# Data Structures



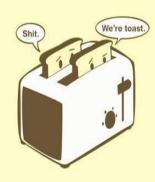
By H10



• 小朋友们大家好,还记得我是谁吗?



- 小朋友们大家好,还记得我是谁吗?
- 好的,今天我们来讲一讲 OI 中的重要考点、难点、<del>易错点</del>:数据结构。



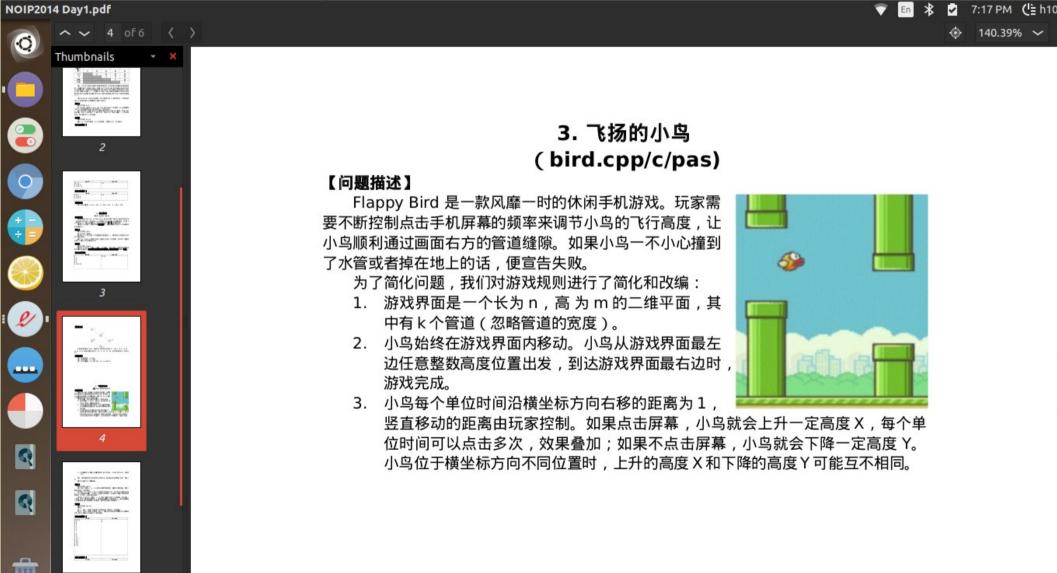
• 然而发现 NOIP 不怎么喜欢考数据结构啊呵呵。



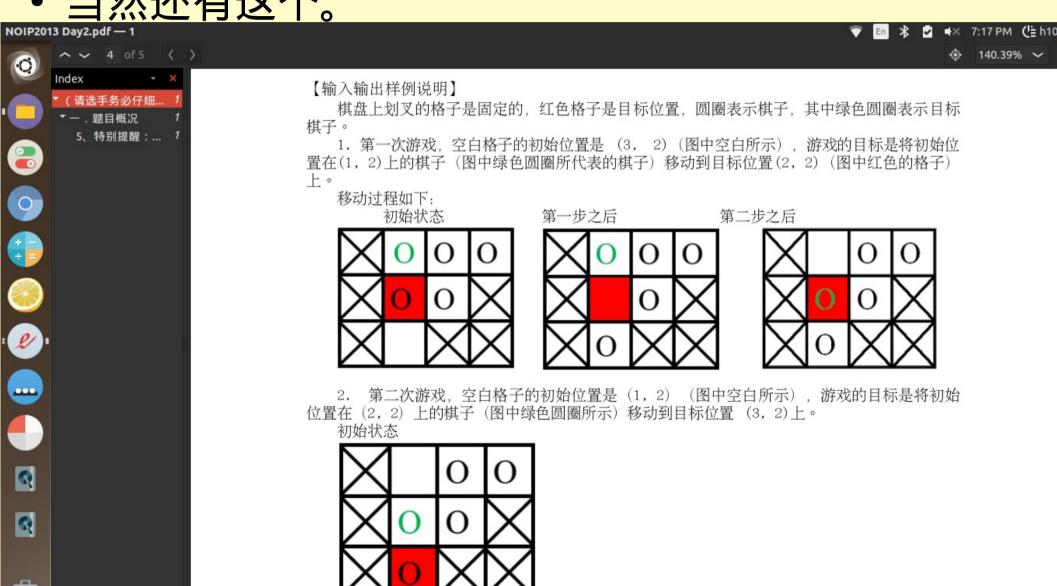
• 更准确的说 NOIP 喜欢考这个。



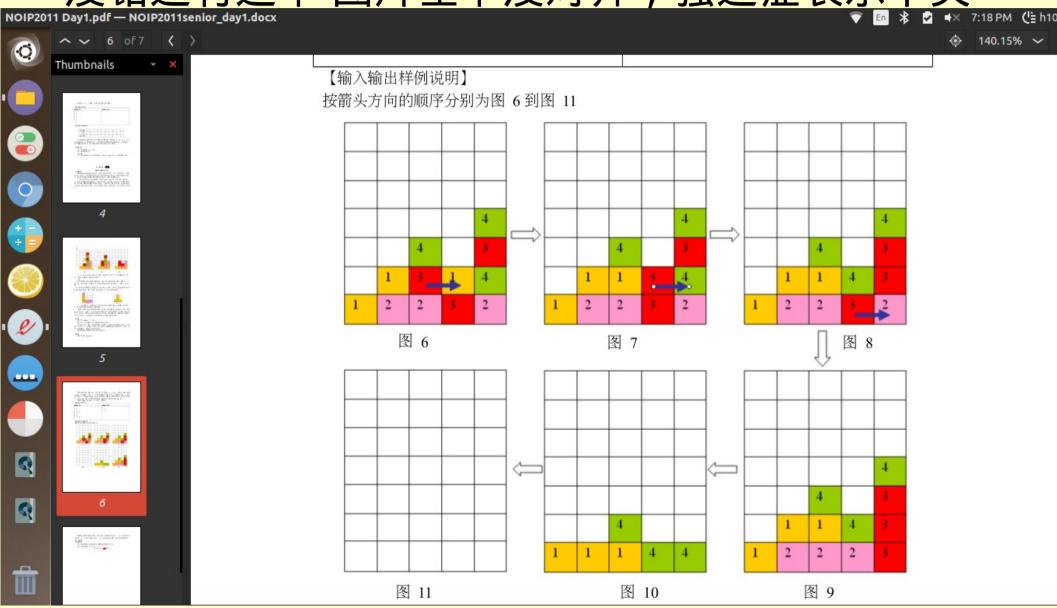
• 也喜欢考这个。



当然还有这个。



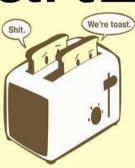
• 没错还有这个 图片左下没对齐,强迫症表示不爽



- · 总而言之 NOIP 最喜欢大型工业项目啊。
- 数据结构什么的好像被丢到后宫里去了。



# Part1 树状数组





• 定义实现什么的大家可自行背诵白书相关内容。



- 定义实现什么的大家可自行背诵白书相关内容。
- 可以支持单点修改啊,区间求和啊什么的,都是 O(n logn)。



- 定义实现什么的大家可自行背诵白书相关内容。
- 可以支持单点修改啊,区间求和啊什么的,都是 O(n logn)。
