提高算法班

卡特兰数、容斥原理、鸽巢原理、[扩展]卢卡斯定理

Mas

卡特兰数

1. 由 n 个结点可构造多少个不同的二叉搜索树？
2. 对角线不相交的情况下，将一个凸 $n\ (n \geq 3)$ 边形区域分成三角形区域的方法数？
3. 在圆上选择 $2n$ 个点，将这些点成对连接起来使得所得到的 n 条线段不相交的方法数？
4. 由 n 个 $+1$ 和 n 个 -1 构成 $2n$ 项 $a_1, a_2, \cdots, a_{2n}$
5. 满足 $a_1 + a_2 + \cdots + a_k \geq 0\ (k = 1, 2, 3, \cdots, 2n)$，有多少种合法的序列？
6. 由 n 个左括号和 n 个右括号组合的合法括号序列有多少种？
7. 有 $2n$ 个人排成一行进入剧场，入场费 5 元。只有 n 个人有一张 5 元钞票，另外 n 人只有 10 元钞票，剧院无其它钞票有多少种方法使得只要有 10 元的人买票，售票处就有 5 元的钞票找零？
8. 在 $n \times n$ 格点中不越过对角线的单调路径的个数是多少？
9. 一个栈 (无穷大) 的进栈序列为 $1 \sim n$ 有多少个不同的出栈序列？

卡特兰数

n 个结点编号 $1 \sim n$ ，可构造多少个不同的二叉搜索树?

设 H_n 为节点数为 n 时二叉搜索树数量

假设选取 k 号点为根

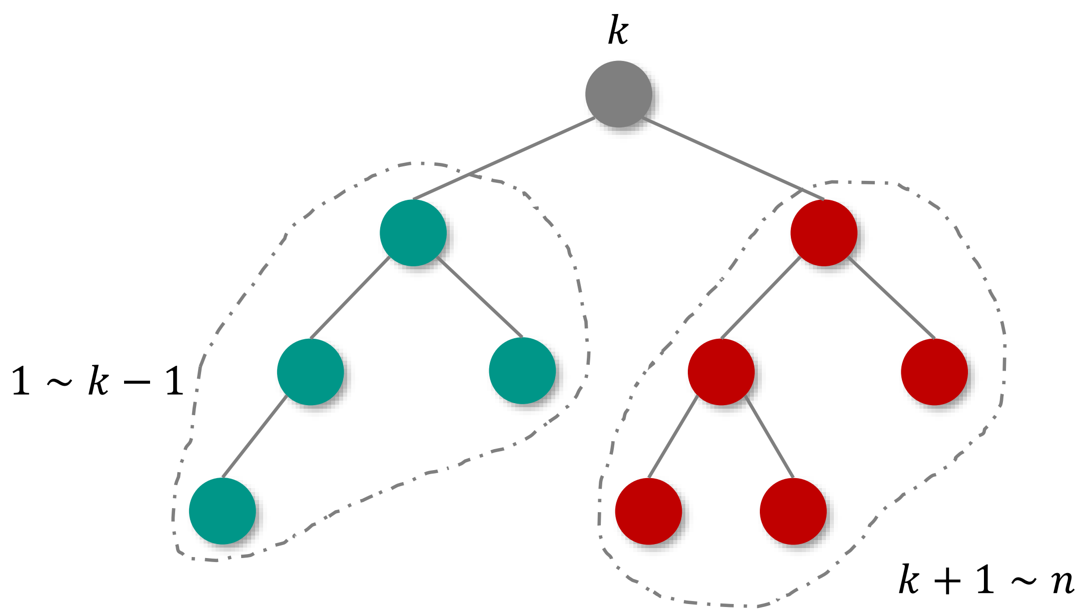
- 节点 $1 \sim k-1$ 作为左子树

  左子树也要为二叉搜索树，方案数为 H_{k-1}

- 节点 $k+1 \sim n$ 作为右子树

  左子树也要为二叉搜索树，方案数为 H_{n-k}

k 可取 $1 \sim n$



对角线不相交的情况下，将一个凸 n ($n \geq 3$) 边形区域分成三角形区域的方法数?

设 F_n 为凸 n 边形区域分成三角形区域的方法数

卡特兰数

将凸 n 边形顶点 $1\sim n$ 编号

从 $3\sim n-1$ 任选一点 k

- 顶点 $1\sim k$ 作为新的凸多边形

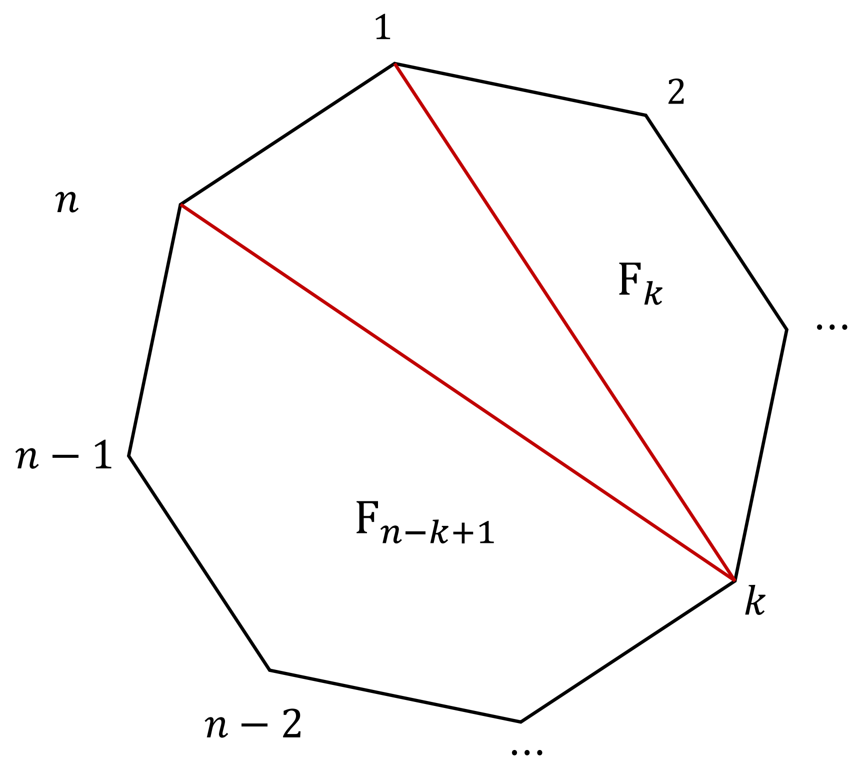
  划分方案数为 F_k

- 顶点 $k\sim n$ 作为新的凸多边形

  方案数为 F_{n-k+1}

令 $F_n=H_{n-3}$

结合上述模型，不难得出递推式



$$H_n=\begin{cases}1, & n=0,1\\ \sum_{i=1}^{n} H_{i-1}H_{n-i}, & n\ge 2\end{cases}$$

卡特兰数

由 n 个 $+1$ 和 n 个 -1 构成 $2n$ 项 $a_1, a_2, \cdots, a_{2n}$，满足 $a_1 + a_2 + \cdots + a_k \geq 0\ (k = 1, 2, 3, \cdots, 2n)$，有多少种合法的序列?

