新世纪计算机专业系列教材

计算机算法引论

——设计与分析技术

刘 璟 编著

1.6

科学出版社
www.sciencep.com

新世纪计算机专业系列教材

计算机算法引论

——设计与分析技术

刘　璟　编著

科学出版社

北　京

内 容 简 介

本书是一本面向计算机、软件工程和网络工程专业及相关专业的本科生(高年级)和研究生教材,根据国内外计算机技术的最新发展,讲述计算机算法的各种设计策略,包括分治技术、贪心技术、动态规划技术、回溯和分支限界技术等;介绍算法分析技术、算法的时间和空间复杂度分析方法,包括最坏情况和平均情况的分析等;讨论各类经典和应用问题的算法,包括排序算法、搜索算法、字符串匹配算法、图论算法、调度算法、组合优化算法、数论算法等。并在计算复杂性理论的基础上,引入近似算法、概率算法等最新内容。

图书在版编目(CIP)数据

计算机算法引论:设计与分析技术/刘璟编著. —北京:科学出版社,2003
(新世纪计算机专业系列教材)
ISBN 7-03-011741-7

Ⅰ.计… Ⅱ.刘… Ⅲ.电子计算机—算法理论—高等学校—教材
Ⅳ.TP301.6

中国版本图书馆 CIP 数据核字(2003)第 054673 号

策划编辑:陈晓萍/责任编辑:丁 波
责任印制:吕春珉/封面设计:王 浩

科学出版社 出版
北京东黄城根北街16号
邮政编码 100717
http://www.sciencep.com
双青印刷厂印刷
科学出版社发行 各地新华书店经销

2003年9月第 一 版 开本:787×1092 1/16
2003年9月第一次印刷 印张:17 3/4
印数:1—5 000 字数:397 000

定价:24.00元

(如有印装质量问题,我社负责调换〈双青〉)

新世纪计算机专业教材
编委会

顾问编委

施伯乐教授　复旦大学
白英彩教授　上海交通大学

主　任

左孝凌教授　上海交通大学

编　委

刘　璟教授　南开大学
宋方敏教授　南京大学
何炎祥教授　武汉大学
余雪丽教授　太原理工大学
阮家栋教授　上海工程技术大学
顾训穰教授　上海大学
徐汀荣教授　苏州大学
曾　明教授　西安交通大学
曹元大教授　北京理工大学
曹文君教授　复旦大学
陶树平教授　同济大学
缪淮扣教授　上海大学
谢康林教授　上海交通大学

总　序

20年来，计算机学科的发展日新月异，促使现代科学在各个领域突飞猛进。目前，计算机科学技术已应用在实时控制、信息处理、通信传输、企事业管理等领域，成为人们工作、学习、生活必不可少的工具。计算机技术的发展瞬息万变，具有以下三方面特点：

（一）传统的工、理、文、医、商、农在计算机的应用方面都有着各自专业的需要，例如，经济、艺术、法律、管理、医学等各种学科都需要依赖于计算机技术的应用。除了各自领域的专业实践外，应用计算机已是各个专业提高效率、发挥潜能、促进发展的必不可少的手段。因此现在很难用传统的工、理、文、医、商、农等去界定学科的分类。

（二）计算机网络改变了计算机通信的时空距离。计算机应用的发展是与计算机网络的发展紧密相连的。从最初的局域网（LAN）到广域网（WAN），以至用一种新的方法将LAN和WAN互联起来，即成为网际网（Internetwork）。这种网际网的实验原型Internetwork，通常缩写为Internet。计算机网格将计算机互连起来，从而使计算机之间可以交换信息，而且这种信息交换可以在几分钟内就影响到世界各地。计算机网络的发展，带动了计算机学科在很多领域的拓展。

