

普通高校本科计算机专业特色教材精选·算法与程序设计

算法设计与分析

张军 钟竞辉 等编著

清华大学出版社



普通高校本科计算机专业特色教材精选·算法与程序设计

算法设计与分析

张 军 钟竞辉

龚月姣 詹志辉

许瑞填 周 琦

林 盈 陈 霓

编著

清华大学出版社
北京

内 容 简 介

本书对算法设计与分析的基本原理、常用的经典算法以及新兴发展的智能算法进行介绍,重点对各种算法的思想、流程结构以及具体的实践应用过程等方面进行介绍。本书内容包括绪论、基本数据结构、蛮力算法、分治算法、贪心算法、动态规划算法、回溯算法、分支限界算法、概率算法等经典算法的思想和原理,同时还介绍了人工神经网络、模糊逻辑、遗传算法、蚁群算法、粒子群优化算法、差分进化算法,以及分布估计算法等现代计算智能算法。本书通俗易懂,图文并茂,深入浅出,避免其他算法书中大量公式、定理、证明等难懂的内容,相反通过大量的图表示例对各个算法进行说明和介绍,不但提供了算法的伪代码,而且通过具体的应用举例对算法的使用方法和使用过程进行说明,以利于读者快速掌握算法分析与设计的原理和精髓。

本书适于作为相关专业本科生和研究生的必修课或选修课教材,同时还可以作为广大算法研究者和工程技术人员的参考书和工具书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

算法设计与分析/张军等编著. —北京:清华大学出版社,2011.8

(普通高校本科计算机专业特色教材精选·算法与程序设计)

ISBN 978-7-302-25401-0

I. ①算… II. ①张… III. ①电子计算机—算法设计—高等学校—教材 ②电子计算机—算法分析—高等学校—教材 IV. ①TP301.6

中国版本图书馆CIP数据核字(2011)第071756号

责任编辑:袁勤勇 薛 阳

责任校对:梁 毅

责任印制:杨 艳

出版发行:清华大学出版社

<http://www.tup.com.cn>

社 总 机:010-62770175

投稿与读者服务:010-62795954,jsjcc@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015,zhiliang@tup.tsinghua.edu.cn

地 址:北京清华大学学研大厦A座

邮 编:100084

邮 购:010-62786544

印 装 者:北京国马印刷厂

经 销:全国新华书店

开 本:185×260 印 张:13.5

字 数:321千字

版 次:2011年8月第1版

印 次:2011年8月第1次印刷

印 数:1~3000

定 价:19.50元

产品编号:039232-01

出版说明

在我国高等教育逐步实现大众化后，越来越多的高等学校将会面向国民经济发展的第一线，为行业、企业培养各级各类高级应用型专门人才。为此，教育部已经启动了“高等学校教学质量和教学改革工程”，强调要以信息技术为手段，深化教学改革和人才培养模式改革。如何根据社会的实际需要，根据各行各业的具体人才需求，培养具有特色显著的人才，是我们共同面临的重大问题。具体地说，培养具有一定专业特色的和特定能力强的计算机专业应用型人才是计算机教育要解决的问题。

为了适应 21 世纪人才培养的需要，培养具有特色的计算机人才，急需一批适合各种人才培养特点的计算机专业教材。目前，一些高校在计算机专业教学和教材改革方面已经做了大量工作，许多教师在计算机专业教学和科研方面已经积累了许多宝贵经验。将他们的教研成果转化为教材的形式，向全国其他学校推广，对于深化我国高等学校的教学改革是一件十分有意义的事情。

清华大学出版社在经过大量调查研究的基础上，决定组织出版一套“普通高校本科计算机专业特色教材精选”。本套教材是针对当前高等教育改革的新形势，以社会对人才的需求为导向，主要以培养应用型计算机人才为目标，立足课程改革和教材创新，广泛吸纳全国各地的高等院校计算机优秀教师参与编写，从中精选出版确实反映计算机专业教学方向的特色教材，供普通高等院校计算机专业学生使用。

本套教材具有以下特点：

1. 编写目的明确

本套教材是在深入研究各地各学校办学特色的基础上，面向普通高校的计算机专业学生编写的。学生通过本套教材，主要学习计算机科学与技术专业的基本理论和基本知识，接受利用计算机解决实际问题的基本训练，培养研究和开发计算机系统，特别是应用系统的基本能力。

