

# 1 One, Two, Three

## 1.1 题目来源

44th Petrozavodsk Programming Camp (2023 Winter), Day 2. GP of ainta, Problem C<sup>1</sup>.

## 1.2 题目大意

给定一个长度为  $N$  的数列  $A_0, A_1, \dots, A_{N-1}$ , 保证  $A_i \in \{1, 2, 3\}$ 。定义有序三元组  $(i, j, k)$  是好的当且仅当  $i < j < k$  且其满足以下两个条件之一:

- $A_i = 1, A_j = 2, A_k = 3$ ;
- $A_i = 3, A_j = 2, A_k = 1$ 。

请选出若干个好的有序三元组  $(a_0, b_0, c_0), (a_1, b_1, c_1), \dots$ , 需要保证每个元素最多只出现在一个三元组中 (即  $a_i, b_i, c_i$  两两不同), 请求出选出的有序三元组的最大数量并输出方案。

## 1.3 数据范围

保证  $1 \leq N \leq 6 \times 10^5$ 。

## 1.4 解题方法一

称满足  $A_i = 1, A_j = 2, A_k = 3$  的三元组  $(i, j, k)$  为第一类三元组, 满足  $A_i = 3, A_j = 2, A_k = 1$  的三元组  $(i, j, k)$  为第二类三元组, 记两类三元组的个数分别为  $X, Y$ 。下文中讨论合法的  $X, Y$  (即存在一种选出  $X$  个第一类三元组与  $Y$  个第二类三元组的方案) 需要满足的条件。

记  $x_i, y_i, z_i$  分别表示数列  $A$  在下标范围内  $[0, i)$  之内 1, 2, 3 的个数, 显然需要满足  $X + Y \leq \min(\{x_N, y_N, z_N\})$ 。

考虑数列  $A$  中的所有 1, 它们可以作为第一类三元组的  $A_i$  出现, 也可以作为第二类三元组的  $A_k$  出现; 数列  $A$  中的所有 3, 它们可以作为第一类三元组的  $A_k$  出现, 也可以作为第二类三元组的  $A_i$  出现。

**定理 1.4.1.** 对于任意  $X, Y$ , 若存在一种选出  $X$  个第一类三元组与  $Y$  个第二类三元组的方案, 则存在一种方案使得第一类三元组的  $A_i$  为  $A$  中的前  $X$  个 1,  $A_k$  为  $A$  中的后  $X$  个 3, 第二类三元组的  $A_i$  为  $A$  中的前  $Y$  个 3,  $A_k$  为  $A$  中的后  $Y$  个 1。

证明. 假设  $A$  中的所有 1 的下标从小到大排序后依次为  $p_0, p_1, \dots, p_{x_N-1}$ , 所有 3 的下标从小到大排序后依次为  $r_0, r_1, \dots, r_{z_N-1}$ 。

<sup>1</sup>QOJ 链接: [QOJ 6537](#)

对于一种方案,记  $X$  个第一类三元组依次为  $(a_0, b_0, c_0), (a_1, b_1, c_1), \dots, (a_{X-1}, b_{X-1}, c_{X-1})$ ,  $Y$  个第二类三元组依次为  $(d_0, e_0, f_0), (d_1, e_1, f_1), \dots, (d_{Y-1}, e_{Y-1}, f_{Y-1})$ , 则将第  $i$  个第一类三元组  $(a_i, b_i, c_i)$  调整为  $(p_{rank a_i}, b_i, r_{z_N - X + rank c_i})$ , 将第  $i$  个第二类三元组  $(d_i, e_i, f_i)$  调整为  $(r_{rank d_i}, e_i, p_{x_N - Y + rank f_i})$ 。其中  $rank a_i$  表示  $a$  中小于  $a_i$  的元素个数,  $rank c_i, rank d_i, rank f_i$  类似分别表示  $c, d, f$  中小于  $c_i, d_i, f_i$  的元素个数。

