20180208模拟赛题解

福建省福州第一中学 钟知闲

February 8, 2018

目录

最大真因数

求$[l, r]$之间所有合数的最大真因数之和

$1 \leq l \leq r \leq 5 \times 10^9$

算法一

- $O(n)$枚举$[l, r]$内的数
- 对于每个数x，$O(n)$枚举x的所有因数，如果x是合数则将最大的因数加进答案

时间复杂度$O(n^2)$，期望得分30

算法二

x最大真因数就是$\frac{x}{p_x}$，p_x为x的最小质因数
枚举x因数只要枚举到$\sqrt{x}$
时间复杂度$O(n\sqrt{n})$，期望得分40

算法三

用筛法解决

- 用埃氏筛法或线性筛法筛质数
- 同时预处理出$1 \sim r$所有数的最小质因数

时间复杂度$O(r \log\log r)$或$O(r)$，期望得分60

算法四

对于测试点7,8，$r - l \le 10^6$，可以只筛这个区间

- 用埃氏筛法或线性筛法筛$\sqrt{r}$以内质数
- 对于$\sqrt{r}$以内的每个质数p，将区间$[l, r]$内所有p的倍数标记上p，即可得到$[l, r]$内每个数的最小质因数

时间复杂度$O((r-l)\log\log(r-l))$，期望得分80

算法五

把问题转成$[2,r]$的和减去$[2,l-1]$的和

用DP模拟埃氏筛法统计$[2,n]$的和

记$f(i,j)$表示：$[2,i]$内的整数，筛掉前j个质数的非自身整数倍，剩下的数的和

假设$[2,i]$已经筛了前$j-1$个质数，则第j个质数p_j筛掉的数形如p_jx

如果$p_j^2>i$，那么筛不掉任何数，$f(i,j)=f(i,j-1)$

算法五

否则：

- $2 \le x \le \lfloor \frac{i}{p_j} \rfloor$
- x不含小于p_j的质因数

这说明x应该在$[2, \lfloor \frac{i}{p_j} \rfloor]$区间筛掉前$j-1$个质数后剩下的数中，但不能是前$j-1$个质数之一

$$f(i,j) = f(i,j-1) - p_j(f(\lfloor \frac{i}{p_j} \rfloor, j-1) - s_{j-1})$$

其中s_j表示前j个质数的和

边界$f(i,0) = \frac{i(i+1)}{2} - 1$

算法五

- 注意到只有$\lfloor \frac{n}{x} \rfloor, p_j \le \sqrt{i}$的状态$f(i,j)$是有用的
- 并且$p_j \le \sqrt{i}$的j只有$\frac{\sqrt{i}}{\log i}$个
- 总复杂度为$O(\frac{n^{\frac{3}{4}}}{\log n})$

期望得分100
注意开连续数组而不是map，否则会TLE

欧拉函数

给定正整数序列$\{a_n\}$，依次执行q个操作，操作有三类：

- 修改$a_i \leftarrow w$
- 查询$\varphi(\sum_{i=l}^{r} a_i) \bmod (10^9+7)$
- 查询$\varphi(\prod_{i=l}^{r} a_i) \bmod (10^9+7)$

$n \le 5\times10^4, q \le 10^5, a_i \le 4\times10^4$，修改不超过20000个，数值和操作参数随机生成

算法一

暴力模拟，根据定义暴力计算欧拉函数

复杂度$O(qa_i^n \log a_i^n)$，期望得分25分

算法二

考虑只有操作1的做法

由于所有数的和不超过$n \cdot a_i \leq 2 \times 10^9$，可以先筛$\sqrt{2 \times 10^9}$以内质数，用质因数分解计算$\varphi$

单次最坏$\frac{\sqrt{2\times 10^9}}{\log(2\times 10^9)}$，由于数据随机复杂度不满

直接维护序列，对于操作1，$O(n)$暴力求和然后质因数分解求φ，时间复杂度$O(qn)$

期望得分20分，结合算法一35分

算法三

只有操作1的部分还可以进一步优化
单点修改，区间求和可以用线段树或树状数组维护
修改和求和复杂度$O(\log n)$，求φ复杂度不变
期望得分40分，结合算法一55分