- · 当然不怕麻烦的适你也可以在树状数组里实现 lazy-tag , 然后就可以区间修改了。

- 定义实现什么的大家可自行背诵白书相关内容。
- 可以支持单点修改啊,区间求和啊什么的,都是 O(n logn)。
- 当然不怕麻烦的适你也可以在树状数组里实现 lazy-tag ,然后就可以区间修改了。
- 基本上树状数组的功能线段树都能实现。

• 那么还用它干嘛呢?



- 那么还用它干嘛呢?
- 因为有一些不良心的出题人会出只能用它做的题。



- 比如这道题 Bzoj 1452
- 一个 n\*m 的方格,初始时每个格子有一个整数 权值,接下来每次有2种操作:
- 1) 改变一个格子的权值;
- 2) 求一个子矩阵中某个特定权值出现次数。
- n,m<=300 , Q<=200000 , 1<= 权值 <=100
- Time Limit: 1s
- Memory Limit : 64MB
  Data Structures

• 考虑用线段树,空间会是 100\*4n\*4m (549MB)



- 考虑用线段树,空间会是 100\*4n\*4m (549MB)
- 用树状数组,空间 100\*n\*m (34MB)



- 考虑用线段树,空间会是 100\*4n\*4m (549MB)
- 用树状数组,空间 100\*n\*m (34MB)
- 卡空间什么的最讨厌了。



• 然后树状数组就讲完了。



- 然后树状数组就讲完了。
- 什么,嫌讲的太少?