（三）现代计算机学科向综合性发展。计算技术发展伊始，每种学科均以软硬件分类，泾渭分明。但自网络发展以来，Internet软件中的两部分变得特别重要和特别具有开创性，即网际协议（Internet Protocol，简称IP）和传输控制协议（Transmission Control Protocol，简称TCP）。这些协议是必不可少的软件系统。但是在网络系统中，网络的互连必须依靠路由器、服务器、接口插座、调制解调器等硬件设施，所以计算机网络很难归结为软件或硬件的单一体系。

随着计算机技术的发展，计算机与通讯、视频、声音等密不可分；随着多媒体的发展和应用，计算机科学已经愈来愈成为与数字传输、视频、声、光、电等综合的学科。

尽管计算机技术的发展如此神速、新异，但像一切新学科的发展一样，计算机教育水平仍滞后于计算机技术的发展。为了适应计算机教学改革的需要，我们国内部分重点院校的教授、学者，在科学出版社的积极鼓励和支持下，成立了新世纪计算机专业教材编委会。自2000年10月以来，我们群策群力，多次探讨了当前教育与技术进展之间的差距，并且仔细研讨了美国ACM/IEEE-CS公布的*Computing Curricula* 2001的优点与不足，结合我国计算机教育的实际情况，提出了编著一套适用于计算机本科专业的励精图治的教材计划。这套教材的选题、定位乃至作者的遴选，都得到了国内很多著名教授和学者的认同，并且有很多选题都争取到了一些著名教授亲自参与编写。这套教材立意着重基础，反映导向，注重实践。

因此我们在基础课目方面，首先列选了数据库原理、操作系统、编译程序原理、智能基础等基础教程。这些基础课教材都由一些国内著名学者执笔，论述内容既注意打好扎实

试读结束：需要全本请在线购买：　www.ertongbook.com

基础，又注意要反映最新导向，高屋建瓴，使读者迅速接近最新领域。

同时，为了反映导向，我们抓住网络课程作为计算机专业学生的应用基础，编写了一本实用性极强的《计算网络教程》。这本教材的编著思想是以基础—理论—应用为主线，通信是基础，协议是核心，互连是重点，应用是目标。

其次，为了拓展学生的网络应用本领，我们还安排了电子商务、多媒体应用以及 Web 数据库技术三门应用课程。电子商务和多媒体应用是计算机应用中最为热门的课程，也是拓展性极广的计算机应用领域，应用前景极为广阔。

Web 数据库技术是一种随着互联网技术发展起来的应用技术。它涉及网络、HTTP 协议、Script 语言、动态网页开发平台、远程数据访问技术等各种网络应用技术。目前国内外还无适合教材，因此，编写 Web 数据库技术的教材，可以说是填补了应用领域的一个空白。

在研究美国公布的“计算 2001-CS 教程”中，我们仔细探讨了数据结构这一课程的变化。在“计算 1991 教程”中，数据结构内容明确放在算法与数据结构之中，而“2001-CS 教程”却无数据结构的课程名称，代之以程序设计基础（Programming Fundamentals）。文件中提到了基本数据结构和抽象数据类型以及面向对象的程序设计等内容。从这里可以看出，数据结构是以程序设计基础作为研究对象的。另外该教程把算法与复杂性作为一个单独课程列出，这一方面说明算法是一种问题求解的策略，另一方面也说明基本算法及复杂性的讨论对于程序设计是多么重要。

为此在这套丛书中我们安排了一个软件课程系列，即开设从语言、数据结构、算法到软件工程的课程。首先我们从面向对象的 C＋＋语言入手，进一步讲解语言学概论。主要内容是分析语法结构，掌握语言构成规律，读懂语言文本。任何计算机语言均可触类旁通，这种从结构规律来学会应用的方法，就是以不变应万变，因为从根本上说，尽管计算机语言千变万化，但万变不离其宗。在搞通语言基础上，我们组编了数据结构，或者说是研究程序设计基础。然后是学习基本算法，也就是为了程序设计需要，而进行问题求解，即进行常用算法讨论。为了使开发软件遵循工程管理方法，软件工程的学习将是计算机专业学生规范软件开发的必不可少的训练课程。

