中国计算机学会
China Computer Federation

NOI冬令营2020 第一课堂

杂题选讲

OI与一些组合问题

CommonAnts
刘承奥
lca19@mails.tsinghua.edu.cn
清华大学 交叉信息研究院

中国计算机学会
China Computer Federation

本课件为杂谈和杂题选讲形式，旨在分享一些组合问题的思路。各节标题只是一种分类和引导，并不表示本节涵盖了OI在这个概念下的大部或全部内容。

Index... 杂谈与杂题选讲

1. OI的组合基础
   1. 集合与集族
   2. 分治
2. 序列杂题
3. 排列
   1. 最值与笛卡尔树
   2. 连续段问题，PQ树与析合树
4. 树
   1. 不交集
   2. 树的重心与中心
   3. 树分治与商集

组合？

- 用集合来描述OI的组合问题……

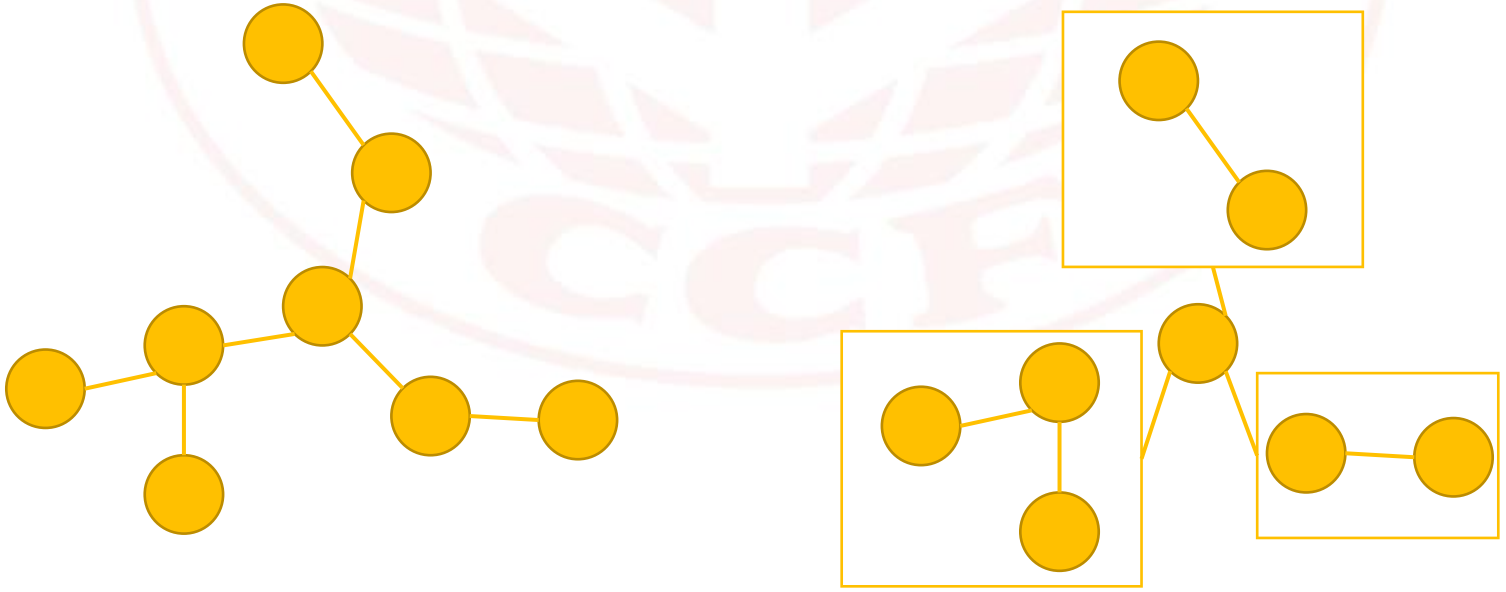
- 组合问题的模式：有一个集合S，和一个关于S子集的性质P（S的某些子集满足P，其它不满足），你需要：
- 判定是否存在满足P的子集A（判定）
- 输出一个满足P的子集A（构造）
- 找到所有满足P的子集A里权值最大的/极大的（优化）
- 求满足P的子集A的总个数（计数）
- ……