设 H_n 为方案数，不考虑限制时方案数为 $\binom{2n}{n}$

对于任意一个由 n 个 $+1$ 和 n 个 -1 构成的非法序列 S，其必然存在 $2p+1 \in [1, 2n]$ 满足

- $S[1 \sim 2p+1]$ 有 p 个 $+1$ 和 $p+1$ 个 -1
- $S[2p+2 \sim 2n]$ 有 $n-p$ 个 $+1$ 和 $n-p-1$ 个 -1

将 $S[2p+2 \sim 2n]$ 绝对值取反，可得到新的序列 S'

那么 $S'[2p+2 \sim 2n]$ 有 $n-p$ 个 -1 和 $n-p-1$ 个 $+1$，S' 中一共有 $n-1$ 个 $+1$ 和 $n+1$ 个 -1

对于任意一个由 $n-1$ 个 $+1$ 和 $n+1$ 个 -1 构成的非法序列 S，其必然存在 $2p+1 \in [1, 2n]$

- $S[1 \sim 2p+1]$ 有 p 个 $+1$ 和 $p+1$ 个 -1
- $S[2p+2 \sim 2n]$ 有 $n-p-1$ 个 $+1$ 和 $n-p$ 个 -1

将 $S[2p+2 \sim 2n]$ 绝对值取反，可得到新的序列 S'

卡特兰数

将 $S[2p+2\sim 2n]$ 绝对值取反，可得到新的序列 S'

那么 $S'[2p+2\sim 2n]$ 有 $n-p-1$ 个 -1 和 $n-p$ 个 $+1$，S' 中一共有 n 个 $+1$ 和 n 个 -1

即 由 n 个 $+1$ 和 n 个 -1 构成的非法序列 与 $n-1$ 个 $+1$ 和 $n+1$ 个 -1 构成的非法序列，构成一一对应(双射)

综上

$$\mathrm{H}_n = \binom{2n}{n} - \binom{2n}{n-1}$$

n 个左括号和 n 个右括号组合的合法括号序列有多少种?

将左括号视为 $+1$ 右括号视为 -1 ，该问题等价于上一问题

有 $2n$ 个人排成一行进入剧场，入场费 5 元。 只有 n 个人有一张 5 元钞票，另外 n 人只有 10 元钞票

剧院无其它钞票有多少种方法使得只要有 10 元的人买票，售票处就有 5 元的钞票找零？

将持有 5 元的人视为 $+1$ 持有 10 元的人视为 -1，该问题同样等价于上一问题

卡特兰数

在 $n \times n$ 格点中不越过对角线的单调路径的个数是多少?

设 H_n 为方案数, 不考虑限制时方案数为 $\binom{2n}{n}$

对于任意一个非法路径，其必然存在与 $y=x+1$ 的交点

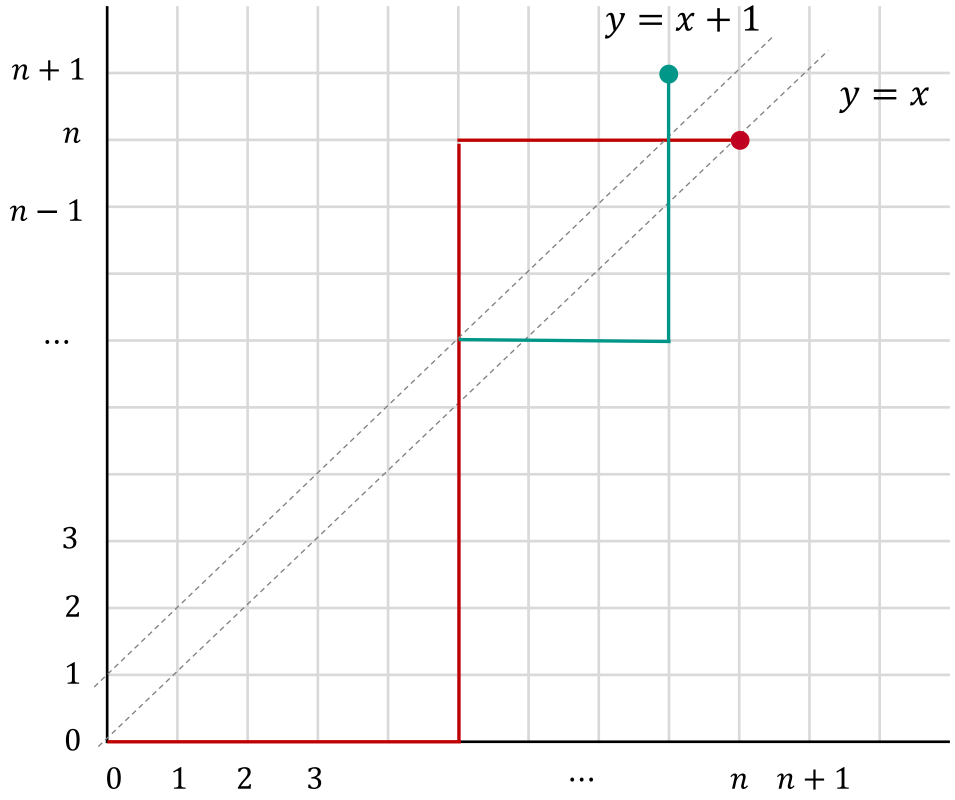
将最早产生的交点，其后路径根据 $y=x+1$ 对称翻转

最终达到点 $(n-1, n+1)$

显然这种翻转后的路径与翻转前的路径一一对应(双射)

即

$$H_n = \binom{2n}{n} - \binom{2n}{n-1}$$



一个栈 (无穷大) 的进栈序列为 $1 \sim n$ 有多少个不同的出栈序列?

