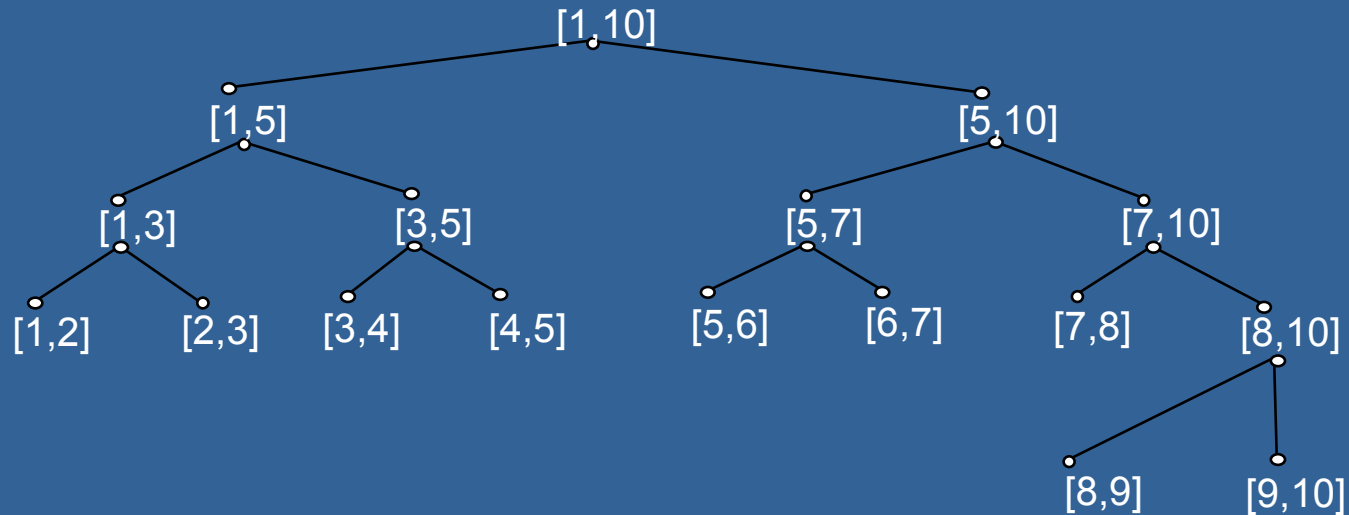


线段树的应用

广西柳铁一中 林涛

线段树的定义

表示区间 $[1, 10]$ 的线段树样例：



线段树是一棵二叉树，记为 $T(a, b)$ ，参数 a, b 表示区间 $[a, b]$ ，其中 $b-a$ 称为区间的长度，记为 L 。

线段树 $T(a, b)$ 也可递归定义为：

若 $L > 1$ ： $[a, (a+b) \div 2]$ 为 T 的左儿子；

$[(a+b) \div 2, b]$ 为 T 的右儿子。

若 $L = 1$ ： T 为叶子节点。

线段树的特征

- 定理一：线段树的深度不超过 $\log L$ 。
- 定理二：线段树把区间上的任意一条线段都分成不超过 $2\log L$ 条线段。
- 这些结论为线段树能在 $O(\log L)$ 的时间内完成一条线段的插入、删除、查找等工作，提供了理论依据

