

# 2021 香农先修班第三次课

---

## 排序算法

# 排序理论

**排序算法** (Sorting algorithm) , 是一种将一组特定的数据按某种顺序进行排列的算法。

例如, 下面的操作是排序:

- 给定若干个实数, 从小到大排列输出
- 给定若干个字符串, 按字典序排序输出
- 给定若干个学生各科分数, 先按总分, 同分依次按每科分数高低排序直到可以区分为止 `OJ`  
`1072` 排名

**稳定性** 指作为比较依据的值相同的多个元素的相对位置在排序前后是否不变。不变的是**稳定排序**, 变的是**不稳定排序**。

例如, 有五个二元数据  $(x, y)$ :  $(1, 10), (2, 5), (5, 10), (4, 12), (3, 10)$ , 以  $y$  为依据排序, 若排序后有:

$(2, 5), (1, 10), (5, 10), (3, 10), (4, 12)$

则, 这样的排序是稳定排序; 若中间三个元素不是按这个顺序的, 是不稳定排序。

常见的稳定和不稳定排序有:

- 稳定排序: 冒泡排序、插入排序、计数排序、归并排序.....
- 不稳定排序: 选择排序、快速排序、堆排序.....

判断排序是否是稳定排序的方法: 若无不作为比较依据的值, 设辅助标记(如下标)形成结构体(若有不作为比较依据的值就直接用就行了), 进行结构体排序; 然后用任意稳定排序算法再排序一次, 逐项比较这两种排序的结果, 若完全一致就是稳定, 不一致就是不稳定。复杂度为两次排序复杂度 + 逐项比较复杂度。

有的排序需要额外开辟内存空间, 有的不需要(即直接使用待排序数组空间)。记排序的空间复杂度是除原数组外所需的额外空间大小。

- `in-place` 规模是  $n$  时, 空间复杂度是与  $n$  无关常数(是常函数), 又称原址操作。

如  $O(1), O(m)$

冒泡选择插入都是 `in-place`

- `out-place` 规模是  $n$  时, 空间复杂度是  $n$  的函数。

如  $O(n + k), O(n \log n)$

一个结论: 基于比较然后换位置的排序(如冒泡插入选择), 时间复杂度的下限是  $O(n \log n)$

常见排序算法表:

方法	平均	最坏	最好	空间	方式	稳定
冒泡	$O(n^2)$	$O(n^2)$	$O(n)$	$O(1)$	in	Y
选择	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(1)$	in	N
插入	$O(n^2)$	$O(n^2)$	$O(n)$	$O(1)$	in	Y
希尔	$O(n^{1.25})$	$O(n^2)$	$O(n)$	$O(1)$	in	N
归并	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$	$O(n)$	out	Y
快速	$O(n \log n)$	$O(n^2)$	$O(n \log n)$	$O(\log n)$	in	N
堆	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$	$O(1)$	in	N
计数	$O(n + k)$	$O(n + k)$	$O(n + k)$	$O(n + k)$	out	Y
桶	$O(n + k)$	$O(n^2)$	$O(n)$	$O(n + k)$	out	Y
基数	$O(nk)$	$O(nk)$	$O(nk)$	$O(n + k)$	out	Y

各经典排序动图过程演示可参考 [这里](#)

**逆序对**：对数组  $a$ ，满足  $i < j, a_i > a_j$  的两个数。逆序对数目(**逆序数**)能体现数组错乱程度。

一种求法：暴力枚举所有满足  $i < j$  的两数之对逐个判定即可。时间复杂度为  $O(n^2)$

实际上常用的求法在下文

性质：冒泡排序的交换次数为逆序数。

## 常用排序

### 计数排序

Counting Sort

#### 常规写法

排序思路：

- 设数组  $c$ ， $c[i]$  代表待排序数组  $a$  里  $i$  出现了多少次  
遍历  $a$ ，统计出现次数，即 `c[a[i]]++`。
- 记  $a$  当前下标为  $k$   
遍历  $c$ ，插入  $c[i]$  个  $i$  到数组  $a$  (即  $a$  下标为  $[k, k + c[i] - 1]$  的范围均赋值为  $i$ )，更新  $k$

参考代码：(设值域是  $[1, k]$  的整数)

```

1  for (int i = 1; i <= n; ++i)
2  {
3      ++c[a[i]]; // c 初始值均 0; a 数组使用下标范围 [1,n] 存值
4  }
5  for (int i = 1; i <= k; ++i) // k 是 a 的最大值
6  {
7      for (int j = 1; j <= c[i]; ++j)
8      {
9          a[++p] = i; // p 初始值为 0
10     }
11 }

```

设待排序值的值域是离散的，值域长度为  $k$ ，则时间复杂度为  $O(n + k)$ ，空间复杂度为  $O(k)$ 。

这是因为  $\sum_{i=1}^k c_i = n$ ，所以 `a[++p]=i` 语句必然只能被执行  $n$  次。

计数排序是非基于比较的排序，所以可以突破  $O(n \log n)$  的下限。

## 前缀和写法

排序思路：

- 设数组  $cnt$ ， $cnt[i]$  最初表示  $i$  出现了多少次，遍历  $a$  累积  $cnt$
- 对  $cnt$  原地求前缀和，使得  $cnt[i]$  表示小于等于  $i$  的数出现了多少次  
此时  $cnt[i] = \sum_{j=1}^i c[j]$ ，满足性质： $cnt[i]$  表示最后一个  $i$  排序后下标是  $cnt[i]$
- 新开数组  $b$  表示排序后数组

倒序遍历  $a$ ，根据性质，最后一个  $a[i]$  出现在下标  $cnt[a[i]]$ ，所以这个下标的值是  $a[i]$ ；将它取走，那么  $cnt[a[i]]$  自减，表示删掉当前最后一个  $a[i]$  后，新的  $a[i]$  的下标是自减后的  $cnt[a[i]]$

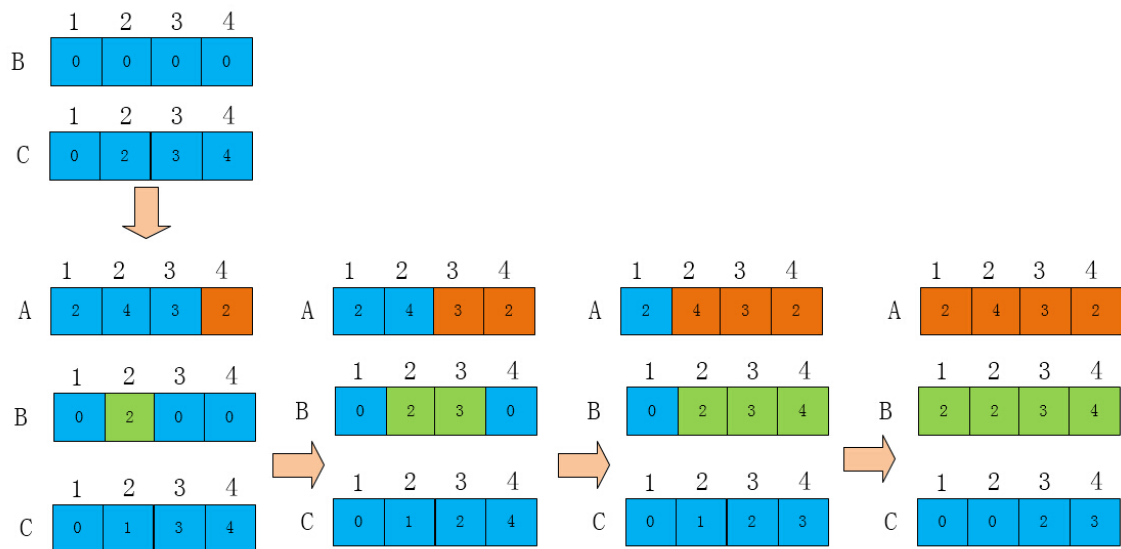
代码：

```

1  for (int i = 1; i <= n; ++i)
2  {
3      ++cnt[a[i]];
4  }
5  for (int i = 1; i <= k; ++i)
6  {
7      cnt[i] += cnt[i - 1];
8  }
9  for (int i = n; i >= 1; --i) //顺着i到n遍历也行
10 {
11     b[cnt[a[i]]--] = a[i];
12 }

