数据结构

第10章　内排序

主讲教师：祝建华

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

10.1 概述

1. 排序：将文件或表中的记录，通过某种方法整理成按关键字大小次序排列的处理过程。

假定n个记录的文件为

$$(R_1, R_2, \ldots, R_n)$$

对应的关键字为

$$(K_1, K_2, \ldots, K_n)$$

则排序是确定如下一个排列

$$p_1, p_2, \ldots, p_n$$

使得：$Kp_1 \leqslant Kp_2 \leqslant \ldots \leqslant Kp_n$

从而得到一个有序文件

$$(Rp_1, Rp_2, \ldots Rp_n)$$

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网课基地（www.e-studysky.com）；咨询QQ：2696670126

学生成绩表

	学号	姓名	数学	外语
1	20051	刘大海	80	75
2	20042	王　伟	90	83
3	20066	吴晓英	82	88
4	20038	刘　伟	80	70
5	20052	王　洋	60	70

(a) 无序表

	学号	姓名	数学	外语
1	20038	刘　伟	80	70
2	20042	王　伟	90	83
3	20051	刘大海	80	75
4	20052	王　洋	60	70
5	20066	吴晓英	82	88

(b) 按学号排列的有序表

	学号	姓名	数学	外语
1	20052	王　洋	60	70
2	20051	刘大海	80	75
3	20038	刘　伟	80	70
4	20066	吴晓英	82	88
5	20042	王　伟	90	83

(c) 按数学成绩排列的有序表

	学号	姓名	数学	外语	总分
1	20042	王　伟	90	83	173
2	20066	吴晓英	82	88	170
3	20051	刘大海	80	75	155
4	20038	刘　伟	80	70	150
5	20052	王　洋	60	70	130

(d) 按总分成绩排列的有序表

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

2. 什么是排序的稳定性

假设在待排序的文件中，存在两个具有相同关键字的记录R(i)与R(j)，其中R(i)位于R(j)之前。在用某种排序法排序之后，R(i)仍位于R(j)之前，则称这种排序方法是稳定的；否则，称这种排序方法是不稳定的。

例 数列

(10, 25, 22, 42, 25, 30, 18) —稳定的排序→ (10, 18, 22, 25, 25, 30, 42)

(10, 25, 22, 42, 25, 30, 18) —不稳定的排序→ (10, 18, 22, 25, 25, 30, 42)

	学 号	姓 名	数学	外语
1	20051	刘大海	80	75
2	20042	王 伟	90	83
3	20066	吴晓英	82	88
4	20038	刘 伟	80	70
5	20052	王 洋	60	70

不稳定的排序 →

	学 号	姓 名	数学	外语
1	20052	王 洋	60	70
2	20038	刘 伟	80	70
3	20051	刘大海	80	75
4	20066	吴晓英	82	88
5	20042	王 伟	90	83

(e) 按数学成绩排列的有序表

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

3. 内部排序(内排序)----在计算机内存中进行的排序
   外部排序(外排序)----借助计算机外存进行的排序

4. 待排序的记录和顺序表(文件)的数据类型

```
#define MAXSIZE 20          //最大长度
typedef int KeyType;        //关键字类型
typedef struct              //记录类型
 { KeyType key;             //关键字
   InfoType otherinfo;      //其它数据类型
 }RecType;                  //记录类型名
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

```
typedef struct
   { RecType r[MAXSIZE+1]; //r[0]用作监视哨
     int length;           //实际表长
    }SeqList;              //记录表类型
```

或表和表长分别定义和说明

```
RecType r[MAXSIZE+1];     //r[0]用作监视哨
int length;               //实际表长
```

length≤MAXSIZE

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

5. 排序算法分析

(1) 时间复杂度

对n个记录排序，所需比较关键字的次数；

最好情况；最坏情况；平均情况

对n个记录排序，所需移动记录的次数；

最好情况；最坏情况；平均情况

(2) 空间复杂度

排序过程中，除文件中的记录所占的空间外，所需的辅助存储空间的大小。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

6. 内排序方法

(1) 对顺序表的排序

插入排序：

- 直接插入排序；
- 折半插入排序；
- 2-路插入排序；
- 表插入排序；
- 希尔(Shell)排序；

交换排序：

- 冒泡排序：单向冒泡排序，双向冒泡排序
- 快速排序

选择排序：简单选择/选择排序；

- 树形选择排序；
- 堆排序

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

归并排序：

2-路归并排序

k-路归并排序

基数排序：

多关键字排序

最高位优先法

最低位优先法

链式基数排序

(2)对单链表的排序：

直接插入, 简单选择, 冒泡排序, 基数排序

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

10.2 插入排序

算法基本思想

将待排序的记录插入到已排序的子文件中去，使得插入之后得到的子文件仍然是有序子文件。插入一个记录，首先要对有序子文件进行查找，以确定这个记录的插入位置。按查找方式的不同，插入排序又可以分为线性插入排序和折半插入排序，前者使用顺序查找，后者使用折半查找。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

1. 直接插入排序(线性插入排序)

设待排序的文件为：(r[1],r[2],...,r[n])

关键字为：(r[1].key,r[2].key,...,r[n].key)

首先,将初始文件中的记录r[1]看作有序子文件；

第1遍：将r[2]插入前面有序子文件中，得到3个记录的有序子文件。

第2遍：将r[3]插入前面有序子文件中，得到3个记录的有序子文件。

以此类推，依次插入r[4],...,r[n],最后得到n个记录的递增有序文件。

r[0]	r[1]	r[2]	r[3]	r[4]	r[5]	r[6]
	43	21	89	15	43	28
	j	j	i	j	j	i

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：考学天地（www.e-studysky.com），咨询QQ：2696670826

例：直接插入排序，设r[0]为“监视哨”

	r[0]	r[1]	r[2]	r[3]	r[4]	r[5]	r[6]
初始关键字：		(43)	21	89	15	43	28
第1遍排序后：	21	(21	43)	89	15	43	28
第2遍排序后：	89	(21	43	89)	15	43	28
第3遍排序后：	15	(15	21	43	89)	43	28
第4遍排序后：	43	(15	21	43	43	89)	28
第5遍排序后：	28	(15	21	28	43	43	89)

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

直接插入排序算法（对数组r[1..n]中的n个记录作插入排序）

```
void InsertSort(RecType r[],int n)
{ int i,j;
  for (i=2; i<=n; i++)
   { r[0]=r[i];                 //待插记录r[i]存入监视哨中
     j=i-1;                     //已排序的范围1 — i-1
                                //从r[i-1]开始向左扫描
     while(r[0].key<r[j].key)
     { r[j+1]=r[j];             //记录后移
       j--;                     //继续向左扫描
      }
     r[j+1]=r[0];               //插入记录r[0],即原r[i]
   }
}
```

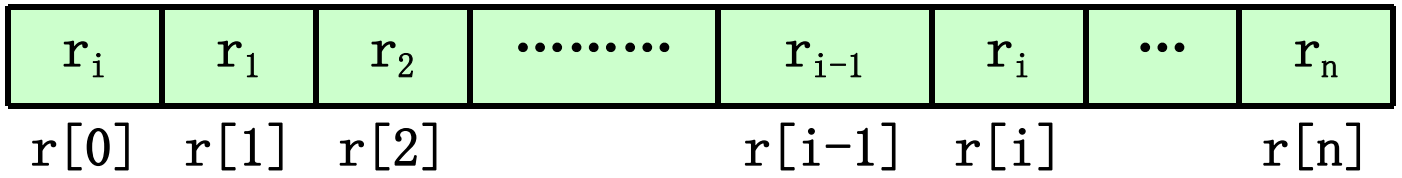
华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

直接插入排序算法分析：

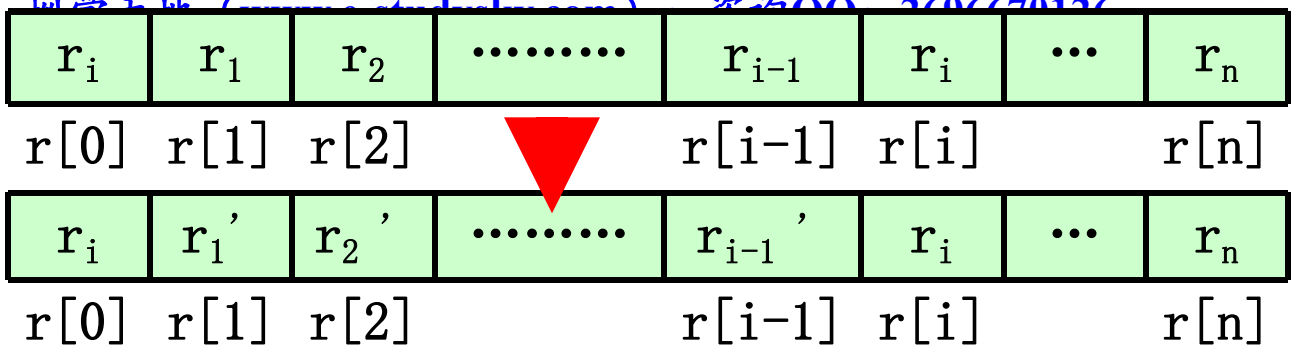
(1)最好情况，原n个记录递增有序：

比较关键字n-1次

移动记录2(n-1)次,(将数据复制到r[0]后又复制回来)



华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.studylocated.com） 咨询QQ：2696670126



(2)最坏情况，原n个记录递减有序：

比较关键字的次数：

$$\sum_{i=2}^{n} i = 2+3+\ldots+n = (n-1)(n+2)/2 = O(n^2)$$

移动记录的次数(个数)：

$$\sum_{i=2}^{n} (i-1+2) = 3+4+\ldots+(n+1)$$

$= (n-1)(n+4)/2$ 次 $= O(n^2)$

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地 www.e-studysky.com 客服QQ：2696670126

(3)平均比较关键字的次数约为:

$$\sum_{i=2}^{n}(1+2+\ldots+i)/i=\sum_{i=2}^{n}(i+1)/2$$

$$=(3+4+\ldots+(n+1))/2=\frac{(n-1)(n+4)}{4}$$

平均移动记录的次数约为:

$$\sum_{i=2}^{n}((0+2)+(1+2)+\ldots+(i-1+2))/i=\sum_{i=2}^{n}(i+3)/2$$

$$=(5+6+\ldots+(n+3))/2=\frac{(n-1)(n+8)}{4}$$

故,时间复杂度为$O(n^2)$。

(4) 只需少量中间变量作为辅助空间。

(5) 算法是稳定的。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

2. 折半插入排序

减少了记录的比较次数平均值，
记录的移动次数不变。
稳定性不变

```
void BInsertSort(RecType r[],int n）
{ int i,j;
 for (i=2；  i<=n；  i++)
  { r[0]=r[i]；              //待插记录r[i]存入监视哨中
    low = 1 ；   high = i - 1;
    while ( low <= high )
          { m = ( low + high ) / 2 ;
     if ( r[0].key <r[m].key )
        high = m - 1;          //插入位置可能在低半区
     else  low = m+1;          //插入位置可能在高半区
     } //结束时表示插入在high和low之间
     for (j = i – 1; j  >= high + 1;- -j ) r[j+1] = r[j];
     r[high + 1] = r[0];
  }
}
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

3 shell排序

shell排序是由P.L.Shell于1959年提出的一种对插入排序的改进算法。其算法思想是将待排序的数据分成若干组：

将待排序的数据，简单处理时,可首先看成n/2组，即间隔为n/2的数据元素在同一组中，每一组使用前面的插入排序算法，实现数据的大跨度前移。

再将待排序的数据看成n/4组，即间隔为n/4的数据元素在同一组中，每一组使用前面的插入排序算法。

余以类推。当比较间隔缩小到0时，整个排序过程结束。由于每比较一趟，就将比较间隔缩小一半.

间隔也称为增量, shell算法又称为缩小增量排序算法。算法分析与增量的选取有关.

