杂题选讲

任轩笛

PKU

2020 年 7 月 26 日

题意

给你一张连通无向图，你要么找一个三染色，要么找一个奇环使得删去后图仍然连通。

范围

$n \leq 10^5, m \leq 2 \times 10^5$。

原题只要求四染色。

那么先跑一棵生成树。考虑非树边，如果有奇环，直接删了输出就行了。否则非树边就是一张二分图，直接根据树上深度奇偶性、黑白染色的颜色来对图 4 染色就可以了。

三染色可以这样：

随便选一个点染黑，把邻居全染白，然后找一个未访问过的与白点相邻的某个点染黑，继续……

三染色可以这样：

随便选一个点染黑，把邻居全染白，然后找一个未访问过的与白点相邻的某个点染黑，继续……

树上只有黑白边：由树的生成过程立即可得。

非树边没有黑黑边：处理其中较早的点时一定会把另一个点标白。

三染色可以这样：

随便选一个点染黑，把邻居全染白，然后找一个未访问过的与白点相邻的某个点染黑，继续……

树上只有黑白边：由树的生成过程立即可得。

非树边没有黑黑边：处理其中较早的点时一定会把另一个点标白。

这说明如果非树边是二分图，黑点一定全在一侧。

于是染成黑、白 1、白 2 即可。

BoardPainting

SRM 577 1000pts

题意

$n \times m$ 的网格里有一些 # 号需要消除，每次可以选一段横向或纵向的连续 # 号一起消掉，不能选择空格或者已消除的格子。问最少需要消除几次。

范围

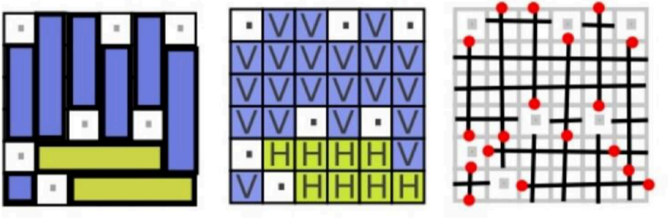
$n, m \leq 50$。

BoardPainting

SRM 577 1000pts

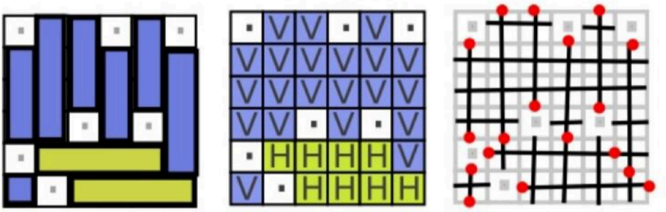
要用尽可能少的 $1 \times k$ 矩形覆盖所有的 #，不能重叠或覆盖到空格。

将横向的记为 H，纵向的记为 V。



要用尽可能少的 $1 \times k$ 矩形覆盖所有的 #，不能重叠或覆盖到空格。

将横向的记为 H，纵向的记为 V。



考虑矩形的个数，那就是每个矩形的头尾总个数 /2。矩形的头尾必然是 V 和 H 之间的边/V 在纵向与空格或边界的边/H 在横向与空格或边界的边（即图中的红点们）。

那么就比较明确了，最小割，属于 S 集表示该格是 V 格，否则是 H 格。

从 s 向每个 # 连边，容量为其横向连接的边界/空格个数。割这条边即表示填成 H。

从每个 # 向 t 连边，容量为其纵向连接的边界/空格个数。割这条边即表示填成 V。

相邻两个 # 之间连双向容量为 1 的边，表示异割的话付出 1 的代价。

求最小割，除以 2 就是答案。

题意

有一个正方形，边长为 m，给出 n 个正方形内的特殊点，求有多少顶点坐标为整数，且都在正方形边上的三角形，覆盖所有特殊点。

范围

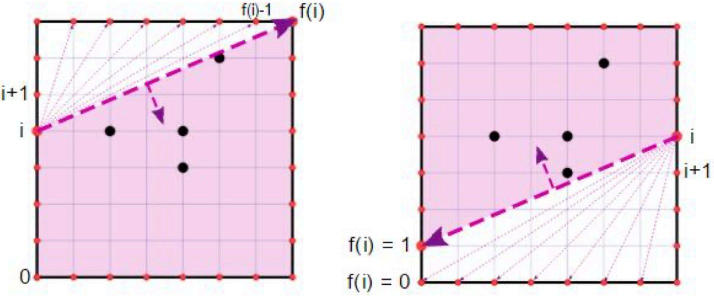
$n \leq 20, m \leq 10^5$.

EnclosingTriangle

SRM 585 1000pts

先把边界点从 0 开始标号，从每个点向前引一条射线，尽量往右拐，满足所有特殊点都在这条射线右侧。设这个点是 a，这条射线另一个端点为 $f(a)$。

求这个的时候把环拉成链再做。



现在就是要求这样的三元组 (a, b, c) 的对数:

