Advanced Algorithms, Fall 2012

Prof. Bernard Moret

Homework Assignment #5

due Sunday night, Oct. 28

Write your solutions in LaTeX using the template provided on the Moodle and web sites and upload your PDF file through Moodle by 4:00 am Monday morning, Oct. 29.

Question 1. (Greedy Algorithms)
You are given an undirected connected graph $G = (V, E)$ such that $|E| = |V| + 5$ (i.e., this is a very sparse graph). Each edge of G has a distinct cost. Design a linear-time algorithm to find the Minimum Spanning Tree of G, and prove the correctness of your algorithm.

Question 2. (Greedy Algorithms)
Given n positive integers, $a_1, a_2, \cdots, a_n$, design an polynomial-time algorithm to determine whether there exists an undirected simple graph (self loops and multi-edges are not allowed) with n nodes such that the degrees of the n nodes correspond to the given n integers; prove the correctness of your algorithm.

Question 3. (Greedy Algorithms)
This problem addresses optimal file allocation for computer networks. You are given a completely connected network of nodes, a set of files to be allocated among the nodes, and a sequence of retrieval and updating requests (the entire sequence is known in advance). A retrieval or updating request is a triple, consisting of the node initiating the request, the file involved, and the number of bytes to be transferred.

   An allocation scheme is an assignment of each file to one or more nodes. Having multiple copies of a file is advantageous in retrieval: the cost of a retrieval is zero when the file is held locally, but equals the number of bytes to be transferred when the file must be accessed remotely. On the other hand, multiple copies of a file increase the cost of an updating operation, because each copy must be updated and thus the number of bytes needed for updating is multiplied by the number of remote copies. Multiple copies are also desirable for reasons of reliability, but only up to a point, since relatively few nodes are expected to fail. The gain in reliability is a function of the number of copies and obeys the law of diminishing returns: each additional copy of a file gives a smaller gain than the previous copy. The cost of an allocation scheme is the cost of the given sequences of retrieval and updating requests minus the gain in reliability.

   Develop a greedy algorithm that constructs the optimal allocation scheme and prove its correctness.

Question 4. (Matching)
We are given a $n \times n$ 0-1 matrix M. We can swap two columns of M, or two rows of M. We say that M can be diagonalized if we can perform a sequence of (row and/or column) swaps to make all the diagonal items of M equal to 1. Design an algorithm to decide whether a given binary matrix M can be diagonalized and prove the correctness of your algorithm.