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com），咨询QQ：2696670126

gap=10/2 ： 5

a[0] a[1] a[2] a[3] a[4] a[5] a[6] a[7] a[8] a[9]

43 21 89 15 43 28 27 26 50 78

gap=gap/2=5/2 ： 2

a[0] a[1] a[2] a[3] a[4] a[5] a[6] a[7] a[8] a[9]

28 21 26 15 43 43 27 89 50 78

gap=gap/2=2/2 ： 1

a[0] a[1] a[2] a[3] a[4] a[5] a[6] a[7] a[8] a[9]

26 15 27 21 28 43 43 78 50 89

gap=gap/2=1/2 ： 0

a[0] a[1] a[2] a[3] a[4] a[5] a[6] a[7] a[8] a[9]

15 21 26 27 28 43 43 50 78 89

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

gap=10/2 : 5

a[0]	a[1]	a[2]	a[3]	a[4]	a[5]	a[6]	a[7]	a[8]	a[9]
26	25	27	25	28	48	43	89	50	78

```
void shellsort(int a[], int n)
{ int gap, i, j, t;
  /*开始时gap的值为n/2，然后每趟间隔缩小一半*/
  for(gap=n/2;gap>0;gap=gap/2)
    for(i=gap;i<n;i++)
      for(j=i-gap;j>=0 && a[j]>a[j+gap];j-=gap)
        t=a[j],a[j]=a[j+gap],a[j+gap]=t;
}
```

- 需要将范围：gap——n-1 内的所有元素作为起点，用a[i]表示，将a[i]和同组前面的有序排列的元素进行比较，必要时进行交换。
- 用元素a[j]从后向前依次表示a[i]前面的同组元素，开始时，j=i-gap，当j>0时，表示a[j]存在，并且这时a[j]和a[j+gap]进行比较。

若a[j]>a[j+gap] 这将这两个数交换，j向前移动，即 j=j-gap；否这结束a[i]的处理。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

```
void ShellSort(int a[],int n,int delta[],int t)
{
 int i,j,k,gap;
  for(k=0;k<t;k++)
   for(i=gap=delta[k];i<n;i++)
       for(j=i-gap;j>=0 && a[j]>a[j+gap];j-=gap)
       t=a[j],a[j]=a[j+gap],a[j+gap]=t;
}
```

delta[t]={…9,5,3,2,1}　　这里：t为排序趟数

$$delta[k]=2^{t-k}+1 \qquad 0\le k\le t\le \lfloor \log_2(n-1)\rfloor$$

delta[t]={…40,13,4,1}

$$delta[k]=\frac{1}{2}(3^{t-k}-1) \qquad 0\le k\le t\le \lfloor \log_3(2n+1)\rfloor$$

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

10.3 交换排序

10.3.1 冒泡排序

基本思想: 设待排序的文件为r[1..n]

第1趟(遍)：从r[1]开始,依次比较两个相邻记录的关键字r[i].key和r[i+1].key,若r[i].key>r[i+1].key，则交换记录r[i]和r[i+1]的位置；否则,不交换。(i=1,2,...n-1)

第1趟之后,n个关键字中最大的记录移到了r[n]的位置上。

第2趟：从r[1]开始,依次比较两个相邻记录的关键字r[i].key和r[i+1].key,若r[i].key>r[i+1].key，则交换记录r[i]和r[i+1]的位置；否则,不交换。 (i=1,2,...n-2)

第2趟之后,前n-1个关键字中最大的记录移到了r[n-1]的位置上。

.....

作完n-1趟,或者不需再交换记录时为止。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

一般情况

第1遍：	10	3	7	8	5	2	1	4	9	6	[0—10-1)
第2遍：	3	7	8	5	2	1	4	9	6	10	[0—10-2)
第3遍：	3	7	5	2	1	4	8	6	9	10	[0—10-3)
第4遍：	3	5	2	1	4	7	6	8	9	10	[0—10-4)
第5遍：	3	2	1	4	5	6	7	8	9	10	[0—10-5)
第6遍：	2	1	3	4	5	6	7	8	9	10	[0—10-6)
第7遍：	1	2	3	4	5	6	7	8	9	10	[0—10-7)
第8遍：	1	2	3	4	5	6	7	8	9	10	[0—10-8)
第9遍：	1	2	3	4	5	6	7	8	9	10	[0—10-9)
	1	2	3	4	5	6	7	8	9	10	交换范围

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
远程一对一辅导班（www.e-studysky.com）；咨询QQ：2696670126

算法分析：

- 最好情况：待排序的文件已是有序文件，只需要进行1趟排序，共计比较关键字的次数为
  n－1　　不交换记录。
- 最坏情况：要经过n-1趟排序，所需总的比较关键字的次数为
  (n-1)+(n-2)+...+1＝n(n-1)/2
  交换记录的次数最多为
  (n-1)+(n-2)+...+1＝n(n-1)/2
  移动记录次数最多为
  3n(n-1)/2 。
- 只需要少量中间变量作为辅助空间。
- 算法是稳定的。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

冒泡排序算法(对n个整数按递增次序作冒泡排序)

```
void  bubble1(int a[],int n)
{ int i,j,temp;
  for(i=0; i<n-1; i++)        //作n-1趟排序
    for(j=0; j<n-1-i; j++)
      if (a[j]>a[j+1])
        { temp=a[j];          //交换记录
          a[j]=a[j+1];
          a[j+1]=temp;
        }
  for(i=0; i<n; i++)
    printf("%d",a[i]);    //输出排序后的元素
}
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

改进的冒泡排序算法

```
void bubblesort(RecType r[],int n)
{ int i,j,swap;  RecType temp;
  j=1;                    //置比较的趟数为1
  do{ swap=0;             //置交换标志为0
      for (i=1; i<=n-j; i++)
       { if (r[i].key>r[i+1].key)
          {  temp=r[i];   //交换记录`
             r[i]=r[i+1];
             r[i+1]=temp;
             swap=1;      //置交换标志为1
          }
         j++;             //作下一趟排序
       }
   } while (j<n && swap);
     //未作完n-1趟,且标志为1
}
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

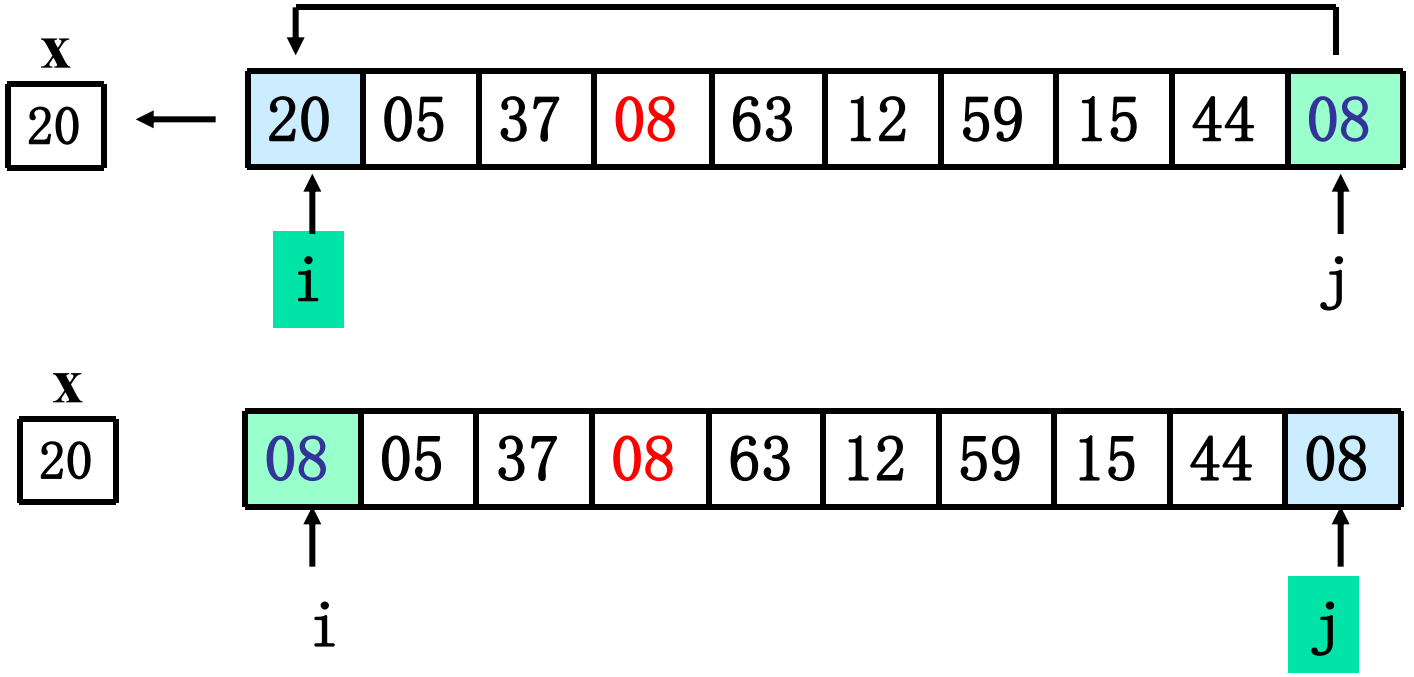
10.3.2 快速排序

基本思想：首先在r[1..n]中，确定一个r[i]，经过比较和移动，将r[i]放到"中间"某个位置上，使得r[i]左边所有记录的关键字小于等于r[i].key，r[i]右边所有记录的关键字大于等于r[i].key。以r[i]为界,将文件划分为左、右两个子文件。

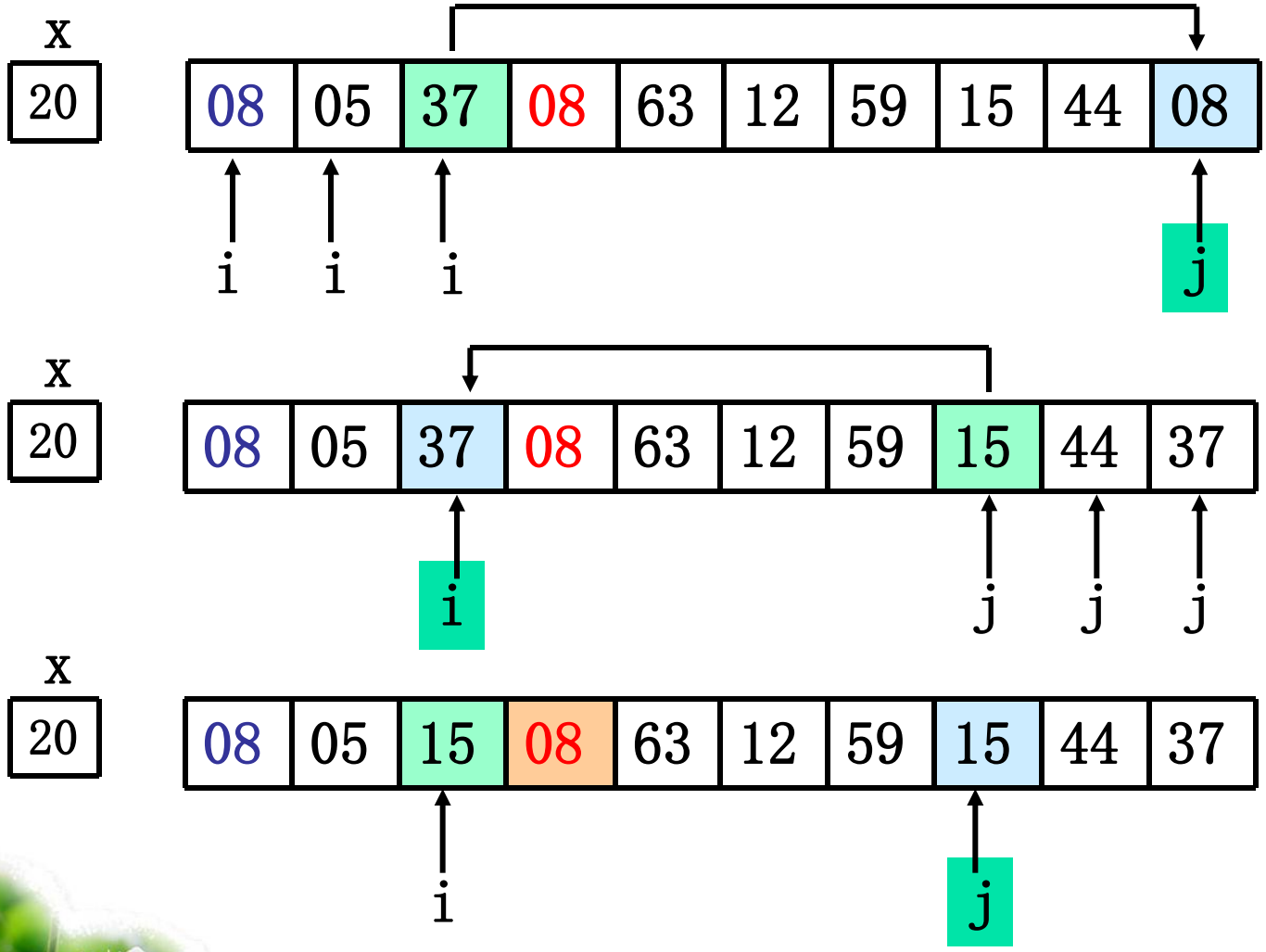
用同样的方法分别对这两个子文件进行划分，得到4个更小的子文件。继续进行下去,使得每个子文件只有一个记录为止,便得到原文件的有序文件。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