$$0 \le a < b < c < N, f(a) \ge b, f(b) \ge c, f(c) \ge a + N$$

现在就是要求这样的三元组 (a, b, c) 的对数：

$$0 \leq a < b < c < N, f(a) \geq b, f(b) \geq c, f(c) \geq a + N$$

显然 f 是单调的，于是枚举 a，关于 a 合法的 $b \in (a, f(a)]$，关于 a 合法的 c 是一个后缀。

现在就是要求这样的三元组 (a, b, c) 的对数：

$$0 \leq a < b < c < N, f(a) \geq b, f(b) \geq c, f(c) \geq a + N$$

显然 f 是单调的，于是枚举 a，关于 a 合法的 $b \in (a, f(a)]$，关于 a 合法的 c 是一个后缀。

每次单调地加/删 b，对一个 b，把 $(b, f(b)]$ 给 ± 1，查询后缀和即可。线段树就能解决。$O(N \log N)$。

BearDestroys

SRM 671 900pts

题意

有一个 $n \times m$ 的矩阵，每个点上写着 S（南）或 E（东），每个格子里种着一棵高度为 2 的树。

一只熊来推树，他推的顺序为第一行从左到右、第二行从左到右⋯⋯

如果他推一棵树时，这个点已经被倒下的树覆盖了，就跳过；如果能按照格子上写的方向推树，就推；否则如果能按另一个方向推树，就推；都不行也跳过。

这样他最后总共会推倒若干棵树。对于所有 2^{nm} 种矩阵，求他推倒的树的数量之和，取模。

范围

$n \leq 13, m \leq 30$。

简单的想法是对行进行状压 DP，一行行处理过去，维护方案数以及数量总和。

简单的想法是对行进行状压 DP，一行行处理过去，维护方案数以及数量总和。

转移时对于没被上一行占用的节点，要枚举它是 S 还是 E。这一步总的复杂度是 $O(3^m)$ 的。

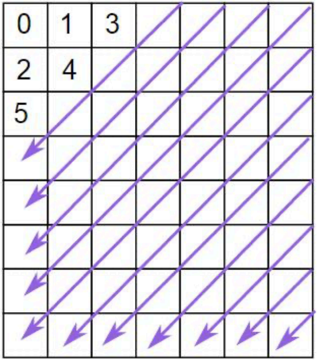
于是复杂度大概是 $O(3^m \times ...)$。

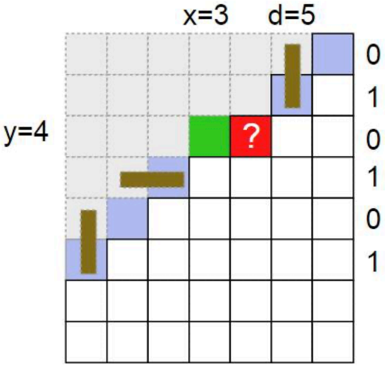
观察这个数据范围，希望状压一列的信息。

然而对于左下、右上两个点，正确顺序是先做右上再做左下，直接一列列做过去会出问题。

BearDestroys

SRM 671 900pts





那么状压左下到右上的对角线，一条条 DP 过去就行了，能正确处理先右上再左下的顺序，且长度是 $\min\{n, m\}$。

复杂度 $O(3^{\min(n,m)} \times ...)$。

题意

有一个 $n \times m$ 的网格，每个格子是草原或者城镇。要用修若干首尾相接的铁路把所有城镇串起来。

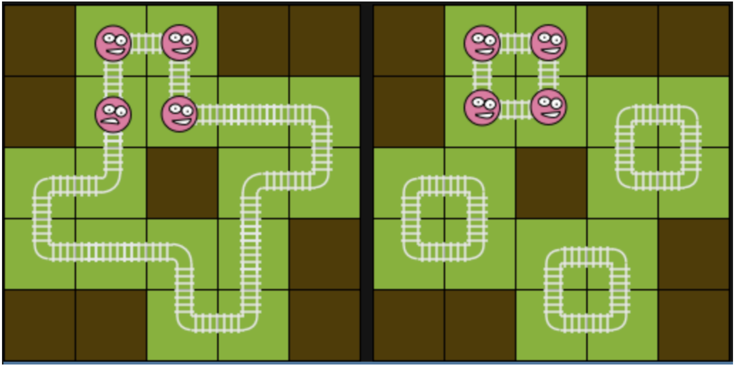
有些城镇里有人，他们希望穿过所在城镇的铁路是弯的。问最多能满足多少人的需求。无解输出-1。

范围

$n, m \leq 25$。

CurvyonRails

SRM 570 900pts



如果不考虑弯不弯的需求，就只要用若干个环把所有城镇串起来，那显然只要每个点度数都是 2 即可。

黑白染色，白点排在左边，黑点排在右边，白点朝相邻的黑点连容量 1 的边，s 朝白点、黑点朝 t 连容量为 2 的边即可。

CurvyonRails

SRM 570 900pts

考虑转角的问题，把每个格子拆成 2 个，一个表示水平方向，一个表示竖直方向。s 朝白点、黑点朝 t 连容量 1 的边，每个白点朝对应方向的黑点连容量为 1 的边。

现在如果能满流的话，所有城镇都是转角了。但是可能会无解。

于是在同一个点拆成的两个点之间连容量为 1、费用为 1 的边，如果这两个点通过这条边“交换”了流量，说明这个城镇的铁路其实是直的。跑最小费用最大流即可。

题意

有一排 n 个数，选出了一个数。
你可以问若干次：某个区间内是否包含了选出的数。
求在所有 $2^{n\times(n+1)/2}$ 种问法中，有多少种能够唯一确定答案。

范围

$n \leq 400.$

考虑问一个区间怎么影响它的相等关系，那就是把区间内的数和区间外的数标成不等。

考虑问一个区间怎么影响它的相等关系，那就是把区间内的数和区间外的数标成不等。

然后观察一下可以发现相等关系是不会交叉的，即不会出现 ABAB 这种情况：这些情况下实际上是 AAAA。

考虑问一个区间怎么影响它的相等关系，那就是把区间内的数和区间外的数标成不等。

然后观察一下可以发现相等关系是不会交叉的，即不会出现 ABAB 这种情况：这些情况下实际上是 AAAA。

所以最终序列的相等关系一定能分成一个类似于树形的结构。比如 $\{1,2,3,2,4,2,5,5,6,6,7,8,8,7,9\}$。

考虑第一层的极长首尾相同段：$\{1,2,5,6,7,9\}$。

现在所有的区间分成了两种：

- 包住了若干个极长首尾相同段
- 极长首尾相同段内部的区间

可以发现第一种就等价于一个原问题，而第二种对于各段是独立的，可以分别计算。

设 f_n 表示 n 个数的答案。枚举最后形成了多少个极长的首尾相同段，设为 i 个。

设 f_n 表示 n 个数的答案。枚举最后形成了多少个极长的首尾相同段，设为 i 个。那么要用第一类区间把这 i 段“区分”成 i 段，方案数就是 f_i。

设 f_n 表示 n 个数的答案。枚举最后形成了多少个极长的首尾相同段，设为 i 个。那么要用第一类区间把这 i 段“区分”成 i 段，方案数就是 f_i。每段内部随便怎么搞都不会影响外面的结构，且都是合法的。设 $g_{i,j}$ 表示把 i 个数划成 j 段，内部的询问选取的方案数。