线段树的基本操作

- 插入一条线段：每个节点增加一个变量记录该节点被插入所有线段覆盖的次数，自根节点往下，直到一个被线段完全覆盖的节点。

- 删除一条插入过的线段，与插入操作是一致的。

节点被插入的线段覆盖

记录变量加 1，
儿子不再处理

- 查找一个区间内线段的总长度。每个节点增加一个变量记录该区间内线段的总长度，并在插入和删除后维护相关节点的这个变量：

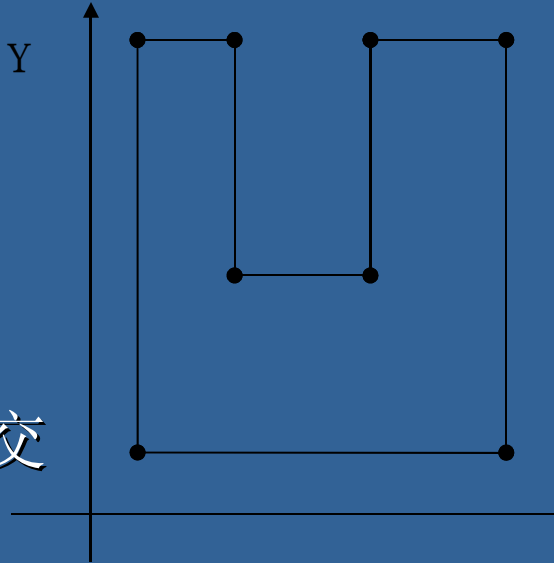
节点与插入线段不相干

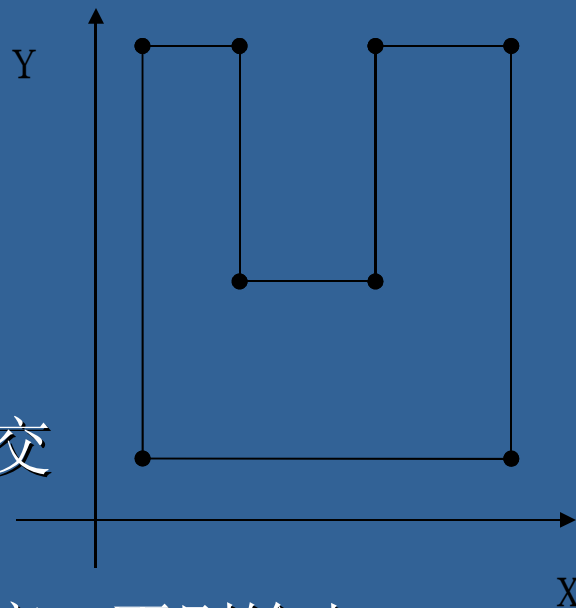
该节点及儿子都不作处理

- 儿子区间内的线段长度 + 覆盖该区间的线段长度
度 相干但不被覆盖

把线段分别插入它的两个儿子

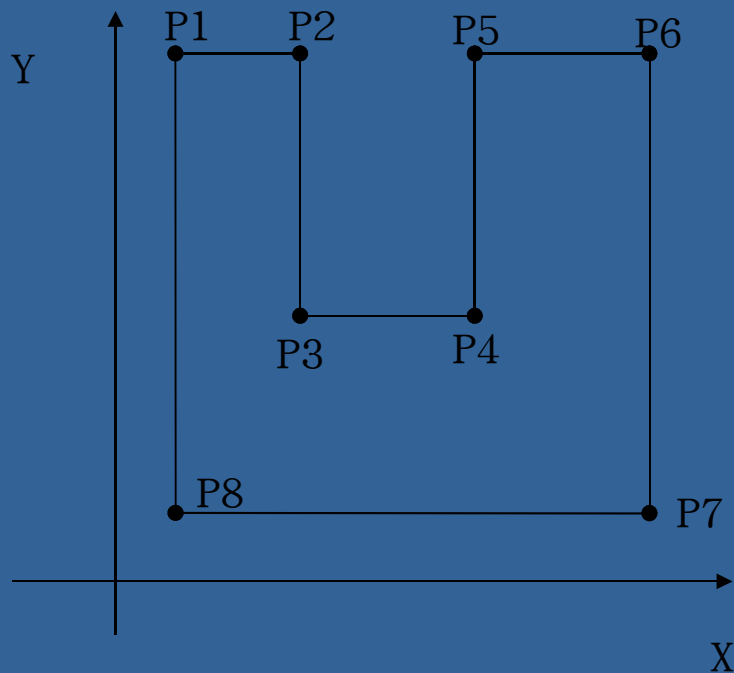
例 1：蛇（SGU 128）

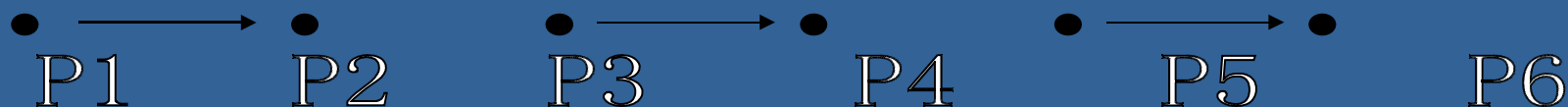
- 在平面上有 N 个点，现在要用一些线段将它们连起来，使其满足以下要求：
- 1 这些线段必须闭合
 - 2 线段的端点只能是这 N 个点。
 - 3 交于一点的两条线段成 90° 度角
 - 4 线段都必须平行于 X 轴或 Y 轴
 - 5 所有线段除了在这 N 个点外不相交
 - 6 所有线段的长度之和必须最短
- 如果存在这样的线段，则输出最小长度，否则输出 0。
- 



【问题分析】

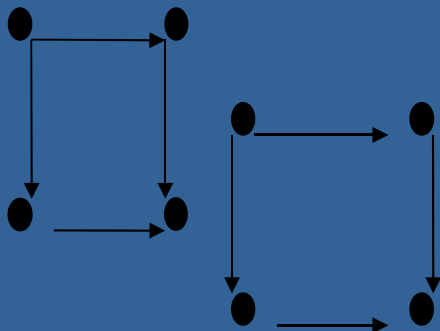
- 所求的图形是以题目所给的 N 个点为顶点的多边形。每条边要和 X 轴或 Y 轴平行。每个顶点与一条平行于 X 轴和一条平行于 Y 轴的线段相连。



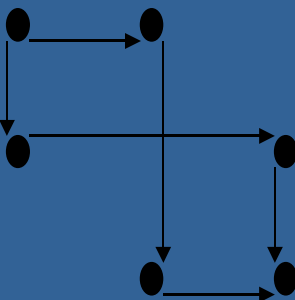


- 将所有点排序后发现，在同一水平线的点中，设为 $P_1, P_2, P_3, P_4, \dots, P_n$ ， P_1 必须连它右边的点—— P_2 ， P_3 只能连 P_4 ， P_5 连 P_6, \dots ，同垂直线上的点也是如此。
- 如果有解，解是唯一的，那么最小长度的问题就解决了。

