杂题选讲

___debug

2017 年 1 月 23 日

#1, Matching

给定两个含小写字母和通配符 ? 的字符串 S, T.
通配符 ? 可以匹配任意一个字符.
请你求出 T 在 S 中的出现次数.

$|T| \leq |S| \leq 2 \times 10^5$

这题应该是不好用普通的字符串算法解决的, 但是我们可以换一种思路.

不妨将 ? 和 26 个字母依次映射到 0 到 26 的整数. 构造这样一个函数

$$f(i) = \sum_{j=0}^{|T|-1} S_i T_i (S_i - T_i)^2$$

那么 T 从 S 的第 i 位开始可以匹配当且仅当 $f(i) = 0$.

然后不妨把 S 串反过来, 发现是一个卷积的形式.

FFT.

#2, 从今以后

定义一个数列是合法的，当且仅当这个数列的每一个子串，都至少存在一个只出现过一次的元素. 给你 T 个序列，要你判断这些序列是否合法.

$$\sum n \leq 2 \times 10^6, n \leq 2 \times 10^5$$

这题的一个线段树做法十分显然. 不过还有另一个比较好的做法.

首先先预处理出 $prev(i)$ 和 $next(i)$ 表示与 A_i 的值的相同的前一个和下一个. 利用这个我们能快速判断 A_i 是否在 $[l, r]$ 中只出现了一次.

现在我们要 $check(l, r)$, 表示要检查 $[l, r]$ 这个区间代表的序列是否合法. 显然我们可以找到在这个区间里只出现过一次的元素 A_p, 然后递归 $check(l, p-1)$ 和 $check(p+1, r)$ 即可. 但是这样无法保证复杂度, 可能退化为 $O(N^2)$

对于区间 $[l, r]$, 我们同时从 l 和 r 出发去找满足条件的 p, 然后一旦找到一个就往下分治, 这样复杂度就对了. 考虑 $[l, r]$ 这个区间的耗时, 与分治分出来的那个较小的区间是同阶的.

启发式合并的逆过程.

#3, 向再见说再见

有两个队伍，每个队伍都有 n 个人，并且每个人有一个能力值. 每个人的能力值互不相同.

现在要你将这些队伍两两配对比赛，对于一组比赛，如果一方的能力值高于另一方，则给这一方属于的队伍加一分.

问你满足最后两队的分差为 K 的配对的方案数有多少种，对 10^9+7 取模.

$K \le n \le 2000$

首先对两个队的数组分别从小到大排序.

考虑这类题的经典思路: 设 $dp(i,j)$ 表示考虑前 i 个人赢了 j 场, 并且不考虑输的人的贡献, 这样的方案数.

然后显然 $dp(n,i)\times(n-i)!$ 表示的是这个队赢了 $\geq i$ 场的方案数, 注意赢了 j 场的方案被算了 $\binom{j}{i}$ 次.

所以容斥一下就好了. 设 $g(i)$ 为赢得场数 $=i$ 的方案数, 则有

$$g(i)=dp(n,i)\times(n-i)!-\sum_{j>i}\binom{j}{i}g(j)$$

#4, BZOJ 3836, POI2014, Tourism

给你一个 N 个点 M 条边的无向图，每个点有点权 A_i. 保证任意两点之间距离不超过 10 (简单路径).

现在要你选取一些点，使得每个点要么自己被选，要么存在至少一个相邻的点被选.

$$N \leq 20000, M \leq 25000, 0 \leq A_i \leq 10000$$

考虑如何利用 “距离不超过 10” 这个限制.

不妨把生成树搞出来, 那么这个树的深度 ≤ 10. 一个很显然的想法是 DP 时, 对每个点设状态 $dp(u,S)$, 表示以 u 为根的子树, 当 u 到根这条路径上的点的选取情况是 S 时的最小代价.

由于, 一个点可能会对通过非树边覆盖祖先, 对于 S 中每个未被选取的点, 我们可以再记录一下它是否已经被覆盖. 对于一个点状态数是 3^{10} 的.

用 3 进制来表示 S. 规定一下 S 中某一位的值的意义:

- 0: 这个点被选取了
- 1: 这个点未被选取, 但已经被覆盖
- 2: 这个点未被选取, 且仍未被覆盖

在 DFS 序上做即可. 转移并不困难, 注意的是当 DFS 序中前一个点不是当前点的父亲时, 要把 dp 值维护一下.

使用滚动数组, 空间复杂度 $O(N+M+3^{10})$, 时间复杂度 $O(M+3^{10}N)$. 记得对每个连通块都要做一遍.

#5, BZOJ 1797, AHOI2009, 最小割

给你一个有向图, 边有权. 同时给定 S, T 表示源和汇.
要你求出每条边是否

- 至少出现在一种最小割的割集中
- 出现在所有最小割的割集中

$N \le 4000, M \le 60000$

先跑一遍网络流. 首先两问的前提都是 (u,v) 满流.

