

Chapter9 排序

概述

排序：将一组杂乱无章的数据按一定的规律顺次排列起来。

数据表(datalist): 它是待排序数据元素的有限集合。

排序码(key): 通常数据元素有多个属性域, 即多个数据成员组成, **其中有一个属性域可用来区分元素, 作为排序依据。该域即为排序码。**每个数据表用哪个属性域作为排序码, 要视具体的应用需要而定。

排序算法的稳定性: 如果在元素序列中有两个元素 $r[i]$ 和 $r[j]$, 它们的排序码 $k[i] == k[j]$, 且在排序之前, 元素 $r[i]$ 排在 $r[j]$ 前面。如果在排序之后, 元素 $r[i]$ 仍在元素 $r[j]$ 的前面, 则称这个排序方法是稳定的, 否则称这个排序方法是不稳定的。

内排序与外排序: 内排序是指在排序期间数据元素全部存放在内存的排序; 外排序是指在排序期间全部元素个数太多, 不能同时存放在内存, 必须根据排序过程的要求, 不断在内、外存之间移动的排序。

排序的时间开销: 排序的时间开销是衡量算法好坏的最重要的标志。排序的时间开销可用算法执行中的**数据比较次数与数据移动次数**来衡量。

算法运行时间代价的大略估算一般都按平均情况进行估算。对于那些受元素排序码序列初始排列及元素个数影响较大的, 需要按**最好情况和最坏情况**进行估算。

算法执行时所需的附加存储: 评价算法好坏的另一标准。

```
1  template <class T>
2  class dataList {                //数据表类定义
3  private:
4      Element<T>* Vector;         //存储排序元素的向量
5      int maxSize;                //向量中最大元素个数
6      int currentSize;            //当前元素个数
7  }
```

9.1 插入排序 (Insert Sorting)

9.1.1 直接插入排序 (Insert Sort)

```
1  template <class T>
2  void InsertSort (dataList<T>& L, int left, int right) {
3      //依次将元素L.Vector[i]按其排序码插入到有序表
4      //L.Vector[left],...,L.Vector[i-1]中,使得
5      //L.Vector[left]到L.Vector[i]有序。
6      Element<T> temp;  int i, j;
7      for (i = left + 1; i <= right; i++)
8          if (L[i] < L[i-1]) {
9              temp = L[i];  j = i-1;
10             do {
11                 L[j+1] = L[j];  j--;
12             } while (j >= left && temp < L[j]);
13             L[j+1] = temp;
14         }
15     };
```

算法分析

设待排序元素个数为currentSize = n，则该算法的主程序执行n-1趟。

排序码比较次数和元素移动次数与元素排序码的初始排列有关。

最好情况下，排序前元素已按排序码从小到大有序，每趟只需与前面有序元素序列的最后一个元素比较1次，总的排序码比较次数为 n-1，元素移动次数为0。

最坏情况下，第 i 趟时第 i 个元素必须与前面 i 个元素都做排序码比较，并且每做1次比较就要做1次数据移动。则总排序码比较次数KCN和元素移动次数RMN分别为

$$KCN = \sum_{i=1}^{n-1} i = n(n-1)/2 \approx n^2/2,$$
$$RMN = \sum_{i=1}^{n-1} (i+2) = (n+4)(n-1)/2 \approx n^2/2$$

9.1.2 折半插入排序 (Binary Insertsort)

基本思想是：设在顺序表中有一个元素序列 V[0], V[1], ..., V[n-1]。其中, V[0], V[1], ..., V[i-1] 是已经排好序的元素。在插入V[i] 时, 利用折半搜索法寻找V[i] 的插入位置。

```
1  template <class T>
2  void BinaryInsertSort (dataList<T>& L, const int left, const int right) {
3  //利用折半搜索，在L.Vector[left]到L.Vector[i-1]中
4  //查找L.Vector[i]应插入的位置，再进行插入。
5      Element<T> temp;
6      int i, low, high, middle, k;
7      for (i = left + 1; i <= right; i++)
8      {
9          temp = L[i]; low = left; high = i-1;
10         while (low <= high)
11         {
12             middle = (low + high)/2; //折半搜索插入位置
13             if (temp < L[middle]) //取中点
14                 high = middle - 1; //插入值小于中点值
15             else low = middle + 1; //向左缩小小区间
16             //否则，向右缩小小区间
17         }
18         for (k = i-1; k >= low; k--) L[k+1] = L[k]; //成块移动，空出插入位置
19         L[low] = temp; //插入
20     }
21 }
```

