药水 解题报告

北京市十一学校 武林

题目大意

题面描述

给定概率序列 $a_l, a_{l+1}, \ldots, a_r$，保证 $\sum a_i = 1$。

考虑所有长为 n 的整数序列 $b_1, b_2, \ldots, b_n$，满足 $b_i \in [l, r]$，定义其出现概率为 $\prod_i a_{b_i}$。

定义序列 b 是好的，当且仅当存在 $c_1, c_2 \ldots, c_n$，满足 $c_i \in \{0, k\}$，使得数列 $s_i = \min(s_{i-1} + b_i, c_i)$ 所有元素 ≥ 0，且任意连续 m 项都有一项为 0，其中 $s_0 = 0$。

求所有好的 b 序列的出现概率之和对 998244353 取模的结果。

输入格式

第一行输入五个整数 n, m, k, l, r。

第二行输入 $r - l + 1$ 个整数 $a'_l \sim a'_r$，其中 a'_i 表示实际的 a_i 对 998244353 取模的结果，保证 $\sum a'_i \equiv 1 \pmod{998244353}$。

输出格式

输出一行一个数，表示答案对 998244353 取模的结果。

数据范围

subtask	n	$r-l+1$	特殊性质	分值
1	≤ 10	≤ 7	无	10
2	≤ 100	≤ 7	无	10
3	$\leq 10^4$	≤ 7	无	20
4	$\leq 1.2 \times 10^5$	≤ 3	$a'_{-1} = a'_1, a'_0 = 0$	15
5	$\leq 1.2 \times 10^5$	≤ 3	无	10
6	$\leq 6 \times 10^4$	≤ 5	无	15
7	$\leq 1.2 \times 10^5$	≤ 7	无	20

对于所有数据：$1 \leq m \leq n \leq 1.2 \times 10^5$，$1 \leq k \leq 10^6$，$-3 \leq l < 0 < r \leq 3$，$a'_i \in [0, 998244353)$，$a'_l, a'_r > 0$，$\sum a'_i \equiv 1 \pmod{998244353}$。

时限 5s，空限 1024MB。

解题过程

先特殊规定第 n 天结束后发生了清空，然后把问题抽象一下。

设 s_i 为第 i 天炼药锅里的药量，那么有 $s_i = \min(k, s_{i-1} + b_i)$，其中 b_i 是根据 a 带权随机的整数。

如果当天发生了清空，那么 $s_i \leftarrow 0$。

设 s'_i 为 $k = +\infty$ 且不考虑清空时的 s_i，我们考察在不清空的情况下，s_i 与 s'_i 有什么关系。

可以发现 $s_i = \min(s'_i, \min_{j=1}^{i} k + s'_i - s'_j) = s'_i - \max(0, \max_{j=1}^{i} s'_j - k)$。

因此一个不清空的段合法，当且仅当每个 s'_i 与其前缀最大值之差不超过 k，也即每个 s'_i 与其后缀最小值之差不超过 k。

再考虑什么位置可以清空，容易发现只有是 s'_i 后缀最小值（非严格）的位置可以清空，且 $s'_0 = 0$ 必须是一个后缀最小值。

因此可以发现，清空所有后缀最小值的位置不会改变原序列的合法性，且考虑到还有每个段长度 $\le m$ 的限制，实际清空的位置一定就是所有后缀最小值的位置。

考虑画出 s'_i 的折线图，然后将其左右翻转，考察（翻转后）一个前缀最小值和下一个前缀最小值之间的段什么时候合法。

不妨把开头平移到 0，可以发现这个段合法当且仅当过程中最大高度 $\le k$，且除了开头外有且仅有结尾高度 ≤ 0。

设 $F = \sum_{i=1}^{m} f_i x^i$ 为一段的生成函数，那么答案为 $[x^n]\frac{1}{1-F}$，因此只需求出 F，之后用分治 ntt 或多项式求逆即可。

