# 数据结构与算法

参考：

1. 《数据结构与抽象 Java语言描述 第4版》
2. 《数据结构与问题求解（Java语言版）》
3. 《数据结构与算法分析 Java语言描述》
4. 《数据结构与算法 python语言描述》
5. 《数据结构（C语言版）》
6. 《算法设计与分析-王晓东》
7. 《算法图解》
8. 《算法导论（原书第3版）》
9. 《算法（第4版）》
10. 《[编程珠玑和续](https://book.douban.com/subject/34855884/" \o "编程珠玑续)》乔恩·本特利（Jon Bentley）
11. 《剑指offer》《编程之美》《程序员面试金典（第6版）》

数据结构与算法 1

1 位运算 2

2 数组和顺序表 3

3 字符串 3

4 链表 3

5 栈 3

6 队列 4

7 散列表 4

8 树 5

9 图 5

10 复杂度 6

11 排序 6

12 查找 6

13 动态规划 7

14 贪心算法 7

15 摊还分析 7

16 回溯算法 7

17 分治递归算法 7

18 枚举算法 8

19 分支限界 8

20 概率算法 8

21 矩阵运算 8

22 多项式与快速傅里叶变换 8

23 数论算法 8

24 线性规划 8

25 集合算法 8

26 计算几何学 8

27 数值计算 9

28 NP完全性理论与近似算法 9

29 算法优化策略 9

30 在线算法设计 9

## 位运算

n = (n - 1) & n:最后一位1变为0

右移运算符代替除以2；

用位与代替判断奇偶数；

## 数组和顺序表

动态数组-顺序表-线性表

顺序存储 随机存取

堆内存上创建数组，栈内存中保存了引用地址；堆内存中保存了具体内容；

三对角矩阵

矩阵的压缩存储 广义表

有序表：数组、链表、继承、组合、抽象基类实现方式

迭代器：

## 字符串

BF算法（普通匹配算法）：时间复杂度O(m\*n)；空间复杂度O(1)

KMP算法：时间复杂度O(m+n)；空间复杂度O(n)；next数组是什么？

Boyer-Moore字符串查找算法

字符串从后往前复制；

BM、

Trie树

Rabin&Karp算法，多子串搜索与AC自动机，

字符串排序，单词查找树

正则表达式: 利用有限自动机进行字符串匹配

数据压缩: 霍夫曼压缩

## 链表

单链表

双向链表

循环链表

双向循环链表

复杂链表；

链式存储 顺序存取

链表双指针

链表头指针前弄一个指针

## 栈

1. 检测中缀表达式是否平衡
2. 中缀表达式转为后缀表达式：操作符栈
3. 计算后缀表达式的值：数字串
4. 计算中缀表达式的值：维护两个栈
5. 程序栈、栈溢出、递归问题（将问题化为更小的同类型问题去解决，可多次递归，递归处理链表等，计算幂指数，递推求时间复杂度并用数学归纳法证明，汉诺塔问题，不要在递归调用中重复解决同一个问题斐波那契数列，最后一个动作是递归调用时发生尾递归）、间接递归和相互递归、用栈替代递归）