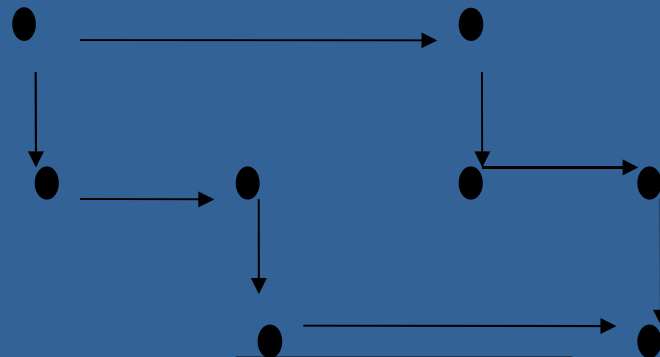
【问题分析】



不相连—不合法



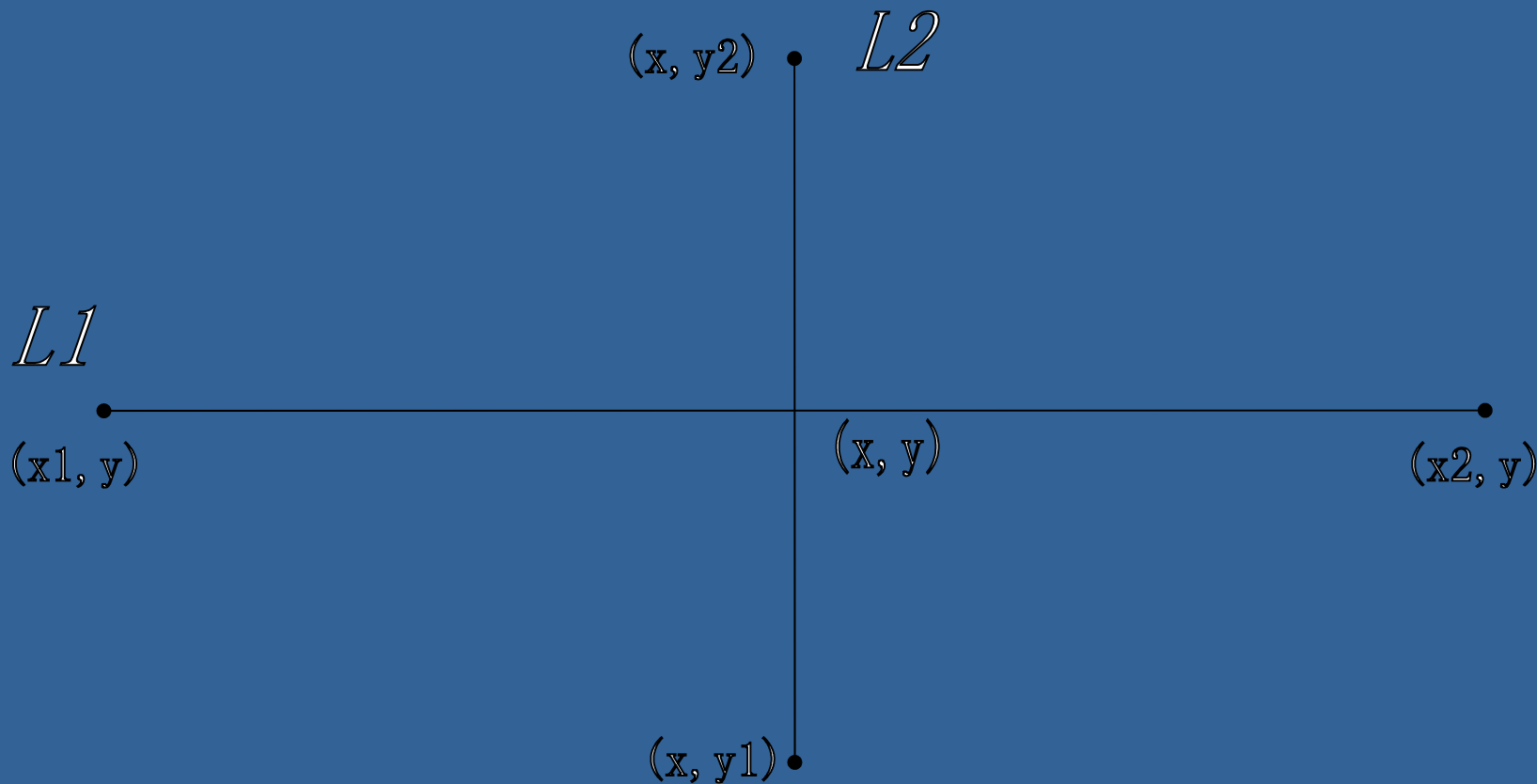
自交—不合法



不自交—合法

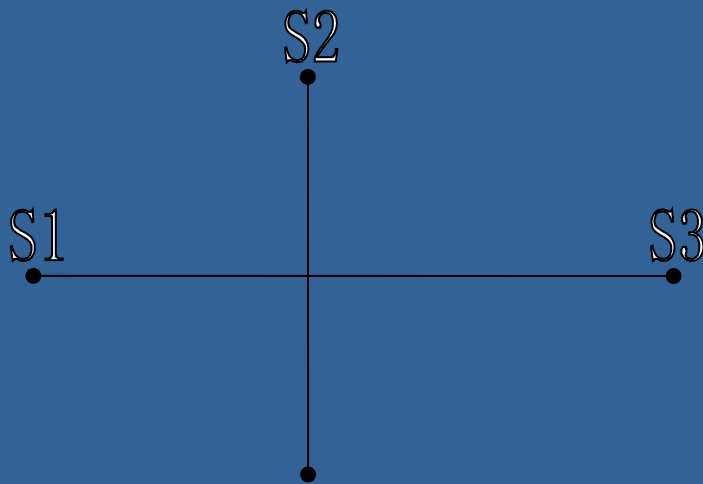
- 由于解是唯一，所以关键在于判断由上述方法构造出的图形是否合法——满足线段不自交：