设 f_n 表示 n 个数的答案。枚举最后形成了多少个极长的首尾相同段，设为 i 个。那么要用第一类区间把这 i 段"区分"成 i 段，方案数就是 f_i。每段内部随便怎么搞都不会影响外面的结构，且都是合法的。设 $g_{i,j}$ 表示把 i 个数划成 j 段，内部的询问选取的方案数。

则

$$f_n = 2^{\binom{n+1}{2}} - \sum_{i=1}^{n-1} f_i \times g_{n,i}$$

g 的话就枚举最后一段长度，乘上 $2^{\binom{k-1}{2}}$ 转移即可。$O(n^3)$。

简单题

来源不明

题意

二维平面上初始时有若干个点，接着会按时间顺序来 n 个以 $(0,0)$ 为左下角的矩形，要计算当前这个矩形围成的区域内有多少点。

范围

$n \le 5000$?

$\sum ans_i \le 10^7$?

$n \le 100000$?

$n \le 1000000$?

$O(n^2)$ 做法：
考虑求出每个矩形在它进来时的最小外接矩形。

$O(n^2)$ 做法：

考虑求出每个矩形在它进来时的最小外接矩形。

对于一个矩形 (x,y)，按时间从前往后找 $x_i \geq x$ 且 $y_i \geq y$ 的矩形，每次往小的框，框到最后那个就是。然后把每个点的最小外接矩形答案 $+1$，最后再把每个点的答案加到父亲上即可。

$O(ans)$ 做法：

拿出时间顺序第一个矩形，把点分成在里面/不在里面两部分，把其它矩形也分成它里面/它外面两部分，递归下去做就可以了。

$O(n\log^2 n)$ 做法：

还是考虑求每个矩形进来时的最小外接矩形，可以发现就是此时它的顶点往右/往上走的第一个。

线段树套 set 算一算就行了。

简单题

来源不明

$O(n \log n)$ 做法：

考虑把所有事件按照 y 坐标而不是时间进行排序，从上往下扫描，维护所有的竖线。进来一个矩形时，往左扫过去，如果碰到一条时间在它之前的竖线就停下来，否则就把那条竖线删了。

lower_bound 一下就可以得到一个点的外接矩形了（只不过时间可能不合法）。

简单题

来源不明

$O(n\log n)$ 做法：

考虑把所有事件按照 y 坐标而不是时间进行排序，从上往下扫描，维护所有的竖线。进来一个矩形时，往左扫过去，如果碰到一条时间在它之前的竖线就停下来，否则就把那条竖线删了。

lower_bound 一下就可以得到一个点的外接矩形了（只不过时间可能不合法）。

最后还要求这么个东西：一个点最深的时间 $<$ 它的祖先。相当于把时间 $>$ 它的所有点都删了之后求一个点往上爬到哪。拿个并查集算下就行了。

题意

给出平面上 n 个点，问有多少三角形和线段的二元组没有公共点。

范围

$n \leq 300$。

triangle

Bytedance Camp 2019 Day 2, Div A, Prob K

固定三角形，线段两个端点要么都在里面，要么都在外面，要么一里一外。

都在里面一定合法，一里一外一定不合法，只要计算都在外面的合法对数。

triangle

Bytedance Camp 2019 Day 2, Div A, Prob K

固定三角形，线段两个端点要么都在里面，要么都在外面，要么一里一外。

都在里面一定合法，一里一外一定不合法，只要计算都在外面的合法对数。

可以发现如果线段两个端点都在三角形外面且不合法的话，一定会穿过三角形的两条边。

固定三角形，线段两个端点要么都在里面，要么都在外面，要么一里一外。

都在里面一定合法，一里一外一定不合法，只要计算都在外面的合法对数。

可以发现如果线段两个端点都在三角形外面且不合法的话，一定会穿过三角形的两条边。设两个端点为 X, Y，夹的那个三角形顶点为 O。我们在 O 那里计算，把所有 O 出发的射线按极角排序，在 OX 到 OY 之间，ΔOXY 之外的每两个点都会产生-1 的贡献。

triangle

Bytedance Camp 2019 Day 2, Div A, Prob K

现在问题变成了求给出的点集中三个点形成的三角形内点数。

可以拆成 3 个以原点为顶点的有向三角形内点数之和。

现在问题变成了求给出的点集中三个点形成的三角形内点数。

可以拆成 3 个以原点为顶点的有向三角形内点数之和。

把原点取成某个顶点，这样能保证原点和任意点的连线上没有别的点。

还需要特判某个顶点在另两个点和原点形成的三角形内的情况。

总复杂度 $O(n^2 \log n + n^3)$。

IncreasingNumber

SRM 452 1000pts

题意

求满足下列条件的数字个数：

- 共有 n 位数码
- 没有前导 0
- 从高位到低位非降
- 模 $m = 0$

范围

$n \leq 10^{18}, m \leq 500$。

IncreasingNumber

SRM 452 1000pts

考虑每一位对数位和的影响。
本质不同的只有 m 类。
如何解决数位递增?

差分。

合法的数字一定能表示成 $\sum A_i \times 111...111$。其中 $\sum A_i \le 8$。

按照 111...111 模 m 的值进行分类，预处理出模 $m = i$ 的有 T_i 位，一类类 DP 过去即可。

$O(m^2)$。

题意

一个长为 n 的序列需要被分成若干连续的段，其中第 i 个元素所在段的长度必须 $\geq a_i, \leq b_i$。求满足限制的情况下，最多能分成几段，以及划分成尽量多段的方案数。

范围

$n \leq 10^6$.

