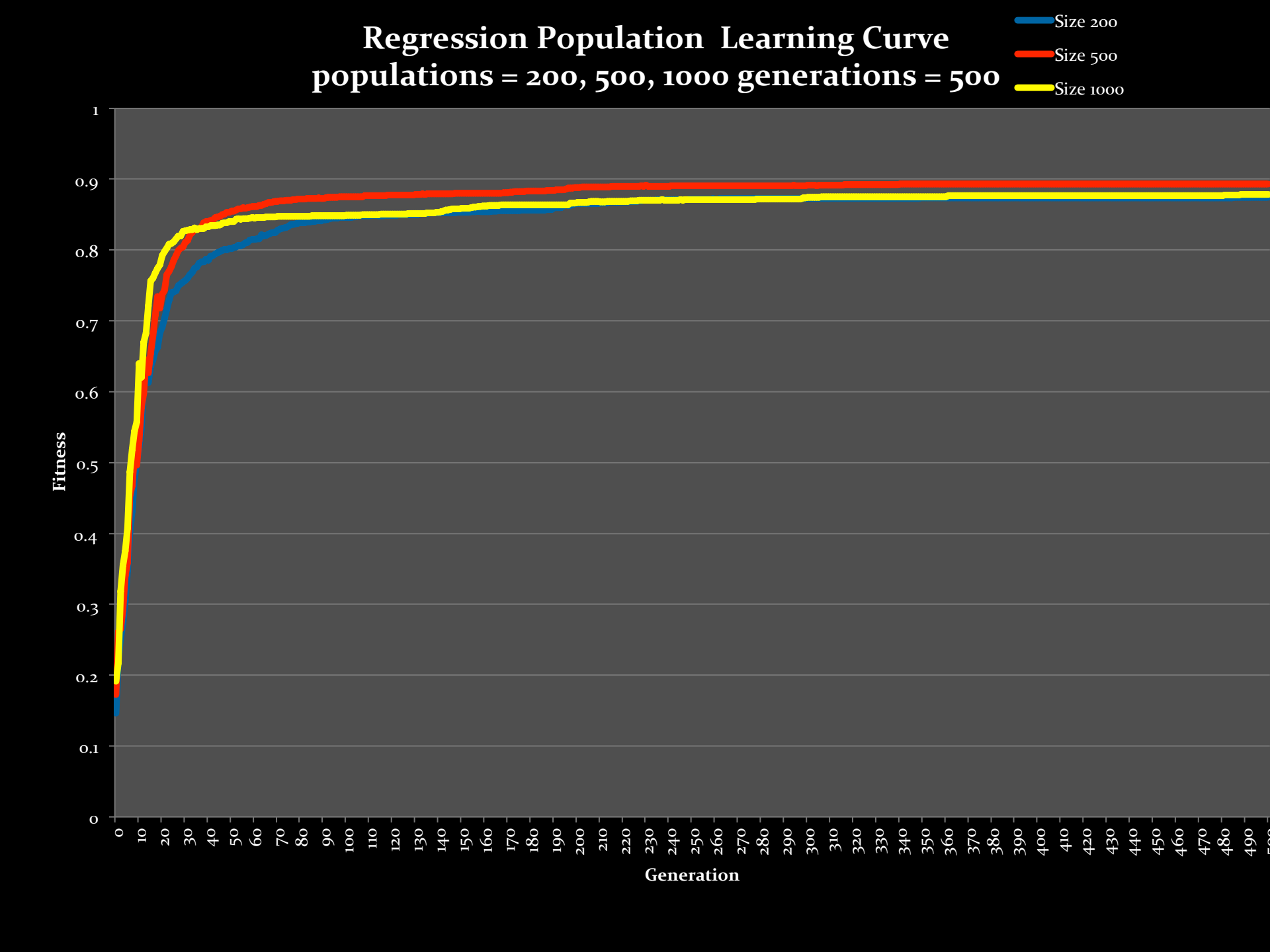
My Experiment: Tested Parameters

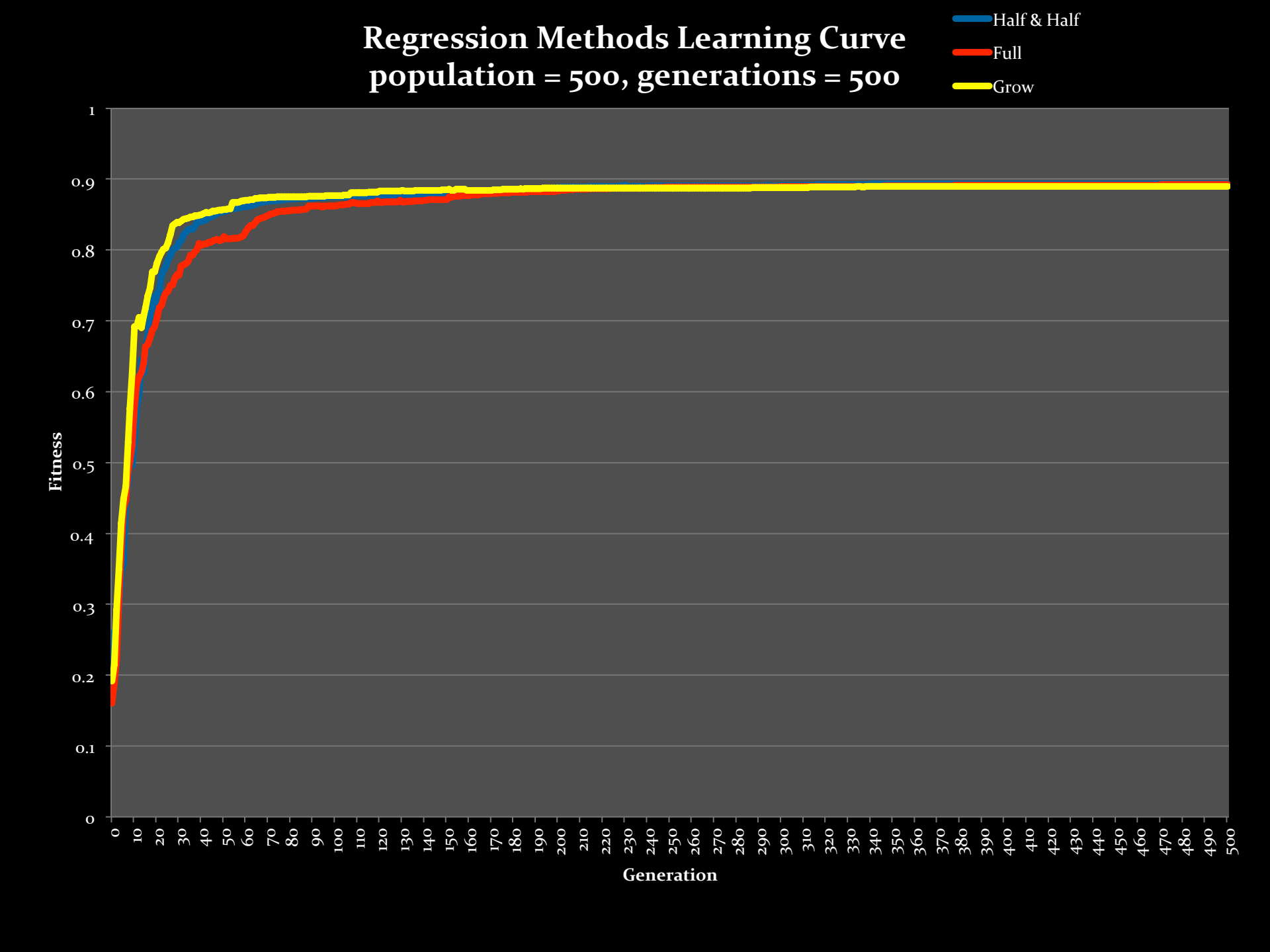
- Purpose of my experiment
  - No set guidelines for parameter selection
- Population size
  - Selecting a starting point for a number of individuals to be chosen from.
  - Population size = 200, 500, 1000
- Method of generating an initial population
  - Creating a variety of program structures for later evolution
  - Methods: Full, Grow, Half-and-Half
- Breeding Operators Probability
  - Evolution proceeds by transforming the initial population by the use of operators
  - 90% C- 5% M; 80% C- 15% M; 70% C-25% M