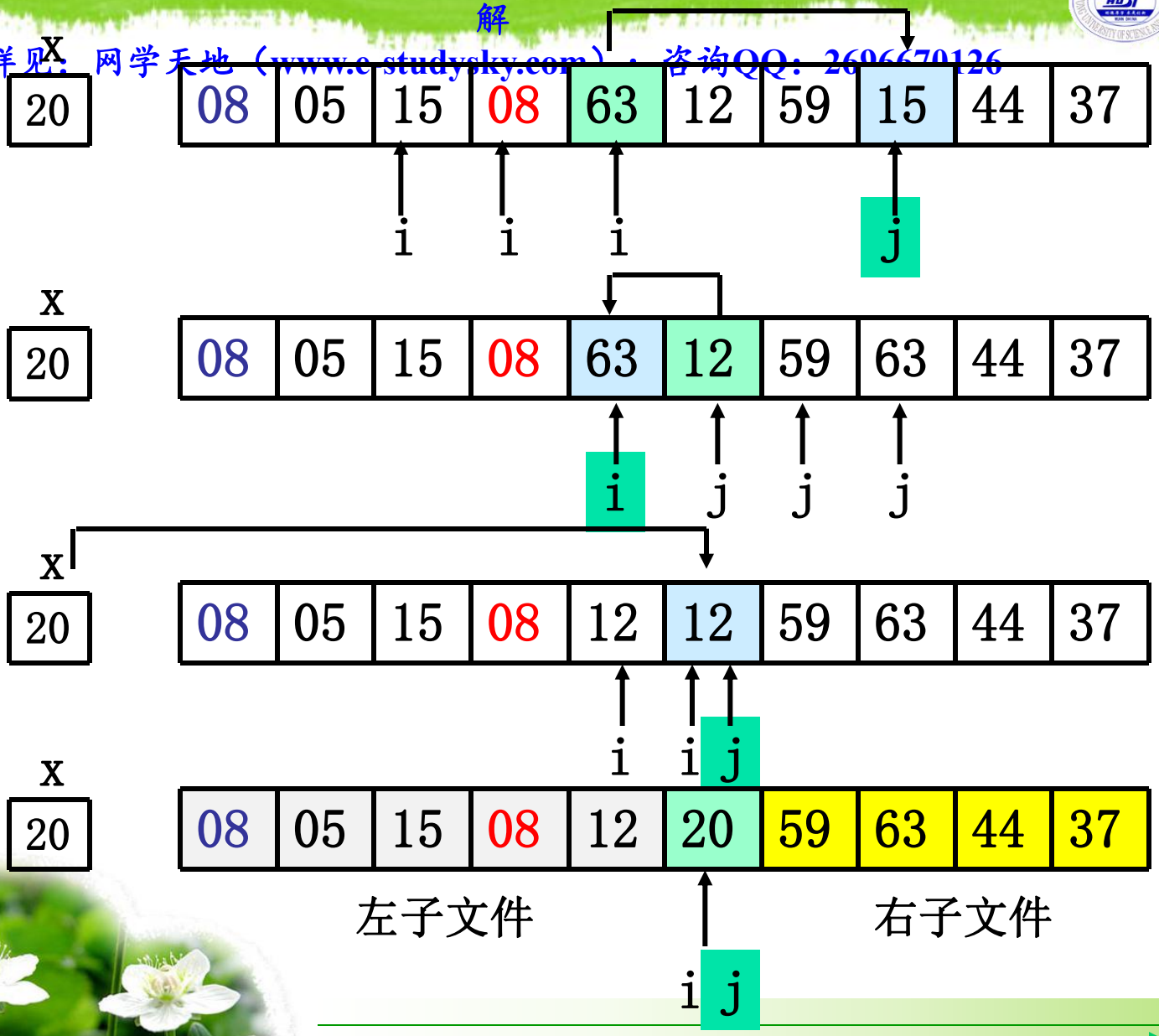
例. 给定文件(20, 05, 37, 08, 63, 12, 59, 15, 44, 08)，选用第1个元素20进行划分：



华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126



华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e studysky.com），咨询QQ：2696670126



华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
华中科技大学
更多：网学天地（www.studysky.com），咨询QQ：2696670126

```
void quksort(RecType r[],int low,int high)
{  RecType x; int i,j;
  if (low<high)                          //有两个以上记录
   { i=low; j=high; x=r[i];              //保存记录到变量x中
     do {                               //此时i指示位置可用
         while (i<j && r[j].key>=x.key)
           j--; //j从右向左端扫描通过key不小于x.key的元素
         if (i<j)                       //i,j未相遇
          { r[i]=r[j];  i++;            //此时j指示位置可用
            while(i<j && r[i].key<=x.key)
              i++;   //i从左向右端扫描通过key不大于x.key的元素
            if (i<j)
              { r[j]=r[i]; j--; }
          }
     } while (i!=j);                    //i,j未相遇
}
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

```
//划分结束，i经过的是key不大于x.key的元素；
                    j经过的是key不小于x.key的元素。
                    i,j至少有一个指示的位置可用
        r[i]=x;
        quksort(r,low,i-1);                    //递归处理左子文件
        quksort(r,i+1,high);                   //递归处理右子文件
      }
  }
```

对文件r[1..n]快速排序：

```
 void quicksort(RecType r[],int n)
 { quksort(r,1,n);  }
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

算法分析

- 就平均速度而言，快速排序是已知内部排序方法中最好的一种排序方法，其时间复杂度为O(nlog(n))。
- 但是，在最坏情况下（有序或基本有序时，此时每次划分都出现一端为空的形式），快速排序所需的比较次数和冒泡排序的比较次数相同，其时间复杂度为$O(n^2)$。
- 快速排序需要一个栈作辅助空间，用来实现递归处理左、右子文件。在最坏情况下，递归深度为n，因此所需栈的空间大小为O(n)数量级。
- 快速排序是不稳定的。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网易云课堂（www.e-studysky.com）；咨询QQ：2696670126

10.4 选择排序

10.4.1. 简单选择(选择排序)

算法思想:设待排序的文件为(r[1],r[2],...,r[n]),关键字为(r[1].key,r[2].key,...,r[n].key),

第1趟(遍)：在(r[1],r[2],...,r[n])中,选出关键字最小的记录r[min].key，若min<>1,则交换r[1]和r[min]；

需要进行n-1次比较。

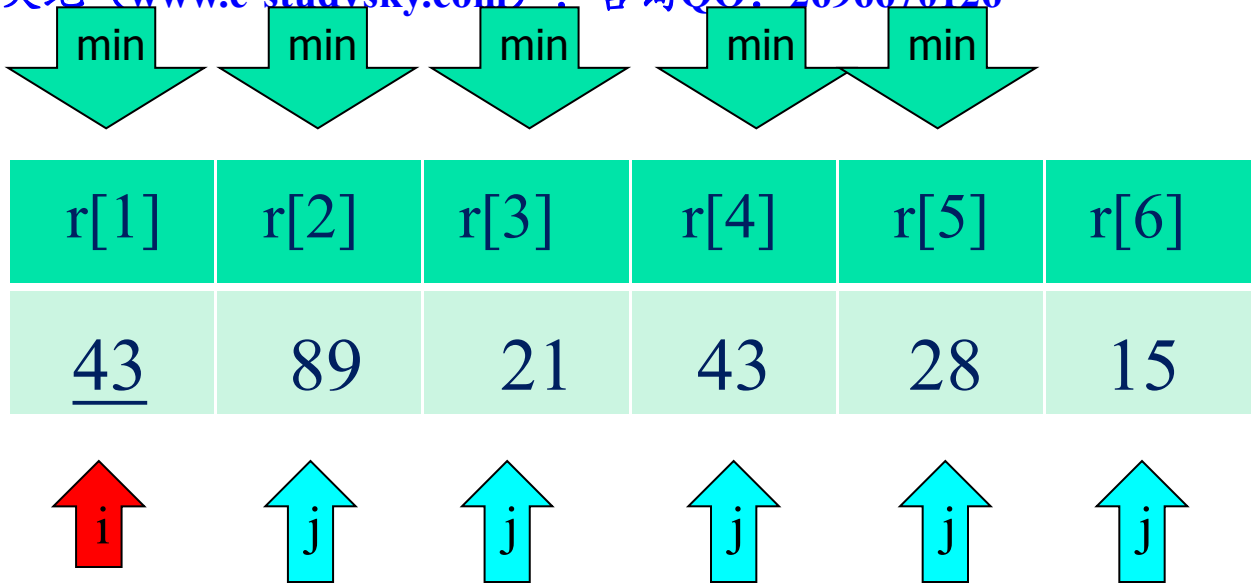
第2趟(遍)：在n-1个记录(r[2],...,r[n])中,选出关键字最小的记录r[min].key，若min<>2,则交换r[2]和r[min]；

需要进行n-2次比较。

.....

第n-1趟(遍)：在最后的2个记录记录(r[n-1],r[n])中,选出关键字最小的记录r[min].key，若min<>n-1,则交换r[n-1]和min]；需要进行1次比较。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126



i的取值范围：　1.....n-1
j的取值范围：　i+1……n
Min定位区间：i……n中最小关键字记录的下标

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
淘宝：网络天地 www.studyskycity.com 咨询QQ：2696670126

