<http://www.jianshu.com/p/42f81846c0fb>

# 概念

## 算法

算法是指解题方案的准确而完整的描述，是一系列解决问题的清晰指令，算法代表着用系统的方法描述解决问题的策略机制。对于同一个问题的解决，可能会存在着不同的算法，为了衡量一个算法的优劣，提出了空间复杂度与时间复杂度这两个概念。

一个算法是由**控制结构**（顺序、分支和循环3种）和**原操作**（指固有数据类型的操作）构成的，则算法时间取决于两者的综合效果。为了便于比较同一个问题的不同算法，通常的做法是，**从算法中选取一种对于所研究的问题（或算法类型）来说是基本操作的原操作，以该基本操作的重复执行的次数作为算法的时间量度。**

## 时间复杂度

<http://blog.csdn.net/zolalad/article/details/11848739>

时间频度 一个**算法执行所耗费的时间**。

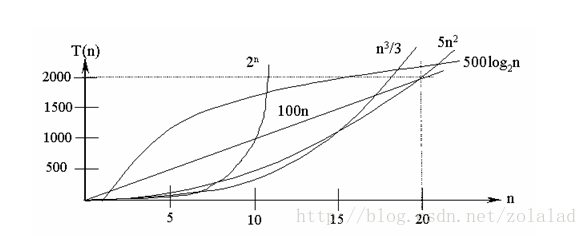
一个算法花费的时间与算法中语句的执行次数成正比例，哪个算法中语句执行次数多，它花费时间就多。一个算法中的语句执行次数称为语句频度或时间频度。记为T(n)

n称为问题的规模，当n不断变化时，时间频度T(n)也会不断变化。一般情况下，算法中基本操作**重复执行的次数**是**问题规模n**的某个函数，用T(n)表示，若有某个辅助函数f(n),使得当n趋近于无穷大时，T(n)/f(n)的极限值为不等于零的常数，则称f(n)是T(n)的同数量级函数。记作T(n)=Ｏ(f(n)),称Ｏ(f(n)) 为算法的渐进时间复杂度，简称时间复杂度。

简单来说，就是T(n)在n趋于正无穷时最大也就跟f(n)差不多大。

**算法中语句执行次数为一个常数，则时间复杂度为O(1)**。常见的时间复杂度有：常数阶O(1),对数阶O(log2n),线性阶O(n), 线性对数阶O(n log2n),平方阶O(n2)，立方阶O(n3),...。

Log28：2为底N的**对数**，即2的几次方等于8，结果为3



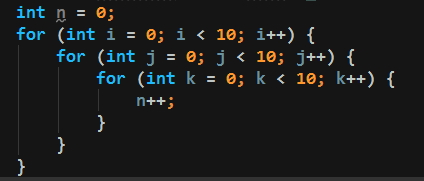
常见的算法时间复杂度由小到大依次为：Ο(1)＜Ο(log2n)＜Ο(n)＜Ο(n log2n)＜Ο(n2)＜Ο(n3)＜…＜Ο(2n)＜Ο(n!)

常数阶 < 对数阶 < 线性阶 < 线性对数阶 < 平方阶 < 立方阶 < … < 指数阶 < 阶乘

### 求解算法的时间复杂度的具体步骤

1. 找出算法中的基本语句；算法中**执行次数最多的那条语句**就是基本语句，通常是**最内层循环的循环体。**
2. 计算基本语句的**执行次数的数量级**；只需计算基本语句执行次数的数量级，这就意味着只要保证基本语句执行次数的函数中的最高次幂正确即可，可以忽略所有低次幂和最高次幂的系数。这样能够简化算法分析，并且使注意力集中在最重要的一点上：增长率。
3. 用大Ο记号表示算法的时间性能。

如：



第一个for循环的时间复杂度为Ο(n)，第二个for循环的时间复杂度为Ο(n2)，则整个算法的时间复杂度为Ο(n1+n2+n3)=Ο(n3)。

Ο(1)表示基本语句的执行次数是一个常数，一般来说，只要算法中不存在循环语句，其时间复杂度就是Ο(1)。其中**Ο(log2n)、Ο(n)、 Ο(nlog2n)、Ο(n2)和Ο(n3)称为多项式时间**，而Ο(2n)和Ο(n!)称为指数时间。计算机科学家普遍认为前者（即多项式时间复杂度的算法）是有效算法，