```

图解演示：



该写法，时间复杂度为  $O(n + k + n) = O(n + k)$ ，空间复杂度为  $O(n + k)$

另一种写法是顺着来，即  $cnt[i]$  表示下一个  $i$  排序后下标是多少，代码为：

```

1  for (int i = 1; i <= n; ++i)
2  {
3      ++cnt[a[i]];
4  }
5  for (int i = 1, s = 0; i <= k; ++i)
6  {
7      s += cnt[i];
8      cnt[i] = s - cnt[i];
9  }
10 for (int i = 1; i <= n; ++i) //倒着也行
11 {
12     b[+cnt[a[i]]] = a[i];
13 }

```

## 离散化

标准定义：**离散化** (discretization) 是把无限空间中有限的个体映射到有限的空间中去。

一种离散化的实现：把绝对大小转化为相对大小。

当值域跨度很大，而真实不同的取值个数很少时，可以用离散化的方法把值域调整为取值个数。

如： $a = (1, 10^3, 10^9 + 0.1, 10^3, 1, 10^3)$ ，离散化后得  $a_h = (1, 2, 3, 2, 1, 2)$

离散化+计数排序的实现：借助 *map*。

定义  $h$ ， $h[i]$  表示  $i$  出现了多少次。

得到排序后原数组的实现代码：

```

1  for (ll i = 1; i <= n; ++i)
2  {
3      h[a[i]]++; //h是map<ll, ll>
4  }
5  for (auto i : h)
6  {
7      for (ll j = 0; j < i.second; ++j)
8      {
9          a[++ai] = i.first; //ai初始值为0
10     }
11 }

```

考虑到 `map` 自带的复杂度，时间复杂度为  $O(n \log n)$

也可以用 `unordered_map`，下同。

得到排序后离散化数组的实现代码：

```

1  for (ll i = 1; i <= n; ++i)
2  {
3      h[a[i]]++;
4  }
5  for (auto i : h)
6  {
7      ++r; //r初始值为0
8      for (ll j = 0; j < i.second; ++j)
9      {
10         a[++ai] = r;
11     }
12 }

```

时间复杂度为  $O(n \log n)$

直接得到离散化数组的实现代码：

```

1  copy(a + 1, a + 1 + n, b + 1);
2  sort(b + 1, b + 1 + n); //任意排序
3  m = unique(b + 1, b + 1 + n) - (b + 1); //相对大小为i的值为b[i]，m为相对最值
4  for (ll i = 1; i <= m; ++i)
5  {
6      s[b[i]] = i; //s是map，代表b里v的下标为s[v]
7  }
8  for (ll i = 1; i <= n; ++i)
9  {
10     printf("%lld ", s[a[i]]);
11 }

```

时间复杂度为排序复杂度 +  $O(n + m \log m + n \log m)$

`copy`, `sort`, `unique` 都是常用的 STL 函数。（`copy` 即 `memcpy`）

## 练习题

完成 [bny 学妹人物传记](#)、[计数排序](#)、[计数排序2](#)

## 归并排序

merge sort

实现思路：

- 将数组划分为平衡的两个子数组 (即长度相差不大于 1)
- 对两个子数组实施归并排序，使两个子数组分别有序
- 合并两个有序子数组为单个有序数组

使用**滑动窗口法**合并两个有序子数组：

- 左右子数组、新数组当前下标为起始值
- 若当前下标左数组元素小于等于右数组元素，把当前左数组元素放到新数组，左数组、新数组下标自增
- 否则，把当前右数组元素放到新数组，右数组、新数组下标自增
- 当一边子数组访问完毕后，将另一边子数组所有元素先后放到新数组后面即可

一种参考手写代码：([洛谷 P1177](#))

```
1  #include <bits/stdc++.h>
2  using namespace std;
3  #define mn 100010
4  typedef long long ll;
5  ll n, a[mn], b[mn];
6  void mergesort(ll lf, ll rf)
7  {
8      if (lf < rf)
9      {
10         ll cf = (lf + rf) >> 1; //left-face; center-face; right-face
11         mergesort(lf, cf);
12         mergesort(cf + 1, rf);
13         ll i = lf, j = cf + 1, je = rf, ie = cf, k = 0; //左[i,ie],右[j,je]
14         while (i <= ie && j <= je)
15         {
16             if (a[i] <= a[j])
17             {
18                 b[k++] = a[i++];
19             }
20             else
21             {
22                 b[k++] = a[j++];
23             }
24         }
25         while (i <= ie)
26         {
27             b[k++] = a[i++];
28         }
29         while (j <= je)
30         {
31             b[k++] = a[j++];
32         }
33     }
```

```

33         for (ll h = 0; h < k; ++h)
34         {
35             a[lf + h] = b[h];
36         }
37     }
38 }
39 signed main()
40 {
41     scanf("%lld", &n);
42     for (ll i = 1; i <= n; ++i)
43     {
44         scanf("%lld", &a[i]);
45     }
46     mergesort(1, n);
47     for (ll i = 1; i <= n; ++i)
48     {
49         printf("%lld ", a[i]);
50     }
51     return 0;
52 }

```

时间复杂度分析：根据主定理有  $a = b = 2, k = 1$ ， $a = b^k$ ，所以  $O(n^k \log_b n) = O(n \log n)$

空间复杂度：  $O(n)$

性质：是稳定排序

用途：以  $O(n \log n)$  求逆序对数目

- 滑动窗口合并时，若发现左数组元素大于右数组元素，说明形成逆序，由于左数组有序，即左数组接下来的元素必然也都大于该右数组元素。

当前左下标到左结尾共有  $ie - i + 1$  个元素，它们都大于当前右元素，贡献这么多个逆序对

即：

```

1  while (i <= ie && j <= je)
2  {
3      if (a[i] <= a[j])
4      {
5          b[k++] = a[i++];
6      }
7      else
8      {
9          ans += ie - i + 1;
10         b[k++] = a[j++];
11     }
12 }

