查找

一.定义

查找表是由同一类型的数据元素构成的集合（元素之间没有本质联系）

查找是根据给定的某个值，在查找表中确定一个其关键字等于给定值的数据元素

二.查找的操作

1 静态查找

查询某个特定的数据是否在查找表中

检索某个特定元素的各种属性

2 动态查找

在查找表中插入一个数据元素

在查找表中删除某个数据元素

三.查找表中的关键字

数据元素在某个数据项的值，用以标示一个数据元素

主关键字：可以唯一标示的一个数据元素的关键字

次关键字：可以识别的不止一个元素的关键字

四.顺序表查找

从线性表中的第一个(或最后一个)数据元素开始，逐个进行数据元素关键字和给定值的比较

若某个数据元素的关键字和给定值相等则查找成功；如果直到最后一个(或第一个)数据元素

都与给定值不匹配，则查找失败

五.二分查找

基本思想

1首先将查找表进行排序

2取中间数据元素进行比较

当给定值与中间数据元素关键字相等时，查找成功

当给定值小于中间元素时，在中间元素左区间进行二分查找

当给定值大于中间元素时，在中间元素右区间进行二分查找

3 当任意区间均无记录时，查找失败

六.插值查找

根据要查找的关键字Key与查找表中最大最小记录的关键字比较的查找方法

核心为：

mid = low + f(key)(high-low)

f(x) = (x-a[low])/(a[high]-a[low])

七.线性索引查找

1.综述

索引是把一个关键字与它对应的记录相关联的过程

索引是由若干个索引项组成

每个索引项包括关键字和其对应记录的位置

线性索引

将索引项组织成线性结构

非线性索引

将索引项组织成树形结构或图形结构

2.稠密索引

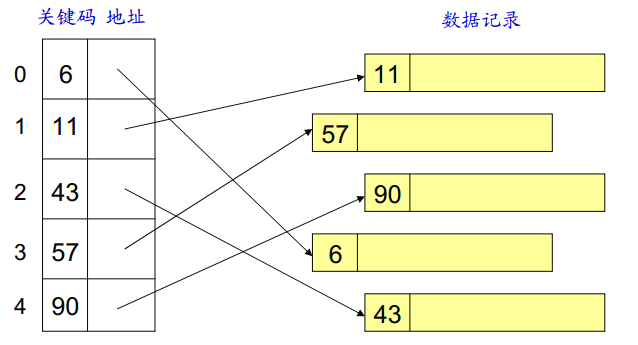
基本思想

1将数据记录的(关键码，地址)组织成索引表，并排序

2当需要通过某个关键字Key的值进行查找时，可以利用二分查找法或

插值查找等算法对（关键码，地址）表进行查找

3 通过查找结果中的地址可以直接得到数据记录



特点

不需要对数据记录本身进行组织或者排序

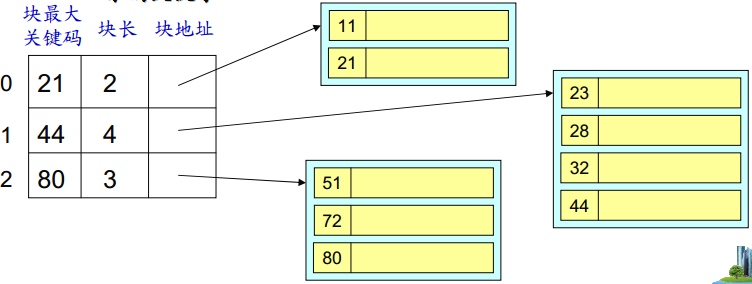
每一条数据记录对应索引表中的一个索引项

3.线性索引

将数据集中的记录分成若干块

块内无序：每一块中的记录可以是无序的

块间有序：第n块中所有记录的关键字必须大于第n-1块中任一记录的关键字



分块基本思想

以(最大关键码，块长，块指针)的方式组织索引表，并排序

利用二分查找或者插值查找等在索引表中查找关键字所在块

根据块首指针找到对应的块，并在块中查找数据记录

分块索引特点

不需要为每个数据记录分配一个索引项

块中的数据记录可以是杂乱的，只要能进行顺序查找即可

块间的数据记录必须是有序的

第n块所有记录的关键字必须大于第n-1块中任一记录的关键字

倒排索引

倒排索引将非关键码组织成（非关键码，地址）索引表

倒排索引的索引表结构类似稠密索引

倒排索引与稠密索引的区别在于

倒排索引的每个索引项可以对应多个记录

倒排索引的索引项不会随着数据记录的增加而有明显的增加

线性索引是线性表和排序算法的和查找算法的组合应用

1首先将索引项组织到线性表中

2其次对索引表进行排序

3最后对排序好的索引表进行查找得到数据记录

八.二叉排序树

1.定义

二叉排序树是一棵空树，或者是具有以下性质的二叉树

1）如左子树不空，则左子树上所有结点的值均小于或等于它的根的结点的值

2）如右子树不空，则右子树上所有结点的值均大于或等于它的根结点的值

3）左,右子树也分别为二叉排序树

2.插入：

插入操总是在叶结点处进行

3.删除：

叶结点：直接删除

非叶结点：查找合适的替代者或者后删除

二叉排序树的所有操作都必须保证其二叉排序性不变

九.哈希

1.哈希含义

在数据元素的存储位置和它的关键字之间建立一个映射关系f

通过f可以直接得到关键字所代表的数据元素

2.哈希表

哈希技术中用于存储数据元素的数据结构

3.哈希函数

哈希技术中的映射关系f