一个经验规则：其中c是一个常量，如果一个算法的复杂度为c 、 log2n 、n 、 n\*log2n ,那么这个算法时间效率比较高 ，如果是2n ,3n ,n!，那么稍微大一些的n就会令这个算法不能动了，居于中间的几个则差强人意。

## 空间复杂度

<http://blog.csdn.net/zolalad/article/details/11848739>

类似于时间复杂度的讨论，一个算法的空间复杂度(Space Complexity)S(n)定义为**该算法所耗费的存储空间**，它也是问题规模n的函数。

包括存储算法本身所占用的存储空间

算法的输入输出数据所占用的存储空间

算法在运行过程中临时占用的存储空间这三个方面。

当一个算法的空间复杂度为一个常量，即不随被处理数据量n的大小而改变时，可表示为O(1)；

# 数据结构

## 线性结构

### 数组

### 链表

### 队列

### 栈

## 非线性结构

### 树

#### 二叉查找树

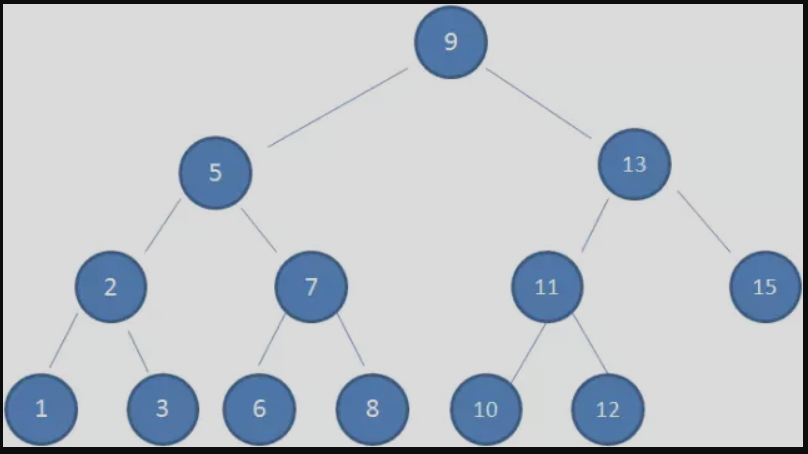
<http://mp.weixin.qq.com/s?__biz=MzIxNTIwOTgxNw==&mid=2650612655&idx=1&sn=09bd44c8bf9978c8beedcdf5a07d1047&chksm=8f924199b8e5c88f45a0ffc8571c820d830c4a332a2c16f162d48295e70ea796ebabc4773fdf&mpshare=1&scene=23&srcid=1125u1TfchKscPuptKpdLlMC#rd>