```

逆序对也可以用线段树、树状数组求，大佬可以自行尝试



# 快速排序

Quick sort

快速排序抽象定义：

- 将数组划分为两个子数组，使得任意左子数组元素小于所有右子数组元素
- 对左右子数组依次执行快速排序，使得两个子数组分别有序

由于上列性质，当两个子数组分别有序时，直接拼接起来就是有序的数组

对所有基于比较的排序，快速排序是理论上平均复杂度最优的排序。

实现思路：(以**三路快速排序**为例)

- 对长为  $n$  的待排序数组  $a$ ，随机选择一个数组元素作基准值  $p$
- 所有小于基准值  $p$  的元素移动到数组的前面  
所有大于基准值  $p$  的元素移动到数组的后面
- 对所有小于基准值元素组成的子数组和所有大于基准值元素组成的子数组作为待排序数组，返回第一步

一种参考手写代码：(洛谷 P1177)

```
1  #include <bits/stdc++.h>
2  using namespace std;
3  #define mn 100010
4  typedef long long ll;
5  void quicksort(ll *a, ll n)
6  {
7      if (n <= 1)
8      {
9          return;
10     }
11     ll lf = 0, rf = n, i = 0, p = a[rand() % n];
12     while (i < rf)
13     {
14         if (a[i] < p)
15         {
16             swap(a[i++], a[lf++]);
17         }
18         else if (a[i] > p)
19         {
20             swap(a[i], a[--rf]);
21         }
22         else
23         {
24             ++i;
25         }
26     }
27     quicksort(a, lf);
28     quicksort(a + rf, n - rf);
29 }
30 ll n, a[mn];
31 signed main()
32 {
33
34     scanf("%lld", &n);
```

```

35     for (ll i = 1; i <= n; ++i)
36     {
37         scanf("%lld", &a[i]);
38     }
39     quicksort(a + 1, n);
40     for (ll i = 1; i <= n; ++i)
41     {
42         printf("%lld ", a[i]);
43     }
44     return 0;
45 }

```

这个实现的空间复杂度是  $O(1)$ ；根据具体实现不同，空间复杂度会有不同

性质：不稳定排序，平均复杂度  $O(n \log n)$ ，最差复杂度  $O(n^2)$

复杂度分析：

1. 最容易分析的情况是每次左右都分到一半(相差不大于 1)，可以近似看成两边相等，此时，根据主定理， $a = b = 2, k = 1, a = b^k$ ，复杂度为  $O(n^k \log_b n) = O(n \log n)$

事实上，这种情况是最优情况

2. 最坏的情况是每次某一边都是空，那么每次长度只减少 1，需要递归  $O(n)$  次，每次跑一层循环，复杂度为  $O(n^2)$

如果不使用随机基准，可以构造特殊数组使得每次基准都是最值；使用随机基准时，每次都取最值的概率约为  $\frac{2}{n!} \approx 0$  (注：计算不严谨，实际概率表达式远比这个复杂)

3. 平均情况的严格证明相对复杂 (具体参见文章 [这里](#)，前置知识：高等数学下-调和级数，各种高中数列知识)

(p.s. 这已经是算法理论证明里相当易懂的证明步骤了，多数算法的证明远比这个复杂)

直观上判断，可以看出平均复杂度介于最优和最坏之间

4. 实践表明，绝大多数情况下平均约等于最优

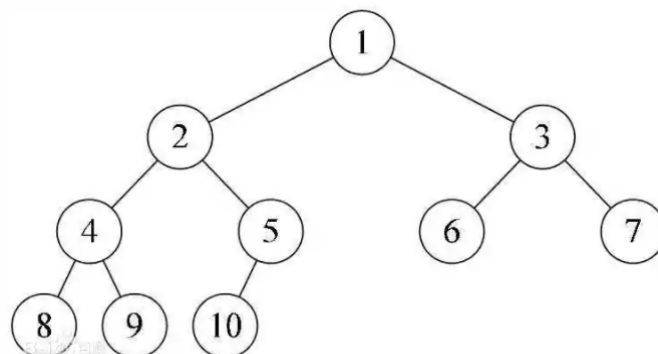
(还记得一开始的那句话吗：快速排序是理论上平均复杂度最优的排序)

## 堆排序

选读

heapsort

从下标 1 开始的数组可以转化为**完全二叉树**。



部分所用到的基本概念：

- 节点：即数组下标(节点下标)-该下标的数组值(结点值)的整体
- 根节点：下标为 1 的节点
- 子节点：有零到两个，该节点往下直接相连的节点；按编号大小依次为左、右子节点
- 父节点：有零到一个(根节点没有)，直接往上相连的节点
- 子树：以某节点为根节点的树(这里是完全二叉树)；有左右子节点为根的左子树、右子树
- 边：两节点间的连线
- 深度：根节点为 1；其他节点是到根节点的最短距离(边数目)
- 树高：所有节点深度的最大值

从下标 1 开始的完全二叉树性质：

- 当前节点编号为  $p$ ，则其父节点编号为  $\lfloor p/2 \rfloor$ ，左子节点编号为  $2p$ ，右子节点编号为  $2p + 1$
- 节点数为  $n$ ，则树高为  $\lfloor \log n \rfloor + 1$

**大根堆**(最大堆)是所有结点的值都大于其子节点值的完全二叉树。性质：

- 根节点必然是最大值
- 所有节点值不小于其子节点值

节点上浮操作：

- 当前节点值小于大儿子值(左右儿子的较大者)时，与大儿子互换，继续上浮；否则停止上浮

一个节点最多上浮  $\log n$  次，因为树高为  $\log n$ 。单次上浮复杂度为  $O(\log n)$

完全二叉树转大根堆：

- 从  $n$  到 1 每个节点依次上浮，复杂度为  $O(n \log n)$

不能从 1 到  $n$  上浮

思路：

1. 从下标 1 开始使用数组建完全二叉树
2. 将完全二叉树转化为大根堆
3. 取走删除堆顶(放到当前最后位置)，将最后一个节点作为新堆顶上浮，维护堆的性质，不断执行此操作直到堆空

一种参考手写代码：(洛谷 P1177)

```

1  #include <bits/stdc++.h>
2  using namespace std;
3  typedef long long ll;
4  #define mn 100010
5  ll a[mn], n;
6  void goup(ll bg, ll len)
7  {
8      ll now = bg, nx = bg * 2; //假设左儿子是大儿子
9      while (nx <= len)
10     {
11         if (nx + 1 <= len && a[nx + 1] > a[nx]) //找大儿子
12         {
13             nx++;
14         }
15         if (a[now] >= a[nx])
16         {
17             return;
18         }
19         swap(a[now], a[nx]);
    
```

```

20     now = nx;    //继续操作
21     nx = nx * 2; //假设左大
22 }
23 }
24 void heapsort()
25 {
26     for (ll i = n; i >= 1; --i)
27     {
28         goup(i, n);
29     }
30     for (ll i = n; i >= 1; --i)
31     {
32         swap(a[1], a[i]);
33         goup(1, i - 1);
34     }
35 }
36 signed main()
37 {
38     scanf("%lld", &n);
39     for (ll i = 1; i <= n; ++i)
40     {
41         scanf("%lld", &a[i]);
42     }
43     heapsort();
44     for (ll i = 1; i <= n; ++i)
45     {
46         printf("%lld ", a[i]);
47     }
48     return 0;
49 }