DP，如果只考虑上界，能转移到每个点的是一个后缀，可以求出能转移到 i 的最小值为 low_i，它是单调不降的。

DP，如果只考虑上界，能转移到每个点的是一个后缀，可以求出能转移到 i 的最小值为 low_i，它是单调不降的。

把 DP 式子列出来：

$$f_i = \min_{i-j+1 \geq mx} f_j + 1$$

其中 mx 是 $j..i$ 这一段的 max。

分治。

做法 1：讨论 mx 在 $[L, Mid]$ 还是 $[Mid+1, R]$。如果在左边，则能转移到的 i 是一个前缀，如果在右边，能用来转移的 j 是一个后缀。再加上 $i-j+1 \geq mx$ 的限制以及 $j \geq low_i$ 的限制，每个 j 能更新的 i、每个 i 能用以更新的 j 都是一个区间。如果用线段树维护区间信息，复杂度是 $O(n\log^2 n)$。

做法 1：讨论 mx 在 $[L, Mid]$ 还是 $[Mid+1, R]$。如果在左边，则能转移到的 i 是一个前缀，如果在右边，能用来转移的 j 是一个后缀。再加上 $i-j+1 \geq mx$ 的限制以及 $j \geq low_i$ 的限制，每个 j 能更新的 i、每个 i 能用以更新的 j 都是一个区间。如果用线段树维护区间信息，复杂度是 $O(n\log^2 n)$。

把序列长度拓展到 2 的幂次，使用猫树，分治做完叶子时把涉及到它的猫树节点去建出来/分治要算一个叶子时把涉及到它的猫树节点对它的影响都计算掉。

做法 1：讨论 mx 在 $[L, Mid]$ 还是 $[Mid+1, R]$。如果在左边，则能转移到的 i 是一个前缀，如果在右边，能用来转移的 j 是一个后缀。再加上 $i-j+1 \geq mx$ 的限制以及 $j \geq low_i$ 的限制，每个 j 能更新的 i、每个 i 能用以更新的 j 都是一个区间。如果用线段树维护区间信息，复杂度是 $O(n \log^2 n)$。

把序列长度拓展到 2 的幂次，使用猫树，分治做完叶子时把涉及到它的猫树节点去建出来/分治要算一个叶子时把涉及到它的猫树节点对它的影响都计算掉。

这里询问的区间不会跨过中点，于是每个节点可以等到做完整个区间时再建立出来。

$O(n \log n)$。

做法 2：考虑区间最大值下标 pos，直接以 pos 为“中点”划分，根据 $pos - L$ 和 $R - pos$ 的大小关系，决定是暴力枚举左边更新右边，还是暴力枚举右边询问左边。这里的复杂度 $T(n) = T(a) + T(b) + \min(a, b)$。

做法 2：考虑区间最大值下标 pos，直接以 pos 为“中点”划分，根据 $pos-L$ 和 $R-pos$ 的大小关系，决定是暴力枚举左边更新右边，还是暴力枚举右边询问左边。这里的复杂度 $T(n)=T(a)+T(b)+\min(a,b)$。

如果直接用线段树，还是 $O(n\log^2 n)$ 的。还是类似地改成猫树就行了。

做法 2：考虑区间最大值下标 pos，直接以 pos 为“中点”划分，根据 $pos-L$ 和 $R-pos$ 的大小关系，决定是暴力枚举左边更新右边，还是暴力枚举右边询问左边。这里的复杂度 $T(n)=T(a)+T(b)+\min(a,b)$。

如果直接用线段树，还是 $O(n\log^2 n)$ 的。还是类似地改成猫树就行了。

注意这样的话修改/询问的区间会跨过中点，以 $O(1)$ 询问区间为例，猫树的右儿子需要做到哪儿建到哪儿，左儿子还是等整个区间插满之后从右往左建过来。具体细节自己脑补下吧。

$O(n\log n)$。

题意

有 N 个信号塔，第 i 个塔的位置是 i，信号强度 X_i（X_i 保证互不相同）。

有 N 个人，第 i 个人的位置是 i，一个人往左走一格要 A 秒，往右走一格要 B 秒。

这些人之间要传递信息，具体地，如果 i 有信息，那么 i 会依次做以下操作：

- 选择一个 j 满足 $1 \le j \le i$，并找到一个 k 使得 $j \le k \le i$ 并且 X_k 最大来保证通信。
- i, j 同时向 k 移动,先到的会等另一个人直到两个人都到达。
- 等到 i, j 都到达 k 时,信息的传递**瞬间完成**，并且 i, j **瞬间**回到**原来的位置**。
- 之后** i 会失去信息**，j 会获得信息。

请对每个 i 计算，如果初始 i 有信息，那么最少多少时间以后信息可以传递到 1，并输出最少时间的方案数，**方案数对 2^{32} 取模**。

一个方案可以被描述成 $P_1 = i, P_2, P_3, \ldots, P_t = 1$，表示信息的传递是 $P_1 \to P_2 \to P_3 \to \cdots \to P_t$。

两个方案被认为是不同的当且仅当 t 不同或者存在一个 $1 \le i \le t$ 使得 P_i 不同。

特殊地,对于 1，我们认为最少时间是 0，方案数为 1。

范围

$$n \le 8 \times 10^5.$$

列一下 DP 式子：

$$f_i = \min_{1\le j<i}\left\{f_j + \max_{k=mx[j..i]}(A\times(i-k), B\times(k-j))\right\}$$

列一下 DP 式子：

$$f_i = \min_{1 \le j < i} \left\{ f_j + \max_{k=mx[j..i]} (A \times (i-k), B \times (k-j)) \right\}$$

跟上一题差不多，分治，考虑这个 k 是在 j 这一侧还是 i 这一侧，再讨论 max 那一项到底是啥，每种情况能更新到/能用来更新的都是一个区间。

线段树，$O(n \log^2 n)$。

猫树，$O(n \log n)$。

题意

给一个 $n \times n$ 的矩阵 A，问 A 能否被置换矩阵以非负系数线性表出。如果不可以输出-1，否则求一组解，要求用这些矩阵表出 A 的方式唯一。

范围

$n \leq 50.$

双随机矩阵：元素非负，每行每列和都是 1。

双随机矩阵：元素非负，每行每列和都是 1。

Birkhoff-von Neumann 定理：每个双随机矩阵都是置换矩阵的凸组合。

双随机矩阵：元素非负，每行每列和都是 1。

Birkhoff-von Neumann 定理：每个双随机矩阵都是置换矩阵的凸组合。

注意到置换矩阵的凸组合一定是双随机矩阵，于是这题有解当且仅当矩阵 A 每行每列的和相等。

建立行-列二分图，行 i 和列 j 有边当且仅当 $A_{i,j} > 0$。以下用 Hall 定理证明该二分图存在一个完备匹配。考虑一个行集合 $\{x_1, ..., x_k\}$，设与它们在二分图中相邻的列集合为 $\{y_1, \ldots, y_l\}$。