卡特兰数

设 H_n 为栈序列 $1\sim n$ 其不同的出栈序列方法数

不妨设 1 为第 k 个出栈元素

1 先入栈，接下来 $2\sim\ k$ 依次入栈并出栈

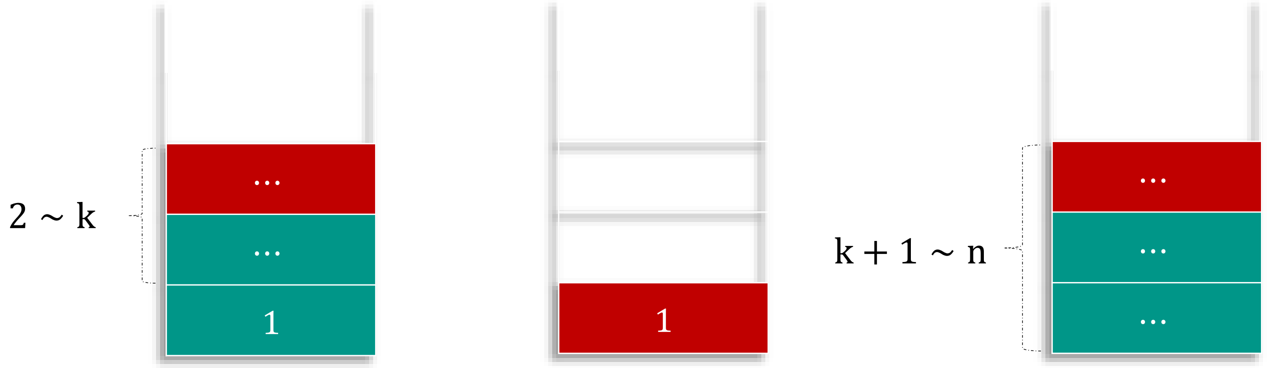
方案数为 H_{k-1}

1 出栈

接下来 $k+1\sim n$ 依次入栈，并出栈

方案数为 H_{n-k}

k 可取 $1\sim n$



若将一个元素入栈视作 +1 出栈视为 −1

问题转化为：**由 n 个 $+1$ 和 n 个 -1 构成 $2n$ 项 $a_1,a_2,\cdots,a_{2n}$，满足 $a_1+a_2+\ \cdots+a_k\geq 0\ (k=1,2,3,\cdots,2n)$ 的合法的序列数**

通过该转化可将递推式与通项式联系起来

卡特兰数

记 H_n 为卡特兰数 第 n 项 , 其递推式为

$$\mathrm{H}_n = \begin{cases} 1, & n = 0,1 \\ \sum_{i=1}^{n} \mathrm{H}_{i-1}\mathrm{H}_{n-i}, & n \geq 2 \end{cases}$$

同时有

$$\mathrm{H}_n = \binom{2n}{n} - \binom{2n}{n-1}$$

$$= \frac{1}{n+1}\binom{2n}{n} = \frac{1}{n+1}\sum_{i=0}^{n}\binom{n}{i}^2$$

$$= \frac{1}{n+1}\mathrm{H}_{n-1}(4n-2)$$

抽屉原理

抽屉原理/鸽巢原理（pigeonhole principle），常被用于证明存在性证明和求最坏情况下的解

将 $n+1$ 个物体划分为 n 组,那么有至少一组有两个(或以上)的物体

证明

若各分组有至多 1 个物体，那么最多有 n 个物体，而实际上有 $n+1$ 个物体,矛盾

推广

将 n 个物体划分为 k 组，至少存在一个分组含有大于或等于 $\left\lceil \frac{n}{k} \right\rceil$ 个物品

证明

若每个分组含有小于 $\left\lceil \frac{n}{k} \right\rceil$ 个物体，则其总和 S

$$S = \left(\left\lceil \frac{n}{k} \right\rceil - 1\right) \times k = k\left\lceil \frac{n}{k} \right\rceil - k \ < \ k(\frac{n}{k} + 1) - k = n$$

矛盾

#1286、倍数问题

题目描述

一个长度为 N 的数组 A

从 A 中选出若干个数,使得这些数的和是 N 的倍数

输入格式

第 1 输入 1 个数 N , N 为数组的长度,同时也是要求的倍数

第 $2 \sim N+1$ 行: 数组 A 的元素

输出格式

如果没有符合条件的组合,输出 `No Solution` ,若存在多组解,输出任意一组即可

第 1 行: 1 个数 S 表示你所选择的数的数量

第 $2-S+1$ 行: 每行 1 个数,对应你所选择的数

数据规模

对于全部的数据 $1 \leq N \leq 10^3, 1 \leq A_i \leq 10^9$

若存在 $A_i \equiv 0 \pmod n$ 直接输出 A_i 即可

否则令 $\mathrm{sum}_i = \sum_{j=0}^{i} A_i \pmod n$, 发现值域为 $[0, n-1]$

若 sum_i 为零 , $\sum_{j=1}^{i} A_j$ 即为一组可行解

根据抽屉原理

必然存在 $i \geq j$ 使得 $\mathrm{sum}_i = \mathrm{sum}_j$ 即

$$\mathrm{sum}_i - \mathrm{sum}_j \equiv 0 \pmod n$$

其中 $\sum_{x=j+1}^{i} A_x$ 即为一组可行解

时间复杂度 $O(n)$

容斥原理

设全集 U 中元素有 n 种不同的属性

第 i 种属性称为 P_i , 拥有属性 P_i 的元素构成集合 S_i

那么

$$\left|\bigcup_{i=1}^{n} S_i\right| = \sum_{i} |S_i| - \sum_{i<j} |S_i \cap S_j| + \sum_{i<j<k} |S_i \cap S_j \cap S_k| + \cdots + (-1)^{n-1} |S_1 \cap S_2 \cap \cdots \cap S_n|$$

即

$$\left|\bigcup_{i=1}^{n} S_i\right| = \sum_{T \subseteq \{1,2,3,\cdots,n\}} \left((-1)^{|T|-1} \left|\bigcap_{i \in T} S_i\right| \right)$$

容斥原理证明

考虑某个元素被 n ($n > 0$) 个集合 $T_1, T_2, \cdots, T_n$ 包含

根据 **容斥原理** 考虑其在并集中出现次数为

容斥原理

$$\sum_{i}|\mathrm{T}_i| - \sum_{i<j}|\mathrm{T}_i \cap \mathrm{T}_j| + \sum_{i<j<k}|\mathrm{T}_i \cap \mathrm{T}_j \cap \mathrm{T}_k| + \cdots + (-1)^{n-1}|\mathrm{T}_1 \cap \mathrm{T}_2 \cap \cdots \cap \mathrm{T}_n|$$

即

$$\binom{n}{1} - \binom{n}{2} + \binom{n}{3} + \cdots + (-1)^{n-1}\binom{n}{n}$$

根据 **二项式定理**

$$\binom{n}{0} - \left(\binom{n}{1} - \binom{n}{2} + \binom{n}{3} + \cdots + (-1)^{n-1}\binom{n}{n}\right) = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \cdots + (-1)^{n}\binom{n}{n} = \sum_{i=0}^{n}(-1)^{i}\binom{n}{i}$$

那么

$$\binom{n}{1} - \binom{n}{2} + \binom{n}{3} + \cdots + (-1)^{n-1}\binom{n}{n} = \binom{n}{0} - \sum_{i=0}^{n}(-1)^{i}\binom{n}{i} = 1 - (1-1)^{n} = 1$$