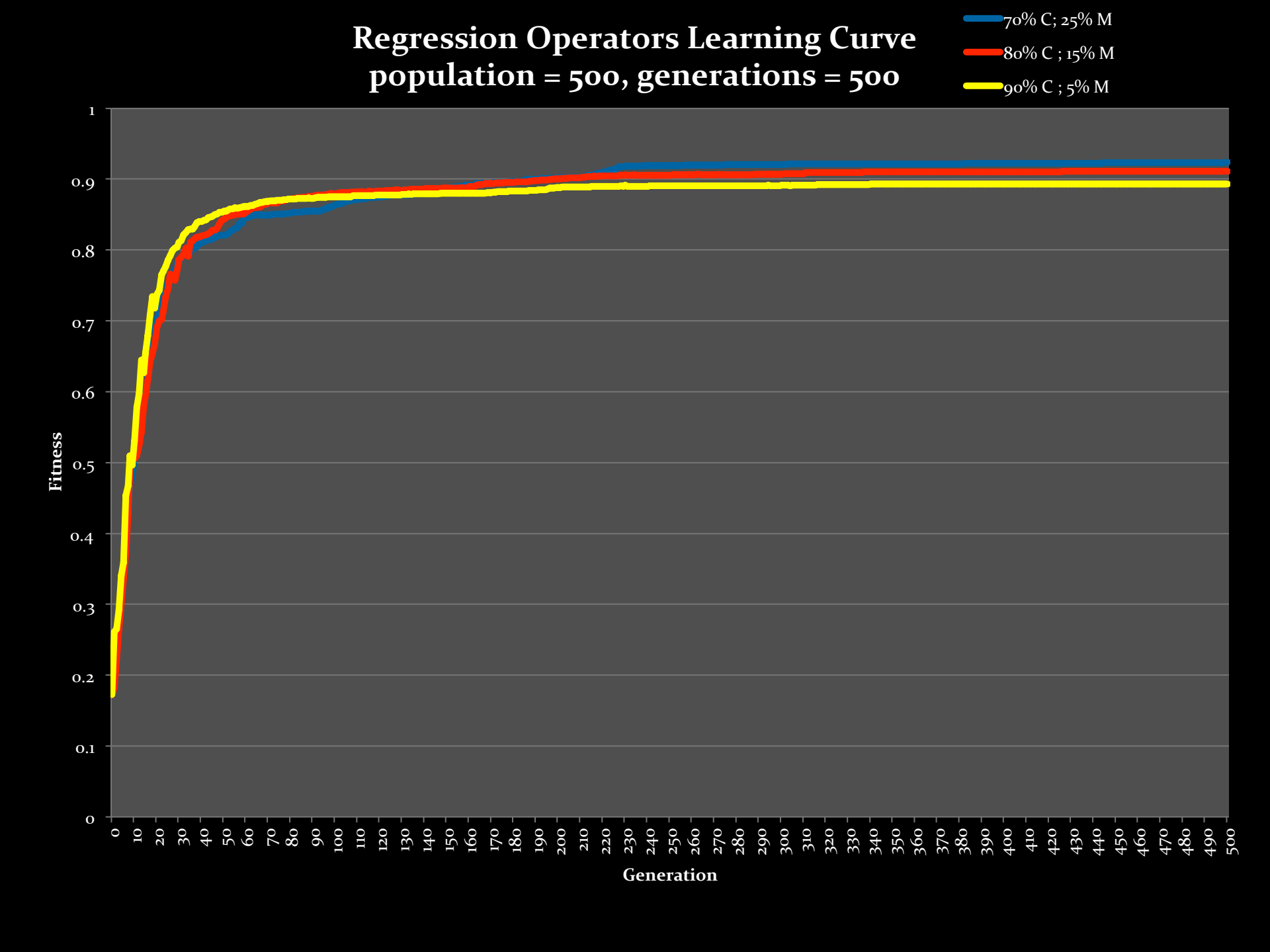

Ant Populations Learning Curve
population = 200,500,1000; generations = 500
Size 200
Size 500
Size 1000
Fitness
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500
Generation


Ant Initial Method Learning Curve
population = 500, generations = 500
Half & Half
Full
Grow
Fitness
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500
Generation


Ant Operator Learning Curve
population = 500, generations = 500
70% C; 25% M
80% C ; 15% M
90 C; 5% M
Fitness
Generation
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500


Regression Population Learning Curve
populations = 200, 500, 1000 generations = 500
Size 200
Size 500
Size 1000
Fitness
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490
Generation


Regression Methods Learning Curve
population = 500, generations = 500
Half & Half
Full
Grow
Fitness
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500
Generation


Regression Operators Learning Curve
population = 500, generations = 500
70% C; 25% M
80% C ; 15% M
90% C ; 5% M
Fitness
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500
Generation

Conclusions

- Artificial Ant Problem
  - The initial population method was most sensitive
  - Population size second most sensitive
  - Breeding operation probability least sensitive
- Regression Problem
  - Difficult to determine which parameters are more sensitive
    - All of the graphs are similarly shaped
  - Convergence occurs earlier
  - Because of the basis of the problem, there are higher optimal solutions

Review

- Genetic Programming
  - Branch of genetic algorithms that is a method for getting computers to automatically solve problems.
- Lil-GP System
  - Executable that contains two problems addressed in my experiment
- Artificial Ant and Regression Algorithms
  - Ant- deals with finding the optimal path for the ant
  - Regression- wants to find a function that matches the given data points
- Tested Parameters
  - Ant results showed major sensitivity to the initialization method
  - Regression's results were similar to Ant results, just not as much sensitivity among the different parameters.

Sources

- “*Genetic Programming*.”Dolan, Kevin. 2009.Web. 13 Dec. 2010. <http://geneticprogramming.us/What_is_Genetic_Programming>
- “Genetic Programming." *Wikipedia*. 14 Oct. 2010. Web. 13 Dec. 2010. <http://en.wikipedia.org/wiki/Genetic_Programming>.
- "What Is Genetic Programming?" *Genetic Programming*. 27 Aug. 2003. Web. 13 Dec. 2010. <http://www.genetic-programming.com/gpanimatedtutorial.html>.
- Zongker, Douglas, Punch, Dr. Bill, and Rand, Bill. *lil-gp 1.01 User's Manual*. Michigan State University, 1995. PDF