

数据结构4

目录

- 序列分块基础
- 操作分块
- 树形结构上的分块
- “根号分治”
- 莫队算法及其变种

序列分块基础

- 序列分块，就是将序列划分为若干较小的部分，通过维护每部分的信息，来快速地进行查询。

序列分块基础

- 区间加区间求和。

序列分块基础

- 常用的技巧：
- $O(\sqrt{n})$ 单点修改， $O(1)$ 前缀和。
- $O(\sqrt{n})$ 前缀和， $O(1)$ 单点修改。
- $O(\sqrt{n})$ 区间加， $O(1)$ 求单点。
- $O(\sqrt{n})$ 求单点， $O(1)$ 区间加。
- 本质相同。

数列分块入门 2、3

- [LOJ 6278](#)、[LOJ 6279](#)
- 区间加，查询区间内小于 x 的元素个数。
- 区间加，查询区间内小于 x 的元素的最大值。
- 经典技巧：归并
- 必备知识：合理设计块长，以达到最优复杂度。

数列分块入门 8

- [LOJ 6284](#)
- 区间查询等于 c 的个数，并将这个区间全部修改为 c 。
- 复杂度分析：一个块是否只有一种数，有多种数的块有多少个。

五彩斑斓的世界

- [洛谷P4117](#)
- 区间将大于 x 的数减去 x 。
- 区间查询 x 的出现次数。
- 值域 500000。

五彩斑斓的世界

- 稍微难一点的题。
- 经典操作：将 x 全部变成 y
- 经典技巧：逐块处理以节省空间。
- 令 k 为最大值，分别 $2x \geq k$ 和 $2x < k$ 的情况。

值域分块

- 简单来说，就是对桶进行分块。
- 和序列分块差不多。
- 要求值域较小，一般和序列长度同阶。

值域分块

- 给定一个集合，支持插入、删除一个数，查询第 k 大。

操作分块

- 对时间序列的分块。
- 将操作和询问划分为若干个块。
- 维护一个块开始前的数据并支持高效查询。
- 处理块内的操作对询问的贡献。
- 块外的操作对询问的贡献直接查询得到。
- 一个块内询问完毕后，更新数据到下一个块开始前。

操作分块

- 有 $k+1$ 个整数序列 $a_0, a_1, a_2, \dots, a_k$ 。
- 其中 a_1 是 a_0 的前缀和数组, a_2 是 a_1 的前缀和数组, 以此类推。
- 对 a_0 单点修改。
- 单点查询任意一个序列。

APIO2019 桥梁

- 给定一张带边权的无向图。
- 修改一条边的边权。
- 查询某个点出发，只走边权不超过 m 的边，能走到的不同的点的个数。
- 思考：修改较少的时候怎么做。

DFS 序分块

- 对树的 DFS 序进行序列分块。
- 局限：在涉及链修改时需要结合树链剖分，复杂度会多 \log 。

轻重链剖分+序列分块

- 对一条长度为 k 的重链按照块长为 \sqrt{k} 分块（假设已经平衡各个操作的复杂度）。
- 考虑跳轻边时子树大小至少减少一半，而重链长度不超过子树大小。
- 因此单次操作的复杂度为 $O(\sqrt{n}) + O\left(\sqrt{\frac{n}{2}}\right) + O\left(\sqrt{\frac{n}{4}}\right) + \dots = O(\sqrt{n})$ 。
- 注意块的个数最多可能达到 n 。
- 局限：仅适用于路径修改。

CF925E May Holidays

- 给定一棵树，每个点有点权 t_i 和颜色（黑/白）
- 若一个点为白点，且这个点为根的子树内，黑点个数超过 t_i ，则这个点是好点。
- 操作为修改一个点的颜色，每次操作完查询好点个数。

CF925E May Holidays

- 问题转化为单点修改，链 ± 1 ，查询 < 0 的点的个数。
- 考虑序列上的做法。

简单树分块

- 在 n 个点的树上随机撒 \sqrt{n} 个关键点。两个相邻关键点之间的距离期望为 \sqrt{n} 。
- 用贪心可以将期望变为严格，但需要额外的 $O(n)$ 空间。
- 若有需要，也可以将关键点的虚树上的点都标记为关键点。
- 局限：仅适用于路径修改

SDOI2022 无处存储

- 给定一棵树，点有点权，要求支持链加链求和。
- $n \leq 7000000$ $q \leq 50000$
- 5s, 64MB

简单树分块

- 将树分为若干个块，每个块满足：
- 存在一个点（这个点可以不在块里），使得所有块内的点都能通过块内的边到达这个点（以菊花图为例）。
- 例题：[王室联邦](#)。
- 局限：似乎只能做树上莫队。

Top Cluster 树分块

- 详见周欣的集训队论文 《浅谈一类树分块的构建算法及其应用》
- 学习难度和代码难度都极高，且算法竞赛中的实用意义不大。

“根号分治”