我们筹组这套丛书时，希望每本教材都有创意，能引起共鸣，能被关注，能被采纳，能被推广。但是我们也注意到，由于各个学校情况不同，各人观点不同，理解角度也有所不同，所以对教材的选用和编著，不易一致认同。不过我们希望这套教材能够反映当前学校动向，在促进学以致用等方面有所促进、有所推动，更希望兄弟院校的教师、学者能够积极使用，参与讨论，以使本套丛书能够不断修改，日臻完善。

最后我要感谢科学出版社的领导对本套丛书的列选、报审、出版所给予的鼓励和支持。

左孝凌

2001 年 7 月 30 日

前　言

“计算机算法设计与分析”作为计算机学科的核心课程，源自 20 世纪 70 年代 D. Knuth 的巨著《计算机程序设计的艺术》(第一卷“基本算法”，第二卷“半数值算法”，第三卷“排序与搜索”)和 A. Aho、J. Hopcroft 与 J. Ullman 合著的《计算机算法的设计与分析》的发表，从那时候开始，“数据结构”、“算法分析”逐渐成为计算机专业本科和研究生的主要课程，有人统计，这两种书在其后发表的论文和专著中被引用次数最多，而 D. Knuth 还因此获得了 1974 年的图灵(Turing)奖，自 1965 年至今，在近四十届图灵奖的颁奖历史中，这是惟一的一次因写了一本影响巨大的书而获奖。现在，算法理论在计算学科领域中的核心地位已逐渐为人们所公认，三十年来，计算机科学技术的几乎每一项新的成就都与算法研究密切相关。

最近几年，我国计算机专业教育得到了前所未有的发展，除了在规模上扩大了几倍之外，在教学内容和课程设置方面，亦有诸多改进，现在人们对于借鉴国外经验尤为注重，例如，IEEE-Computer & ACM 发表的《计算教程 2001》(CC2001)对我国的计算机专业教育产生了很大的影响，它把“算法及其复杂性”列为计算学科的核心课程，这使得算法设计与分析技术的讲授更受重视。一方面在研究生培养计划中，算法及复杂性的知识成为主要必修内容，由于研究生的多数研究工作与算法相关，高水平的算法理论基础成了高质量计算机人才培养的必要条件。另一方面，有关算法的理论与技术的内容正在向本科教学中扩展，这种趋势在近几年越来越明显，过去在大多数计算机本科专业中，有关算法设计的初步知识是在“数据结构”课程中介绍的，目前，这种情况正在发生变化，一种做法是把传统的“数据结构”课程中算法及复杂性的内容充实加强，为了适应这一变化，一些“数据结构与算法”之类的教材已经在国内外出现；许多计算机科学技术专业则采取另一方案，即在“数据结构”课程之外，再开设一门本科生的算法课程，这是一种趋势。除此之外，许多新建的“软件工程”(包括示范性软件学院)、“网络工程”等专业的教学计划中，也多把算法课程放在重要地位。

为适应计算机教育的这种形势，本书的编写遵从下列宗旨。

1. 深度

算法的设计与分析技术包含着深入的理论内容，是为培养中高级计算机人才所必需的知识基础，本书不是一种算法查找手册，而是一本要深入浅出地讲授有效算法的设计与分析技术，引导读者了解国内外算法理论和应用的主流和最新发展，并通过对于大量的著名问题的优秀算法实例的讲解，使读者经过消化、理解，获得设计、分析和创新能力的书籍。