2. 理论知识与实践训练相结合

根据计算学科的三个学科形态及其关系,本套教材力求突出学科的理论与实践紧密结合的特征,结合实例讲解理论,使理论来源于实践,又进一步指导实践。学生通过实践深化对理论的理解,更重要的是使学生学会理论方法的实际运用。在编写教材时突出实用性,并做到通俗易懂,易教易学,使学生不仅知其然,知其所以然,还要会其如何然。

3. 注意培养学生的动手能力

每种教材都增加了能力训练部分的内容,学生通过学习和练习,能比较熟练地应用计算机知识解决实际问题。既注重培养学生分析问题的能力,也注重培养学生解决问题的能力,以适应新经济时代对人才的需要,满足就业要求。

4. 注重教材的立体化配套

大多数教材都将陆续配套教师用课件、习题及其解答提示,学生上机实验指导等辅助教学资源,有些教材还提供能用于网上下载的文件,以方便教学。

由于各地区各学校的培养目标、教学要求和办学特色均有所不同,所以对特色教学的理解也不尽一致,我们恳切希望大家在使用教材的过程中,及时地给我们提出批评和改进意见,以便我们做好教材的修订改版工作,使其日趋完善。

我们相信经过大家的共同努力,这套教材一定能成为特色鲜明、质量上乘的优秀教材。同时,我们也希望通过本套教材的编写出版,为“高等学校教学质量和教学改革工程”做出贡献。

清华大学出版社

前言

算法是计算机科学领域最重要的研究内容之一。著名的算法大师 Donald Knuth 甚至认为“计算机科学就是算法的研究”，因为计算机科学的各个应用领域都高度依赖于算法设计。而且，随着社会的发展，人们对计算机信息处理能力的要求越来越高，算法在计算机的普及、应用和发展中的核心地位愈加突出。在互联网诞生后，人们每天增加的信息量（包括文本、音频和视频等）呈指数级增长，仅仅依靠计算机运算能力的增强已无法满足人们日益增长的信息处理需求，只有设计出更高效的算法才是满足这种需求的关键。事实上，设计高效的算法从而提高计算机解决现实应用问题的效率是当前计算机在各应用领域都不可回避的重要研究课题。

自计算机问世后的半个多世纪以来，计算机已经在社会生活中的广泛领域取得了成功的应用。从航空航天到生物医疗、从工业控制到农业生产，都有计算机应用的例子，其中涉及的算法更是不可胜数。然而，算法设计与分析这一门学科并不是包罗万象地设计一些求解具体问题的方法，而是从方法学的角度来研究算法设计的一般原理以及算法分析的一般方法。算法设计与分析的研究内容主要包括算法设计与分析的一般步骤、算法的复杂性定义和复杂性分析、经典的确定性算法（如贪心法、分治法、动态规划、回溯算法等）、常用的概率性算法等。算法是沟通人和计算机之间的桥梁，只有熟练地掌握算法设计与分析的原理，才能更好地利用计算机强大的计算能力解决问题。

本书将对算法设计与分析的一般原理进行介绍，重点是对各种经典算法的思想、算法流程以及具体实践应用的步骤等方面进行介绍。内容共分为 10 章，第 1 章是绪论，主要是对算法的基本概念、算法设计的基本步骤以及复杂性分析的基本原理等进行概要介绍；第 2 章介绍基本数据结构（Data Structure）；第 3 章介绍蛮力算法（Brute Force Search）；第 4 章介绍分治算法（Divide and Conquer）；第 5 章介绍贪心算法（Greedy Algorithm）；第 6 章介绍动态规划算法（Dynamic Programming）；

第7章介绍回溯算法 (Backtracking Algorithm); 第8章介绍分支限界算法 (Branch and Bound Algorithm); 第9章介绍概率算法 (Probabilistic Algorithm); 第10章介绍计算智能 (Computational Intelligence)。