集合与集族

- **基础集S**：问题的基本元素的集合（例如序列里的所有位置、一棵树上的所有边和点……）
- **集族$I \subseteq 2^S$**：S的若干子集的集合（例如序列里的若干区间，树上的若干联通块……）
- **集族$I_P = \{A | P(A) \wedge A \subseteq S\}$**：我们需要讨论的集族。

分治？

- 分治过程：原来的基础集被划分成了若干部分，对每个部分分别处理，最后合并结果。
- 称分治关系$R=\{(a,b)|a,b$分到了同一部分$\}$，所有“部分”（“分治到同一部分”的等价类）构成的集合是这个分治对应的**商集**S/R。



分治之后？

- 集族I中的每个集都被划成了若干部分……

- I_P呢？
- 按照分治的要求，I_P对于每个部分应当是独立的。

- 设划分出来的部分是$S/R = \{S_1, S_2, \ldots\}$，可能有若干种情况：
- I_P的直和（如图上不同的联通块）
- I_P在每个部分的子集，并上I_P'的直和（如点分治、边分治等）
- 更复杂的情况

直和

- $I_P = \{\cup A_i | A_i \in I_P \cap 2^{S_i}\}$
- 也就是说，每个集与每个分治部分的交仍然满足P；反之，从每个分治部分中任取一个满足P的集，它们的并仍然满足P。
- 这种情况下我们称I_P是$\{I_P \cap 2^{S_i}\}$的**直和**，称A_i是$A = (\cup A_i)$在S_i上的**投影**。

- 写成直和的问题只需要分别讨论每个部分即可。
- （优化问题只需分别优化，计数问题即乘法原理）

- 例如图的每个联通块，线性基的每个元素张成的空间。

中国计算机学会
China Computer Federation

Independent and Invariant

- 直和表达了部分之间的某种独立性。也就是说，
- $f(\cup A_i) =\circ_i F(A_i)$

- 但这个性质很少成立。我们考虑一种更弱的情况：
- $f(\cup A_i) =\circ_i F(\cup_{j\leq i} A_j)$
- 或者
- $f(\cup A_i) =\circ_i F(\{A_j|j \leq i\})$
- 也就是，存在某个序使得每个部分只和之前的部分有关。这表示原问题可以递推计算。
- 对最优化问题来说，两式分别对应贪心和动态规划的方法。

中国计算机学会
China Computer Federation

「十二省联考 2019」皮配

- 有n（≤1000）组学生（总人数≤2500）分属c座城市。
- 每组学生要选择四位导师之一。四位导师组成阵营和派系：

	红阵营	蓝阵营
Y派系	A	B
R派系	C	D

- 同一城市的学生必须加入相同阵营。
- 有k（≤30）组学生各有一个不能选择的导师。
- 每个阵营和派系有一个人数上限。
- 求方案数mod 998244353。

[BJ United Round #3] 押韵

- 一个长nd的序列，每个位置要染成k种颜色之一，要求每种颜色出现次数都是d的倍数。
- 求方案数mod 1049874433。
- $n \le 10^9$ $k \le 2000$ $d=1,2,3,4,6$

- $[\mathbb{Z}:\mathbb{Z}(\omega_x)]$有限，当且仅当$x=1,2,3,4,6$

中国计算机学会 China Computer Federation

「清华集训 2017」无限之环

- 接水管游戏：一个n×n的网格状棋盘，每格是15种水管之一。

- 你可以逆时针或顺时针旋转任意非直线型水管任意次（13种）。
- 给定初始局面，要求你旋转水管使得最终不存在断头，最小化旋转总度数。

【CSP-S 2019】树上的数

- 一棵n个节点的树编号1到n，同时每个点上还有一个数字，构成排列。
- 每次删一条边并交换两端节点上的数字。
- 求删完所有边后可能出现的字典序最小的数字排列。

图的直和

- 定义无向图$G_1=(V_1,E_1)$，$G_2=(V_2,E_2)$的直和为G：

- G的点集为所有$u\in V_1$，$v\in V_2$构成的二元组(u,v)。
- G的边集包含$((u_1,v_1),(u_2,v_2))$当且仅当$u_1=u_2,(v_1,v_2)\in G_2$，或者$v_1=v_2,(u_1,u_2)\in G_1$。