2. 结构

全书的安排(与目前国内外最好的教材一致)不拘泥于按问题或按设计技术分章分节，这有利于读者由浅入深逐步深入地学习全面的算法知识。第 1 章为绪论，介绍算法理

论的基本概念、基本思考方法、算法的评估标准，以及算法研究的目标和与解决实际问题的关系。第2章以读者比较熟悉的排序问题为例，全面介绍算法的分析技术，并以此为基础，指导读者在后面各章节中用分析的方法评价算法的优劣，从分析结果中得到设计更优算法的启示。第3、5、7、8章，介绍传统的算法设计技术，通过对于大量重要问题算法(包括图论问题、调度问题、选择问题和一些组合问题的算法)的设计，使读者掌握有效算法的主要设计策略。第2、4、6、11章则重点介绍排序、数据检索、字符串匹配和与数据加密相关的数论问题的算法设计问题。在针对具体问题讨论算法设计的同时，又逐步加深整个算法理论的学习，例如，讲排序问题重点结合算法分析技术，讨论检索问题时注意揭示算法与数据结构的密切相关性，数论问题算法的学习则引入了一种称为概率算法(随机算法)的全新技术。第9章概述NP－完全性理论，用较小的篇幅对计算复杂性理论的基本思想给出了一个清楚明晰的严谨描述。在第10和11章中对于近似算法和概率算法的讨论，为读者解决"计算机难解问题"开辟了新的思路和新的途径。

3. 层次

为了适应不同的教学需求(如不同的课时数、不同的读者程度等)，也为了引导有兴趣的读者深入学习，我们把全部内容分为下面几个层次。

(1)普通章节为必选的基本内容。

(2)标"*"的章节为可选的内容，教师可根据情况选择一部分讲授。

(3)标"**"的章节为比较深入的内容，可在研究生课程中介绍。

(4)附录中的内容，可作为有兴趣的读者参考。

例如，我们把第8章"回溯与分技限界技术"、第10章"近似算法"，第11章"数论算法及其应用"等都列为可选章节，是因为在有些算法教材中也可以不包括这些内容，主讲者可以根据情况有所取舍。有些标"**"的内容，例如"Hash技术的新发展"、"随机二叉搜索树"等内容更适合研究生学习。在第9章，我们采用了一种新的方法引入"不确定算法"与NP类的概念，这有利于避免过多的理论叙述，而把图灵机(TM)和Cook定理的证明放在附录中，以便读者深入学习。

我能够把这本书奉献给读者，仰赖于多方面的帮助，特别应感谢四川大学李万学教授、上海交通大学左孝陵教授、复旦大学朱洪教授的鼓励和指点，感谢研究生王颖、江标为本书编选了各章的习题并完成了全书的图文编辑工作，感谢王刚博士阅读了本书并提出了有益的修改意见，感谢科学出版社陈晓萍编辑为本书的出版所付出的努力。

本书的编写吸收了国内外优秀教材的精华和作者二十年讲授算法设计与分析课程的教学实践经验，力图贡献出一本适应当前教育形势所需要的高质量的教材，然而限于水平和精力，错误和不妥之处在所难免，恳请同行专家和阅读本书的老师和同学多多指正。

刘　璟

2003.4

目　　录

试读结束：需要全本请在线购买： www.ertongbook.com

1 绪　论

算法的理论处于计算机科学的核心地位，同时，它与计算机应用的许多实际问题有着直接的关系，例如，网络技术、多媒体技术、计算机安全技术等，其最关键的部分都是算法设计问题。因此，让我们从一个最常见的应用问题开始，体会算法理论是如何在实际问题中发挥作用的，以及人们如何是进行计算机算法的设计与分析的。

