算法导论习题选集

作业 6-1

节选自《算法导论》教材第三版

课程网站：https://algorithm.cuijiacai.com

Problem 1

（**Hoare 划分的正确性**）第 6 讲 PPT 第 6 页的 PARTITION 算法并不是其最初的版本。下面给出的是最早由 C.R.Hoare 所设计的划分算法：

```
HOARE-PARTITION(A, p, r)
1   x = A[p]
2   i = p - 1
3   j = r + 1
4   while TRUE
5       repeat
6           j = j - 1
7       until A[j] ≤ x
8       repeat
9           i = i + 1
10      until A[i] ≥ x
11      if i < j
12          exchange A[i] with A[j]
13      else return j
```

1. 试说明 HOARE-PARTITION 在数组 $A = \langle 13, 19, 9, 5, 12, 8, 7, 4, 11, 2, 6, 21 \rangle$ 上的操作过程，并说明在每一次执行第 4 到 13 行 **while** 循环时数组元素的值和辅助变量的值。

后续的三个问题要求读者仔细论证 HOARE-PARTITION 的正确性。在这里假设子数组 $A[p..r]$ 至少包含 2 个元素，试证明下列问题：

2. 下标 i 和 j 不会使我们访问在子数组 $A[p..r]$ 以外的数组 A 的元素。

3. 当 HOARE-PARTITION 结束时，它返回的值 j 满足 $p \leq j < r$ 。

4. 当 HOARE-PARTITION 结束时，$A[p..j]$ 中的每一个元素都小于或等于 $A[j+1..r]$ 中的元素。

在第 6 讲 PPT 第 6 页的 PARTITION 过程中，主元（原来储存在 $A[r]$ 中）是与它所划分的两个分区分离的。与之对应，在 HOARE-PARTITION 中，主元（原来储存在 $A[p]$ 中）是存在于分区 $A[p..j]$ 或 $A[j+1..r]$ 中的。因为有 $p \leq j < r$ ，所以这一划分总是非平凡的。

Problem 1 continued on next page…

5. 利用 HOARE-PARTITION，重写 QUICKSORT 算法。

Problem 1 continued on next page…

(续页)

(续页)

Problem 2

（针对相同元素值的快速排序） 在第 6 讲 PPT 第 28 页对随机化快速排序的分析中，我们假设输入元素的值是互异的，在本题中，我们将看看如果这一假设不成立会出现什么情况。

1. 如果所有输入元素的值都相同，那么随机化快速排序的运行时间会是多少?

2. PARTITION 过程返回一个数组下标 q ，使得 $A[p..q-1]$ 中的每个元素都小于或等于 $A[q]$ ，而 $A[q+1..r]$ 中的每个元素都大于 $A[q]$ 。请修改 PARTITION 代码来构造一个新的 PARTITION’(A, p, r)，它排列 $A[p..r]$ 的元素，返回值是两个数组下标 q 和 t ，其中 $p \leq q \leq t \leq r$ ，且有

- $A[q..t]$ 中的所有元素都相等。
- $A[p..q-1]$ 中的每个元素都小于 $A[q]$ 。
- $A[t+1..r]$ 中的每个元素都大于 $A[q]$ 。

与 PARTITION 类似，新构造的 PARTITION’ 的时间复杂度是 $\Theta(r-p)$ 。

3. 将 RANDOMIZED-PARTITION 过程修改为调用 PARTITION’，并重新命名为 RANDOMIZED-PARTITION’。请修改 QUICKSORT 的代码构造一个新的 QUICKSORT’(A, p, r)，它调用 RANDOMIZED-PARTITION’，并且只有分区内的元素互不相同的时候才做递归调用。

4. 在 QUICKSORT’ 中，应该如何改变第 6 讲 PPT 第 28 页中的分析方法，从而避免所有元素都是互异的这一假设?

Problem 2 continued on next page…

(续页)

Problem 2 continued on next page…

(续页)

Problem 3

（另一种快速排序的分析方法）对随机化版本的快速排序算法，还有另一种性能分析方法，这一方法关注每一次单独递归调用的期望运行时间，而不是比较的次数。

1. 证明：给定一个大小为 n 的数组，任何特定元素被选为主元的概率为 $1/n$ 。利用这一点来定义指示器随机变量 $X_i = I\{H_i\}$ ，其中事件 H_i 表示“第 i 小的元素被选为主元为事件”，$E[X_i]$ 是什么？

2. 设 $T(n)$ 是一个表示快速排序在一个大小为 n 的数组上运行时间的随机变量，试证明：

$$E[T(n)] = E\left[\sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n))\right]$$

3. 证明第 2 问中的公式可以重写为：

$$E[T(n)] = \frac{2}{n}\sum_{q=2}^{n-1} E[T(q)] + \Theta(n)$$

4. 证明下面的等式（提示：可以将累加式分成两个部分，一部分是 $k = 2, 3, \cdots, \lceil n/2 \rceil - 1$ ，另一部分是 $k = \lceil n/2 \rceil, \cdots, n-1$ ）。

$$\sum_{k=2}^{n-1} k \log k \le \frac{1}{2}n^2 \log n - \frac{1}{8}n^2$$

5. 利用第 4 问中给出的界来证明：第 3 问中的递归式有解 $E[T(n)] = O(n \log n)$ 。（提示：使用代入法，证明对于某个正常数 a 和足够大的 n ，有 $E[T(n)] \le an \log n$ 。）

Problem 3 continued on next page…

(续页)

Problem 3 continued on next page…

(续页)