- 如图，两条线段在内部相交，则必须满足 $x1 < x < x2$ 和 $y1 < y < y2$ 。

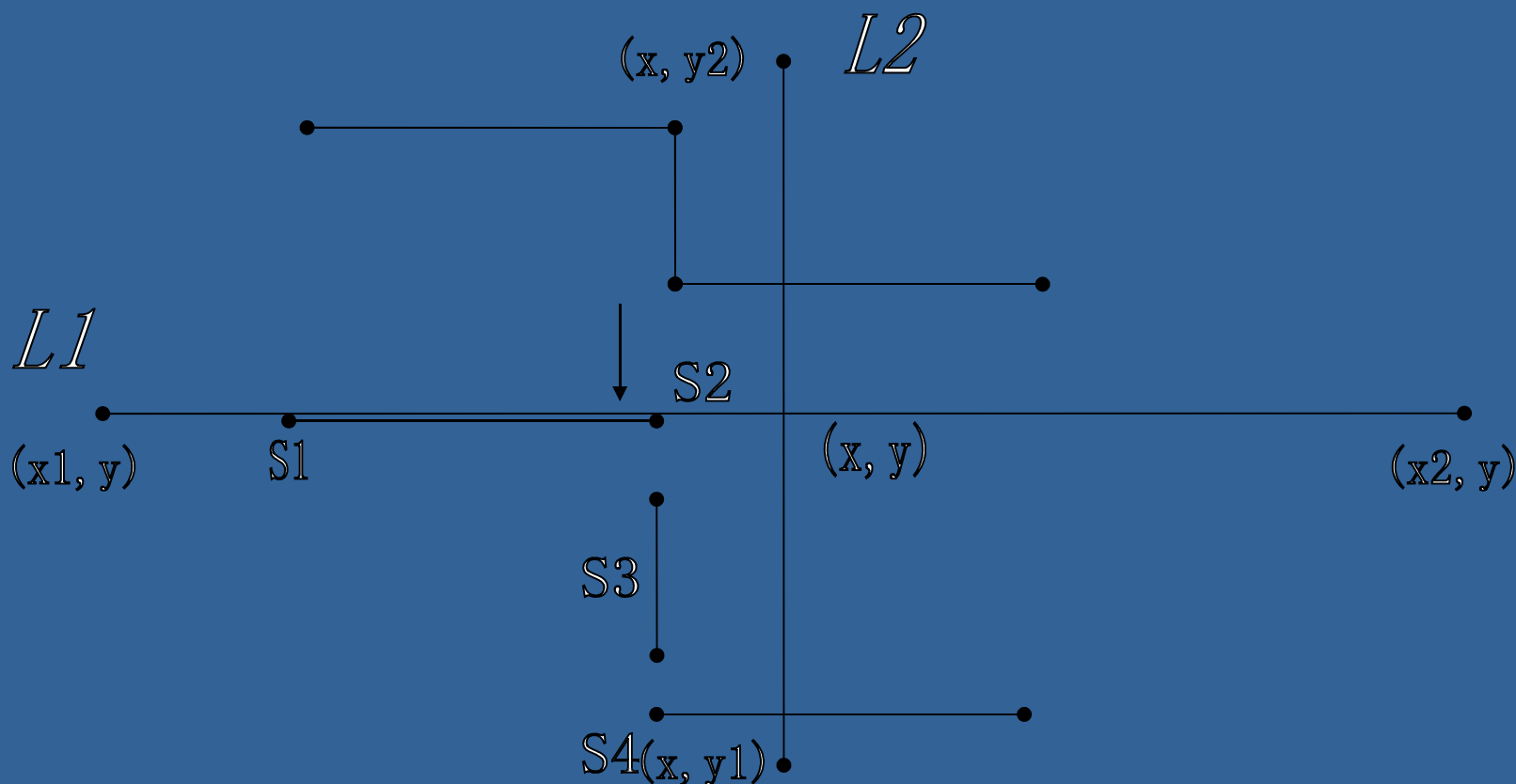


【问题解法】

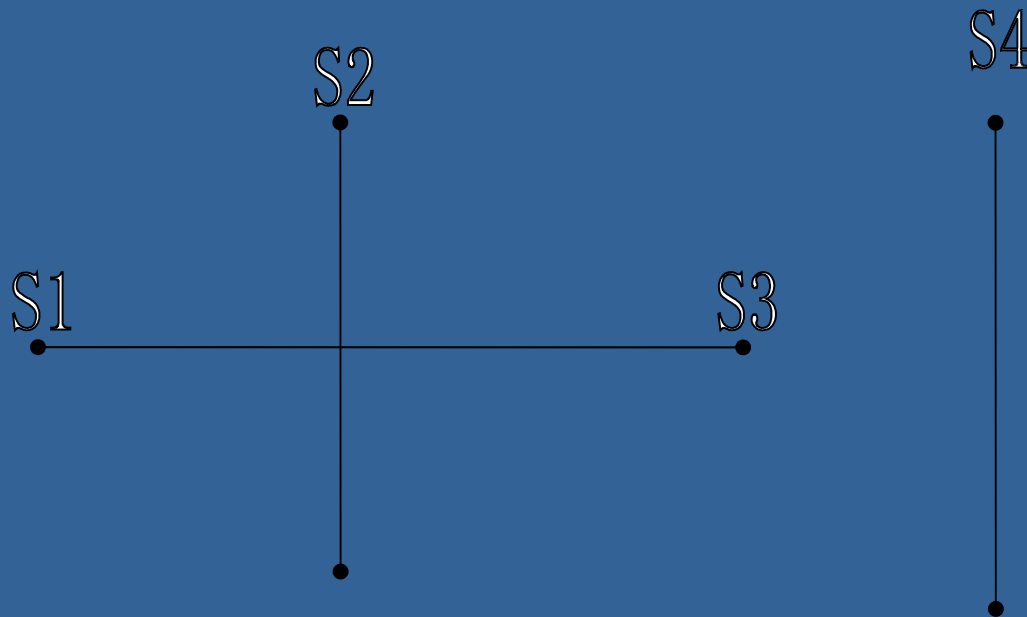
- 由于这题的区间性很明显，可以用线段树来解决。
- 将所有线段按 X 轴坐标排序，每条平行于 Y 轴的线段和每个平行于 X 轴线段的端点称为一个事件



- 本题要注意的是，线段在端点重合不算自交，所以 x 轴坐标相同时，事件的顺序要恰当处理。



- 将 Y 轴代表的区间建成线段树，并且每个节点记录它的区间内插入的点数。按顺序，扫描所有事件：如果遇到平行于 X 轴线段的左端点，则插入到线段树中，遇到右端点，则从线段树中删除。如果遇到平行于 Y 轴的线段，则在线段树中查找。



- 将 Y 轴坐标离散后，每次插入、删除、查找的复杂度是 $O(\log n)$ 级别的，由于所有线段数量是 $O(n)$ 级别的，所以整题的复杂度是 $O(n \log n)$ 级别。
- 思考：本题删除的点与插入的点一一对应，所以删除实现很方便。如果删除的线段不一定插入过该怎么办？

例 2：空心长方体（ POI99 Cuboid ）