**特性**

**1. 左子树上所有结点的值均小于或等于它的根结点的值。**

**2. 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。**



问：找出值为10的节点

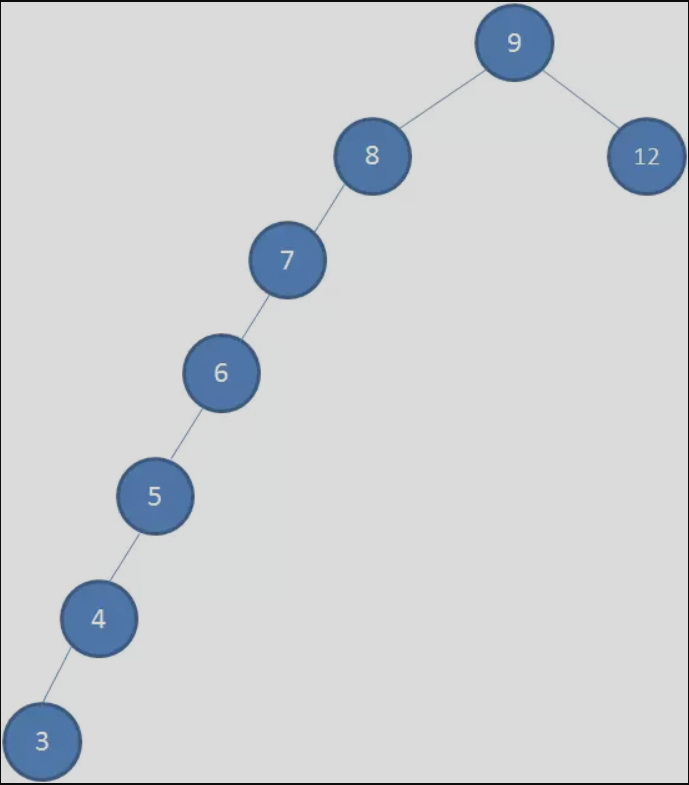
先看9,9小于10，根据二叉查找树的特性，10应该在节点9的右节点子树上。

依次查看9-13-11->9，可找到9所在节点。

这种方式正是二分查找的思想，查找所需的最大次数等于二叉树的高度。插入节点时也是基于该特性。

**缺陷**

依次插入如下五个节点：7,6,5,4,3。依照二叉查找树的特性，结果会变成什么样呢？



虽然符合二叉查找数的特性，但查找效率大打折扣，几乎变成线性。

为了**解决二叉查找树多次插入新节点导致的不平衡**，红黑树应运而生。

#### 红黑树

自平衡的二叉树，符合二叉查找树的基本特性，此外还具有如下特性：

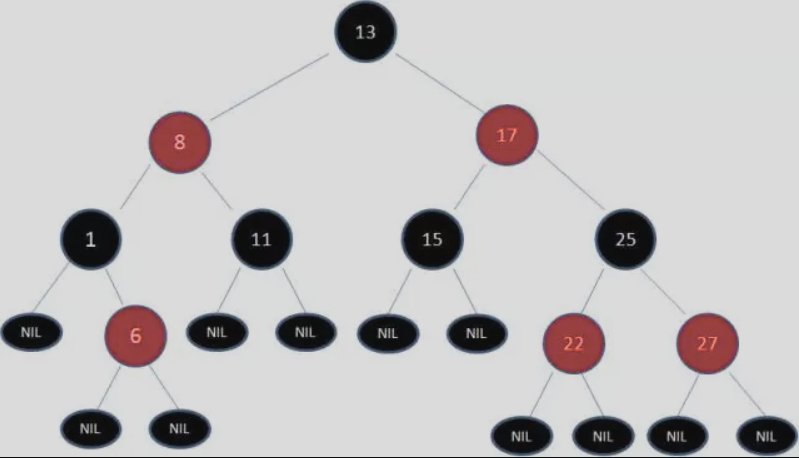
**1. 节点是红色或黑色。**

**2. 根节点是黑色。**

**3. 每个叶子节点都是黑色的空节点（NIL节点）。**

**4. 每个红色节点的两个子节点都是黑色。(从每个叶子到根的所有路径上不能有两个连续的红色节点)**

**5. 从任一节点到其每个叶子的所有路径都包含相同数目的黑色节点。**

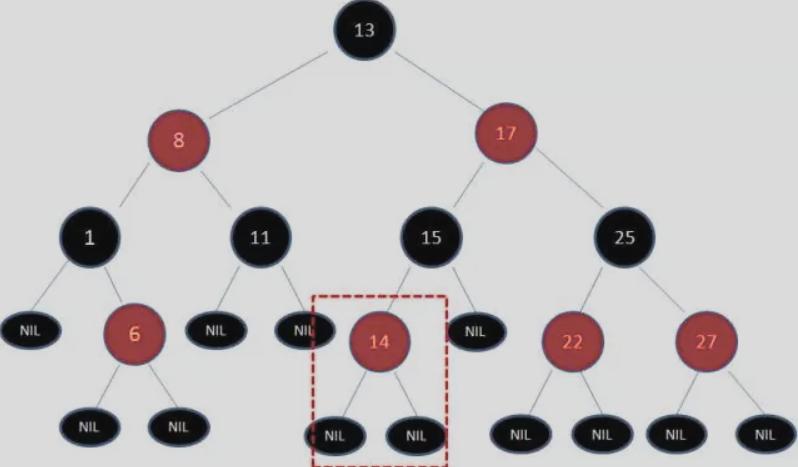


