算法导论　第二周作业　9 月 24 日　周四

PB18151866 龚小航

3.1 下面的排序算法中哪些是稳定的：插入排序、归并排序、堆排序、快速排序和计数排序？

给出一个能使任何排序算法都稳定的方法。你所给出的方法带来的额外时间和空间开销是多少？

解：在给出的排序算法中，稳定的排序算法为插入排序、归并排序、计数排序。

为了使所有排序算法都是稳定的，只能使用多条件排序，在原本的主条件的基础上附加次要条件。可以这样构造次要条件：在读取被排数组时，不将它们存入线性数组而是将它们存入结构体或是二维数组中，同时存入读取到的元素值以及读取到的顺序序号，例如需要排序 n 个数就将它们的序号标为 $(1,2,3,\cdots,n)$。这么做会额外带来 100% 的空间开销，以及按照次要条件排序相同元素所需的时间开销，平均时间开销渐近不变。

3.2 假设用 Random − Select 去选择数组 $A = \langle 3,2,9,0,7,5,4,8,6,1\rangle$ 的最小元素，给出能够导致 Random − Select 最坏情况发生的一个划分序列。

```
Random-Select(A,p,r,i){
  if p == r then
    return A[p]
  q = Randomized-Partition(A,p,r)
  k = q - p + 1
  if i == k then
    return A[q]
  else
   if i < k then
    return Random-Select(A,p,q-1,i)
   else
    return Random-Select(A,q+1,r,i-k)
}
```

解：代码中 q 是通过随机选择得到的元素。根据算法，当用 Random − Select 去选择数组 A 的最小元素时，若随机到的 q 就是最小元素则立刻得到结果，而若每次都随机到最大元素则每次只能确定一个最大元素，所有比它小的数都无法确定位置。因此这种情况对应的就是最差情况。

此时的划分系列为：

$$9 \to 8 \to 7 \to 6 \to 5 \to 4 \to 3 \to 2 \to 1 \to 0$$

3.3 因为在基于比较的排序模型中，完成 n 个元素的排序，其最坏情况下需要 $\Omega(n\lg n)$ 时间。试证明：任何基于比较的算法从 n 个元素的任意序列中构造一棵二叉搜索树，其最坏情况下需要 $\Omega(n\lg n)$ 的时间。

解：可以从二叉搜索树的特性出发来证明这个结论。

由于中序遍历一次二叉搜索树就可以得到树上所有元素的升序系列，相当于是完成了一次排序，遍历二叉树所需要的时间为 $O(n)$。由提示，任何基于比较的排序模型要完成 n 个数的排序在最坏情况下需要 $\Omega(n\lg n)$ 的时间，因此就可以用反证法来说明题中的结论：

反证：假设某个基于比较的算法从 n 个元素的任意系列中构造一棵二叉搜索树，在最坏情况下所需要的时间小于 $\Omega(n\lg n)$。

设想这样的实例：有一个含 n 个元素的任意系列待排序，就可以采用这个基于比较的算法，先将其构造成一棵二叉搜索树，所需的时间为 $T(n) < cn\lg n$ ，构建完成后再对其进行中序遍历，立刻得到排序好的数组。这整个过程是一个基于比较的排序模型，但这个模型完成对 n 个数的排序所需要的最坏时间 $T(n) + O(n) < cn\lg n$，与已知结论相违。假设不成立。因此任何基于比较的算法从 n 个元素的任意序列中构造一棵二叉搜索树，其最坏情况下需要 $\Omega(n\lg n)$ 的时间

3.4 证明：在一棵高度为 h 的二叉搜索树上，无论从哪个结点开始，k 次连续的 TREE − SUCCESSOR 调用所需时间为 $O(k+h)$.

解：TREE − SUCCESSOR 过程实现返回节点的后继，即将节点值按大小排序后的后一个节点值。

假设开始的结点为 X，结束的结点为 Y。

① 若 X 和 Y 在同一条路径上，即 X 是 Y 的祖先或者 Y 是 X 的祖先,此时只需要沿着这条路访问，一共访问 k 条边，所需时间为 $O(k)$。

② 若 X 和 Y 不在同一路径上；则令 A 是 X 和 Y 的最小公共祖先，则从 X 到 Y 过程必经过 A。从 X 到 A 和 A 到 Y 的过程的时间必不大于 $2h$，即 $O(h)$。另外的在 X 到 A 和 A 到 Y 经过的分支的时间复杂度之和为 $O(k)$。

所以总时间复杂度为 $O(h+k)$。