# 《算法导论》学习笔记

最笨的方法往往是最好最快的方法 by 谭川奇

## 第一部分：基础知识

### 第1章：算法在计算中的作用

* 1. 算法即是一系列的计算步骤，用来将一个有效的输入转换成一个有效的输出。
  2. 计算机的有限的资源必须被有效的利用，算法就是来解决这些问题的方法。

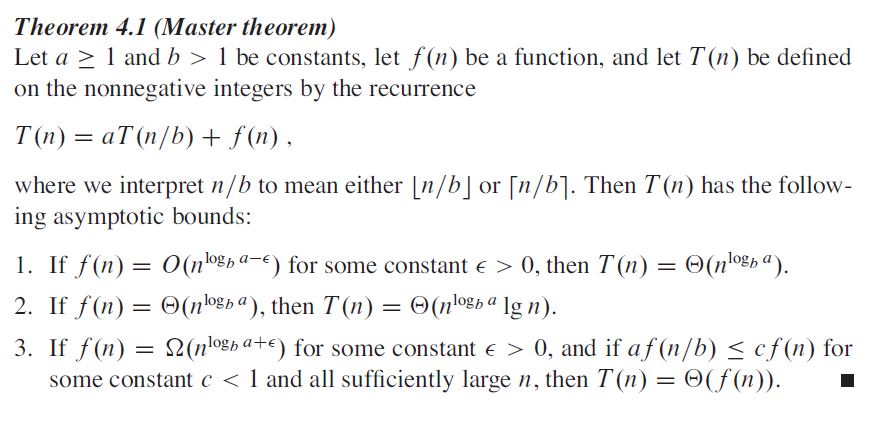
### 第2章：算法入门

* 1. 循环不变式的三个性质：（循环不变式通常用来证明递归的正确性）
     1. 初始化：它在循环的第一轮迭代开始之前，应该是正确的。
     2. 保持：如果在循环的某一次迭代开始之前它是正确的，那么，在下一次迭代开始之前，它也应该保持正确。
     3. 终止：当循环结束时，不变式给了我们一个有用的性质，它有助于表明算法是正确的。
  2. 伪代码中的约定：
     1. 书写上的“缩进”表示程序中的分程序（程序块）结构。
     2. while,for,repeat等循环结构和if,then,else条件结构与Pascal中相同。
     3. 符号 "▷”表示后面部分是个注释。
     4. 多重赋值i←j←e是将表达式e的值赋给变量i和j；等价于j←e，再进行赋值i←j。
     5. 变量（如i,j和key等）是局部给定过程的。
     6. 数组元素是通过“数组名[下标]”这样的形式来访问的。
     7. 复合数据一般组织成对象，它们是由属性(attribute)和域(field)所组成的。
     8. 参数采用按值传递方式：被调用的过程会收到参数的一份副本。
     9. 布尔运算符"and”和"or”都是具有短路能力。
  3. 算法分析即指对一个算法所需要的资源进行预测。
  4. 对于一个算法，一般只考察其最坏情况的运行时间，理由有三：
     1. 一个算法的最坏情况运行时间是在任何输入下运行时间的一个上界。
     2. 对于某些算法来说，最坏情况出现得还是相当频繁的。
     3. 大致上看来，“平均情况”通常和最坏情况一样差。
  5. 分治策略：将原问题划分成ｎ个规模较小而结构与原问题相似的子问题；递归地解决这些小问题，然后再合并其结果，就得到原问题的解。
  6. 分治模式在每一层递归上都有三个步骤：
     1. 分解(Divde)：将原问题分解成一系列子问题；
     2. 解决(Conquer)：递归地解答各子问题。若子问题足够小，则直接求解；
     3. 合并(Combine)：将子问题的结果合并成原问题的解。

### 第3章：函数的增长

* 1. 对几个记号的大意：o（非渐近紧确上界） ≈ <； O（渐近上界） ≈ ≤；  
     Θ（渐近紧界）≈ =； Ω（渐近下界）≈ ≥； ω（非渐近紧确下界）≈  >;  
     这里的<,≤,=,≥,>指的是规模上的比较，即o(g(n))的规模比g(n)小。
  2. o(g(n))={ f(n):   
     对任意正常数c，存在常数n0>0，使对所有的n≧n0，有0≦f(n)<cg(n) }
  3. O(g(n))={ f(n):   
     存在正常数c和n0，使对所有n≧n0，有0≦f(n)≦cg(n) }
  4. Θ(g(n))={ f(n):   
     存在正常数c1，c2和n0，使对所有的n≧n0，有0≦c1g(n)≦f(n)≦c2g(n) }
  5. Ω(g(n))={ f(n):   
     存在正常数c和n0，使对所有n≧n0，有0≦cg(n)≦f(n) }
  6. ω(g(n))={ f(n):   
     对任意正常数c，存在常数n0>0，使对所有的n≧n0，有0≦cg(n)<f(n) }

### 第4章：递归式

* 1. 递归式是一组等式或不等式，它所描述的函数是用在更小的输入下该函数的值来定义的。例如Merge-Sort的最坏情况运行时间T(n)可以用以下递归式来表示：  
                                         2T(n/2) + n,  if n > 11
  2. 解递归式的方法主要有三种：代换法、递归树方法、主方法。
  3. 代换法(Substitution method)(P38~P40)  
     定义：先猜测某个界的存在，再用数学归纳法去证明该猜测的正确性。  
     缺点：只能用于解的形式很容易猜的情形。  
     总结：这种方法需要经验的积累，可以通过转换为先前见过的类似递归式来求解。
  4. 递归树方法(Recursion-tree method)  
     起因：代换法有时很难得到一个正确的好的猜测值。  
     用途：画出一个递归树是一种得到好猜测的直接方法。  
     分析(重点)：在递归树中，每一个结点都代表递归函数调用集合中一个子问题的代价。将递归树中每一层内的代价相加得到一个每层代价的集合，再将每层的代价相加得到递归式所有层次的总代价。  
     总结：递归树最适合用来产生好的猜测，然后用代换法加以验证。  
     递归树的方法非常直观，总的代价就是把所有层次的代价相加起来得到。但是分析这个总代价的规模却不是件很容易的事情，有时需要用到很多数学的知识。
  5. 主方法(Master method)   
     主方法是最好用的Cookbook方法，总结了常见的情况并给出了一个公式。  
     优点：针对形如T(n)=aT(n/b)+f(n)的递归式  
     缺点：并不能解所有形如上式的递归式的解。因为主方法在第1种情况与第2种情况之间、第2种情况与第3种情况之间都存在着一条沟，所以会存在着不能适用的情况。  
     直觉上：实际上主方法一直在比较f(n)与NLogBA的规模，然后选取规模大的作为最后的递归式的规模。  
     

### 第5章：概率分析与随机算法

一：随机算法：如果一个算法的行为不只是由输入决定，同时也由随机数生成器所产生的数值决定，则称这个算法是随机的。

二：指示器随机变量I(A)的定义很简单：

I(A)= {0 如果A不发生的话1 如果A发生的话

       事件A对应的指示器随机变量的期望期等于事件A发生的概率。

三：介绍了两种随机排列数组的生成方法：

1. 随机优先级法：为数组的每个元素赋一个随机的优先级，再根据这个优先级对数组中的元素进行排序。可证这样得到的数字满足随机的性质。
2. 原地交换法：依次把A[i]与A[Random(i+1, Length(A))]进行swap，得到的新数组也满足随机性。  
   for i ← 1 to n  
       do swap A[i] ↔ A[Random(i, Length(A))]

四：有时间，在真正的环境中的输入可能并不是随机的，所以我们可以采用先将输入进行随机打乱的方法来保证输入数据的随机性，这点在很多算法中得以体现，比如快排有其随机选取种子数来向输入中加入随机化的成分。

五：===5.4节带\*未看===  
 不过值得提一下的是“在线雇佣问题”与“苏格拉底的择偶观”很相似。  
 先用三分之一的时间，即分出大、中、小三类，再用三分之一的时间验证自己的观点是否正确，等到最后三分之一时，选择了属于大类中的一支美丽的麦穗。

## 第二部分：排序和顺序统计学

这一部分将要给出几个解决以下排序问题的算法：  
输入：n个数的序列<a1,a2, … an>  
输出：输入序列的一个重排<a1’,a2’,...,an’>，使a1’≦a2’≦...≦an’

* 1. 原地排序算法：只有线性个数的元素会被移动到集合之外的排序算法。
  2. 第6章介绍堆排序
  3. 第7章介绍快速排序
  4. 第8章介绍了基于“比较”排序的算法的下界为Ω(nlgn)。并介绍了几种不基于比较的排序方法，它们能突破Ω(nlgn)的下界。计数排序、基数排序、桶排序。
  5. 第9章介绍了顺序统计的概念：第i个顺序统计是集合中第i小的数。  
     并介绍了两个算法：
     1. 最坏情况为O(n2)，但平均情况下为线性O(n)的算法
     2. 最坏情况下为线性O(n)的算法

### 第6章：堆排序

* 1. 堆排序是一个时间复杂度为O(nlgn)、原地排序算法。
  2. “堆”数据结构不只在推排序时有用，还可以构成一个有效的优先队列。
  3. 堆的定义是这样的：
     1. 一个堆是一颗完全二叉树
     2. 对于大（小）根堆，每个节点的值都比它的子节点要大（小）
  4. 虽然堆排的理论效率好，但是往往一个好的快排的实现要优于堆排，所以堆更常见于作为高效的优先级队列。一个堆可以在O(lgn)的时间内，支持大小为n的集合上的任意优先队列的操作。

