IOI2026 中国国家队选拔 第二试《三色花园》解题报告

IOI2026 中国国家集训队工作组

2025 年 2 月 11 日

$k=0$ 定向

- 对于 $k=0$ 的情况，删去任意一种颜色的边得到的图都是一个 DAG。
- 令 G_c 为删去所有颜色为 c 的边得到的图，p_c 为 G_c 的拓扑序，rk_c 为拓扑序的排名序列。
- 对于原来的竞赛图中的边 $x \xrightarrow{c} y$，在 $p_{(c+1) \bmod 3}, p_{(c+2) \bmod 3}$ 两个排列中 x 均在 y 前面，因此至少存在至少两个颜色 c' 满足 $rk_{c',x} < rk_{c',y}$。
- 考虑采取如下的摩尔投票策略：令

$$F(x,y) = [rk_{1,x} < rk_{1,y}] + [rk_{2,x} < rk_{2,y}] + [rk_{3,x} < rk_{3,y}],$$

，则原图中连接 x, y 的边的方向可以唯一确定：
  - 若 $F(x,y) \geq 2$ 且 $F(y,x) \leq 1$，则方向为 $x \to y$；
  - 若 $F(x,y) \leq 1$ 且 $F(y,x) \geq 2$，则方向为 $y \to x$。

$k=0$ 染色

- 考虑如何还原一组符合要求的边的颜色。
- 仅通过 rk_1, rk_2, rk_3 进行还原实际上是可行的，采取如下策略：
- 对于还原出的不含颜色的竞赛图 G，令 $x \to y$ 是其中的一条边，那么有 $F(x,y) \geq 2$。
  - 若 $F(x,y)=2$，则将该边的颜色设置为唯一满足 $rk_{c,x} > rk_{c,y}$ 的颜色 c。
  - 若 $F(x,y)=3$，则任意设置该边的颜色均可。
- 考虑证明上述构造的正确性：若存在一个简单环 $a_1 \to \cdots \to a_l \to a_1$ 只包含颜色 $c, (c+1) \bmod 3$，记 $q = rk_{(c+2) \bmod 3}$，则有 $q_{a_1} < \cdots < q_{a_l} < q_{a_1}$，矛盾。故上述构造一定满足所有简单环均包含三种颜色。
- 传输信息量为

$$\log_2\left((n!)^3\right) = 3n\log_2 n - 3(\log_2 e)n + o(n),$$

当 $n=300$ 时需要 6123.83 个比特。

$k=1$ 定向

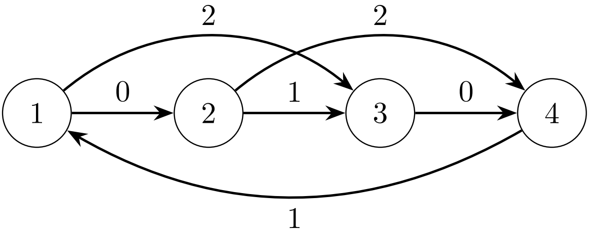
- 对于 $k=1$ 的情况，删去任意一种颜色的边，得到的图的每一个强连通分量都是一个环。
- 令 G_c 为删除所有颜色为 c 得到的图。对于图 G_c 的所有环，若反转每个环的一条边的方向，则可以得到一个 DAG。因此存在一个拓扑序满足，每个环上的点在拓扑序上构成一段区间 $[l,r]$，且反转方向的边恰为 $r\to l$。
- 考虑扩展 $k=0$ 时的做法，在传递拓扑序后额外传递每个环的信息用于修正答案，可以采用如下方式：传递一个 01 序列，其中每段取值为 1 的区间 $[l,r)$ 对应一个环的拓扑序区间 $[l,r]$。由于每个环均满足 $r-l+1\geq 3$，因此可以通过 a_c 唯一确定所有环。
- 直接传输 a_0,a_1,a_2 的信息量为

$$\log_2\left((n!)^3\right)+3n=3n\log_2 n+3(1-\log_2 e)n+o(n),$$

当 $n=300$ 时需要约 7023.83 个比特。

$k = 1$ 染色

- 考虑如何还原一组符合要求的边的颜色。如果直接沿用 $k=0$ 的做法，对于 $F(x,y)=3$ 的情况任意设置边的颜色，则可能出现如下情况：



- 上图中，若 G_0 取 $[2,4,1,3]$，G_1, G_2 均取 $[1,2,3,4]$ 作为拓扑序，则将边 $1\to3$ 与 $2\to4$ 的颜色均设置为 1 也是一组解，然而额外增加了一个只包含两种颜色的环 $1\to3\to4\to1$。