均衡

- K_2表示两个点的无向完全图。
- 令D_k表示k个K_2的直和。
- 给定H，求最大的k使得存在无向图G满足H同构于D_k与G的直和。

- H的规模不超过2×10^5

集族的对偶

- 把 I_P 作为新的基础集，原来基础集的每个元素对应一个新的子集，即原来包含该元素的所有子集。

- $S^* \sim I_P$
- $I_P^* \sim \{\{A|A \in I_P \wedge i \in A\}|i \in S\}$

- 例如，在有根树上，子树与到根路径是对偶的。

集族的对偶和数据结构

- 假设有一个二分图，左侧是所有基础集内的元素，右侧是集族内的所有元素（也就是若干子集），每个子集向它的所有元素连边。
- 那么集族的对偶就是交换此图的左右两侧。

- 例如：若数据结构维护的是，单点修改左侧某个元素（基础集），查询集族内某个元素包含的元素和。那么在对偶问题上，这个数据结构就变成了修改一个集族内的所有元素，查询单点。

数据结构和对偶问题？

- 对偶：$D \to R$和$D \leftarrow R$
- 数据结构问题的常见对偶：
  - 交换维护的权值的值域和定义域（下标）；
  - 交换询问和修改（如上页所述）；

- 单点修改区间查询和区间修改单点查询
- 区间修改区间查询
  - 线段树
- 区间k大与区间排名
  - 划分树与归并树
- 连续段问题
  - PQ树与析合树
- 对偶问题通常难度类似。

序列上的问题

- 以下是几道序列组合题目……（也就是说，I_P是一些子段或者不连续的子序列）
- 这些序列组合问题通常可以通过寻找I_P中的某种序解决（例如端点跨度尽量大或端点的字典序等等），往往有$A \leq B \Leftrightarrow P(A) \to P(B)$。

「LibreOJ Round #9」Menci 的序列

- 有一个×和+组成的序列。×表示乘2，+表示加1。
- 你需要选出一个子序列使得0在依次执行操作后对2^k取模的结果尽可能大。
- 序列长度$\leq 10^6$

中国计算机学会
China Computer Federation

调整法

- 对于序列问题，考虑把合法方案调整成特定形式。
- $++\times++\times++ \rightarrow \times\times++\times++\times++ \rightarrow \times\times++\times+++\times \rightarrow \times\times+++\times+\times \rightarrow \times+\times+\times+\times (1110_2)$
- $\times$右侧的两个$+$ $\rightarrow$左侧的一个$+$?
- $+\times++$有子序列$+\times+(11)$
- $++\times$无法得出11

【NOI2019】序列

- 给定两个长度为n的正整数序列。你需要从两个序列各取K个数，要求至少有L个下标在两个序列中对应的数都被取出，使得取出的所有数总和最大。

- $n \le 2 \times 10^5$

「LibreOJ NOIP Round #1」序列划分

- 有一个长度为n的整数列。
- 请你将其划分成k的倍数个子序列（不一定连续），同时每个子序列以k的倍数开头，k的倍数结尾，长度也为k的倍数。

- $n \le 10^6$

【CSP-S 2019】划分

- 给定一个长为n的整数列。
- 要求将其划分成若干段，每段的总和依次递增。最小化每段总和的平方和。

- $n \le 10^7$

最值与笛卡尔树

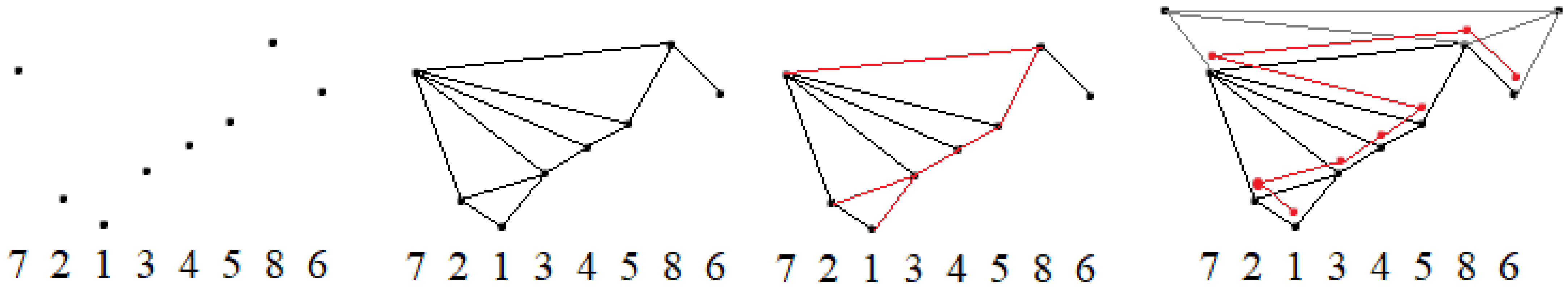
- 区间最值问题：给定一个排列，每次询问一个区间的最值。（不妨认为是最大值）
- 可以注意到这个问题的结构：整个序列的最大值在任意包含它的区间中都是最大值，所以可以从最大值处将序列切开分治。
- 也就是说，笛卡尔树是一个满足搜索树性质的堆（每个子树对应一个区间，并且祖先权值大于子孙）。
- 据此可以把笛卡尔树推广到树上等。