该章的代码如下：

|  |
| --- |
| //////////////////////////////////////////////////////////////////////////  /// COPYRIGHT NOTICE  /// Copyright (c) 2009  /// All rights reserved.  ///  /// @file HeapSort.cpp  /// @brief 堆排序的学习，使用大头堆  ///  ///  /// @author 谭川奇 chuanqi.tan@gmail.com  /// @date 2011-5-25  ///  ///  /// 修订说明：最初版本  //////////////////////////////////////////////////////////////////////////  #include"stdafx.h"  #include<iostream>  #include<vector>  #include<algorithm>  #include<iterator>  usingnamespacestd;  namespacechapter6  {  /// 保持堆的性质  ///  /// 将to\_make的[0,length)元素视为一棵完全二叉树，以第i个元素为根的子树除了第i个元素之外都满足大堆的性质  /// 调用此方法之后，这棵完全二叉树以第i个元素为根的子树都满足大堆的性质  /// @to\_make 保存数据的数组  /// @length 标记to\_make的[0,length)元素视为一个完全二叉树  /// 第length个元素之后[length, n)的元素不包括在这棵完全二叉树里  /// @i 需要处理的第i个元素  /// @note to\_make的前length个元素并不一定是一个堆（因为它不满足大堆的性质），但可以映射为完全二叉树  voidMakeHeap(vector<int>&to\_make, size\_tlength, size\_ti)  {  size\_tleft = 2 \* i + 1;  size\_tright = 2 \* i + 2;  size\_tthe\_max = i;  if (left<length&&to\_make[left] >to\_make[i])  {  the\_max = left;  }  if (right<length&&to\_make[right] >to\_make[the\_max])  {  the\_max = right;  }  if (the\_max != i)  {  std::swap(to\_make[i], to\_make[the\_max]);  MakeHeap(to\_make, length, the\_max);  }  }  /// 建堆  ///  /// 将to\_built数组改建成一个大头堆  voidBuildHeap(vector<int>&to\_built)  {  //这里只需要从to\_built.size() / 2 - 1开始的原因在于：  //对叶子结点来说，它和它的子结点（为空）总是满足堆的定义的，所以只需要处理非叶子结点  for (inti = to\_built.size() / 2 - 1; i>= 0; --i)  {  MakeHeap(to\_built, to\_built.size(), i);  }  }  /// 堆排序  voidHeapSort(vector<int>&to\_sort)  {  BuildHeap(to\_sort);  for (inti = to\_sort.size() - 1; i> 0; --i)  {  std::swap(to\_sort[0], to\_sort[i]);  MakeHeap(to\_sort, i, 0);  }  }  /// 优先队列  ///  /// 堆的主要应用之一：优先队列  classPriorityQueue  {  public:  PriorityQueue()  {  intto\_init[] = {8,5,78,45,64,987,45,34,23,4,23};  \_queue.assign(to\_init, to\_init + sizeof(to\_init) / sizeof(int));  BuildHeap(\_queue);  }  intGetMaxPriority()  {  return\_queue[0];  }  intGetAndPopMaxPriority()  {  intmax\_priority = GetMaxPriority();  std::swap(\_queue[0], \_queue[\_queue.size() - 1]);  \_queue.erase(\_queue.end() - 1);  MakeHeap(\_queue, \_queue.size(), 0);  returnmax\_priority;  }  voidChangePriority(intindex, intnew\_priority)  {  if (\_queue[index] == new\_priority)  {//优先级不变，直接返回  return;  }  if (\_queue[index] >new\_priority)  {//降底优先级  \_queue[index] = new\_priority;  MakeHeap(\_queue, \_queue.size(), index);  }  else  {//提升优先级  \_queue[index] = new\_priority;  while(index> 0)  {  intparent\_index = static\_cast<int>((index + 1) / 2) - 1;  if (\_queue[parent\_index] <\_queue[index])  {  std::swap(\_queue[parent\_index], \_queue[index]);  index = parent\_index;  }  else  {  break;  }  }  }  }  voidInsert(intpriority)  {  \_queue.push\_back(std::numeric\_limits<int>::min());  ChangePriority(\_queue.size() - 1, priority);  }  boolIsEmpty()  {  return\_queue.empty();  }  voidDisplay()  {  copy(\_queue.begin(), \_queue.end(), ostream\_iterator<int>(cout, " "));  cout<<endl;  }  private:  vector<int>\_queue;  };  inttest()  {  intto\_init[] = {8,5,78,45,64,987,45,34,23,4,23};  vector<int>to\_sort(to\_init, to\_init + sizeof(to\_init) / sizeof(int));  cout<<"原始数组，准备进行堆排序：";  copy(to\_sort.begin(), to\_sort.end(), ostream\_iterator<int>(cout, " "));  HeapSort(to\_sort);  //reverse(result.begin(), result.end());  cout<<endl<<"堆排序结束：";  copy(to\_sort.begin(), to\_sort.end(), ostream\_iterator<int>(cout, " "));  cout<<endl;  cout<<"初始化一个优先队列：";  PriorityQueuequeue;  queue.Display();  cout<<"开始不断的取最高优先级的任务出列："<<endl;  while(!queue.IsEmpty())  {  cout<<queue.GetAndPopMaxPriority() <<":\t";  queue.Display();  }  cout<<"开始添加任务入列："<<endl;  for (size\_ti = 0; i<to\_sort.size(); ++i)  {  queue.Insert(to\_sort[i]);  queue.Display();  }  cout<<"更改一个任务的优先级：任务优先级调整为-9并自动调整堆："<<endl;  queue.ChangePriority(2, -9);  queue.Display();  return 0;  }  }  Output:  987: 78 64 45 45 23 8 5 34 23 4  78: 64 45 45 34 23 8 5 4 23  64: 45 34 45 23 23 8 5 4  45: 45 34 8 23 23 4 5  45: 34 23 8 5 23 4  34: 23 23 8 5 4  23: 23 5 8 4  23: 8 5 4  8: 5 4  5: 4  4:  开始添加任务入列：  4  5 4  8 4 5  23 8 5 4  23 23 5 4 8  34 23 23 4 8 5  45 23 34 4 8 5 23  45 45 34 23 8 5 23 4  64 45 34 45 8 5 23 4 23  78 64 34 45 45 5 23 4 23 8  987 78 34 45 64 5 23 4 23 8 45  更改一个任务的优先级：任务2优先级调整为-9并自动调整堆：  987 78 23 45 64 5 -9 4 23 8 45  请按任意键继续. . . |

### 第7章：快速排序

1. 快速排序的最坏运行时间为O(n2)，期望运行时间为O(nlgn)，且由于O(nlgn)中隐含的常数因子很小，所以快排通常是用于排序的最佳的实用选择（因为其平均性能非常好）。
2. 快排真的太棒了：平均性能非常好、原地排序不需要额外的空间、算法简单只需要寥寥几行就搞定（比冒泡还少）。
3. 对10W个随机数进行排序比较，快排平均在600MS，而堆排平均在900MS，性能差距可见一斑啊。
4. 快排的平均情况运行时间与其最佳情况运行时间很接近，而不是非常接近于其最坏情况运行时间，所以一般来说快排效率是最高的，这是快排在现代得以大规模的使用的根本原因。
5. 快速排序的随机化版本：正如第5章所说的，由于工程中的输入可能不随机的，所以我们要将其随机化。有两种可选方案（1）直接对输入数据进行随机化排列（2）采用随机取样的随机化技术。随机取样的效率更高一些，所以在快速排序的随机化版本中采用随机取样的技术。  
   方法很简单，就是在每趟sort之前随机选取一个数与最未尾的元素进行交换操作，这样简单高效的实现了随机化。  
   //加入随机取样的随机化技术  
   intrandom\_swap = (rand() % (EndIndex - BeginIndex + 1)) + BeginIndex;  
   std::swap(ToSort[random\_swap], ToSort[EndIndex]);  
   这个技术太有用啊，因为快速排序在输入数据已经有序时的性能是最差的，但是输入数据已经有序的情况又会经常发生，所以这个随机取样就显得异常的重要。如果没有这个随机取样，快排绝得不到这样的应用。  
   在我做的实验中，对2000个有序的数据进行排序，在未没采用随机化的情况下，平均耗时860MS，而使用了随机取样之后平均耗时8MS，效率提高了100倍。由于输入有序的情况是非常常见的，所以这个随机化才更显示重要了！