```

复杂度分析：时间  $O(n \log n + n \log n) = O(n \log n)$ ，空间  $O(1)$

性质：不稳定排序

## std::sort

### 复杂度

库 `algorithm`；当然万能头即可。

实现方法：**内省排序** (intro sort)，是快速排序和堆排序、插入排序的结合。当递归深度超过  $\lfloor \log_2 n \rfloor$  时使用堆排序，当元素数量少于阈值(如 16)时，用插入排序；否则使用快速排序。由于堆排序最差复杂度是  $O(n \log n)$ ，所以内省排序的最差复杂度是  $O(n \log n)$ 。

基本数据类型比较的复杂度是  $O(1)$ ，所以复杂度如上；若每次比较复杂度是  $f$ ，时间复杂度是  $O(fn \log n)$ 。

由于最大递归层数是  $\log n$ ，每层递归起码需要一个变量参数，所以空间复杂度起码是  $O(\log n)$

小数据时用插入排序，是因为插入排序在几乎有序时的复杂度约为  $O(n)$

具体原理参见 [这里](#)

对于大的、乱数列表一般相信是最快的已知排序

## 使用方法

```
1 | sort(头迭代器, 尾迭代器[, 比较依据函数]);
```

将会对迭代器范围 [ 头, 尾 ) 进行排序, 依据是比较函数或已定义的比较运算符。默认升序排序。

1. 对于数组, 头迭代器是首个元素指针地址; 尾迭代器是最后一个元素指针地址的下一个地址
2. 对于迭代器, 通常使用 `.begin()`, `.end()` 代表上述两个含义
3. 对于结构体等没有定义大小关系运算符的数据类型, 可以增加比较函数, 或者重载运算符

例:

1. 给一维整型数组  $a$  下标范围  $[0, n)$  的元素升序排序

```
1 | sort(a, a + n);
```

2. 给一维双精度浮点型数组  $a$  下标范围  $[1, n]$  的元素升序排序

```
1 | sort(a + 1, a + 1 + n)
```

3. 给二维长整数数组  $b$  第  $[1, n]$  行每行下标范围  $[1, m]$  的元素升序排序

```
1 | for (ll i = 1; i <= n; ++i)
2 | {
3 |     sort(b[i] + 1, b[i] + 1 + m);
4 | }
```

4. 给一维整型向量  $v$  所有元素升序排序

```
1 | sort(v.begin(), v.end());
```

5. 给一维整型数组  $a$  下标范围  $[1, n]$  的元素逆序排序

```
1 | sort(a + 1, a + 1 + n);
2 | reverse(a + 1, a + 1 + n);
```

```
1 | sort(a + 1, a + 1 + n, [](const ll &x, const ll &y)
2 |     { return x > y; }); //重定义 x < y 关系
```

```
1 | bool cmp(const ll &x, const ll &y) // ll x, ll y 也行, 上同
2 | {
3 |     return x > y; //不要写成 x-y因为负数也是true
4 | }
5 | //...
6 | sort(a + 1, a + 1 + n, cmp);
```

6. 给二元整数对  $(x, y)$  向量升序  $d$  排序, 先按  $x$  大小,  $x$  相同时按  $y$  大小

```

1 // vector<pair<ll, ll>> d = {{2, 2}, {1, 2}, {-1, -2}, {1, 1}, {2,
  3}}};
2 vector<pair<ll, ll>> d; //pair本身的比较运算符就是如此定义的
3 sort(d.begin(), d.end());
4 // for (auto i : d)
5 // {
6 //     printf("%lld %lld\n", i.first, i.second);
7 // }

```

```

1 struct node
2 {
3     ll x, y;
4     //bool operator<(node r) 但是不可以仅定义>不定义<
5     bool operator<(const node &r) const
6     {
7         if (x != r.x)
8         {
9             return x < r.x;
10        }
11        return y < r.y;
12    }
13 };
14 // vector<node> d = {{2, 2}, {1, 2}, {-1, -2}, {1, 1}, {2, 3}};
15 vector<node> d;
16 sort(d.begin(), d.end());
17 // for (auto i : d)
18 // {
19 //     printf("%lld %lld\n", i.x, i.y);
20 // }

```

```

1 struct node
2 {
3     ll x, y;
4 };
5 bool cmp(const node &l, const node &r)
6 { //即 l.x != r.x ? l.x < r.x : l.y < r.y;
7     if (l.x != r.x)
8     {
9         return l.x < r.x;
10    }
11    return l.y < r.y;
12 }
13 vector<node> d = {{2, 2}, {1, 2}, {-1, -2}, {1, 1}, {2, 3}};
14 signed main()
15 {
16     sort(d.begin(), d.end(), cmp);
17     for (auto i : d)
18     {
19         printf("%lld %lld\n", i.x, i.y);
20     }
21     return 0;
22 }

```

```

1 struct node
2 {

```

```

3     ll x, y;
4 };
5 vector<node> d = {{2, 2}, {1, 2}, {-1, -2}, {1, 1}, {2, 3}};
6 signed main()
7 {
8     sort(d.begin(), d.end(), [](const node &l, const node &r)
9         { return l.x != r.x ? l.x < r.x : l.y < r.y; });
10    for (auto i : d)
11    {
12        printf("%lld %lld\n", i.x, i.y);
13    }
14    return 0;
15 }

```

## 7. OJ1072 排名

```

1  #include <bits/stdc++.h>
2  using namespace std;
3  typedef long long ll;
4  ll n, m;
5  struct student
6  {
7      string name;
8      ll sum;
9      vector<ll> subj;
10     bool operator<(student r)
11     {
12         return sum != r.sum ? sum > r.sum : subj > r.subj;
13     }
14 } s[1010];
15 signed main()
16 {
17     cin >> n >> m;
18     for (ll i = 0; i < n; ++i)
19     {
20         cin >> s[i].name;
21         for (ll j = 0, v; j < m; ++j)
22         {
23             cin >> v;
24             s[i].sum += v;
25             s[i].subj.emplace_back(v);
26         }
27     }
28     sort(s, s + n);
29     for (ll i = 0; i < n; ++i)
30     {
31         cout << s[i].name << ' ';
32     }
33     return 0;
34 }

```

```

1  #include <bits/stdc++.h>
2  using namespace std;
3  typedef long long ll;
4  ll n, m;
5  struct student

```

```

6  {
7      string name;
8      ll sum;
9      vector<ll> subj;
10     bool operator<(student r)
11     {
12         if (sum != r.sum)
13         {
14             return sum > r.sum;
15         }
16         for (ll i = 0; i < m; ++i)
17         {
18             if (subj[i] != r.subj[i])
19             {
20                 return subj[i] > r.subj[i];
21             }
22         }
23     }
24 } s[1010];
25 signed main()
26 {
27     cin >> n >> m;
28     for (ll i = 0; i < n; ++i)
29     {
30         cin >> s[i].name;
31         for (ll j = 0, v; j < m; ++j)
32         {
33             cin >> v;
34             s[i].sum += v;
35             s[i].subj.emplace_back(v);
36         }
37     }
38     sort(s, s + n);
39     for (ll i = 0; i < n; ++i)
40     {
41         cout << s[i].name << ' ';
42     }
43     return 0;
44 }

