

21世纪高等学校规划教材 | 计算机科学与技术

算法分析与设计 及案例教程

师智斌 靳雁霞 井超 梁志剑 雷海卫 编著



清华大学出版社

21世纪高等学校规划教材 | 计算机



算法分析与设计 及案例教程

师智斌 靳雁霞 井超 梁志剑 雷海卫 编著

清华大学出版社
北京

内 容 简 介

本书介绍了算法的概念,算法分析的基本理论、过程和方法以及算法设计的基本策略。主要内容包括算法概述、算法效率分析基础、蛮力法、分治法、分治策略变体——减治策略和变治策略、动态规划、时空权衡技术、贪心算法、回溯法和分支限界法、NP 完全性理论等。本书最后对 ACM 竞赛精选案例进行了分析和讲解,书中既有新创趣题,也有经典名题,难度适宜,使读者能够沿着一条清晰的、一致的、连贯的思路来探索算法设计与分析这一领域。

本书内容丰富,逻辑性强,既注重理论知识,又强调工程实用,既可以作为高等院校计算机及相关专业本科生、研究生“算法分析与设计”课程的教材,也可以作为广大计算机工程与应用行业的读者的参考书。此外,本书还可以作为参加 ACM 程序设计大赛的爱好者的参考书或培训教材。

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

算法分析与设计及案例教程/师智斌等编著. --北京:清华大学出版社,2014

21 世纪高等学校规划教材·计算机科学与技术

ISBN 978-7-302-38349-9

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

中国版本图书馆 CIP 数据核字(2014)第 244473 号

责任编辑:闫红梅 薛 阳

封面设计:傅瑞学

责任校对:焦丽丽

责任印制:何 芊

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

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

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 装 者:三河市少明印务有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:15.75 字 数:381 千字

版 次:2015 年 1 月第 1 版 印 次:2015 年 1 月第 1 次印刷

印 数:1~2000

定 价:29.00 元

出版说明

随着我国改革开放的进一步深化,高等教育也得到了快速发展,各地高校紧密结合地方经济建设发展需要,科学运用市场调节机制,加大了使用信息科学等现代科学技术提升、改造传统学科专业的投入力度,通过教育改革合理调整和配置了教育资源,优化了传统学科专业,积极为地方经济建设输送人才,为我国经济社会的快速、健康和可持续发展以及高等教育自身的改革发展做出了巨大贡献。但是,高等教育质量还需要进一步提高以适应经济社会发展的需要,不少高校的专业设置和结构不尽合理,教师队伍整体素质亟待提高,人才培养模式、教学内容和方法需要进一步转变,学生的实践能力和创新精神亟待加强。

教育部一直十分重视高等教育质量工作。2007年1月,教育部下发了《关于实施高等学校本科教学质量与教学改革工程的意见》,计划实施“高等学校本科教学质量与教学改革工程”(简称“质量工程”),通过专业结构调整、课程教材建设、实践教学改革、教学团队建设等多项内容,进一步深化高等学校教学改革,提高人才培养的能力和水平,更好地满足经济社会发展对高素质人才的需要。在贯彻和落实教育部“质量工程”的过程中,各地高校发挥师资力量强、办学经验丰富、教学资源充裕等优势,对其特色专业及特色课程(群)加以规划、整理和总结,更新教学内容、改革课程体系,建设了一大批内容新、体系新、方法新、手段新的特色课程。在此基础上,经教育部相关教学指导委员会专家的指导和建议,清华大学出版社在多个领域精选各高校的特色课程,分别规划出版系列教材,以配合“质量工程”的实施,满足各高校教学质量和教学改革的需要。

为了深入贯彻落实教育部《关于加强高等学校本科教学工作,提高教学质量的若干意见》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院、系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上。精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展,顺应并符合21世纪教学发展的规律,代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版

社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。推出的特色精品教材包括:

(1) 21 世纪高等学校规划教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。

(2) 21 世纪高等学校规划教材·计算机科学与技术——高等学校计算机相关专业的教材。

(3) 21 世纪高等学校规划教材·电子信息——高等学校电子信息相关专业的教材。

(4) 21 世纪高等学校规划教材·软件工程——高等学校软件工程相关专业的教材。

(5) 21 世纪高等学校规划教材·信息管理与信息系统。

(6) 21 世纪高等学校规划教材·财经管理与应用。

(7) 21 世纪高等学校规划教材·电子商务。

(8) 21 世纪高等学校规划教材·物联网。

清华大学出版社经过三十多年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会

联系人:魏江江

E-mail: weijj@tup.tsinghua.edu.cn

前言

根据教育部高等学校计算机科学与技术教学指导委员会对高等学校计算机科学与技术专业人才培养能力构成与培养的主题的阐述,计算机专业人才的专业基本能力包括计算思维能力、算法设计与分析能力、程序设计与实现能力、系统能力。算法是系统工作的基础,作为一名优秀的计算机专业人才,关键是建立算法的概念,具备算法设计与分析的能力。

本教材按照“算法基本知识—经典算法思想—算法应用实践”的顺序进行了内容的组织及编写。读者通过阅读算法基础部分,可了解算法的由来及其发展过程,理解算法的含义及问题分类,掌握算法的分析表示方法及算法效率的评价手段。面对日益复杂的问题,可将算法分为蛮力法、分治法及其变体算法、动态规划、时空权衡、贪心算法、回溯和分支限界法等几种。在经典算法思想部分,基本按照“算法思想—算法特点—算法实例—效率分析”的体例分别描述了各种算法,目的是使读者能够深入浅出地理解并掌握算法,能够分析并比较相同问题采用不同算法时的效率。为了提高读者的算法应用能力,本书结合 ACM 竞赛,从中选取了 12 个竞赛题目,例如果园篱笆问题、旅游预算问题等,并对各类问题进行了分析和讨论,加强了读者理论和实践相结合的意识。

全书共分为 11 章。

第 1 章介绍算法的概念、由来与发展,对基本问题类型、数据结构简要阐述。然后介绍算法求解的框架和步骤。

第 2 章介绍算法效率分析基础。介绍算法分析的框架、三种渐进符号和基本效率类型。然后介绍针对非递归算法和递归算法的数学分析方法。

第 3 章介绍蛮力法。它是解决问题的最直接的方法,基于问题的描述和所涉及的概念、定义直接求解。

第 4 章介绍分治法。分治法 is 问题求解采用的最常用的算法策略之一,非常重要。

第 5 章介绍分治策略的变体。介绍了分治法的两种变形:减治和变治策略,并通过实例介绍这两种策略在实际中的应用。

第 6 章介绍动态规划算法。以实例详述动态规划的算法思想、特点和求解问题的方法步骤。

第 7 章介绍时空权衡技术。介绍牺牲时间效率换取空间效率和牺牲空间效率换取时间效率的算法设计方法。

第 8 章介绍贪心算法。它也是非常重要的算法策略,且效率较高。介绍了几种典型的采用贪心算法求解最优问题的方法。

第 9 章介绍搜索算法。介绍回溯法和分支限界法,这两种算法适宜解决数据量较大且难解的问题。

第 10 章介绍 NP 完全性理论。简单介绍了 NP 完全性理论,以引起读者进一步学习和研究的兴趣。

第 11 章精心挑选了 12 道 ACM 竞赛题目,对各问题进行了分析和讲解,并在电子资源中提供了程序清单,以供读者学习和参考。

本教材可以作为计算机科学与技术、软件工程、网络工程等专业本科生及研究生的教材使用,同时也可作为有关专业软件开发人员的参考书。

本教材由师智斌等编著,其中师智斌编写了第 1 章,井超编写了第 2、5、6 章,王东编写了第 3 章,靳雁霞编写了第 4 章,梁志剑编写了第 7 章,雷海卫编写了第 8~10 章,秦品乐编写了第 11 章。本教材还参阅了大量国内外专家、学者发表的著作、论文,在此向这些同行们表示衷心的感谢!

