# 高级数据结构与算法

# 绪论

# 线性存储结构

## 单调队列/栈。

所谓单调队列，就是队里元素是单调递增或递减的队列,一旦入队元素使得队列不单调，就把队尾元素删除直到满足单调性后再插入那个新元素，这个实际是栈的性质，本来队列是尾入头出，而不能删除尾部。所以有的地方叫做单调栈，但因为队列支持从队头出队的功能(在滑动窗口问题有用到)，所以习惯叫单调队列，它确实不是真正的队列，它同时具有栈和队列的性质。

编码上，c++有deque的双端队列可用，也可以用来手动实现。(有人说c++库慢，但之后实测一点不慢)。但是最快的是手动实现的数组双端队列。这么做在算法竞赛中只要开够了内存(2倍于峰值)，就不需要释放。

### 模型一,找字典序：

给定一个序列，找字典序最小的子序列。从头到尾构建单调队列即可

### 模型二,滑动窗口问题：

所谓滑动窗口就是给一个长度是n的数列A,对于数列每一段长度是m的区间，都要求得一个最值，**单调递减队列**可以用来求滑动窗口的**最大值(这里习惯队尾在前，队头在后，所以单调递减队列是维护最大值)**。其算法复杂度是O(n)。这是正招，如果用堆或线段树等结构，复杂度是n\*logn，单调队列的一个用途是利用其滑动窗口最值优化动态规划问题的时间复杂度。

例如poj2823，以滑动最大值为例

用单调递减队列记录最大值，构建单调队列整个过程队列头部的数所在位置和队列尾部所在位置一定不能超过m,如果超过了就要把队列头部出队(保证了求长度是m的区间内最值)。如何知道队列每个元素位置，在队列额外维护一个属性即可。

先把A序列前m个元素按照单调队列规则入队，同时保证队列长度，处理完前m个数，队尾就是最大值(很好理解，单调队列尾永远是最大的)，之后按顺序每次入队A的一个元素A[i]，每次输出队列尾部，就是区间[i-m+1,i]最大值，同时一定保证队列长度.这里要注意边界问题，因为是区间最大值，所以单调队列出队时，相等元素也要出队，这样可以更新所存的位置变量，使其更靠后

## ST表

## 扩展习题

# 树和森林

## 线段树

## 树状数组

## 二叉查找树

AVL树，treap，fhq\_treap，伸展树 ,红黑树，替罪羊树，sbt树，2-3树，01树

## 并查集

## 可并堆

## 动态森林(LCT)

# 图

## 适应性问题

## K短路问题

## K小生成树问题

## 联通分量问题

## 网络流问题

## 二分图问题

## 公共祖先问题

## 跳舞链

# 字符串和有限状态机

## 后缀数组

后缀数组是处理字符串的有效工具

## AC自动机

Ac自动机是Aho-Corasick 提出的一种dfa

概念去维基看，简单来讲DFA是一张图，有一个开始节点，代表空串，其他各个点并没有实际意义(准确来讲部分自动机，但是不是所有的自动机都要求节点有意义)，点之间有边。

DFA的图的边是有实际意义，一条边代表了**一种字符**，比如在纯英文小写字符构建的dfa中，每个节点有26个分叉，但分叉不一定指向某个实际的点，也可能指向空，如果从起点出发，走出一条路径，那么这个路径可以代表一个字符串。

Ac自动机解决了多模式串匹配的问题。

给定6个模式串字符串

6

beta

alpha

haha

delta

dede

tata

主串是：dedeltalphahahahototatalpha

利用模式串构建ac自动机，完了用主串在ac自动机里跑一边即可。

下面说怎么构建ac自动机，每个节点维护三个属性，son[26]代表26个叉，fail代表失败指针，cnt是代表以这个点为结尾的字符串数量初始是0，注意ac自动机一个点只可能是一个字符串的结尾，所有构建的新节点son和fail指向NULL，事先构建好根节点root，其他节点构建分为三步：

1. 构建字典树，比较容易，每加入一个字符串，就遍历一遍每个字符，字典树从根开始如果没有这个叉，就建立一个新节点，并且连上，向这个方向走；如果有这个叉，就直接向这个方向走即可。走完后最后一个点的cnt+=1
2. 建立fail指针，从root利用广度优先搜索顺序建立，首先处理root，root->fail=root,遍历根节点所有存在的叉，把它们入队，并且把它们的fail连到root。处理完root，对于非root节点，开始广搜的循环，每次取出一个点，找它的所有叉now->son[i]：如果now->son[i]存在, now->son[i]->fail= now->fail->son[i],这里如果now->fail->son[i]不存在，now->son[i]->fail就连到root；如果now->son[i]不存在，就now->son[i]= now->fail->son[i]
3. 对于主串str，从i=0开始，ac自动机从root开始，匹配一遍：如果now->son[str[i]]存在说明匹配成功，正常的前进一步；

如果now->son[str[i]]不存在说明匹配未成功，沿着now->fail反复跳，跳几次取决于问题本身，不同问题有着不同思路，列举几个

1.求模式串在主串中出现几个，同一个多次出现也记录：

沿着fail一直跳，直到跳到根为止，每次加上cnt的值，总和就是答案。

2.求模式串在主串中出现几个，同一个多次出现不记录：

沿着fail一直跳，途中访问的节点的cnt做上标记，比如让他是-1，跳到root了或者跳到是-1的节点就退出，每次加上cnt的值

* Ac自动机指向空的指针，在代码实现时候都视为指向根，有助于简化代码

## 后缀自动机

## 回文自动机

# 多项式和序列变换

## 序列变换

### 傅里叶变换

### 数论变换

### 拉普拉斯变换

### 沃尔什变换

## 多项式的表示

系数表示法：P(x)=

点值表示法：P(x)=

## 非数论的多项式运算

设

A数组是a[n]

B数组是b[m]

它们代表多项式在系数表示法下的系数

### 多项式加减法

逐项相加减无特别算法。

### **多项式对多项式取模**

意义：P(x)%意味着只取多项式的前t项

## 数论意义下的多项式运算

### FFT/NTT计算多项式乘法：

多项式乘法模型，就是所谓卷积。对于朴素算法就是直接模拟，复杂度O(n\*n).

