Problem 1.

(Siegel 5.23) Consider the following sorting problem. The input is a sequence of \(n\) integers with many duplications, such that the number of distinct integers in the sequence is \(O(\log(n))\). Design a sorting algorithm to sort such sequences using at most \(O(n \log \log n)\) comparisons in the worst case.

Problem 2.

In a tree, the tree distance between nodes \(X\) and \(Y\) is the number of steps it takes to get from \(X\) to \(Y\), going up and down the tree. The least common ancestor of nodes \(X\) and \(Y\) is the lowest node in the tree that is an ancestor of both \(X\) and \(Y\). Note that if \(Z\) is the least common ancestor of \(X\) and \(Y\), then the tree distance from \(X\) to \(Y\) is the number of steps from \(X\) to \(Z\) plus the number of steps from \(Y\) to \(Z\).

For example, in the figure below, the tree distance from \(F\) to \(K\) is 3 (\(F \rightarrow G \rightarrow J \rightarrow K\)). The least common ancestor of \(F\) and \(K\) is \(G\). The least common ancestor of \(C\) and \(B\) is \(B\) itself.

Describe a method that takes as input two tree nodes \(X\) and \(Y\) and returns their least common ancestor, and runs in time proportional to the tree distance from \(X\) to \(Y\). You may add additional fields to the nodes, and you may assume that these are initialized to null. Assume that the implementation of the tree includes upward pointers, from each node to its parent. Hint: Climb upward from \(X\) and \(Y\) in alternating steps.
Problem 3

Suppose that we modify the standard definition of a binary search tree to add a field N.size at each node, which records the size of the subtree under N (including N itself).

A. Explain how to modify the procedure for adding an element X to a tree. Be sure to consider both the case where X is not yet in the tree and is added, and the case where X is already in the tree, and the tree remains unchanged.

B. Explain how to modify the procedure for deleting an element. As in (A), consider both cases.

C. Describe a procedure for finding the kth largest element in the tree.

D. Describe a procedure for finding the number of elements in the tree less than X.

All of these procedures should run in time proportional to the height of the tree.