- 核心思想是，设定一个阈值，对超过和不超过阈值的数据采用不同的算法处理，以达到复杂度平衡。

哈希冲突

- [洛谷P3396](#)
- 给定一个序列。
- 单点修改
- 查询模 x 余数为 y 的所有位置的和。

哈希冲突

- 暴力处理询问，查询复杂度是 $O(\frac{n}{x})$ ，修改复杂度是 $O(1)$ 。适用于 x 较大的情况。
- 维护数组 $f[x][y]$ 表示模 x 余 y 位置的总和。查询复杂度是 $O(1)$ ，修改复杂度是 $O(n)$ 。需要 $O(n^2)$ 的空间。
- 当 $x > \sqrt{n}$ 时用前者，反之用后者。这样单次修改、查询的最坏时间复杂度均为 $O(\sqrt{n})$ ，空间复杂度 $O(n)$ 。

初始化

- [洛谷P5309](#)
- 给定一个序列。
- 将模 x 余数为 y 的所有位置加上 z 。
- 区间求和。

初始化

- 直接暴力做的话，一次操作会产生 $\frac{n}{x}$ 次单点加，还需要支持区间求和。
- 序列分块可以 $O(1)$ 单点加， $O(\sqrt{n})$ 区间求和。
- 维护数组 $f[x][y]$ 表示模 x 余 $0 \sim y$ 增加的总数，查询一个 x 可以 $O(1)$ 。

莫队算法

- k 维莫队的一个状态可以用一个 k 元组表示。
- 当得知一个状态的信息时，可以以较小的代价得知这个状态的相邻状态 (k 元组的某一维 ± 1) 的信息。
- 这样就可以从前一个目标状态，通过不断转移到相邻状态，来得到后一个目标状态。
- 莫队算法对所有目标状态按照某种方式排序，以尽可能减小“转移到相邻状态”这一操作的次数。

莫队算法

- 常见的二维莫队的形式为：不带修区间查询。
- 长度为 n 的序列， m 次查询，转移到相邻状态的复杂度为 $O(1)$ 。
- 最优复杂度为 $O(n\sqrt{m})$ 。
- 排序方式为，将序列按照块长为 $\frac{n}{\sqrt{m}}$ 分块，对询问排序，第一关键字为左端点所在块，第二关键字为右端点。
- 另外，三维莫队的最优复杂度是 $O(n^{\frac{5}{3}})$ (n, m 同阶)，按照 $n^{\frac{2}{3}}$ 分块可取到最优。

小 Z 的袜子

- 有 n 个物品，每个物品有一个颜色。
- m 次询问，每次给定一个区间，求区间里等概率随机抽取两个不同物品，得到颜色相同的概率。
- 经典题。

组合数前缀和

- 多次询问，每次给定 n, m ，求
- $\sum_{i=0}^m \binom{n}{i}$
- 对质数取模。

组合数前缀和

- 不常见的莫队形式。
- $\binom{n}{m} = \binom{n-1}{m} + \binom{n-1}{m-1}$

树上莫队

- 一般以询问树上路径信息的形式出现。
- 考虑写出树的括号序列（dfs 时，进入该点添加左括号，退出该点添加右括号），在该括号序列上进行莫队即可。
- 维护的信息需要支持可删。

树上莫队

- 进行树分块，然后对询问进行排序。

Count on a tree II

- [SPOJ 10707](#)
- 给定一棵树，每个点有一个颜色。
- 多次询问树上一条路径上的颜色个数。
- 树上莫队入门题。

不删除莫队/不插入莫队

- 有时维护的信息不支持删除或不支持插入，则需要对莫队算法进行改进。
- 以不删除莫队为例。
- 考虑对左端点在一个块内的询问同时处理，按照右端点排序，这样保证了右端点单调。
- 对每个询问，先移动右端点，再移动左端点。处理完后再将左端点移回来（撤销左端点移动带来的影响）。
- 需要保证左端点单次移动的代价。

JOISC2014 历史研究

- 给定一个序列，每个位置有一个颜色 c_i 。
- 每次询问给定一个区间。定义颜色 k 的价值为， k 乘上 k 在这个区间内的出现次数。一个区间的价值为所有颜色价值的最大值。求区间价值。
- max/min 是典型的不支持删除的信息。

WC2022 秃子酋长

- 给一个长为 n 的排列，有 m 次询问，每次询问区间 $[l,r]$ 内，排序后相邻的数在原序列中的位置的差的绝对值之和。

WC2022 秃子酋长

- 链表支持 $O(1)$ 前驱、后继、删除。
- 每次删除一个元素，我们可以直接知道它对应链表中的位置，然后删掉它。
- 不插入莫队。

莫队二次离线

- 可以解决单次转移代价较大的序列莫队问题。
- 适用范围：
- 一个数对区间的贡献与区间内的数有关。
- 设 $F(x, l, r)$ 为 x 对区间 $[l, r]$ 的贡献，则满足 $F(x, l, r) = F(x, 1, r) - F(x, 1, l-1)$
- 将端点移动的贡献进行上述拆分，对两部分分别讨论。

【模板】莫队二次离线

- 给定一个序列 a ，每次询问给定一个区间 $[l,r]$ ，问有多少对 (i,j) 满足 $l \leq i < j \leq r$ 且 $\text{popcount}(a_i \text{ xor } a_j) = k$ 。
- 值域 16384