|  |
| --- |
| //////////////////////////////////////////////////////////////////////////  /// COPYRIGHT NOTICE  /// Copyright (c) 2009  /// All rights reserved.  ///  /// @file quick\_sort.cpp  /// @brief quick\_sort.cpp的简短描述  ///  /// quick\_sort.cpp的详细描述  ///  /// @author 谭川奇 chuanqi.tan@gmail.com  /// @date 2011/05/26  /// @version 1.0  ///  ///  /// 修订说明：最初版本  //////////////////////////////////////////////////////////////////////////  #include"stdafx.h"  #include<vector>  #include<iostream>  #include<iterator>  #include<ctime>  #include<algorithm>  usingnamespacestd;  namespacechapter7  {  /// 采用了随机取样技术的快速排序  ///  /// 快速排序的平均效率为O(nlgn)，最坏情况为O(n^2)  voidQuickSort(vector<int>&ToSort, intBeginIndex, intEndIndex)  {  if(BeginIndex<EndIndex)  {  //加入随机取样的随机化技术  //一定要使用，对平均性能的提升作用太大了  intrandom\_swap = (rand() % (EndIndex - BeginIndex + 1)) + BeginIndex;  std::swap(ToSort[random\_swap], ToSort[EndIndex]);  //i代表的是比ToSort[EndIndex]小的元素的上界，即ToSort[0...i)的元素值都比ToSort[EndIndex]要小  //也意味着下一个比ToSort[EndIndex]小的元素要放置的位置；但是在当前ToSort[i] >= ToSort[EndIndex]  inti = BeginIndex;  //j代表已经检查过的元素的上界  for (intj = BeginIndex; j != EndIndex; ++j)  {  //找到满足比ToSort[EndIndex]小的元素  if (ToSort[j] <ToSort[EndIndex])  {  swap(ToSort[i], ToSort[j]); //将这个比ToSort[EndIndex]小的元素移到第i个去,满足了i代表的意义  ++i; //由于新找到了一个比ToSort[EndIndex]小的元素,所以上界应该+1  }  }  swap(ToSort[i], ToSort[EndIndex]);  QuickSort(ToSort, BeginIndex, i - 1);  QuickSort(ToSort, i + 1, EndIndex);  }  }  /// 对模糊区间的快速排序  ///  /// 问题描述：（算法导论-6题）\n  /// 考虑这样的一种排序问题，即无法准确地知道待排序的各个数字到底是多少。对于其中的每个数字，  /// 我们只知道它落在实轴上的某个区间内。亦即，给定的是n个形如[a(i), b(i)]的闭区间（这里小括  /// 后起下标的作用，后同），其中a(i) <= b(i)。算法的目标是对这些区间进行模糊排序  /// （fuzzy-sort），亦即，产生各区间的一个排列<i(1), i(2), ..., i(n)>，使得存在一个c(j)属于  /// 区间[a(i(j)), b(i(j))]，满足c(1) <= c(2) <= c(3) <= ... <= c(n)。\n  /// a) 为n个区间的模糊排序设计一个算法。你的算法应该具有算法的一般结构，它可以快速排序左部  /// 端点（即各a(i)），也要能充分利用重叠区间来改善运行时间。（随着各区间重叠得越来越多，  /// 对各区间进行模糊排序的问题会变得越来越容易。你的算法应能充分利用这种重叠。）\n  /// b) 证明：在一般情况下，你的算法的期望运行时间为Θ(nlgn)，但当所有的区间都重叠时，期望的  /// 运行时间为Θ(n)（亦即，当存在一个值x，使得对所有的i，都有x∈[a(i), b(i)]）。你的算法  /// 不应显式地检查这种情况，而是应随着重叠量的增加，性能自然地有所改善。  voidSmoothQuickSort(vector<pair<int, int>>&to\_sort, intbegin\_index, intend\_index)  {  if (begin\_index<end\_index)  {  //取最后一个区间为主元  autolast = to\_sort[end\_index];  //获取要比较的区间（除去主元）为[begin\_index, end\_index) => [i,j]  //区间[i,j]意思是：其中所有的元素要不还未处理，要不相互重叠有至少一个重叠值，并且该值还与to\_sort[end\_index]重叠  //即在题目中规定的语义下与to\_sort[end\_index]绝对相等  inti = begin\_index;  intj = end\_index - 1;  for (intk = begin\_index; k<= j; )  {  if (to\_sort[k].second<= last.first)  {//严格小于主元  swap(to\_sort[i], to\_sort[k]);  ++i;  ++k;  }  elseif (to\_sort[k].first>= last.second)  {//严格大于主元  swap(to\_sort[j], to\_sort[k]);  --j;  }  else  {  // 与主元区间有重叠，则更新主元为重叠区间（交集）  // 此方法参考了http://blogold.chinaunix.net/u/18517/showart\_487873.html  //  // 这种想法真的好，因为缩小了主元的区间（交集），所以就可以认为以后任何与缩小之后主元有重叠的区间都一定与当前区间to\_sort[k]重叠（因为它完全包括缩小后的主元）\n  // 因此这样就可以确定最后在[i,j]中的所有元素在本题的约定下与主元绝对相等（即所有的元素相互重叠），所以不需要再处理。这就符合了题目中的“充分利用重叠区间来改善运行时间”\n  // 如果没有这步缩小区间，就只能认为[i,j]中的元素各自与主元有重叠而无法判断为绝对相等。  // @note 这里很容易弄错的一点是：区间重叠并没有传递性，重叠区间的元素并不能认为是已序的  last.first = max(to\_sort[k].first, last.first);  last.second = min(to\_sort[k].second, last.second);  ++k;  }  }  swap(to\_sort[j + 1], to\_sort[end\_index]);  SmoothQuickSort(to\_sort, begin\_index, i - 1);  SmoothQuickSort(to\_sort, j + 2, end\_index);  }  }  inttest()  {  cout<<"==========================快速排序============================="<<endl;  vector<int>ToSort;  for (inti = 0; i< 100; ++i)  {  ToSort.push\_back(rand());  }  cout<<"随机填充个数："<<endl;  copy(ToSort.begin(), ToSort.end(), ostream\_iterator<int>(cout, " "));  QuickSort(ToSort, 0, ToSort.size() - 1);  cout<<endl<<"快速排序的结果如下："<<endl;  copy(ToSort.begin(), ToSort.end(), ostream\_iterator<int>(cout, " "));  cout<<endl<<"======================模糊区间的快速排序========================="<<endl;  vector<pair<int, int>>to\_sort\_smooth;  for (inti = 0; i< 10; ++i)  {  intb = rand() % 100;  to\_sort\_smooth.push\_back(make\_pair(b, b + rand() % 100));  }  SmoothQuickSort(to\_sort\_smooth, 0, to\_sort\_smooth.size() - 1);  for\_each(to\_sort\_smooth.begin(), to\_sort\_smooth.end(), [](pair<int, int>&p){  cout<<p.first<<"\t --> \t"<<p.second<<endl;  });  return 0;  }  }  Output:  ==========================快速排序=============================  随机填充100个数：  41 18467 6334 26500 19169 15724 11478 29358 26962 24464 5705 28145 2  3281 16827 9961 491 2995 11942 4827 5436 32391 14604 3902 153 292 1  2382 17421 18716 19718 19895 5447 21726 14771 11538 1869 19912 25667  26299 17035 9894 28703 23811 31322 30333 17673 4664 15141 7711 28253  6868 25547 27644 32662 32757 20037 12859 8723 9741 27529 778 12316  3035 22190 1842 288 30106 9040 8942 19264 22648 27446 23805 15890 6  729 24370 15350 15006 31101 24393 3548 19629 12623 24084 19954 18756  11840 4966 7376 13931 26308 16944 32439 24626 11323 5537 21538 16118  2082 22929 16541  快速排序的结果如下：  41 153 288 292 491 778 1842 1869 2082 2995 3035 3548 3902 4664 482  7 4966 5436 5447 5537 5705 6334 6729 6868 7376 7711 8723 8942 9040  9741 9894 9961 11323 11478 11538 11840 11942 12316 12382 12623 12859  13931 14604 14771 15006 15141 15350 15724 15890 16118 16541 16827 1  6944 17035 17421 17673 18467 18716 18756 19169 19264 19629 19718 1989  5 19912 19954 20037 21538 21726 22190 22648 22929 23281 23805 23811  24084 24370 24393 24464 24626 25547 25667 26299 26308 26500 26962 274  46 27529 27644 28145 28253 28703 29358 30106 30333 31101 31322 32391  32439 32662 32757  ======================模糊区间的快速排序=========================  41 --> 43  21 --> 69  34 --> 87  36 --> 110  20 --> 116  83 --> 178  99 --> 167  84 --> 165  50 --> 141  99 --> 117  请按任意键继续. . . |

### 第8章：线性时间排序

1. 任何比较的排序在最坏的情况下都要用Ω(nlgn)次比较来进行排序，所以合并排序和堆排序是渐近最优的。  
   注意的是：快排不是渐近最优的，因为它在最坏的情况下是O(n2)。
2. 三种以线性时间运行的排序算法：计数排序、基数排序和桶排序。它们都是非比较的。
3. 计数排序
   1. 计数排序的一个重要性就是它是稳定的排序算法，这个稳定性是基数排序的基石。
   2. 计数排序的想法真的很简单、高效、可靠
   3. 缺点在于：
      1. 需要很多额外的空间（当前类型的值的范围）
      2. 只能对离散的类型有效比如int（double就不行了）
      3. 基于假设：输入是小范围内的整数构成的。
4. 基数排序
   1. 基数排序时对每一维进行调用子排序算法时要求这个子排序算法必须是稳定的。
   2. 基数排序与直觉相反：它是按照从底位到高位的顺序排序的。  
      我觉得原因在于：高有效位对底有效位有着决定性的作用。
5. 桶排序
   1. 桶排序也只是期望运行时间能达到线性，对于最坏的情况，它的运行时间取决于它内部使用的子排序算法的运行时间，一般为O(nlgn)。
   2. 桶排序基于假设：输入的的元素均匀的分布在区间[0, 1]上。
   3. 感觉桶排没有什么大的实现价值，因为它限定了输入的区间，还要求最好是均匀分布，它的最坏情况并不好。
6. 所有的线性时间内的排序算法，都作出了一定的假设，是建立在一定的假设基础上的。