由于  $X + Y \leq \min(\{x_N, y_N, z_N\})$ , 所以在调整后仍不存在一个元素被使用两次。且在调整后,  $b$  和  $e$  没有发生变化而  $a$  和  $d$  每一项均不增加,  $c$  和  $f$  每一项均不减少。因此调整后题目所述性质仍成立, 且在调整后符合定理内容, 原命题得证。

在下文中, 我们记  $p_0, p_1, \dots, p_{X-1}$  表示  $A$  中前  $X$  个 1 的下标从小到大排序排序后的结果,  $q_0, q_1, \dots, q_{X-1}$  表示  $A$  中后  $X$  个 3 的下标从小到大排序排序后的结果,  $r_0, r_1, \dots, r_{Y-1}$  表示  $A$  中前  $Y$  个 3 的下标从小到大排序排序后的结果,  $s_0, s_1, \dots, s_{Y-1}$  表示  $A$  中后  $Y$  个 1 的下标从小到大排序排序后的结果,

**定理 1.4.2.** 对于任意  $X, Y$ , 若存在一种选出  $X$  个第一类三元组与  $Y$  个第二类三元组的方案, 一定存在一种方案使得对于任意  $i \in [0, X)$ ,  $p_i$  与  $q_i$  在同一个第一类三元组中; 且对于任意  $i \in [0, Y)$ ,  $r_i$  与  $s_i$  在同一个第二类三元组中。

证明. 考虑当存在第一类三元组  $(p_i, a, q_j)$  与  $(p_k, b, q_l)$  且  $i < k, j > l$  时, 将  $p_i$  改为与  $q_l$  配对, 将  $p_k$  改为与  $q_j$  配对:

- 当  $a < q_l$  时, 将三元组调整为  $(p_i, a, q_l)$  以及  $(p_k, b, q_j)$ ;
- 当  $a > q_l$  时, 将三元组调整为  $(p_i, b, q_l)$  以及  $(p_k, a, q_j)$ 。

容易证明在调整过后方案仍符合条件。

当  $p_i$  与  $q_j$  配对时, 记  $r_i = j$ , 则上述  $(i, j)$  和  $(k, l)$  的对数就等价于  $r$  的逆序对数量。而一次交换等价于交换一对逆序对, 因此  $r$  的逆序对数量一定会减少。因此, 在有限次调整内, 必然可以使得  $r$  的逆序对数变为 0。此时, 对于任意  $i \in [0, X)$ ,  $p_i$  与  $q_i$  在同一个第一类三元组中。

同样的, 可以证明存在一种方案使得对于任意  $i \in [0, Y)$ ,  $r_i$  与  $s_i$  在同一个第二类三元组中。

此时, 我们已经得到了所有三元组的始末元素, 只有中间的元素需要确定。由于仍需确定的元素只有一种, 且约束条件仅有位置之间的相对顺序。因此考虑一张二分图, 左部点对应  $X + Y$  个已经确定的区间 (对于第一类三元组, 对应区间为  $[p_i, q_i]$ ; 对于第二类三元组, 对应区间为  $[r_i, s_i]$ ), 右部点对应  $y_N$  个元素 2 的下标, 两个点之间有连边当且仅当右部点对应下标在左部点对应的区间内。那么有解当且仅当存在一个包含了全部左部点的匹配。

**引理 1.4.3** (Hall 定理). 对于一张有着  $N$  个左部点和  $M$  个右部点的二分图  $G$ , 对左部点集合  $S$ , 记  $N(S)$  表示右部点中所有与至少一个  $S$  相连的点构成的集合, 那么图  $G$  存在一个所有左部点的匹配当且仅当对于任意左部点集合  $S$ , 均有  $|N(S)| \geq |S|$ 。

**引理 1.4.4** (Hall 定理的等价形式). 对于一张有着  $N$  个左部点和  $M$  个右部点的二分图  $G$ , 对右部点集合  $T$ , 记  $f(T)$  表示满足相连节点是  $T$  的子集的左部点的数量, 那么图  $G$  存在一个所有左部点的匹配当且仅当对于任意右部点集合  $T$ , 均有  $f(T) \leq |T|$ 。