实现方式：数组（静态数组，动态数组）实现，链表实现，向量实现；

## 队列

普通队列：循环数组或者链表实现，带尾指针的链表

阻塞队列

并发队列

双端队列：双向链表，循环链表

优先队列：用堆实现，数组或者链表实现（把优先级最高的排容易操作的位置）

可用于解决广度优先问题：

## 散列表

字典：实现方式（链表[一条链有两种实现方式、两条链]、数组[kv作为一个数组放一起]），有序字典和无序字典（散列表实现）；

集合：

散列技术：如何存储和查找；

（1）选择一个Hash函数，使得每个键字能有一个唯一的地址（散列地址，存储位置）；要求计算简单；分布均匀；hashcode，不相等的对象可能有相同的散列码；

0直接定址法

①数字分析法

②平方取中法

③分段叠加法

④除留取余法

5.全域散列法

要考虑：

关键字的分布；

关键字的长度；

散列表的大小；

哈希函数计算量；

（2）当不同键字具有相同的地址，此时选择一种冲突处理方法至关重要：

①开放地址法（包括线性探测法、平方探测法、随机探测法）

②再散列法

③链地址法

④公共溢出区

位图

动态扩容

完全散列

## 树

二叉树：4大遍历算法，线索二叉树

平衡查找树：平衡节点

二叉查找树：创建、查找、中序遍历是排序的、添加、删除节点，操作效率O(logn)，可以实现字典；

平衡二叉查找树：AVL树、红黑树，kd树，顺序统计量，AA树，treap树

完全二叉树：

多路查找树：B树，B+树, B\*树，B-树，2-3树，2-3-4树，

伸展树

哈希树，

森林 普通树（使用二叉树转换转换表示，前序遍历一样，一般树的后序遍历与二叉树的中序遍历一样；

树的路径长度是从树根到树中每一结点的路径长度之和

最优二叉树(赫夫曼树)：

哈夫曼树的路径长度，表达式树的求值，决策树，语法树，游戏树，

堆：（最大，最小，优先队列的实现方式，二项堆，斐波那契堆，d堆，左式堆，二项队列，配对堆）

1. 使用数组实现堆, 添加项的时间复杂度O(logn)，删除根；
2. 构建堆的时间复杂度为O(n)，重建堆的时间复杂度为nlogn；
3. 堆排序的时间复杂度为nlogn：初始堆的建立，reheap()过程；

归并优先级队列（斜堆，偶堆 ）

区间树，van Emde Boas树，

## 图

两个存储：邻接矩阵（无向图，有向图，网），邻接表存储（无向图，有向图，网），

补：十字链表，邻接多重表，边集数组

稀疏图和稠密图，深度优先（双连通性，欧拉回路，查找强分支）和广度优先，

寻找路径（深度优先），无权图最短路径（广度优先），带权图最短路径（dijkstra, floyd负权），从某个源点到其余各顶点的最短路径即单源最短路径（Bellman-Ford算法，差分约束和最短路径 ），每一对顶点之间的最短路径（Floyd-Warshall算法，用于稀疏图的Johnson算法）

连通性问题：无向图的连通分量和生成树，有向图的强连通分量，最小生成树（prim kruscal），

有向无环图：拓扑排序（邻接矩阵存储O(n^3)，邻接表O(n+e)）；关键路径

最大流，网络流问题：流网络，Ford\Fulkerson方法，最大二分匹配，推送重贴标签算法，前置重贴标签算法

AOV/AOE网拓扑排序和关键路径

## 复杂度

空间复杂度

时间复杂度：最好、最坏、平均

主方法和主定理

## 排序

稳定排序（不改变相等对象的相对次序）：插入；冒泡；归并；基数（多关键字的排序，链式基数排序）；

外部排序（多路平衡归并排序，最佳归并树，多相归并，置换选择）

时间复杂度：

O(n^2):冒泡（垃圾） 选择（递归和非递归，半垃圾） 插入（递归和非递归，对链表可行，一般） 希尔（大数组，ok）

O(nlogn):归并（分治，常常用递归表示，也可以迭代归并，对象数组java.util是归并，总是nlogn, 消耗空间） 快排（分治，平均nlogn,创建划分有不同的策略，基本类型数组java.util是快排，插入排序处理量小的情况，三元中指枢轴选择，随机化版本） 堆排

O(n+k):基数（O(n\*d)） 计数 桶排序

空间复杂度：

logn: 快排

n：归并

中位数和顺序统计量：

分配排序

## 查找

七大查找：

静态查找：顺序表的查找；有序表的查找；静态树表的查找；索引顺序表的查找；插值查找；

动态查找：二叉排序树和平衡二叉树，B树和B+树，键树

哈希表查找

线性索引查找（稠密，分块，倒排）

块内无序；块间有序；索引存取；索引表；

斐波那契查找；

<http://blog.csdn.net/xuezhisdc/article/details/52432325>

## 动态规划

将待求解问题分解为若干子问题，先求解子问题，然后从这些子问题的解得到原问题的解。这些子问题不独立，需要保存已解决的子问题的答案，在需要时再找出来，从而避免大量重复计算，得到了多项式时间算法。动态规划用于求解最优问题，有下面特点：