则意味着该元素仅在并集中被计入一次

命题得证

#318、能被整除的数

实验舱 青少年编程 走近科学 走进名校

题目描述

给定一个整数 n 和 m 个不同的质数 $p_1, p_2, \sim, p_m$

请你求出 $1 \sim n$ 中能被 $p_1, p_2, \cdots, p_m$ 中的至少一个数整除的整数有多少个

输入格式

第一行包含整数 n 和 m

第二行包含 m 个质数

输出格式

输出一个整数，表示满足条件的整数的个数

数据范围

对于全部的数据 $1 \leq m \leq 16, 1 \leq n, p_i \leq 10^9$

输入样例

```
10 2
2 3
```

输出样例

```
7
```

#318、能被整除的数

对于$1 \sim n$ 设能被 p_i 整除的数的集合为 S_{p_i}

答案为

$$\left|S_{p_1} \cup S_{p_2} \cup \cdots \cup S_{p_m}\right|$$

其中

$$|S_i| = \left\lfloor \frac{n}{p_i} \right\rfloor$$

由于 p_i 均为质数, 质数的乘积即为最小公倍数所以

$$|S_{p_1} \cap S_{p_2} \cdots \cap S_{p_i}| = \left\lfloor \frac{n}{p_1 p_2 \cdots p_i} \right\rfloor$$

二进制枚举出 p_i 的选取方式容斥求解

但选出质因子乘积超过 n 时应当忽略，否则将出现数据溢出

时间复杂度 $O(2^m - 1)$

#2470、互质

题目描述

对于两个正整数 n 和 k

求与 n 互质的第 k 个正整数

输入格式

输入两个正整数 n 和 k

输出格式

一个正整数表示答案

输入样例

```
10 5
```

输出样例

```
11
```

数据范围

对于 20% 的数据, $1 \le n \le 200, 1 \le k \le 50$

对于 40% 的数据, $1 \le n \le 10^3, 1 \le k \le 500$

对于 60% 的数据, $1 \le n \le 10^5, 1 \le k \le 10^4$

对于 100% 的数据, $1 \le n \le 10^6, 1 \le k \le 10^8$

思路一

若有

$$(x,n) = 1$$

显然

$$(x+n,n) = 1$$

可发现 $(x,n) = 1$ 存在周期性，周期为 $\varphi(n)$

求出第 $k \bmod \varphi(n)$ 个互质的数,加上 $\left\lfloor\frac{n}{\varphi(n)}\right\rfloor k$

需要特判 $k \bmod \varphi(n)$为 0

时间复杂度 $O(n \log n)$

#2470、互质

思路二

在 $1 \sim nk$ 范围内二分答案,求出 $1 \sim$ mid 中与 n 互质数的个数

容易求出 $1 \sim \text{mid}$ 中不与 n 互质数的个数

将 n 唯一分解为 $\prod_{i=1}^{k} p_i^{c_i}$

对于 $1 \sim \text{mid}$ 设能被 p_i 整除的数的集合为 S_{p_i}, 不与 n 互质数个数为

$$\left|S_{p_1} \cup S_{p_2} \cup \cdots \cup S_{p_m}\right|$$

其中$|S_i| = \left\lfloor \frac{\text{mid}}{p_i} \right\rfloor$, 所以 $|S_{p_1} \cap S_{p_2} \cdots \cap S_{p_i}| = \left\lfloor \frac{\text{mid}}{p_1 p_2 \cdots p_i} \right\rfloor$

二进制枚举出 p_i 的选取方式

使用容斥原理求解 , 在极限数据的情况下发现 n 最多有 8 个质因子

单次检查答案最坏时间复杂度 $O(2^8 - 1)$

容斥原理&欧拉函数

根据唯一分解定理 $n=\prod_{i=1}^{k} p_i^{c_i}$

把 $p_i \nmid x$ 作为属性对应的集合为 S_i，因此有

$$\varphi(n)=\left|\bigcap_{i=1}^{k} S_i\right|=|U|-\left|\bigcup_{i=1}^{k} \overline{S_i}\right|$$

$|U|$ 为 n，$\overline{S_i}$ 为 S_i 的补集（$p_i \mid x$ 构成的集合）

显然

$$|\overline{S_i}|=\frac{n}{p_i}$$

不难推出

$$\left|\bigcap_{a_i<a_{i+1}} \overline{S_{a_i}}\right|=\frac{n}{\prod p_{a_i}}$$

容斥原理 & 欧拉函数

根据定义

$$\varphi(n) = n - \left(\sum_i \frac{n}{p_i} - \sum_{i<j} \frac{n}{p_i p_j} + \cdots + (-1)^k \frac{n}{p_1 p_2 \cdots p_k}\right)$$

$$= n\left(\frac{(p_1-1)(p_2-1)\cdots(p_k-1)}{p_1 p_2 \cdots p_k}\right)$$

通分,提取公因式

$$= n\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_k}\right) = n\prod_{i=1}^{k}\left(\frac{p_i-1}{p_i}\right)$$

进一步的若 n, m 互质 , $n = \prod_{i=1}^{k_1} p_i^{c_i}$, $m = \prod_{j=1}^{k_2} q_j^{c_j}$ 显然 $p_i \neq q_j$

$$\varphi(n \times m) = \left(n\prod_{i=1}^{k_1}\left(\frac{p_i-1}{p_i}\right)\right) \times \left(m\prod_{j=1}^{k_2}\left(\frac{q_j-1}{q_j}\right)\right)$$

即

$$\varphi(n \times m) = \varphi(n) \times \varphi(m)$$

容斥原理 & 不定方程解的数量

给定非负整数$a_1, a_2, \cdots, a_k$,求 $x_1 + x_2 + \cdots + x_k = n$ 的非负整数解个数,要求满足 $x_1 \le a_1, x_2 \le a_2, \cdots, x_k \le a_k$

若无限制时方案数为 $\binom{n+k-1}{k-1}$ 即

$$|\mathrm{U}| = \binom{n+k-1}{k-1}$$

把 $x_i \le a_i$ 作为属性对应的集合为 S_i, 最终答案为 $\left|\bigcap_{i=1}^{k} S_i\right|$

那么

$$\left|\bigcap_{i=1}^{k} S_i\right| = |\mathrm{U}| - \left|\bigcup_{i=1}^{k} \overline{S_i}\right|$$

其中 $\overline{S_i}$ 表示满足 $x_i \ge a_i + 1$ 解的数目

不妨枚举出所有 $\left|\bigcup_{i=1}^{k} \overline{S_i}\right|$ 的情况

容斥原理 & 不定方程解的数量

实验舱 青少年编程 走近科学 走进名校 LAB

对于 t 个满足 $x_i \geq a_i + 1$ 的集合交集，问题可转化为

$$\sum_{i=1}^{k} x_i = n - \sum_{i=1}^{t}(a_i + 1)$$

显然答案为

$$\binom{n + k - 1 - \sum_{i=1}^{t}(a_i + 1)}{k - 1}$$

综上答案为

$$\binom{n + k - 1}{k - 1} + \sum_{i=1}^{n}\left((-1)^{i-1} \sum_{b_j < b_{j+1}}^{|b|=i} \binom{n + k - 1 - \sum_{x=1}^{i}(a_{b_x} + 1)}{k - 1}\right)$$