查找值为10的节点：13-8-11-NULL

红黑树从根到叶子的最长路径不会超过最短路径的2倍。

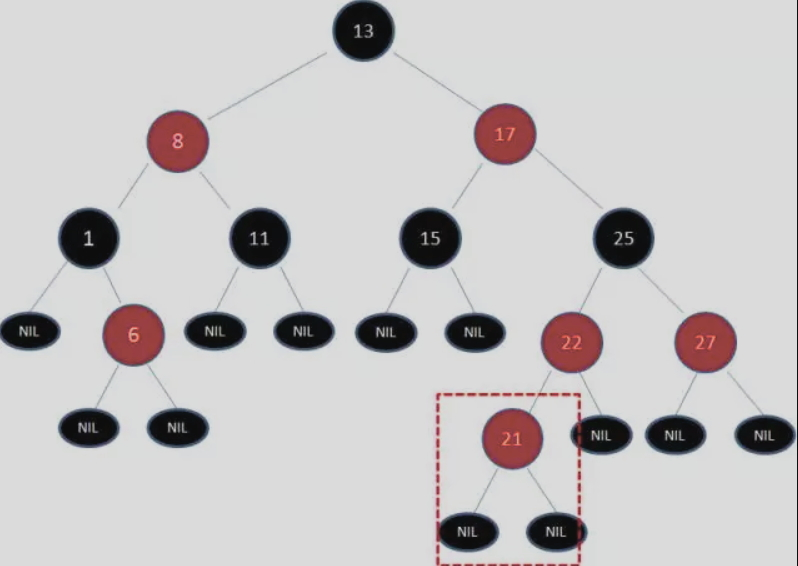
当插入或删除节点时，红黑树的规则有可能被打破，此时需要做出调整，来继续维持红黑树的规则。

1. 向原红黑树插入值为14的新节点：



由于父节点15是黑色节点，因此这种情况并不会破坏红黑树的规则，无需做任何调整。

1. 向原红黑树插入值为21的新节点：



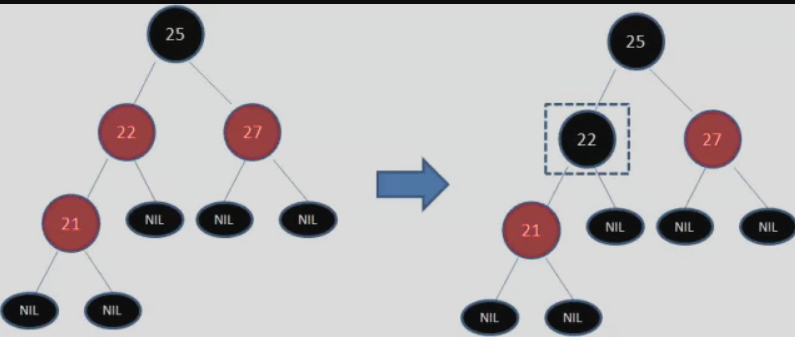
由于父节点22是红色节点，因此这种情况打破了红黑树的规则4（每个红色节点的两个子节点都是黑色），必须进行调整，使之重新符合红黑树的规则。

调整有两种方法：**变色**和**旋转**（旋转又分左旋转和右旋转）。

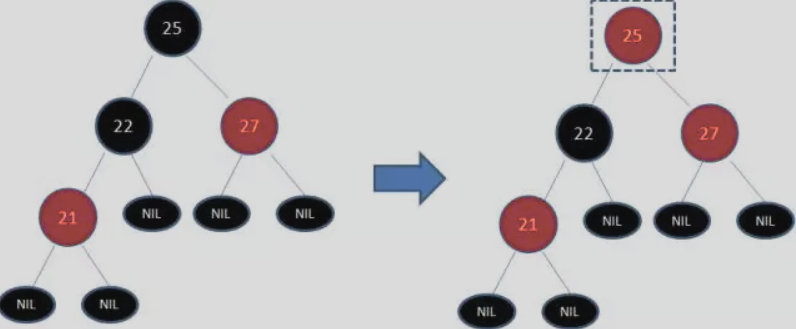
* 变色

为了重新符合红黑树的规则，尝试把红色节点变为黑色，或者把黑色节点变为红色。

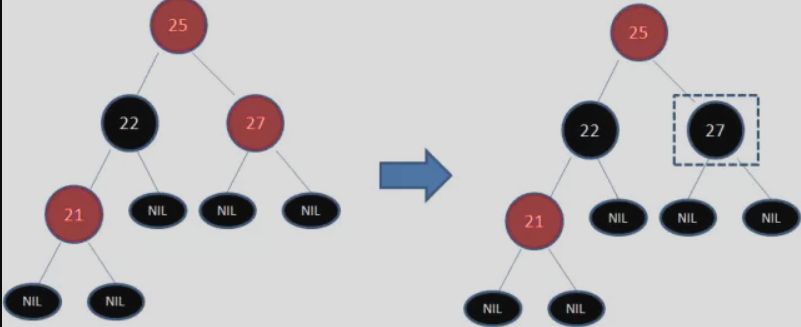
下图所表示的是红黑树的一部分，需要注意节点25并非根节点。因为节点21和节点22连续出现了红色，不符合规则4，所以把节点22从红色变成黑色：



