

GP Assignment Report

Terminal set a)

The terminal set is the x variable.

Function set b)

The addition, subtraction, multiplication, and protected division operators should be enough for the simple problem of describing the relationship between variables x and y.

This function set satisfies the closure requirement, because all functions can accept any two numbers, including zero.

Fitness function c)

The fitness function is the mean-squared error, that is, the mean of $(\hat{y}_i - y_i)^2$ for all i. It is the mean squared difference between the predicted y-value and the actual value. Programs that produce a lower mean-squared error are better.

Parameters and stopping criteria d)

The population size is set to 500 because it is large while the code still runs in a reasonable amount of time. The stopping criteria is when reaching generation 20 because there appears to be little improvement afterwards.

Best solutions e)

Seed	Best solution	Fitness
0	add(protectedDiv(x, x), mul(sub(x, mul(x, sub(add(x, x), mul(x, x)))), x))	7.03125e-12
1	sub(mul(add(add(x, x), x), add(x, x)), add(add(x, x), add(x, add(x, add(add(x, add(protectedDiv(x, x), mul(x, x))), x))))))	10.1336
2	sub(mul(add(add(x, add(x, x)), add(x, x)), x), add(x, add(add(x, add(x, add(x, x))), x)))	10.9047