(CS1.301) Algorithm Analysis and Design (Monsoon 2023)

Practice Problem Set 3

Instructions:

- Discussions amongst the students are not discouraged.
- Referring sources other than the lecture notes and textbooks is discouraged as the corresponding solutions available on the internet need not be accurate.
- Please attend tutorials to ensure that you understand the problem correctly.

Problems for the course

Question 1 Given a sorted array of distinct integers $A[1, n]$, we want to find out whether there is an index i for which $A[i] = i$. Give a divide-and-conquer algorithm that runs in time $O(\log(n))$.

Question 2 Consider the following algorithm given an array as input. Divide the input array into $[\frac{n}{5}]$ blocks, each containing exactly 5 elements, except possibly the last. (If the last block isn't full, just throw in a few 1s.) Then compute the median of each block by brute force, collect those medians into a new array $M[1 \dots [\frac{n}{5}]]$, and then recursively compute the median of this new array. Show that this algorithm computes an element that is at least $\frac{3}{10}^{th}$ largest and at most $\frac{7}{10}^{th}$ largest.

Question 3 Describe and analyze a variant of Karatsuba's algorithm that multiplies any m-digit number and any n-digit number, for any $n \geqslant m$, in $O(nm^{\log_2 3-1})$ time.

Question 4 Suppose you're consulting for a bank that's concerned about fraud detection, and they come to you with the following problem. They have a collection of n bank cards that they've confiscated, suspecting them of being used in fraud. Each bank card is a small plastic object, containing a magnetic stripe with some encrypted data, and it corresponds to a unique account

in the bank. Each account can have many bank cards corresponding to it, and we'll say that two bank cards are equivalent if they correspond to the same account.

It's very difficult to read the account number off a bank card directly, but the bank has a high-tech "equivalence tester" that takes two bank cards and, after performing some computations, determines whether they are equivalent.

Their question is the following: among the collection of n cards, is there a set of more than $\frac{n}{2}$ of them that are all equivalent to one another? Assume that the only feasible operations you can do with the cards are to pick two of them and plug them in to the equivalence tester. Show how to decide the answer to their question with only $O(n \log(n))$ invocations of the equivalence tester.

Question 5 Given an array of $n \geqslant 2$ integers, say $[x(1), \ldots, x(n)]$, we want to find the largest step d, which is defined to be the max of $x(j) - x(i)$ over all $j > i$. For example, for $x = [22, 5, 8, 10, -3, 1]$ and in this case, $d = x(4) - x(2) = 10 - 5 = 5$.

Additional Practice Problems[1]

Question 6 You are interested in analyzing some hard-to-obtain data from two separate databases. Each database contains n numerical values—so there are $2n$ values total—and you may assume that no two values are the same. You'd like to determine the median of this set of $2n$ values, which we will define here to be the n^{th} smallest value.

However, the only way you can access these values is through queries to the databases. In a single query, you can specify a value k to one of the two databases, and the chosen database will return the k^{th} smallest value that it contains. Since queries are expensive, you would like to compute the median using as few queries as possible. Give an algorithm that finds the median value using at most $O(\log(n))$ queries.

Question 8 Given a set, P , of n teams in some sport, a round-robin tournament is a collection of games in which each team plays each other team exactly once. Such round-robin tournaments are often used as the first round for establishing the order of teams (and their seedings) for later single- or double- elimination tournaments. Design an efficient algorithm for constructing a round-robin tournament for a set, P , of n teams assuming n is a power of 2.

[1] These won't be part of the exams.

Question 7 Imagine that you are getting ready to go on a vacation and are looking for a hotel. Suppose that for you, the only criteria that matter when you judge a hotel are the size of its pool and the quality of its restaurant, as measured by a well-known restaurant guide. You have gone to the restaurant guide website and several hotel websites and have discovered, for each of n hotels, the size of its pool and the quality of its restaurant. With so many hotels to choose from, it would be useful if you could rule out some possibilities. For instance, if there is a hotel whose pool is smaller and restaurant is of lower quality than another hotel, then it can be eliminated as a possible choice, as both of your criteria for the first hotel are dominated by the second hotel. We can visualize the various trade-offs by plotting each hotel as a two-dimensional point, (x, y), where x is the pool size and y is the restaurant quality score. We say that such a point is a maximum point in a set if there is no other point, (x', y'), in that set such that $x \leqslant x'$ and $y \leqslant y'$. Given such a set of points, then, design an efficient divide-and-conquer algorithm that can identify the maxima set.