- 在一个三维正坐标系中，存在 N ($N \leq 5000$) 个点，现在要求一点 $P(x, y, z)$ ，使得 $O(0, 0, 0)$ 与 $P(x, y, z)$ 两个顶点构成的长方体内不包括 N 个点中的任何一个点（在长方体边缘不算包括），并使这个长方体的体积最大。 x, y, z 均不得超过 1000000。

【问题分析】

- $P(x, y, z)$ 代表的长方体包含一点 Q ，那么 P 的所有坐标值，都大于 Q 点的坐标值。
- $P_x > Q_x \quad P_y > Q_y \quad P_z > Q_z$
- 体积最大的长方体，其 P 点任意轴的坐标，都与 N 个点中的一个相同或者和边界相同。

【问题分析】

- 在已经确定 P 的 X 坐标情况下，将所有点的 y 坐标排序，得到序列 $y[1], y[2], y[3], \dots$ 。 $\max[i]$ 记录 P 的 Y 轴坐标为 $y[i]$ 时， Z 轴坐标的最大取值，数组 \max 的值是单调不增的。

【问题分析】

- 把所到点按坐标值排序，当这点的坐标要第断维按数组 \max 第因及第的以后的点都第可能被覆盖第 $i+1$ 个点的，第 i 个点的坐标就会增加对数组 \max 的限制。
- 考虑第 i 个点 $Q(x, y, z)$ 增加对数组的限制