容斥原理 & 多重集的组合数

设 $S=\{n_1\cdot a_1,n_2\cdot a_2,\cdots,n_k\cdot a_k\}$ 表示由 n_1 个 a_1,n_2 个 a_2,⋯,n_k 个 a_k 组成的多重集

从 S 中选择 r 个元素组成一个多重集的方案数就是 **多重集的组合数**

若对于整数 r 且 $\forall 1\le i\le k$ 都有 $r\le n_i$ 求解多重集的组合数

等价于 $x_1+x_2+\cdots+x_k=r$ 的非负整数解的数目

可以用插板法解决，答案为

$$\binom{r+k-1}{k-1}$$

对于整数 $r\le\sum_{i=1}^{k}n_i$，从 S 中选出 r 个元素组成一个多重集的方案数

该问题等价于：$\forall\ 1\le i\le k$ 都有 $x_i\le n_i$ 且 $x_1+x_2+\cdots+x_k=r$ 的非负整数解的数目

容斥原理 & 多重集的组合数

记全集为 U 那么

$$|\mathrm{U}| = \binom{r+k-1}{k-1}$$

满足 $x_i \le n_i$ 的集合为 S_i ，不满足 $x_i \le n_i$ 的集合为 $\bar{\mathrm{S}_i}$ 可将其转化为 $x_i \ge n_i + 1$

显然

$$\bigcap_{i=1}^{k} |\mathrm{S}_i| = |\mathrm{U}| - \bigcup_{i=1}^{k} |\bar{\mathrm{S}_i}|$$

根据容斥原理

$$\bigcup_{i=1}^{k} |\bar{\mathrm{S}_i}| = \sum_{i} \binom{r+k-n_i-2}{k-1} - \sum_{i<j} \binom{r+k-n_i-n_j-3}{k-1} + \cdots + (-1)^k \binom{r+k-n_i-\sum_{i=1}^{k} n_i - k - 1}{k-1}$$

答案即为 $\cap_{i=1}^{k} |\mathrm{S}_i|$

二项式反演

记 $\mathbf{F}_n$ 表示**恰好**由 n 个不同元素形成特定结构的方案数

$\mathbf{g}_n$ 表示从 n 个不同元素选 $i \geq 0$ 个元素形成特定结构的总方案数

形式1

若已知 $\mathbf{F}_n$ 求 $\mathbf{g}_n$ 有

$$\mathbf{g}_n = \sum_{i=0}^{n} \binom{n}{i} \mathbf{F}_i \Leftrightarrow \mathbf{F}_n = \sum_{i=0}^{n} \binom{n}{i} (-1)^{n-i} \mathbf{g}_i$$

证明

若已知

$$\mathbf{g}_n = \sum_{i=0}^{n} \binom{n}{i} \mathbf{F}_i$$

将右式中 $\mathbf{g}_i$ 以左式代入展开

二项式反演

$$\mathbf{F}_n = \sum_{i=0}^{n} \binom{n}{i} (-1)^{n-i} \sum_{j=0}^{i} \binom{i}{j} \mathbf{F}_j$$

$$= \sum_{i=0}^{n} \sum_{j=0}^{i} \binom{n}{i} \binom{i}{j} (-1)^{n-i} \mathbf{F}_j$$

交换求和顺序

对于数对 (i,j) 需保证 $j \le i$

$$\mathbf{F}_n = \sum_{j=0}^{n} \mathbf{F}_j \sum_{i=j}^{n} \binom{n}{i} \binom{i}{j} (-1)^{n-i}$$

根据 **三项式系数恒等式** 有

$$\mathbf{F}_n = \sum_{j=0}^{n} \mathbf{F}_j \sum_{i=j}^{n} \binom{n}{j} \binom{n-i}{i-j} (-1)^{n-i} = \sum_{j=0}^{n} \binom{n}{j} \mathbf{F}_j \sum_{i=j}^{n} \binom{n-j}{i-j} (-1)^{n-i}$$

二项式反演

令 $k=i-j$ 那么 $i\ =\ k+j$ 那么

$$\mathbf{F}_n=\sum_{j=0}^{n}\binom{n}{j}\mathbf{F}_j\sum_{k=0}^{n-j}\binom{n-j}{k}(-1)^{n-j-k}\,1^k$$

根据 **二项式定理**

$$\mathbf{F}_n=\sum_{j=0}^{n}\binom{n}{j}(1-1)^{n-j}=\sum_{j=0}^{n}\binom{n}{j}\mathbf{F}_j[n-j=0]=\sum_{j=0}^{n}\binom{n}{j}\mathbf{F}_j[n=j]$$

$$=\mathbf{F}_n$$

若已知

$$\mathbf{F}_n=\sum_{i=0}^{n}\binom{n}{i}(-1)^{n-i}\mathbf{g}_i$$

将右式中的 $\mathbf{F}_i$ 代入左式展开可证：已知 $\mathbf{F}_n$ 求 $\mathbf{g}_n$

二项式反演

实验舱 青少年编程 走近科学 走进名校 LAB

形式2

$$\mathbf{g}_n = \sum_{i=n}^{m} \binom{i}{n} \mathbf{F}_i \Leftrightarrow \mathbf{F}_n = \sum_{i=n}^{m} \binom{i}{n} (-1)^{i-n} \mathbf{g}_i$$

证明

将右式代入左式

$$\mathbf{g}_n = \sum_{i=n}^{m} \binom{i}{n} \sum_{j=i}^{m} \binom{j}{i} (-1)^{j-i} \mathbf{g}_i$$

$$= \sum_{i=n}^{m} \sum_{j=i}^{m} \binom{i}{n} (-1)^{j-i} \binom{j}{i} \mathbf{g}_j$$

考虑交换枚举顺序

对于数对 (i,j) 需保证 $j \leq i$

二项式反演

$$\mathbf{g}_n = \sum_{j=n}^{m} \mathbf{g}_j \sum_{i=n}^{j} \binom{i}{n}\binom{j}{i}(-1)^{j-i}$$

根据 **三项式系数恒等式、对称恒等式**

$$\mathbf{g}_n = \sum_{j=n}^{m} \mathbf{g}_j \sum_{i=j}^{m} \binom{j}{n}\binom{j-n}{i-n}(-1)^{j-i} = \sum_{j=k}^{n} \binom{j}{n} \mathbf{g}_j \sum_{i=j}^{n} \binom{j-n}{j-i}(-1)^{j-i}$$

令 $k\ = j - i$ 那么 $i = k + j$

$$\mathbf{g}_n = \sum_{j=n}^{m} \binom{j}{n} \mathbf{g}_j \sum_{k=0}^{j-n} \binom{j-n}{k}(-1)^{j-n-k} = \sum_{j=n}^{m} \binom{j}{n} \mathbf{g}_j \sum_{k=0}^{j-n} \binom{j-n}{k}(-1)^{j-n-k}1^k$$