$k=1$ 染色

- 这是因为传递的拓扑序忽略了环的信息。考虑在 $F(x,y)$ 中特殊处理同一个环的情况。具体地，定义 $F(x,y,c)$ 如下：
  - 若 x,y 在 G_c 中属于不同的环，则 $F(x,y,c)=[rk_x<rk_y]$。
  - 若 x,y 在 G_c 中属于同一个环，则 $F(x,y,c)=[(x,y)\in E_c]$，其中 E_c 为 G_c 的边集。
- 令 $F(x,y)=\sum_c F(x,y,c)$，则仍然有 $F(x,y)\geq 2$ 或 $F(y,x)\geq 2$，这是因为 $x\xrightarrow{c}y$ 会使 $F(x,y,(c+1)\bmod 3)=F(x,y,(c+2)\bmod 3)=1$。
- 不妨设 $F(x,y)\geq 2$。此时再根据 $F(x,y)$ 的值设置颜色即可，即 $F(x,y)=3$ 时任选一种颜色，否则选择满足 $F(x,y,c)=0$ 的颜色 c。

$k=1$ 染色

- 考虑证明上述构造的正确性：若存在一个简单环 $a_1 \to \cdots \to a_l \to a_1$ 只包含颜色 $c, (c+1) \bmod 3$，令 $c' = (c+2) \bmod 3$，则有

$$F(a_1, a_2, c') = \cdots = F(a_{l-1}, a_l, c') = F(a_l, a_1, c') = 1.$$

- 如果存在 $1 \le x < y \le l$ 使得 a_x 与 a_y 不属于 $E_{c'}$ 中的同一个环，则由拓扑序可以得到矛盾，因此 $a_1, \ldots, a_l$ 一定属于同一个环。此时有

$$[(a_1, a_2) \in E_{c'}] = \cdots = [(a_{l-1}, a_l) \in E_{c'}] = [(a_l, a_1) \in E_{c'}] = 1,$$

即在原图中这个环同样只包含 c 与 $(c+1) \bmod 3$ 两种颜色。故上述构造正确。

$k=1$ 压缩信息

- 上述解法使用的比特数量超出了题目的限制，核心问题在于使用 $3n$ 个比特传递所有 a_0,a_1,a_2 过于浪费。
- 考虑进一步压缩以上信息，有两种方式，一种是直接构造，另一种则是使用动态规划计算合法状态数量。

$k = 1$ 压缩信息：构造

- 考虑进一步观察题目结构，可以发现如下性质：由于每个点仅被至多一个只有不超过两种颜色的简单环经过，G_0, G_1, G_2 中经过一个点的简单环的个数之和不超过 2，其中 G_0, G_1, G_2 中经过一个点的简单环的个数之和为 2 当且仅当这个点在一个同色简单环中。进一步地，不难证明同色简单环必定是三元环。
- 对于个数之和恰好为 1 的情况，只需要记录这个点被哪个图中的环经过，一共四种情况，需要两个比特。
- 对于个数之和恰好为 2 的情况，考虑只记录经过一个点的其中一个环，然后再根据这两个环是同样的三元环的特性，利用拓扑序中这个三元环的顺序来标记与识别。

