Imbalance 解题报告

安徽师范大学附属中学 杨珖

2021.10.3

题目大意

定义一个 01 串是平衡的当且仅当它的 01 个数相同。

现在给定 n,k 和一个长度为 m 的 01 串 S，求有多少长度为 n 且以 S 为前缀的 01 串满足其不存在长度为 k 的平衡子串，对 998244353 取模。

数据范围

Subtask1(10pts): $n \le 20$。

Subtask2(10pts): $k \le 20$。

Subtask3(30pts): $n \le 66$。

Subtask4(20pts): $m = 0$。

Subtask5(30pts): 无特殊限制。

对于所有数据均保证 $1 \le m+1 \le k \le n \le 114$ 且 k 是偶数。

解题过程

算法一

爆搜，枚举有所有的 01 串并判断它所有长度为 k 的子串的 01 的个数是否有相等的。复杂度 $O(2^n n)$。

期望得分：10。

算法二

考虑状压，设状态 $dp_{i,j}$ 表示目前 DP 到第 i 个位置，从 i 往前 k 个位置填数的状态为 j 时，i 之前合法的填数方案数。

通过枚举 $i+1$ 填的数进行转移，复杂度 $O(2^k n)$。

期望得分 20。

算法三

直接状压看起来不太能优化了，我们考虑寻找一些性质。

不难发现：如果同时存在两段长度为 k 的子串，其中一段有大于 $\frac{k}{2}$ 个 1，而另一段只有小于 $\frac{k}{2}$ 个 1，那么一定存在一个长度为 k 的子串满足其有恰好 $\frac{k}{2}$ 个 1。

考虑证明，设这两个串分别为 $[a,a+k-1]$ 和 $[b,b+k-1]$。由于对于串 $[l+1,l+k]$ 和串 $[l,l+k-1]$，其中 1 的个数至多相差 1。所以在从 $[a,a+k-1]$ 平移到 $[b,b+k-1]$ 的过程中，一定存在某时刻恰好有 $\frac{k}{2}$ 个 1。

于是，我们可以统计满足对于所有长度为 k 的子串都有小于 $\frac{k}{2}$ 个 1 的串的数量，然后翻转 01 再统计一次。

设 s_i 表示前 i 个字符中有多少个 1。那么一个合法的串只需满足 $\forall k \le i \le n, s_i - s_{i-k} < \frac{k}{2}$。

我们换一种思路进行 DP，如果我们 k 个 s_i 一组，排列成一个矩阵的形式的话：

$$
\begin{matrix}
s_0 & s_1 & \cdots & s_{n \bmod k} & \cdots & s_{k-1} \\
s_k & s_{k+1} & \cdots & s_{k+(n \bmod k)} & \cdots & s_{2k-1} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
s_{k\lfloor \frac{n}{k} \rfloor} & s_{k\lfloor \frac{n}{k} \rfloor+1} & \cdots & s_n & &
\end{matrix}
$$

那么矩阵上下两个相邻元素的差必须小于 $\frac{k}{2}$。

当 k 较小的时候，我们可以使用算法二，而当 k 较大的时候，我们发现这个矩阵的列数是很少的。

于是我们先枚举 $s_0, s_k, s_{2k}, \ldots, s_{k\lfloor \frac{n}{k} \rfloor}$，然后我们按列进行轮廓线 DP。设 $dp_{i,j,S}$，i,j 表示扫描到了矩阵的第 i 行，第 j 列的位置，S 表示 $(i-1,j),(i-2,j),\ldots,(1,j),(n,j-1),(n-1,j-1),\ldots,(i,j-1)$ 这些位置的 s 的值的状压。并且在 S 中我们只需要记录该位置的 s 的值减去该行开头的 s 的值即可。

转移我们可以枚举 (i,j) 所填的数，并且检查其与上方元素之差是否满足条件。

最后，我们要求 $\forall l, 0 \le s_{lk} - s_{lk-1} \le 1$，因为我们枚举了开头元素，所以可以直接检查。

对于这个做法，我们的复杂度为 $O(n\left(\frac{k}{2}\right)^{\frac{2n}{k}})$。结合算法二后期望得分 50。

算法四

考虑优化算法三，我们考虑建立一个模型：我们从 $(0,0)$ 出发，对于第 i 步，若 $s_i - s_{i-1} = 1$，那么我们向右上走一步，否则我们向右走一步。

那么一个 01 串就对应了一条折线段，折线上点 (i,j) 就对应了 $s_i = j$。若要 k 个 s 一组，那么我们规定每一步都是从 (x,y) 走到 $((x+1) \bmod k, y)$ 或者 $((x+1) \bmod k, y+1)$，那么这就产生了 $\lceil \frac{n}{k} \rceil$ 个折线段，这样我们的折线就与 s 形成的矩阵所对应。

由于矩阵的上下相邻元素要满足差的限制，那么对于相邻的两条折线段上横坐标相同的两点 (x,i) 和 (x,j)，要满足 $|i-j| < \frac{k}{2}$。