简单选择排序算法：（对数组r[1..n]中的记录作简单选择排序）

```
void SelectSort(RecType r[],int n)
{ int i,j,min;
  RecType x;                          //交换记录的中间变量
  for (i=1;i<n;i++)                   //共n-1趟(遍)
   { min=i;                           //r[i]为最小记录r[min]
     for (j=i+1;j<=n;j++)
       if (r[j].key<r[min].key)
          min=j;                      //修改min
     if (min!=i)                      //若r[min]不是r[i]
      { x=r[min];                     //交换r[min]和r[i]
        r[min]=r[i]; r[i]=x;
       }
   }
}
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

算法分析：

(1) 比较次数，在任何情况下，均为

$$\sum_{i=1}^{n-1}(n-i)=(n-1)+(n-2)+\ldots+1$$

$$=\ n(n-1)/2\ =O(n^2)$$

(2) 交换记录的次数

在最好情况下，原n个记录递增有序：

不移动记录。

在最坏情况下，每次都要交换数据（不是递减有序）

共交换记录n-1对，移动记录数3(n-1)次。

故，时间复杂度为$O(n^2)$。

(3) 只需少量中间变量作为辅助空间。

(4) 算法是不稳定的。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

10.4.2. 堆排序（Heap Sort）

堆的定义：n个元素的序列$\{k_1, k_2, \cdots, k_n\}$ 当且仅当满足下关系时，称之为堆。

$$\begin{cases} k_i \leqslant k_{2i} \\ k_i \leqslant k_{2i+1} \end{cases} \quad \text{或} \quad \begin{cases} k_i \geqslant k_{2i} \\ k_i \geqslant k_{2i+1} \end{cases} \qquad (i=1, 2, \cdots, \lfloor n/2 \rfloor)$$

其中：　前面一种称为小顶堆；

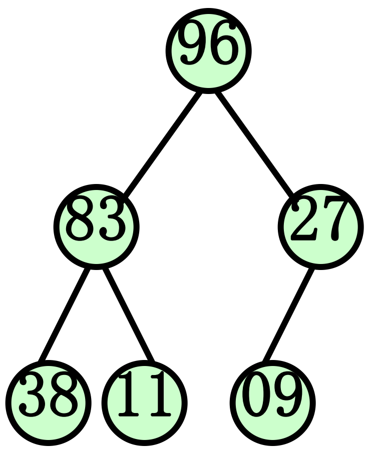
后面一种称为大顶堆。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

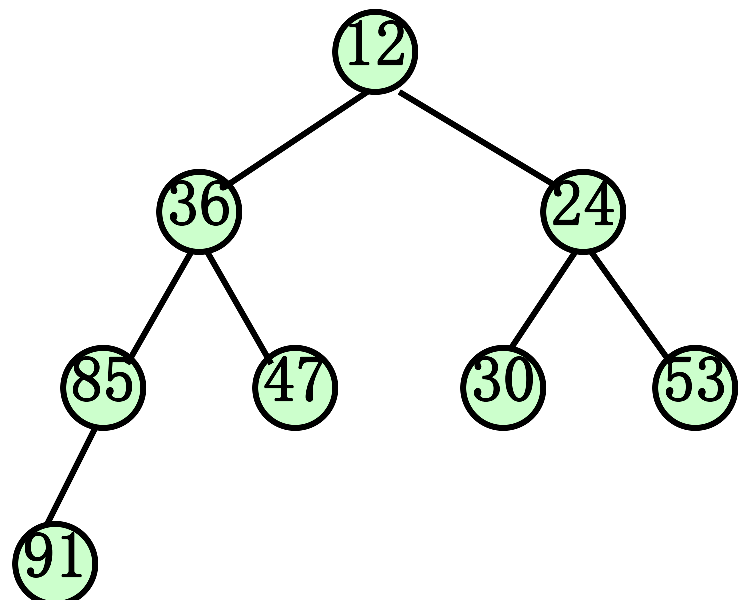
n个元素的序列{k_1,k_2,…,k_n}可看成是一个结点个数为n的完全二叉数，若其为大顶堆，则k_1最大；若其为小顶堆，则k_1最小。

例：　序列1：{96，83，27，38，11，09}

　　　序列2：{12，36，24，85，47，30，53，91}



序列1的二叉树(大顶堆)



序列2的二叉树(小顶)堆

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

通常，n个元素的序列$\{k_1, k_2, \cdots, k_n\}$ 不符合堆的定义，所以，面临的第一个问题：

问题1：

如何将序列$\{k_1, k_2, \cdots, k_n\}$ 处理成（大顶）堆（初始化）？

问题1一旦解决，得到规模为n的堆，则k_1最大，将k_1与k_n互换，则最大的数已放置到最后，同时，剩下的序列$\{k_1, k_2, \cdots, k_{n-1}\}$不是堆，如何将其重新处理成规模为n-1的堆，求取第二大的数据，以此类推，堆的规模逐步减小，直到求出第n-1大的数据，完成递增排序。所以，面临的第二个问题是：

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

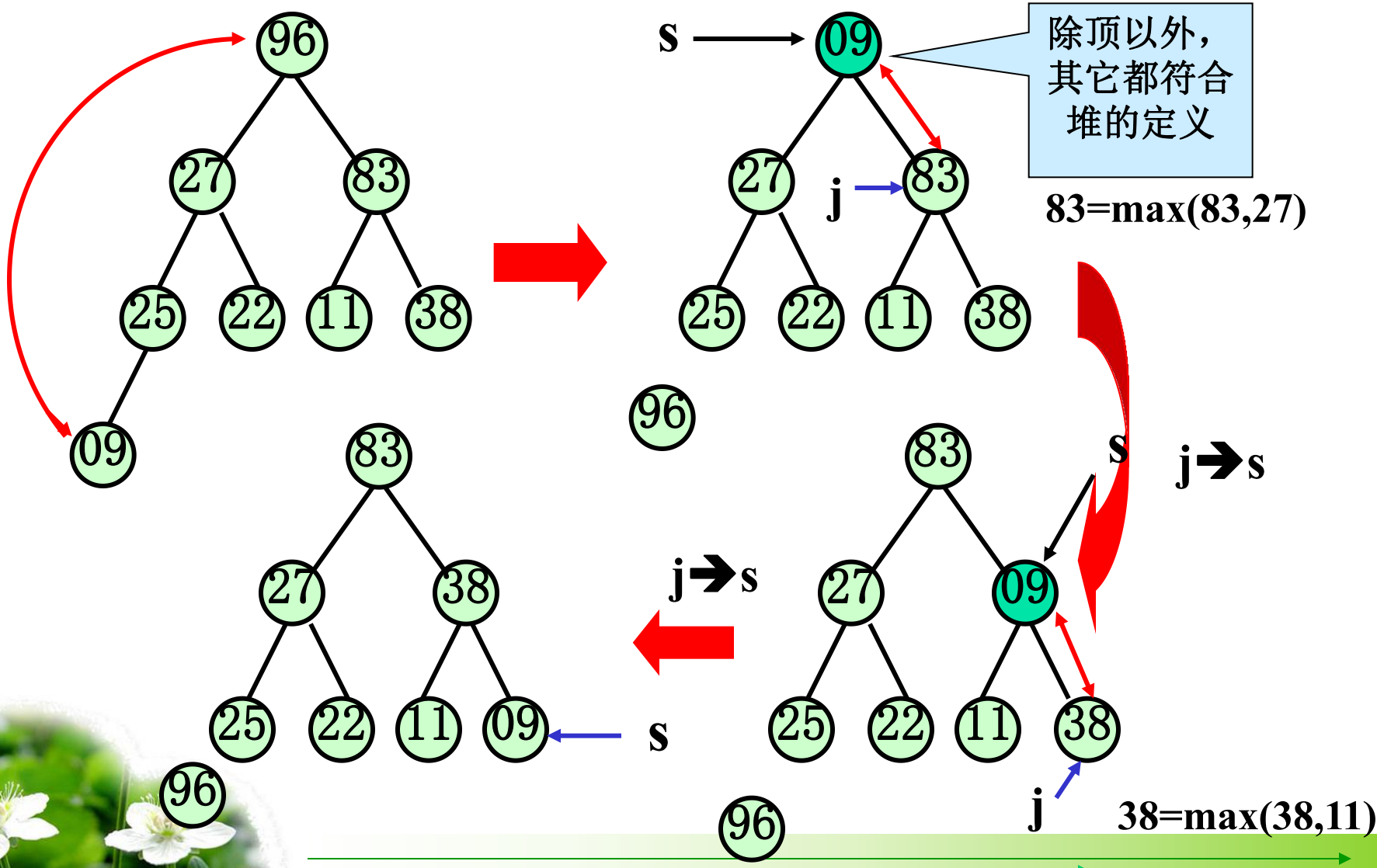
问题2：

如何在堆顶元素被替换后，调整剩余元素成为一个新的堆。

提示：根据上述过程描述， 借助大顶堆可实现序列的递增排序；借助小顶堆可实现序列的递减排序。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：考研天堂（www.studytx.com）咨询QQ：2663670126

问题2方法：某序列的堆形式：{96，27，83，25，22，11，38，09}



华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

```
typedef   SqList      HeapType;
//采用顺序表存储表示
void   HeapAdjust(HeapType   &H,
                  int s, int   m)
```

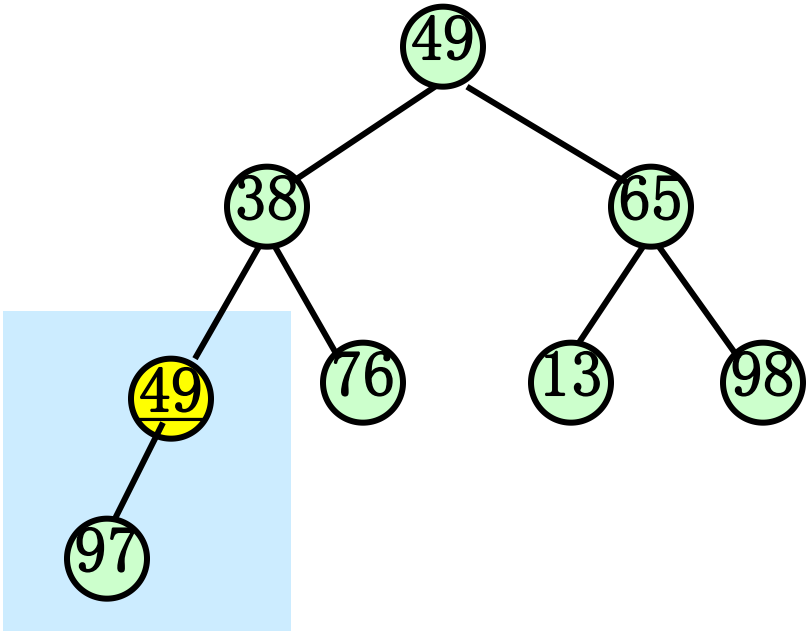
//已知H.r[s…m]中记录的关键字除H.r[s].key之外均满足堆的定义，本函数调整H.r[s]的关键字,使H.r[s…m]成大顶堆。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com），咨询QQ：2696670126

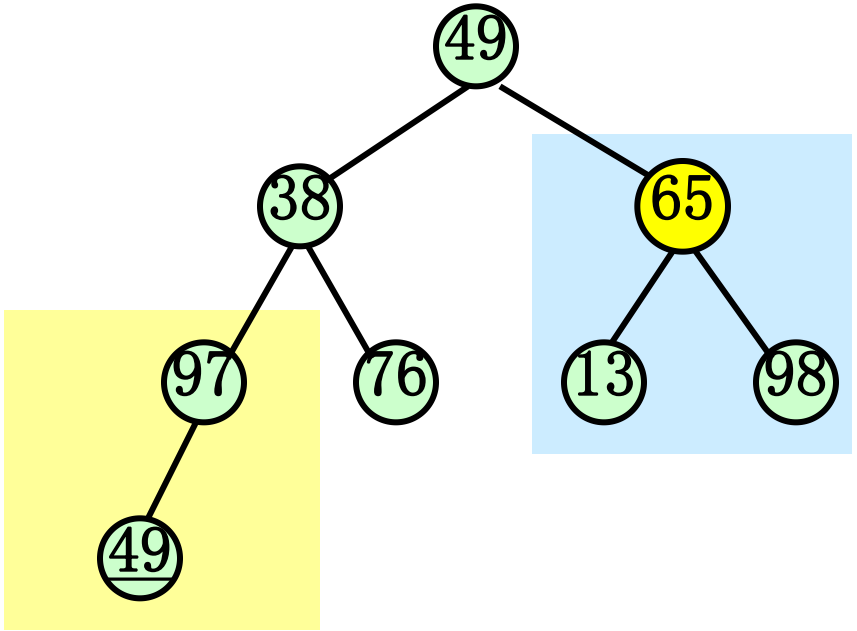

```
void  HeapAdjust(HeapType  &H, int s, int  m)
{ rc=H.r[s];                    //保存需调整的数据元素，空出s的记录位置
  for(j=2*s; j<=m;j*=2) {   //j<=m表示s有左孩子序号     j=2*s
    if (j<m&&H.r[j].key<H.r[j+1].key) //j<m表示s有右孩子j+1
          j++;                      //计算s的具有较大关键字的孩子的序号j
    if (rc.key> H.r[j].key)//rc在s的结点时满足结点定义,调整完毕
          break;
    H.rs[s]=H.r[j];s=j;       //s的较大孩子上移，修改s下移
   }                                //正常结束时，  s的结点无孩子结点。
  H.r[s]=rc;
}
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