|  |
| --- |
| //////////////////////////////////////////////////////////////////////////  /// COPYRIGHT NOTICE  /// Copyright (c) 2009  /// All rights reserved.  ///  /// @file liner\_sort.cpp  /// @brief liner\_sort.cpp的简短描述  ///  /// liner\_sort.cpp的详细描述  ///  /// @author 谭川奇 chuanqi.tan@gmail.com  /// @date 2011/05/26  /// @version 1.0  ///  ///  /// 修订说明：最初版本  //////////////////////////////////////////////////////////////////////////  #include"stdafx.h"  #include<iostream>  #include<algorithm>  #include<vector>  #include<iterator>  usingnamespacestd;  namespacechapter8  {  /// 计数排序  voidcounting\_sort()  {  intconstk\_max\_size = 100; //待排序的所有元素都必须位于区间[0, k\_max\_size)  //初始化区间为[0, k\_max\_size)之间的随机数作为输入  std::vector<int>v;  for (inti = 0; i< 100; ++i)  {  v.push\_back(rand() % k\_max\_size);  }  //进行计数c[i] = j代表着i在输入数据中出现了j次  std::vector<int>c(k\_max\_size, 0);  std::for\_each(v.begin(), v.end(), [&](inti){++c[i];});  //对所有的计数从依次总结出最后的排序，并没有使用原书上的方法  //比书上的算法更直接，将二步合成了一步，效率上是同样的渐近时间复杂度的，似乎更好点。  v.clear();  for (inti = 0; i<k\_max\_size; ++i)  {  for (intk = 0; k<c[i]; ++k)  {  v.push\_back(i);  }  }  copy(v.begin(), v.end(), std::ostream\_iterator<int>(std::cout, " "));  }  /// 基数排序时值最大的维数：意味着元素的区间为[0, 999]  intconstk\_max\_dim = 3;  /// 基数排序  voidradix\_sort(std::vector<int>&v, intdim)  {  //根据第dim维的大小进行“稳定”的排序，这个stable很重要，是基数排序的基石  stable\_sort(v.begin(), v.end(), [&](intleft, intright) ->bool{  //得到一个数的某一维的值  //eg: GetDim(987, 1) = 9  // GetDim(987, 2) = 8  // GetDim(987, 3) = 7  autoGetDim = [](intnumber, intd) ->int  {  for (inti = 1; i<= k\_max\_dim - d; ++i)  {  number /= 10;  }  returnnumber % 10;  };  returnGetDim(left, dim) <GetDim(right, dim);  });  if (dim> 1)  {  radix\_sort(v, dim - 1);  }  }  /// 基础排序算法的初始化和调用  voidradix\_sort\_caller()  {  //初始化[0,999]之间的随机数作为输入  std::vector<int>v;  for (inti = 0; i< 10; ++i)  {  v.push\_back(rand() % static\_cast<int>(pow(10.0, static\_cast<double>(k\_max\_dim))));  }  std::copy(v.begin(), v.end(), std::ostream\_iterator<int>(std::cout, " ")); std::cout<<std::endl;  radix\_sort(v, k\_max\_dim);  std::copy(v.begin(), v.end(), std::ostream\_iterator<int>(std::cout, " "));  }  /// 桶排序  voidbucket\_sort()  {  //初始化[0,1)之间的随机数  vector<double>v;  for (inti = 0; i< 10; ++i)  {  v.push\_back((rand() % 100) \* 1.0 / 100.0);  }  copy(v.begin(), v.end(), ostream\_iterator<double>(cout, " "));  cout<<endl;  //构建桶，并将所有的元素放入到相应的桶中去  vector<vector<double>>bucket(10); //10个桶  for\_each(v.begin(), v.end(), [&](doubled){  bucket[d \* 10].push\_back(d);  });  //对每一个桶里的元素进行排序  for\_each(bucket.begin(), bucket.end(), [](vector<double>&sub\_v){  sort(sub\_v.begin(), sub\_v.end());  });  //依次把每个桶中的元素提取出来并组合在一起  v.clear();  for\_each(bucket.begin(), bucket.end(), [&](vector<double>&sub\_v){  v.insert(v.end(), sub\_v.begin(), sub\_v.end());  });  copy(v.begin(), v.end(), ostream\_iterator<double>(cout, " "));  }  inttest()  {  cout<<endl<<"===========开始计数排序==========="<<endl;  counting\_sort();  cout<<endl<<"===========开始基数排序==========="<<endl;  radix\_sort\_caller();  cout<<endl<<"===========开始桶排序==========="<<endl;  bucket\_sort();  getchar();  return 0;  }  }  Output:  ===========开始计数排序===========  0 1 2 3 4 5 5 6 6 8 11 11 12 16 16 18 18 21 22 23 23 23 24  26 26 27 27 29 29 29 29 31 33 34 35 35 36 37 37 38 38 39 40  40 41 41 41 41 42 42 42 44 44 45 46 47 47 48 48 50 53 53 54  56 57 58 59 61 62 62 64 64 64 66 67 67 68 69 69 70 71 73 76  78 78 81 82 82 84 88 90 90 91 91 92 93 94 95 95 99  ===========开始基数排序===========  833 115 639 658 704 930 977 306 673 386  115 306 386 639 658 673 704 833 930 977  ===========开始桶排序===========  0.21 0.45 0.24 0.72 0.7 0.29 0.77 0.73 0.97 0.12  0.12 0.21 0.24 0.29 0.45 0.7 0.72 0.73 0.77 0.97 |

### 第9章：中位数和顺序统计学

1. 第i个顺序统计量是该集合中第i小的元素。  
   最小值是第1个顺序统计量(i=1)最大值是第n个顺序统计量(i=n)
2. 中位数是它所在集合的“中点元素”
3. 找最大最小值的算法，一般人可能以为需要2n次比较，实际上只需要最多3次比较，使用的技巧是：  
   将一对元素比较，然后把较大者于max比较，较小者与min比较，这样就只需要3次比较就能得到最后的结果。
4. 以期望线性时间选择顺序统计量的方法是以快速排序为模型。如同在快速排序中一样，此算法的思想也是对输入数组进行递归划分。但和快速排序不同的是，快速排序会递归处理划分的两边，而randomized-select只处理划分的一边。并由此将期望的运行时间由O(nlgn)下降到了O(n)。  
   我觉得C++ STL中的nth\_element用的可能就是这个算法，所以它的效率应该很高。
5. 最坏线性时间选择顺序统计量的方法的核心在于：要保证对数组的划分是一个好的划分。  
   于是方法使用了一个很奇怪的取主元的方法，虽然看起来很奇怪，但是该方法被这样的提出就肯定有它的理论基础的。  
   不过这种取巧的方法不太值得去写一遍，而且明显写出来也很容易的，没有什么新技术和新想法。

|  |
| --- |
| //////////////////////////////////////////////////////////////////////////  /// COPYRIGHT NOTICE  /// Copyright (c) 2009  /// All rights reserved.  ///  /// @file nth\_element.cpp  /// @brief nth\_element.cpp的简短描述  ///  /// nth\_element.cpp的详细描述  ///  /// @author 谭川奇 chuanqi.tan@gmail.com  /// @date 2011/05/26  /// @version 1.0  ///  ///  /// 修订说明：最初版本  //////////////////////////////////////////////////////////////////////////  #include"stdafx.h"  #include<iostream>  #include<algorithm>  #include<vector>  #include<iterator>  usingnamespacestd;  namespacechapter9  {  /// 寻找v数组的子集的第i个顺序统计量  ///  /// 寻找v数组的子集[begin\_index, end\_index]中的第i个元素顺序统计量，<=i<size  int\_NthElement(vector<int>&v, intconstbegin\_index, intconstend\_index, intconstn)  {  //这个判断纯粹只是一个加速return的技巧，没有这个判断算法也是正确的！  if (begin\_index == end\_index)  {  returnv[begin\_index];  }  //随机取样  intswap\_index = rand() % (end\_index - begin\_index + 1) + begin\_index;  swap(v[swap\_index], v[end\_index]);  //根据最后一个主元进行分割成两部分  inti = begin\_index;  for (intj = begin\_index; j<end\_index; ++j)  {  if (v[j] <v[end\_index])  {  swap(v[i++], v[j]);  }  }  swap(v[i], v[end\_index]);  //主元是本区间的第k个元素顺序统计量，<=k<size  intk = i - begin\_index;  if (n == k)  {//找到了  returnv[i];  }  if (n<k)  {//在左区间继续找  return\_NthElement(v, begin\_index, i - 1, n);  }  else  {//在右区间继续找：由于主元是第k个元素顺序统计量(0<=k<size)，所以小于等于主元的元素有k+1个（包括主元），因此寻找右区间的第n-(k+1)个顺序统计量  return\_NthElement(v, i + 1, end\_index, n - k - 1);  }  }  /// 寻找v数组中的第i个顺序统计量，<=i<size  intNthElement(vector<int>&v, intconstn)  {  return\_NthElement(v, 0, v.size() - 1, n);  }  inttest()  {  vector<int>v;  for (inti = 0; i< 10; ++i)  {  v.push\_back((rand() % 1000));  }  copy(v.begin(), v.end(), ostream\_iterator<int>(cout, " "));  cout<<endl;  for (inti = 0; i< 10; ++i)  {  cout<<i<<"th element is:"<<NthElement(v, i) <<endl;  }  return 0;  }  }  Output:  41 467 334 500 169 724 478 358 962 464  0th element is:41  1th element is:169  2th element is:334  3th element is:358  4th element is:464  5th element is:467  6th element is:478  7th element is:500  8th element is:724  9th element is:962  请按任意键继续. . . |