本书在对各种算法进行介绍时,力求做到通俗易懂,图文并茂,深入浅出。本书的每一章都以生动的图示开头,尝试用最直观、最通俗的形式去展示各种算法的思想原理和基本特征,以利于读者快速、形象、深刻地对算法进行认识和把握。在叙述的过程中,本书避免其他算法书中大量公式、定理、证明等难懂的内容,而是通过大量的图表示例对各个算法的思想、算法原理以及具体的实践应用过程等方面进行说明和介绍,不但提供了算法实现的伪代码,而且通过具体的应用举例对算法的使用方法和使用过程进行说明,同时精选了一些具有代表性的经典习题供读者进一步深入学习和理解。本书的部分插图来源于网络开放资源和 Microsoft Office 软件提供的剪贴画素材,我们对此表示感谢。

本书编写的分工如下:中山大学的张军教授负责全书的编写与统稿工作,钟竞辉参与编写了第1、6、10章,龚月姣参与编写了第2章,詹志辉参与编写了第3、4章,许瑞填参与编写了第5章,周琦参与编写了第7章,林盈参与编写了第8章,陈霓参与编写了第9章。

由于编者水平有限,书中难免存在不足和疏漏之处,希望广大读者批评指正。

张 军
2011年5月
于中山大学

目 录

| | |
|---------------------------|-----------|
| 第 1 章 绪论 | 1 |
| 1.1 算法的基本概念 | 2 |
| 1.1.1 算法的重要性 | 2 |
| 1.1.2 算法设计与分析的流程 | 3 |
| 1.2 算法设计与分析的重要问题类型 | 4 |
| 1.2.1 排序问题 | 4 |
| 1.2.2 查找问题 | 4 |
| 1.2.3 图问题 | 5 |
| 1.2.4 组合问题 | 5 |
| 1.2.5 数值问题 | 5 |
| 1.2.6 几何问题 | 6 |
| 1.3 算法复杂性分析基础 | 6 |
| 1.3.1 算法复杂性分析的原理 | 6 |
| 1.3.2 渐进符号 | 8 |
| 1.4 本章小结 | 9 |
| 1.5 习题 | 9 |
| | |
| 第 2 章 基本数据结构 | 11 |
| 2.1 数据结构的概念 | 12 |
| 2.2 线性结构 | 15 |
| 2.2.1 线性表 | 15 |
| 2.2.2 栈 | 18 |
| 2.2.3 队列 | 20 |
| 2.2.4 串 | 21 |
| 2.3 树形结构 | 23 |
| 2.3.1 树的定义与性质 | 23 |
| 2.3.2 二叉树 | 24 |
| 2.3.3 多叉树 | 27 |

| | | |
|------------|-------------|-----------|
| 2.4 | 图状结构 | 28 |
| 2.4.1 | 图的定义 | 28 |
| 2.4.2 | 图的存储结构 | 29 |
| 2.4.3 | 图的遍历 | 31 |
| 2.5 | 集合与字典 | 33 |
| 2.5.1 | 集合 | 33 |
| 2.5.2 | 字典 | 34 |
| 2.6 | 本章小结 | 35 |
| 2.7 | 习题 | 35 |
| 第3章 | 蛮力算法 | 37 |
| 3.1 | 算法设计思想 | 38 |
| 3.2 | 排序问题中的蛮力算法 | 39 |
| 3.2.1 | 选择排序 | 39 |
| 3.2.2 | 冒泡排序 | 40 |
| 3.3 | 查找问题中的蛮力算法 | 41 |
| 3.3.1 | 顺序查找算法 | 41 |
| 3.3.2 | 串匹配算法 | 42 |
| 3.4 | 组合问题中的蛮力算法 | 43 |
| 3.4.1 | 旅行商问题 | 43 |
| 3.4.2 | 背包问题 | 44 |
| 3.4.3 | 任务分配问题 | 45 |
| 3.5 | 几何问题中的蛮力算法 | 46 |
| 3.5.1 | 最近点对问题 | 46 |
| 3.5.2 | 凸包问题 | 46 |
| 3.6 | 本章小结 | 47 |
| 3.7 | 习题 | 48 |
| 第4章 | 分治算法 | 51 |
| 4.1 | 算法设计思想 | 52 |
| 4.2 | 排序问题中的分治算法 | 53 |
| 4.2.1 | 归并排序 | 53 |
| 4.2.2 | 快速排序 | 55 |
| 4.3 | 查找问题中的分治算法 | 56 |
| 4.3.1 | 折半查找 | 56 |
| 4.3.2 | 二叉树遍历算法 | 57 |
| 4.4 | 组合问题中的分治算法 | 58 |
| 4.4.1 | 最大子段和问题 | 58 |