根据 **二项式定理**

二项式反演

$$\mathbf{g}_n = \sum_{j=n}^{m}\binom{n}{j}\mathbf{g}_j(1-1)^{j-n} = \sum_{j=n}^{m}\binom{n}{j}\mathbf{g}_j[j-n=0] = \sum_{j=n}^{m}\binom{n}{j}\mathbf{g}_j[j=n]$$

$$= \mathbf{g}_n$$

将左式中的 $\mathbf{g}_n$ 代入右式展开可证：已知 $\mathbf{g}_n$ 求 $\mathbf{F}_n$

$\mathbf{F}_n$ 对应 **恰好** 的方案数，$\mathbf{g}_n$ 对应 **至少/至少** 的方案数

一般而言 $\mathbf{g}_n$ 往往方便求解，$\mathbf{F}_n$ 不好求解

已知两者其一求解另一个的过程被为 **二项式反演**

二项式反演 & 错位排列

令 $\mathbf{F}_n$ 表示**恰有** n 封错误的方案数，$\mathbf{g}_n$ 表示**至多** n 封装错的方案数

显然有 $\mathbf{g}_n = n!$ 且有

$$\mathbf{g}_n = \sum_{i=0}^{n} \binom{n}{i} \mathbf{F}_i$$

根据 **二项式反演 形式1** 有

$$\mathbf{F}_n = \sum_{i=0}^{n} \binom{n}{i} (-1)^{n-i} \mathbf{g}_i = \sum_{i=0}^{n} \binom{n}{i} (-1)^{n-i} i!$$

$$= \sum_{i=0}^{n} (-1)^{n-i} \frac{n!}{i!\,(n-i)!} i! = \sum_{i=0}^{n} (-1)^{n-i} \frac{n!}{(n-i)!} = \sum_{j=0}^{n} (-1)^{j} \frac{n!}{j!}$$

#3856、 集合计数

题目描述

一个有 N 个元素的集合有 2^N 个不同子集(包含空集)

现在要在这 2^N 个集合中取出若干集合(至少一个)

使得它们的交集的元素个数为 K

求取法的方案数,答案 $\bmod 1000000007$ 后输出

输出格式

一行两个整数 N, K

输入格式

一行为答案

数据规模

对于 100% 的数据 $1 \le N \le 1000000, 0 \le K \le N$

令 $\mathbf{F}_i$ 表示选出集合交集 **恰为** i 的方案数

令 $\mathbf{g}_i$ 表示选出集合交集 **至少** i 的方案数

显然有

$$\mathbf{g}_k = \sum_{i=k}^{n} \binom{i}{k} \mathbf{F}_i$$

考虑 $\mathbf{g}_i$ 的求解

可先选出 i 个元素方案数为 $\binom{n}{i}$

剩余元素任选，可与选出的 i 个元素搭配

剩余元素构成集合个数为 2^{n-i}

#3856、 集合计数

对于每个集合都可选可不选，但不能一个也不选

方案数为 $2^{2^{n-i}}-1$

即

$$\mathbf{g}_i=\binom{n}{i}\left(2^{2^{n-i}}-1\right)$$

根据 **二项式反演 形式2**

$$\mathbf{F}_k=\sum_{i=k}^{n}\binom{i}{k}(-1)^{i-k}\mathbf{g}_i$$

对于 $2^{2^{n-i}}$ 可使用 **费马小定理** 降幂

若不考虑预处理组合数的代价，时间复杂度 $O(n)$

Lucas定理

Lucas 定理用于求解大组合数取模的问题,其中模数必须为素数

$$\binom{n}{m} \equiv \binom{\left\lfloor \frac{n}{p} \right\rfloor}{\left\lfloor \frac{m}{p} \right\rfloor} \binom{n \bmod p}{m \bmod p} \pmod p$$

观察上述表达式

$n \bmod p$ 和 $m \bmod p$ 一定是小于 p 的数，可直接求解 $\binom{n \bmod p}{m \bmod p}$

对于 $\binom{\left\lfloor \frac{n}{p} \right\rfloor}{\left\lfloor \frac{m}{p} \right\rfloor}$ 可继续用 Lucas 定理求解，当下指标为 0 时返回 1

Lucas 定理

引理1

对于素数 p , $0 < k < p$ 有

$$p \mid \binom{p}{k}$$

证明

根据 **吸收恒等式** 有

$$\binom{p}{k} = \frac{p}{k} \times \binom{p-1}{k-1}$$

$$\Rightarrow k \times \binom{p}{k} = p \times \binom{p-1}{k-1}$$

由于 $(p,k) = 1 \wedge 0 < k < p$ 显然 $p \nmid k$

得证

引理2

对于素数 p

$$(1+x)^p \equiv 1+x^p \pmod p$$

证明

根据 **二项式定理** 展开

$$(1+x)^p = \sum_{i=0}^{p} \binom{p}{i} a^i$$

根据 **引理1**

$$(1+x)^p \equiv \binom{p}{0} 1 + \binom{p}{p} x^p = 1 + x^p \pmod p$$

得证 (也可用 **费马小定理** 证明)

Lucas 定理

证明

根据 **二项式定理** 有

$$\sum_{i=0}^{n} x^i \binom{n}{i} = (1+x)^n$$

将 n 写成带余数形式

$$(1+x)^n = (1+x)^{\left\lfloor \frac{n}{p} \right\rfloor \times p + n \bmod p} = \left((1+x)^p\right)^{\left\lfloor \frac{n}{p} \right\rfloor} \times (1+x)^{n \bmod p}$$

根据 **引理2**

$$\left((1+x)^p\right)^{\left\lfloor \frac{n}{p} \right\rfloor} \times (1+x)^{n \bmod p} \equiv (1+x^p)^{\left\lfloor \frac{n}{p} \right\rfloor} \times (1+x)^{n \bmod p} \pmod p$$

根据 **二项式定理** 展开

$$(1+x^p)^{\left\lfloor \frac{n}{p} \right\rfloor} \times (1+x)^{n \bmod p} = \left(\sum_{j=0}^{\left\lfloor \frac{n}{p} \right\rfloor} x^{j \times p} \times \binom{\left\lfloor \frac{n}{p} \right\rfloor}{j} \right) \times \left(\sum_{k=0}^{n \bmod p} x^k \times \binom{n \bmod p}{k} \right)$$

Lucas 定理

即

$$\sum_{i=0}^{n} x^i \binom{n}{i} \equiv \left(\sum_{j=0}^{\lfloor \frac{n}{p} \rfloor} x^{j\times p} \times \binom{\lfloor \frac{n}{p} \rfloor}{j}\right) \times \left(\sum_{k=0}^{n \bmod p} x^k \times \binom{n \bmod p}{k}\right)$$

$$= \sum_{j=0}^{\lfloor \frac{n}{p} \rfloor} \sum_{k=0}^{n \bmod p} x^{j\times p+k} \times \binom{\lfloor \frac{n}{p} \rfloor}{j} \times \binom{n \bmod p}{k} \pmod p$$

