动态规划与计数问题

北京大学 洪华敦

概率

- 概率的定义：一个事件发生的情况的占比

- 假设所有情况构成的集合为 S，事件 A 发生的情况构成的集合为 T，那么记 A 的概率 P(A)=|T|/|S|

- 例如掷骰子问题，设事件 A 为：投出来的点数为偶数

- 则 S={1,2,3,4,5,6}

- T={2,4,6}

- P(A)=|T|/|S|=1/2

概率

- 在算法竞赛中研究的都是离散概率，即 S 的大小是有限的
- 假设有一个整数随机变量 X，则有：
- $P(X=K)=P(X>=K)-P(X>K)$

例子

- 求投 n 次骰子，最小的点数等于 2 的概率

期望

- 假设事件 A 的收益为 W(A)
- 则 $E[X] = \sum P(A)W(A)$

期望的性质

- 期望的线性性：E[X+Y]=E[X]+E[Y]
- 同理，如果 X1+X2=X，则 E[X]=E[X1]+E[X2]
- 假设 X 是一个随机正整数变量，则：
  - E[X]=sum(P(X>=i))   (i=1..inf)

一些概率期望的计算

- 每次随机一个 [1,n] 的数，问期望随机几次能随机出所有数。
- 随机一个长度为 n 的排列p，p[i] 是 p[1..i] 中最大的的概率。

[NOIP2016]换教室

- 一天一共有 n 节课，第 i 节课默认在第 c[i] 间教室上，可以申请第 i 节课到第 d[i] 间教室上，申请有 k[i] 的概率通过，从第 x 间教室到第 y 间教室有距离 dis[x][y]，你最多可以申请 m 门课，要求跑的距离尽量少

- n,m<=2000

[NOIP2016]换教室

- E[跑动距离]=Sum(E[第 i 节课下课后的跑动距离])
- 第 i 节课下课后的跑动距离，只与第 i , i+1 节课是否申请换教室有关
- O(nm) DP

随机游走

- 一条长度为 n 的链，从一端走到另一端的期望时间
- 一个 n 个点的完全图，求从 S 走到 T 的期望时间

硬币游戏

- 给定 n 个硬币，第 i 个硬币的价值是 w[i]，每次随机取走一个硬币，获得收益是左右两个硬币的价值的乘积，求期望总价值

- 1<=n<=200, 1<=w[i]<=10^9

硬币游戏

- 根据期望的线性性，变成算硬币 i 和硬币 j 的乘积产生贡献的概率
- 设 i<j，则概率是 2/((j-i+1)(j-i))
- O(n^2)
- 甚至能做 n<=10^5

线性与等价性的应用

- 很多时候，我们可以利用题目里对象的等价性和期望的线性性来计算一些东西
- 例如：有 n 个变量 X[1…n]，如果他们都是同分布的，且有 E[Sum(X)]=1，那么 E[X[1]]=1/n

例子

- 随机一个 1…n 的排列 p，p[i] 是 p[1…i] 中最大值的概率
- 随机一个 1…n 的排列 p，p[i] 是 p[1…n] 中前 k 大的概率
- 给定一棵 n 个点的有根树，一开始每个点都是白色，每次等概率随机选择一个白点，将他的整个子树都染成黑色，问期望随机几次能把整棵树染黑 (n<=10^5)
- 有 n 堆石头，每堆个数为 a[i]，现在每次等概率随机选一个石头，然后将那个石头所在的那堆石头全扔了，直到第一堆石头被扔掉，问期望扔几次 (n<=10^5)
- 对于一个 1…n 的排列 p，定义他的价值为：所有满足 p[i]>max(p[i-1],p[i+1]) 的 i 的和，现在随机一个排列 p ，求价值的期望

例子

- 等概率随机一个长度为 n 的 01串，定义 01 串的价值为：所以极长全1连续子串的长度的平方和，求价值的期望

- 等概率随机一个排列 p[1…n]，定义 f(p) 为有多少个 i 满足 p[i] 是 p[1…i] 的最大值，求 f(p)，f(p)^2 的期望

字符串拼接

- 给定一个字符串 T，和一个字符串序列 S[1..n]
- 你将选择 k 次来构造一个字符串，每次随机从 S[1..n] 中随机选一个字符串拼接到你已构造的串后面
- 求 T 在你构造的串里的期望出现次数
- |T|<=50，n<=50, |S[i]|<=50, k<=10^(12)