1.1 交通信号灯问题

1.1.1 问题

城市中用交通信号灯来管理交叉路口的车辆的通行。现代城市的交通信号管理应达到如下两个目标。

(1) 每一车辆通过路口时都不会与其他车辆发生冲突。

(2) 车辆在路口等待通过的时间尽可能少。

一般信号灯的设置方法是确定几个固定相位，在同一相位下允许通过的路线不相交，同时，为了减少等待时间，应使相位数最少。

1.1.2 实例

图 1.1 是一个五叉路口的实例，五条路 a,b,c,d,e 相交在一起，其中 c,e 为单行线，车辆通过该路口时共有 13 种不同路线：

ab,ac,ad,

ba,bc,bd,

da,db,dc,

ea,eb,ec,ed。

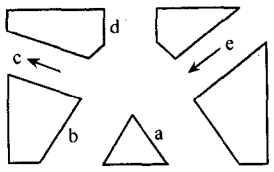

图 1.1　一个五叉路口

各条线路可能相交，也不能不相交，例如，ab 与 dc 不相交，而 ac 与 eb 是相交的。

把信号灯设置为 13 个相位，每次只允许一条路线通行，当然不会产生冲突，但显然这样效率太低，车辆等待时间过长。把不相冲突的路线放在同一相位，允许同时通过是个好办法。那么，这 13 条路线最少可以分为几个不同相位呢？故问题归结为，把所有路线分为尽可能少的组，组中的路线互不相交。当然，有的路线可同时出现在不同的组。熟悉算法知识的人会想到，这样的问题实际上是一个图着色问题。

1.1.3 图着色问题

图着色(Graph Coloring)问题是一个经典的组合算法问题，它来源于地图着色问题。

在绘制地图时，总是要求相邻的国家着上不同的颜色以示区别。1852 年一位美国大学生古德里(F. Guthrie)提出了一个猜测：为了给任意一个平面地图着色，并使任何有公共边界的区域颜色不同，至多需要四种颜色。这就是四色问题，又称四色猜想。古德里仅仅是提出了这个猜想，他和他的老师都未能证明。二十几年后的 1878 年，著名数学家凯莱(A. Kayley)发表了一篇“论地图着色”的文章，虽然仍没有解决这个问题，却掀起了四色猜想的研究热。经过一百年的探索，美国伊利诺大学的哈肯(W. Haken)与阿佩尔(K. Appel)通过改进算法，借助计算机(共用了 1200 个机时)终于在 1976 年完成了四色猜想的证明。当地的邮局在寄出的信上除了通常的邮戳外，还要加盖“四色是足够的”一句话以表示自豪。

当把地图(Map)上的一个国家与图(Graph)上的一个顶点对应，两个国家的相邻关系对应于无向图上的边，于是，上面关于地图的“四色问题”实际上是无向图的顶点着色问题的特例。

已知：n 点无向图 $G=\langle V,E\rangle$，$||V||=n$。

求：G 的最小色数 k，使得用 k 种颜色对 G 的顶点着色，可令任意两个相邻顶点着色不同。

着色问题是一个有名的 NP-完全问题，我们将在第 9 章中指出，这类问题属于“计算机难解”问题，关于着色问题算法的研究，许多学者已有大量成果。

类似于地图着色问题，交通信号灯问题也可以归结为图的顶点着色问题。把路口的所有路线表示为无向图的顶点，两个顶点之间有一边连接(当且仅当对应的两条路线在路口相交)，而把各个路线划分为尽可能少的互不相交的 k 组，恰恰等价于求相应的无向图的最小色数 k。

具体的解决方案如下。

无向图顶点着色问题：顶点，边，最小色数 k。

地图着色问题：国家，相邻关系，最小色数 k。

交通信号灯问题：路线，交叉关系，最小相位(组)数 k。

图 1.2 给出的无向图对应于我们的五叉路口实例，该图有 13 个结点和 20 条边，其中 ba，dc，ed 三个顶点是孤立点，说明这三条路线与其他所有路线不相交，就是所谓的“拐小弯”。

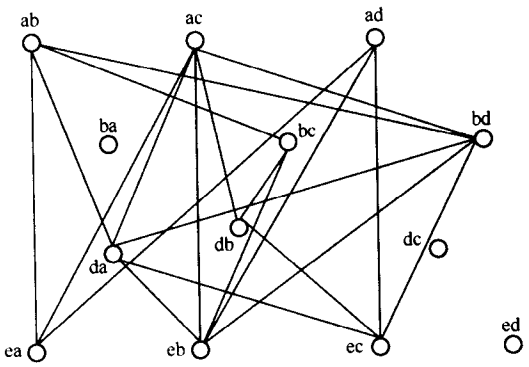

图 1.2　五叉路口信号灯问题对应的无向图

试读结束：需要全本请在线购买：　www.ertongbook.com