但这样并不算完，因为凭空多出的黑色节点打破了规则5，所以发生连锁反应，需要继续把节点25从黑色变成红色：



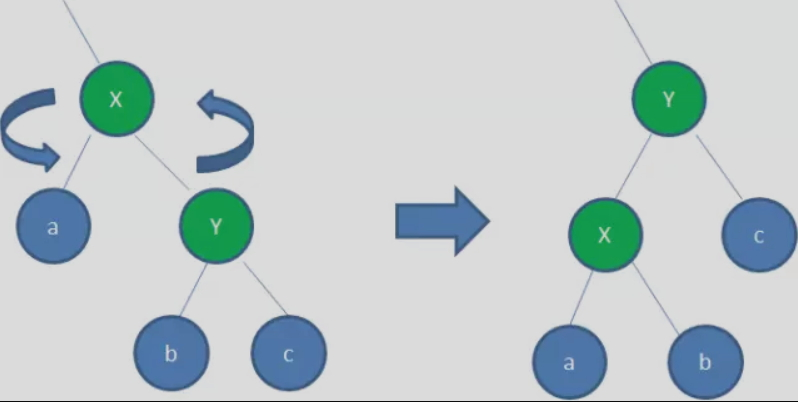
此时仍然没有结束，因为节点25和节点27又形成了两个连续的红色节点，需要继续把节点27从红色变成黑色：



* 旋转

左旋转

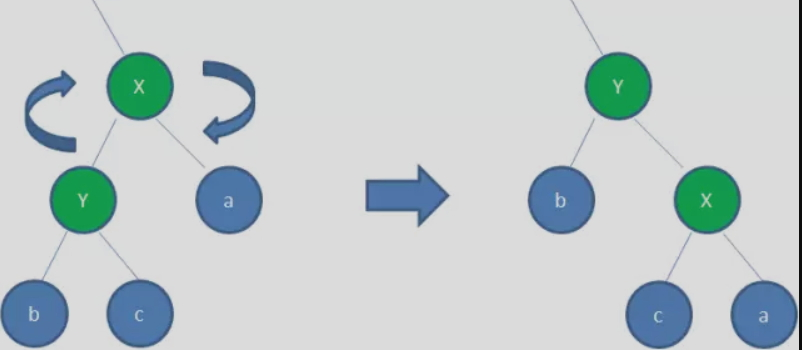
逆时针旋转红黑树的两个节点，使得父节点被自己的右孩子取代，而自己成为自己的左孩子。说起来很怪异，大家看下图：