「LibreOJ Round #8」MIN&MAX I

- 对于一个n阶排列p，建立一张无向简单图G(p)，有n个节点，标号从1到n，每个点分别向排列中对应位置左右两侧最近的比它大的元素以及比它小的元素连边。
- 在所有的n阶排列中等概率随机选择一个排列p，求G(p)中三元环的期望个数，答案对998244353取模。

笛卡尔树的结构

- 从每个点向排列中左，右两侧最近的比它大的元素连边，可以得到一张外平面图。
- 其对偶图与笛卡尔树同构（在两端添加无穷大，忽略无限区域）。
- 只考虑其中向一侧的边，可以得到左树和右树。
- 一个点在笛卡尔树上的祖先是左树和右树祖先的并。

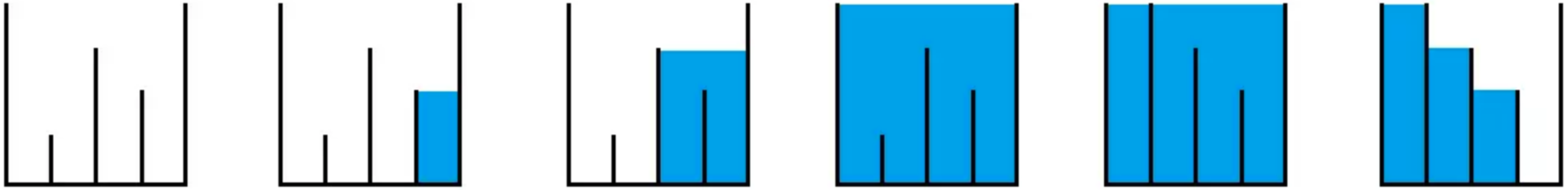


「SNOI2020」水池

有一个长条形的水池，可以划分成 n 格。其中第 i 格和 $i+1$ 格之间相邻，由一块高度为 h_i 的可调节挡板隔开。第 1 格左侧和第 n 格右侧是无限高的池壁。初始时水池中没有水。现在进行 q 次操作，操作有以下四种：

- `0 i x h` 在第 x 格灌水直到该格的水面高度不低于 h（若当前水面高度已经达到 h 则无事发生）；
- `1 i x` 打开第 x 格底部的排水口直到该格的水流干，再关闭排水口；
- `2 i x h` 将第 x 格右侧的挡板高度增加到 h（不改变现有水面，保证挡板高度不会下降）；
- `3 i x` 查询第 x 格的水面高度。

其中，i 表示这次操作是基于第 i 次操作之后的情况，$i=0$ 表示基于初始状态。也就是说，这个问题要求对操作可持久化。



「2018 集训队互测 Day 3」北校门外的未来

对于一棵树 $T=(V,E)$，V 中每个点有一个互不相同的正整数标号。我们用点 i 表示编号为 i 的点。

定义这棵树的谷图为 $G(T)=(V,E')$。$G(T)$ 是无向简单图。存在边 $(u,v)\in E'$ 当且仅当在 T 中，不存在一个异于 u,v 的点 x 满足 x 在从 u 到 v 的简单路径上且其编号大于 $\min(u,v)$。