因此, 我们只需要判断是否存在一个右部点集合  $T$  使得  $f(T) > |T|$  即可。事实上, 我们只需要考虑一部分  $T$ 。

为了方便描述, 对于一个右部点集合  $T$ , 定义  $\min(T)$  表示  $T$  中每个节点对应下标的最小值,  $\max(T)$  表示  $T$  中每个节点对应下标的最大值。

**引理 1.4.5.** 对于上文中提到的二分图以及非空右部点集合  $T_1, T_2$ , 若存在对应下标为  $x$  的右部点  $x$  使得  $\max(T_1) < x < \min(T_2)$ , 则  $f(T_1 \cup T_2) = f(T_1) + f(T_2)$ 。

上述结论的证明仅需对于  $S$  中的每一个  $x$  分别证明  $[N(\{x\}) \subseteq (T_1 \cup T_2)] = [N(\{x\}) \subseteq T_1] + [N(\{x\}) \subseteq T_2]$  即可, 在此略去。

定义一个非空右部点集合  $T$  是**连续的**当且仅当不存在  $x$  使得  $x$  为某一右部点对应下标且下标  $x$  对应右部点不在  $T$  中出现, 并且满足  $\min(T) < x < \max(T)$ 。

**定理 1.4.6.** 对于上文中提到的二分图, 若存在右部点集合  $T$  满足  $f(T) > |T|$ , 则必然存在一个**连续的**集合  $T$  满足  $f(T) > |T|$ 。

**证明.** 任取一极小的满足  $f(T) > |T|$  的右部点集合  $T$ , 即不存在  $T$  的真子集  $T'$  满足  $f(T') > |T'|$ 。

若  $T$  不是**连续的**, 任取一上述定义中的  $x$ , 将  $T$  中对应下标小于  $x$  的节点划分为集合  $T_1$ , 其余节点划分为  $T_2$ 。由 1.4.5, 有  $f(T) = f(T_1) + f(T_2)$ , 由于  $T_1, T_2$  均不能作为  $T'$  满足  $f(T') > |T'|$ , 因此  $f(T_1) \leq |T_1|, f(T_2) \leq |T_2|$ , 故  $f(T) \leq |T|$ , 矛盾!

因此, 此时  $T$  必然是**连续的**。

所以, 我们只需考虑所有**连续的**集合  $T$  即可。显然, 存在这样的  $T$  满足  $f(T) > |T|$  等价于存在区间  $[L, R)$  使得下标在  $[L, R)$  内的 2 的个数小于对应区间完全包含在  $[L, R)$  内的左部点的个数。

对于前者, 下标在  $[L, R)$  内 2 的个数等于  $y_R - y_L$ 。考虑后者的计算, 其可以被分为区间  $[p_i, q_i]$  和区间  $[r_i, s_i]$  两个部分分别计算。

对于前半部分, 由于  $p, q$  的单调性, 其等于区间总数减去左端点小于  $L$  的区间数量再减去右端点大于等于  $R$  的区间数量并对 0 取  $\max$  后的结果, 即  $\max(X - x_L - (z_N - z_R), 0)$ 。

同理, 对于后半部分, 满足条件的区间共有  $\max(Y - z_L - (x_N - x_R), 0)$  个, 因此  $X, Y$  合法当且仅当对于任意区间  $[L, R)$ , 均有  $y_R - y_L \geq \max(X - x_L - (z_N - z_R), 0) + \max(Y - z_L - (x_N - x_R), 0)$ 。

将最大值运算展开, 上式可被拆为四个限制:

- $y_R - y_L \geq 0$ ;
- $y_R - y_L \geq X - x_L - (z_N - z_R)$ ;
- $y_R - y_L \geq Y - z_L - (x_N - x_R)$ ;