它所需的排序码比较次数与待排序元素序列的初始排列无关，仅依赖于元素个数。在插入第 i 个元素时，需要经过 $\log_2 i + 1$ 次排序码比较，才能确定它应插入的位置。因此，将 n 个元素(为推导方便，设为 $n=2k$) 用折半插入排序所进行的排序码比较次数为：

$$\sum_{i=1}^{n-1} (\lfloor \log_2 i \rfloor + 1) \approx n \cdot \log_2 n$$

当 n 较大时，总排序码比较次数比直接插入排序的最坏情况要好得多，但比其最好情况要差。

在元素的初始排列已经按排序码排好序或接近有序时，直接插入排序比折半插入排序执行的排序码比较次数要少。折半插入排序的元素移动次数与直接插入排序相同，依赖于元素的初始排列。

9.1.3 希尔排序 (Shell Sort)

希尔排序方法又称为缩小增量排序。该方法的基本思想是：

1) 设待排序元素序列有 n 个元素, 首先取一个整数 $gap < n$ 作为间隔, 将全部元素分为 gap 个子序列, 所有距离为 gap 的元素放在同一个子序列中。

2) 在每一个子序列中分别施行直接插入排序。

3) 然后缩小间隔 gap , 例如取 $gap = \lceil gap/2 \rceil$, 重复上述的子序列划分和排序工作。直到最后取 $gap == 1$, 将所有元素放在同一个序列中排序为止。

开始时 gap 的值较大, 子序列中的元素较少, 排序速度较快; 随着排序进展, gap 值逐渐变小, 子序列中元素个数逐渐变多, 由于前面工作的基础, 大多数元素已基本有序, 所以排序速度仍然很快。

```
1 void shellSort (dataList<T>& L, const int left, const int right)
2 {
3     int i, j, gap = right-left+1; //增量的初始值
4     Element<T> temp;
5     do {
6         gap = gap/2; //求下一增量值
7         for (i = left + gap; i <= right; i++)
8             if (L[i] < L[i - gap]) //逆序
9                 {
10                     temp = L[i]; j = i-gap;
11                     do {
12                         L[j + gap] = L[j]; j = j - gap;
13                     } while (j >= left && temp < L[j]);
14                     L[j + gap] = temp; //将vector[i]回送
15                 }
16     } while (gap > 1);
17 };
```

Gap的取法有多种。最初 shell 提出取 $gap = n/2$, $gap = gap/2$, 直到 $gap = 1$ 。knuth 提出取 $gap = gap/3 + 1$ 。还有人提出都取奇数为好, 也有人提出各 gap 互质为好。

对特定的待排序元素序列, 可以准确地估算排序码的比较次数和元素移动次数。

9.2 交换排序 (Exchange Sort)

基本思想是两两比较待排序元素的排序码, 如果发生逆序(即排列顺序与排序后的次序正好相反), 则交换之。直到所有元素都排好序为止。

9.2.1 起泡排序 (Bubble Sort)

基本方法是: 设待排序元素序列中的元素个数为 n 。最多作 $n-1$ 趟, $i = 1, 2, \dots, n-1$ 。在第 i 趟中从后向前, $j = n-1, n-2, \dots, i$, 顺次两两比较 $V[j-1].key$ 和 $V[j].key$ 。如果发生逆序, 则交换 $V[j-1]$ 和 $V[j]$ 。

```
1 template <class T>
2 void BubbleSort (dataList<T>& L, const int left,
3                 const int right) {
4     int pass = left+1, exchange = 1;
5     while (pass <= right && exchange) {
6         exchange = 0; //标志为0假定未交换
```

```

7      for (int j = right; j >= pass; j--)
8          if (L[j-1] > L[j]) { //逆序
9              swap (L[j-1], L[j]); //交换
10             exchange = 1; //标志置为1,有交换
11         }
12         pass++;
13     } // end of while
14 };