有一棵树 T，初始时只有一个点，编号为 1，接下来有 q 次操作，操作有以下两种：

- `1 u v` 表示加入一个编号为 v 的节点并与当前编号为 u 的节点相连（保证任何时刻不会有两个编号相同的节点）；
- `2 u v` 表示查询 $G(T)$ 中点 u 到 v 的最短路（每条边长度均为 1）。

请你回答所有查询。

中国计算机学会
China Computer Federation

树与不交集族

- 如果I_P中任意两个元素至少满足下列之一：
- 交集为空；
- 其中一个包含另一个；
- 并集为S（或者说，二者补集的交集为空）

- 那么一定存在一棵点数与$|I_P|$同阶的无根树使得I_P与其一部分子树同构。称这样的集族I_P是不交集族。

- 如果都满足前两条之一，那么一定存在一棵点数与$|I_P|$同阶的有根树使得I_P与其一部分子树同构。称这样的集族I_P是嵌套集族。

「LibreOJ NOI Round #2」小球进洞

有若干个小球放在数轴上，第 i 号小球的坐标参数为 a_i。

有两种操作：

1. 输入 i, v，修改第 i 号小球的坐标参数为 $a_i \leftarrow v$；
2. 输入 l, r，询问下述内容：

   按照 a_i 从小到大的顺序（a_i 相等时按 i 从小到大的顺序）依次将小球放在数轴上。第 i 号小球放在 $\leq a_i$ 的没有被之前放置的小球占据的最大的整点处，设第 i 号小球放置的位置为 b_i。

   请你输出 $\sum_{i=1,2,\ldots,n,[l,r]\subseteq[b_i,a_i]}(a_i+b_i)$ 的值。也即，所有满足 $b_i \leq l, a_i \geq r$ 的小球的 $a_i + b_i$ 之和。

最小割树

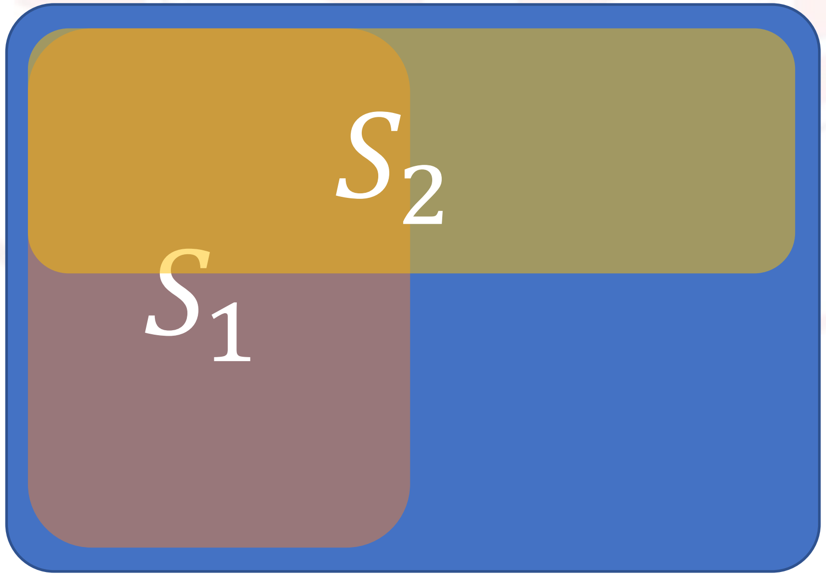
- 无向图$G = \langle V, E\rangle$的割是对点集的二分，将点集V划分成$V = S \cup (V \setminus S)$（令$V \setminus S = \bar{S}$）。称$S$是割集。
- 割的权值定义为跨越两个部分之间的边权和。

- 对于$u, v \in V$，称满足$u \in S, v \in \bar{S}$的割集S是u到v的割集；称所有u到v的割集中割权值最小的是u到v的最小割集。
- 在每个点对之间取一个最小割集可以组成G的最小割集族。

- 我们有引理：
- 存在某个最小割集族是不交集族；

最小割树

- 因而，我们可以定义出无向图的最小割树（Gomory-Hu），树上的边表示该边两侧之间的割，边权是割的权值。
- 两点之间的最小割等于它们树上路径上任一最小边对应的割。