建立行-列二分图，行 i 和列 j 有边当且仅当 $A_{i,j} > 0$。以下用 Hall 定理证明该二分图存在一个完备匹配。考虑一个行集合 $\{x_1, ..., x_k\}$，设与它们在二分图中相邻的列集合为 $\{y_1, \ldots, y_l\}$。

考虑 $S = \sum_{i,j} A_{x_i, y_j}$，按行算，$S = k \times C$，按列算，$S \leq l \times C$，由此知 $k \leq l$。

建立行-列二分图，行 i 和列 j 有边当且仅当 $A_{i,j} > 0$。以下用 Hall 定理证明该二分图存在一个完备匹配。考虑一个行集合 $\{x_1, ..., x_k\}$，设与它们在二分图中相邻的列集合为 $\{y_1, \ldots, y_l\}$。

考虑 $S = \sum_{i,j} A_{x_i,y_j}$，按行算，$S = k \times C$，按列算，$S \le l \times C$，由此知 $k \le l$。

对这张二分图求完美匹配，得到一个置换矩阵，调系数使减掉后仍然非负，且至少有一个元素减成 0，重复这个过程 $O(n^2)$ 次即可。

这组矩阵为什么是线性无关的？设每次选出置换矩阵为 $B_1, \ldots, B_k$，选出时会让 (x_i, y_i) 元素由非零变成零。假设存在一组系数 $\{a_1, \ldots, a_k\}$ 使得 $a_1 B_1 + \ldots + a_k B_k = 0$，考虑 (x_1, y_1)，它只在 B_1 中出现，于是 $a_1 = 0 \cdots \cdots$如此归纳可证所有系数均为 0，故 B_i 线性无关。

定理

设 $G=(V,E)$ 是一个二分图，设 $P_{\text{match}}(G)$ 是所有表示匹配的 $\chi_M \in \{0,1\}^E$ 这些向量构成的凸包，设

$$Q(G)=\left\{\mathbf{x}\in\mathbb{R}_+^E : \forall v\in V, \sum_{e\in\delta(v)} x_e \le 1\right\}$$

则 $P_{\text{match}}(G)=Q(G)$。

$\subseteq$ 比较显然，因为合法的匹配肯定满足每个点度数 ≤ 1，凸组合一下还是满足。

$\supseteq$ 用归纳法：对 $F = \mathrm{supp}(\mathbf{x}) = \{e \in E : x_e > 0\}$ 进行归纳。$|F| \leq 1$ 时显然是对的。$|F| > 1$ 时分 3 种 case：

- F 包含一个环。由于是二分图，一定是偶环。考虑取一个偶环，设 $\mathbf{d} = (\cdots + 1 \cdots - 1 \cdots + 1 \cdots - 1 \cdots)$，即让相邻的边分别加减一个小值，由于这个环上每条边都在 $(0, 1)$ 内，一定可找到 $\epsilon > 0, \mathbf{x} \pm \epsilon\mathbf{d}$ 的某条边系数变成 0，于是 $\in Q(G)$。于是可以用 $\in Q(G)$ 的东西组合出来。
- F 不含环（是森林），但不是匹配。于是考虑一条最长的路径，同样进行扰动，组合出来。
- F 是匹配，则直接把非零元素从小到大排序，差分下，这就是这些“前缀和”的凸组合。

定理

设 $G=(V,E)$ 是一个二分图，设

$$Q(G)=\left\{\mathbf{x}\in\mathbb{R}_{+}^{E}:\forall v\in V,\sum_{e\in\delta(v)}x_e=1\right\}$$

则 $P_{\text{perfect-match}}(G)=Q(G)$。

证明

$\subseteq$ 还是显然的。只证 $\supseteq$：首先由上一个定理，这个东西是在 $P_{\text{match}}(G)$ 里面，然后如果某个非完美匹配的系数 >0，凸组合后就没法使得所有点度数和都是 1 了。

Birkhoff-von Neumann 定理

每个双随机矩阵都是置换矩阵的凸组合。

证明

把行列分别看成左侧和右侧节点，则一个双随机矩阵就是一个 $\left\{\mathbf{x} \in \mathbb{R}_+^E : \forall v \in V, \sum_{e\in\delta(v)} x_e = 1\right\}$ 里的点，也就是完美匹配的凸组合，也就是排列矩阵的凸组合。

给三个 $n \times n$ 矩阵 A, B, C，问是否 $A \times B = C$.

给三个 $n \times n$ 矩阵 A, B, C，问是否 $A \times B = C$.

做法：在数集 S 中随一个 random vector $r = (r_1, ..., r_n)$，看 ABr 是否 $= Cr$。这是单面错误且错误率 $\leq \frac{1}{|S|}$。

给三个 $n \times n$ 矩阵 A, B, C，问是否 $A \times B = C$.

做法：在数集 S 中随一个 random vector $r = (r_1, ..., r_n)$，看 ABr 是否 $= Cr$。这是单面错误且错误率 $\leq \frac{1}{|S|}$。

证明：$AB - C$ 非 0，秩 ≥ 1，核空间维数 $\leq n - 1$，$(AB - C)r = 0$ 的概率最多 $\frac{|S|^{n-1}}{|S|^n}$。或者考虑 $D = AB - C$ 至少有一个非 0 元素，设为 $D_{1,1}$。则为使 $(AB - C)r = 0$，r_1 就会被限制住。

以字符串形式读入一个多项式，问它是否是零多项式。

以字符串形式读入一个多项式，问它是否是零多项式。
做法（Schwartz-Zippel Algorithm）：随机代值检验。

以字符串形式读入一个多项式，问它是否是零多项式。
做法（Schwartz-Zippel Algorithm）：随机代值检验。
多项式 $P \neq 0$，则对于数集 S 中随机的 $r_1, ..., r_n$，
$Pr[P(r_1, ..., r_n) = 0] \leq \frac{d}{|S|}$，$d$ 是 P 的次数。