由于编者水平有限,书稿虽数次修改,仍会有不妥甚至错误之处,诚盼专家和广大读者不吝指正。联系方式:电子邮件 1637350520@qq.com。

编 者

2014 年 6 月

目 录

第 1 章 绪论	1
1.1 什么是算法	1
1.1.1 算法的由来	1
1.1.2 算法的发展	3
1.1.3 算法的例子	4
1.2 重要的问题类型	5
1.2.1 排序	6
1.2.2 查找	6
1.2.3 字符串匹配	6
1.2.4 图问题	7
1.2.5 组合问题	7
1.2.6 几何问题	7
1.2.7 数值问题	8
1.3 基本数据结构	8
1.3.1 线性结构	8
1.3.2 树结构	8
1.3.3 图结构	9
1.3.4 集合	9
1.3.5 数据的物理结构	10
1.4 算法问题求解基础	10
1.4.1 算法求解框架	10
1.4.2 算法设计步骤	10
1.5 算法的表示	14
1.6 为什么学习算法	15
总结	16
习题 1	16
第 2 章 算法效率分析基础	17
2.1 算法分析框架	17
2.1.1 算法分析概述	17
2.1.2 算法正确性分析	17
2.1.3 时空效率分析	17

2.1.4	算法分析过程	19
2.2	渐进符号和基本效率类型	20
2.2.1	三种渐进符号	20
2.2.2	渐进符号的特性	20
2.2.3	基本效率类型	21
2.3	非递归算法的数学分析方法	22
2.4	递归算法的数学分析	26
2.4.1	递归算法的数学分析方法	26
2.4.2	斐波那契数列	29
2.5	算法的其他分析方法	31
	总结	32
	习题 2	32
第 3 章	蛮力法	33
3.1	概述	33
3.2	排序问题	35
3.2.1	选择排序	35
3.2.2	冒泡排序	36
3.3	查找问题	37
3.3.1	顺序查找	38
3.3.2	字符串匹配	38
3.4	几何问题	39
3.4.1	最近对问题	40
3.4.2	凸包问题	40
3.5	组合问题	42
3.5.1	旅行商问题	42
3.5.2	背包问题	44
	总结	45
	习题 3	45
第 4 章	分治法	46
4.1	概述	46
4.2	分治法的基本策略及步骤	46
4.2.1	分治法的基本策略	46
4.2.2	分治法的基本步骤	47
4.3	排序问题	48
4.3.1	合并排序	48
4.3.2	快速排序	49
4.4	查找问题	51

4.4.1 折半查找	51
4.4.2 二叉树遍历及其相关特性	53
4.5 数值计算问题	54
4.5.1 大整数乘法	54
4.5.2 Strassen 矩阵乘法	56
4.6 几何问题	58
4.6.1 用分治法解最近对问题	58
4.6.2 用分治法解凸包问题	59
4.7 分析分治法在安排循环赛中的应用	60
总结	62
习题 4	62
第 5 章 分治策略变体——减治策略和变治策略	64
5.1 减治策略	64
5.1.1 插入排序	65
5.1.2 拓扑排序	67
5.1.3 生成组合对象的算法	68
5.1.4 减常因子算法	71
5.1.5 减可变规模算法	73
5.2 变治策略	75
5.2.1 排序问题	75
5.2.2 平衡查找树	80
5.2.3 霍纳法则和二进制幂	86
5.2.4 问题化简	88
总结	89
习题 5	89
第 6 章 动态规划	91
6.1 概述	91
6.2 算法特点	93
6.2.1 备忘录方法	94
6.2.2 最优化原理	98
6.2.3 求解步骤	98
6.3 矩阵连乘问题	99
6.4 最长公共子序列	104
6.5 0-1 背包问题	107
6.6 最大子段和	113
6.7 最优二叉查找树	116
总结	122

习题 6	122
第 7 章 时空权衡技术	125
7.1 时空权衡策略	125
7.2 计数排序	126
7.3 字符串匹配	128
7.4 散列法	134
总结	139
习题 7	139
第 8 章 贪心算法	141
8.1 概述	141
8.1.1 贪心算法的基本要素	142
8.1.2 贪心算法的求解过程	142
8.2 活动安排问题	143
8.3 背包问题	146
8.4 最小生成树问题	148
8.4.1 Prim 算法	150
8.4.2 Kruskal 算法	153
8.5 单源(点)最短路径问题	158
8.6 哈夫曼编码	162
总结	168
习题 8	168
第 9 章 回溯法和分支限界法	171
9.1 回溯法	171
9.1.1 概述	171
9.1.2 子集和问题	176
9.1.3 n 皇后问题	178
9.1.4 哈密顿回路	180
9.1.5 装载问题	182
9.2 分支限界法	186
9.2.1 概述	186
9.2.2 0-1 背包问题	191
9.2.3 任务分配问题	194
9.2.4 多段图的最短路径问题	199
9.2.5 旅行商问题	204
总结	210
习题 9	210