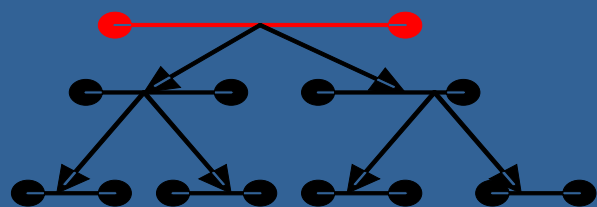
把 \max 看成一个区间



- 可以看出每增加一个点的限制，需要修改的是一个区间，要高效的进行区间操作就可以用到线段树。

【问题解法】

- 建立关于数组 \max 的线段树，每增加一个节点的限制，就相当于修改线段树中的一个区间内的值



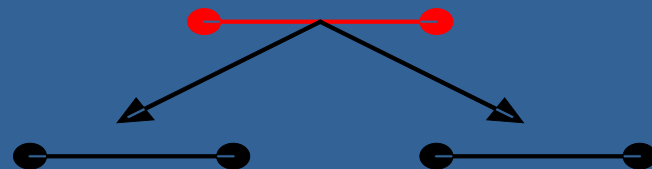
要修改的子树

收缩成一个
叶子节点



一个叶子节点

释放出两个儿子



- 修改的区间如果包含我们那个点，那么我们就把这个点表示的区间，那个点释放出来，然后那个点就变成叶子节点。修改的区间如果包含我们那个点，那么我们就把这个点表示的区间，那个点释放出来，然后那个点就变成叶子节点。

【问题解法】

- 线段树中，每个节点要记录其表示的区间内 $\max[i] * y[i]$ 的最大值，并且在数组被修改时，维护这个最大值。根节点的最大值与 P 点当前 X 坐标相乘，就是当前最大体积。
- 由于利用子树收缩的方法，离散后，每次修改的节点数是 $O(\log n)$ 级别，相关的维护也是这个级别，所以本题的复杂度为 $O(n \log n)$ 。

- 总结这题的成功之处：不把线段树死板的按看成固定的，而是抓住线段树叶子节点的本质——区间内的各种数据是单一，根据情况把子树收缩为叶子或让叶子释放出儿子，从而避免重复的操作。