最优子结构性质：

子问题重叠性质：

备忘录方法：

矩阵连乘问题，最长公共子序列，凸多边形三角剖分，多边形游戏，图像压缩，电路布线，流水作业调度，0|1背包问题，最优二叉搜索树，钢条切割，所有点对最短路径

后缀数组与最长公共子串（后缀数组的基本概念，构造后缀数组的倍前缀算法，构造后缀数组的DC3分治法，最长公共前缀数组与最长公共扩展算法，最长公共子串算法）  
序列比较算法（编辑距离算法，最长公共单调子序列，有约束最长公共子序列）

## 贪心算法

最优子结构性质：当一个问题的最优解包含其子问题的最优解时。

自顶向下，以迭代的方式做出相继的贪心选择。

贪心选择性质：所求问题的最优解可以通过一系列局部最优选择达到。

活动安排问题，装载，哈夫曼编码，单源最短路径，最小生成树，多机调度问题，任务时间表问题，集合覆盖问题，拟阵和贪心算法，用拟阵求解任务调度问题

## 摊还分析

聚合分析，核算法，势能法

二项队列，斜堆，斐波那契堆，伸展树

## 回溯算法

深度优先方式系统搜索问题解，搜索过程用剪枝函数避免无效搜索。搜索整个解空间。

回溯法的算法框架：递归回溯；迭代回溯；子集树与排列树

装载问题，批处理作业调度，符号三角形问题，n后问题，01背包问题，最大团问题，图的m着色问题，旅行售货员问题，圆排列问题，电路板排列问题，连续邮资问题，收费公路重建问题，博弈

## 分治递归算法

将难以解决的大问题分割成较小规模的问题，这些问题相互独立且与原问题相同。然后各个击破，分而治之，并将子问题的解合并得到原问题的解

用代入法求解递归式 ，用递归树方法求解递归式

大整数的乘法，Strassen矩阵乘法，棋盘覆盖，合并排序，快速排序，最接近点对问题，循环赛日程表，最大子数组问题，

## 枚举算法

## 分支限界

以广度优先方式搜索解空间树。解决离散化最优问题

单源最短路径问题，装载问题，布线问题，01背包问题，最大团问题，旅行售货员问题，电路板排列问题，批处理作业调度184

## 概率算法

随机数，数值概率算法（用随机投点法计算π值，计算定积分，解非线性方程组），舍伍德算法（线性时间选择算法，跳跃表，确定性跳跃表），拉斯维加斯算法（n后问题，整数因子分解），蒙特卡罗算法（蒙特卡罗算法的基本思想，主元素问题，素性测试），雇用问题，指示器随机变量，生日悖论，球与箱子，特征序列，在线雇用问题

## 矩阵运算

求解线性方程组，矩阵求逆，对称正定矩阵和最小二乘逼近，

## 多项式与快速傅里叶变换

多项式的表示，DFT与FFT，高效FFT实现

## 数论算法

基础数论概念，最大公约数，模运算，求解模线性方程，中国余数定理，元素的幂，RSA公钥加密系统，素数的测试，整数的因子分解

## 线性规划

标准型和松弛型，将问题表达为线性规划，单纯形算法，对偶性，初始基本可行解

## 集合算法

不相交集合类：快速并查找，灵巧求并，带路径压缩的按秩合并

## 计算几何学

确定任意一对线段是否相交，寻找凸包，寻找最近点对

## 数值计算

牛顿迭代

## NP完全性理论与近似算法

1. P类与NP类问题:非确定性图灵机,P类与NP类语言，多项式时间验证，NP完全性与可归约性
2. NP完全问题：多项式时间变换，Cook定理
3. 一些典型的NP完全问题：合取范式的可满足性问题，3元合取范式的可满足性问题，团问题，顶点覆盖问题，子集和问题，哈密顿回路问题，旅行售货员问题
4. 近似算法的性能：顶点覆盖问题的近似算法，旅行售货员问题近似算法，集合覆盖问题的近似算法，子集和问题的近似算法，子集和问题的指数时间算法，子集和问题的完全多项式时间近似格式，随机化和线性规划

## 算法优化策略

算法设计策略的比较与选择（最大子段和问题的简单算法，分治算法，动态规划算法）

动态规划加速原理（货物储运问题，算法及其优化）

优化数据结构（带权区间最短路问题，算法设计思想，算法实现方案，并查集，可并优先队列）

优化搜索策略

## 在线算法设计

在线算法设计的基本概念：页调度问题，势函数分析，k服务问题，Steiner树问题，在线任务调度，负载平衡