| | | |
|------------|------------------------------|-----------|
| 4.4.2 | 棋盘覆盖问题 | 59 |
| 4.5 | 几何问题中的分治算法 | 60 |
| 4.5.1 | 最近点对问题 | 60 |
| 4.5.2 | 凸包问题 | 61 |
| 4.6 | 本章小结 | 62 |
| 4.7 | 习题 | 62 |
| 第5章 | 贪心算法 | 65 |
| 5.1 | 算法设计思想 | 66 |
| 5.1.1 | 贪心算法的设计思想 | 66 |
| 5.1.2 | 贪心算法的求解过程 | 66 |
| 5.2 | 图问题中的贪心算法 | 67 |
| 5.2.1 | 单源最短路径问题: Dijkstra 算法 | 67 |
| 5.2.2 | 最小生成树问题: Prim 算法和 Kruskal 算法 | 70 |
| 5.2.3 | 哈夫曼树 | 74 |
| 5.3 | 组合问题中的贪心算法 | 76 |
| 5.3.1 | 背包问题 | 76 |
| 5.3.2 | 活动安排问题 | 77 |
| 5.3.3 | 多机调度问题 | 78 |
| 5.4 | 本章小结 | 81 |
| 5.5 | 习题 | 81 |
| 第6章 | 动态规划算法 | 85 |
| 6.1 | 算法设计思想 | 86 |
| 6.1.1 | 动态规划算法的基本要素 | 86 |
| 6.1.2 | 动态规划算法的基本步骤 | 87 |
| 6.2 | 查找问题中的动态规划算法 | 90 |
| 6.2.1 | 最优二叉查找树 | 90 |
| 6.2.2 | 近似串匹配问题 | 92 |
| 6.3 | 图问题中的动态规划算法 | 94 |
| 6.3.1 | 多段图的最短路径问题 | 94 |
| 6.3.2 | 多源最短路径问题: Floyd 算法 | 95 |
| 6.4 | 组合问题中的动态规划算法 | 97 |
| 6.4.1 | 0/1 背包问题 | 97 |
| 6.4.2 | 最长公共子序列问题 | 98 |
| 6.5 | 本章小结 | 100 |
| 6.6 | 习题 | 100 |

| | |
|---------------------------|------------|
| 第 7 章 回溯算法 | 103 |
| 7.1 算法设计思想 | 104 |
| 7.1.1 问题的解空间与解空间树..... | 104 |
| 7.1.2 解空间树的动态搜索..... | 106 |
| 7.1.3 回溯算法的求解过程..... | 106 |
| 7.1.4 回溯算法的时间性能..... | 107 |
| 7.2 图问题中的回溯算法 | 108 |
| 7.2.1 深度优先搜索..... | 108 |
| 7.2.2 TSP 问题 | 110 |
| 7.3 组合问题中的回溯算法 | 113 |
| 7.3.1 0/1 背包问题 | 113 |
| 7.3.2 八皇后问题..... | 117 |
| 7.3.3 图着色问题..... | 119 |
| 7.4 本章小结 | 121 |
| 7.5 习题 | 121 |
| 第 8 章 分支限界算法 | 123 |
| 8.1 算法的设计思想 | 124 |
| 8.1.1 解空间树的动态搜索..... | 124 |
| 8.1.2 分支限界算法的设计思想..... | 126 |
| 8.1.3 分支限界算法的时间性能..... | 128 |
| 8.2 图问题中的分支限界算法 | 128 |
| 8.2.1 TSP 问题 | 128 |
| 8.2.2 单源最短路径问题..... | 130 |
| 8.3 组合优化问题中的分支限界算法 | 134 |
| 8.3.1 0/1 背包问题 | 134 |
| 8.3.2 任务分配问题..... | 136 |
| 8.3.3 活动安排问题..... | 139 |
| 8.4 本章小结 | 142 |
| 8.5 习题 | 142 |
| 第 9 章 概率算法 | 145 |
| 9.1 概率算法设计思想与实现基础 | 146 |
| 9.1.1 确定性与随机性..... | 146 |
| 9.1.2 各种概率算法的设计思想..... | 147 |
| 9.1.3 随机数和伪随机数..... | 147 |
| 9.2 数值概率算法 | 149 |
| 9.2.1 投点法计算 π 值 | 149 |

