T-79.250 **Spring 2005**

Combinatorial Models and Stochastic Algorithms

Tutorial 12, April 22

Problems

1. [*Computer problem.*] Consider the following simple genetic algorithm: A population consists of two 2-bit strings, and the values of the objective function c are determined as $c(00) = 1$, $c(01) = c(10) = 2$, $c(11) = 4$. At each computation step, a one-point crossover occurs between the two strings with probability 25%, and results in simply replacing the given two strings by the ones obtained by interchanging their first and second bits. Following that, each bit of each string is mutated (flipped) with a 10% probability, independent of the other bits. Selection is performed based on a relative cost fitness measure, i.e. two new strings are selected from the two existing ones (with replacement), so that both existing strings have a probability of being selected proportional to their objective function value. Determine the stationary population distribution in this setting using the Vose-Liepins canonical GA model (lecture notes p. 115). What would the stationary distribution be if no selection between individuals were performed? (*Note:* You will definitely need the help of a computer algebra system such as Maple, Mathematica, or Matlab to work on this problem.)

2. Consider the following *k-Set Splitting* problem: Given a collection $\mathcal{C}$ of k-element subsets of a finite set S, is there a subset $S' \subseteq S$ such that no $C \in \mathcal{C}$ is contained in either S' or $S - S'$ (i.e., S' "splits" all the sets in $\mathcal{C}$ in two pieces). The problem is NP-complete for $k \geq 3$. Make an educated guess concerning the location of "hard instances" for this problem.