字符串拼接

- f[k]表示拼接 k 次后的期望次数
- 当 $k > |T|$ 时有 $f[k+1]-f[k]=f[k+2]-f[k+1]$

拓扑序

- 给定一张 n 个点的有向图，f(E) 表示边集 E 的拓扑序个数，令 E 取编该图的边集的所有子集，求 f(E) 的和
- n<=20

与位运算有关的计数

- 一般是利用位运算的两个性质：
- 1. 独立性
- 2. 高位优先

XOR之和

- 给定数组 a[1..n]，求 sum(a[i]^a[j]) 其中 1<=i<j<=n
- 1<=n<=10^5，0<=a[i]<2^30

XOR之和

- 枚举第 i 位，假设 a[1..n] 里第 i 位为 0 的有 x 个，那么贡献就是 x*(n-x)*2^i
- 时间复杂度：O(nlogn)

序列的值

- 定义 f(b[1..m]) 为：有几个 [0,m-1] 中的 i 满足：b[1..i]的异或和 < b[1..i+1] 的异或和

- 给定一个数组 a[1..n]，对于 a[1..n] 的每个子序列 b[1..m]，求 f(b[1..m]) 的和

- 1<=n<=10^5，0<=A[i]<2^31，答案对 998244353 取模

序列的值

- 考虑一下 x<(x^S) 的条件
- 那么肯定是 : 对于 S 的最高位，x 这一位必须为 0
- 考虑 dp ， 令 f[i][j][v] 表示 a[1..i]有几个子序列满足异或和的第 j 位为 v
- 对于 a[i]，设最高位为第 k 位，那么给答案加上 f[i-1][k][0]*2^(n-i)
- 时间复杂度：O(nlogn)

异或排序

- 给定一个长度为 n 的非负整数序列 A[1..n]，求有多少个不同的整数 S 满足以下条件：
- (1).0<=S< (2^60)
- (2).对于所有 [1,n-1] 中的 i ，有 (A[i]^S)<=(A[i+1]^S)
- 1<=n<=20，0<=A[i]< (2^60)

异或排序

- 考虑一下对于 a,b ，有哪些 S 满足 (a^S)<=(b^S)
- 找到 a^b 的最高位，如果这一位里 a 是0，b 是 1，那么 S 这一位必须是0，否则必须是 1
- 所以每个 (A[i]^S)<=(A[i+1]^S) 其实是限制了 S 的某一位的值
- 把限制全部搞出来算一下就行了
- 时间复杂度：O(nlogn)

和的XOR

- 给定数组 a[1..n]，求所有的 (a[i]+a[j]) xor起来的值，其中 1<=i<j<=n
- 给定数组 a[1..n]，求所有的 (a[i]+a[j]) xor起来的值，其中 1<=i,j<=n
- n<=10^5. a[i]<=10^9

和的XOR

- 考虑对于第 i 位，计算有多少 (a[x]+a[y]) 的第 i 位为 1
- 那么一定有 (a[x]+a[y]) mod 2^(i+1) >=2^i
- 所以 (a[x] mod 2^(i+1))+(a[y] mod 2^(i+1)) 要么在 [2^i,2^(i+1)-1]里，要么在 [3*(2^i),2^(i+2)-1]里
- 用个two pointer扫一下就好了
- 时间复杂度：O(nlognlogn)
- 用归并排序的话可以做到 O(nlogn)

区间统计

- 给定 a[1…n]，l[1…n]，r[1…n]
- 求有几个数组 b[1…n]，满足 l[i]<=b[i]<=r[i]，且 a[i]&b[i]=a[i]，且 b[i]>b[i-1]
- n<=100

XOR之和加强版

- 给定数组 a[1..n]，求sum((a[i]^a[j])*(a[j]^a[k])*(a[i]^a[k]))
- 其中 1<=i<j<k<=n
- 1<=a[i]<=10^9, 1<=n<=10^5

XOR之和加强版

$$\sum_{i<j<k}(a_i \oplus a_j)(a_j \oplus a_k)(a_i \oplus a_k)$$

设 $bit(x,i)$ 表示 x 二进制下第 i 位是 0 还是 1

原式可以写成：

$$\sum_{i<j<k}(a_i \oplus a_j)(a_j \oplus a_k)\sum_{l=0}^{29}2^l(bit(a_i,l) \oplus bit(a_k,l))$$