$k=1$ 压缩信息：构造

- 具体地，对于每一个三元环，
  - 若仅包含一种颜色 c，则要求其在 $G_{(c+1) \bmod 3}$ 与 $G_{(c+2) \bmod 3}$ 的拓扑序中顺序相同；
  - 若恰好包含两种颜色 $c, (c+1) \bmod 3$，则要求其在 $G_{(c+2) \bmod 3}$ 中的拓扑序不能以连续子串出现在 G_c 或 $G_{(c+1) \bmod 3}$ 的拓扑序中，由抽屉原理知一定存在至少一种顺序满足要求。
- 此时对于每一个点，若其在某个简单环中出现，只需枚举相邻两个在同一个图的简单环中出现的点，然后判断是否存在另一个图中这三个点的拓扑序也连续出现即可分辨。
- 由于每个点需要传输两个比特，总信息量为

$$\log_2\left((n!)^3\right) + 2n = 3n\log_2 n + (2 - 3\log_2 e)n + o(n),$$

当 $n = 300$ 时需要约 6723.83 个比特。

$k = 1$ 压缩信息：动态规划

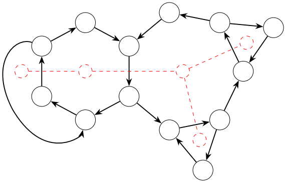
- 对于每一个环，在可以任意指定其中一条边反转方向，因此可以任意“旋转”环，即可以对拓扑序任意循环移位。
- 不妨指定每个环的拓扑序为循环移位能得到的字典序最小的序列，即将编号最小的元素移动至最前面。
- 此时对于一个大小为 l 的环，表示其的方案数能够除以 l。
- 使用动态规划计算总方案数，得到的结果以 2 为底的对数约为 $0.483n$，而每种颜色均需要记录以上结构，因此总共需要约 $1.449n$ 个额外比特。
- 总信息量约为

$$\log_2\left((n!)^3\right) + 1.449n = 3n\log_2 n + (1.449 - 3\log_2 e)n + o(n),$$

  当 $n = 300$ 时需要约 6556.20 个比特。
- 命题时将限制设置为了 $\left\lceil \log_2\left((n!)^3\right) + cn \right\rceil$ $(c \in \mathbb{N})$ 的形式，因此两种压缩信息的方式均可通过。

$k = 2$

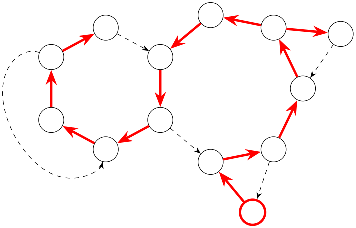
- 对于 $k = 2$ 的情况，图的结构更加复杂。
- 考虑先分析 G_1, G_2, G_3 的每一个强连通分量，即每个点仅被不超过 2 个简单环经过的强连通图。可以发现其满足以下性质：
  - 将每条有向边视为无向边后是一个平面图，其中每个环对应一个面；
  - 将有公共点的环对应的面连边后，得到的对偶图是一棵树，且每个环的边的方向由对应点在树上深度的奇偶性决定。
- 一种可能的结构如下图所示：



- 具体证明可以在耳分解结构上使用归纳法。每一个耳分解的添加相当于恰好在这个原来的树上长出一个叶子面，并且旋转的方向与父亲刚好相反。

$k = 2$

- 进一步地，这个图一定存在一个点满足其是这个树形图的根，即到其他所有点恰好有一条简单路径。可以证明，对偶图的任意一个叶子面上的终止点即是树形图的根。如下图所示，红色边构成一棵 DFS 树，且所有黑色边均为返祖边。



- 因此，原图中的每个简单环恰好对应一条返祖边，则原限制转化为该图上的每一个点至多被两条返祖边路径覆盖，即将相交的返祖边连边后得到的图是二分图。且由于图强连通，每一个结点至多有两个儿子。

$k=2$

- 考虑当树确定时如何传递返祖边。将返祖边二染色后，对于其中一种颜色，其构成结构类似栈匹配。具体地，存在一种树链剖分方案，使得剖分出来的每一条链均满足所有返祖边无交。因此只需要记录一个点是否成为了返祖边的端点，还原时根据奇偶性匹配即可得到所有返祖边。
- 由于有两种颜色，因此需要传递两种树链剖分的方案。但可以发现，第二种颜色的剖分方案恰好为第一种颜色的剖分方案的反序（即剖分时均选择另一条边）。因此只需要传递其中一种颜色的 DFS 序，便可还原出两种颜色的剖分方案。
- 还原一组符合要求的边的颜色直接沿用 $k=1$ 的做法即可。

