SEMESTER 1 EXAMINATION 2012-13

EVOLUTION OF COMPLEXITY

Duration: 90 mins

This paper contains 5 questions - Answer **only THREE** questions**.**
(25 marks each, 75 marks total).

An outline marking scheme is shown in brackets to the right of each question.

Only University approved calculators may be used.

A foreign language translation dictionary (paper version) is permitted provided it contains no notes, additions or annotations.

Question 1.

a) Describe in detail (e.g. using pseudocode) an algorithm for implementing fitness proportionate selection. Assume a fitness function is provided. [8 marks]

b) Describe in detail (e.g. using pseudocode) a genetic algorithm. Use generational reproduction. Assume a fitness function and selection routine are provided (as per part a). Assume individuals are represented with bit-strings, and assume crossover and mutation functions are provided. [11 marks]

c) *Reproduction* and *fitness differences* are two of the four necessary and sufficient conditions that Ridley describes for evolution by natural selection. What are the other two and how are they incorporated in the genetic algorithm? [6 marks]

Question 2.

a) Consider the schema **110*1***** where '*' can be either 0 or 1. What is the order of this schema? How many strings are in this schema? What is its defining length? If the fitness function was 'counting-ones' or 'one-max', would this schema be a "building block"? Explain. [10 marks]

b) Describe two ways in which a genetic algorithm could, in principle, be superior to local search (e.g. a mutation hill-climber). How does the possibility of premature convergence affect these possible advantages? [9 marks]

c) Why can't *uniform* crossover be useful for the operation of the genetic algorithm when it comes to combining building blocks? Give one reason why uniform crossover might (nonetheless) be useful in the genetic algorithm. Explain. [6 marks]

TURN OVER

Question 3.

a) What is a fitness landscape? What is a local optimum?
[5 marks]

b) i) Describe (e.g. provide a well-labelled sketch or a verbal description) one example problem/fitness landscape where a genetic algorithm can find the global optimum quickly but other stochastic local search processes (e.g. a stochastic hill-climber) cannot. [9 marks]

ii) Explain why this problem requires exponential time for a hill-climber. [5 marks]

iii) Explain how a genetic algorithm can solve this type of problem in polynomial time. [6 marks]

Question 4.

a) Consider the four genotypes that are possible for two loci, each with two alleles. If their fitnesses are f(ab)=2, f(aB)=4, f(Ab)=5, f(AB)=7; Do these genotypes exhibit epistasis? Show your working. [5 marks]

b) Consider the following population of four 5-bit genotypes:

00001
00000
00000
10000

Is this population at linkage equilibrium? Explain. [4 marks]

c) What is the significance of linkage disequilibrium for the benefit of sex? Use the Fisher/Muller model as an example. [10 marks]

d) Describe some of the differences between natural fitness landscapes and the fitness landscapes common to engineering problems. [6 marks]

TURN OVER

Question 5.

a) One of the reasons that coevolution is appealing in artificial evolution is the possibility of an ‘open-ended arms race’ where one population provides a constant selection pressure on a second population’s improvement and that population in turn provides a constant selection pressure on the first’s improvement. One might hope that coevolution thereby provides a guaranteed method of open-ended improvement and a route to evolved complexity. Explain some ways in which it might fail.

[8 marks]

b) Describe a coevolutionary method for finding teams of individuals that solve a problem collectively. How is this approach different from adversarial or competitive coevolution? [5 marks]

c) “The complexity of living things is very high (by any reasonable measure) therefore it must be the case that evolution by natural selection produces a generic trend toward increasing complexity in nature.” Comment on the validity of this argument. [6 marks]

d) “The genome of organisms, even for bacteria, is at least 10^6 nucleotides long each of which can be one of 4 bases (A, C, T or G). The number of possible genomes is therefore at least $4^{1000000}$. A single mutation to any of the bases in an organism can be fatal or result in non-viable offspring. It is therefore inconceivable that a random trial and error process like evolution could have produced the genomes of living organisms.” Comment on the validity of this argument.

[6 marks]

END OF PAPER