以字符串形式读入一个多项式，问它是否是零多项式。

做法（Schwartz-Zippel Algorithm）：随机代值检验。

多项式 $P \neq 0$，则对于数集 S 中随机的 $r_1, ..., r_n$，$Pr[P(r_1, ..., r_n) = 0] \leq \frac{d}{|S|}$，$d$ 是 P 的次数。

证明：归纳，一元时显然成立，多元时考虑 x_1 的最高次，$P(x_1, \ldots, x_n) = M(x_2, \ldots, x_n) x_1^k + N(x_1, \ldots, x_n)$，要么 $M(x_2, ..., x_n) = 0$ 使得其关于 x_1 可能是个零多项式（概率 $\leq \frac{d-k}{|S|}$），要么不是零多项式，最多 k 个根（概率 $\leq \frac{k}{|S|}$）。

Checking associativity

大小为 n 的集合 X 上有一种运算 $\circ$，要判断这个运算是否满足结合律。

大小为 n 的集合 X 上有一种运算 $\circ$，要判断这个运算是否满足结合律。

做法：令 $\mathcal{X}=2^X$，从中随 3 个元素（表示成 01 向量）。然后元素的加法定义为：下标运算为 $\circ$ 的卷积。看这随机出来的 3 个元素是否满足结合律。显然这只会犯单面错误，下证错误率 $\geq \frac{1}{8}$。

如果有一对 (i,j,k) 不满足结合律，考虑容斥原理，其它位置任意固定，A_i,B_j,C_k 取 0/1 总共 8 个元素加加减减得出了错误结果，则必有一类中是错误结果。

另一种做法：设元素是 $g_1,...,g_n$，考虑比较 $((\sum g_ix_i)(\sum g_iy_i))(\sum g_iz_i)$ 和 $(\sum g_ix_i)((\sum g_iy_i)(\sum g_iz_i))$，其中 g 的乘法用 $\circ$。本质上就是把上面的 $(0..1,0..1,...,0..1)$ 改成了 $(x_1,x_2,...,x_n)$。

另一种做法：设元素是 $g_1, ..., g_n$，考虑比较 $((\sum g_i x_i)(\sum g_i y_i))(\sum g_i z_i)$ 和 $(\sum g_i x_i)((\sum g_i y_i)(\sum g_i z_i))$，其中 g 的乘法用 $\circ$。本质上就是把上面的 $(0..1, 0..1, ..., 0..1)$ 改成了 $(x_1, x_2, ..., x_n)$。

这样算出来每个元素前的系数都是个 3 次多项式（变量是 x, y, z）。满足结合律当且仅当所有元素前的系数都是零多项式。

由 Schwartz-Zippel 引理，每个 x_i, y_i, z_i 都在模 p 域下随，正确率就能达到 $1 - \frac{3}{p}$ 了。

设 G 是个 n 个元素的集合，有个 G 上的运算 $\circ$，想知道它满不满足交换律。

设 G 是个 n 个元素的集合，有个 G 上的运算 $\circ$，想知道它满不满足交换律。

做法：考虑 $h = g_1^{b_1} \circ g_2^{b_2} \circ \cdots \circ g_k^{b_k}$，每个 b_i 随机取 $\{0, 1\}$。以这个方式随 h, h'，输出是否 $h \circ h' = h' \circ h$。

Suppose H is any *proper* subgroup of G. For a random product h, show that

$$\Pr[h \notin H] \geq \frac{1}{2}.$$

[Hint: Construct $h' \notin H$ from any $h \in H$.]

Answer. Consider the first element $g_i \notin H$, denote $g_1^{b_1} \circ \ldots \circ g_{i-1}^{b_{i-1}}$ by u and $g_{i+1}^{b_{i+1}} \circ \ldots \circ g_n^{b_n}$ by v. We set $b_i = 1$ iff $v \in H$, then the result $u \circ g_i^{b_i} \circ v$ must be $\notin H$, since $a \in H, b \notin H \Rightarrow a \circ b \notin H, b \circ a \notin H$. (otherwise $b = a^{-1} \circ (a \circ b) \in H$ or $b = (b \circ a) \circ a^{-1} \in H$). So we have $\Pr[h \notin H] \geq \frac{1}{2}$. $\triangleleft$

Show that when G is not abelian, the algorithm reports a correct answer with probability at least $1/4$.

[Hint: Consider the elements of G commute with all elements of G.]

Answer. When G is commutative, the algorithm will always output correct answers. When it is not, consider the subset $H_1 = \{g_k | \forall g_i \in G, g_i \circ g_k = g_k \circ g_i\}$,

- $\forall a, b \in H_1, \forall g_i \in G, g_i \circ (a \circ b) = (g_i \circ a) \circ b = (a \circ g_i) \circ b = a \circ (g_i \circ b) = a \circ (b \circ g_i) = (a \circ b) \circ g_i$, so we have $ab \in H_1$.
- $\forall a \in H_1, \forall g_i \in G, a \circ g_i = g_i \circ a$. Left multiply and right multiply both sides by a^{-1} we obtain $a^{-1} \circ g_i = g_i \circ a^{-1}$, so $a^{-1} \in H_1$.

