算法基础 2023 春 | Homework 2

任课老师：陈雪 & 邵帅 | due: Mar 23, 23:59

**问题 1** (30). 分析程序，证明下面的问题：

**Algorithm 1** Quick Sort algorithm

**function** QSort$(l, r)$
  **if** $l \geq r$ **then**
    **exit**
  pick a random element in $A[l], ..., A[r]$ as *pivot*
  $i = l, j = r$
  **while** $i \leq j$ **do**
    **while** $A[i] < pivot$ **do** $i = i + 1$
    **while** $A[j] > pivot$ **do** $j = j - 1$
    **if** $i \leq j$ **then**
      SWAP$(A[i], A[j])$
      $i = i + 1$
      $j = j - 1$
  QSort$(l, j)$
  QSort$(i, r)$

(a) 证明 QSort$(1, n)$ 算法的正确性。

(b) 用递归公式分析算法的期望运行时间为 $O(n \log(n))$。

**问题 2** (20). 归并排序算法如下，试改进归并算法，计算逆序对数目 (存在正整数 $i, j$ 使得 $1 \leq i < j \leq n$ 而且 $A[i] > A[j]$，则 $\langle A[i], A[j] \rangle$ 称为 $A$ 的一个逆序对)。

**Algorithm 2** Merge Sort algorithm

**function** MERGE$(p, j, k)$
    $n_1 = j - p$, $n_2 = k - j + 1$
    Let $L[1, ..., n_1 + 1]$ and $R[1, ..., n_2 + 1]$ be new arrays to avoid shift
    $L[i] = A[p + i - 1]$ for $i \leq n_1$ and $L[n_1 + 1] = \infty$
    $R[i] = A[j + i - 1]$ for $i \leq n_2$ and $R[n_2 + 1] = \infty$
    $i_l = 1$, $i_r = 1$
    **for** $i = p, ..., k$ **do**
        **if** $L[i_l] \leq R[i_r]$ **then**
            $A[i] = L[i_l]$
            $i_l = i_l + 1$
        **else**
            $A[i] = R[i_r]$
            $i_r = i_r + 1$
**function** SORT$(p, r)$
    **if** $p < r$ **then**
        $q = \frac{p+r}{2}$
        SORT$(p, q)$
        SORT$(q + 1, r)$
        MERGE$(p, q + 1, r)$

**问题 3** (30). 给定两个排好序的长为 $n$ 的数组 $X[1 \dots n]$ 和 $Y[1 \dots n]$:

(a) 给定一个键值 $key$，设计算法找出 $X$ 中小于等于 $key$ 的元素个数；并分析算法的正确性。

(b) 给定一个整数 $k \in [1, n]$，设计算法在 $O(\log n)$ 时间内找到 $X$ 和 $Y$ 一共 $2n$ 个元素中第 $k$ 大的元素；并分析算法的正确性。

**问题 4** (20). 参考算法导论第 8.1 章，考虑 $\sigma$ 为 $n$ 个未知元素所有 $n!$ 个排列中均匀随机产生的。

(a) 对任何基于比较排序的确定性算法 $A$，证明对至少 99%的全排列，$A$ 的运行时间为 $\Omega(n \log n)$。也就是

$$\Pr_{\sigma}\left[Time\Big(A(\sigma)\Big) \geq 0.5n\log n\right] \geq 0.99.$$

其中 $Time(A(\sigma))$ 表示算法 $A$ 的运行时间。

(b) 考虑任一基于比较排序的随机化算法 $A$（永远输出正确答案，但运行时间为随机变量），证明类似的结论：

$$\Pr_{\sigma}\left[\mathop{\mathbb{E}}_{r}\left[Time\big(A(\sigma, r)\big)\right] \geq 0.5n\log n\right] \geq 0.99.$$

**提示：** 假定算法 $A$ 采用固定长度 $l$ 的随机数 $r$，那么 $r \in \{0,1\}^l$。在随机数 $r$ 确定之后，$A(\cdot, r)$ 就是一个确定性的算法。