不难发现 $j \times p + k$ 各值只出现一次

$\binom{n}{\mathrm{m}}$ 即为左式 x^m 项的系数，对比系数得出

$$\binom{n}{m} \equiv \binom{\lfloor \frac{n}{p} \rfloor}{\lfloor \frac{m}{p} \rfloor} \binom{n \bmod p}{m \bmod p} \pmod p$$

Lucas 定理

Lucas 定理的另一描述形式

若 p 为质数

$$n = a_0 + a_1 p + a_2 p^2 + \cdots + a_k p^k$$

$$m = b_0 + b_1 p + b_2 p^2 + \cdots + b_k p^k$$

有

$$\binom{n}{m} \equiv \prod_{i=0}^{k} \binom{a_i}{b_i} \pmod p$$

上述为 p 进制视角描述 Lucas 定理

若单次求解组合数的的复杂度为 $g(n)$

使用 Lucas 定理求解组合数的复杂度为 $O(g(n)\log n)$

#813、Lucas定理

题目描述

这是一道模板题

给定 n, m 求

$$\binom{n}{m} \bmod 10007$$

输入格式

第一行一个整数 t ,表示有 t 组数据

接下来 t 行每行两个整数 n, m ,如题意

输出格式

t 行

每行一个数为 $\binom{n}{m} \bmod 10007$ 的答案

数据范围

对于全部数据, $1 \le t \le 200, 1 \le m \le n \le 2 \times 10^8$

扩展 Lucas 定理

对于模数 M 不为素数时需要用到 **扩展 Lucas 定理**

根据唯一分解定理

$$\mathrm{M} = \prod_{i=1}^{k} p_i^{c_i}$$

对于任意的 $i \neq j$ 有 $(p_i^{c_i}, p_j^{c_j}) = 1$,可构造 k 个线性同余方程

$$\begin{cases} c_1 \equiv \binom{n}{m} \pmod{p_1^{c_1}} \\ c_2 \equiv \binom{n}{m} \pmod{p_2^{c_2}} \\ \vdots \\ c_k \equiv \binom{n}{m} \pmod{p_k^{c_k}} \end{cases}$$

扩展Lucas定理

问题转化为求解

$$\binom{n}{m} \bmod p^a$$

根据组合数定义

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$$

需求解 $m!^{-1}$ 与 $(n-m)!^{-1}$

此时无法保证 $m!^{-1}$ 与 $(n-m)!^{-1}$ 存在

将原式化为

$$\frac{\frac{n!}{p^x}}{\frac{m!}{p^y} \cdot \frac{(n-m)!}{p^y}} p^{x-y-z} \bmod p^a$$

扩展 Lucas 定理

需保证

$$\gcd\left(\frac{n!}{p^x},p^a\right)=\gcd\left(\frac{m!}{p^y},p^a\right)=\gcd\left(\frac{(n-m)!}{p^z},p^a\right)=1$$

那么 $\frac{m!}{p^y}$ 与 $\frac{(n-m)!}{p^z}$ 逆元必然存在

问题进一步转化为求解

$$\frac{n!}{p^x} \bmod p^a$$

先考虑 $n! \bmod p^a$

将 $n!$ 展开

$$n!=1\times2\times3\times\cdots\times n=\left(p\times 2p\times\cdots\times\left\lfloor\frac{n}{p}\right\rfloor p\right)\times(1\times2\times\cdots\times(p-1)\times(p+1)\times\cdots)$$

由于 $1\sim n$ 中存在 $\left\lfloor\frac{n}{p}\right\rfloor$ 个 p 的倍数

扩展 Lucas 定理

$$n! = p^{\left\lfloor \frac{n}{p} \right\rfloor} \times \left\lfloor \frac{n}{p} \right\rfloor! \times \left(\prod_{\substack{i=1 \\ (i,p)=1}}^{n} i \right)$$

对于 $k \in \mathbb{Z}$ 有

$$\prod_{\substack{i=1 \\ (i,p)=1}}^{p^a} i \equiv \prod_{\substack{i=1 \\ (i,p)=1}}^{p^a} (k\,p^a + i) \pmod{p^a}$$

那么

$$\prod_{\substack{i=1 \\ (i,p)=1}}^{n} i \equiv \left(\prod_{\substack{i=1 \\ (i,p)=1}}^{p^a} i \right)^{\left\lfloor \frac{n}{p^a} \right\rfloor} \times \left(\prod_{\substack{i=1 \\ (i,p)=1}}^{n \bmod p^a} i \right)$$

$\prod_{i=1,(i,p)=1}^{n} i$ 存在循环，循环了 $\left\lfloor \frac{n}{p^a} \right\rfloor$ 次，剩余尾部为 $\prod_{i=1,(i,p)=1}^{n \bmod p^a} i$

即

扩展 Lucas 定理

$$n! \equiv \left(p^{\lfloor \frac{n}{p} \rfloor} \times \left\lfloor \frac{n}{p} \right\rfloor! \times \left(\prod_{\substack{i=1 \\ (i,p)=1}}^{p^a} i \right)^{\lfloor \frac{n}{p^a} \rfloor} \times \left(\prod_{\substack{i=1 \\ (i,p)=1}}^{n \bmod p^a} i \right) \right) \pmod{p^a}$$

$$\Rightarrow \frac{n!}{p^{\lfloor \frac{n}{p} \rfloor}} \equiv \left(\left\lfloor \frac{n}{p} \right\rfloor! \times \left(\prod_{\substack{i=1 \\ (i,p)=1}}^{p^a} i \right)^{\lfloor \frac{n}{p^a} \rfloor} \times \left(\prod_{\substack{i=1 \\ (i,p)=1}}^{n \bmod p^a} i \right) \right) \pmod{p^a}$$

对于 $\left(\left\lfloor \frac{n}{p} \right\rfloor\right)! \pmod{p^a}$ 继续递归求解，但 $\left\lfloor \frac{n}{p} \right\rfloor!$ 中依然可能包含 p 的因子

令 $\mathbf{g}(n)$ 为 $n!$ 中包含 p 的因子数量

$$\mathbf{g}(n) = \left\lfloor \frac{n}{p} \right\rfloor + \mathbf{g}\left(\left\lfloor \frac{n}{p} \right\rfloor\right)$$

扩展 Lucas 定理

令

$$\mathbf{F}(n,p,p^a) = \frac{n!}{p^{\mathrm{g}(n)}} \pmod{p^a}$$

即 $\mathbf{F}(n,p,p^a)$ 在计算 $n!$ 以及递归时，忽略 **所有** p 的倍数

在计算 $n!$ 忽略了 $\left\lfloor\frac{n}{p}\right\rfloor$ 次，计算 $\left\lfloor\frac{n}{p}\right\rfloor!$ 时忽略了 $\left\lfloor\frac{\left\lfloor\frac{n}{p}\right\rfloor}{p}\right\rfloor$ 次