```

时间复杂度分析：每次比较最差复杂度  $O(m)$ ，故时间复杂度为  $O(nm \log n)$

## 其他函数

选读

也是在 `algorithm` 库

### stable\_sort

用法同 `std::sort`，区别在于是稳定排序。复杂度与 `std::sort` 相近。

对稳定有要求的，一般都是结构体，一般而言可以直接重载时规定一下就好了



## nth\_element

```
std::nth_element(first, nth, last[, compare])
```

使得迭代器 `nth` 指向的第  $nth - first + 1$  大元素恰在该位置

时间复杂度  $O(n)$

如：洛谷 P1923 求第  $k$  小的数

```
1 #include<bits/stdc++.h>
2 using namespace std;
3 long long n,k,a[5000010];
4 int main()
5 {
6     scanf("%d%d",&n,&k);
7     for(int i=0;i<n;i++)
8         scanf("%d",&a[i]);
9     nth_element(a,a+k,a+n); //使第k小整数就位
10    printf("%d",a[k]); //调用第k小整数
11 }
```

想看具体实现思路找洛谷题解区别的题解

## 其他排序

选读

### 桶排序

bucket sort

适用于待排序数据值域较大但分布比较均匀的情况。追求更快的效率，可以使用桶排序。使用分块思想。

抽象步骤：

- 设置足够的空二维数组
- 将原数组平均拆分到二维数组里，保证任意前一个数组值小于后面的数组所有值
- 对每个数组排序
- 合并这些数组

桶排序稳定与否取决于第三个抽象步骤所用排序是否稳定。

一种实现是：值域为  $m$ ，将原数据分成  $\lceil \sqrt{m} \rceil$  个二维数组，每块值域跨度为  $s = \lfloor \sqrt{m} + 1 \rfloor$ ，第  $i$  个数组值域为  $[is, (i+1)s)$ ，可知最后一个数组右边界为  $\lceil \sqrt{m} \rceil \cdot \lfloor \sqrt{m} + 1 \rfloor \geq \sqrt{m} \cdot \sqrt{m} = m$ 。且各数组值域不重复。

数值  $v$  放在第  $\lfloor \frac{v}{s} \rfloor$  个数组里。对这  $k$  个二维数组分别排序(可以使用快排)。假设数据分布均匀，期望每个二维数组有  $\sqrt{n}$  个元素。那么，拆分复杂度为  $O(n)$ ，用 `std::sort` 对  $\sqrt{m}$  个数组排序复杂度为  $\sqrt{n} \log \sqrt{n}$ ，合并复杂度为  $O(n)$ 。总时间复杂度为  $O(n + \sqrt{mn} \log \sqrt{n} + n)$ ，其中  $\log \sqrt{n} = \frac{1}{2} \log n$ 。空间复杂度为  $O(m + n)$ 。

一种实现代码：(以 OJ 1577 违抗熵增之人 为例)

```
1 #include <bits/stdc++.h>
```

```

2 using namespace std;
3 typedef int ll;
4 #define il inline
5 typedef double db;
6 #define re register
7 il ll read()
8 {
9     re char p = 0;
10    re ll r = 0, o = 0;
11    for (; p < '0' || p > '9'; o |= p == '-', p = getchar())
12        ;
13    for (; p >= '0' && p <= '9'; r = (r << 1) + (r << 3) + (p ^ 48), p =
14        getchar())
15        ;
16    return o ? (~r) + 1 : r;
17 }
18 #define sc(x) x = read() //快读
19 #define mn 8000010
20 ll n, a[mn], mx, q, x, span;
21 vector<ll> b[65536]; //int的max:2^32开根是2^16=65536
22 signed main()
23 {
24     sc(n);
25     for (ll i = 1; i <= n; ++i)
26     {
27         sc(a[i]);
28         mx = max(mx, a[i]);
29     }
30     span = sqrt(mx) + 1; //块长
31     for (ll i = 1; i <= n; ++i)
32     {
33         b[a[i] / span].emplace_back(a[i]);
34     }
35     for (ll i = 0, k = 0; i < span; ++i)
36     {
37         sort(b[i].begin(), b[i].end());
38         for (ll j = 0, je = b[i].size(); j < je; ++j)
39         {
40             a[++k] = b[i][j];
41         }
42     }
43     sc(q);
44     while (q--)
45     {
46         sc(x);
47         printf("%d\n", a[x]);
48     }
49     return 0;
50 }

```

## 基数排序

学习后缀排序的前置知识

radix sort

是非比较排序。

抽象思想思路：将待排序元素拆分为若干个子元素(关键字)，逆序(从最不重要到最重要)依次稳定排序比较这些关键字。设有  $k$  个关键字，则需要进行  $k$  次排序。通常使用  $k$  次计数排序，那么复杂度为  $O(kn)$ 。

一种思想思路：

以整数为例，把  $k$  位十进制整数拆分为  $k$  个关键字，依次是从大到小的各数位。那么值域为  $m$  的整数有  $k = \lceil \log_{10} m \rceil$  位。所以时间复杂度为  $O(n \log_{10} m)$ 。由于计数排序，所以

329	720	720	329
457	355	329	355
657	436	436	436
839	457	839	457
436	657	355	657
720	329	457	720
355	839	657	839

十进制是为了理解方便。但对计算机实现方便而言(和享受更快的位运算)，通常用二的整数进制，如下面实现代码使用了 256 进制。

一种实现代码：(以 oj 1577 违抗熵增之人 为例)

```
1  #include <bits/stdc++.h>
2  #define re
3  using namespace std;
4  typedef int ll;
5  #define il inline
6  il ll read()
7  {
8      re char p = 0;
9      re ll r = 0, o = 0;
10     for (; p < '0' || p > '9'; o |= p == '-', p = getchar())
11         ;
12     for (; p >= '0' && p <= '9'; r = (r << 1) + (r << 3) + (p ^ 48), p =
13         getchar())
14         ;
15     return o ? (~r) + 1 : r;
16 }
17 #define sc(x) x = read()
18 #define mn 8000010
19 ll n, a[mn], q, x, b[mn], cnt[1 << 8];
20 void radixsort(ll n)
21 {
22     ll mask = (1 << 8) - 1;
```

```

22     11 *x = a, *y = b;
23     for (11 i = 0; i < 32; i += 8) //每八位每八位拆分，int拆为四个关键字
24     {
25         for (11 j = 0; j != (1 << 8); ++j) //[0,255)
26         {
27             cnt[j] = 0; //计数排序归零
28         }
29         for (11 j = 0; j != n; ++j)
30         {
31             ++cnt[x[j] >> i & mask]; //当前的八个位，注意位运算优先级
32         }
33         for (11 sum = 0, j = 0; j != (1 << 8); ++j)
34         { //计数排序变式，cnt代表同值元素段的开头下标;即减去前缀和前自己
35             sum += cnt[j], cnt[j] = sum - cnt[j];
36         }
37         for (11 j = 0; j != n; ++j)
38         { //计数排序变式
39             y[cnt[x[j] >> i & mask]++] = x[j];
40         }
41         swap(x, y); //y是排序好的,x是未排序的;换过来，继续操作
42     }
43 }
44 signed main()
45 {
46     sc(n);
47     for (11 i = 0; i < n; ++i)
48     {
49         sc(a[i]);
50     }
51     radixsort(n);
52     sc(q);
53     while (q--)
54     {
55         sc(x);
56         printf("%d\n", a[x - 1]);
57     }
58     return 0;
59 }