- 然后树状数组就讲完了。
- 什么,嫌讲的太少?
- 那么就多加道题好了。



- Poj 2481
- 给你n个闭区间,问每一个区间被多少个区间完全包含(完全重合不算)。

n<=100000</li>

• Time Limit :18

Memory Limit: 256MB

• 按 Li 排个序, 把 Ri 加入树状数组里面维护一下就好了对吧。经典的二维数点。



- 按 Li 排个序, 把 Ri 加入树状数组里面维护一下就好了对吧。经典的二维数点。
- 这题就是来凑个数的。



• 那么树状数组到这里就结束了吧。



- 那么树状数组到这里就结束了吧。
- 才怪咧 .....



- 那么树状数组到这里就结束了吧。
- 才怪咧 .....
- 我怎么可能只讲两道水题啊。



- 那么树状数组到这里就结束了吧。
- 才怪咧 .....
- 我怎么可能只讲两道水题啊。
- 肯定要来一道非常 sang) 常 (xin) 有 (bing) 趣 (kuang) 的题啊。

Bzoj 4240

给你一个数列,相邻两个数之间可以交换,要求 使每一个数要么左侧没有比它大的数,要么右侧 没有比它大的数,求最少交换次数。

• n<=300000

• Time Limit : 1s

Memory Limit: 256MB

• 考虑对于每一个数,它的移动是不会影响到比它更大的数的。



• 考虑对于每一个数,它的移动是不会影响到比它更大的数的。

 因此我们可以从小到大移动数,并且贪心地移动 到需要移动次数较少的一边。

- 考虑对于每一个数,它的移动是不会影响到比它更大的数的。
- 因此我们可以从小到大移动数,并且贪心地移动 到需要移动次数较少的一边。
- 于是每一个数对答案的贡献就是 min(左侧比它 大的数的个数,右侧比它大的数的个数)。

- 考虑对于每一个数,它的移动是不会影响到比它 更大的数的。
- 因此我们可以从小到大移动数,并且贪心地移动 到需要移动次数较少的一边。
- 于是每一个数对答案的贡献就是 min(左侧比它 大的数的个数,右侧比它大的数的个数)。
- 那么可以从大到小添入节点然后用树状数组<del>或线</del> <del>段树</del>维护即可。

• 总结:

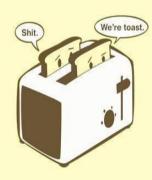


• 总结:

• 树状数组其实不算难,毕竟它支持的操作不多。



- 总结:
- 树状数组其实不算难,毕竟它支持的操作不多。
- 单单只考树状数组的题基本是水题(第一题);



- 总结:
- 树状数组其实不算难,毕竟它支持的操作不多。
- 单单只考树状数组的题基本是水题(第一题);
- 树状数组也可以解决一些经典问题(第二题);

- 总结:
- 树状数组其实不算难,毕竟它支持的操作不多。
- 单单只考树状数组的题基本是水题(第一题);
- 树状数组也可以解决一些经典问题(第二题);
- OI 中数据结构常常会与其它考点放一起考,但 多半情况下数据结构是用于辅助(如第三题的贪心+树状数组<del>或线段树</del>)。

# Part2 RMQ





• 用 O(n logn) 时间预处理,以后每次查询区间极值都是 O(1)。



- 用 O(n logn) 时间预处理,以后每次查询区间极值都是 O(1)。
- 可以求区间最大最小值,还可求 LCA。



- 用 O(n logn) 时间预处理,以后每次查询区间极值都是 O(1)。
- 可以求区间最大最小值,还可求 LCA。
- 用处基本就这么多了,还是看看在题目中怎么灵(xia)活(ji)地(ba)使用吧。

- Poj 3264
- 给你一个序列,每次询问第L个至第R个数之间的极差(最大值-最小值)



- Poj 3264
- 给你一个序列,每次询问第L个至第R个数之间的极差(最大值-最小值)

· 你 TM 放模板题是在鄙视我的智商吗?