- $y_R - y_L \geq X - x_L - (z_N - z_R) + Y - z_L - (x_N - x_R)$ 。

显然第一个式子恒成立，而我们可以通过第二个式子求出  $X$  的上界，通过第三个式子求出  $Y$  的上界，通过第四个式子求出  $X + Y$  的上界。具体的，这三个上界均可通过移项将  $L, R$  独立后在  $\mathcal{O}(N)$  的时间复杂度内计算。

此时，我们即可求出  $X + Y$  的最大可能值并求出此时一组可能的  $X, Y$ ，再使用上文所提及的方法构造方案即可，总时间复杂度  $\mathcal{O}(N)$ 。

## 1.5 解题方法二

构造根据序列  $A$  构造长度同为  $N$  的数组  $A'$ ，对于  $i \in [0, N)$ ：

- 若  $A_i = 2$ ，则  $A'_i = 2$ ；
- 若  $A_i \neq 2$ ，记  $C_1 = \sum_{j=0}^{N-1} [A_j = 1]$ ,  $C_2 = \sum_{j=0}^i [A_j \neq 2]$ ：
  - 若  $C_2 \leq C_1$ ，则  $A'_i = 1$ ；
  - 若  $C_2 > C_1$ ，则  $A'_i = 3$ 。

上述变换，可以简单的理解为  $A'$  中值 2 的元素不变，对于其余元素从小到大排序。记对于原问题的任意一组最优解为  $(a_0, b_0, c_0), (a_1, b_1, c_1), \dots, (a_{r-1}, b_{r-1}, c_{r-1})$ 。

考虑以下问题：在  $A'$  中选出尽可能多的不交有序三元组  $(i, j, k)$ ，满足  $a_i = 1, a_j = 2, a_k = 3$ 。将该问题的任意一组最优解记为  $(a'_0, b'_0, c'_0), (a'_1, b'_1, c'_1), \dots, (a'_{r'-1}, b'_{r'-1}, c'_{r'-1})$ 。在下文中，我们证明  $r = r'$  并给出关于双向转化的构造。

**定理 1.5.1.**  $r' \geq r$ 。

证明. 只需证若原问题有解  $(a_0, b_0, c_0), (a_1, b_1, c_1), \dots, (a_{r-1}, b_{r-1}, c_{r-1})$ ，则新问题有  $r' = r$  的解。

考虑长度为  $N$  的数列  $B$ ，其由如下方式生成：

- 初始时，令  $B = A$ ；
- 依次考虑  $i = 0, 1, \dots, r - 1$ ，如果  $A_{a_i} = 3$  且  $A_{c_i} = 1$  则交换  $B_{a_i}, B_{c_i}$ 。

那么，三元组  $(a_0, b_0, c_0), (a_1, b_1, c_1), \dots, (a_{r-1}, b_{r-1}, c_{r-1})$  在数列  $B$  上即对应了若干个  $A_i = 1, A_j = 2, A_k = 3$  的有序三元组。

记所有  $B_i = 3$  且  $C_i = 1$  的下标排序后构成了数列  $p_0 < p_1 < \dots < p_{|p|}$ ，所有  $B_i = 1$  且  $C_i = 3$  的下标排序后构成了数列  $q_0 < q_1 < \dots < q_{|q|}$ ，显然有  $|p| = |q|$ 。

由于对于任意前缀  $i \in [1, n]$ ， $B$  中  $[0, i)$  中 1 的个数均小于等于  $C$  中  $[0, i)$  中 1 的个数且 2 的个数相等，因此  $[B_i = 3 \wedge C_i = 1] - [C_i = 1 \wedge B_i = 3]$  的前缀和非负。

因此，可以证明必然有  $p_i < q_i (0 \leq i < |p|)$ 。下文中，我们称  $p_i$  对应  $q_i$ ， $q_i$  对应  $p_i$ 。显然每个位置最多只对应一个数且仅被自己所对应的数对应（如果存在的话）。简记  $x$  对应  $y$  为  $f(x) = y$ 。