算法四

考虑乘积的φ怎么求

虽然直接乘起来的数可能很大，但乘积的质因数分解中每个数依然不超过40000

- 初始化和操作0时，将a_i质因数分解
- 操作1的求法不变
- 操作2将区间内a_i包含的每个质因数的幂加起来
- 根据$\varphi(x)=\prod_i (p_i-1)p_i^{c_i-1}$在模$10^9+7$意义下计算答案

质因数分解可以用预处理每个数最小质因数的方法$O(\log a_i)$，同样随机数据不满

复杂度$O(qn\log a_i)$，期望得分60分，结合算法三80分

算法五

操作1的复杂度已经可以通过最大数据范围了，以下只考虑优化操作2

观察

$$\varphi(x) = x\prod_i(1-\frac{1}{p_i})$$

可以发现，操作2只要求出区间积、区间包含的所有不同质因数p_i的$1-\frac{1}{p_i}$的乘积

这里的$1-\frac{1}{p_i}$可以用乘法逆元

单点修改、维护区间积可以用线段树或树状数组$O(\log n)$

关键在于维护不同质因数怎么做

算法五

思路一：莫队
莫队算法，通过加入一维时间可以实现带修改
复杂度$O(nq^{\frac{2}{3}})$，期望得分$80 \sim 100$

算法五

思路二：线段树套bitset

开一个线段树维护区间内不同质因子，查询时暴力查出每个质因数是否出现

但直接维护的复杂度是$\log n \times$(40000以内不同质数个数)，即$\log n \cdot \frac{a_i}{\log a_i}$

用std::bitset压位，复杂度可以除掉一个$w = 64$

复杂度$O(q(\log n \cdot \frac{a_i}{w \log a_i} + \frac{a_i}{\log a_i}))$，期望得分$80 \sim 100$

算法五

思路三：分治

假设在某个时刻数列中所有质因子p分布在位置$x_1, x_2, \cdots, x_k$

建立坐标系xOy，那么每个(x_i, x_i)左上方区域的并集包含的每个点(l, r)对应的询问区间$[l, r]$都包含p，其它点对应的询问都不包含p

运用差分与前缀和技巧打标记：

- 在点$(x_1, x_1)(x_2, x_2)\cdots(x_k, x_k)$处打$\frac{p-1}{p}$标记
- 在点$(x_1, x_2)(x_2, x_3)\cdots(x_{k-1}, x_k)$处打$(\frac{p-1}{p})^{-1} = \frac{p}{p-1}$标记

那么对于一个询问$[l, r]$，p对它的贡献就是点(l, r)右下方所有点标记的乘积

算法五

修改怎么办?
修改等价于先去掉旧的质因数再插入新的质因数
假设在某个时刻数列中所有质因子p分布在位置$x_1, x_2, \cdots, x_k$

- 去掉x_i位置的p：(x_i, x_i)和(x_{i-1}, x_{i+1})插入标记$\frac{p}{p-1}$，(x_{i-1}, x_i)和(x_i, x_{i+1})插入标记$\frac{p-1}{p}$
- 插入x_i位置的p：(x_i, x_i)和(x_{i-1}, x_{i+1})插入标记$\frac{p-1}{p}$，(x_{i-1}, x_i)和(x_i, x_{i+1})插入标记$\frac{p}{p-1}$

动态插入平面点，查询所有$x \geq l, y \leq r$的点(x, y)的点权积
离线询问，分治+树状数组解决
时间复杂度$O(q \log^2 q)$，期望得分100

残缺的算式

给一个只包含空位、$+$、$-$、$*$和括号的表达式

在所有空位中随机填入$[1, n]$之间的互不相同的整数（保证能填满），求表达式结果的期望值模10^9+7

表达式长度$|s| \le 10^4$，$n \le 10^9$

算法一

什么都不会怎么办?
对于前4个测试点，$n \le 8$，只有空位、$+$、$*$
暴力枚举每个空位填的数，然后线性扫一遍求值
时间复杂度$O(n! \cdot n)$，期望得分16

算法二

如何解析表达式?