问题1方法:建立序列：{49, 38, 65, 49, 76, 13, 98, 97}的初始堆。

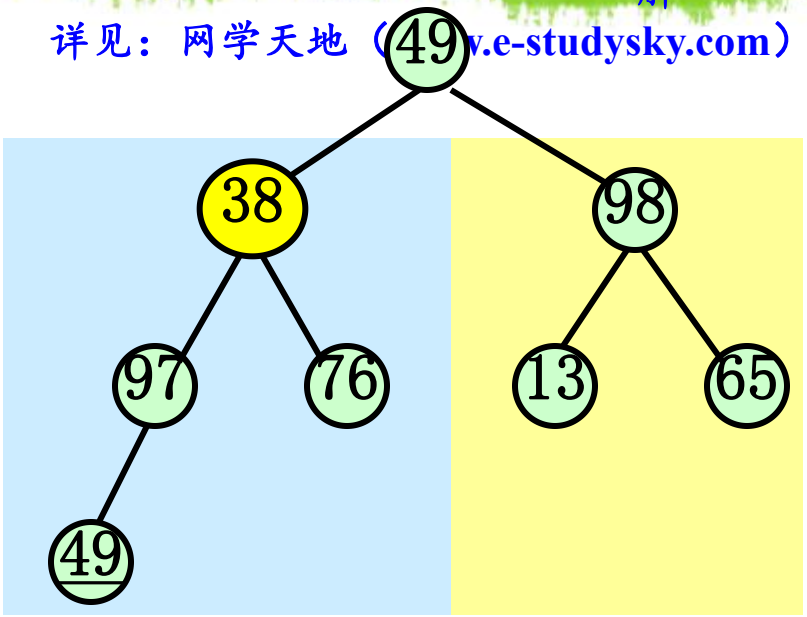


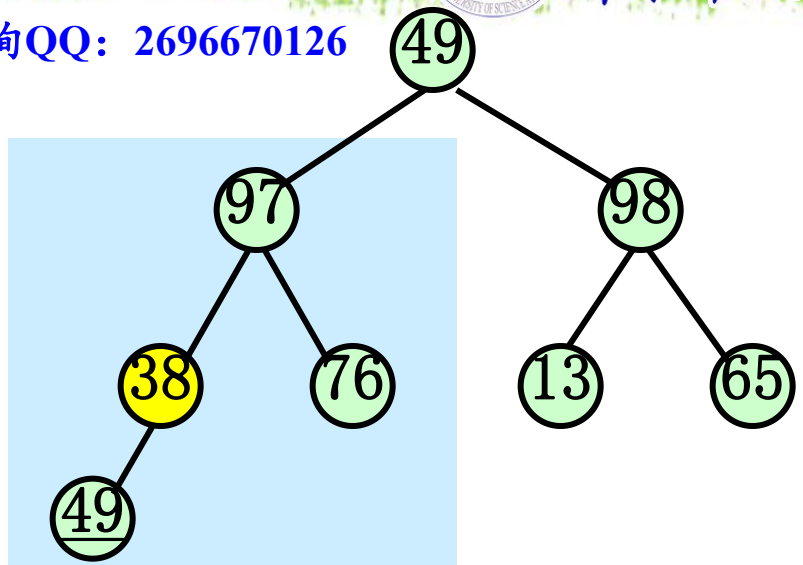
1。对以最后一个结点（序号n）的双亲结点（序号i=⌊n/2⌋）为根的二叉树，进行堆调整。



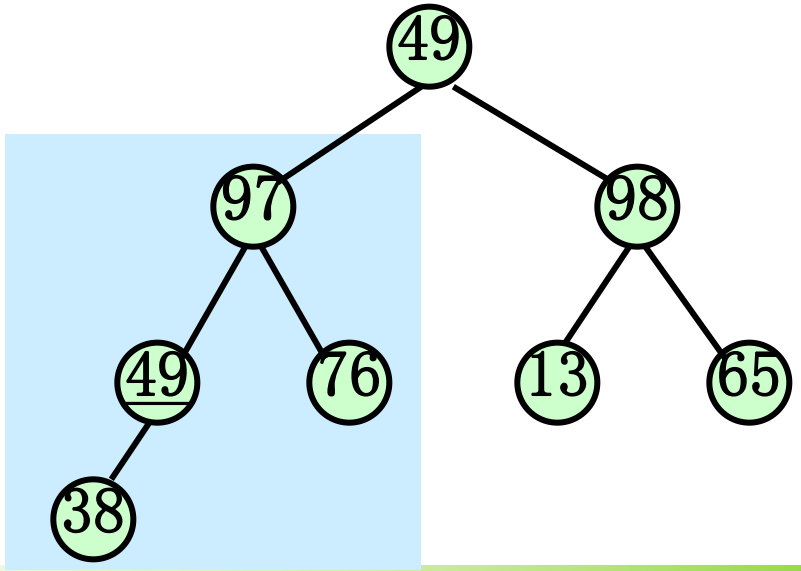
2。对以序号序号i=i-1的结点为双亲结点为根的二叉树，进行堆调整。

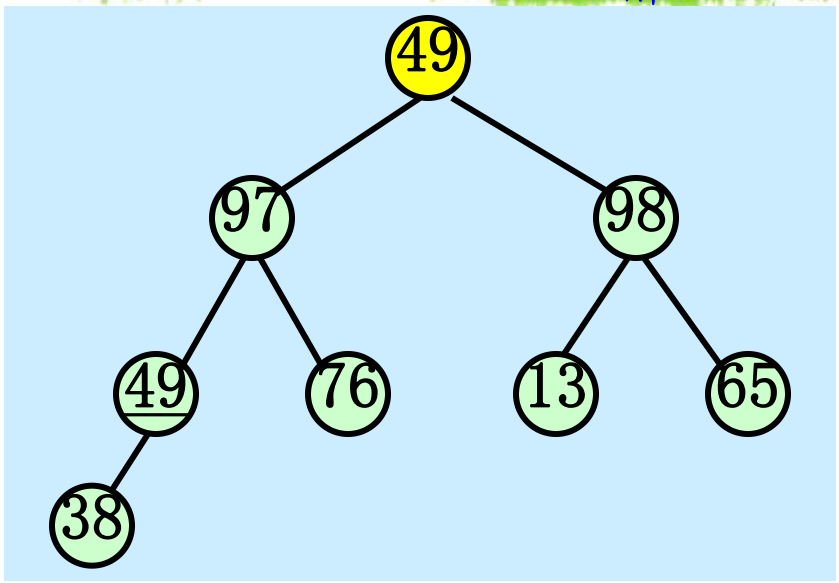
华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126



