联合省选模拟赛题解

$cz\_xuyixuan$

January 30, 2021

1 异或

省队参考得分：100**，子任务** $1 \sim 7$

A **队参考得分：**100**，子任务** $1 \sim 7$

记 x 为最小的满足 $2^i \geq k$ 的 i ，可以按照除去最低的 i 位后剩余的数对 a_i 进行分组，将除去最低的 i 位均相同的数放在一起考虑，组与组之间的方案数是相乘的。

在每一组中，所选的数的第 i 位必须两两不同，因此，所选的数不会超过 2 个。

分别考虑选取 $0, 1, 2$ 个的方案数，选取 2 个数的方案数可以用线段树计算。

时间复杂度 $O(N \log V)$ 。

子任务 $5, 6$ 分别提示了分组和线段树的思想。

另外，正解的做法也保证了如下做法的正确性：将数据集按大小排序，则一个子集是合法的当且仅当其相邻的两个元素异或得到的结果不小于 x ，这一点由正解的分析是显然的。对满足转化后的条件的序列 DP 计数，同样可以得到 $O(N \log V)$ 的做法。

2 计数

省队参考得分：40+**，测试点** $1 \sim 6, 10, 14$

A **队参考得分：**70+**，测试点** $1 \sim 14$

首先将价值的贡献拆开，即

$$\max\{L(a) - x, 0\} = \sum_{i=x+1}^{N} [L(a) \geq i]$$

由此，考虑对于 $i \in [1, N]$ ，计算最长相同子段的长度大于等于 i 的序列数。

补集转化，等价地考虑计算最长相同子段的长度小于 x 的序列数。

考虑一个动态规划解法，令 dp_i 表示满足最后一个元素与最后第二个元素颜色不

同，且符合条件的序列的个数。则容易得到转移：

$$dp_i = \begin{cases} k & i = 1 \\ (k-1)\sum_{j=1}^{x-1} dp_{i-j} & i \geq 2 \end{cases}$$

显然转移可以使用部分和优化，由于我们需要运行 N 次上述 dp，我们得到了一个时间复杂度为 $O(N^2)$ 的做法，可以通过测试点 $1 \sim 13$，参考实现 **interwoven - force**。

观察上述转移代码的计算过程，可以发现，对于每个 $i \in [x+1, N]$ ，我们需要计算对应 dp 数组的两个点值，以更新答案。而 dp 数组的转移方式较为单一，可以将其等价地看为如下问题的解：

有一张含有 N 个点的图，从 1 号点出发，点 j 到 $j+1$ 号点有 a 条路径，点 j 到 $j+i$ 号点有 b 条路径，求出到达 N 号点的方案数。

枚举走第二类路径的次数，上述问题的答案应为：

$$\sum_{j=1}^{\frac{N}{x}} \binom{j+N-ji-1}{j} a^{N-ji-1} b^j$$

由此，我们可以在 $O(\frac{N}{i})$ 的时间内算得所需位置的 dp 值，从而总的时间复杂度可由调和级数分析为 $O(N \log N)$ 。

3 优化

省队参考得分：35+，**子任务** $1, 2, 6$

A **队参考得分：**60+，**子任务** $1 \sim 4, 6$

首先，考虑单组询问时的解法，这是一个经典的问题，可以通过拆点得出一个费用流的建模。此后，传统的做法是用线段树优化费用流，达到**模拟费用流**的效果，详情可见该问题：`https://blog.csdn.net/qq_39972971/article/details/80818719`。但是该做法可拓展性差，无法做到快速支持多组询问，因此并不能应用于本题。

直接实现单组询问时的解法可以通过子任务 2 ，其时间复杂度为 $O(Q \times N \log N)$，配合容易取得的子任务 6 即可得到省队参考得分。

注意到费用流的建模保证了如下性质：

引理：令 $f(x)$ 表示对于特定区间，$k = x$ 时的答案，$f(x)$ 在其定义域上是凸函数。

因此，单组询问时，同样可以考虑用凸优化 DP 解决问题。二分斜率 S ，规定每选择一段，便获得 S 的收益，可以将问题转化为一个简单的线性 DP 问题，同样可以得到 $O(Q \times N \log N)$ 的解法。

注意到二分斜率 S 后，我们可以用 $O(1)$ 的信息描述一个区间，即 $info[0/1][0/1]$ ，分别表示该区间的左右两端取/不取，最优收益的值。可以考虑在线段树上维护所需的信息，询问时在定位到的区间上快速查询这些信息，将问题拓展至多组询问。

首先预处理线段树各个区间对应的 $f(x)$ ，由于 $f(x)$ 的凸性，合并两个子树信息的过程可以看做一次对凸包的闵可夫斯基和。分开处理每一个询问，则需要在线段树各个节点的 $f(x)$ 上二分，得出各个区间的信息，时间复杂度 $O(Q \log^2 N \log V)$ 。

若将二分改为整体二分，可以避免在线段树各个节点的 $f(x)$ 上二分，时间复杂度降为 $O((N+Q)\log N \log V)$ 。