先考虑第一问. 一条边 (u,v) 有可能在最小割中, 当且仅当 u 有可能在 $\mathcal{S}$ 集中, **同时** v 有可能在 $\mathcal{T}$ 集中.

不妨假设, 我们通过连 $(S,u,+\infty),(v,T,+\infty)$, 强制 u 在 $\mathcal{S}$ 集中, v 在 $\mathcal{T}$ 集中, 那么 u 到 v 一定不能有路径.

注意到此时在残量网络中 (u,v) 一定满流, 所以 (u,v) 的反向弧会出现在残量网络中. 所以如果 u 到 v 有路径, 那么将会形成一个强连通分量.

所以只要 u,v 不在同一个强连通分量中, (u,v) 就可能在最小割里.

再考虑第二问. 第二问实际上就是说 u 一定在 $\mathcal{S}$ 集中, 而 v 一定在 $\mathcal{T}$ 集中. 也就是说如果连边 $(u,T,+\infty)$ 强制 u 在 $\mathcal{T}$ 集中, 那么一定会出现新的增广路, 也就是说残量网络中 S 到 u 有路径.

注意到原图 S 到 u 一定有一条增广路, 所以在残量网络中 u 到 S 一定有一条路径. v 同理.

结合上面两个条件, 如果 u 和 S 在一个强连通分量中, v 和 T 在一个强连通分量中, 那么 (u,v) 一定任何一个最小割里.

#6, SPOJ TRANSP2

$2^a 2^b$ 的二维数组在内存中的存储方式为先存第一行，再存第二行，以此类推.

现在你需要将其转置，举个例子

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

问最少需要交换多少次元素. 对 1000003 取模. 共 T 组数据.

$T \leq 4 \times 10^5, 0 \leq a + b \leq 10^6$

矩阵转置时, 每个位置的数都会有一个唯一的目标位置, 那么将这个点数为 2^{a+b} 的有向图画出来一定是由很多环组成的. 显然最少的交换次数为 $2^{a+b}-c$, 其中 c 表示这张图中环的个数.

但是暴力找环显然不可行, 得好好利用 2 的次幂这个条件.

不妨将每个点在内存中的位置表示为一个 pair, 如 3 行 5 列表示为 $(11, 101)$. 那么转置一下, 它就变成了 $(101, 11)$, 可以理解为交换横纵坐标, 但是如果把横纵坐标连在一起的话 (即 $11101 \to 10111$), 也可以看成是循环左移 a 位.

然后由于坐标的范围是 2^a 和 2^b, 也就是说每个点都可以对应 $[0, 2^{a+b})$ 中的一个二进制数, 可以将问题转化为:

一个长度为 $a+b$ 的环, 给每个点染上黑白两种颜色. 每次可以将一个环逆时针旋转 a 个单位, 如果一种染色方案能通过旋转变为另一种染色方案, 那么这两种方案视为同一种. 求不同的方案数.

上面这个问题的答案就是环的个数 c.

这就是一个比较明显的 Burnside 了, 置换群 G 的大小为 $\frac{\operatorname{lcm}(a+b,a)}{a} = \frac{a+b}{\gcd(a,b)}$.

为了叙述方便, 记 $g = \gcd(a,b)$, $n = \frac{a+b}{g}$.

根据 Burnside 引理, 有

$$\begin{aligned} N(G,\mathcal{C}) &= \frac{1}{|G|}\sum_{f\in G}\mathcal{C}(f) \\ &= \frac{1}{n}\sum_{k=1}^{n} 2^{\gcd(ka,n)} \\ &= \frac{1}{n}\sum_{k=1}^{n} (2^g)^{\gcd(k,n)} \end{aligned}$$

然后就是非常套路的枚举约数再用 φ 来算贡献了.

#7, BZOJ 3202, SDOI2013, 项链

定义每种颜色可以用一个三元组 (x, y, z) 表示，满足 $0 < x \le y \le z \le A$ 且 $\gcd(x, y, z) = 1$.

现在要用这些颜色对一个 N 个点的环染色. 循环同构，并且相邻的点的颜色必须不同.

求不同染色方案的个数. 对 $10^9 + 7$ 取模. 共 T 组数据.

$T \le 10, N \le 10^{14}, A \le 10^7$

首先我们要求出不同的颜色有多少种. 不难求出 $g(d)$ 表示 $d \mid \gcd(x,y,z)$ 的三元组有多少个:

$$g(d) = \binom{\left\lfloor \frac{A}{d} \right\rfloor + 2}{3}$$

然后设 $f(d)$ 为 $d = \gcd(x,y,z)$, 显然有 $g(n) = \sum_{n|d} f(d)$. 莫比乌斯反演一下就是:

$$f(n) = \sum_{n|d} \mu\left(\frac{n}{d}\right) g(d)$$

我们要求的就是 $f(1)$. 预处理 μ 之后可以 $O(\sqrt{A})$ 来求.