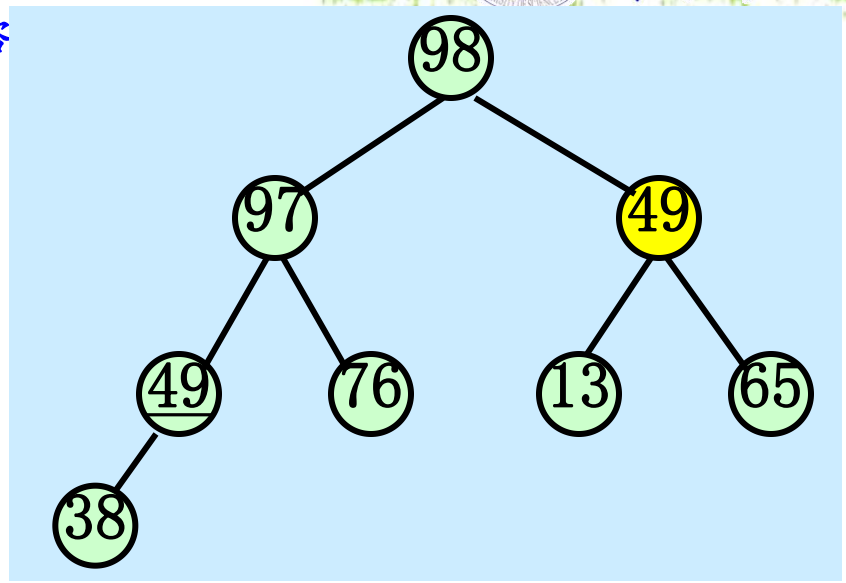


3。对以序号序号i=i-1的结点为双亲结点为根的二叉树，进行堆调整。

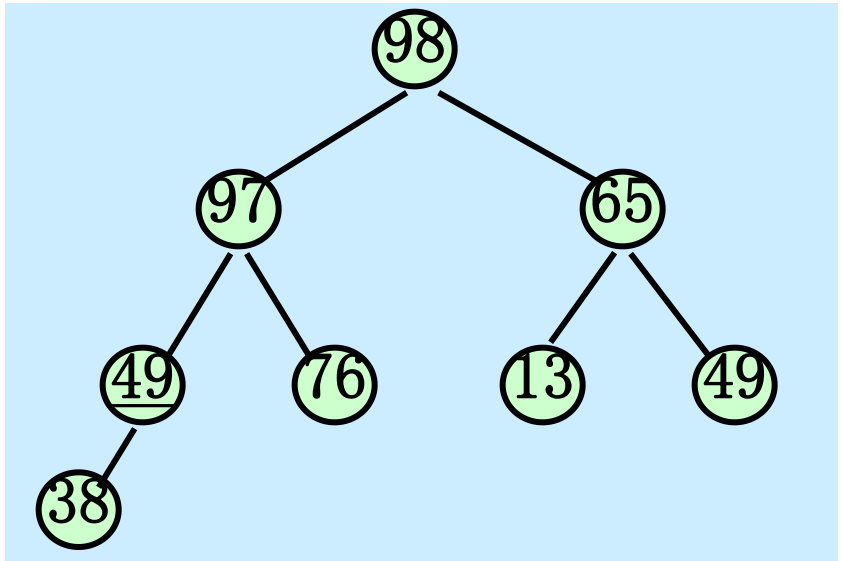


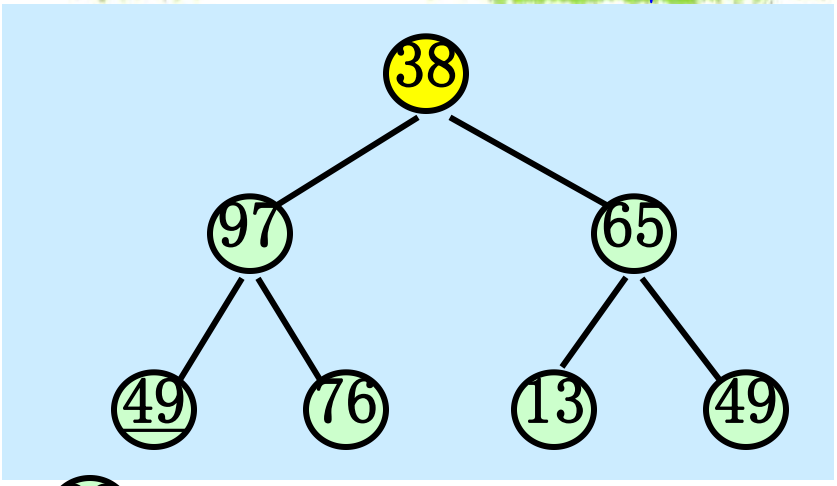


咨

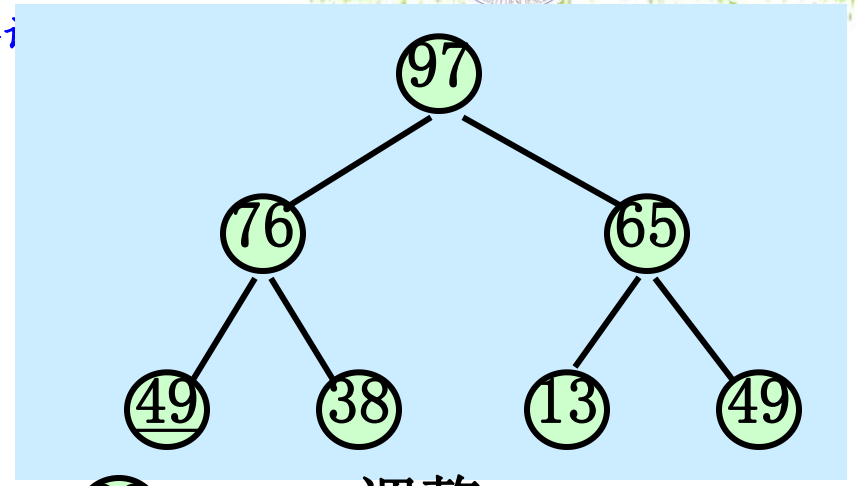


4。对以序号序号i=i-1的结点为双亲结点为根的二叉树，进行堆调整。此时i已等于1，调整完后，初始大顶堆建成。

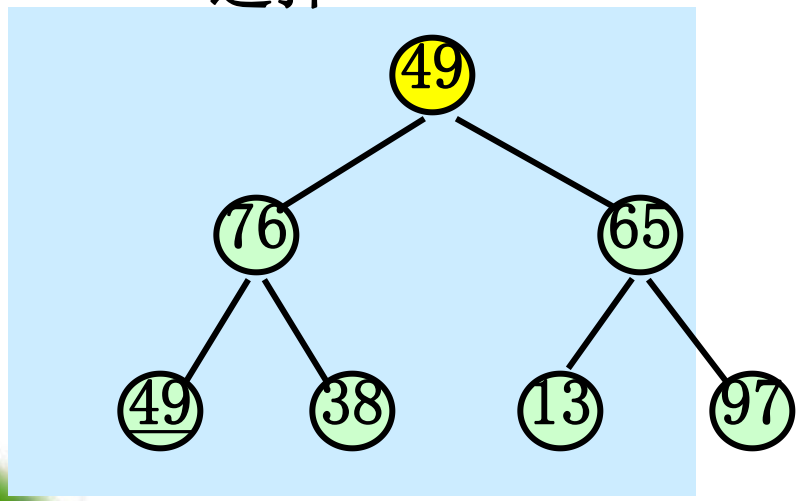




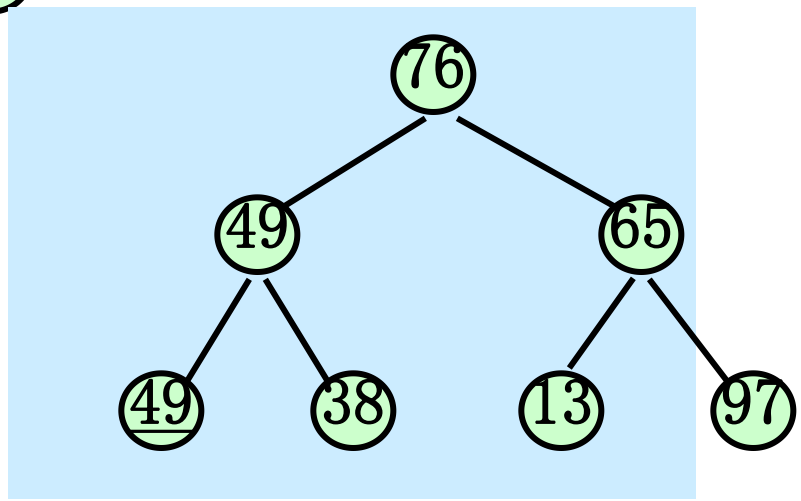
98 选择



98 调整

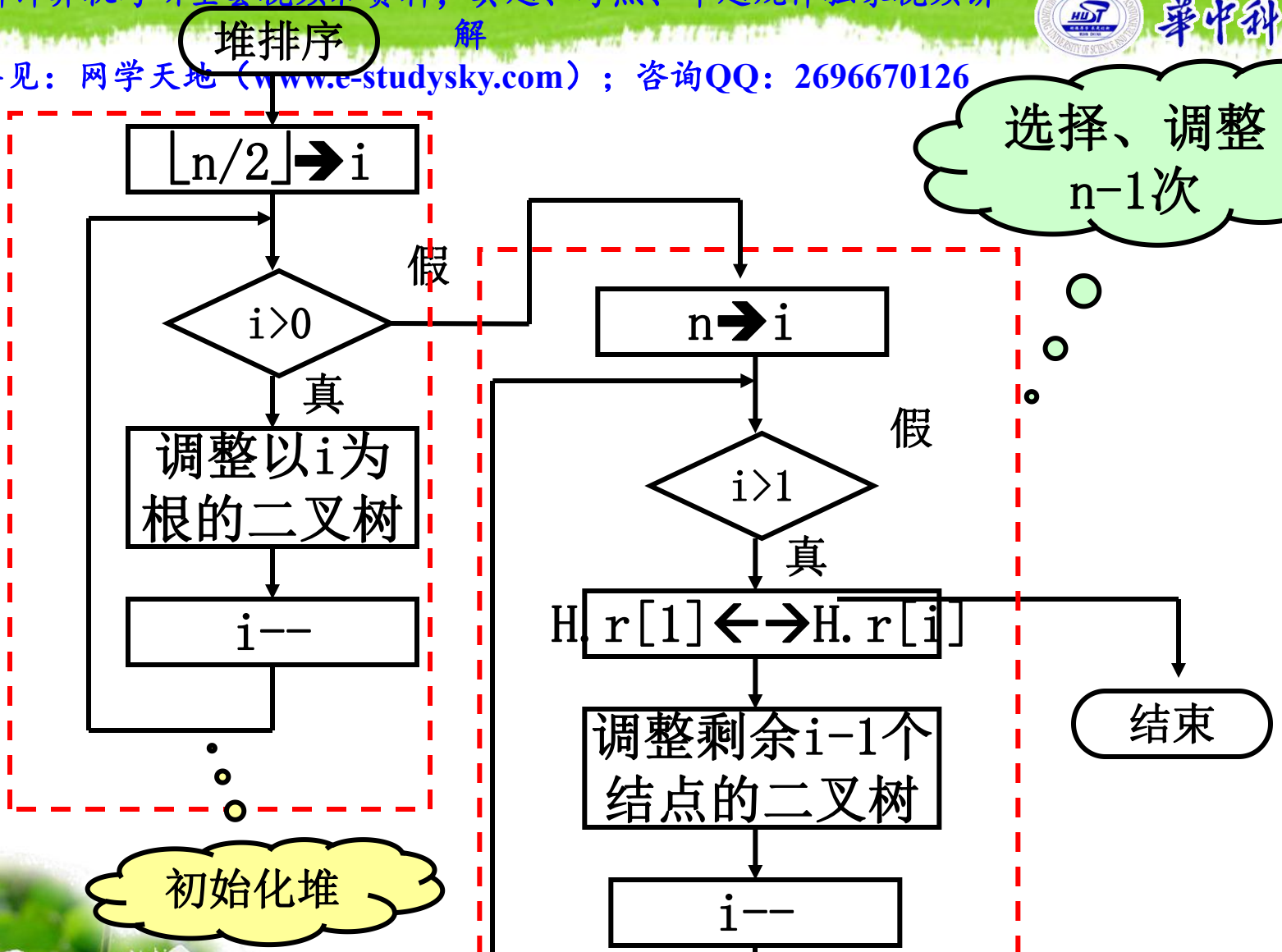


98 较小范围选择



98 较小范围调整

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126



华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com），咨询QQ：2696670126

算法分析与评价：

- 对深度为k的堆，调整算法进行的关键字比较次数至多为2(k-1)次，
- 建立n个元素的初始堆，调用HeapAdjust算法 $\lfloor n/2 \rfloor$次，总共进行的关键字比较次数不超过4n次。

n个结点的完全二叉树深度为$h=\lfloor \log_2 n \rfloor+1$

需要调整的层为h-1层至1层，以第i层某结点为根的二叉树深度对应为h-i+1,第i层结点最多为2^{i-1}个，故调整时比较关键字最多为：

$$\sum_{i=h-1}^{1} 2^{i-1}*2*(h-i+1-1)=\sum_{i=h-1}^{1} 2^{i}*(h-i)$$

令j=h-i，　当i=h-1时　j=1　　当i=1时　j=h-1

$$\sum_{j=1}^{h-1} 2^{h-j}*j=2^{h-1}*1+\ 2^{h-2}*2+\cdots+\ 2^{1}*(h-1)=2^{h+1}-2h-2$$

$$< \ 2^{h+1}\ =2^{\lfloor \log 2n \rfloor+2}<=4*\ 2^{\log_2 n}\ =4n$$

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.exstudy.com）；咨询QQ：2696670126

算法分析与评价(续)：

- n个结点的完全二叉树深度为$\lfloor \log_2 n \rfloor+1$，选择调整过程n-1次，总共比较次数至多为：

$$2\left(\lfloor \log_2(n-1) \rfloor+\lfloor \log_2(n-2) \rfloor+\cdots+\lfloor \log_2 2 \rfloor\right)<2n\left(\lfloor \log_2 n \rfloor\right)$$

- 最坏情况下，算法的时间复杂度为$O(4n+n\log n) \rightarrow O(n\log n)$；
- 仅需要一个记录大小供交换用的辅助存储空间；
- 堆排序是不稳定排序；
- 堆排序对记录较少的文件不提倡，对较大的文件很有效。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

10.5 归并排序

基本思想：把k(k≥2)个有序子文件合并在一起，形成一个新的有序文件。同时归并k个有序子文件的排序过程称为k-路归并排序。

2-路归并排序： 归并2个有序子文件的排序。

例. 将有序文件A和A归并为有序文件C。

A＝(2, 10, 15, 18, 21, 30)

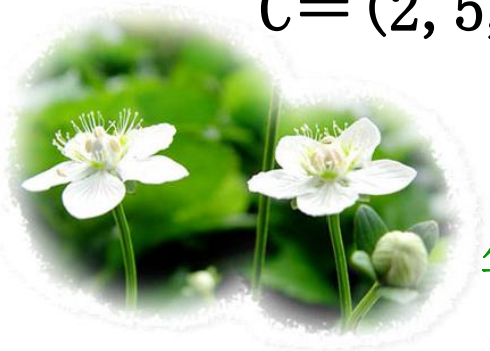
B＝(5, 20, 35, 40)

按从小至大的次序从A或B中依次取出

2, 5, 10, 15, . . . , 40,

顺序归并到C中, 得:

C＝(2, 5, 10, 15, 18, 20, 21, 30, 35, 40)



华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

一般地，2-路归并过程为：

假定文件r[low..high]中的相邻子文件(子表)(r[low],r[low+1],...,r[mid])和(r[mid+1],...,r[high])为有序子文件,其中:low≤mid<high 。

将这两个相邻有序子文件归并为有序文件y[low..high],即：(y[low],y[low+1],...,y[high])

low=9
mid=12
high=16

r[9]	r[10]	r[11]	r[12]	r[13]	r[14]	r[15]	r[16]
06	08	20	30	02	20	33	57

i　　j

y[9]	y[10]	y[11]	y[12]	y[13]	y[14]	y[15]	y[16]

k

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见考研天地 http://www.csstudy.com.cn 咨询QQ 2656670146

将两个有序子文件归并为有一个有序文件的算法

```
void merge(RecType r[], RecType y[],low,mid,high)
{
  int  k=i=low,j=mid+1;
  while (i<=mid && j<=high)
  { if (r[i].key<=r[j].key)
     { y[k]=r[i];      //归并前一个子文件的记录
       i++; }
    else
     { y[k]=r[j];      //归并后一个子文件的记录
       j++; }
    k++;
  }
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

```
while (j<=high)      //归并后一个子文件余下的记录
    { y[k]=r[j];
      j++;  k++;
     }
while (i<=mid)       //归并前一个子文件余下的记录
   { y[k]=r[i];
     i++;  k++;
    }
 } // merge
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

2-路归并排序

假定文件(r[1],r[2],...,r[n])中记录是随机排列的,进行2-路归并排序，首先把它划分为长度均为1的n个有序子文件，然后对它们逐步进行2一路归并排序。其步骤如下：

第1趟：从r[1..n]中的第1个和第2个有序子文件开始,调用算法merge,每次归并两个相邻子文件,归并结果放到y[1..n]中。在y中形成 ⌈n/2⌉ 个长度为2的有序子文件。若n为奇数,则y中最后一个子文件的长度为1。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

第2趟：把y[1..n]看作输入文件,将⌈n/2⌉个有序子文件两两归并,归并结果回送到r[1..n]中,在r中形成⌈⌈n/2⌉/2⌉个长度为4的有序子文件。若y中有奇数个子文件，则r中最后一个子文件的长度为2。

……

共计经过⌈log2n⌉趟归并,最后得到n个记录的有序文件。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