• 抱歉刚刚搞错题了啊,应该是这道。



- Poj 1785
- 给你N个结点建立一颗树,每个结点有两个关键字,要求第一个满足BST性质,第二个满足堆的性质。
- N<=50000
- Time Limit
- Memory Limit: 32MB

• <del>不就是 treap</del> 啊



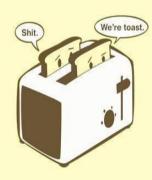
- <del>不就是 treap 啊</del>
- 既然讲 RMQ , 就用 RMQ 做吧。



• 其实也就是道傻逼题了。



- 其实也就是道傻逼题了。
- Treap 做法:把点按第一关键字排序,按顺序插入右子树,不满足堆的性质就左旋。

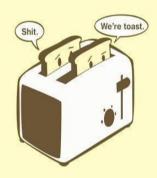


- 其实也就是道傻逼题了。
- Treap 做法:把点按第一关键字排序,按顺序插入右子树,不满足堆的性质就左旋。
- RMQ 做法:还是首先要排序,每次递归处理一个区间,找出第二关键字最大的做根。其左边的就是左子树,右边的是右子树。

• 题目难度够良心了吧。



- 题目难度够良心了吧。
- 下面来一道有意思点的题。



- Nkoj 1752
- 给你一个有序数列,对于每一个询问 Q(i,j),求出 [i,j] 中出现次数最多的数的出现次数。
- N<=200000, Q<=200000 , Ai<=1000000
- Time Limit
- Memory Limit: 256MB

• 我会莫队



- 我会莫队
- 既然讲 RMQ , 就用 RMQ 做吧。(怎么感觉这句话讲过)



- 举个例子
- 序列 -1 -1 1 1 1 1 3 10 10 10
- 可以转换为 2 4 1 3 ,利用前缀和二分查找位置,左右两介端点暴力,中间 RMQ。

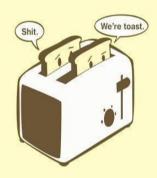
- 举个例子
- 序列 -1 -1 1 1 1 1 3 10 10 10
- 可以转换为 2 4 1 3 , 利用前缀和二分查找位置, 左右两介端点暴力, 中间 RMQ。
- · <del>怎么尽是些</del>水题啊



• 下面来讲黑科技



- 下面来讲黑科技
- 只用 O(n) 预处理就能 O(1) 查询的 ±1 RMQ



- 下面来讲黑科技
- 只用 O(n) 预处理就能 O(1) 查询的 ±1 RMQ
- <del>又不是 NOIP 考点</del>



- 下面来讲黑科技
- 只用 O(n) 预处理就能 O(1) 查询的 ±1 RMQ
- <del>又不是</del> NOIP 考点
- · 常数都能比 log(n) 大

 先分块,每一块的大小设为 sz=log(n)/2,则一 共有 l=n/sz 块



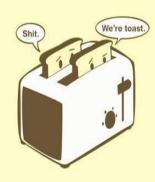
 先分块,每一块的大小设为 sz=log(n)/2,则一 共有 l=n/sz 块

把每一块当成一个数字,做普通RMQ,时间为O(l\*log(l)) = O(n/log(n)\*(log(n)-log(sz))) = O(n)

- 先分块,每一块的大小设为 sz=log(n)/2,则一 共有 l=n/sz 块
- 把每一块当成一个数字,做普通RMQ,时间为O(l\*log(l)) = O(n/log(n)\*(log(n)-log(sz))) = O(n)
- · 连续的几块的查询就与普通 RMQ 一毛一样了

- 先分块,每一块的大小设为 sz=log(n)/2,则一 共有 l=n/sz 块
- 把每一块当成一个数字,做普通RMQ,时间为O(l\*log(l)) = O(n/log(n)\*(log(n)-log(sz))) = O(n)
- · 连续的几块的查询就与普通 RMQ 一毛一样了
- 重点是块内查询

考虑到块的大小为 log(n)/2 ,每一个数与上一个数的差只能是 1 或 -1 ,那么本质不同的块就只有 2^(log(n)/2) = sqrt(n) 个了。



- 考虑到块的大小为 log(n)/2 ,每一个数与上一个数的差只能是 1 或 -1 ,那么本质不同的块就只有 2^(log(n)/2) = sqrt(n) 个了。
- 对于一种种块,预处理出所有有可能的块内查询,都只有 20(log(n)/2) = sqrt(n) 种,所以块内查询的时间 为 0(sqrt(n)\*sqrt(n)) = O(n)。

• 讲完了,想来道题吗?