```

第 i 趟对待排序元素序列 $V[i-1], V[i], \dots, V[\text{right}]$ 进行排序, 结果将该序列中排序码最小的元素交换到序列的第一个位置($i-1$)。

最多做 $n-1$ 趟起泡就能把所有元素排好序。

在元素的初始排列已经按排序码从小到大排好序时, 此算法只执行一趟起泡, 做 $n-1$ 次排序码比较, 不移动元素。这是最好的情形。

最坏的情形是算法执行 $n-1$ 趟起泡, 第 i 趟 ($1 \leq i < n$) 做 $n-i$ 次排序码比较, 执行 $n-i$ 次元素交换。在最坏情形下总的排序码比较次数 KCN 和元素移动次数 RMN 为:

$$KCN = \sum_{i=1}^{n-1} (n-i) = \frac{1}{2} n(n-1)$$

$$RMN = 3 \sum_{i=1}^{n-1} (n-i) = \frac{3}{2} n(n-1)$$

起泡排序需要一个附加元素以实现元素值的对换。

起泡排序是一个稳定的排序方法。

9.2.2 快速排序 (Quick Sort)

基本思想是任取待排序元素序列中的某个元素 (例如取第一个元素) 作为基准, 按照该元素的排序码大小, 将整个元素序列划分为左右两个子序列:

左侧子序列中所有元素的排序码都小于或等于基准元素的排序码

右侧子序列中所有元素的排序码都大于基准元素的排序码

基准元素则排在这两个子序列中间(这也是该元素最终应安放的位置)。

然后分别对这两个子序列重复施行上述方法, 直到所有的元素都排在相应位置上为止。

算法描述

```

QuickSort ( List ) {
    if ( List的长度大于1 ) {
        将序列 List 划分为两个子序列
            LeftList 和 Right List;
        QuickSort ( LeftList );
        QuickSort ( RightList );
        将两个子序列 LeftList 和 RightList
            合并为一个序列 List;
    }
}

```

```

1  template <class T>

```

```

2  int dataList<T>::Partition (const int low, const int high) {
3  //数据表类的共有函数
4      int pivotpos = low;
5      Element<T> pivot = Vector[low];          //基准元素
6      for (int i = low + 1; i <= high; i++)//检测整个序列，进行划分
7          if (Vector[i] < pivot) {
8              pivotpos++;
9              if (pivotpos != i)
10                 Swap(Vector[pivotpos],Vector[i]);
11         }          //小于基准的交换到左侧去
12     Vector[low] = Vector[pivotpos];
13     Vector[pivotpos] = pivot;//将基准元素就位
14     return pivotpos;          //返回基准元素位置
15 };

```

Partition补充算法

```

1  int dataList<T>::Partition (int s, int t) {
2      int x = a[s]; i=s; j=t;
3      while (i<j)
4      {
5          while ((i<j) &&(a[j]>=x)) {j=j-1; }
6          if (i<j) {a[i]=a[j];}
7
8          while ((i<j) &&(a[i]<=x)) {i=i+1; }
9          if (i<j) {a[j]=a[i];}
10     }
11     a[i] = x; return i;
12 }

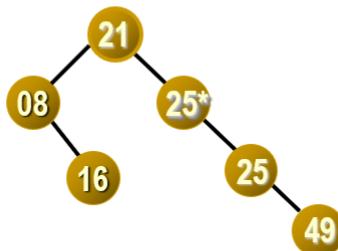
```

```

1  template <class T>
2  void QuickSort (dataList<T>& L, const int left, const int right) {
3      //对元素Vector[left], ..., Vector[right]进行排序,
4      //pivot=L.Vector[left]是基准元素，排序结束后它的
5      //位置在pivotPos，把参加排序的序列分成两部分，
6      //左边元素的排序码都小于或等于它，右边都大于它
7      if (left < right) {          //元素序列长度大于1时
8          int pivotpos = L.Partition (left, right); //划分
9          QuickSort (L, left, pivotpos-1);
10         QuickSort (L, pivotpos+1, right);
11     }
12 };