下面我们把一个段的问题看作，从原点开始每步以 a_i 的概率走 $(1, i)$ 这个向量，求在 j 位置第一次 y 坐标 ≤ 0 的概率（除去原点）。

由于 $|l|, |r|$ 都很小，所以可以发现第一次超过 k 的位置一定是 $[k+1, k+r]$，第一次不高于 0 的位置一定是 $[l, 0]$，都是很小的区间。

因此考虑在 $y \in [l, 0] \cup [k+1, k+r]$ 的行放置障碍点，可以把原问题转化成对每个障碍点求出“第一次碰到的障碍点是它”的概率。

这是一个经典问题，设 f_p 是 p 这个障碍点的答案，那么我们有 $f_p = w(o, p) - \sum_q f_q w(q, p)$，其中 o 是原点，q 是另一个障碍点，$w(a, b)$ 是从 a 走到 b 的概率。

容易发现 $w(a, b) = w(o, b - a)$，设 $d = \max(-l, r), D = r - l + 1$，那么可以发现 y 的差只有 $[-d, d]$，$[k, k + D]$，$[-k - D, -k]$ 这 $O(D)$ 种。

设 $G = \sum_{i=l}^{r} a_{l+r-i} x^i$（系数翻转是因为考虑的序列被翻转了），从 o 走到 (u, v) 的概率就是 $[x^u]G^v$
。

因为 G 是短多项式，设 $H = G^v$，利用 $vG'H = GH'$ 进行对比系数有：

$$
\begin{aligned}
[x^t]vG'H &= [x^t]GH' \\
\sum_{i=l}^{r} v i g_i h_{t+1-i} &= \sum_{i=l}^{r} (t + 1 - i) g_i h_{t+1-i} \\
0 &= \sum_{i=l}^{r} h_{t+1-i} g_i (t + 1 - (v + 1)i) \\
h_t &= \frac{\sum_{i=l+1}^{r} h_{t+l-i} g_i (t + l - (v + 1)i)}{g_l (vl - t)} \\
h_t &= \frac{\sum_{i=l}^{r-1} h_{t+r-i} g_i (t + r - (v + 1)i)}{g_r (vr - t)}
\end{aligned}
$$

可以在有 G^v 连续 D 项系数的时候 $O(D)$ 递推上/下一项。

并且 G^v 的系数递推到 G^{v+1} 的系数也是容易的，可以直接拓展两侧长度然后暴力对 G 做卷积。

因此我们可以在 $O((D+R-L+1)Dm)$ 的复杂度内对所有 $i\in[L,R], j\in[1,m]$ 算出所有 $[x^i]G^j$。

具体来说，对于每个 j 都维护一段包含 $[L,R]$ 的，长度至少为 D 的区间的项的系数，不妨设这个区间为 $[l_0,r_0]$，然后把系数拓展到 $[l_0-r,r_0-l]$，再暴力和 G 做卷积即可得到 $j+1$ 处 $[l_0,r_0]$ 区间的系数。

那么用上述方法 $O(mD^2)$ 求出需要的三个区间 $[-d,d]$，$[k,k+D]$，$[-k-D,-k]$ 的转移系数之后，我们可以直接使用分治 ntt 优化 dp 来算出每个障碍点的答案，暴力做是 $O(mD^2\log^2 m)$ 的，不可接受。

但是注意到分治时处理左对右的贡献不需要每个卷积都先 DFT，再对位乘，再 IDFT 回去，而是可以先对这 $O(D)$ 个多项式做 DFT，做 $O(D^2)$ 次点值的乘法和加法，最后再 IDFT 回去，这样总复杂度就是 $O(mD\log m(D+\log m))$。

为了减小常数，可以 $O(mD\log m)$ 预处理转移系数的点值表达式，并利用循环卷积的性质优化分治 ntt 的长度。

多叉卷积可能也会有明显的优化，但出题人并没有尝试过。

这个 dp 也存在一个多项式求逆解多项式线性方程组的做法，可以做到 $O(mD^2(D+\log m))$，可惜常数太大，比上述做法慢了将近一倍。

参考资料

本题有一部分参考了 qoj7634 的做法。

感谢吴畅学长与我的讨论。