```

## 练习题

完成 [逆序对](#)、[结构体排序](#)

## 有序的维护

要求对于动态的数组(如不断插入、删除)，一直保持该数组有序，从而能够利用有序完成一系列题目要求的事情。

根据各题具体要求不同，可能会采取各种不同的数据结构。这里只讲一小部分可以直接调用库函数的知识，剩下大部分其他知识的在后续课程讲或自学。

# STL

## xx\_bound

对升序序列：

`lower_bound(begin, end, v)` 在  $[begin, end)$  找第一个大于等于  $v$  的值，返回该值的迭代器

`upper_bound(begin, end, v)` 在  $[begin, end)$  找第一个大于  $v$  的值，返回该值的迭代器

若找不到返回  $end$

对降序序列：

`lower_bound(begin, end, v, greater<type>())` 在  $[begin, end)$  找第一个小于等于  $v$  的值，返回该值的迭代器

`upper_bound(begin, end, v, greater<type>())` 在  $[begin, end)$  找第一个小于  $v$  的值，返回该值的迭代器

若找不到返回  $end$

时间复杂度：  $O(\log n)$

使用测试：

```
1 #include <bits/stdc++.h>
2 using namespace std;
3 typedef long long ll;
4 vector<ll> a = {1, 3, 3, 3, 3, 5}, b(a);
5 ll v = 3;
6 signed main()
7 {
8     ll p1 = lower_bound(a.begin(), a.end(), 3) - a.begin();
9     auto p2 = upper_bound(a.begin(), a.end(), 3) - a.begin(); //11
10    cout << p1 << endl;
11    cout << p2 << endl;
12    ll p3 = lower_bound(a.begin(), a.end(), -1) - a.begin();
13    ll p4 = lower_bound(a.begin(), a.end(), 6) - a.begin();
14    cout << p3 << endl;
15    cout << p4 << endl;
16    // ll p5 = upper_bound(a.begin(), a.end(), -1) - a.begin();
17    // ll p6 = upper_bound(a.begin(), a.end(), 6) - a.begin();
18    // cout << p5 << endl;
19    // cout << p6 << endl;
20
21    // copy(a.begin(), a.end(), b.begin()); (error)
22    // b.assign(a.begin(), a.end());
23    reverse(b.begin(), b.end());
24
25    ll p7 = lower_bound(b.begin(), b.end(), 3, greater<ll>()) - b.begin();
26    ll p8 = upper_bound(b.begin(), b.end(), 3, greater<ll>()) - b.begin();
27    cout << p7 << endl;
28        << p8 << endl;
```

```

29
30     ll p9 = lower_bound(a.begin(), a.end(), -1, greater<ll>()) -
a.begin();
31     ll p10 = lower_bound(a.begin(), a.end(), 6, greater<ll>()) -
a.begin();
32     cout << p9 << endl;
33     cout << p10 << endl;
34     // ll p11 = lower_bound(a.begin(), a.end(), -1, greater<ll>()) -
a.begin();
35     // ll p12 = lower_bound(a.begin(), a.end(), 6, greater<ll>()) -
a.begin();
36     // cout << p11 << endl;
37     // cout << p12 << endl;
38
39     return 0;
40 }

```

```

1  #include <bits/stdc++.h>
2  using namespace std;
3  typedef long long ll;
4  struct node
5  {
6      ll x, y;
7      bool operator<(const node &r) const { return x < r.x; }
8  };
9  vector<node> v = {{1, 3}, {2, 6}, {2, 4}, {2, 8}, {3, 0}};
10 signed main()
11 {
12     auto a = *lower_bound(v.begin(), v.end(), (node){2, 2});
13     cout << a.x << ' ' << a.y << endl;
14     a = *upper_bound(v.begin(), v.end(), (node){2, 2});
15     cout << a.x << ' ' << a.y << endl;
16     auto pos = upper_bound(v.begin(), v.end(), (node){2, 2}) - v.begin() -
1;
17     cout << v[pos].x << ' ' << v[pos].y << endl;
18     return 0;
19 }

```

例：SCNUOJ 1216 二分查找(区间计数)

请尝试手写实现二分查找。

给定一个长度为  $n$  的序列  $a$ ，你需要处理  $m$  个询问。

对于每一个询问，给定两个整数  $l$  和  $r$ ，你需要回答序列中有多少项满足  $l \leq a_i \leq r$ 。

输入

第一行包含两个整数  $n, m$  ( $1 \leq n, m \leq 10^5$ )，含义如题目所示。

第二行包含  $n$  个用空格间隔的整数  $a_1, a_2, \dots, a_n$  ( $-10^9 \leq a_i \leq 10^9$ )，表示序列  $a$ 。

接下来  $m$  行每行包含两个用空格间隔的整数  $l, r$  ( $-10^9 \leq l \leq r \leq 10^9$ )，含义如题目描述所示。

输出

对于每个询问，输出一行，包含一个整数，表示答案。

```
1 #include <bits/stdc++.h>
2 using namespace std;
3 typedef long long ll;
4 ll n, m, a[100010], l, r;
5 signed main()
6 {
7     scanf("%lld%lld", &n, &m);
8     for (ll i = 1; i <= n; ++i)
9     {
10         scanf("%lld", &a[i]);
11     }
12     sort(a + 1, a + 1 + n);
13     while (m--)
14     {
15         scanf("%lld%lld", &l, &r);
16         auto pos1 = lower_bound(a + 1, a + 1 + n, l);
17         auto pos2 = upper_bound(a + 1, a + 1 + n, r);
18         printf("%lld\n", pos2 - pos1);
19         // printf("%lld %lld\n", pos1 - a, pos2 - a);
20     }
21     return 0;
22 }
```

## 序列式容器

使用 `vector` 或 `deque` 来维护：

- 插入一个元素，找位置 `lower_bound` + 插入 `insert`  $O(\log n + n)$
- 删除一个元素，找位置 `lower_bound` + 删除 `erase`  $O(\log n + n)$
- 查找特定位置元素(第  $k$  大) `[]`  $O(1)$
- 查找特定值元素 `lower_bound, upper_bound`  $O(\log n)$

更多其他操作及其复杂度见 [这里](#)

以 洛谷 P3369 为例:

您需要写一种数据结构 (可参考题目标题) , 来维护一些数, 其中需要提供以下操作:

1. 插入  $x$  数
2. 删除  $x$  数(若有多个相同的数, 因只删除一个)
3. 查询  $x$  数的排名(排名定义为比当前数小的数的个数  $+1$ )
4. 查询排名为  $x$  的数
5. 求  $x$  的前驱(前驱定义为小于  $x$ , 且最大的数)
6. 求  $x$  的后继(后继定义为大于  $x$ , 且最小的数)

## 输入格式

第一行为  $n$ , 表示操作的个数,下面  $n$  行每行有两个数  $opt$  和  $x$ ,  $opt$  表示操作的序号( $1 \leq opt \leq 6$ )