```

- 算法quicksort是一个递归的算法, 其递归树如图所示。



- 从快速排序算法的递归树可知，快速排序的趟数取决于递归树的高度。

最大递归调用层数与递归树高度一致，理想情况为 $\text{Up_Edge}(\log_2(n+1))$ 。存储开销为 $O(\log_2 n)$ 。

在最坏的情况，即待排序元素序列已经按其排序码从小到大排好序的情况下，其递归树成为单支树，每次划分只得到一个比上一次少一个元素的子序列。必须经过 $n-1$ 趟才能把所有元素定位，而且第 i 趟需要经过 $n-i$ 次排序码比较才能找到第 i 个元素的安放位置，总的排序码比较次数将达到

$$\sum_{i=1}^{n-1} (n-i) = \frac{1}{2} n(n-1) \approx \frac{n^2}{2}$$

快速排序是一种不稳定的排序方法。

对于 n 较大的平均情况而言，快速排序是“快速”的，但是当 n 很小时，这种排序方法往往比其它简单排序方法还要慢

因此，当 n 很小时可以用直接插入排序方法

9.3 选择排序 (Select Sort)

基本思想是：每一趟（例如第 i 趟， $i = 0, 1, \dots, n-2$ ）在后面 $n-i$ 个待排序元素中选出排序码最小的元素，作为有序元素序列的第 i 个元素。待到第 $n-2$ 趟作完，待排序元素只剩下1个，就不用再选了。

9.3.1 直接选择排序 (Select Sort)

直接选择排序的基本步骤是：

①在一组元素 $V[i] \sim V[n-1]$ 中选择具有最小排序码的元素；

②若它不是这组元素中的第一个元素，则将它与这组元素中的第一个元素对调；

在这组元素中剔除这个具有最小排序码的元素。在剩下的元素 $V[i+1] \sim V[n-1]$ 中重复执行第①、②步，直到剩余元素只有一个为止

```
1  #include "dataList.h"
2  template <class T>
3  void SelectSort (dataList<T>& L, const int left, const int right) {
4      for (int i = left; i < right; i++)
5      {
6          int k = i;          //在L[i]到L[n-1]找最小排序码元素
7          for (int j = i+1; j <= right; j++)
8              if (L[j] < L[k]) k = j;
9          if (k != i) swap (L[i], L[k]); //交换
10     }
11 }
```

直接选择排序的排序码比较次数 KCN 与元素的初始排列无关。设整个待排序元素序列有 n 个元素，则第 i 趟选择具有最小排序码元素所需的比较次数总是 $n-i-1$ 次。总的排序码比较次数为

$$KCN = \sum_{i=0}^{n-2} (n-i-1) = \frac{n(n-1)}{2}$$

元素移动次数与元素序列初始排列有关。

当这组元素初始状态是按其排序码从小到大有序的时候，元素的移动次数达到最少 $RMN = 0$ ，

最坏情况是每一趟都要进行交换，总的元素移动次数为 $RMN = 3(n-1)$ 。

直接选择排序是一种不稳定的排序方法。

9.3.2 锦标赛排序 (Tournament Tree Sort)

- 它的思想与体育比赛时的淘汰赛类似。首先取得 n 个元素的排序码，进行两两比较，得到 $\lceil n/2 \rceil$ 个比较的优胜者(排序码小者)，作为第一步比较的结果保留下来。然后对这 $\lceil n/2 \rceil$ 个元素再进行排序码的两两比较, ..., 如此重复，直到选出一个排序码最小的元素为止。
- 由于每次两两比较的结果是把排序码小者作为优胜者上升到双亲结点，所以称这种比赛树为**胜者树**。