The above two properties guarantee that H_1 is a subgroup of G, and it is actually a proper subgroup since G is not commutative. With probability $\geq \frac{1}{2}$, a random product h is not in H_1, then consider $H_2 = \{g_k | \ g_k \circ h = h \circ g_k\}$, similarly H_2 is a proper subgroup of G, too. So another random product h' has probability $\geq \frac{1}{2}$ to be not in H_2, i.e. with probaiblity at least $\frac{1}{4}$, we will find a witness (h, h') if G is not commutative. $\triangleleft$

要检验两个 n 位的数是否相同。

要检验两个 n 位的数是否相同。

做法：在 $\{2, ..., T\}$ 内随机一个质数 p，看是否 $a \equiv b \pmod p$。

错误当且仅当 $p|a-b$。一个 n 位数的不同素因子个数是 $O(\pi(n))$ 的（即对于数 $n, \omega(n) = O\left(\frac{\log n}{\log\log n}\right)$）。于是错误概率是 $\frac{\pi(n)}{\pi(T)}$。

要检验两个 n 位的数是否相同。

做法：在 $\{2, \ldots, T\}$ 内随机一个质数 p，看是否 $a \equiv b \pmod{p}$。

错误当且仅当 $p | a - b$。一个 n 位数的不同素因子个数是 $O(\pi(n))$ 的（即对于数 $n, \omega(n) = O\left(\frac{\log n}{\log\log n}\right)$）。于是错误概率是 $\frac{\pi(n)}{\pi(T)}$。

Prime Number Theorem：

$$\frac{x}{\ln x} \leq \pi(x) \leq 1.26 \frac{x}{\ln x} \quad \forall x \geq 17$$

于是选 $T = cn$ 就能得出错误概率 $\leq \frac{1.26}{c}\left(1 + \frac{\ln c}{\ln n}\right)$。

问一个小串是否在大串中出现过。

问一个小串是否在大串中出现过。

设 $X(j)$ 表示大串从 j 开始的哈希值，Y 表示小串的哈希值，错误当且仅当 p 整除某个 $X(j)-Y$，即 p 整除 $\prod(X(j)-Y)$，这是个 mn 位的数，于是错误概率不超过 $\frac{\pi(mn)}{\pi(T)}$。取 $T=cmn$ 即可。

给一个数，判它是否是质数。
费马小定理：如果 p 是质数，则

$$\forall a \in \{1, ..., p-1\}, a^{p-1} \equiv 1 \pmod p$$

Carmichael numbers：对所有跟它互质的 a 都能通过测试，如 $561 = 3 \times 11 \times 17$。

直接搞只能指望随到跟它不互质的 a，而 $\phi(n)$ 可以很接近于 n，不现实。

如果 n 是合数并且不是 Carmichael number，则错误率 $\leq \frac{1}{2}$。

如果 n 是合数并且不是 Carmichael number，则错误率 $\leq \frac{1}{2}$。

考虑 $S_n = \left\{a \in \mathbb{Z}_n^* : a^{n-1} = 1 \bmod n\right\}$ 是模 n 缩系的一个子群，如果不是 Carmichael number，那就是真子群，由拉格朗日定理，子群的 size 是它的约数，即 $|S_n| \leq \frac{1}{2}\left|\mathbb{Z}_n^*\right|$。

如果 p 是素数，则 1 有其仅有两个平方根 ± 1。

如果 p 是素数，则 1 有其仅有两个平方根 ± 1。
Miller-Rabin 算法就是假如 $p-1=2^R\cdot Q$，就去求 $a^Q, a^{2Q}, \ldots, a^{2^R Q}=a^{n-1}$，看有没有 1 的 non-trivial root。

设 s^* 是形如 $2^i \cdot Q$ 的最大的使存在 $a, a^{s^*} \equiv -1 \pmod{p}$ 的数。

设 s^* 是形如 $2^i \cdot Q$ 的最大的使存在 $a, a^{s^*} \equiv -1 \pmod p$ 的数。

考虑集合 $\{a | a^{s^*} \equiv \pm 1 \pmod p\}$。显然是个群，且所有 non-witness 都在里面。下证是个真子群。

设 s^* 是形如 $2^i \cdot Q$ 的最大的使存在 $a, a^{s^*} \equiv -1 \pmod p$ 的数。

考虑集合 $\{a | a^{s^*} \equiv \pm 1 \pmod p\}$。显然是个群，且所有 non-witness 都在里面。下证是个真子群。

考虑分解 p 为 $n_1 n_2$（它们互质），希望构造 $a^{s^*} \equiv 1 \pmod{n_1}, a^{s^*} \equiv -1 \pmod{n_2}$，这样 a^{s^*} 就不可能 $\equiv \pm 1 \pmod p$。

设 s^* 是形如 $2^i \cdot Q$ 的最大的使存在 $a, a^{s^*} \equiv -1 \pmod p$ 的数。

考虑集合 $\{a | a^{s^*} \equiv \pm 1 \pmod p\}$。显然是个群，且所有 non-witness 都在里面。下证是个真子群。

考虑分解 p 为 $n_1 n_2$（它们互质），希望构造 $a^{s^*} \equiv 1 \pmod{n_1}, a^{s^*} \equiv -1 \pmod{n_2}$，这样 a^{s^*} 就不可能 $\equiv \pm 1 \pmod p$。

注意到 $\exists b, b^{s^*} \equiv -1 \pmod p$，于是构造 $a \equiv 1 \pmod{n_1}, a \equiv b \pmod{n_2}$。CRT 起来就完成了证明。

结合拉格朗日定理，知 non-witness 不超过 $\frac{1}{2}$。

这么分析的道理：要用一个真子群把所有 non-witness 包起来。设 s^* 是最后一个存在 -1 的位置（$\exists a, a^{s^*} \equiv 1 \pmod{n}$）。

这么分析的道理：要用一个真子群把所有 non-witness 包起来。设 s^* 是最后一个存在 -1 的位置（$\exists a, a^{s^*} \equiv 1 \pmod{n}$）。

这个真子群必须把所有 non-witness 包住，于是不能取得太小；同时又必须确实是个真子群，不能取得太大（如 $\{a|a^{2s^*} \equiv 1 \pmod{n}\}$，当 $2s^* = n - 1$ 时，Carmichael number 就怎么都 check 不出来）。

最后取成 $\{a|a^{s^*} \equiv \pm 1 \pmod{n}\}$。

一句话概括：用随机期望 $\geq 1(< 1)$ 来证明至少有一种方案 $\geq 1(< 1)$。

一句话概括：用随机期望 $\geq 1(<1)$ 来证明至少有一种方案 $\geq 1(<1)$。

另一个角度理解：考虑一个方案中所有的元素，它坏的概率是 p，union bound 得到至少有一个坏的概率 <1，则存在方案没有坏的元素。

Ramsey number $R_{n,m}$ 是说至少多少个点的图，能保证边二染色时要么有一个 n 个点的红色团，要么有一个 m 个点的蓝色团。$n = m$ 时简写成 R_n。

Ramsey number $R_{n,m}$ 是说至少多少个点的图，能保证边二染色时要么有一个 n 个点的红色团，要么有一个 m 个点的蓝色团。$n = m$ 时简写成 R_n。

$R_3 = 6.$

$$R_k > 2^{k/2}.$$

$R_k > 2^{k/2}$.