## 第三部分：数据结构

### 第10章：基本数据结构

没发现有什么值得看的，貌似就是下面这些基本的知识，这些知识都不知道就没法混啦。

1. 栈
2. 队列
3. 链表
4. 树的“左孩子、右兄弟”表示法

### 第11章：散列表

1. 在散列表中查找一个元素的时间与在链表中查找一个元素的时候相同，在最坏情况为O(n)，但期望时间为O(1)。  
   在实践中，散列表的效率是很高的，一般可认为是O(1)。
2. 散列是一种极其有效和实用的技术，基本的字典操作只需要O(1)的平均时间。
3. 当待排序的关键字集合是静态的（即当关键字集合一旦存入后不再改变），“完全散列能够在O(1)的最坏情况时间内支持关键字查找。
4. 散列是一种极其有效和实用的技术：基本的字典操作只需要O(1)的平均时间。而且当待排序的关键字的集合是静态的（即当关键字集合一旦存入后不需要再改变），“完全散列”能够在O(1)的最坏时间内支持查找操作。
5. 在众多的简单的解决碰撞的方法中，我觉得比较好的是通过链表法解决碰撞，虽然这个方法的理论最坏效率为O(n)，但是在平均情况下，它的性能也是非常好的，实现简单又高效。
6. 装载因子：给定一个能存放n个元素的、具有m个槽位的散列表T，定义T的装载因子α为n/m，即一个链中平均存储的元素数。
7. 多数的散列函数都假定关键字域为自然数集N，如果所给关键字不是自然数，则必须有一种方法来将它们解释为自然数。
   1. 除法散列法：  
      一般选取m的值为与2的整数幂不大接近的质数
   2. 乘法散列法：  
      构造散列函数的乘法方法包含两个步骤：首先用关键字剩上常数A(0<A<1)，并抽取kA的小数部分；然后用m剩以这个值，再取结果的底。  
      Knuth认为是一个比较理想的值。
   3. 全域散列：全域散列的基本思想是在执行开始时，就从一族仔细设计的函数中，随机地选择一个作为散列函数。
      1. 首先：全域散列表是一种使用“键接法”来解决碰撞问题的散列表方法。
      2. 随机化保证了对于任何输入，算法都具有较好的平均性能。
      3. 全域的散列函数组：设H为一组散列函数，它将给定的关键字域U映射到{0,1,…,m-1}中，这样的一个函数组称为是全域的。如果从H中随机地选择一个散列函数，当关键字K≠J时，两者发生碰撞的概率不大于1/m。
      4. 常用的一个全域散列函数类：  
         首先选择一个足够大的质数p ，使得每一个可能的关键字k 都落到0 到p-1 的范围内，包括首尾的0 和p-1。这里我们假设全域是0 – 15，p 为17。设集合Zp 为{0, 1, 2, …, p-1}，集合Zp\* 为{1, 2, 3, …, p-1}。由于p 是质数，我们可以定义散列函数h(a, b, k) = ((a\*k + b) mod p) mod m。其中a 属于Zp，b 属于Zp\*。由所有这样的a 和b 构成的散列函数，组成了函数簇。即全域散列。
      5. 明白这个散列函数的选取是在“执行开始”随机的选取一个是很重要的，要不然就会不明白到时候怎么进行查找。这里所谓的随机性应该这样理解：对于某一个散列表来说，它在初始化时已经把a,b固定了，但是对于一个还未初始化的全域散列表来说，a,b是随机选取的。
8. 开放寻址法：所有的元素都放在散列表里
   1. 开放寻址法的好处就在于它根本不用指针，而是计算出要存取的各个槽。这样一来，由于不用存储指针就节省了空间，从而可以用同样的空间来提供更多的槽，其潜在的效果就是可以减少碰撞，提高查找速度。
   2. 感觉上开放寻址法很像一个启发式的搜索，它的最坏性能也是O(n)，只不过散列函数为它提供了启发信息从而使得一般的平均性能会很好。
   3. 在开放寻址法中，对散列元素的删除操作执行起来比较困难，因为删除操作会影响查找操作。解决办法是在槽里的值被删除后置一个特定的值DELETED，而不是删除后不管，查找的时候处理一下就可以了
   4. 线性探查、二次探查和双重散列都是对最基本的数组法的改进，虽然它们很漂亮，但是思想上并没有太大的革新，看起来很容易懂的。
   5. 双重散钱是用于开放寻址的最好方法之一，因为它所产生的排列具有随机选择的排列的许多特性。
9. 完全散列：如果某一种散列技术在进行查找时，其最坏情况内存访问次数为O(1)的话，则称其为完全散列。  
   书上使用了一种两级的散列方案，每一级都采用全域散列。  
   通常利用一种两级的散列方案，每一级上都采用全域散列。为了确保在第二级上不出现碰撞，需要让第二级散列表Sj的大小mj为散列到槽j中的关键字数nj的平方。如果利用从某一全域散列函数类中随机选出的散列函数h，来将n个关键字存储到一个大小为m=n的散列表中，并将每个二次散列表的大小置为mj=nj2 (j=0, 1, …, m-1)，则在一个完全散列方案中，存储所有二次散列表所需的存储总量的期望值小于2n。  
   完全散列的关键在于：二次散列表中要求没有碰撞。这是通过确保槽的个数是关键字的个数的平方来实现的。

|  |
| --- |
| //////////////////////////////////////////////////////////////////////////  /// COPYRIGHT NOTICE  /// Copyright (c) 2009  /// All rights reserved.  ///  /// @file hash\_table.cpp  /// @brief hash\_table.cpp的简短描述  ///  /// hash\_table.cpp的详细描述  ///  /// @author 谭川奇 chuanqi.tan@gmail.com  /// @date 2011/05/30  /// @version 1.0  ///  ///  /// 修订说明：最初版本  //////////////////////////////////////////////////////////////////////////  #include"stdafx.h"  #include<iostream>  #include<algorithm>  #include<vector>  #include<iterator>  #include<iomanip>  #include<limits>  usingnamespacestd;  namespacechapter11  {  /// 全域散列表  ///  /// 首先：全域散列表是一种使用“键接法”来解决碰撞问题的散列表方法。  /// 全域散列函数类，首先选择一个足够大的质数p ，使得每一个可能的关键字k 都落到到p-1 的  /// 范围内，包括首尾的和p-1。这里我们假设全域是–15，p 为。设集合Zp 为  /// {0, 1, 2, …, p-1}，集合Zp\* 为{1, 2, 3, …, p-1}。由于p 是质数，我们可以定义散列  /// 函数h(a, b, k) = ((a\*k + b) mod p) mod m。其中a 属于Zp，b 属于Zp\*。  /// 由所有这样的a 和b 构成的散列函数，组成了函数簇。即全域散列。  /// 全域散列的基本思想是在执行<b>开始</b>时，从一族仔细设计的函数中，随机的选择一个作为散列函数。  /// 使之独立于要存储的关键字。不管对手选择了怎样的关键字，其平均性态都很好。  template<typenameT> ///< T是散列表里要存储的元素类型  classUniversalHashTable  {  public:  UniversalHashTable()  {  \_p = 101; //一个足够大的质数  \_m = 10; //槽的个数  \_items.resize(\_m, NULL);  for (inti = 0; i<\_m; ++i)  {//全部先设置好头结点  \_items[i] = new\_Node();  \_items[i]->Next = NULL;  }  // 全域散列的基本思想是在执行<b>开始</b>时，从一族仔细设计的函数中，随机的选择一个作为散列函数。  \_a = rand() % (\_p - 1) + 1;  \_b = rand() % \_p;  }  ~UniversalHashTable()  {  for\_each(\_items.begin(), \_items.end(), [](\_Node \*item){  while (item)  {  autonext = item->Next;  deleteitem;  item = next;  }  });  }  voidInsert(Tconst&new\_value)  {//始终插入在键表的头，头结点之后的第个位置  \_Node \*new\_item = new\_Node;  new\_item->Item = new\_value;  new\_item->Next = NULL;  inthash\_value = \_HashFunction(new\_value);  new\_item->Next = \_items[hash\_value]->Next;  \_items[hash\_value]->Next = new\_item;  }  voidDelete(Tconst&delete\_value)  {  inthash\_value = \_HashFunction(delete\_value);  \_Node \*root = \_items[hash\_value];  while (root->Next)  {  if (root->Next->Item == delete\_value)  {  autotemp = root->Next;  root->Next = root->Next->Next;  deletetemp;  break;  }  else  {  root = root->Next;  }  }  }  T \* Search(Tconst&search\_value)  {  inthash\_value = \_HashFunction(search\_value);  \_Node \*root = \_items[hash\_value]->Next;  while (root)  {  if (root->Item == search\_value)  {  return&(root->Item);  }  root = root->Next;  }  returnNULL;  }  voidDisplay()  {  for (inti = 0; i<\_m; ++i)  {  \_Node \*item = \_items[i]->Next; //跳过头结点  cout<<"槽["<<setw(3) <<i<<setw(3) <<"]";  while (item)  {  cout<<" -> "<<item->Item;  item = item->Next;  }  cout<<endl;  }  }  protected:  private:  /// 节点（使用单键表）  ///  /// 要是用双键表就会方便很多啊  struct\_Node  {  TItem;  \_Node \*Next;  };  /// 全域散列函数  ///  /// 本函数、一开始时随机选取\_a,\_b、再加上选取\_p,\_m的方法，就是全域散列的核心所在！！！  /// h(a, b, k) = ((a\*k + b) mod p) mod m  int\_HashFunction(Tk)  {  returnstatic\_cast<int>(\_a \* k + \_b) % \_p % \_m;  }  int\_p, \_m, \_a, \_b;  vector<\_Node \*>\_items; ///< 存储真正的键表，使用的是带头结点的单键表  };  inttest()  {  UniversalHashTable<int>table;  cout<<"开始往UniversalHashTable里添加内容[0,100)："<<endl;  for (inti = 0; i< 100; ++i)  {  table.Insert(i);  }  table.Display();  cout<<"开始删除内容[0,5)："<<endl;  for (inti = 0; i< 5; ++i)  {  table.Delete(i);  }  table.Display();  for (inti = 0; i< 10; ++i)  {  autofinded = table.Search(i);  cout<<"开始检索结点["<<i<<"]：";  if (finded)  {  cout<< \*finded<<endl;  }  else  {  cout<<"未找到"<<endl;  }  }  return 0;  }  }  Output:  开始往UniversalHashTable里添加内容[0,100)：  槽[ 0 ] -> 98 -> 92 -> 79 -> 73 -> 60 -> 54 -> 41 -> 35 -> 22 -> 16 -> 10  槽[ 1 ] -> 99 -> 86 -> 80 -> 67 -> 61 -> 48 -> 42 -> 29 -> 23 -> 4  槽[ 2 ] -> 93 -> 87 -> 74 -> 68 -> 55 -> 49 -> 36 -> 30 -> 17 -> 11  槽[ 3 ] -> 81 -> 75 -> 62 -> 56 -> 43 -> 37 -> 24 -> 18 -> 5  槽[ 4 ] -> 94 -> 88 -> 69 -> 63 -> 50 -> 44 -> 31 -> 25 -> 12 -> 6  槽[ 5 ] -> 95 -> 82 -> 76 -> 57 -> 51 -> 38 -> 32 -> 19 -> 13 -> 0  槽[ 6 ] -> 89 -> 83 -> 70 -> 64 -> 45 -> 39 -> 26 -> 20 -> 7 -> 1  槽[ 7 ] -> 96 -> 90 -> 77 -> 71 -> 58 -> 52 -> 33 -> 27 -> 14 -> 8  槽[ 8 ] -> 97 -> 84 -> 78 -> 65 -> 59 -> 46 -> 40 -> 21 -> 15 -> 2  槽[ 9 ] -> 91 -> 85 -> 72 -> 66 -> 53 -> 47 -> 34 -> 28 -> 9 -> 3  开始删除内容[0,5)：  槽[ 0 ] -> 98 -> 92 -> 79 -> 73 -> 60 -> 54 -> 41 -> 35 -> 22 -> 16 -> 10  槽[ 1 ] -> 99 -> 86 -> 80 -> 67 -> 61 -> 48 -> 42 -> 29 -> 23  槽[ 2 ] -> 93 -> 87 -> 74 -> 68 -> 55 -> 49 -> 36 -> 30 -> 17 -> 11  槽[ 3 ] -> 81 -> 75 -> 62 -> 56 -> 43 -> 37 -> 24 -> 18 -> 5  槽[ 4 ] -> 94 -> 88 -> 69 -> 63 -> 50 -> 44 -> 31 -> 25 -> 12 -> 6  槽[ 5 ] -> 95 -> 82 -> 76 -> 57 -> 51 -> 38 -> 32 -> 19 -> 13  槽[ 6 ] -> 89 -> 83 -> 70 -> 64 -> 45 -> 39 -> 26 -> 20 -> 7  槽[ 7 ] -> 96 -> 90 -> 77 -> 71 -> 58 -> 52 -> 33 -> 27 -> 14 -> 8  槽[ 8 ] -> 97 -> 84 -> 78 -> 65 -> 59 -> 46 -> 40 -> 21 -> 15  槽[ 9 ] -> 91 -> 85 -> 72 -> 66 -> 53 -> 47 -> 34 -> 28 -> 9  开始检索结点[0]：未找到  开始检索结点[1]：未找到  开始检索结点[2]：未找到  开始检索结点[3]：未找到  开始检索结点[4]：未找到  开始检索结点[5]：5  开始检索结点[6]：6  开始检索结点[7]：7  开始检索结点[8]：8  开始检索结点[9]：9  请按任意键继续. . . |

