**数据结构研讨课报告（2）**

计算机一班

1652228

王哲源

这次研讨课主要是关于前段时间所学的树的实际应用讨论。讨论问题如下：

1、二叉排序树的建立与遍历

2、树形结构在文件管理中的应用

3、从大量有重复，以文本文件形式进行存储的短信中，找出重复最多的10条短信

结合讨论结果，我对问题的回答如下：

1. 根据二叉排序树的特性，即对每一课子树，其根节点永远比其左子树的所有元素大，而比其右子树的所有元素小，可才用递归建树的方式，即顺序遍历所有元素，当访问某一元素时，从根节点向下递归，根据该元素与当前子树的根节点元素大小决定向左子树或右子树进行遍历，直到访问叶子节点为止，再进行插入操作。

除此之外，考虑到最坏情况下可能使二叉排序树退化为链使得树的特性退化，因此可以采用在二叉排序树上的变种例如红黑树，伸展树等进一步优化

而根据其特性，分别才用中序遍历与逆中序遍历便可以获取递增/递减的数列

1. （1）文件夹和文件的区分可以通过对树中每个节点记录一个标记tag用于区分

（2）通过（1）的tag，可以自下而上进行递归统计，即父节点的文件/文件夹个数=Σ（子节点下的文件/文件夹个数+1），通过该遍历可以实现统计

（3）文件/文件夹的删除即直接将以其为根的子树删除即可。复制则是通过遍历其所在子树进行拷贝，将新子树直接接到目标父节点上。移动则跳过拷贝部分，直接将子树移动到目标父节点上即可

（4）由于（1）的tag可以对文件和文件夹进行区分，因此获取路径也可以通过tag进行获取。而对于给定路径，只需要在每棵子树的所有孩子中寻找与文件夹/文件名匹配的即可

1. 该题可以采用两种方法，一种是利用哈希表，对所有的文本进行哈希处理，丢入哈希表进行统计，最后遍历哈希表寻找出现次数最多的文本

还有一种方法即是利用trie树，将文本拆为一棵树上的节点，文本结尾所对的节点打上tag表示其为某一条文本的结尾，并于此进行计数统计，最后寻找统计最高的前十条即可。

最后是关于研讨课我个人的想法：

随着学习知识的深入，这次研讨课的讨论思路与程度相比上一次有了很明显的提高。同时这次接触到的后两个问题都是比较实际应用的问题，也是树形结构在现实中一些比较广泛的应用举例，对于树形结构知识的加深有很大的帮助。