Helsinki University of Technology
Laboratory for Theoretical Computer Science
Pekka Orponen (tel. 5246)

T-79.250 Combinatorial Models and Stochastic Algorithms (4 cr)
Exam Wed 14 May 2003, 9–12 a.m.

Permitted material at exam: lecture notes, any personal handwritten notes, tutorial problems and their solutions; calculator.

Write down on each answer sheet:
- Your name, department, and study book number
- The text: "T-79.250 Combinatorial Models and Stochastic Algorithms 14.5.2003"
- The total number of answer sheets you are submitting for grading

1. Prove that the graph property "G contains a connected subgraph of at least k nodes" has a threshold function for any fixed $k \geq 2$, and compute it. *6p.*

2. Let m and n be positive integers with $\gcd(m, n) = 1$, and consider the Markov chain defined on state space $S = \{0, \dots, m-1\} \times \{0, \dots, n-1\}$ by the transition rule: from state (i, j), move with probability $\frac{1}{2}$ to state $(i+1 \bmod m, j)$, and with probability $\frac{1}{2}$ to state $(i, j+1 \bmod n)$. (I.e. the chain always moves either "right" or "up" on the rectangular grid S, with periodic boundary conditions.) Show that the chain is regular and determine its stationary distribution. *6p.*

3. Give some nontrivial upper bound on the mixing time of the chain defined in Problem 2. (You don't need to try to get a best possible result.) *6p.*

4. Consider an NK model with $K = 2$, binary alleles ($a \in \{0, 1\}$), and each locus i being influenced by loci $i \pm 1 \pmod N$. The local fitness function for the model is defined as $f(a_i; a_{i-1}, a_{i+1}) = 1$, if $a_{i-1} \neq a_i = a_{i+1}$, and 0 otherwise. Design a simulated annealing approach to optimising the corresponding genotype fitness function, and give a (theoretical) bound on a cooling schedule that guarantees asymptotic convergence to the set of fitness-optimal genotypes. *6p.*

5. Recall the following results from Tutorial Problem 6/4:

   "Consider a random walk on an undirected graph $G = (V, E)$, where transitions are made from each node u to an adjacent node with uniform probability β/d, where d is the maximum degree of any node in G and $\beta \leq 1$ is a positive constant. In addition, each node u has a self-loop probability of $1 - \beta \deg(u)/d$. If G is connected and $\beta < 1$, then the corresponding Markov chain $\mathcal{M}_G$ is regular and reversible, with uniform stationary distribution. Moreover, the conductance of $\mathcal{M}_G$ is given by the formula
   $$\Phi = \beta\mu(G)/d,$$
   where $\mu(G)$, the *edge magnification* of G, is defined as
   $$\mu(G) = \min_{0<|U|\leq|V|/2} \frac{|\partial(U)|}{|U|},$$
   where $\partial(U) = \{\{u, v\} \in E \mid u \in U, v \notin U\}$."

   Based on these results, design a Markov chain for sampling (approximately) uniformly at random a node from a "caveman graph" consisting of m r-node "caves" connected together in an m-cycle, as discussed in the lectures. Estimate the number of steps required for the sampling process to converge within a relative pointwise distance of ε of the uniform distribution. Can you relax in the design of the Markov chain the assumption of knowing the number of the caves? *6p.*

Total 30p.