例1. 对8个记录作2-路归并排序，第1趟归并：

r[1]	r[2]	r[3]	r[4]	r[5]	r[6]	r[7]	r[8]
06	44	20	10	02	20	08	07

y[1]	y[2]	y[3]	y[4]	y[5]	y[6]	y[7]	y[8]

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

例1. 对8个记录作2-路归并排序，第2趟归并：

r[1]	r[2]	r[3]	r[4]	r[5]	r[6]	r[7]	r[8]

y[1]	y[2]	y[3]	y[4]	y[5]	y[6]	y[7]	y[8]
06	44	10	20	02	20	07	08

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

例1. 对8个记录作2-路归并排序，第3趟归并：

r[1]	r[2]	r[3]	r[4]	r[5]	r[6]	r[7]	r[8]
06	10	20	44	02	07	08	20

y[1]	y[2]	y[3]	y[4]	y[5]	y[6]	y[7]	y[8]

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

例1. 第4趟归并，共$\lceil \log_2 8 \rceil$+1=4趟。

r[1]	r[2]	r[3]	r[4]	r[5]	r[6]	r[7]	r[8]

y[1]	y[2]	y[3]	y[4]	y[5]	y[6]	y[7]	y[8]
02	06	07	08	10	20	20	44

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

例2. 对11个记录作2-路归并排序，第1趟归并。

r[1]	r[2]	r[3]	r[4]	r[5]	r[6]	r[7]	r[8]	r[9]	r[10]	r[11]
06	44	20	10	02	20	08	07	05	32	14

y[1]	y[2]	y[3]	y[4]	y[5]	y[6]	y[7]	y[8]	y[9]	y[10]	y[11]

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

例2. 对11个记录作2-路归并排序，第2趟归并。

r[1]	r[2]	r[3]	r[4]	r[5]	r[6]	r[7]	r[8]	r[9]	r[10]	r[11]
06	44	20	10	02	20	08	07	05	32	14`

y[1]	y[2]	y[3]	y[4]	y[5]	y[6]	y[7]	y[8]	y[9]	y[10]	y[11]
06	44	10	20	02	20	07	08	05	32	14

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

例2. 对11个记录作2-路归并排序，第3趟归并。

r[1]	r[2]	r[3]	r[4]	r[5]	r[6]	r[7]	r[8]	r[9]	r[10]	r[11]
06	10	20	44	02	07	08	20	05	14	32

y[1]	y[2]	y[3]	y[4]	y[5]	y[6]	y[7]	y[8]	y[9]	y[10]	y[11]
02 06	07 10	08 20	20 44		20			05	14 32	

例2. 对11个记录作2-路归并排序，进行第$\lceil \log_2 11 \rceil = 4$趟归并。

r[1]	r[2]	r[3]	r[4]	r[5]	r[6]	r[7]	r[8]	r[9]	r[10]	r[11]
02	06	07	08	10	20	20	44	05	14	32

一般情况，需要进行「 （$\lceil \log_2 n \rceil$ +1）/2 」*2 趟归并。

y[1]	y[2]	y[3]	y[4]	y[5]	y[6]	y[7]	y[8]	y[9]	y[10]	y[11]
02	06	07	08	10	20	20	44	05	14	32

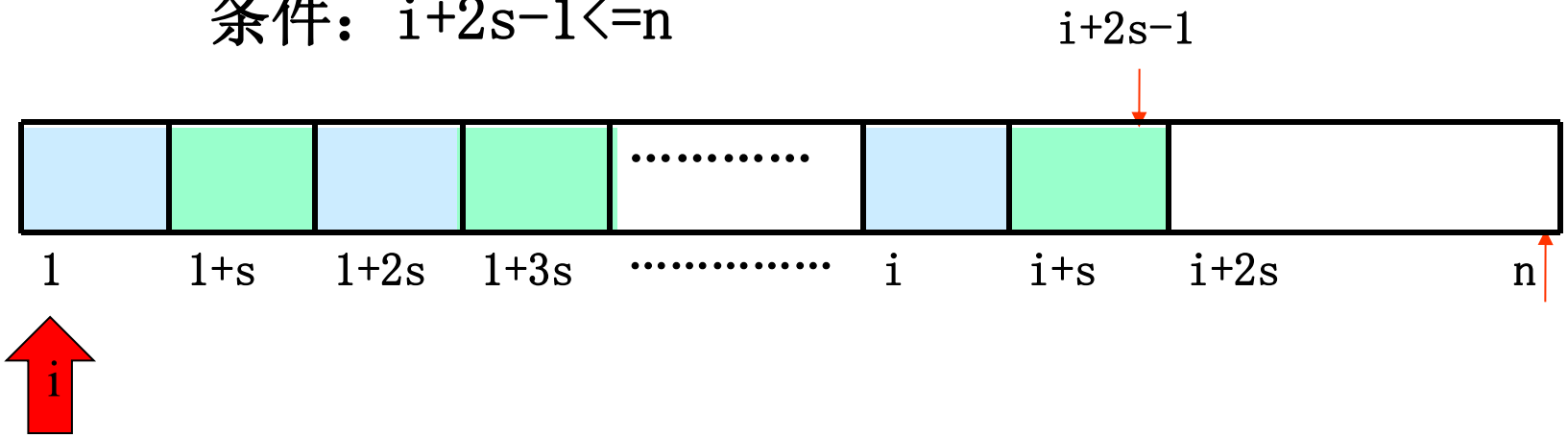
华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

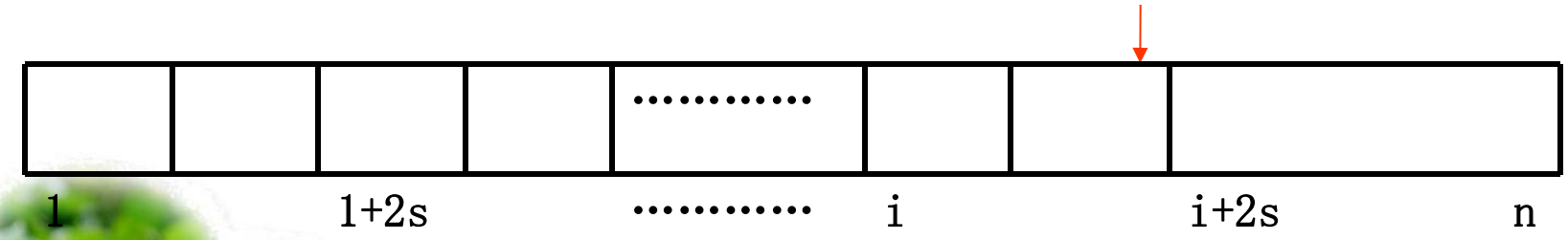
一趟归并排序算法：

假设：s为子文件的长度,将r中的子文件归并到y中

（1） 两等长子文件合并

条件：i+2s-1<=n

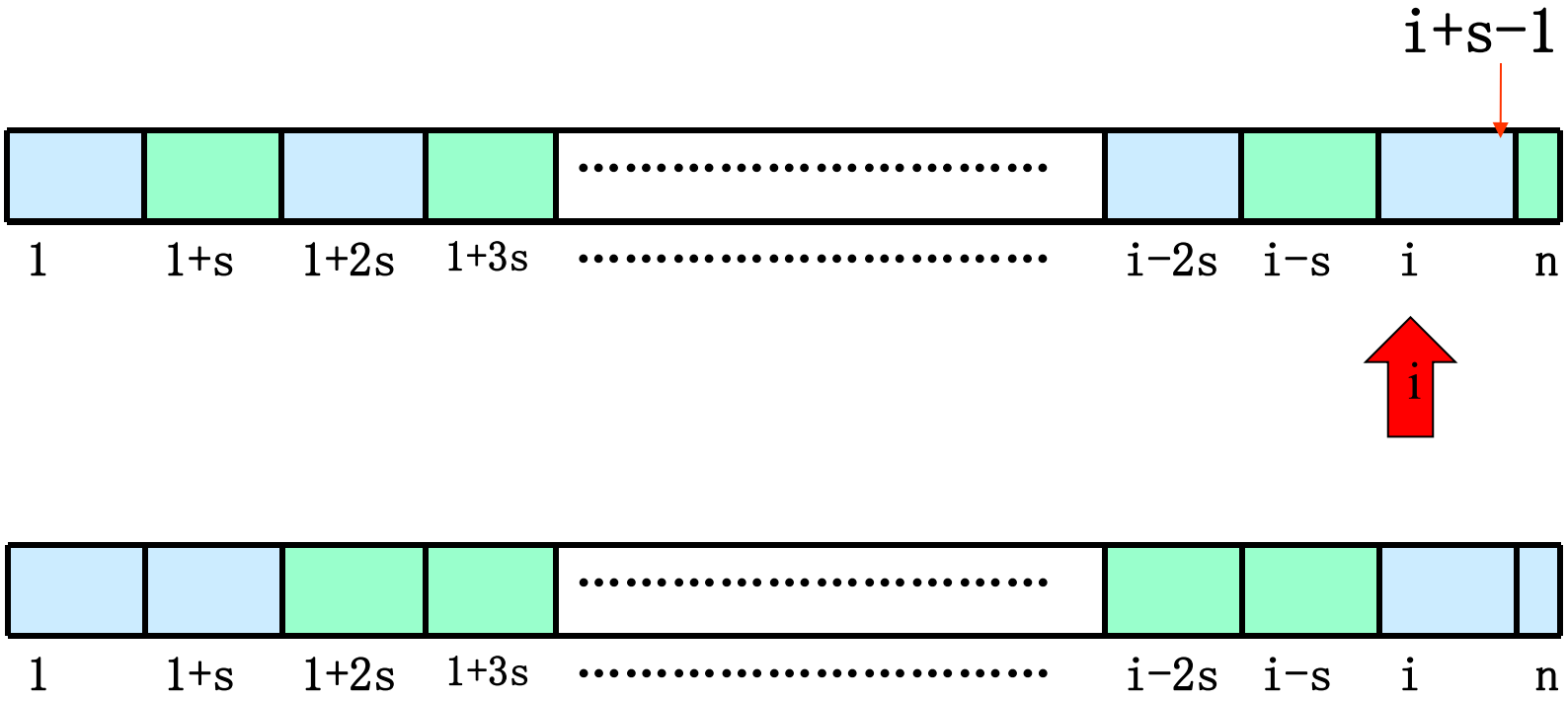




华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）：查询QQ：2696670126

（2）长度为S的子文件与长度为[1，s)子文件合并

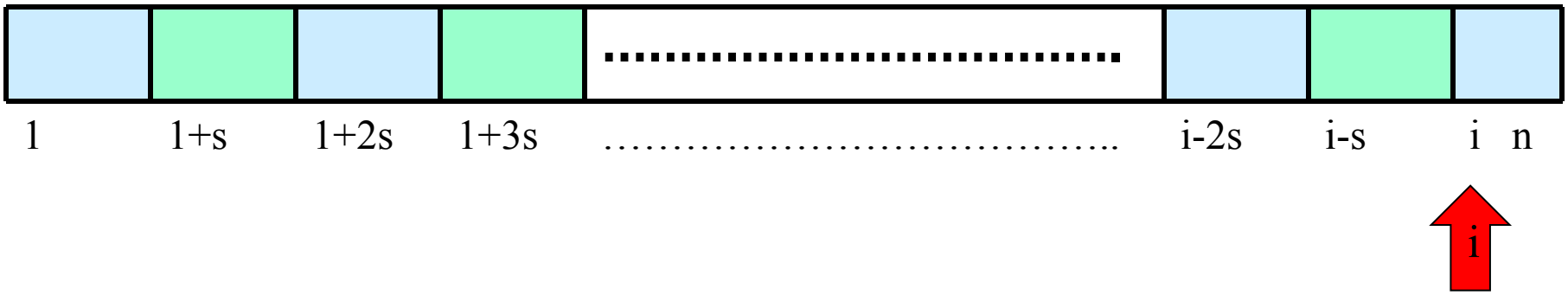
条件：　!(i+2s-1<=n)　并且　i+s-1<n

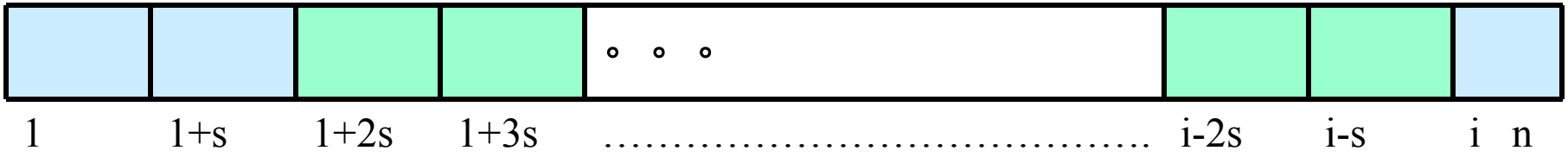


华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

（3） 剩余一个长度为［1，s］子文件时，直接复制。

条件： i+s-1>=n





华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.studysky.com）；咨询QQ：2696670126