第 10 章 NP 完全性理论	212
10.1 判定问题和最优化问题	212
10.2 P 类问题	214
10.3 NP 类问题	215
10.4 NP 完全问题	216
10.5 典型的 NP 完全问题	218
10.6 其他 NP 完全问题	220
10.7 NP 完全问题的计算机处理	221
总结	222
习题 10	222
第 11 章 案例精选	223
11.1 果园篱笆问题	223
11.2 空中飞行管理问题	223
11.3 去数问题	224
11.4 极差问题	224
11.5 最优合并问题	225
11.6 在棋盘上实现从初始布局到目标布局的转变	226
11.7 商店购物问题	227
11.8 旅游预算问题	229
11.9 防卫导弹问题	231
11.10 钓鱼问题	232
11.11 胖男孩问题	234
11.12 护卫队问题	235
参考文献	237

第1章

绪论

1.1 什么是算法

1.1.1 算法的由来

算法,简言之就是解决问题的方法。人们解决问题的过程一般由若干步骤组成,通常把解决问题的确定方法和有限步骤称为算法。如果相关问题的解决最终由计算机来实现,又由于计算机不具备思考能力以及人的“跳跃性思维”等因素,因此方法的确定和对步骤的描述尤为重要。

从中国古代的算筹(如图 1.1 所示)、算盘(如图 1.2 所示),到外国人发明的计算尺(如图 1.3 所示),最终到电子计算机,在这个复杂而漫长的过程中,人们想通过发明机器把人从繁杂的重复工作中解脱出来,这种时而简单时而复杂的重复工作就是计算。

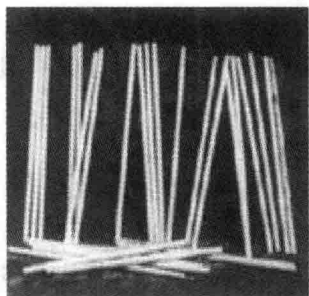


图 1.1 算筹

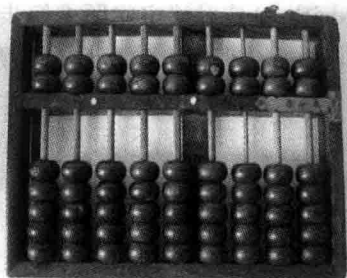


图 1.2 算盘

目前人们普遍认为第一台真正意义上的数字电子计算机是在美国宾夕法尼亚大学诞生的 ENIAC (Electronic Numerical Integrator And Computer, 电子数字积分计算机, 中文译名为埃尼阿克), 如图 1.4 所示。它主要用于第二次世界大战期间的导弹弹道科学计算, 但第二次世界大战期间并未研制成功, 直到 1954 年才成功服役, 计算机业由此迅速发展。

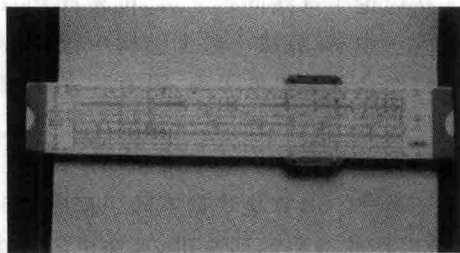


图 1.3 计算尺

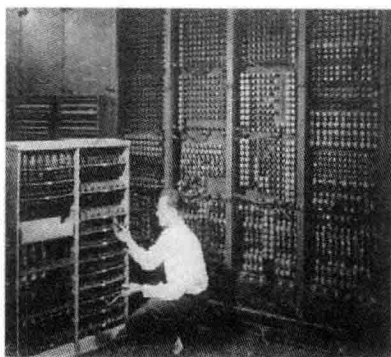


图 1.4 ENIAC

在大多数人的思想中,认为计算就是数值计算。数值计算在工业领域有着广泛的应用,人们发明了诸多工具来帮助解决一些繁琐的计算问题。而随着计算机的普及,人们更多地希望利用计算机解决生活中遇到的问题,这些问题从范围到复杂度都不是数值计算所能覆盖的,因此,非数值计算应运而生。