我们变化一下，将矩阵第 i 行的元素全部减去 $\frac{ik}{2}$，那么我们的条件就变成了矩阵的列要递减。再考虑对应的若干折线段，满足的条件就变成了对于第 i 条和第 $i+1$ 条折线段，$i+1$ 必须完全在 i 的下方。

我们枚举 $s_0, s_k, s_{2k}, \ldots, s_{k\lfloor \frac{n}{k} \rfloor}, s_n$，第 i 条折线段就变成了从 $(0, s_{ik} - \frac{ik}{2})$ 到 $(0, s_{(i+1)k} - \frac{ik}{2})$ 的一条每次向右或右上走的路径（除了最后一条可能不完整）。并且要满足这些路径互不相交。

这就提示我们使用 LGV 引理进行不相交路径计数。起点分别为 $(0, s_0), (0, s_k - \frac{k}{2}), \ldots, (0, s_{k\lfloor \frac{n}{k} \rfloor} - \frac{k}{2}\lfloor \frac{n}{k} \rfloor)$，终点分别为 $(k, s_k), (k, s_{2k} - \frac{k}{2}), \ldots, (k, s_{k\lfloor \frac{n}{k} \rfloor} - \frac{k}{2}(\lfloor \frac{n}{k} \rfloor - 1)), (n \bmod k, s_n - \frac{k}{2}\lfloor \frac{n}{k} \rfloor)$。由于最后一个终点和其他终点并不对齐，所以可能会导致某些不合法的配对方式产生的路径也不相交，从而影响答案。例如下图 1，这是一个错位的配对，但是这两条路径不相交，会被统计入答案。因此我们将 $(n \bmod k, s_n - \frac{k}{2}\lfloor \frac{n}{k} \rfloor)$ 右下方的点全部删除，这样我们就让所有错位的配对起点终点的方式中，每一种都存在两条路径相交。

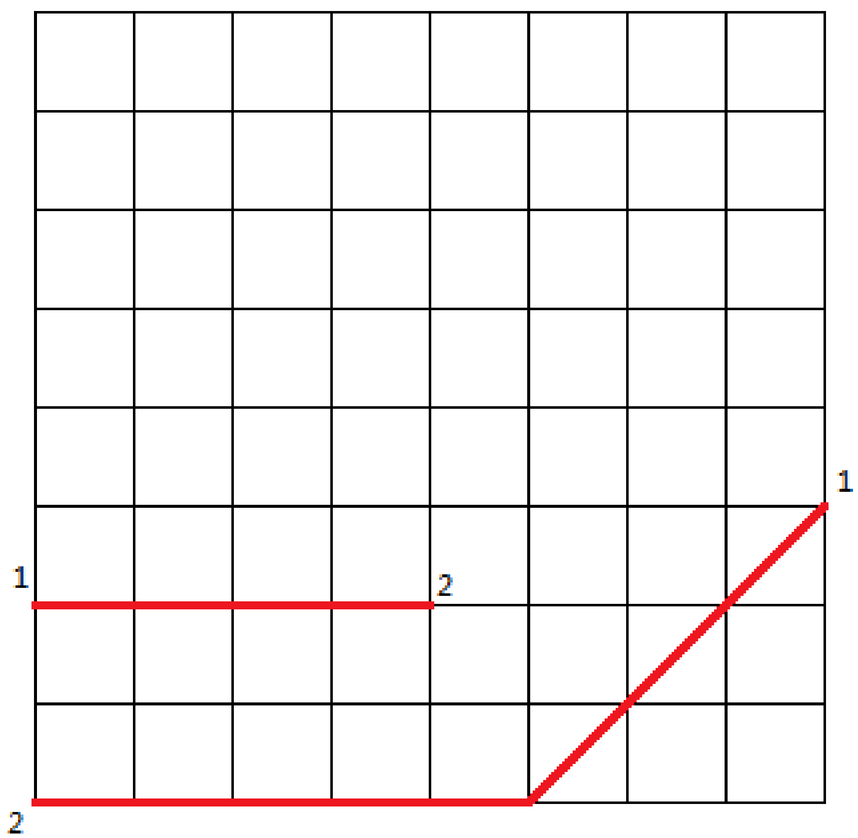

图 1: 反例

我们使用组合数计算每一个起点到每一个终点的答案，被挖掉的部分我们直接容斥，我们枚举是从哪一个位置进入不合法区域的，减掉这些路径的贡献即可。

复杂度变为 $O((\frac{n}{k})^3 (\frac{k}{2})^{\frac{n}{k}})$。结合算法三后期望得分为 70。

算法五

现在考虑串 S 的限制，我们发现这其实等价于第一个起点强制走若干步。于是我们将第一个起点直接移动到走完 S 后的位置。同样为了让不合法配对方式的路径强制相交，我们需要将 $(0, 0)$ 到第一个起点的路径上方和这条路径上的点全部挖掉（起点除外）。

但这样就不能组合数了，我们考虑用一个简单的 DP 计算路径数。我们发现要算的答案中，很多起点/终点的配对是重复的。于是我们将答案记忆化一下，这样就不会提高复杂度了。

结合算法二后期望得分：100。

参考资料

[1] Lindström-Gessel-Viennot lemma, Wikipedia

[2] 【NOI2021】路径交点