接下来就是 Burnside 了. Burnside 完了之后, 转化为一个子问题:

给你一个 n 个点的环, 用 m 种颜色染色, 有多少种方案满足相邻的两个点不同色?

这貌似也是一个经典问题.
设 a_n 为 n 个点时的答案, 显然有

$$\begin{aligned} a_1 &= 0 \\ a_2 &= m(m-1) \\ a_n &= (m-2)a_{n-1} + (m-1)a_{n-2}, (n \geq 3) \end{aligned}$$

解特征方程得到通项

$$a_n = (m-1)^n + (-1)^n(m-1)$$

#8, BZOJ 3864, Hero meet devil

给你一个只由 ACGT 组成的字符串 S, 对于每个 $0 \le k \le |S|$, 问有多少个只由 ACGT 组成的长度为 M 的字符串 T, 使得 $\mathrm{LCS}(S, T) = k$.

$|S| \le 15, M \le 10^3$

DP 套 DP.

考虑算 LCS 的 DP

$$dp(i,j) = \max\left(dp(i-1,j-1) + [T_i = S_j], dp(i-1,j), dp(i,j-1)\right)$$

这题的 $|S|$ 很小, 不妨对于每个 i, 将 j 不同时的 DP 值都记录下来, 计个数即可.

但是直接记 DP 值状态数会爆炸, 但是注意到相邻的 DP 值只会差 1, 所以我们可以用 $2^{|S|}$ 的状态数将这些值记录下来.

#9, Codeforces 603E, Pastoral Oddities

给你 N 个点的一张图，初始没有边.

每次多给你一条带权无向边，询问是否可以选择某些当前给你的这些边，使得每个点的度数都为奇数，且边权最大值最小. 若不存在输出 -1 .

一共加 M 次边，每加一次输出一次. 允许离线.

$N \leq 10^5, M \leq 3 \times 10^5$

点数为奇数时无解, 试一下发现偶数的连通块似乎都有解. 其实就是这样. 如果点数是偶数, 搞出生成树以后从下往上贪心地选, 最后只有根有可能度数为偶, 很容易可以证明这是不可能的.

所以只要当前所有连通块的点数为偶数就有解.

用 LCT 维护最小生成树即可. 但是最小生成树上的某些边不一定被用到. 我们再维护一个堆, 表示哪些边在 LCT 中. 我们只需取出边权最大的边, 看看两端是不是大小都为偶数, 如果是的话就可以删了.

用 LCT 维护虚边信息.

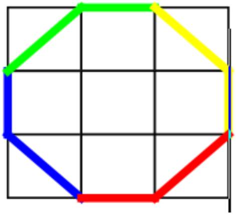
#10, 题目

请你在 $n \times n$ 的网格中画一个点都在格点上的凸多边形，要求凸多边形按任意方向旋转 90 度以后和原来一样.

询问最多能画多少边的凸多边形.

$n \le 10^{18}$

首先考虑如何暴力.



如上图，根据题目要求，四种颜色代表的那一部分必须完全相同.

现在问题就转化为了用尽可能多的斜率递增的线段去覆盖一个部分 (其他的是对称的). 注意 “一个部分” 需要满足长 $+$ 宽 $= N$.

我们将一种线段描述为 (x, y) (长 x 宽 y), 那么容易发现我们一定要选 $gcd(x, y) = 1$ 的线段.

既然只要求 $\sum x + y \le N$ (注意可以小于, 因为一定会选一个 $x = 1, y = 0$ 的线段, 我们可以将这条线段延长); 所以我们每次选出 $gcd(x, y) = 1$ 且 $x + y$ 最小的数对即可.

可以先枚举一个 $sum = x + y$, 然后从 1 到 sum 枚举 x, 检查 $gcd(x, sum - x)$ 是否等于 1 即可; 加的时候如果当前的 $\sum x + y$ 即将 $> N$ 了, 那么就 `break`.

这题只要计数，考虑优化.
注意到

$$gcd(x, sum - x) = 1 \iff gcd(x, sum) = 1$$

所以对于每个 sum，最多可以贡献答案 $\varphi(sum)$ 条线段，同时也会对 $\sum x + y$ 贡献 $\varphi(sum) \times sum$.
这样，我们只需二分出最大的 P 满足

$$\sum_{sum=1}^{P} \varphi(sum) \times sum \le N$$

答案即为

$$\left[\left(\sum_{i=1}^{P} \varphi(i) \right) + \frac{N - \sum_{i=1}^{P} \varphi(i) \times i}{P+1} \right] \times 4$$