例 3：战场统计系统

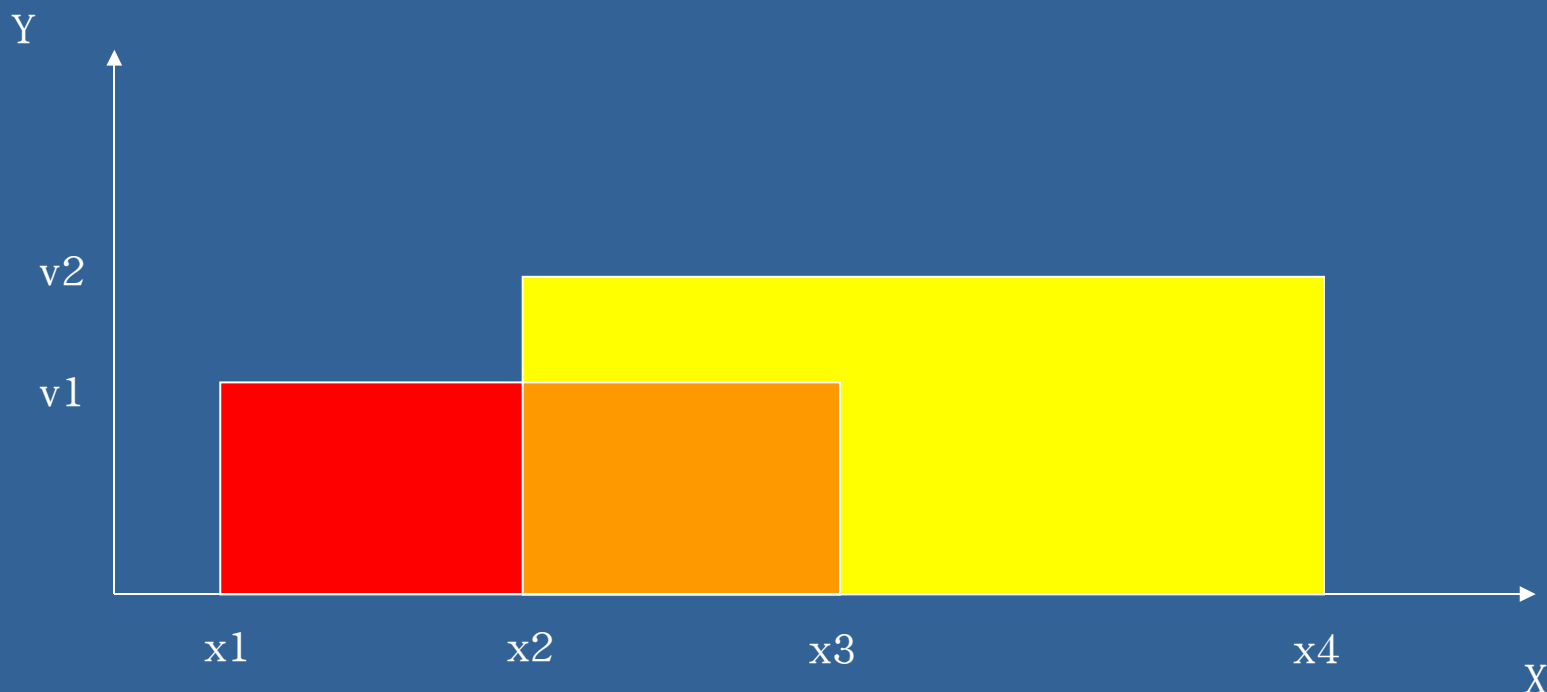
【问题描述】

人类与外星人之间的战争已趋于白热化请你尽快设计出这样一套系统。这套系统需要具备能够处理如下 2 类信息的能力：

- 外星人向 $[x1, x2]$ 内的每个位置增援一支防御力为 v 的部队。
- 人类使用超级武器对 $[x1, x2]$ 内的所有位置进行一次攻击力为 v 的打击。系统需要返回在这次攻击中被消灭的外星人个数。
- 注：防御力为 i 的外星人部队由 i 个外星人组成，其中第 j 个外星人的防御力为 j 。数据范围：
 $1 \leq x1 \leq x2 \leq 1000, v \leq 1000$ ；信息总数 $m \leq 2000$ 。

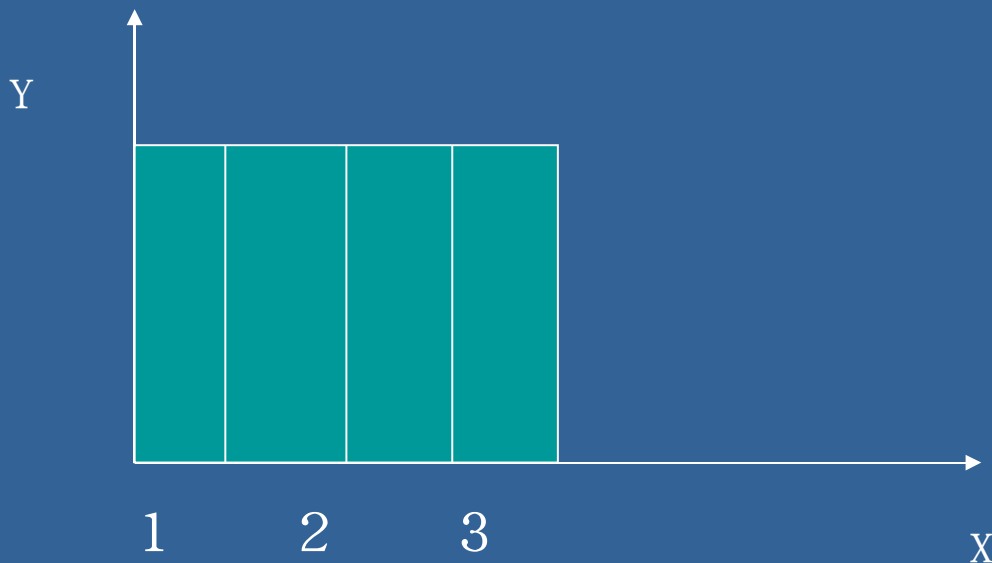
【问题分析】

- 把两类信息抽象化： $T[x, y]$ 表示单位区间 $[x, x]$ 上防御力为 y 的部队数目，如果在 $[x1, x2]$ 增加防御力为 v 的部队，就是在这个二维数组内，添加一个 $[x1, x2][1, v]$ 的矩形；要使用一次武器就相当于查找一个矩形 $[x1, x2][1, v]$ 内的部队数目，并将此矩形内的所有部队删除。



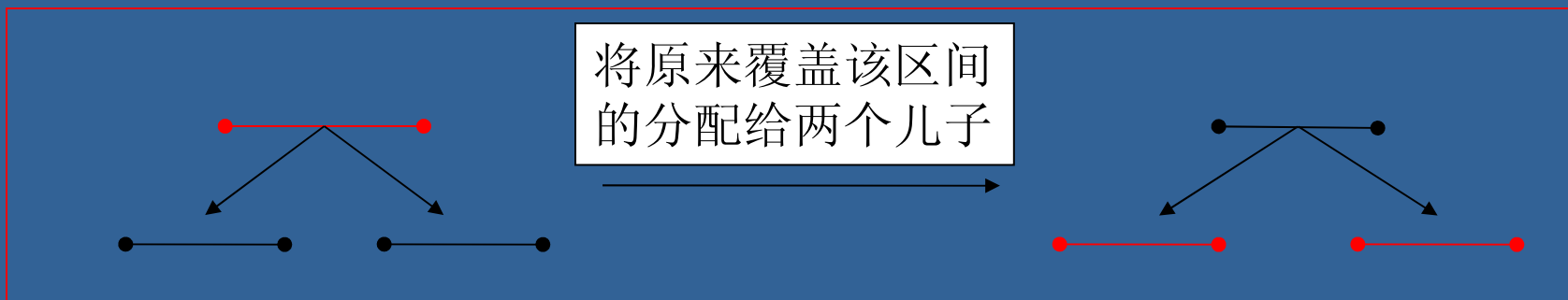
【问题分析】

- 如果采用最简单的方法：对矩形内的每个点进行操作，复杂度高达 $O(m*n*v)$ 。
- 首先可以考虑在一维上优化，在每个单位区间 $[x, x]$ 上，分别用一棵线段树记录各种防御力的部队数。