最大子段和问题

- 给定一个长为n的数列，每次查询一个区间的最大子段和（这个区间的子区间中，权值和最大的那个）。
- 要求支持**区间加一个非负整数**。（假设询问和修改次数$O(n)$）
- 无修改/单点修改：线段树维护每个段的最大前缀、后缀和子段和。
- 区间修改的问题：区间整体加一个数后不知道最大子段是否改变。（可能会变成一个更长的子段）
- 朴素做法：对每个段，线段树T直接维护区间内最大子段改变需要整体加的最小数，发生改变就递归计算。
- 时间复杂度等于$O(n\log n + \log n \times$ 线段树最大子段改变次数总和$)$

最大子段和问题

- 全局加导致的最大子段改变次数？

- 考察最大子段的结构，有引理：
- 对于任一区间，区间内不同时刻的极长最大子段构成嵌套集族
- 这是因为一个最大子段的每个前缀与后缀的和都是非负的，所以两个最大子段的并在较晚的时刻一定也是最大的。

- 于是，每个段的最大子段都形成一棵树。一次操作可能在叶子间切换，或者合并相邻叶子。后者的总次数显然不超过所有线段树节点总长$O(n\log n)$。
- 切换叶子次数似乎也不超过$O(\text{区间长度}+\text{区间内部修改次数})$，这样总复杂度是$O(n\log^2 n)$。但是，严格证明？

「SNOI2020」区间和

有一个长度为 n 的整数数列 $a_1, a_2, \ldots, a_n$（可能含有负数）。现在对其进行 q 次操作，每次操作是以下二者之一：

- `0 l r x` 表示对于 $i = l, l+1, \ldots, r$，将 a_i 赋值为 $\max(a_i, x)$；
- `1 l r` 求区间 $[l, r]$ 的最大子段和。即：$\max(0, \max_{l \le u \le v \le r}(\sum_{i=u}^{v} a_i))$。

中国计算机学会
China Computer Federation

嵌套集族

- 对于一个集族I_P，我们找到I_P中所有不与其它元素“相交且互不包含”的：
- $I_T = \{A | A \in I_P \wedge (\nexists B \in I_P \text{ s.t. } A \cap B \wedge A \setminus B \neq \emptyset \wedge B \setminus A \neq \emptyset)\}$
- 它们显然是一个嵌套集族。我们在这个嵌套集族形成的树的基础上分析I_T的结构，只需考虑树上每个点和孩子的关系即可。

- 下面我们来看一些例子。

连续段问题

- 有一个排列P，令I_P表示P中所有连续段的集合。
- 连续段是一个区间，满足区间长度等于区间最大值减最小值（或者说区间的值域是连续的）。

1. 连续段的交是连续段；
2. 相交连续段的并是连续段；
3. 相交但互不包含连续段的差是连续段；

- 连续段（连通块）的常用判定方法：

1. $R - L = \max - \min$（维护历史最值）
2. $V - E = 1$（维护$V - E$的最小值）
3. 考虑包含每条边的极小连续段（传递闭包）

- 连续段的结构是怎样的？

CERC2017 I. Intrinsic Interval.

- 给定排列p，每次询问一个区间，求最小的包含它的连续段。

- 离线询问
- 在线询问？

连续段的结构

- 按照上述提出嵌套集的方法，我们可以定义：
- **本原连续段**：对于一个连续段$X \in I_P$，如果不存在与之相交且互不包含的连续段，就称X是**本原连续段**，所有本原连续段形成I_X。

- 那么在这些本原连续段的树上，每个点的孩子之间有什么关系呢？
- 我们有引理：
- I_X上每个点X的孩子至少满足以下二者之一：

1. 所有非平凡的连接都是连续段；
2. 所有非平凡的连接都不是连续段。

- 其中非平凡意为长度不是0、1或$|X|$。

析合树

- 按照这两种情况我们可以给树上的点分类，满足“所有非平凡的连接都是连续段”的称为合点，否则称为析点。

- **连续段集合可以等价地由一棵具有两种节点的有根有序树表示。**

ECFinal2018 B. Mysterious … Host.

