算法导论　第四周作业　10 月 8 日　周四

PB18151866 龚小航

2.1 堆排序:

对于一个按升序排列的包含 n 个元素的有序数组 A 来说,堆排序的时间复杂度是多少? 如果 A 是降序的呢? 请简要分析并给出结果。

解:对于堆排序,时间开销可以分为两部分:建堆开销以及维护最大堆的性质的开销。

函数 $\mathrm{MAX-HEAPIFY}(A,i)$ 的作用是维护堆 A 中第 i 个节点的最大堆性质,即将第 i 个节点和它的左右孩子这三者中最大的做交换。由于维护最大堆会有递归调用的情况发生,维护最大堆的开销包括:调整第 i 个节点和它左右孩子之间的关系开销 $\Theta(1)$,加上在一颗 i 的一个孩子节点为根的子树上运行 $\mathrm{MAX-HEAPIFY}$ 的代价。每个孩子的子树大小至多为 $2n/3$,最坏情况发生在树底层恰好半满的状态。因此 $\mathrm{MAX-HEAPIFY}$ 的运行时间可以用以下递归式刻画:

$$T(n) \le T\left(\frac{2}{3}n\right) + \Theta(1)$$

对其应用主定理,$f(n) = \Theta(1) = \Theta\left(n^{\log_{\frac{3}{2}} 1}\right) = \Theta\left(n^{\log_b a}\right)$,符合主定理情况 2。

因此 $\mathrm{MAX-HEAPIFY}$ 的时间复杂度为 $T(n) = O(\lg n)$,即对高度为 h 的节点复杂度为 $O(h)$,即在建好最大堆后进行堆排序所需的时间是 $n-1$ 次调用 $\mathrm{MAX-HEAPIFY}$,每次的时间为 $O(\lg n)$,因此排序过程的时间复杂度为 $O(n\lg n)$,这部分所需的时间与输入数组的顺序是无关的;

再考虑建堆开销:

含有 n 个元素的堆高度为 $\lfloor \lg n \rfloor$,高度为 h 的层中最多包含 $\lceil n/2^{h+1} \rceil$ 个节点。因此在一个高度为 h 的节点上运行 $\mathrm{MAX-HEAPIFY}$ 的时间开销为 $O(h)$ 。由此建堆的总时间开销为:

$$\mathrm{BUILD\_MAX\_HEAP} = \sum_{h=0}^{\lfloor \lg n \rfloor} \left\lceil \frac{n}{2^{h+1}} \right\rceil O(h) = O\left(n \sum_{h=0}^{\lfloor \lg n \rfloor} \frac{h}{2^h}\right) = O(h)$$

其中利用了错位相减法,求和号下和为 2。

综上,不论输入数组的组成是什么样的,堆排序的建堆开销都可以在线性时间内完成,堆排序的主要时间开销在于排序算法消耗的 $O(n\lg n)$ 。因此不论数组 A 的输入是升序或是降序,堆排序的时间复杂度都为 $O(n\lg n)$

2.2 快速排序:

(a) 假设快速排序的每一层所做的划分比例都是 $1-\alpha : \alpha$,其中 $0 < \alpha \le 1/2$ 且是一个常数. 试证明:在相应的递归树中,叶结点的最小深度大约是 $-\lg n/\lg\alpha$,最大深度大约是 $-\lg n/\lg(1-\alpha)$ (无需考虑舍入问题).

(b) 试证明:在一个随机输入数组上,对于任何常数 $0 < \alpha \le 1/2$,划分产生比 $1-\alpha : \alpha$ 更平衡的划分的概率约为 $1-2\alpha$.

解:分析说明:

(a): 设最大深度为 m ,叶节点的最小深度对应于每次都被分配 α 的那一支,m 次分割后仅剩一个元素。因此有: $n \cdot \alpha^m = 1 \quad \Rightarrow \quad \alpha^m = 1/n$ 两边取对数,立刻得到:

$$\lg \alpha^m = \lg \frac{1}{n} \quad \Rightarrow \quad m = -\frac{\lg n}{\lg \alpha}$$

最大深度同上,对应于每次都被分配 $1-\alpha$ 的那一支,m 次分割后仅剩一个元素。因此有:

$$n \cdot (1-\alpha)^m = 1 \quad \Rightarrow \quad (1-\alpha)^m = \frac{1}{n} \quad \Rightarrow \quad m = -\frac{\lg n}{\lg(1-\alpha)}$$

(b): 由于数组是随机输入的,假设主元也是随机选取的。$1-\alpha : \alpha$ 的划分表示有 $(1-\alpha)n$ 个元素比主元小,αn 个元素比主元大,由题意,$0 < \alpha \le 1/2$ 即小于主元的元素更多。因此若要产生更平衡的划分,主元应该取更小,但不能小于数组中第 αn 个小的数。因此可得产生更平衡的划分概率:

$$P = \frac{n-2\alpha n}{n} = 1-2\alpha$$