### 第12章：二叉查找树

1. 二叉查找树的定义：对任何结点X，其左子树中的关键字最大不超过key[X]；其右子树中的关键字最小不小于key[x]。
2. 首先先明显：二叉查找树上的基本操作的时间都与树的高度成正比的，所以高度越小的树性能越高。
3. 查询二叉查找树可以考虑使用非递归的版本，它运行要快得多而且也很容易理解  
   SEARCH(x, k):  
   while (x != NULL && x.Key != k){  
    if (k < x.Key) x = x.Left;  
    else x = x.Right;  
   }  
   return x;
4. 前趋和后继：  
   前趋：左一次，然后右到头；  
   后继：右一次，然后左到头。  
   错！不止这么简单，以后继为例：当结点的右子树不存在时，应该一路向上传递，直到找到根结点（没有后继）或者是找到一次非右子树传递（后继找到）为止。我的代码就在这里犯一次错误了，本以为很简单的！
5. 对二叉查找树的插入和删除操作也不复杂，唯一有点难度的地方就是在删除同时存在左右子树的结点时需要进行一些处理。  
   书上叙述的有点过度的复杂，其实可以很简单地说明白：对于这样的结点x，找到x结点的前趋（或后继）y，将x的值替换为y的值，然后递归删除y结点就可以了。因为y一定没有右子树（后继对应没有左子树），所以递归删除的时候就是很简单的情况了。
6. 可以证明：随机构造的二叉村在平均情况下的行为更接近于最佳情况下的行为，而不是接近最坏情况下的行为。所以一棵在n个关键字上随机构造的二叉查找树的期望高度为O(lgn)。

|  |
| --- |
| //////////////////////////////////////////////////////////////////////////  /// COPYRIGHT NOTICE  /// Copyright (c) 2009  /// All rights reserved.  ///  /// @file binary\_search\_tree.cpp  /// @brief binary\_search\_tree.cpp的简短描述  ///  /// binary\_search\_tree.cpp的详细描述  ///  /// @author 谭川奇 chuanqi.tan@gmail.com  /// @date 2011/05/30  /// @version 1.0  ///  ///  /// 修订说明：最初版本  //////////////////////////////////////////////////////////////////////////  #include"stdafx.h"  #include<iostream>  #include<algorithm>  #include<vector>  #include<iterator>  #include<iomanip>  #include<limits>  usingnamespacestd;  namespacechapter12  {  /// 二叉查找树  classBinarySearchTree  {  private:  /// 二叉查找树中的结点  struct\_Node  {  intValue;  \_Node \*Parent;  \_Node \*Left;  \_Node \*Right;  };  public:  BinarySearchTree() : \_root(NULL)  {}  ~BinarySearchTree()  {  //从根结点开始递归的析构  \_RecursiveReleaseNode(\_root);  }  /// 插入一个结点  ///  /// 向二叉查找树中插入一个值  /// @param new\_value 要插入的值  /// @return 是否插入成功，失败意味着树中已经存在该值  boolInsert(intconstnew\_value)  {  if (Search(new\_value))  {//已经存在  returnfalse;  }  if (!\_root)  {//插入的是第个节点  \_root = new\_Node();  \_root->Value = new\_value;  returntrue;  }  //非第个节点  \_Node \*current\_node = \_root;  while (current\_node)  {  if (new\_value>current\_node->Value)  {  if (current\_node->Right)  {  current\_node = current\_node->Right;  }  else  {  current\_node->Right = new\_Node();  current\_node->Right->Value = new\_value;  current\_node->Right->Parent = current\_node;  break;  }  }  if (new\_value<current\_node->Value)  {  if (current\_node->Left)  {  current\_node = current\_node->Left;  }  else  {  current\_node->Left = new\_Node();  current\_node->Left->Value = new\_value;  current\_node->Left->Parent = current\_node;  break;  }  }  }  returntrue;  }  /// 删除结点  ///  /// 在二叉查找树中删除一个值  /// @param delete\_value 要删除的值  /// @return 是否删除成功，删除失败意味着树中没有这个值的结点  boolDelete(intconstdelete\_value)  {  \_Node \*delete\_node = \_Search(\_root, delete\_value);  if (!delete\_node)  {//未找到该点  returnfalse;  }  else  {  \_DeleteNode(delete\_node);  returntrue;  }  }  /// 查找元素  ///  /// 在当前二叉查找树中查找某一值  /// @param search\_value 要查找的值  /// @return 是否在二叉树中找到值为search\_value的结点  /// @retval true 查找到了该元素  /// @retval false 找不到该元素  boolSearch(intconstsearch\_value) const  {  return\_Search(\_root, search\_value) != NULL;  }  /// 显示当前二叉查找树的状态  ///  /// 使用Graphviz的dot语言画出当前二叉查找树  voidDisplay() const  {  cout<<"digraph graphname"<< (rand() % 1000) <<"{"<<endl  <<" node [shape = record,height = .1];"<<endl;  \_Display(\_root);  cout<<"}"<<endl;  }  private:  void\_DeleteNode( \_Node \* delete\_node )  {  if (delete\_node->Left&&delete\_node->Right)  {//要删除的结点同时存在左子树和右子树  //前驱结点：前驱一定存在，因为该结点同时存在左右子树  \_Node \*previous\_node = \_GetPreviousNode(delete\_node);  delete\_node->Value = previous\_node->Value;  //previous\_nde一定没有右子树，所以再递归调用一定是走这个if的else分支  \_DeleteNode(previous\_node);  }  else  {//要删除的结点至少有一个子结点为空  //sub\_node为delete\_node的子树  //sub\_node要么为delete\_node的左子树，要么为delete\_node的右子树，或者在delete\_node无子结点时为空  \_Node \*sub\_node = (delete\_node->Left ? delete\_node->Left : delete\_node->Right);  if (delete\_node->Parent == NULL)  {//是第个结点  \_root = sub\_node;  }  else  {  if (delete\_node->Parent->Left == delete\_node)  {  delete\_node->Parent->Left = sub\_node;  }  else  {  delete\_node->Parent->Right = sub\_node;  }  if (sub\_node)  {//在delete\_node有子结点时需要设置子结点的Parent指针  sub\_node->Parent = delete\_node->Parent;  }  }  deletedelete\_node;  }  }  //得到一个节点的前驱  \_Node \* \_GetPreviousNode( \_Node \* node )  {  \_Node \*previous\_node = node;  //求结点的前驱：先左一下，再右到头  previous\_node = previous\_node->Left;  while (previous\_node->Right)  {  previous\_node = previous\_node->Right;  }  returnprevious\_node;  }  void\_RecursiveReleaseNode(\_Node \*node)  {  if (node)  {  \_RecursiveReleaseNode(node->Left);  \_RecursiveReleaseNode(node->Right);  deletenode;  }  }  /// 非递归查找一个结点  \_Node \* \_Search(\_Node \*node, intconstsearch\_value) const  {  while (node&&node->Value != search\_value)  {  if (search\_value<node->Value)  {  node = node->Left;  }  else  {  node = node->Right;  }  }  //到这里如果node为空就是未找到  returnnode;  }  void\_Display(\_Node \*node) const  {  if (node)  {  cout<<" node"<<node->Value<<"[label = \"<f0>|<f1>"<<node->Value<<"|<f2>\"];"<<endl;  if (node->Left)  {  cout<<" \"node"<<node->Value<<"\":f0 -> \"node"<<node->Left->Value<<"\":f1;"<<endl;  \_Display(node->Left);  }  if (node->Right)  {  cout<<" \"node"<<node->Value<<"\":f2 -> \"node"<<node->Right->Value<<"\":f1;"<<endl;  \_Display(node->Right);  }  }  }  \_Node \*\_root; ///< 二叉查找树的根结点  };  inttest()  {  BinarySearchTreebst;  //用随机值生成一棵二叉查找树  for (inti = 0; i< 100; ++i)  {  bst.Insert(rand() % 100);  }  bst.Display();  //删除所有的奇数值结点  for (inti = 1; i< 100; i += 2)  {  if(bst.Delete(i))  {  cout<<"### Deleted ["<<i<<"] ###"<<endl;  }  }  bst.Display();  //查找以内的数，如果在二叉查找树中，则显示  cout<<endl;  for (inti = 0; i< 100; i += 1)  {  if (bst.Search(i))  {  cout<<"搜索["<<i<<"]元素：\t成功"<<endl;  }  }  return 0;  }  }  Outputs:  digraph graphname705{  node [shape = record,height = .1];  node41[label = "<f0>|<f1>41|<f2>"];  "node41":f0 -> "node34":f1;  node34[label = "<f0>|<f1>34|<f2>"];  "node34":f0 -> "node0":f1;  node0[label = "<f0>|<f1>0|<f2>"];  "node0":f2 -> "node24":f1;  node24[label = "<f0>|<f1>24|<f2>"];  "node41":f2 -> "node67":f1;  node67[label = "<f0>|<f1>67|<f2>"];  "node67":f0 -> "node58":f1;  node58[label = "<f0>|<f1>58|<f2>"];  "node58":f2 -> "node62":f1;  node62[label = "<f0>|<f1>62|<f2>"];  "node62":f2 -> "node64":f1;  node64[label = "<f0>|<f1>64|<f2>"];  "node67":f2 -> "node69":f1;  node69[label = "<f0>|<f1>69|<f2>"];  "node69":f2 -> "node78":f1;  node78[label = "<f0>|<f1>78|<f2>"];  }  ### Deleted [41] ###  ### Deleted [67] ###  ### Deleted [69] ###  digraph graphname145{  node [shape = record,height = .1];  node34[label = "<f0>|<f1>34|<f2>"];  "node34":f0 -> "node0":f1;  node0[label = "<f0>|<f1>0|<f2>"];  "node0":f2 -> "node24":f1;  node24[label = "<f0>|<f1>24|<f2>"];  "node34":f2 -> "node64":f1;  node64[label = "<f0>|<f1>64|<f2>"];  "node64":f0 -> "node58":f1;  node58[label = "<f0>|<f1>58|<f2>"];  "node58":f2 -> "node62":f1;  node62[label = "<f0>|<f1>62|<f2>"];  "node64":f2 -> "node78":f1;  node78[label = "<f0>|<f1>78|<f2>"];  }  搜索[0]元素：成功  搜索[24]元素：成功  搜索[34]元素：成功  搜索[58]元素：成功  搜索[62]元素：成功  搜索[64]元素：成功  搜索[78]元素：成功  请按任意键继续. . . |