常见的非数值计算问题有排序、查找、字符串处理、排列组合、图问题、几何问题等,人们同样希望能利用计算机帮助解决和处理这些问题。因此与这些非数值计算问题相关的计算机科学越来越得到人们的重视,这其中就包括大家已经了解的数据结构等内容。

目前现代计算机仍基于图灵机模型,即采用冯·诺依曼体系结构来进行运算。从计算机的工作机制上看,计算机利用二进制数表示数据,因此所有存储在计算机内部的数据都是数值。这表示如果人们想使用计算机解决非数值问题,就需要用数值来表示非数值的信息;同时,人们还必须对问题进行分析和处理,找到解决问题的方法,并将求解方法和步骤转换成计算机能够理解并执行的程序;最终,计算机把需要执行的程序放在内存中运行,由运算器、控制器、存储器、输入设备和输出设备共同完成相应的工作。因此,人们在使用计算机解决问题时必须遵循这些要求。这个过程就是算法的形成过程。

由于计算机学科的发展时间还比较短,对于算法的概念,计算机领域还没有统一的定义,但人们达成了基本共识。本书给出了一种简单的定义:

算法(Algorithm)是对特定问题求解步骤的一种描述,是指令的有限序列。

算法具有以下 5 大特征:输入、输出、确定性、有穷性、可行性。对于某一具体问题的算法,需要把握以下 5 个方面。

- (1) 算法中每一个步骤都不能有歧义,这是对算法确定性的要求。
- (2) 必须确定算法所处理问题的输入定义域,这是对算法输入的要求。
- (3) 同一算法可用多种不同的形式来描述,这是对算法的描述方式不唯一。
- (4) 同一问题,可能存在多种不同的求解算法,这是针对同一问题算法有多多样性。
- (5) 针对同一问题的多种算法可能会基于完全不同的思路,而且解题速度也会有显著不同,这是多样的算法会有多样的效率。

在讨论了这些问题之后,再来看 Pascal 之父、结构化程序设计的先驱 Niklaus Wirth 最著名的一本书《算法+数据结构=程序》,算法与数据结构在程序设计中的重要性不言自明。

程序是计算机的灵魂,没有程序,计算机就不能完成人们交给它的任务,它的优势也就显示不出来了。同样,算法是程序的灵魂,算法是解决问题时的方法和步骤。因此,程序、算法都需要遵循计算机的要求和约定,所以上述的特性要求不仅是针对算法的,在编写程序时也需要遵守。但算法毕竟不同于程序,算法是让程序设计者读的,当然也需要易于转化成程序,所以用句通俗的话来讲,算法更应该接近于人,而程序更接近于计算机,算法是将要转化为程序的人类语言。由此,描述算法可以用自然语言、程序设计语言,而更多的是介于二者之间的伪代码。

1.1.2 算法的发展

虽然计算机的发展历史并不久,但是人类对于算法的研究却可以追溯较长的一段时间。

“算法”这一术语是从 Algorithm 翻译而来的,但直至 1957 年,西方著名的 Webster's New World Dictionary(《韦伯斯特新世界词典》)仍未将其收录其中。据西方数学史家的考证,古代阿拉伯的一位学者写了一部著作——*Hisab al-jabr Wa'l muqabalah*(《复原和化简的规则》),作者的署名是 Abū Abd Allāh Muhammad ibn Mūsā al-Khwārizmī,从字面上看,其含义是“穆罕默德(Muhammad)的父亲,摩西(Moses)的儿子, Khwārizmī 地方的人”。后来,这部著作流传到了西方,结果,从作者署名中的 al-Khwārizmī 派生出 Algorithm(算法)一词,而从作品名字中的 al-jabr 派生出 Algebra(代数)一词。随着时间的推移,现今 Algorithm(算法)这个词的含义,已经与它原来的含义完全不同。

在算法发展的过程中,从有人类文明的记录起,就有关于算法的记录。人类约公元前 4000 年,聪明的苏美尔人发明了人类最早的文字——楔形文字,以及“一周七天”、“一年十二个月”等历法。在公元前的人类历史中,就有关于数制、计算、密码和加密、矩阵等问题的记载,说明当时人类对于计算和算法已经开始涉及。随着时间的推移,历史记载了人类对于算术、运算、数制等问题的进一步的研究。

