# 查找

* 1. 概述

平均查找长度：ASL

|  |  |  |
| --- | --- | --- |
| 算法 | ASL | 时间复杂度 |
| 顺序查找 | (n+1)/2 | O(n) |
| 二分查找 | lg(n+1) | O(lgn) |
| 分块查找 |  |  |
| 二叉排序树上的查找 | lg(n+1) 到(n+1)/2 | O(lgn)到O(lgn) |
|  |  |  |

* 1. 分块查找

以增加空间复杂度为代价(存储每块中的最大值已经最大值的位置)，为原数组做一个索引(索引本身是递增有序的)，这样先查索引，再查块内位置。如果索引的选择科学有效，则可以获得比顺序查找快的速度

平均查找长度在顺序查找和二分查找之间，并且当结点数为元素数量的平方根时，查找长度最小。一般用待查的关键字对索引关键字进行二分查找

* 1. 二叉排序树上的查找

二叉排序树上的平均查找长度取决于二叉树的深度。和二叉排序树的形态有关。在极端情况下，二叉树只有单一的左或右分支，则查找长度为(n+1)/2；如果是平衡二叉树，则查找长度为lgn(树的层次)。

* 1. 散列技术下的查找

  将元素的值和其位置直接对应，对应的方法就是散列函数(如平方取中，直接定址法（自身函数法），数字分析法，折叠法，除余法等等)；然而再好的散列函数也会引起冲突，

* + 1. 则解决冲突的方法
* 拉链法
* 开放地址法

发生冲突时，按照某种方法继续探测哈希表中其他存储单元。

H i ( key ) = ( H ( key )+ d i ) mod m ( i = 1,2,…… ， k ( k ≤ m – 1))

增量 d 可以有不同的取法，并根据其取法有不同的称呼：   
（ 1 ） d i ＝ 1 ， 2 ， 3 ， …… 线性探测再散列；   
（ 2 ） d i ＝ 1^2 ，－ 1^2 ， 2^2 ，－ 2^2 ， k^2， -k^2…… 二次探测再散列；   
（ 3 ） d i ＝ 伪随机序列 伪随机再散列；

# 内部排序

* 1. 总结

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 算法 | 时间复杂度 | | | 空间 | 稳定性 |
| 平均 | 最坏 | 最好 |
| 直接插入 | O（n^2） | O（n^2） | O（n） | O（1） | 稳定 |
| 希尔 | O（n^1.5） | O（n^2） | O（nlgn） | O（1） | 不稳定 |
| 简单选择 | O（n^2） | O（n^2） | O（n^2） | O（1） | 不稳定 |
| 堆排序 | O（nlgn） | O（nlgn） | O（nlgn） | O（1） | 不稳定 |
| 冒泡排序 | O（n^2） | O（n^2） | O（n） | O（1） | 稳定 |
| 快速排序 | O（nlgn） | O（n^2） | O（nlgn） | O（nlgn） | 不稳定 |
| 归并排序 | O（nlgn） | O（nlgn） | O（nlgn） | O（n） | 稳定 |
| 基数排序 | O(d(r + n)) | O(d(r + n)) | O(d(r + n)) | O（n＋r） | 稳定 |

基数排序的复杂度中，r代表关键字的基数，d代表长度，n代表关键字的个数

* 1. 插入排序
     1. 直接插入排序
     2. 希尔排序

希尔排序是1959 年由D.L.Shell 提出来的，相对直接排序有较大的改进。希尔排序又叫缩小增量排序

* 基本思想：

先将整个待排序的记录序列分割成为若干子序列分别进行直接插入排序，待整个序列中的记录“基本有序”时，再对全体记录进行依次直接插入排序。

* 操作方法：
* 选择一个增量序列t1，t2，…，tk，其中ti>tj，tk=1；
* 按增量序列个数k，对序列进行k 趟排序；
* 每趟排序，根据对应的增量ti，将待排序列分割成若干长度为m 的子序列，分别对各子表进行直接插入排序。仅增量因子为1 时，整个序列作为一个表来处理，表长度即为整个序列的长度。
* 时间复杂度

希尔排序时效分析很难，关键码的比较次数与记录移动次数依赖于增量因子序列d的选取。增量因子中除1 外没有公因子，且最后一个增量因子必须为1。

希尔排序的时间复杂度是：O（nlogn）～O（n2），平均时间复杂度大致是O(n√n)

* 1. 选择排序
     1. 简单选择排序
     2. 二元选择排序

在比较的时候同时选出最大值和最小值

* + 1. 堆排序
  1. 交换排序
     1. 冒泡排序
     2. 冒泡排序的改进
* 加入一标志性变量exchange，用于标志某一趟排序过程中是否有数据交换，如果进行某一趟排序时并没有进行数据交换，则说明数据已经按要求排列好，可立即结束排序
* 在每趟排序中进行正向和反向两遍冒泡的方法一次可以得到两个最终值
  + 1. 快速排序

快排又称分区交换排序，是冒泡排序的一种改进。

快速排序是通常被认为在同数量级（O(nlog2n)）的排序方法中平均性能最好的。但若初始序列按关键码有序或基本有序时，快排序反而蜕化为冒泡排序。为改进之，通常以“三者取中法”来选取基准记录，即将排序区间的两个端点与中点三个记录关键码居中的调整为支点记录。

* 1. 归并排序
  2. 基数排序