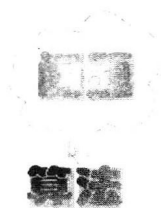
| | | |
|---------------|-----------------|------------|
| 9.2.2 | 拉普拉斯方程狄利克雷问题的求解 | 150 |
| 9.3 | 蒙特卡罗算法 | 151 |
| 9.3.1 | 蒙特卡罗算法正确率的提升 | 151 |
| 9.3.2 | 串相等性测试问题 | 153 |
| 9.3.3 | 素数性测试 | 154 |
| 9.4 | 拉斯维加斯算法 | 156 |
| 9.4.1 | 随机抽牌问题 | 156 |
| 9.4.2 | 整数因子分解 | 158 |
| 9.5 | 舍伍德算法 | 159 |
| 9.5.1 | 舍伍德型的快速排序 | 159 |
| 9.5.2 | 随机化的选择算法 | 160 |
| 9.6 | 本章小结 | 162 |
| 9.7 | 习题 | 162 |
| 第 10 章 | 计算智能 | 165 |
| 10.1 | 神经网络 | 166 |
| 10.1.1 | 思想来源和发展历程 | 166 |
| 10.1.2 | 神经网络的基本原理 | 167 |
| 10.1.3 | ANN 小结 | 170 |
| 10.2 | 模糊逻辑 | 170 |
| 10.2.1 | 模糊逻辑概述 | 170 |
| 10.2.2 | 模糊逻辑的基本原理 | 171 |
| 10.2.3 | 模糊逻辑技术小结 | 173 |
| 10.3 | 遗传算法 | 174 |
| 10.3.1 | 遗传算法的思想起源 | 174 |
| 10.3.2 | 遗传算法的基本原理 | 175 |
| 10.3.3 | 遗传算法的特点及其发展趋势 | 176 |
| 10.4 | 蚁群算法 | 177 |
| 10.4.1 | 蚁群算法的思想来源 | 177 |
| 10.4.2 | 蚁群优化的基本原理 | 178 |
| 10.4.3 | 蚁群优化小结 | 180 |
| 10.5 | 粒子群优化算法 | 181 |
| 10.5.1 | 粒子群优化算法的思想来源 | 181 |
| 10.5.2 | 粒子群优化算法的基本原理 | 181 |
| 10.5.3 | 粒子群优化算法的发展趋势 | 182 |
| 10.6 | 差分进化算法 | 183 |
| 10.6.1 | 差分进化概述 | 183 |
| 10.6.2 | 差分进化算法的基本原理 | 184 |

| | | |
|------------------|-------------|------------|
| 10.6.3 | 差分进化算法小结 | 185 |
| 10.7 | 分布估计算法 | 185 |
| 10.7.1 | 分布估计算法概述 | 185 |
| 10.7.2 | 分布估计算法的基本原理 | 187 |
| 10.7.3 | 分布估计算法的发展趋势 | 187 |
| 10.8 | 本章小结 | 188 |
| 10.9 | 习题 | 189 |
| 附录 A 名词索引 | | 191 |
| 索引 | | 196 |
| 参考文献 | | 199 |

第 1 章

绪 论

“计算机科学=算法的研究”

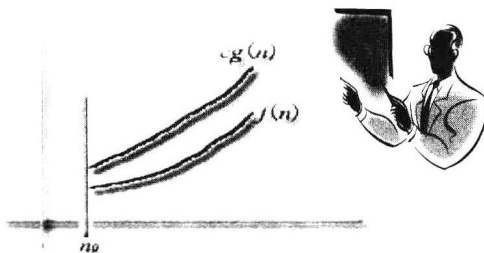
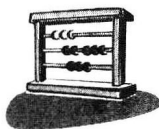