胜者树是完全二叉树, 其高度为 $\text{up_limit}(\log_2 n)$, 其中 n 为待排序元素个数。

除第一次选择具有最小排序码的元素需要进行 $n-1$ 次排序码比较外, 重构胜者树选择次小、再次小排序码所需的比较次数均为 $O(\log_2 n)$ 。总排序码比较次数为 $O(n \log_2 n)$ 。

n 这种排序方法虽然减少了许多排序时间, 但是使用了较多的附加存储。

n 如果有 n 个元素，必须使用至少 $2n-1$ 个结点来存放胜者树。

锦标赛排序是一个稳定的排序方法。

9.3.3 堆排序 (Heap Sort)

利用堆及其运算, 可以很容易地实现选择排序的思路。

堆排序分为两个步骤

- 1) 根据初始输入数据，利用堆的调整算法 `siftDown()` 形成初始堆;
- 2) 通过一系列的元素交换和重新调整堆进行排序。

为了实现元素按排序码从小到大排序，要求建立最大堆。

最大堆的向下调整算法

```

1  template <class T>
2  siftDown (dataList<T>& L, const int start, const int m){
3  //私有函数：从结点start开始到m自上向下比较，
4  //如果子女的值大于双亲的值，则相互交换，将一
5  //个集合局部调整为最大堆。
6      int i = start; int j = 2*i + 1;    //j是i的左子女
7      Element<T> temp = L[i];           //暂存子树根结点
8      while (j <= m) {                   //逐层比较
9          if (j < m && L[j] < L[j+1]) j++; //让j指向两子女中的大者
10         if (temp >= L[j]) break;         //temp排序码大不调整
11         else {                           //否则子女中的大者上移
12             L[i] = L[j];
13             i = j; j = 2*j+1;           //i下降到子女位置
14         }
15     }
16     L[i] = temp;                        //temp放到合适位置
17 };
18 
```

最大堆堆顶 `L.Vector[0]` 具有最大的排序码，将 `L.Vector[0]` 与 `L.Vector[n-1]` 对调，把具有最大排序码的元素交换到最后，再对前面的 $n-1$ 个元素, 使用堆的调整算法 `siftDown(L, 0, n-2)`，重新建立最大堆，具有次最大排序码的元素又上浮到 `L.Vector[0]` 位置。

再对调L.Vector[0]和L.Vector[n-2]，再调用siftDown(L, 0, n-3), 对前面的n-2个元素重新调整，...

如此反复执行，最后得到全部排序好的元素序列。这个算法即堆排序算法，其细节在下面的程序中给出。

```
1  template <class T>
2  void HeapSort (dataList<T>& L) {
3      //对表L.Vector[0]到L.Vector[n-1]进行排序，使得表
4      //中各个元素按其排序码非递减有序
5      int i, n = L.length();
6      for (i = (n - 2)/2; i >= 0; i--)    //将表转换为堆
7          siftDown (L, i, n-1);
8      for (i = n - 1; i >= 0; i--) {    //对表排序
9          L.Swap(0, i);  siftDown (L, 0, i-1);
10     }
11 };
```

第二个 for 循环中调用了n-1次siftDown()算法, 该循环的计算时间为 $O(n\log_2 n)$ 。因此, 堆排序的时间复杂度为 $O(n\log_2 n)$ 。

该算法的附加存储主要是在第二个 for 循环中用来执行元素交换时所用的一个临时元素。因此, 该算法的空间复杂度为 $O(1)$ 。

堆排序是一个不稳定的排序方法

9.4 归并排序 (Merge Sort)

归并，是将两个或两个以上的有序表合并成一个新的有序表。

元素序列L1中有两个有序表Vector[left..mid]和Vector[mid+1..right]。它们可归并成一个有序表, 存于另一元素序列L2的Vector[left..right] 中。

这种方法称为两路归并 (2-way merging)。

变量 i 和 j 分别是表Vector[left..mid]和Vector [mid+1..right]的检测指针。k 是存放指针

9.4.1 迭代的归并排序算法

- 迭代的归并排序算法就是利用两路归并过程进行排序。其基本思想是：
- 设初始元素序列有 n 个元素，首先把它看成是 n 个长度为 1 的有序子序列（归并项），做两两归并，得到 $\lceil n/2 \rceil$ 个长度为 2 的归并项（最后一个归并项的长度为1）；再做两两归并，得到 $\lceil n/4 \rceil$ 个长度为 4 的归并项（最后一个归并项长度可以短些）...，如此重复，最后得到一个长度为 n 的有序序列。




```

1  template <class T>
2  void MergeSort (dataList<T>& L) {
3  //归并排序,按元素排序码非递减的顺序对表L中元素排序
4      int n = L.Length();
5      dataList<T> & tempList (n);    //创建临时表
6      int len = 1;
7      while (len < n) {
8          MergePass (L, tempList, len);    len *= 2;
9          MergePass (tempList, L, len);    len *= 2;
10     }
11 };