## 输出格式

对于操作 3, 4, 5, 6 每行输出一个数, 表示对应答案

### 【数据范围】

对于 100% 的数据,  $1 \leq n \leq 10^5$ ,  $|x| \leq 10^7$

```
1 #include <bits/stdc++.h>
2 using namespace std;
3 vector<int> v;
4 int n, cmd, x;
5 signed main()
6 {
7     scanf("%d", &n);
8     while (n--)
9     {
10         scanf("%d%d", &cmd, &x);
11         if (cmd == 1)
12             v.insert(lower_bound(v.begin(), v.end(), x), x); //或upper
13         else if (cmd == 4)
14             printf("%d\n", v[x - 1]);
15         else if (cmd == 2)
16             v.erase(lower_bound(v.begin(), v.end(), x));
17         else
```



```

18     {
19         int pos = lower_bound(v.begin(), v.end(), x) - v.begin();
20         if (cmd == 3)
21             printf("%d\n", pos + 1);
22         else if (cmd == 5)
23             printf("%d\n", v[pos - 1]);
24         else if (cmd == 6)
25         {
26             int pos2 = upper_bound(v.begin(), v.end(), x) - v.begin();
27             printf("%d\n", v[pos2]);
28         }
29     }
30 }
31 return 0;
32 }

```

分析：理论最优复杂度是  $O(n \log n)$ ，最差复杂度是  $O(n^2 \log n)$ ，能过题因为数据弱。

## 关联式容器

使用 `map`, `multimap`, `set`, `multiset` 来维护，可以用  $O(\log n)$  实现：

- 插入一个元素 `insert`
- 删除一个元素 `erase`
- 查找特定值元素 `find`, `lower_bound`, `upper_bound`

支持遍历。对 `*_bound` 得到的

无法实现查询当前迭代器是第  $k$  大(除非暴力遍历)。使用 `nth_element` 复杂度为  $O(n)$

注：需要使用成员方法才能确保复杂度，不能使用外部函数(复杂度是  $O(n)$ )

内部实现：红黑树(一种平衡树)

使用示例：

```

1  #include <bits/stdc++.h>
2  using namespace std;
3  typedef long long ll;
4  vector<ll> v = {1, 2, 3, 2, 2, 2};
5  multiset<ll> s{v.begin(), v.end()}; //数组同理
6  signed main()
7  {
8      for (auto i : s)
9      {
10         cout << i << endl;
11     }
12     cout << endl;
13     auto p1 = s.lower_bound(2);
14     auto p2 = s.upper_bound(2);
15     cout << *p1 << endl;
16     p1--;
17     cout << *p1 << endl;
18     cout << *p2 << endl;
19     p2--;

```

```

20 cout << *p2 << endl;
21 return 0;
22 }

```

```

1  #include <bits/stdc++.h>
2  using namespace std;
3  typedef long long ll;
4  struct node
5  {
6      ll x, y;
7      bool operator<(const node &r) const { return x < r.x; }
8  } a[] = {{1, 1}, {2, 4}, {2, 8}, {3, 3}, {2, 3}};
9  void print(node a) //成员函数也行
10 {
11     cout << a.x << ' ' << a.y << endl;
12 }
13 set<node> s1{a, a + 5};
14 multiset<node> s2{a, a + 5};
15 signed main()
16 {
17     for (auto i : s1)
18     {
19         print(i);
20     }
21     cout << endl;
22     for (auto i : s2)
23     {
24         print(i);
25     }
26     cout << endl;
27
28     auto p1 = s2.find({2, 0}); //首个插入的{2,*}
29     print(*p1);
30     print(*s2.lower_bound({2, 0}));
31     auto p2 = s2.upper_bound({2, 0});
32     // cout << p2 - s2.begin() << endl; //不可行;p2-p1也不可行
33     p2--;
34     print(*p2);
35     p2--;
36     print(*p2);
37     return 0;
38 }

```

```

1  #include <bits/stdc++.h>
2  using namespace std;
3  typedef long long ll;
4  typedef pair<ll, ll> pr;
5  pr a[] = {{1, 1}, {2, 4}, {2, 8}, {3, 3}, {2, 3}};
6  multimap<ll, ll> m{a, a + 5};
7  void print(pr x)
8  {
9      cout << x.first << ' ' << x.second << endl;
10 }
11 signed main()
12 {
13     for (auto i : m)

```

```

14 {
15     print(i);
16 }
17 auto p1 = m.lower_bound(2);
18 auto p2 = m.upper_bound(2);
19 print(*p1);
20 p2--;
21 print(*p2);
22 return 0;
23 }

```

## pb\_ds

Policy-Based Data Structures , 是一个库。

只在使用 libstdc++ 为标准库的编译器下可以用。

命名空间: `__gnu_pbds` (区分: `bits/stdc++` 对应命名空间 `std`)

仅重点介绍 `pb_ds` 的以红黑树为基础的平衡树结构的使用。

另外一种比较有用的结构是配对堆, 功能是维护最值(最大或最小之一), 能够以  $O(1)$  的复杂度插入元素、查询最值, 以均摊  $O(\log n)$  (最坏  $O(n)$ ) 的复杂度删除最大元素、修改元素。对于只需要查询动态最值时, 是比较实用的数据结构。(因为很多时候 `priority_queue` 就可以了, 所以这里不重点介绍配对堆)。配对堆原理见 [这里](#)

`pb_ds` 的进阶应用这里不细讲, 感兴趣自学。

更多信息资料参考: [oi-wiki](#) , [洛谷日报](#)

创建:

```

1 #include <ext/pb_ds/assoc_container.hpp>
2 #include <ext/pb_ds/tree_policy.hpp>
3 __gnu_pbds::tree<类型, __gnu_pbds::null_type, less<类型>,
  __gnu_pbds::rb_tree_tag, __gnu_pbds::tree_order_statistics_node_update> 变量名;

```

用法和复杂度同 `set` , 但在 `set` 的基础上, 增加了:

- `order_of_key(x)` 返回值  $x$  在 `set` 的下标
- `find_by_order(x)` 返回下标为  $x$  的迭代器

复杂度为  $O(\log n)$

由此可以实现查找不重复的动态区间第  $k$  大。

要实现查找重复的动态区间第  $k$  大, 可以参考如下思路, 把重复值化为不重复值:

- 记最多元素数为  $n$  , 其二进制位数为  $d = \lceil \log n \rceil$  位; 可以把第  $i$  个数  $a$  变为  $2^d \times a + i$  那么就构造出了互不相同的数, 且  $a$  不同时排序严格不同。取出来还原时即  $\lfloor \frac{a}{d} \rfloor$
- 例如 `0b01`, `0b01`, `0b11`, `0b00` 可以转化为 `0b0100`, `0b0101`, `0b1110`, `0b0011`

- 直接  $na + i$  也是可以的，还原即  $\lfloor \frac{a}{n} \rfloor$ 。

转化后具体如何进行各操作参考下面例题。

以 [洛谷 p6136](#) 为例：

## 题目描述

您需要写一种数据结构（可参考题目标题），来维护一些整数，其中需要提供以下操作：

1. 插入一个整数  $x$ 。
2. 删除一个整数  $x$ （若有多个相同的数，只删除一个）。
3. 查询整数  $x$  的排名（排名定义为比当前数小的数的个数 + 1）。
4. 查询排名为  $x$  的数（如果不存在，则认为是排名小于  $x$  的最大数。保证  $x$  不会超过当前数据结构中数的总数）。
5. 求  $x$  的前驱（前驱定义为小于  $x$ ，且最大的数）。
6. 求  $x$  的后继（后继定义为大于  $x$ ，且最小的数）。