输入 计算机 输出

计算机科学的各个应用领域都高度依赖于算法设计。设计高效的算法来解决现实应用问题是计算机各领域的重要研究课题。



算法(Algorithm)是由若干条指令组成的有穷序列,它用于规定计算机完成人们交给它的任务所要执行的准确而完整的步骤。



算法(Algorithm)是由指令集组成的有穷序列,用于规定计算机完成人们交给它的任务所要执行的准确而完整的步骤。正如 Donald E. Knuth 所说,“计算机科学就是算法的研究”,因为计算机科学的各个应用领域都高度依赖于算法设计。设计高效的算法从而提高计算机解决现实应用问题的效率是当前计算机在各应用领域都不可回避的重要研究课题。

本章从算法的基本概念开始,逐步介绍算法设计的基本步骤以及复杂性分析的基本原理。具体包括如下内容:

- 算法的基本概念
- 算法设计与分析的重要问题类型
- 算法复杂性分析基础

1.1 算法的基本概念

1.1.1 算法的重要性

什么是算法?在不同的应用领域可能会有不同的答案。在计算机科学领域,算法指的是由若干条指令组成的有穷序列,这些指令序列满足以下 4 条性质:

- (1) 输入:有零个或多个输入。
- (2) 输出:算法至少产生一个输出。
- (3) 确定性:组成算法的指令序列是清晰的。
- (4) 有穷性:算法中每条指令的执行次数和时间是有限的。

由于计算机不会思考,我们要计算机完成一定的任务,就需要将外部的信息输入到计算机中,并将完成任务的具体步骤编码成指令让计算机机械地去执行。如此可见,算法是人和计算机之间沟通的桥梁,算法与计算机之间的关系如图 1.1 所示。

比如,需要计算机计算两个数 m 和 n 的最大公约数(不妨设 $m > n$)。可以利用欧几里得原理设计如下计算过程:

- (1) 如果 $n = 0$,则输出 m 作为最终结果,结束算法;
- (2) 求出 m 除以 n 的余数 r ;
- (3) 令 $m = n, n = r$,返回步骤(1)。

下面考察一下上述计算过程是否满足算法的 4 个条件:有两个输入 m 和 n ,因此第一个条件满足;显然上述 3 条指令的定义都非常明确,不存在歧义,因此第 3 个条件满足;由于每经过一次循环, n 的值都会不断变小,而且不会变成负数,因此上述计算过程在有限步内会停止,即满足第 4 个条件;计算过程在停止前总输出 m 作为最终结果,有一个输出,因此第 2 个条件也满足。如此说明上述计算过程是一个算法。

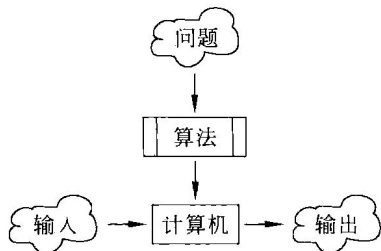


图 1.1 算法与计算机之间的关系

可以有多种不同的方案来完成同一个任务,而每一种方案都对应一个算法。因此,可以设计不同的算法来完成同一个任务。有的算法求解的效率较低,而有的算法则可能效率较高。当面对一个问题的时候,如何设计出高效的算法来让计算机完成任务是算法这门学科所要研究的主要内容。如此可见,算法对计算机的普及和应用具有至关重要的作用。

1.1.2 算法设计与分析的流程

在着手设计算法之前,通常需要先分析清楚问题,还要考虑实际的运行环境等因素,这样才能设计出可行且高效的算法。一般地说,设计算法的过程主要包含如下5个步骤:①理解待求解的问题;②确定算法运行的环境;③设计算法;④分析算法;⑤编程实现。各个步骤具体的内容如图1.2所示。

第一步:理解待求解的问题。

解决问题是设计算法的最终目标,因此分析清楚待求解的问题是首要的任务。在这一步,除了需要分析清楚问题的求解目标、输入数据和限制条件之外,还要判断清楚待求解问题的种类,是否有现成的算法可以直接应用。



应用于航天领域!太空环境下的硬件工作很特殊,算法的鲁棒性要求很高

小型嵌入式环境?实时性要求高不高?内存够不够用?