此时，我们即可构造如下三元组  $(a'_0, b'_0, c'_0), (a'_1, b'_1, c'_1), \dots, (a'_{r'-1}, b'_{r'-1}, c'_{r'-1})$ ，其中  $r' = r$ ，依次考虑每一个  $i \in [0, r)$ ：

- 若  $C_{a_i} = 1, C_{c_i} = 3$ ，则令  $(a'_i, b'_i, c'_i) = (a_i, b_i, c_i)$ ；

- 若  $C_{a_i} = 1, C_{c_i} = 1$ , 则令  $(a'_i, b'_i, c'_i) = (a_i, b_i, f(c_i))$ ;
- 若  $C_{a_i} = 3, C_{c_i} = 3$ , 则令  $(a'_i, b'_i, c'_i) = (f(a_i), b_i, c_i)$ 。

显然调整后的数列仍满足  $a'_i < b'_i < c'_i$  且  $A_{a'_i} = 1, A_{b'_i} = 2, A_{c'_i} = 3$ , 并且可对  $a', b', c'$  分别证明其中元素两两不同, 故原命题成立。

**定理 1.5.2.**  $r \geq r'$ 。

证明. 只需证若新问题有解  $(a'_0, b'_0, c'_0), (a'_1, b'_1, c'_1), \dots, (a'_{r-1}, b'_{r-1}, c'_{r-1})$ , 则原问题有  $r = r'$  的解。我们用如下方式生成  $r'$  个三元组  $(a_i, b_i, c_i)$ 。

将所有满足  $A_{a'_i} \neq A_{c'_i}$  的数对  $(a'_i, b'_i, c'_i)$  提取出来作为原问题的三元组, 并将剩余三元组分为  $A_{a'_i} = A_{c'_i} = 1$  以及  $A_{a'_i} = A_{c'_i} = 3$  两类。

对于两个三元组  $(a'_i, b'_i, c'_i)$  以及  $(a'_j, b'_j, c'_j)$  且  $A_{a'_i} = A_{c'_i} = 1, A_{a'_j} = A_{c'_j} = 3$ , 则有大关系  $a'_i < b'_i < c'_i < a'_j < b'_j < c'_j$ , 故可构造两个新的三元组  $(a'_i, b'_i, a'_j)$  与  $(c'_i, b'_j, c'_j)$ 。

重复上述操作至只剩一种三元组, 可以不失一般性的假设为仅剩  $A_{a'_i} = A_{c'_i} = 1$  一类, 假设共有  $c$  个, 则显然此时剩余至少  $c$  个位置  $x$  满足  $A_x = 3$ 。

任取  $c$  个位置并将它们与这些三元组一一对应, 假设三元组  $(a'_i, b'_i, c'_i)$  对应位置  $x_i$ :

- 若  $x_i < b'_i$ , 则添加有序三元组  $(x_i, b'_i, c'_i)$ ;
- 若  $x_i > b'_i$ , 则添加有序三元组  $(a'_i, b'_i, x_i)$ 。

显然生成的三元组的数量等于  $r'$  且符合题目所述条件, 原命题得证。

**定理 1.5.3.**  $r = r'$ 。

证明. 由 1.5.1 和 1.5.2 立刻得出。

此时, 我们将原问题的求解转化为了  $r'$  以及对应的  $a', b', c'$  的求解。这是信息学竞赛中的经典问题, 可以使用贪心在  $\mathcal{O}(N)$  的时间复杂度内简单求解。

上述根据  $a', b', c'$  得到  $a, b, c$  的构造也可以在  $\mathcal{O}(N)$  内简单求得。综上所述, 本题在  $\mathcal{O}(N)$  的时间复杂度内得解。

## 1.6 参考资料

- [1] Felix Gott. Lecture 30: Matching and Hall's Theorem MIT: 18.211: Combinatorial Analysis.
- [2] Sunghyeon Jo (ainta), Jeyeon Si (tlwpdus). Editorial of Grand of ainta, 44th Petrozavodsk Programming Camp (2023 Winter), Day 2.
- [3] Editorial of Playf and ABC, Hunan University 2023 the 19th Programming Contest.