Natural Computing: Tutorial 1 (week 3)

1. Go (step-by-step and then fast) through example at (cf. lecture 3) http://www.obitko.com/tutorials/genetic-algorithms/example-function-minimum.php

2. The knapsack problem is as follows: Given a set of weights *W* and a target weight *T*, find a subset of *W* whose sum is as close to *T* as possible.

   Example: $W = \{5, 8, 10, 23, 27, 31, 37, 41\}$
   $T = 82$

   - Solve the instance of the knapsack problem given above.
   - Consider solving the knapsack problem using the canonical GA. How can a solution be encoded as a chromosome?
   - What fitness function can be used for the knapsack problem, so that better solutions have higher fitness?
   - Given your answer to question 2, what selection methods would be appropriate?
   - Try out the online demo at http://www.aridolan.com/ga/gaa/Knapsack01.html

3. Assume you have a lot of data points that seem to fall into clusters, e.g. the 2D position of mushrooms in a forest. Instead of applying the *k*-means algorithm directly, you decide to use GA to get the center positions of the clusters. How might you use a canonical GA to solve this, and what are the problems you might run into, particularly regarding the representation?

4. Run through a simple GA, applying fitness proportionate selection and single-point crossover. It could be the "maximize f(x) = x-squared" problem from the lecture notes. Set up a population of individuals by tossing a coin to get the initial chromosomes and use coin-tossing wherever you need to generate random numbers. Note how the average fitness, sum of individual fitness values, and maximum fitness change over the generations. You will need a calculator for this so bring a laptop or a mobile phone (or even a calculator if they still exist!). Or brush up on long multiplication and division.

5. What are negative side effects of crossover (and mutation)? Assuming that crossover and mutations are important for good performance, what can be done to reduce the side effects?

6. Why is crossover usually applied before mutation in GA?

7. Discuss implications of the schema theorem for the following cases (recall the definition of the fitness of a schema):
   - a single instance of a high-fitness schema
   - two different non-overlapping schemas with the same fitness
   - two partially overlapping schemas with a fitness that are both high but not the same
   - a fitness function that depends on the presence of other individuals, such as in the evolution of an ecosystem consisting of rabbits and foxes.