第二步:确定算法运行的环境。

当完成第一步骤之后,还需要了解算法的运行环境,这样才能设计出可行且高效的算法。比如,在一些小型的嵌入式环境中,通常只能运行需要较小内存的算法;而对于并行分布式的运行环境,则可以设计高效的并行算法。

第三步:设计算法。

一旦了解完算法的运行环境后,就可以开展算法的具体设计工作了。在这一步骤中,首先需要确定算法所需要的数据结构。充分利用各种数据结构的特性对算法的设计是非常重要的。在第2章中,将会具体介绍一些经典的数据结构。选择好数据结构后,就可以结合具体问题的特性来选择求解该问题的算法设计技术(比如贪心法、动态规划法和分支限界法等,一些经典的算法设计技术将在后面的章节中详细介绍)。然后,就可以根据算法设计技术的原理描述算法的具体流程。描述算法的流程可以采用流程图、伪代码甚至是自然语言。在本书中,统一采用伪代码来描述算法的流程。



第四步:分析算法。

经过上述3个步骤,即已经设计出了一个可行的算法。在这一步骤,将对所设计出的算法进行复杂性分析,考察其在时间和空间方面的计算开销。我们的目标是通过复杂度分析来提高算法的效率。若算法在某些环节的计算开销较大,可以有针对性地改进该环节;若整个算法的计算开销太大的时候,甚至需要返回第三步重新考虑采用新的算法设计技术来求解该问题。



第五步:编程实现。

至此,已经确定了最终的算法原理。这一步骤要做的是采用某种程序设计语言将设计好的算法实现出来,如C、C++或Java等。

图1.2 算法设计的流程

1.2 算法设计与分析的重要问题类型

计算机目前已经被用于越来越广泛的应用领域,其求解问题所涉及的算法也越来越复杂。但不论最终的算法如何复杂,它们通常都可以由一些求解基本问题的算法组合而成。下面介绍计算机在求解问题过程中经常遇到的一些基本问题。在后续章节中也将以这些基本的问题为例介绍具体的算法设计技术。

1.2.1 排序问题

排序(Sorting)指的是将给定数据集中的元素按照一定的标准来安排先后次序的过程。由于秩序是人们在日常生活中频繁遇到的问题,因此排序问题在算法设计与分析中占有非常重要的地位。计算机科学家 D. E. Knuth 在他的著作《排序与查找》中指出:“我确信,程序设计的每一个重要方面,实际上都离不开排序或查找”,“计算机运行时间的1/4以上是花在排序上,有许多计算机装置,排序竟用去一半以上的计算机时间”。由此可见排序在算法设计中的重要地位。计算机科学家对排序算法的研究也是经久不衰,目前已经提出了成百上千种排序算法,如插入排序、选择排序、归并排序和快速排序等。每个排序算法在时空开销以及其他方面各有特点。在实际应用中,通常需要结合具体的问题选取最合适的排序算法。

如图 1.3 所示为两个排序问题的例子。

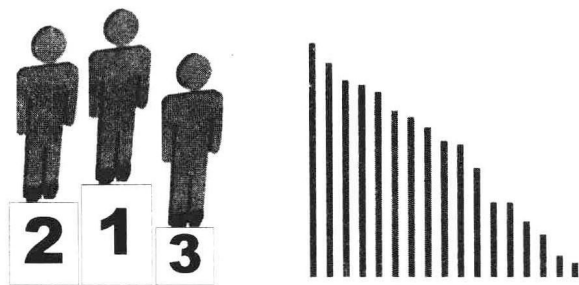


图 1.3 两个排序问题的例子

1.2.2 查找问题

查找(Searching)指的是按照指定的关键字,从给定的数据集中找出指定数据元素的过程(图 1.4)。查找在计算机应用领域中同样占有举足轻重的地位。在宏观方面,随着知识大爆炸时代的到来,从海量数据中查找有用信息的要求日益迫切,且已经成为制约社会发展的关键技术。目前快速发展的搜索引擎技术所要解决的主要问题就是信息的查找问题。在微观方面,数据的查找在算法设计中也是频繁出现的子问题,比如从数组中查找某个元素的值等。目前,计算机科学家也已提出了许多种查找算法,如顺序查找、二分查找和哈希查找等。每种查找算法都有自己的优缺点,需要结合具体的应用实际来选择最合适的算法。