$22\mathbf{C} : 031$ $(CS : 3330 : 0001)$ **Algorithms**
Homework 2

This homework is based on our discussion of greedy algorithms for the interval scheduling problem from Section 4.1. The homework is worth 10 points,

- Suppose that we are given a pattern S', which is a sequence of m items, and a much bigger data sequence S consisting of n items. A given item can occur multiple times in both the pattern and the data sequence. We say that S' is a *subsequence* of S if there is a way to delete certain items from S so that the remaining items, in order, are equal to the sequence S'. For example the sequence of items $\langle Y, E, Y, O \rangle$ is a subsequence of the sequence $\langle A, Y, E, Y, Y, O \rangle$, but is not a subsequence of $\langle A, Y, E, A, O, Y \rangle$.

  Give an efficient (polynomial time) algorithm to determine if a given sequence S' is a subsequence of given sequence S.

- We are given the locations of houses along a long road. Modeling the road as the real line, we are given the house locations as numbers on the line. For example, House 1 is at 1.5, House 2 is at 3, and so on. We want to place cell phone base stations at certain locations on the road (line) so that each house is within 5 miles of some base station. (Assume that a unit on the real line corresponds to a mile – so in the above example, houses 1 and 2 are 1.5 miles apart.) Describe an efficient (polynomial time) algorithm that places the smallest possible number of base stations.

- In the process of finding a correct greedy algorithm for the interval scheduling problem (Section 4.1), we also showed that several other rather natural looking greedy algorithms do not always work. However, some nice things can be said about some of these algorithms.

  Let us revisit one of these greedy algorithms: Pick the shortest (that is, the minimum length) interval, remove it and all incompatible intervals, and repeat this process till no more intervals remain.

  For this greedy algorithm, argue that the number of intervals picked is at least as large as half the size of the optimal solution.

  Hint: Let us recall the analysis of the greedy algorithm that picks the interval with the earliest finish time, rephrasing it slightly so that it suggests a direction for the argument in the present question. Suppose that the optimal solution consists of the intervals $o_1, o_2, \ldots, o_m$. The analysis works by claiming that after the first iteration, when the greedy algorithm has picked its first interval and removed incompatible intervals, *at most one* of the intervals in the optimal solution has been discarded, so *at least* $m-1$ of the intervals in the optimal solution are still in the set of intervals that have not been discarded. After the first two iterations, *at most two* of the intervals in the optimal solution have been discarded, so *at least* $m-2$ of the intervals in the optimal solution are still in the set of intervals that have not been discarded.

- A k-coloring of an undirected graph is an assignment of a number (called a color) from the set $\{1, 2, \ldots, k\}$ to each vertex of the graph, with the property that for any edge, the two incident vertices receive different colors. We talked about graph coloring in the context of the interval partitioning problem from Section 4.1. Recall that the *degree* of a vertex u in a graph is the number of edges incident to u. Describe an algorithm that $(\Delta + 1)$-colors a given undirected graph $G = (V, E)$, where Δ is the maximum degree of a vertex in the graph. (We are looking for an efficient algorithm, in particular, it should run in time that is polynomial in the number of vertices plus edges of the graph. Try some examples to see how you would do the coloring.)

The homework is due Monday, January 30, in class; if you can't make it to class on that day, just make sure you get it to me by that time. Our policy on homeworks is reproduced for your convenience:

"No collaboration is allowed on the exams and quizzes. For homework problems, collaboration is alright. We even encourage it, assuming each of you has first spent some time (about 45 minutes) thinking about the problem yourself. However, no written transcript (electronic or otherwise) of the collaborative discussion should be taken from the discussion by any participant. It will be assumed that each of you is capable of orally explaining the solution that you turn in, so do not turn in something you don't understand."