```

```

1  template <class T>
2  void MergePass (dataList<T>& L1, dataList<T>& L2, const int len) //一趟对数据
   表完整的归并
3  {
4      int i = 0, j, n = L1.Length();
5      while (i+2*len <= n-1)
6      {
7          merge (L1, L2, i, i+len-1, i+2*len-1);
8          i += 2 * len;    //循环两两归并
9      }
10     if (i+len <= n-1)
11         merge (L1, L2, i, i+len-1, n-1);
12     else for (j = i; j <= n-1; j++) L2[j] = L1[j];
13 };

```

```

1  #include "dataList.h"
2  template <class T>
3  void merge (dataList<T>& L1, dataList<T>& L2,
4             const int left, const int mid, const int right) {
5  //L1.Vector[left..mid]与L1.Vector[mid+1..right]是两个有序表, 将这两个有序表归并成
   一个有序表
6  //L2.Vector[left..right]
7      int k, i, j;
8      i = left; j = mid+1; k = left;    //s1, s2是检测指针, t是存放指针
9      while (i <= mid && j <= right)    //两两比较
10         if (L1[i] <= L1[j])
11             L2[k++] = L1[i++];
12         else
13             L2[k++] = L1[j++];
14     while (i <= mid) L2[k++] = L1[i++];    //若第一个表未检测完,复制
15     while (j <= right) L2[k++] = L1[j++]; //若第二个表未检测完,复制
16 };
17

```

9.4.2 递归的表归并排序

与快速排序类似, 归并排序也可以利用划分为子序列的方法递归实现。

递归的归并排序方法:

- 1) 首先, 要把整个待排序序列划分为两个长度大致相等的部分, 分别称之为左子表和右子表。
- 2) 对左、右子表分别递归地进行排序
- 3) 把排好序的左右两个子表进行归并。

```

1  template <class Type>
2  int staticlinkList<Type> :: MergeSort (const int left,  const int right )
3  {
4      if ( left >= right ) return left;
5      int mid = ( left + right ) / 2;
6      return ListMerge (MergeSort ( left, mid ), MegerSort ( mid+1, right ));
7      //以中点mid为界，分别对左半部和右半部进行表归并排序
8  }
9

```

```

1  template <class Type>
2  int staticlinkList<Type> :: ListMerge (const int start1,  const int start2
3  )
4  {
5      int k = 0,  i = start1,  j = start2;
6      while ( i && j )
7      {
8          if ( Vector[i].key <= Vector[j].key )
9          {
10             vector[k].link = i;  k = i;
11             i = Vector[i].link;
12          }
13          else {
14             vector[k].link = j;  k = j;
15             j = Vector[j].link;
16          }
17      }
18      if ( i == 0 ) vector[k].link = j;
19      else vector[k].link = i;
20      return vector[0].link;
21  }
22

```

S排序的性能分析

排 序 方 法	比较次数		移动次数		稳定性	附加存储	
	最好	最差	最好	最差		最好	最差
直接插入排序	n	n^2	0	n^2	√	1	
折半插入排序	$n \log_2 n$		0	n^2	√	1	
起泡排序	n	n^2	0	n^2	√	1	
快速排序	$n \log_2 n$	n^2	$< n \log_2 n$	n^2	×	$\log_2 n$	n
简单选择排序	n^2		0	n	×	1	
锦标赛排序	$n \log_2 n$		$n \log_2 n$		√	n	
堆排序	$n \log_2 n$		$n \log_2 n$		×	1	
归并排序	$n \log_2 n$		$n \log_2 n$		√	n	