### 第13章：红黑树

满足下面几个条件(红黑性质)的二叉搜索树，称为**红黑树**：

1. 每个结点或是红色，或是是黑色。

2. 根结点是黑的。

3. 所有的叶结点(NULL)是黑色的。（NULL被视为一个哨兵结点，所有应该指向NULL的指针，都看成指向了NULL结点。）

4. 如果一个结点是红色的，则它的两个儿子节点都是黑色的。

5. 对每个结点，从该结点到其子孙结点的所有路径上包含相同数目的黑结点。

1. 黑高度的定义： 从某个结点出发(不包括该结点)到达一个叶结点的任意一条路径上，黑色结点的个数成为该结点x的黑高度。红黑树的黑高度定义为其根结点的黑高度。
2. 红黑树是真正的在实际中得到大量应用的复杂数据结构：C++STL中的关联容器map,set都是红黑树的应用（所以标准库容器的效率太好了，能用标准库容器尽量使用标准库容器）；Linux内核中的用户态地址空间管理也使用了红黑树。
3. 红黑权势是许多“平衡的”查找树中的一种（首先：红黑树是一种近似平衡的二叉树），它能保证在最坏的情况下，基本的动态集合操作的时间为O(lgn)。
4. 通过对任何一条从根到叶子的路径上各个结点着色方式的限制，红黑树确保没有一条路径会比其它路径长出两倍，因而是接近平衡的。
5. 一要全是黑结点的满二叉树也满足红黑树的定义。满二叉树的效率本身就非常高啊，它是效率最好的二叉树了，所以说它是红黑树的一个特例；普通的红黑树要求并没有满二叉树这么严格。
6. 旋转操作（左旋和右旋）：旋转操作是一种能保持二叉查找树性质的查找树局部操作。  
   所有对红黑树结构的修改都只能通过左右旋来完成，这样才能保证修改后的红黑树首先是一棵二叉查找树。
7. 红黑树的插入操作：将结点Z插入树T中，就好像T是一棵普通的二叉查找树一样，然后将Z着为红色。为保证红黑性质能继续保持，我们调用一个辅助程序来对结点重新着色并旋转。  
   这么做是有它的智慧的：首先，插入结点Z的位置的确应该和普通二叉查找树一样，因为红黑树本身就首先是一棵二叉查找树；然后将Z着为红色，是为了保证性质5的正确性，因为性质5如果被破坏了是最难以恢复的；到这里，有可能被破坏的性质就只剩下性质2和性质4了，这都可以通过后来的辅助程序进行修复的。  
   插入操作可能破坏的性质：
   1. 性质2：当被一棵空树进行插入操作时发生；
   2. 性质4：当新结点被插入到红色结点之后时发生；
8. 红黑树的删除操作：和插入操作一样，先用BST的删除结点操作，然后调用相应的辅助函数做相应的调整。  
   首先只有被删除的结点为黑结点时才需要进行修补，理由如下：
   1. 树中各结点的黑高度都没有变化
   2. 不存在两个相邻的红色结点
   3. 因为如果被删除的点是红色，就不可能是根，所以根仍然是黑色的
9. 当被删除了黑结点之后，红黑树的性质5被破坏，上面说过了性质5被破坏后的修复难度是最大的。所以这里的修复过程使用了一个很新的思想，即视为被删除的结点的子结点有额外的一种黑色，当这一重额外的黑色存在之后，性质5就得到了继续。然后再通过转移的方法逐步把这一重额外的黑色逐渐向上转移直到根或者红色的结点，最后消除这一重额外的黑色。  
   删除操作中可能被破坏的性质：
   1. 性质2：当y是根时，且y的一个孩子是红色，若此时这个孩子成为根结点；
   2. 性质4：当x和p[y]都是红色时；
   3. 性质5：包含y的路径中，黑高度都减少了；
10. 红黑树是平衡查找树，还有B树也是另一类平衡查找树。

### 第14章：数据结构的扩张

1. 实际的工程中，极少会去创造新的数据结构，通常是对标准的数据结构附加一些信息，并添加一些新的操作以支持应用的要求。
2. 数据结构的扩张：指在实际应用数据结构时对标准的数据结构中增加一些信息、编入一些新的操作等等。附加的信息必须能够为该数据结构上的常规操作所更新和维护。
3. 对一种数据结构的扩张过程可以分为四个步骤：
   1. 选择基础的数据结构
   2. 确定要在基础数据结构中添加哪些信息
   3. 验证可用基础数据结构上的基本修改操作来维护这些新添加的信息
   4. 设计新的操作
4. 红黑树的扩张定理：当结点中新添加的信息可以由该结点和它的左右子树来决定，那么就可以在不影响时间复杂度的前提下在插入和删除等操作中对红黑树的这些附加信息进行维护。

## 第四部分：高级设计和分析技术

设计和分析高效算法的三种重要技术：动态规划、贪心算法和平摊分析。

动态规划通常应用于最优化问题，即要做出一组选择以达到一个最优解时。在做选择的同时，经常出现同样形式的子问题。关键技术是存储这些子问题每一个的解，以备它重复出现。

贪心算法通常也是应用于最优化问题，该算法的思想是以局部最优的方式来做每一个选择。采用贪心算法可以比动态规划更快的得出一个最优解，但是关键是不容易判断贪心算法所得到的是否真的是最优解。  
贪心算法可以在一定的理论之下，通过只考虑局部最优解就可以保证得到的一定是全局最优解。如赫夫曼编码。

平摊分析是一种用来分析执行一系列类似操作的算法的工具。在一个操作序列中，不可能每一个都以其已知的最坏情况运行，某些操作的代价高些，而其它的低一些。

所有的最优化问题都可以通过穷举法来解决，但是这在时间上是不可接受的。

所有的高效算法都是为了加快速度：

1. 动态规划：保存子问题的解，以备重复使用
2. 贪心算法：用局部最优解来求全局最优解

平摊分析是用来分析算法的工具，它本身并不是一种算法。

### 第15章：动态规划

1. 动态规划与分治法之间的区别：
   1. 分治法是指将问题分成一些独立的子问题，递归的求解各子问题
   2. 动态规划适用于这些子问题不是独立的情况，也就是各子问题包含公共子问题
2. 动态规划算法的设计可以分为4个步骤：
   1. 描述最优解的子结构
   2. 递归定义最优解的值
   3. 按自底向上的方法计算最优解的值
   4. 由计算出的结果反向构造出一个最优解
3. 动态规划最最最重要的就是要找出最优解的子结构！
4. 最优子结构在问题域中以两种方式变化（在找出这两个问题的解之后，构造出原问题的最优子结构往往就不是难事了）：
   1. 有多少个子问题被用在原问题的一个最优解中
   2. 在决定一个最优解中使用哪些子问题有多少个选择
5. 动态规划说白了就是一个递归的反向展开的过程：在满足①最优子结构②重叠子问题这2个条件下，通过把递归从下至上的进行展开以避免重复计算子问题从而加速了最终问题的求解的过程。
6. 再次强调“动态规划最关键的一步就是：寻找最优子结构”
7. 动态规划能够消除重复计算子问题是因为它与普通递归相反，它是通过自下而上的方式来进行求解的。
8. 正确使用动态规划方法的2个关键要素：最优子结构 和 重叠子问题。  
   如果问题可以由递归来解决，并且在递归的过程中会不断的出现重复的子问题需要解决，那毫不犹豫的采用动态规划吧！
9. 如果问题的一个最优解中包含了子问题的最优解，则该问题具有最优子结构；而当一个问题具有最优子结构时，提示我们动态规划可能会适用。
10. 剪贴技术：用来证明在问题的一个最优解中，使用的子问题的解本身也必须是最优的
11. 为了描述子问题空间，可以遵循这样一条有效的经验规则，就是尽量保持这个空间简单，然后在需要时再扩充它
12. 非正式地：一个动态规划算法的运行时间依赖于两个因素的乘积：子问题的总个数和每个子问题中有多少种选择。  
    问题解的代价通常是子问题的代价加上选择本身带来的开销。
13. 贪心算法与动态规划有一个显著的区别：就是在贪心算法中，是以自顶向下的方式使用最优子结构的。贪心算法会先做选择，在当时看起来是最优的选择，然后再求解一个结果子问题，而不是先寻找子问题的最优解，然后再选择。
14. 要注意：在不能应用最优子结构的时间，就一定不能假设它能够应用。  
    警惕使用动态规划去解决缺乏最优子结构的问题！
15. 使用动态规划时：子问题必须是相互独立的！可以这样理解，N个子问题域互不相干，属于完全不同的空间。  
    子问题必须独立！
16. 重叠子问题：不同的子问题的数目是输入规模的一个多项式。  
    这样，动态规划算法才能充分利用重叠的子问题，减少计算量。即通过每个子问题只解一次，把解保存在一个需要时就可以查看的表中，而每次查表只需要常数时间。  
    从这段描述可以看出：动态规划与递归时做备忘录的本质是完全相同的，所以说备忘录方法与普通的动态递归本质完全相同，没有孰优孰劣之分，哪个方便用哪个。
17. 由计算出的结果反向构造一个最优解：把动态规划或者是递归过程中作出的每一次选择（记住：保存的是每次作出的选择）都保存下来，在最后就一定可以通过这些保存的选择来反向构造出最优解。
18. 做**备忘录**的递归方法：这种方法是动态规划的一个变形，它本质上与动态规划是一样的，但是比动态规划更好理解！
    1. 使用普通的递归结构，自上而下的解决问题。
    2. 当在递归算法的执行中每一次遇到一个子问题时，就计算它的解并填入一个表中。以后每次遇到该子问题时，只要查看并返回表中先前填入的值即可。
19. 备忘录方法与动态递归方法的比较：如果所有的子问题都至少要被计算一次，则一个自底向上的动态规划算法通常比一个自顶向下的做备忘录算法好出一个常数因子。因为动态规划没有使用递归的代价，只用到了循环，所以常数因子肯定比递归要好一些。  
    此外，在有些问题中，还可以用动态规划算法中的表存取模式来进一步的减少时间和空间上的需求；或者，如果子问题中的某些子问题根本没有必须求解，做备忘录的方法有着只解那些肯定要求解的子问题的优点。（而且这点是自动获得的，那些不必要计算的子问题在备忘录方法中会被自动的抛弃）
20. 备忘录方法总结：由“是否所有的子问题都至少需要被计算一次”来决定使用动态规划还是备忘录。  
    再次下定义：这两种方法没孰优孰劣之分，因为它们的本质思想是完全一样的；消除重复子问题。
21. 动态规划：最重要最重要的就是找到最优子结构。在找到最优子结构之后的消除重复子问题，这点我太容易处理了，无论是动态规划的自底向上的递推，还是备忘录，或者是备忘录的变型，都可以轻松的应付。关键就是最优子结构。