下面让我们回顾一些在算法发展中重要的事件和人物吧。

- 公元 1202 年,斐波那契的传世之作《算术之法》出版。在这部名著中,斐波那契提出了以下饶有趣味的问题:假定一对刚出生的小兔一个月后就能长成大兔,再过一个月便能生下一对小兔,并且此后每个月都生一对小兔,一年内没有发生死亡。问一对刚出生的兔子,一年内能繁殖多少对兔子?
- 1666 年莱布尼茨所著的《论组合术》一书问世,这是组合数学的第一部专著。在计算机问世和普遍应用之后,组合数学获得了蓬勃的发展,在对算法的运行时间和存储的需求估计以及社会科学、生物学、信息论等领域,组合数学都得到了广泛的应用。
- 1700 年前后,伟大的德国科学家莱布尼茨提出了二进制算法,这可以说是为现代计算机奠定了算法基础。同时,通过对中国古老“易经”的研究,莱布尼茨也在中国的传统文化中印证了二进制的思想。
- 1735 年,欧拉综合了莱布尼茨的微分与牛顿的流数,解决了长期悬而未决的贝塞尔问题,也是在 1735 年,他定义了微分方程中有用的欧拉-马歇罗尼常数。这些研究为在计算机领域中广泛使用的 RSA 公钥加密算法奠定了坚实的理论基础。
- 1815 年,软件行业的奇才 Ada Lovelace(艾达·洛夫莱斯 1815—1852)为巴贝奇析机拟定了“算法”,然后写作了一份“程序设计流程图”。这份珍贵的规划,被人们视为“第一件计算机程序”,她也被认为是世界上第一位软件工程师。
- 1847 年,英国数学家 George Boole 制定完成的一套逻辑数学计算方法,用来表示两个数值相结合的所有结果。后来人们以他的名字命名这套算法,称为 Boolean(布尔运算)。
- 1934 年,哥德尔在普林斯顿高等研究所报告中对递归函数进行了定义,这是历史上第一次给出算法的数学定义。

- 1965 年,库利和图基在《计算数学》杂志上发表了《机器计算傅里叶级数的一种算法》,该文最早提出了 FFT(快速傅里叶变换)算法,使得数字信号处理能够应用于实时场合,推动了数字信号处理地发展和应用。
- 复杂度的概念首先由 Kolmogorov 于 1965 年提出来,后由 Lempel 和 Ziv、Kasper 和 Schuster 给出了实现这种复杂度的具体算法。

在算法发展的过程中,具有里程碑性质的事件远远不止以上罗列的内容,尤其是进入 20 世纪以后,算法更加受到大家的重视,各种成就也就接踵而至。如今,人们在不同的领域继续研究并改进着各种算法,使之能为人类提供更好的服务。

1.1.3 算法的例子

下面举例说明什么是算法,供读者体会。

例 1.1 计算两个正整数 m 和 n 的最大公约数 $\text{GCD}(m, n)$ 。

【方法 1: 质因数分解法】

这种方法是我们在学习初等代数进行手工计算时常用到的方法。

比如,对于 72 和 48 这两个数,可以通过分解质因数得到:

$$72 = 2 \times 2 \times 2 \times 3 \times 3$$

$$48 = 2 \times 2 \times 2 \times 2 \times 3$$

$$\text{GCD}(72, 48) = 2 \times 2 \times 2 \times 3 = 24$$

由此可以概括出质因数分解法求解最大公约数的步骤。

- (1) 找出 m 的所有质因数。
- (2) 找出 n 的所有质因数。
- (3) 寻找 m 和 n 的所有公因数。需要注意的是, m 和 n 的质因数中同一个质因数可能多次出现(如 72 和 48 的质因数 2 分别有 3 次和 4 次),公因数应该包含相同质因数的较低次方。
- (4) 公因数相乘则为 $\text{GCD}(m, n)$ 。