$k=2$

- 树的形态可以使用 $0,1,2$ 进行传递，额外的信息量为

$$3(\log_2 3+2)n = 10.75n.$$

- 实际上有效的状态并不是很多。经过粗略简枝后，额外的信息量为

$$(\log_2 972)n = 9.92n.$$

- 总信息量为

$$\log_2\left((n!)^3\right) + (\log_2 972)n = 3n\log_2 n + (\log_2 972 - 3\log_2 e)n + o(n),$$

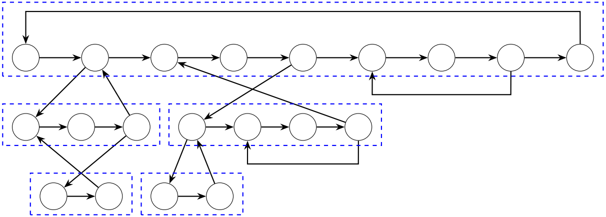
当 $n=300$ 时需要约 9101.28 个比特。

$k=2$ 压缩信息

- 与 $k=1$ 的情况相同，$k=2$ 时仍需要压缩信息，同样存在构造与动态规划两种方式。

$k=2$ 压缩信息：构造

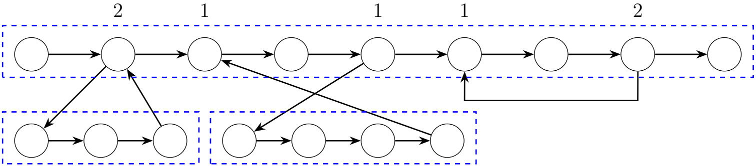
- 考虑直接传递对偶图的结构。如下图所示，先选取一个环作为树根，然后将每个耳分解添加的链作为一个儿子结点加入：



- 整棵树的结构由三部分组成：树的 DFS 序，每个结点对应的 DFS 序区间，以及每个结点与父亲结点的相交区间。其中 DFS 序即拓扑序，区间划分可以使用 $k=1$ 的做法，用 01 序列标记每个环的位置。

$k=2$ 压缩信息：构造

- 接下来记录每个结点与父亲结点的相交区间。考虑在父亲结点上另外记录一个标记序列：
  - 若相交区间长度恰好为 1，则在该区间处标记 2；
  - 若相交区间长度大于 1，则在区间左右端点处均标记 1；
  - 若添加的链仅包含一条边，则该结点直接被父亲结点吸收，此时在区间左端点处标记 1，右端点处标记 2。



$k=2$ 压缩信息：构造

- 进一步地，类比 $k=1$ 时的压缩方式，可以发现每个结点只会在至多一个图中被两个简单环覆盖，且恰为标记序列中所有标记位置。
- 由此可以将三个图的标记序列压缩为一个，只需记录七种情况即可。
- 总信息量为

$$\log_2\left((n!)^3\right)+(3+\log_2 7)n=3n\log_2 n+(5.391-3\log_2 e)n+o(n),$$

当 $n=300$ 时需要约 7866.04 个比特。

$k=2$ 压缩信息：动态规划

- 上述结构只具有必要性，不具有充分性。可以猜测原图具有更强的性质，具体为：
  - 存在一个常数 d 使得对于任意一个树形图，存在一个根满足每一个结点要么只有 ≤ 1 个儿子，要么有两个儿子且其中一个儿子的子树大小 $\leq d$。
- 目前暂未找到简单的证明方法，但可以通过搜索证明 $d \geq 3$ 时的正确性。
- 因此可以通过动态规划求出符合条件的树和符合条件的匹配方案，再计算强连通分量的划分方案。
- 在动态规划得到的结果中，每个图需要约 $1.797n$ 个比特传递信息。因此总信息量约为

$$\log_2\left((n!)^3\right) + 5.391n = 3n\log_2 n + (5.391 - 3\log_2 e)n + o(n),$$

当 $n = 300$ 时需要约 7741.13 个比特。

$k=2$ 压缩信息：动态规划

- 令 $F(d)$ 为子树大小限制为 d 的情况下最坏每一个结点所需的平均比特数。实际计算结果表明，随着 d 的增加，$F(d)$ 始终向着 2 逼近但是无法超过 2，其中 $F(300)$ 的结果为 1.971，只需要约 7897.73 个比特。因此任取 d 均是可行的，只是可能无法在规定时间限制内得到答案。