Homework Set 3

Due date: Monday, Jan. 27.

Written Exercises

HW3.1. (a) Referring to the paper by Krawiec ("Genetic Programming-based Construction of Features..."), explain in one or two sentences the problem that led the author to develop the "extended method." (b) List **two** possible disadvantages of the "extended method" with respect to the original method (in terms of computational costs of the GP algorithm, or problems created for GP evolution of solutions). (10 points)

HW3.2. Consider "uniform crossover" as discussed on p. 58 of the textbook. Uniform crossover in bit strings is defined as follows: For each bit position in Parent 1 and Parent 2, with probability 0.5, exchange the bits at those positions. List one possible advantage of uniform crossover as opposed to traditional "single-point crossover" (as described in class and implemented in the Simple GA in C), and one possible disadvantage. (10 points)

HW3.3. Consider the following time series data:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8			
x^0	=	{0.45,	0.31,	0.49,	0.47,	0.84,	0.22,	0.54,	0.53 },	y^0	=	0.64
x^1	=	{0.97,	0.63,	0.26,	0.46,	0.94,	0.24,	0.34,	0.91 },	y^1	=	0.87
x^2	=	{0.91,	0.41,	0.41,	0.66,	0.26,	0.18,	0.48,	0.42 },	y^2	=	0.27
x^3	=	{0.74,	0.71,	0.90,	0.79,	0.27,	0.00,	0.92,	0.74 },	y^3	=	0.80
x^4	=	{0.56,	0.36,	0.25,	0.78,	0.85,	0.32,	0.49,	0.38 },	y^4	=	0.38
x^5	=	{0.04,	0.74,	0.55,	0.71,	0.47,	0.31,	0.28,	0.30 },	y^5	=	0.85
x^6	=	{0.55,	0.59,	0.66,	0.16,	0.48,	0.72,	0.89,	0.58 },	y^6	=	0.90
x^7	=	{0.98,	0.66,	0.24,	0.44,	0.34,	0.88,	0.86,	0.51 },	y^7	=	0.05
x^8	=	{0.92,	0.49,	0.84,	0.27,	0.97,	0.22,	0.55,	0.13 },	y^8	=	0.29
x^9	=	{0.43,	0.86,	0.69,	0.17,	0.60,	0.48,	0.28,	0.92 },	y^9	=	0.91

This data is also given in

```
nfs:  /home/projects/class/cse580ec/HW3.4
cifs: \\hopper\clss-prj\cse580ec\HW3.4
```

Consider the following two condition sets:

$C_1 = \{(x_2 \leq 0.5) \wedge (0.2 \leq x_4 \leq 0.7)\}$

$C_2 = \{(0.1 \leq x_1 \leq 0.8) \wedge (0.2 \leq x_4 \leq 0.7) \wedge (0.5 \leq x_8)\}$

a. Calculate the fitness of both C_1 and C_2 using the fitness function given on P. 58 of the textbook. Let $\alpha = 1$. Let $\log_2(0) = 0$ by definition.

b. Using this same fitness function, define a condition set C_3 that has a higher fitness than C_1 and C_2, and give its fitness.
(10 points)

HW3.4. Describe in words how you might implement the "random mutation hill climbing" algorithm (defined for bit strings on p.32 of the textbook, computer exercise 4) for the conditions sets used in Meyer and Packard's GA. (There are many possible correct answers to this problem.)
(10 points)

Computer Exercises

HW3.5. Implement the following functions, which will serve as the basis for a genetic programming system for evolving strategies for the Prisoner's Dilemma.

a. A data structure for parse trees. (30 points)

b. A function that generates random parse trees up to a depth MAX_TREE_DEPTH, choosing from the function set {AND, OR, NOT} and the terminal set $\{s_0, s_1, s_2, s_3, s_4, s_5\}$, where s_i is the ith bit in the six bits representing three previous moves in the Prisoner's dilemma. Recall that *strategy* for the Prisoner's dilemma is a lookup table that maps the six bits

$$s_0 s_1 \ s_2 s_3 \ s_4 s_5$$

into a move to be made in the current game. For example, in the case "CC CC CC", $s_i = 0$ for all i; in the case "DD DD DD", $s_i = 1$ for all i.
(30 points)

Hand in hard copy of your code and also e-mail it to the TA, weilin@ece.ogi.edu.