Problems: Order Statistics and String Matching

Palash Dey
Indian Institute of Technology, Kharagpur

1. [CLRS] Design an algorithm to compute the second smallest of n elements with $(n+\lceil \log_2 n \rceil - 2)$ comparisons in the worst case.

2. Design an algorithm to compute the smallest and largest of n elements with $3\lfloor n/2 \rfloor$ comparisons in the worst case.

3. [CLRS] Prove the lower bound of $(\lceil 3n/2 \rceil - 2)$ comparisons in the worst case to find both the maximum and minimum of n numbers.

   Solution sketch. Initially, all n numbers are potentially either the maximum or minimum. Let MAX be the set of numbers which are potentially the maximum and MIN denote the set of numbers which are potentially the minimum. Say we compare two elements a and b. If $a \leqslant b$, we can remove a from MAX and remove b from MIN, so we reduce the counts of both sets by 1. If we compare two elements in MIN, we reduce the size of MIN by 1, and if we compare two elements in MAX, we can reduce the size of MAX by 1. Thus, the first type of comparison is optimal. There are $\lceil n/2 \rceil$ such comparisons which can be made until MIN and MAX are disjoint, and the sets will have sizes $\lceil n/2 \rceil$ and $\lfloor n/2 \rfloor$. Within each of these, only the second and third types of comparisons can be made, each reducing the set size by 1, so we require a total of $\lceil n/2 \rceil - 1 + \lfloor n/2 \rfloor - 1 = n - 2$comparisons. Adding this to initial $\lceil n/2 \rceil$ comparisons gives $(\lceil 3n/2 \rceil - 2)$. □

4. [CLRS] In the algorithm S ELECT, the input elements are divided into groups of 5. Will the algorithm work in linear time if they are divided into groups of 7? Argue that S ELECT does not run in linear time if groups of 3 are used.

5. [CLRS] Suppose that you have a "black-box" worst-case linear-time median subroutine. Give a simple, linear-time algorithm that solves the selection problem for an arbitrary order statistic.

6. [CLRS] The k-th quantiles of an n-element set are the k - 1 order statistics that divide the sorted set into k equal-sized sets (to within 1). Give an $\mathcal{O}(n \log k)$-time algorithm to list the k-th quantiles of a set.

7. [CLRS] Give a linear-time algorithm to determine whether a text T is a cyclic rotation of another string T'.