总共忽略了$\mathbf{g}(n)$ 次

那么

$$\binom{n}{m} \equiv \frac{\mathbf{F}(n,p,p^a)}{\mathbf{F}(m,p,p^a)\,\mathbf{F}(n-m,p,p^a)}\, p^{\mathbf{g}(n)-\mathbf{g}(m)-\mathbf{g}(n-m)} \bmod p^a$$

即只在最后求出组合数时考虑 p 的幂次贡献

对于一次询问时间复杂度约为 $O(\mathrm{M}\log\mathrm{M})$

Kummer 定理

Kummer 定理 (Kummer's theorem) 指出，给定 $n \geq m \geq 0$ 和质数 p

p 在 $\binom{n}{m}$ 中的幂次，恰为 p 进制下 $n-m$ 的 **借位次数**

p 在 $\binom{n+m}{m}$ 中的幂次，也恰为 $n+m$ 在 p 进制下的 **进位次数**

同时有

$$v_p\left(\binom{n}{m}\right) = \frac{S_p(m) + S_p(n-m) - S_p(n)}{p-1}$$

证明

先证等式

上文 **扩展 Lucas 定理** 证明中 $\mathbf{g}(n)$ 即为 $v_p(n!)$ 那么

$$v_p\left(\binom{n}{m}\right) = \mathbf{g}(n) - \mathbf{g}(m) - \mathbf{g}(n-m) = v_p(n!) - v_p(m!) - v_p\big((n-m)!\big)$$

根据 **Legendre 公式**

Kummer 定理

$$v_p\left(\binom{n}{m}\right)=\frac{n-S_p(n)}{p-1}-\frac{m-S_p(m)}{p-1}-\frac{(n-m)-S_p(n-m)}{p-1}$$

$$=\frac{S_p(m)+S_p(n-m)-S_p(n)}{p-1}$$

考虑 $n-m$ 在 p 进制下的借位情况

当考虑到第 i 个 p 进制位时，若第 i 位需要借位给第 $i-1$ 位时有

$$\left\lfloor\frac{n}{p^i}\right\rfloor-\left\lfloor\frac{m}{p^i}\right\rfloor=\left\lfloor\frac{n-m}{p^i}\right\rfloor+1\Rightarrow\left\lfloor\frac{n}{p^i}\right\rfloor-\left\lfloor\frac{m}{p^i}\right\rfloor-\left\lfloor\frac{n-m}{p^i}\right\rfloor=1$$

$\left\lfloor\frac{n}{p^i}\right\rfloor$ 表示 n 截去第 i 位后的数位（显然至多借位 1）

考虑所有的 i，借位次数即为

$$\begin{array}{cccccccc} & 2 & 3 & 2 & ? & ? & ? \\ - & 1 & 0 & 4 & ? & ? & ? \\ \hline & 1 & 2 & 3 & ? & ? & ? \end{array}$$

如上为 $p=5$ 时的情况

在绿分数位产生借位

此时

$(23)_5-(10)_5=(12)_5+1$

更低位借位并不影响该判定

Kummer 定理

$$\sum_{i=0}^{\infty}\left(\left\lfloor\frac{n}{p^i}\right\rfloor-\left\lfloor\frac{m}{p^i}\right\rfloor-\left\lfloor\frac{n-m}{p^i}\right\rfloor\right)=v_p\left(\binom{n}{m}\right)$$

类似的考虑 $n+m$ 的进位情况

当考虑到第 i 个 p 进制位时，若第 i 位获得第 $i-1$ 位进位时有

$$\begin{array}{cccccccc} & 2 & 3 & 2 & ? & ? & ? \\ + & 1 & 0 & 4 & ? & ? & ? \\ \hline & 3 & 4 & 1 & ? & ? & ? \end{array}$$

$p=5$

$$\left\lfloor\frac{n}{p^i}\right\rfloor+\left\lfloor\frac{m}{p^i}\right\rfloor=\left\lfloor\frac{n+m}{p^i}\right\rfloor-1\Rightarrow\left\lfloor\frac{n+m}{p^i}\right\rfloor-\left\lfloor\frac{n}{p^i}\right\rfloor-\left\lfloor\frac{m}{p^i}\right\rfloor=1$$

考虑所有的 i，借位次数即为

$$\sum_{i=0}^{\infty}\left(\left\lfloor\frac{n+m}{p^i}\right\rfloor-\left\lfloor\frac{n}{p^i}\right\rfloor-\left\lfloor\frac{m}{p^i}\right\rfloor\right)=v_p\left(\binom{n+m}{m}\right)$$

Kummer 定理

在考虑 $n+m$ 在 p 进制下进位时有

$$\left\lfloor\frac{n+m}{p^i}\right\rfloor-\left\lfloor\frac{n}{p^i}\right\rfloor-\left\lfloor\frac{m}{p^i}\right\rfloor=1$$

这与 $(n+m)-n$ 在 p 进制下 **借位次数** 相同

所以 p 在 $\binom{n+m}{m}$ 中的幂次，恰为 $n+m$ 在 p 进制下需要 **进位次数**

也为 $(n+m)-n$ 在 p 进制下 **借位次数**

Kummer 定理 可推广到多项式系数

$$v_p\left(\binom{n}{m_1,m_2,m_3,\cdots,m_k}\right)=v_p\left(\frac{n!}{\prod_{i=1}^{k}(m_i!)}\right)=\frac{1}{p-1}\sum_{i=1}^{k}\left(S_p(m_i)-S_p(n)\right)$$

#3226、扩展Lucas定理

题目描述

求

$$\binom{n}{m} \bmod p$$

其中 $\binom{n}{m}$ 为 n 选 m 的组合数

输入格式

一行三个整数 n, m, p ,含义由题所述

输出格式

一行一个整数

数据规模

对于 100% 的数据, $1 \le m \le n \le 10^{18}, 2 \le p \le 10^6$

不保证 p 是质数

```cpp
LL calc(LL n, LL p, LL pa)
{
    if (!n)
        return 1;
    LL res1 = 1, res2 = 1;
    for (int i = 1; i <= pa; i++) // 计算循环节
        if (i % p)
            res1 = (res1 * i) % pa;
    for (int i = 1; i <= n % pa; i++) // 计算多余尾部
        if (i % p)
            res2 = (res2 * i) % pa;
    return calc(n / p, p, pa) * qpow(res1, n / pa, pa) % pa * res2 % pa;
}
LL C(LL n, LL m, LL p, LL pa)
{
    if (m > n)
        return 0;
    if (m == 0 || n == m)
        return 1;
    LL fn = calc(n, p, pa), fm = calc(m, p, pa), fnm = calc(n - m, p, pa), tnm = n - m;
    int k = 0;
    while (n) // g(n)
        n /= p, k += n;
    while (m) // g(m)
        m /= p, k -= m;
    while (tnm) // g(n-m)
        tnm /= p, k -= tnm;
    return fn % pa * inv(fm, pa) % pa * inv(fnm, pa) % pa * qpow(p, k, pa) % pa;
}
```

LAB
实验舱
青少年编程
走近科学 走进名校

谢谢观看