证明：随机给边染色，一个团是同色的概率是 $2^{1-\binom{k}{2}}$。union bound 得到存在一个团是同色的概率 $= \binom{n}{k} 2^{1-\binom{k}{2}}$，代入 $n = 2^{k/2}$，在 $k \geq 3$ 时概率 < 1。于是存在方案没有一个团是同色的。或者理解为同色团数的期望 < 1，则必有方案是 $= 0$ 的。

存在一个割，至少 $\frac{|E|}{2}$ 条边。

存在一个割，至少 $\frac{|E|}{2}$ 条边。

证明：随机取 subset S 和 $V \backslash S$，一条边作为割边概率为 $\frac{1}{2}$，期望 $\frac{|E|}{2}$ 条割边，于是一定有方案至少 $\frac{|E|}{2}$ 条割边。

存在独立集大小 $\geq \sum_{v\in V} \frac{1}{\deg(v)+1}$。

存在独立集大小 $\geq \sum_{v\in V}\frac{1}{\deg(v)+1}$。

证明：随一个排列，把局部最小值拿出来，构成一个独立集。期望 size 就是上面这个，于是一定有方案 $\geq$ 这个。

定义 $c(G)$ 是把 G 画在平面上，边交叉的最少数量。
对于平面图有 $m \leq 3n - 6$。

定义 $c(G)$ 是把 G 画在平面上，边交叉的最少数量。

对于平面图有 $m \leq 3n - 6$。

考虑在最优画法中把交点都新建成点，这样加了 c 个点 $2c$ 条边变成了平面图，于是

$m' \leq 3n' - 6 \Rightarrow m + 2c \leq 3n + 3c - 6 \Rightarrow c \geq m - 3n + 6$。

现在用 probabilistic method 试图得到一个不一样的界。

对 $m \geq 4n$ 的图有 $c(G) \geq \frac{m^3}{64n^2}$。

对 $m \geq 4n$ 的图有 $c(G) \geq \frac{m^3}{64n^2}$。

证明：考虑从一个最优解出发，取随机导出子图 G_p，即每个点以 p 概率留着，$1-p$ 概率删了。

显然不管怎么删点，始终有 $c(\ldots) \geq c(G_p) \geq m_p - 3n_p$。

对 $m \geq 4n$ 的图有 $c(G) \geq \frac{m^3}{64n^2}$。

证明：考虑从一个最优解出发，取随机导出子图 G_p，即每个点以 p 概率留着，$1-p$ 概率删了。

显然不管怎么删点，始终有 $c(\ldots) \geq c(G_p) \geq m_p - 3n_p$。

那么两边对所有方案取期望，$c(G)p^4 \geq mp^2 - 3np, \forall p$。

取一下 p 即可。可以证到 $c(G) \geq \frac{4m^3}{243n^2}$（此时要保证 $p = \frac{9n}{2m} \leq 1$ 即 $m \geq 4.5n$）。

$n \times n$ 网格，每行每列有个开关可以切换整行/整列状态，求最大能点亮多少盏灯。

$n \times n$ 网格，每行每列有个开关可以切换整行/整列状态，求最大能点亮多少盏灯。

考虑行随机，列贪心，这样每列都相当于一坨 i.i.d. 的 ± 1 的和的绝对值。这个期望 $\sim \sqrt{\frac{2}{\pi}} \cdot \sqrt{n}$。

$n\times n$ 网格，每行每列有个开关可以切换整行/整列状态，求最大能点亮多少盏灯。

考虑行随机，列贪心，这样每列都相当于一坨 i.i.d. 的 ± 1 的和的绝对值。这个期望 $\sim \sqrt{\frac{2}{\pi}}\cdot\sqrt{n}$。

于是加起来期望是 $\sqrt{\frac{2}{\pi}}\cdot n^{3/2}$，即至少有一种方案能点亮 $\frac{n^2}{2}+\sqrt{\frac{1}{2\pi}}\cdot n^{3/2}$ 的灯（当然要 $n\to\infty$）。

想要把高维空间中的 n 个点映射到低维，尽可能地保持距离。

想要把高维空间中的 n 个点映射到低维，尽可能地保持距离。

Johnson-Lindenstrauss Lemma

对 $\mathbb{R}^d$ 中任意 n 个点的点集 X，$\forall\epsilon \in (0,1)$，存在映射 $\varphi : \mathbb{R}^d \to \mathbb{R}^k$，其中 $k = \left\lceil \frac{4\ln n}{\varepsilon^2/2-\varepsilon^3/3} \right\rceil \leq \lceil \frac{24}{\epsilon^2} \ln n \rceil$，使得 $\forall u, v \in X$，$(1-\varepsilon)\|u-v\|_2^2 \leq \|\varphi(u)-\varphi(v)\|_2^2 \leq (1+\varepsilon)\|u-v\|_2^2$.

想要把高维空间中的 n 个点映射到低维，尽可能地保持距离。

Johnson-Lindenstrauss Lemma

对 $\mathbb{R}^d$ 中任意 n 个点的点集 X，$\forall \epsilon \in (0,1)$，存在映射 $\varphi : \mathbb{R}^d \to \mathbb{R}^k$，其中 $k = \left\lceil \frac{4 \ln n}{\varepsilon^2/2 - \varepsilon^3/3} \right\rceil \leq \lceil \frac{24}{\epsilon^2} \ln n \rceil$，使得 $\forall u, v \in X$，$(1-\varepsilon)\|u - v\|_2^2 \leq \|\varphi(u) - \varphi(v)\|_2^2 \leq (1+\varepsilon)\|u - v\|_2^2$.

证明这里略。实际做的时候，这个映射可以就是乘一个每个元素是 i.i.d. $N(0,1)$ 的矩阵。