GAGP Tutorial 3

1. A population consists of the following strings. The probability of crossover is 0.75 and the probability of mutation is 0.1. How many instances of the schema *0***0 would you expect in the next population? [From 2005 exam paper.]

String	Fitness
100100	20
001000	20
110111	30
100101	20
100010	10

2. Investigate whether binary tournament selection (i.e. tournament size 2) is equivalent to linear ranking selection (i.e. selection in which the fittest of N gets N chances, the next-fittest gets (N-1) chances, etc., and the least fit gets one chance).
   Tournament selection is selection where m individuals are chosen randomly from the population and the best n of those m are selected. So in binary tournament selection $m = 2$ and $n = 1$.

3. If
   $f(****) = e_0$
   $f(***1) = e_0 + e_1$
   $f(**1*) = e_0 + e_2$
   $f(**11) = e_0 + e_1 + e_2 + e_3$
   (these equations DEFINE $e_0, \ldots, e_3$), then what is $f(**01)$ in terms of $e_0, \ldots, e_3$?

4. It has been observed that some organisms seem to pass on behaviours learned during their lifetime to their offspring. Lamarck's hypothesis was that traits acquired during the lifetime of an individual could somehow be passed on genetically to the individual's children. However, since there is no obvious biological mechanism for this, Lamarck's hypothesis is universally rejected.
   One proposal for a non-Lamarckian mechanism explaining the passing on of learned behaviours was given by Baldwin, who pointed out that if learning helps survival, then the organisms best able to learn will have the most offspring. Further, if the environment remains constant, so that the best things to learn remain constant, then this can lead, via selection, to a genetic encoding of a trait that previously had to be learned. Describe how you could use evolutionary computation as a model system to demonstrate the truth (or otherwise) of Baldwin's hypothesis.

5. (Mitchell) Design a fitness function (in terms of schemas, as in R_1, see 5. Lecture 9/10) on which you believe the GA should outperform RMHC. Test your hypothesis numerically.