- 讲完了,想来道题吗?
- 好消息是我并没有找到专门用 ±1 RMQ 做的题。



• 总结:



- 总结:
- RMQ 其实也就是一个求极值的小工具。



• 总结:

• RMQ 其实也就是一个求极值的小工具。

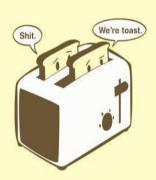
• 同树状数组一样,没有太多的变形,用起来十分

方便。

- 总结:
- RMQ 其实也就是一个求极值的小工具。
- 同树状数组一样,没有太多的变形,用起来十分方便。
- 但是部分 RMQ 题会伪装得很好,难以短时间内 看出来是要用 RMQ 来做,具体的例子会在练习 题中出现。(这里是在剧透哦)

# Part3 线段树





• 重要的内容总要用于最后的压轴对不对。



• 重要的内容总要用于最后的压轴对不对。

线段树, OI 中最常用的数据结构之一,它能支持单点修改,区间修改,区间查询,时间复杂度都是 logn。

• 然而线段树最有意思的一点是:



- 然而线段树最有意思的一点是:
- 凡是树状数组与 RMQ 能做的,线段树都能做;
   凡是支持合并的元素,线段树都能维护。



- 然而线段树最有意思的一点是:
- 凡是树状数组与 RMQ 能做的,线段树都能做;
   凡是支持合并的元素,线段树都能维护。
- 凡是毛主席做出的决策,我们都坚决维护;凡是毛主席的指示,我们都矢志不渝地遵循。

• 第一个凡是不必多说。



• 第一个凡是不必多说。

 关与第二个凡是,请参考 NOI2016 湖南省省队 集训 Day9T2 闷声刷大题暨雅礼中学 2016 湖南 省省选集训 Day2T2 光盘。

- 第一个凡是不必多说。
- 关与第二个凡是,请参考 NOI2016 湖南省省队 集训 Day9T2 闷声刷大题暨雅礼中学 2016 湖南 省省选集训 Day2T2 光盘。
- "光盘"的出题人曾评论过"闷声刷大题":出这道题的人心态很不好啊,因为我出"光盘"时心态就不好。

• 废话就讲到这里吧,下面上题。



- Poj 3667
- 旅馆里有一列连续的房间,要求支持两种操作:
- 1 a : 来了 a 名客人, 询问是不是有连续长度为 a 的空房间。有的话住进最左边;
- 2 l r : 将 [I, 下的房间清空。
- 1 <= n,m <= 50000
- Time Limit: 1s
- Memory Limit: 256MB

• 热身傻逼题 ...



- 热身傻逼题 ...
- 线段树记录区间中最长的空房间;



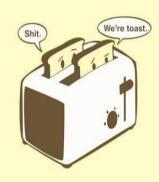
- 热身傻逼题 ...
- 线段树记录区间中最长的空房间;
- 查询一个区间时,先看左子树是否合法,不合法 就再看左子树靠在+右子树靠左是否合法,还不 合法就看右子树。

• 下一题



- Bzoj 2957
- 有 n 栋楼房在一个二维平面上。第 i 栋楼房可以用线段 (i,0)-(i,Hi) 表示。最开始楼房的高度全为 0 , m 次操作 , 每次把第 xi 栋楼房的高度修改为 yi ,并询问在 (0,0) 处可见楼房数量。
- 1 <= xi,n,m <= 100000 , 1 <= yi <= 10^9
- Time Limit: 1s
- Memory Limit: 256MB

显而易见,可见的数字是单调递增的。修改一个数只会对后面的数造成影响。那么考虑线段树中的每一条线段。只有两种情况:



显而易见,可见的数字是单调递增的。修改一个数只会对后面的数造成影响。那么考虑线段树中的每一条线段。只有两种情况:

• 1、最大值小于等于修改的数,那么这个线段的贡献为 0,无需处理;

- 显而易见,可见的数字是单调递增的。修改一个数只会对后面的数造成影响。那么考虑线段树中的每一条线段。只有两种情况:
- 1、最大值小于等于修改的数,那么这个线段的贡献为 0,无需处理;
- 2、否则将这个线段分成两段,如果左侧的最大值大于修改的数,那么不影响右侧贡献,递归处理左侧;否则就变成了第一种情况,递归左侧。

• 下一题

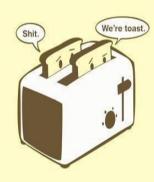


- 某学校某年某日某比赛的某模拟题
- 如果字符串 A 是字符串 B 的后缀,那么称 B 是 A 的 XQ 串。给你 n 个只包含小写字母的字符串,编号为 1-n。对于每个串,求其所有的 XQ 串的编号集合《包括自己》中第 Ki 小的编号。
- Time Limit: 1s
- Memory Limit: 256MB

• 怎么开始讲字符串题了啊?



- 怎么开始讲字符串题了啊?
- 把所有的串反过来建个 trie ,那么某个串的结束 结点的子树里的串就是所有该串的 XQ 串。



• 怎么开始讲字符串题了啊?