- 给定n，问n阶排列的连续段集合可能的种类数。

- $n \le 10^5$

对偶问题与PQ树

- 对偶问题：已知I_P，构造原来的排列p。
- 弱对偶问题：已知I_P的一个子集（即钦定某些区间是连续段），构造原来的排列p。

- 类似地考虑树结构。但在弱对偶问题中，不能直接区分出析点和合点。（因为一个析点改成合点会使I_P变得更大）
- 我们重新分出两类点：
- P点（对应析合树上的一个连通块）：孩子可以任意排列；
- Q点（对应合点的一段孩子）：孩子只能正序或者逆序排列；
- 新的树即为PQ树。(Kellogg S. Booth & George S. Lueker, 1976)

连续段问题的推广？

- 树上编号连续的连通块？

- 两棵树的点之间有一一映射，在两棵树上同时是连通块的点集？
- 利用两棵树规约到传递闭包；

重心和中心

- 树的重心是删去该点后最大联通块最小的点；
- 即以该点为根时最大子树最小的点。
- 树的中心是到最远的点距离最近的点；
- 即以该点为根时树深度最小的点。
- 重心和中心可能在一条边的中间。（一般认为此时两端都是重心）

- 树上所有直径（最长简单路径）都通过中心。

- 如果令$f(u,L)$表示树上所有距离u不超过L的点，那么$I_P=\{f(u,L)\}$对交封闭。

中国计算机学会
China Computer Federation

「2018-2019 集训队作业 Day 1」蜀道难

对于一棵有标号有根树 $T=(V,E)$，标号 $p:v\to p(v), v\in V, p(v)\in[1,|V|]\cap\mathbb{Z}$ 是一个一一映射。令一条边 $e=(u,v), e\in E$ 的边权为：$w:e\to w(e)=|p(u)-p(v)|, e\in E, w(e)\in\mathbb{Z}$。令整棵树的权为：$W:T=(V,E)\to W(T)=\sum_{e\in E}w(e)$。

另外定义一个图 $G(T)=(V,E')$，其中 $(u,v)\in E'$ 当且仅当在 T 中 u 到 v 路径上点的标号 $p_1,p_2,\cdots,p_l$，要么单调递增，要么单调递减。则 p 必须使得 $G(T)$ 的直径不超过 2，即 $\max_{i,j\in V} SP(i,j)\le 2$，其中 $SP(i,j)$ 表示 $G(T)$ 中 i,j 的最短路经过的边数。

现在给定 T，求 $M(T)=\min_p W(T)$。

并且有若干次操作：在 T 中加入一个新的叶子 v（$V\leftarrow V\cup\{v\}, E\leftarrow E\cup\{(x,v)\}, x\in V_{old}$），每次操作后也要求 $M(T)$。这些操作是一脉相承的。

图和树的分治？

- ~~树分治有四种分法，你知道么？~~

- 我们可以按照商集对树分治的方法进行分类：

	基础集的形态	商集的形态	不动点	联通块
点分治	树	星形树	点	点
边分治	树	边	边	点
点双 （圆方树）	图 （仙人掌）	仙人掌 （树）	点	环 （点）
边双 （树分块）	图 （树）	树	边	点
虚树 （Topo Cluster）	树	树	点	边

中国计算机学会
China Computer Federation

「SNOI2019」网络

- 有一棵n个节点的树（单位边权），每次询问一个点u，要求选出一个包含u的点集，满足点集内两两距离不超过d。求点集内两两距离和的最大可能值。

「十二省联考 2019」希望

- 有一棵n个节点的树，你需要依次选出k个点集，满足这些点集相交，且存在一个点到每个点集的任意点距离都不超过L。
- 求方案数，对998244353取模。

参考文献

- B. Korte, J. Vygen, 《组合最优化：理论与算法》，越民义等译。
- R. L. Graham, D. E. Knuth, O. Patashnik, 《具体数学》，张明尧等译。
- S. H. Friedberg, A. J. Insel, L. E. Spence, *Linear Algebra*.
- L. Lovász, J. Pelikán, K. Vesztergombi, *Discrete Mathematics*.
- Booth, Kellogg S. & Lueker, George S. (1976). "Testing for the consecutive ones property, interval graphs, and graph planarity using PQ-tree algorithms". *Journal of Computer and System Sciences*. 13 (3): 335–379.

参考资料

- EntropyIncreaser，《区间增量最大子段和的 polylog 做法》，http://entropyincreaser.blog.uoj.ac/blog/5217
- 刘承奥，《简单的连续段数据结构》，WC2019营员交流。
- 毛啸，《CSP2019划分的简要题解》，http://matthew99.blog.uoj.ac/blog/5299

Q&A

GL&HF