用FFT/NTT加速方法如下(理论不再证明)：

对于多项式数组xa和xb；设changeN(x)代表转变成比x大的第一个2的次方数. 设它们傅里叶变换后数组是fa,fb,有如下性质：

Xa\*xb=)

Xa+xb=)

这个也适用于大整数乘法，具体操作来说：

把A数组的NTT或者FFT得到F1

把B数组的NTT或者FFT得到F2

用F1和F2每一项相乘 加和得到F3

F3[i]=

把F3数组逆NTT或者逆FFT得到x数组

X数组等于n\*n复杂度的暴力算法所计算的x结果数组

FFT或者NTT时候应该统一长度L

L=比len(A)+len(B)-1大的第一个2的整数次方

因为原本两个多项式乘法结果就应该是len(A)+len(B)-1个数

最后结果的x数组也是L长度的，但是我们只需要输出到

len(A)+len(B)-1

一般NTT和FFT都是解决卷积问题的算法，理论上都能用，NTT只能用在数论意义下，但是代码效率等存在差异：

* 在大数运算范围内，FFT在高精度乘法上的误差分析有人做过，结论就是完全可以忽略，原因我没查过。
* 在大数运算范围内，FFT比NTT稍微快一点，虽然FFT是浮点数运算，但总运算次数比NTT少，NTT要频繁取模。
* 大数里，不论FFT还是NTT都之能压位2位，FFT要满足：

HEX\*len\*len<double极限。NTT要满足：HEX\*len\*len<大素数P

* 加减法也可以用，但没有时间优势，在求逆里有应用
* 大整数乘法就是多项式乘法，属于同一模型换个名字。但是注意，如果取P=998244353, 应该满足HEX\*HEX\*len<P HEX是大整数乘法的进制大小，len是要被NTT数组长度。不满足就会出错，为什么呢？因为NTT一直是在源根g的多少次方模P下的，从来不是相乘结果总体模P，你要保证数组累加后，再算乘法时候两个两个乘不会爆P；

### 多项式求逆：

洛谷4238

定义deg：代表多项式最高次数，称之为多项式的度

对于一个多项式 A(x)，如果存在 B(x)满足 degB≤degA并且A(x)B(x) mod =1

那么称 B(x)为 A(x)在 模多项式意义下的**逆元**，记作

由于是数论意义下，所有系数也要模一个给定数P，也就是计算数据时要模P，计算多项式逆时要模

现在考虑如何求：

1. 当 n=1 时，A(x)次数任意时，A(x)%=c，c是常数，由于是数论意义下，c不一定有模P逆元，当c有模P逆元时,左右乘以,得到A(x)\*%=1

；当c没有模P逆元时, A(x)没有模意义下的逆元

由此可知，易求

1. n>1时，求用递归思路，假设已知了

则,且

两式相减得到：

写成同余式：

平方左乘A(x):

此时因为平方了，可以变为

乘进去：

//移项再整理

结论是：

//这里”2”当作一个常数项是2其他项为0的多项式

**按照多项式长度从小到大后倍增递推即可**

用FFT/NTT优化乘法

这个算法时间复杂度是logn\*n\*logn

网上说是n\*;logn的，但其实要做多次FFT/NTT运算，而且NTT/FFT写得好与坏，常数差距太大。要不断优化模板

注意求FFT/NTT运算次数可以优化

### 多项式除法和取模：

给定一个n项多项式A(x)，最高次项n-1

给定一个m项多项式B(x)，最高次项m-1

找到一个D(x)和R(x)使得：

A(x)=(D(x)\*B(x)+R(x) )

D(x)代表多项式的商，R(x)代表余数，除法就是求它们。

商D(x)的项数 (n-m)，余数R(x)的项数 m，如何计算除法？

令Ar(x)=，实际上Ar(x)就是A(x)的系数翻转

比如说A(x)=x3+2x2+4x+1 则AR(x)=1+2x+4x2+x3

现在看A(x)=(D(x)\*B(x)+R(x) )，令Dr(x)和 Br(x)是对应多项式的系数翻转的多项式

且由D(x)和R(x)项数得到：Dr(x)= Br(x)=

把带入得A()=(D()\*B()+R() ) 然后两边乘以带入 Dr(x)和 Br(x)得到

A()=(D()\*B()+R() ) => Ar(x)= Dr(x) Br(x)+

* Ar(x)= Dr(x) Br(x)% x^(n-m+1) //两边取模,下一步求逆

Dr(x)= Ar(x)(x) % x^(n-m+1) 求出Dr(x)直接翻转一下得D(X) R(x)= A(x)-D(x)\*B(x)

整体运算过程要对x^(n-m+1)取模，取模的做法是每次运算只计算前n-m+1项

# 可持久化数据结构

## 可持久化静态线段树

## 可持久化动态线段树

## 可持久化可并堆

## 可持久化并查集

## 可持久化fhq\_treap

# 查找

## 莫队分块

## CDQ分治

## 整体二分

# 数论

## 快速幂与扩展快速幂

## 素数的线性筛法

## 积性函数的亚线性筛法

# 计算几何

# 数值分析

## 方程数值解

## 方程组

## 定积分