 把所有的串反过来建个 trie , 那么某个串的结束 结点的子树里的串就是所有该串的 XQ 串。

· 以 dfs 序排序,问题就变成求区间第 K 小了。

- 怎么开始讲字符串题了啊?
- 把所有的串反过来建个 trie , 那么某个串的结束 结点的子树里的串就是所有该串的 XQ 串。
- · 以 dfs 序排序,问题就变成求区间第 K 小了。
- 用主席树维护维护即可。

怎么开始讲字符串题了啊?

把所有的串反过来建个 trie , 那么某个串的结束 结点的子树里的串就是所有该串的 XQ 串。

以dfs序排序。通知或就是间第K小了。用主席树维护即可。

怎么开始讲字符串题了啊?

把所有的串反过来建个 trie , 那么某个串的结束 结点的子树里的串就是所有该串的 XQ 串。

以dfs序排序,远题就变成求区间第K小了。

用主席树

· 不是说好了的 NOIP 考点么?

怎么开始讲字符串题了啊?

把所有的串反过来建个 trie , 那么某个串的结束 结点的子树里的串就是所有该串的 XQ 串。

以dfs序排序,短题就变成求区间第K小了。

用主席树

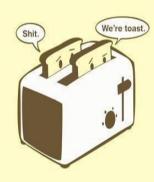
· 不是说好了的 NOIP 考点么?

• 虚假宣传啊!

• 来个简单点的方法吧。



- 来个简单点的方法吧。
- 由于可以离线,区间第 K 小可以直接在 trie 上自 底向上查询。用线段树合并维护结点信息即可。



来个简单点的方法吧。

由于可以离线,区间第K小可以直接在 trie 上自底向上查询。用线段树合并维护结点信息即可。



来个简单点的方法吧。

由于可以离线,区间第K小可以直接在 trie 上自底向上查询。用线段树合并维护结点信息即可。

• 你确定不是。O(n^2) ?

• 确定, 你们可以先看看线段树合并的代码。



- 确定, 你们可以先看看线段树合并的代码。
- int merge(int x,int y){
- if (!x) return y;
- if (!y) return Xiver toast
- Is[x]=merge(ls[x],ls[y]);
- rs[x]=merge(rs[x],rs[y]);
- return x;
- }

• 可以看出线段树合并的复杂度与两颗树重复的部分大小成正比。



可以看出线段树合并的复杂度与两颗树重复的部分大小成正比。

每个结点记录的信息只是这个结点的子树里所有的串,当我们合并两个结点时,它们包含的串是没有交集的。

可以看出线段树合并的复杂度与两颗树重复的部分大小成正比。

每个结点记录的信息只是这个结点的子树里所有的串,当我们合并两个结点时,它们包含的串是没有交集的。

• 由于总共只有 n 个串,时间就是 O(n logn) 了。

• 终于讲完了。



- 终于讲完了。
- 额,好像还没有总结。



• 总结:



• 总结:

• 比起前两种数据结构,线段树有更强的灵活性, 这也就是为什么 OI 常考线段树的主要原因。

(自主) (A) Shit. We're toast

• 总结:

• 比起前两种数据结构,线段树有更强的灵活性, 这也就是为什么 OI 常考线段树的主要原因。 <del>(句式杂糅)</del>

We're toast.

• 正是由于其灵活性,所以能熟练地使用线段树也是作为一名 Cler必须要 get 的技能。 (介词掩盖主语)



Poj 2482

在一个平面内有 n 个整点,每个点都有一个权值,用一个 w\*h 的矩形去围这些点(边上的不算),求能得到的最大权值和。

• n <= 100000 , w,h,|xi|,|yi| <= 2^31

• Time Limit :1s

Memory Limit: 256MB

Poj 1981

• 在一个平面内有 n 个点, 求用一个单位圆最多能 围住多少点。

Memory Limit: 256MB

- 某 Acm 趣题
- 定义完美序列:连续的互不相同的序列,给你一个序列,对于每次询问,回答区间[L,R]之间最长的完美序列的长度。
- n,m <= 200000 ai <= 1000000
- Time Limit : 1
- Memory Limit: 256MB

Bzoj 3339

给你一个序列,对于每一次询问Q(I,r),请回答mex{a[I],a[I+1],...,a[r-1],a[r]}的值。

• n,m <= 200000 ai <= 1000000

• Time Limit

Memory Limit: 256MB

# 谢谢大家