### 第16章：贪心算法

1. 贪心算法使所作的选择看起来都是当前最佳的，期望通所做的局部最优选择来产生一个全局最优解。  
   对算法中的每一个决策点，做一个当时看起来是最佳的选择，这种启发式策略并不是总能产生最优解的。
2. 贪心算法对大多数优化问题能产生最优解，但很多问题也不能通过这样的局部最优解得到全局最优解（0-1背包）。实际上，使用贪心算法最关键的就是在于如何证明贪心选择性质以证明当前的问题可以由贪心算法得到最优解
3. 在活动选择问题中，非常关键的一点先决条件就是：活动要先按照结束时间的单调递增顺序排序，这样才能贪心的选择第一个。  
   从直觉上来看，这种活动选择方法是一种“贪婪的”选择方法，它给后面剩下的待调度任务留下了尽可能多的机会。也就是说，此处的贪心选择使得剩下的、未调度的时间最大化
4. 可以说：动态规划其实是“安全的贪心算法”的基础。无论如何，在每一个贪心算法的下面，几乎总是会有一个更加复杂的动态规划解。  
   所谓安全的贪心算法就是指一定能产生出全局最优解的贪心算法
5. 贪心算法的一般步骤（定义）
   1. 将优化问题转化成这样的一个问题，即先做出选择（对应于动态规划的先解决子问题再选择），再解决剩下的一个子问题。
   2. 证明原问题总是有一个最优解是做贪心选择得到的，从而说明贪心选择的安全。
   3. 说明在做出贪心选择后，剩余的子问题具有这样的一个性质。即如果将子问题的最优解和我们所做的贪心选择联合起来，就可以得出原问题的一个最优解。
6. 正确使用贪心算法的2个关键要素：贪心选择性质 和 最优子结构。
   1. 贪心选择性质：一个全局最优解可以通过局部最优（贪心）选择来达到；即当考虑做何选择时，我们只考虑对当前问题最佳的选择而不考虑子问题的结果。
   2. 最优子结构：一个问题的最优解包含了其子问题的最优解。
7. 要使用贪心算法就必须先证明以下两个性质：
   1. 每一步所做的贪心选择最终能产生一个全局最优解。  
      在证明中先考察一个全局最优解，然后证明对该解加以修改，使其采用贪心选择，这个选择将原问题变为一个相似的、但更小的问题。
   2. 子问题的最优解与所做的贪心选择合并后，的确可以得到原问题的一个最优解。
8. 贪心算法所做的当前选择可能要依赖于已经做出的所有选择，但不依赖于有待于做出的选择或子问题的解。
9. 前缀编码：在所有的编码方案中，没有一个编码是另一个编码的前缀。
10. 一般地，就算证明不出来贪心算法能给出最优解，但是它一般都至少能给出次优解。所以贪心算法在实际的应用中是非常的普及的。

### 第17章：平摊分析

1. 在平摊分析中，执行一系列数据结构操作所需要的时间是通过对执行所有操作所花费的时间求平均而得到的。  
   平摊分析与平均情况分析的不同之处在于它不牵涉到概率；平摊分析保证在最坏情况下，每个操作具有平均性能。
2. 平摊分析表明：并不能总以最坏情况来衡量算法，因为最坏的情况并不总会发生，而且在绝大多数应用中最坏情况出现的次数一定是很少很少的。  
   最明显的例子就是C++ STL中vector.push\_back操作。
3. 通过平摊分析，可以获得对某种特定数据结构的认识，这种认识有助于优化设计。
4. 三种方法进行平摊分析：
   1. 聚集分析：就是指分析一系列操作的总时间
   2. 记账法：对每次操作的对象进行预先记账
   3. 势能方法：与记账法类似，但是将每次预留的势能视为是整个数据结构共享的
5. 聚集分析：由N个操作所构成的序列的总时间在最坏的情况下为T(n)，则每个操作的平均代价（平摊代价）为T(n)/n。  
   以前的时间复杂度分析都是以单次操作为对象的分析它的最坏时间复杂度，而聚集分析所分析的是N次操作的总时间的最坏情况。应注意其与平均情况分析的不同之处！  
   聚集分析：由序列的总最坏时间🡺单次的平摊时间
6. 记账法：对序列操作中的每一个操作收取一定的费用，当所收取的费用比它实际应支付的费用多时就把多余的部分当作存款存起来，一个操作的平摊代价可以看作两部分：实际代价和存款（或被储蓄或被用完）。
   1. 存款可以用来在以后补偿那些其平摊代价低于其实际代价的操作。
   2. 如果希望通过对平摊代价的分析来说明每次操作的最坏情况平均代价较小，则操作序列的总平摊代价就必须是该序列的总的实际代价的一个上界。因为平摊代价是最坏时的平均代价，这意味着：只能有存款而不能有欠账！
   3. 明白这个存款不能为负很重要，因为这种分析方法是不允许欠账的，否则就不能满足平摊代价是操作总时间的最坏情况的平均这个定义。
   4. Eg: 对于栈的操作，当每次入栈时记账支付2元，1元支付该PUSH操作的实际代价，还有1元用于支付该元素被POP出来时的代价。则可以保证在任何时间内都不会有欠账，POP操作可以不收取任何费用。
7. 势能方法：将已预付的工作作为一种“势能”保存，它在需要时可以释放出来，以支付后面的操作。势能是与整个数据结构而不是其中的个别对象发生联系的。  
   （记账法中的账与个别对象发生联系，比如栈操作时支付的2元就记在入栈的那个元素上）
8. 动态表：比如C++ STL中的vector
   1. 可以通过平摊分析证明：插入和删除操作的平摊代价都仅为O(1)
   2. 表扩张时：常用的启发式（启发信息就是容器原来的大小）方法是分配一个原表两倍大小的新表。
   3. 可以通过聚集分析来证明插入操作的平摊代价：  
      所以每次操作的平摊代价为O(1)=3
   4. 通过记账法来证明插入、删除操作的平摊代价：
      1. 插入操作：每次插入操作需要支付3元的代价，分别是：将其自身插入到当前表中、当表扩张时其自身的移动、以及对另一个在扩张表时已经移动过的另一项的移动。于1+1+1=3元的代价

|  |  |
| --- | --- |
| 原有的元素…… | NewItem |
| ↑在原有的元素里还要再支付一个可能移动的代价 | ↑支付自身以后可能的移动操作的代价 |

* + 1. 并且可以证明：无论扩张时分配多少倍的新空间，每次插入操作的平摊代价都大于2。所以标准库中取每次扩张时进行翻倍，这是一个均衡的很好的选择。
    2. 删除操作：
       1. C++STL的vector不会在删除元素时进行自动收缩，因为如果在元素不足1/2时进行收缩，最坏时会产生在这的情况（来回收缩和扩张）：在一次扩张之后，还没有做足够的删除来支付一次收缩的代价。类似地，在一次收缩之后，也有可能没有做足够的插入来支付一次扩张的代价。  
          所以vector选择了在删除元素时不自动收缩来回避这个问题。如果要进行收缩，应该是元素减少为1/4时进行自动收缩，这样才能保证在表收缩之后有足够的存款进行支付一次收缩的代价。
       2. 当删除至1/4个元素时进行自动收缩的平摊代价分析：  
          每次删除操作需要支付的代价为：其自身的删除操作代价 和 剩下的元素可能进行的移动操作的代价。  
          即最坏为：1 + 1 = 2元的代价，最好为1 + 1.33333 = 2.3333元的代价，取最坏情况为2元。

|  |  |  |  |
| --- | --- | --- | --- |
| 剩下的元素…… |  |  |  |
| ↑还要支付一个可能的移动 |  | ↑最早可能从这开始删除 | 最晚可能从这开始删除↑ |

* + 1. 因此：对于 动态表 的插入和删除操作而言，每个操作的平摊代价都有一个常数的上界。

## 第五部分：高级数据结构

**参考：**

* 1. 《算法导论》第2版，机械工业出版社
  2. <http://staff.ustc.edu.cn/~lszhuang/alg/>
  3. <http://www.wutianqi.com/>
  4. <http://www.cnblogs.com/timebug/>
  5. <http://blog.csdn.net/v_JULY_v/>