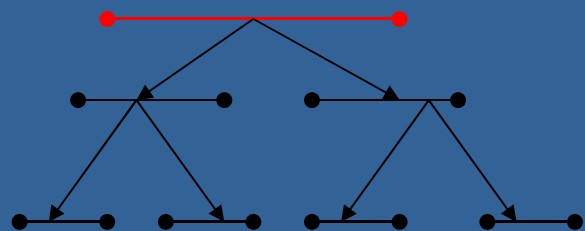


【问题解法】

- 插入一个矩形 $[x1, x2][1, v]$ ，就变成了执行 $x2 - x1 + 1$ 次在线段树上插入线段 $[1, v]$ 的操作，这和一般插入是一样的。
- 统计一个矩形内的部队数目并删除该矩形 $[x1, x2][1, v]$ 内的部队，也是要在每个单位区间执行一次操作。如果 v 不被 $[1, v]$ 包含，则要把本来覆盖它区间的分配给两个儿子。



- 而如果要删除 V 区间，那么可以用前面的收缩子树的方法，如果在插入时遇到叶子节点同样运用前面的方法释放出两个儿子节点。



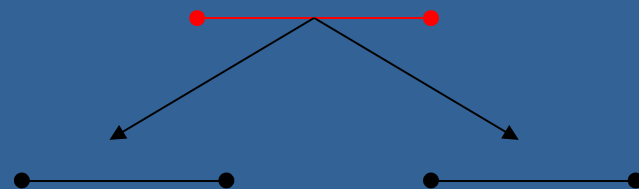
要删除的子树

收缩成一个
叶子节点



一个叶子节点

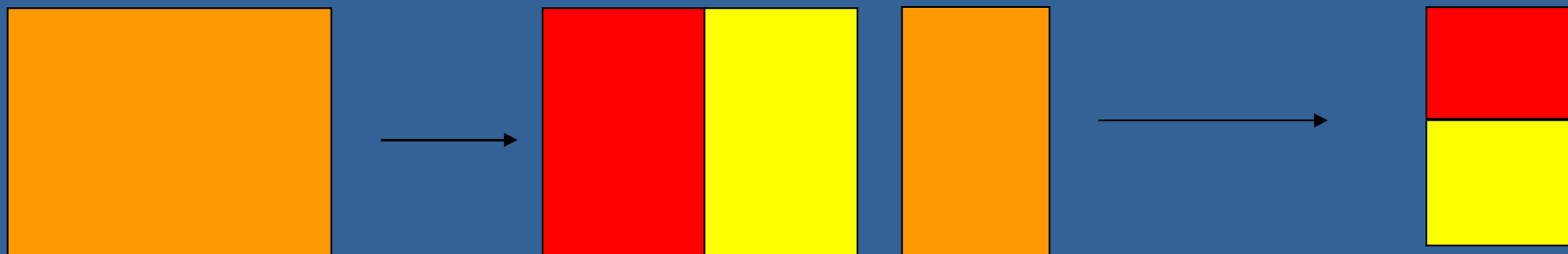
释放出两个儿子



- 优化后每次删除、统计、插入的复杂度都是 $O(n \cdot \log v)$ ，所以总的复杂度降为 $O(m \cdot n \cdot \log v)$ ；

【问题分析】

- 由于这是个二维的区间，不妨将整个二维的区间建成类似线段树的面积树：
- 假设整个二维区间是 $[1..2^k][1..2^k]$ ，递归定义一棵树 T ：
 - 1 T 的根节点为整个二维区间。
 - 2 区间 $[x, x][y, y]$ 为叶子节点
 - 3 儿子的划分：



- 建立面积树后，任意一个矩形在树中，都被划分为不超过 $O(2^k)$ 块区域，仿造线段树的方法进行插入、查找、和删除，就可以把本题复杂度降为 $O(m^*(n+v))$

总结用线段树解题的方法

- 根据题目要求将一个区间建成线段树，一般的题目都需要对坐标离散。建树时，不要拘泥于线段树这个名字而只将线段建树，只要是表示区间，而且区间是由单位元素（可以是一个点、线段、或数组中一个值）组成的，都可以建线段树；不要拘泥于一维，根据题目要求可以建立面积树、体积树等等。

- 树的每个节点根据题目所需，设置变量记录要求的值。
- 用树形结构来维护这些变量：如果是求总数，则是左右儿子总数之和加上本节点的总数，如果要求最值，则是左右儿子的最大值再联系本区间。利用每次插入、删除时，都只对 $O(\log L)$ 个节点修改这个特点，在 $O(\log L)$ 的时间内维护修改后相关节点的变量。

- 在非规则删除操作和大规模修改数据操作中，要灵活的运用子树的收缩与叶子节点的释放，避免重复操作。

谢谢