图中，身为右孩子的Y取代了X的位置，而X变成了自己的左孩子。此为左旋转。

右旋转

顺时针旋转红黑树的两个节点，使得父节点被自己的左孩子取代，而自己成为自己的右孩子。大家看下图：



图中，身为左孩子的Y取代了X的位置，而X变成了自己的右孩子。此为右旋转。

实际应用

JDK集合类，TreeMap和TreeSet底层用的是红黑树，java8中HashMap也用到了红黑树。

### 图

### 表

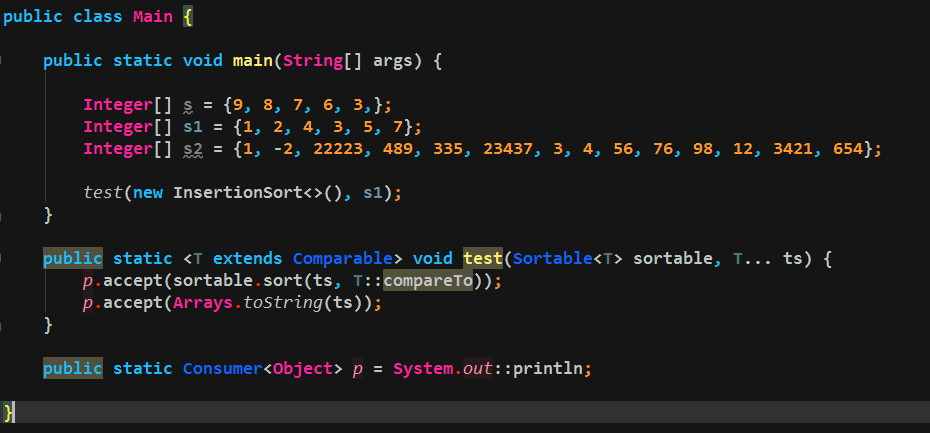
## 哈希结构

# 算法

## 排序

<http://www.jianshu.com/p/42f81846c0fb>

测试方法



### 冒泡排序

<https://zh.wikipedia.org/wiki/%E5%86%92%E6%B3%A1%E6%8E%92%E5%BA%8F>

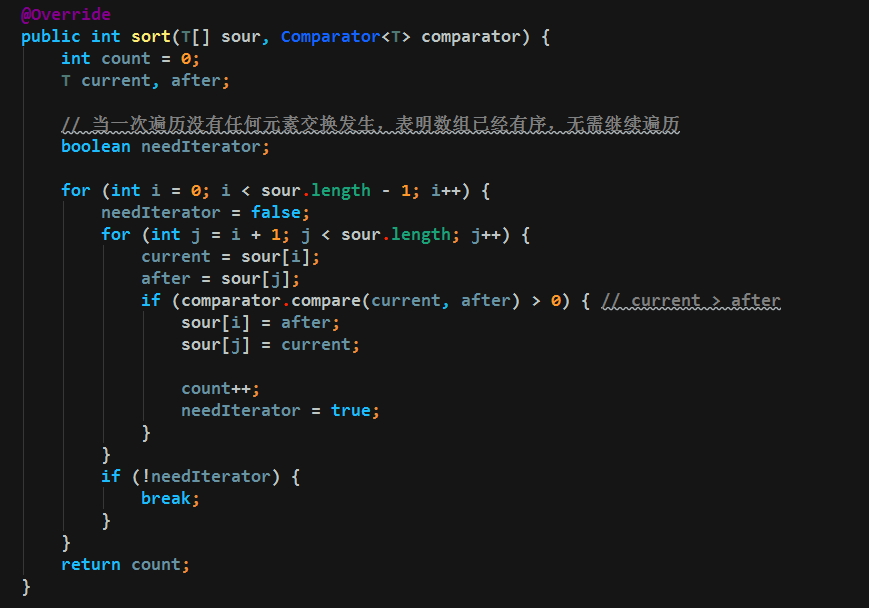
时间复杂度O(n2)，最好情况O(n)，最坏情况O(n2)，稳定。

适合数据规模很小的时候，而且它的效率也比较低。

通过重复地走访要排序的数列，一次比较两个元素，如果他们的顺序错误就把他们交换过来。走访数列的工作是重复地进行直到没有再需要交换，也就是说该数列已经排序完成。

**算法描述如下：**

1. 比较相邻的元素。如果第一个比第二个大，就交换他们两个。
2. 对每一对相邻元素作同样的工作，从开始第一对到结尾的最后一对。这步做完后，最后的元素会是最大的数。
3. 针对所有的元素重复以上的步骤，除了最后一个。
4. 持续每次对越来越少的元素重复上面的步骤，直到没有任何一对数字需要比较。



### 简单选择排序

### 插入排序

<https://zh.wikipedia.org/wiki/%E6%8F%92%E5%85%A5%E6%8E%92%E5%BA%8F>

时间复杂度O(n2)，最好情况O(n)，最坏情况O(n2)，稳定。（同冒泡）

是一种简单直观的排序算法。适用于量级小于千，或者若已知输入元素大致上按照顺序排列。

插入排序不适合对于数据量比较大的排序应用。

工作原理是通过构建有序序列，对于未排序数据，在已排序序列中从后向前扫描，找到相应位置并插入。

**算法描述如下：**

1. 从第一个元素开始，该元素可以认为已经被排序
2. 取出下一个元素（第二个元素）作为目标元素，在已经排序的元素序列中**从后向前**扫描
3. 如果当前元素（扫描到的元素）大于目标元素，将当前元素与目标元素交换
4. 重复步骤3，直到找到目标元素的位置（当前元素小余或等于目标元素）
5. 重复步骤2-4

先对数组前两个元素进行排序，之后将第三个元素与前两个比较，按排序规则插入正确的位置，之后将第四个元素与前三个，比较插入正确的位置，以此类推直至最后一个元素

如果比较操作的代价比交换操作大的话，可以采用二分查找法来减少比较操作的数目。该算法可以认为是插入排序的一个变种，称为**二分查找插入排序**。



### 归并排序

### 希尔（Shell）排序

### 快速排序

### 堆排序



## 查找

### 顺序查找

### 二分查找

## Dijkstra （最短路径算法）

## 递归

## 分治算法

## 动态规划

## 贪心算法

## 回溯算法

## 匹配算法

## 正则表达式和字符串匹配

# 习题

<http://www.codeceo.com/article/15-algorithms-question.html>

<http://dongxicheng.org/structure/structure-algorithm-summary/>