枚举一下 $bit(a_i,l)$ 和 $bit(a_k,l)$，使得 $bit(a_i,l) \oplus bit(a_k,l) = 1$

则原式可以写成：

$$\sum_{l=0}^{29}2^l\sum_{i<j<k}(a_i \oplus a_j)(a_j \oplus a_k) = \sum_{l=0}^{29}2^l(\sum_{i<j}(a_i \oplus a_j))(\sum_{j<k}(a_j \oplus a_k))$$

其中 a_i, a_k 要满足我们枚举的 $bit(a_i,l)$ 和 $bit(a_k,l)$

相当于要对于每个 j 计算，前面和后面满足条件的数和它异或后的和，这个是可以$O(n\log n)$ 的

时间复杂度：$O(n\log^2 n)$

某xor统计

- 给定 n,k，求有多少长度为 n 的数组，满足 a[i]<=k，且 a[1…n] 异或起来为 0
- 普通版：1<=n<=10^5，1<=k<=10^9
- 加强版：1<=n,k<=10^9

某xor统计2

- 给定 n,S，求有几个非负整数数组 a[1…n]，满足 Sum(a)=S，且 a[1…n] 的异或和为 0
- 1<=n<=50
- 1<=S<=10^18

容斥原理

- 容斥原理： $$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{S\subseteq\{1,2..n\}} (-1)^{|S|-1} \left|\bigcap_{i\in S} A_i\right|$$

容斥原理

- 于是根据期望的线性性，有：
- $$E\left[\left|\bigcup_{i=1}^{n} A_i\right|\right] = E\left[\sum_{S\subseteq\{1,2..n\}} (-1)^{|S|-1}\left|\bigcap_{i\in S} A_i\right|\right]$$
  $$E\left[\left|\bigcup_{i=1}^{n} A_i\right|\right] = \sum_{S\subseteq\{1,2..n\}} (-1)^{|S|-1} E\left[\left|\bigcap_{i\in S} A_i\right|\right]$$

Min Max容斥

- Min Max容斥：
- $E[Max(x_1, x_2 \ldots x_n)] = \sum_{S \subseteq \{1,2\ldots n\}} (-1)^{|S|-1} E[Min_{i \in S}(x_i)]$

[PKUWC2018]随机游走

- 给定一棵树，有 Q 次询问，每次询问给定起点 x 和一个点集 S，求从 x 开始随机游走经过 S 中所有点的期望时间

- 1<=n<=18 , 1<=Q<=5000

[PKUWC2018]随机游走

- Min Max 容斥后变成水题
- O(2^n*n^2+Q)

二维凸包

- 给定二维平面 n 个点，随机取一个点集，求期望凸包点数
- n<=100

三维凸包

- 给定三维空间 n 个点，随机取一个点集，求期望凸包点数(没有四点共面)
- n<=100

三维凸包

- 利用欧拉公式
- 面+点-边=2

区间的统计

- 给定 n 个区间，从中选出三个区间，使得他们有公共部分，求方案数
- n<=100000

random

- 给定一个长度为 n 的排列 A，每次在所有 A[i+1]>A[i] 的 i 中等概率随机一个 i，然后交换 A[i] 和 A[i+1]，这一步的代价是 i，求排完序后的期望代价和
- 1<=n<=18

random

- 分别考虑每个数贡献的代价和，这时对于其他数我们只关心他们与这个数的大小关系

- 可以O(n^2*2^n)递推

背包问题

- 有一个大小为 n 的背包，有 n 个物品，第 i 个物品的大小为 i，且有 i 个，求装满这个背包的方案数有多少
- 1<=n<=10^5

背包问题

- 分大于根号和小于根号的物品去做
- 小于根号的暴力背包
- 大于根号的等价于是个无限背包
- f[i][j]表示选了 i 个物品，和为 j 的方案数
- O(n^1.5)

集合统计

给定三个正整数n, m, k，考虑所有大小为k的**正整数可重集合**$\{a_1, a_2 ... a_k\}$，要求$\sum_{i=1}^{k} a_i = n$，定义这样的一个可重集合的权值为：

$$\sum_{i=1}^{k} a_i^m$$

求所有满足条件的可重集合的权值之和，由于答案可能过大，你只需要输出答案对$10^9 + 7$取模后的值。

- n,m,k<=5000