本题强制在线，保证所有操作合法（操作 2 保证存在至少一个  $x$ ，操作 4, 5, 6 保证存在答案）。

## 输入格式

第一行两个正整数  $n, m$ ，表示初始数的个数和操作的个数。

第二行  $n$  个整数  $a_1, a_2, a_3, \dots, a_n$ ，表示初始的数。

接下来  $m$  行，每行有两个整数  $opt$  和  $x'$ ， $opt$  表示操作的序号（ $1 \leq opt \leq 6$ ）， $x'$  表示加密后的操作数。

我们记  $last$  表示上一次 3, 4, 5, 6 操作的答案，则每次操作的  $x'$  都要异或上  $last$  才是真实的  $x$ 。初始  $last$  为 0。

## 输出格式

输出一行一个整数，表示所有 3, 4, 5, 6 操作的答案的异或和。

对于 100% 的数据， $1 \leq n \leq 10^5$ ， $1 \leq m \leq 10^6$ ， $0 \leq a_i, x < 2^{30}$ 。

本题输入数据较大，请使用较快的读入方式。

```
1 #include <bits/stdc++.h>
2 using namespace std;
3 typedef long long ll;
4 #include <ext/pb_ds/assoc_container.hpp>
5 #include <ext/pb_ds/tree_policy.hpp>
6 __gnu_pbds::tree<ll, __gnu_pbds::null_type, less<ll>,
  __gnu_pbds::rb_tree_tag, __gnu_pbds::tree_order_statistics_node_update>
  tr;
7 ll n, m, lastans, dig, ans;
8 signed main()
9 {
10     scanf("%lld%lld", &n, &m);
11     dig = n + m;
12     for (ll i = 1; i <= n; ++i)
13     {
14         scanf("%lld", &a);
```

```

15     tr.insert(a * dig + i);
16 }
17 for (ll i = n + 1, opt, x; i <= m + n; ++i)
18 {
19     scanf("%lld%lld", &opt, &x);
20     x ^= lastans;
21     if (opt == 1)
22     {
23         tr.insert((x * dig) + i);
24     }
25     else if (opt == 2)
26     {
27         tr.erase(tr.lower_bound(x * dig));
28     }
29     else
30     {
31         if (opt == 3)
32         {
33             lastans = tr.order_of_key(x * dig) + 1;
34         }
35         else if (opt == 4)
36         {
37             lastans = (*tr.find_by_order(x - 1)) / dig;
38         }
39         else if (opt == 5)
40         {
41             lastans = (*--tr.lower_bound(x * dig)) / dig;
42         }
43         else if (opt == 6)
44         {
45             lastans = (*tr.upper_bound(x * dig + dig)) / dig;
46         }
47         ans ^= lastans;
48     }
49 }
50 printf("%lld\n", ans);
51 return 0;
52 }

```

附：手写平衡树的参考代码如下(以 [Treap](#) 为例，其他平衡树参考洛谷题解)

```

1  #include <bits/stdc++.h>
2  using namespace std;
3  typedef long long ll;
4  #define sc(x) scanf("%lld", &x)
5  #define mn 1100010 //这是因为可能插入m个点所以1e6(插入)+1e5(本来)
6  typedef pair<ll, ll> pr;
7  ll key[mn], w[mn], siz[mn], son[mn][2], root, seed = 1;
8  ll n, m, cnt, ans, last, c, x;
9  ll rand1() { return seed *= 19260817; } //手写随机
10 void pushup(ll u) { siz[u] = siz[son[u][0]] + siz[son[u][1]] + 1; }
11 pr split(ll u, ll k)
12 {
13     if (!u)
14     {
15         return {0, 0};

```

```

16 }
17 if (key[u] < k)
18 {
19     pr t = split(son[u][1], k);
20     son[u][1] = t.first;
21     pushup(u);
22     return {u, t.second};
23 }
24 else
25 {
26     pr t = split(son[u][0], k);
27     son[u][0] = t.second;
28     pushup(u);
29     return {t.first, u};
30 }
31 }
32 ll merge(ll u, ll v)
33 {
34     if (!u || !v)
35     {
36         return u + v;
37     }
38     if (w[u] < w[v])
39     {
40         son[u][1] = merge(son[u][1], v);
41         pushup(u);
42         return u;
43     }
44     else
45     {
46         son[v][0] = merge(u, son[v][0]);
47         pushup(v);
48         return v;
49     }
50 }
51 void insert(ll k)
52 {
53     key[++cnt] = k, w[cnt] = rand1(), siz[cnt] = 1;
54     pr t = split(root, k);
55     root = merge(merge(t.first, cnt), t.second);
56 }
57 void del(ll k)
58 {
59     pr x = split(root, k), y = split(x.second, k + 1);
60     y.first = merge(son[y.first][0], son[y.first][1]);
61     root = merge(x.first, merge(y.first, y.second));
62 }
63 ll find1(ll k)
64 {
65     pr t = split(root, k);
66     ll v = siz[t.first] + 1;
67     root = merge(t.first, t.second);
68     return v;
69 }
70 ll find2(ll k)
71 {
72     ll pos = root;
73     while (pos)

```

```

74 {
75     if (k == siz[son[pos][0]] + 1)
76     {
77         return key[pos];
78     }
79     if (k <= siz[son[pos][0]])
80     {
81         pos = son[pos][0];
82     }
83     else
84     {
85         k -= siz[son[pos][0]] + 1, pos = son[pos][1];
86     }
87 }
88 return -1; //to erase warning; in fact non-reachable
89 }
90 ll lst(ll k) { return find2(find1(k) - 1); }
91 ll nxt(ll k) { return find2(find1(k + 1)); }
92 signed main()
93 {
94     sc(n), sc(m);
95     for (ll i = 1; i <= n; ++i)
96     {
97         sc(x), insert(x);
98     }
99     for (ll i = 1; i <= m; ++i)
100     {
101         sc(c), sc(x);
102         if (c == 1)
103         {
104             insert(x ^ last);
105         }
106         else if (c == 2)
107         {
108             del(x ^ last);
109         }
110         else if (c == 3)
111         {
112             last = find1(x ^ last), ans ^= last;
113         }
114         else if (c == 4)
115         {
116             last = find2(x ^ last), ans ^= last;
117         }
118         else if (c == 5)
119         {
120             last = lst(x ^ last), ans ^= last;
121         }
122         else if (c == 6)
123         {
124             last = nxt(x ^ last), ans ^= last;
125         }
126     }
127     printf("%lld\n", ans);
128     return 0;
129 }

```

# 提升练习

---

完成：

- Easy String Problem
- 小朋友排队
- 后缀排序(BF实现)
- Left 4 Dead 2
- 排兵布阵
- Ares, Toilet Ares
- 次小值
- 普通平衡树
- 寄蒜几盒