一趟归并排序算法：

```
void mergepass(RecType r[], RecType y[],int s,int n)
{ int i=1;
  while(i+2*s-1<=n)          //两两归并长度均为s的子文件
   { merge(r,y,i,i+s-1,i+2*s-1);
     i=i+2*s; }
  if (i+s-1<n)    //最后两个子长度为s和长度不足s的文件
    merge(r,y,i,i+s-1,n);
  else
    while(i<=n)              //复制最后一个子文件,长度≤s
     { y[i]=r[i];
       i++; }
}
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

对文件r[1..n]归并排序的算法(调用算法mergepass)

```
void mergesort(RecType r[],int n)
{
  RecType y[n+1] ;
  int s=1;                          //子文件初始长度为1
  while (s<n)
   { mergepass(r,y,s);              //将r[1..n]归并到y[1..n]
     s=2*s;                         //修改子文件长度
     mergepass(y,r,s);              //将y[1..n]归并到r[1..n]
     s=2*s;                         //修改子文件长度
   }
}
```

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

算法分析

- 对n个记录的文件进行归并排序，共需⌈$\log_2 n$⌉趟，每趟所需比较关键字的次数不超过n，共比较$O(n\log_2 n)$次。
- 每趟移动n个记录，共移动$O(n\log_2 n)$个记录。
- 归并排序需要一个大小为n的辅助空间y[1..n]。
- 归并排序是稳定的。

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

10.6 基数排序

前面的各种排序算法中，需要进行关键字间的比较。基数排序不同于前面的各种算法，仅通过分析关键字自身每位的值来实现排序。

算法假定记录的关键字为整型（实际并不限制）。

基数排序有两种方法：

（1）首先根据最高有效位进行排序，然后根据次高有效位进行排序，直到根据最低有效位进行排序，产生一个有序序列。

012	0	015	15	036			321	3	325	3	326	03	710			717	9	780		789

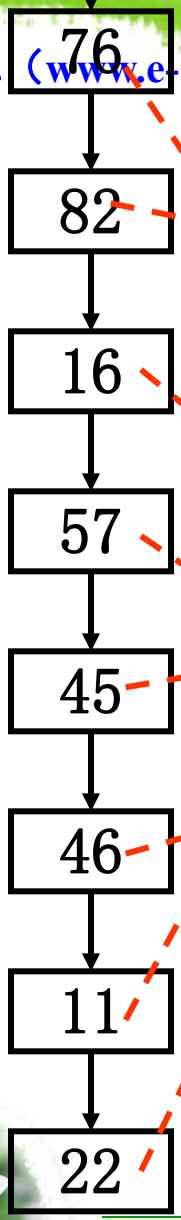
（2）首先根据最低有效位进行排序，然后根据次低有效位进行排序，直到根据最高有效位进行排序，产生一个有序序列。

710	012	015	717	321	325	036	356	780	789

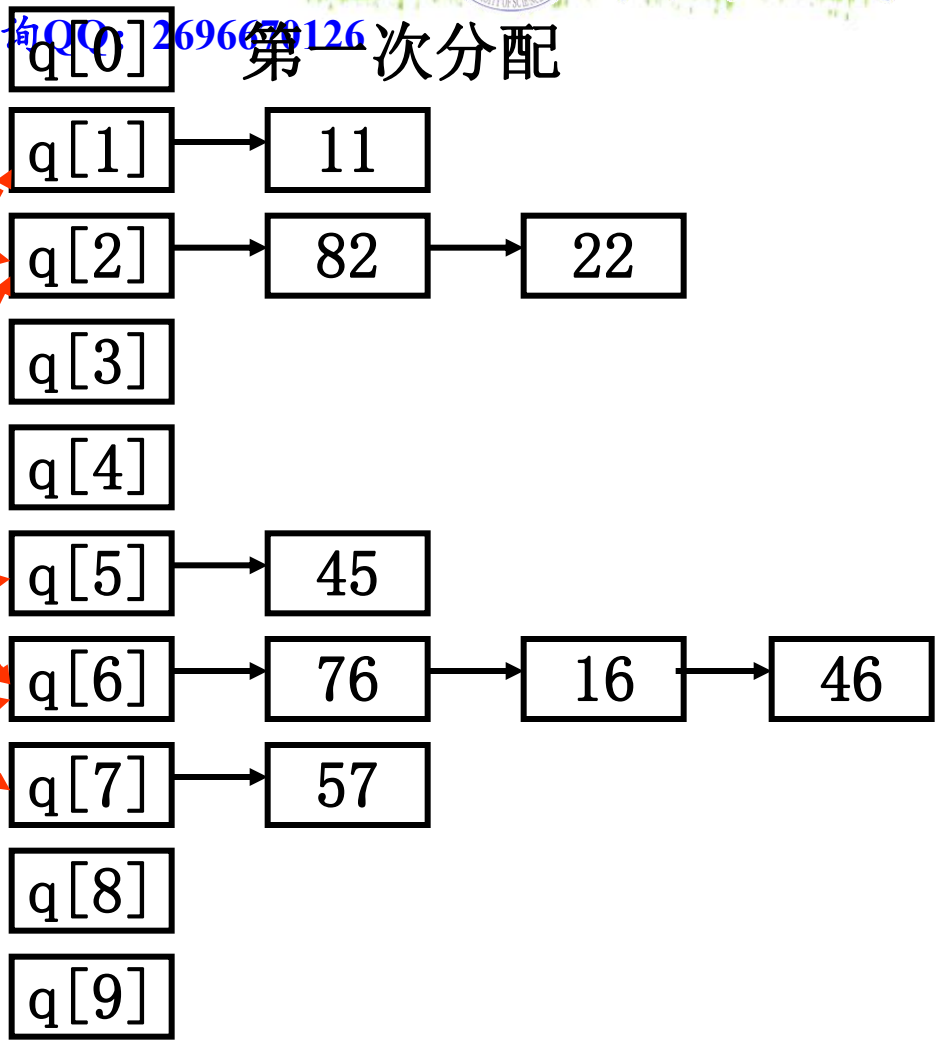
华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

链式基数排序

原始序列

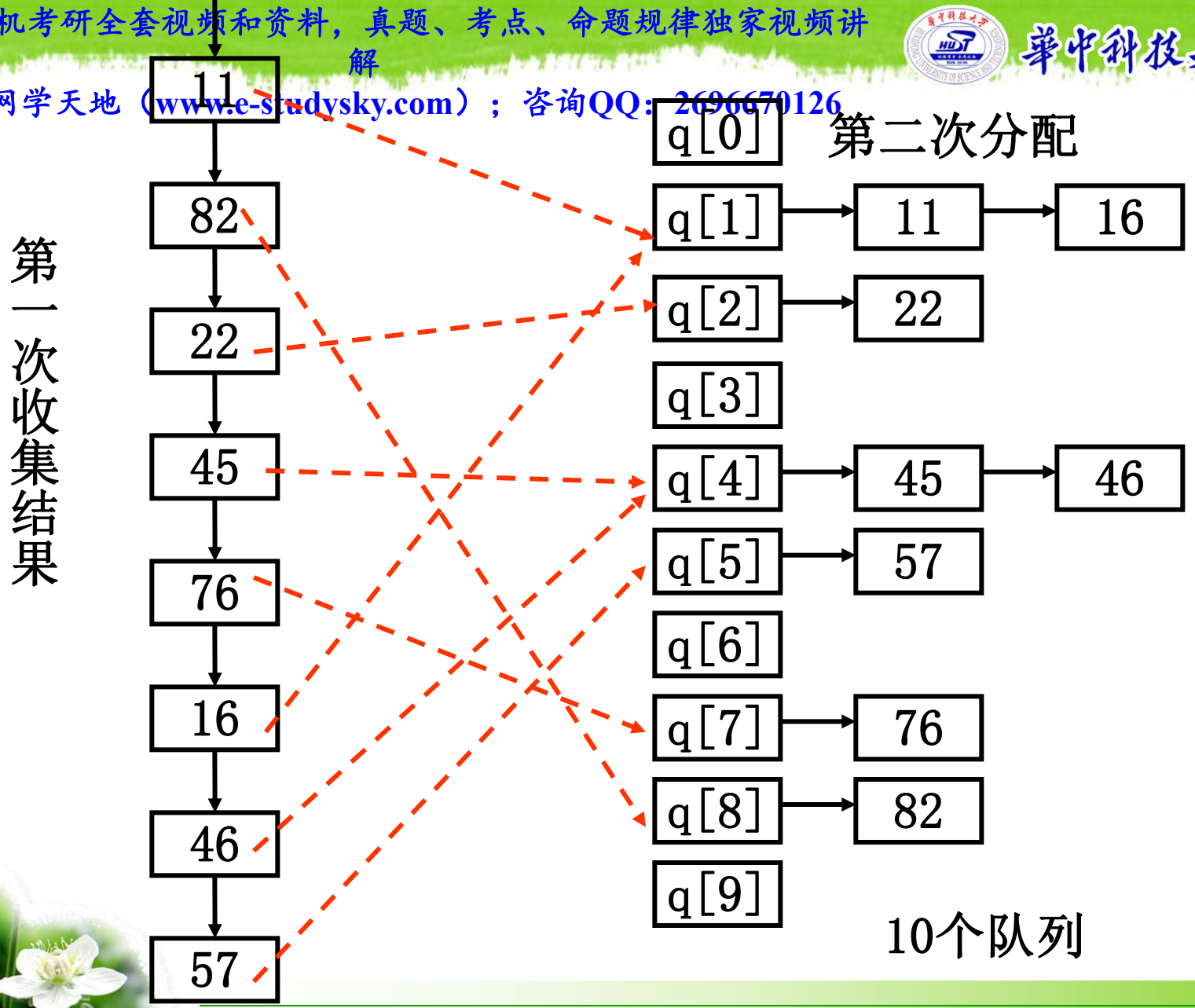


第一次分配



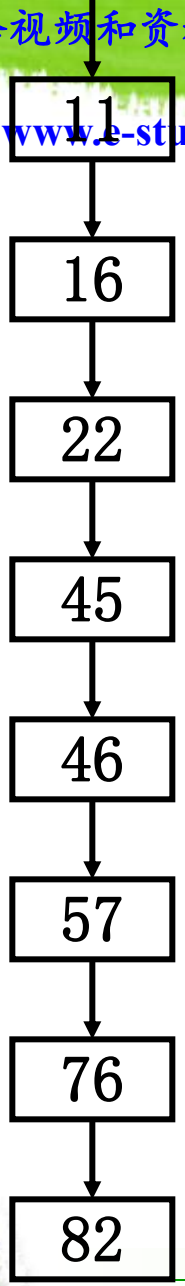
10个队列

华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126



华科计算机考研全套视频和资料，真题、考点、命题规律独家视频讲解
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

第二次收集结果



算法分析：

（1）设数字有效位为d位，需要d趟分配、回收。

（2）每趟分配运算时间O(n)

（3）收集：基数为rd,即rd个队列。从rd个队列中收集，运算时间O(rd)

（4）一趟分配、回收运算时间O(n+rd)， 时间复杂度O(d*(n+rd))

（5）基数排序是稳定的。

（6）辅助空间：每个队列首尾2个指针，共2rd个指针；n个记录需要n个指针域。