对该方法进行分析我们发现,如果采用计算机实现该过程,有以下步骤是无法完成的。

在第(1)和第(2)步骤,对正整数进行质因数分解。如果要让计算机实现质因数分解,则需要生成一张质因数表,如表 1.1 所示。而对这张表的生成步骤没有给出相应的描述。另外,在该表的生成过程中又需要判定一个正整数是否为质数,该实现步骤也没有给出。

表 1.1 72 和 48 的质因数表

正整数	质因数一	质因数二	质因数三	质因数四	质因数五
72	2	2	2	3	3
48	2	2	2	2	3

其次,步骤(3)寻找 m 和 n 的所有公因数。具体怎样求解相同质因数的较低次方的方法和步骤无法得知,计算机无法自动实现。

综上所述,以这种形式表述求解最大公约数的过程,从严格意义上讲还不能称之为算法。真正的算法描述必须给计算机提供一个具体的、可操作的、能实现的方法和步骤。另外,从执行效率方面考虑,上述算法也不是一个高效的方法,对正整数进行质因数分解的过

程比较复杂,求解相同质因数也需要一定的时间,且需要相应的空间进行存储。请读者试着实现这个算法。

【方法 2: 循环测试法】

对求解最大公约数问题分析可得,正整数 m 和 n 的最大公约数 $\text{GCD}(m, n)$ 的取值范围在 1 到 $\min\{m, n\}$ 之间,并且能同时整除 m 和 n 。因此,如果我们从 $\min\{m, n\}$ 开始递减进行测试,第一次能同时整除 m 和 n 的数就是要找的最大公约数。

由此概括出采用循环测试法进行最大公约数求解的步骤如下。

- (1) 将 $\min\{m, n\}$ 的值赋给 a 。
- (2) 进行 m 对 a 的取余运算,如果余数为 0,进入第(3)步;否则进入第(4)步。
- (3) 进行 n 对 a 的取余运算,如果余数为 0,返回 a 的值作为结果;否则进入第(4)步。
- (4) a 值减 1,返回到第(2)步。

上述步骤描述可以很方便地转换成计算机可执行的操作。

以 72 和 48 这两个数为例,循环测试法求最大公约数的过程如下:将 $\min\{72, 48\}$ 的值赋给 a ,即 $a=48$,分别将 72 和 48 对 a 取余数,余数不都为 0; a 值减 1,得 $a=47$,再对 a 取余数,余数都不为 0; a 值继续减 1……直到 $a=24$,对 a 取余数,余数都为 0,找到最大公约数为 24,过程结束。

【方法 3: 欧几里得法】

欧几里得法是采用辗转相除法求两个正整数 m 和 n 的最大公约数。步骤如下。

- (1) 求 m 除以 n 的余数 r 。
- (2) 若 r 等于 0,则 n 为最大公约数,算法结束;否则执行第(3)步。
- (3) 将 n 的值放在 m 中,将 r 的值放在 n 中。
- (4) 转到第(1)步。

上述算法可以证明:设 $t=\text{GCD}(m, n)$,则 $m = a * t, n = b * t$,这里的 a 和 b 均为正整数,并且 a 和 b 互质。求 n 和 $m \% n$ 的最大公约数,则 $n = b * t, m \% n = (a - x * b) * t$,这里的 x 为正整数,如果 b 和 $a - x * b$ 不互质,我们很容易得到 a 和 b 不互质,从而可以证明 t 就是最大公约数。

采用欧几里得法求 72 和 48 的最大公约数:先将 72 除以 48 得余数为 24,再以 48 为 m ,24 为 n ,48 除以 24 得余数为 0,则 $n=24$ 即为所求的最大公约数。只需要 2 次迭代,就可以得到结果。

比较一下上面的三种算法,从算法的复杂度、运行效率等方面考虑,质因数分解法是最复杂的,循环测试法次之,而欧几里得法是最简单的。

本节通过一个例子说明了算法的定义和特性,需要大家注意的是,虽然这里举的例子是以数学中的内容为原型的,但是当今很多算法都不仅仅涉及数学问题,而与生活中可以接触到的事情有广泛的联系,算法是无处不在的。

1.2 重要的问题类型

在继续讨论其他内容前,先回顾一下几种重要的问题类型,包括排序、查找、字符串匹配、图问题、组合问题、几何问题、数值问题。这些问题一方面具有重要的使用价值,另一方