从左到右读字符，遇到左括号暴力递归进去，遇到右括号回溯，这样回溯的时候就是右括号出来的地方

只有+、-、*的表达式仍然可以线性处理

复杂度$O(|s|)$

仍然枚举每个空位填的数，然后线性求值

时间复杂度$O(n! \cdot n)$，期望得分32

算法三

尝试把表达式展开，比如

$$(x_1+x_2x_3)(x_4x_5-x_6)=x_1x_4x_5-x_1x_6+x_2x_3x_4x_5-x_2x_6$$

由期望的性质

$$E((x_1+x_2x_3)(x_4x_5-x_6))=E(x_1x_4x_5)-E(x_1x_6)+E(x_2x_3x_4x_5)-E(x_2x_6)$$

即可以计算每一项的期望

实际上次数一样的$x_1x_2\cdots x_k$项期望也是一样的，所以只要处理出c_k表示所有k次项的系数和

$O(\binom{n}{k})$暴力求期望，期望得分52

算法三

如何求展开式中所有k次项的系数和?

仍然用递归求值，只是这里的值换成了一个struct，存有系数数组c[]

```
struct item{int m,*c;};
item opt(item a,char op,item b){
    item x;
    x.m=op=='*'?a.m+b.m:a.m>b.m?a.m:b.m;
    x.c=new int[x.m+1];
    for(int i=0;i<=x.m;i++)x.c[i]=op=='*'?0:((i<=a.m?a.c[i]:0)+(i<=b.m?op=='+'?b.c[i]:mod-b.c[i]:0))%mod;
    if(op=='*')
        for(int i=0;i<=a.m;i++)
            for(int j=0;j<=b.m;j++)x.c[i+j]=(x.c[i+j]+1ll*a.c[i]*b.c[j])%mod;
    return x;
}
```

算法三

之后表达式处理的写法十分简便

```
item rex(char*&s);
item rnum(char*&s){
    item r;
    if(*s=='_')r.m=1,r.c=new int[2],r.c[0]=0,r.c[1]=1;
    else r=rex(++s);
    return ++s,r;
}
item rit(char*&s){
    item x=rnum(s);
    while(*s=='*')x=opt(x,'*',rnum(++s));
    return x;
}
item rex(char*&s){
    item x=rit(s);
    for(char o;(o=*s)=='+'||o=='-';)x=opt(x,o,rit(++s));
    return x;
}
int main(){
    scanf("%d%s",&n,s);
    char*p=s;
    item a=rex(p);
    // ...
}
```

算法四

对于$|s| \le 10^4$，$n \le 5000$的测试点，用DP更快地计算$x_1 x_2 \cdots x_k$的期望

$f(i,j)$表示从$1,2,\cdots,i$中选j个数的所有方案中j个数的乘积之和

$$f(i,j) = f(i-1,j) + i(i-1,j-1)$$

时间复杂度$O(|s| \cdot n)$
空间复杂度：

- 处理表达式最多进行$\frac{|s|}{2}$次运算，第i次得到的结果次数最多$i+1$，所以总空间$1+2+\cdots+\frac{|s|}{2} = \frac{|s|^2}{8} \le 1.25 \times 10^7$
- DP部分，所以f数组开5000×5000就够了（事实上还可以滚动数组）

期望得分80

算法五

对于$n \le 10^9$，$|s| \le 10^4$的测试点，瓶颈在于n
找一找规律吧：

$$f(i,1)=\sum_{j=1}^{i} j=\binom{i+1}{2}$$

$$f(i,2)=\sum_{j=1}^{i} jf(j-1,1)=\sum_{j=1}^{i} j\binom{j}{2}=\sum_{j=1}^{i}(\binom{j}{3}+2\binom{j}{2})$$

$$=\binom{i+1}{4}+2\binom{i+1}{3}$$

算法五

因为：

- $\binom{x}{k}x = \binom{x}{k+1} + k\binom{x}{k}$
- $\sum_{i=0}^{x} \binom{i}{k} = \binom{i+1}{j+1}$

所以$f(i,j)$具有$\sum_{k=0}^{2j} a_{j,k}\binom{i+1}{k}$的形式，其中$a_{j,k}$为常系数
暴力递推求出$a_{j,k}$，然后将$f(n,k)$代入计算即可
复杂度$O(|s|^2)$，期望得分100

算法五

实际上这题有多种解法：

- $f(i,k)$是一个$2k$次多项式，可以用拉格朗日插值法$O(k^2)$求出
- $f(i,j)$实际上是第一类斯特林数i行倒数j个数，可以用FFT计算
- ……

End

祝大家接下来的比赛取得优异成绩！