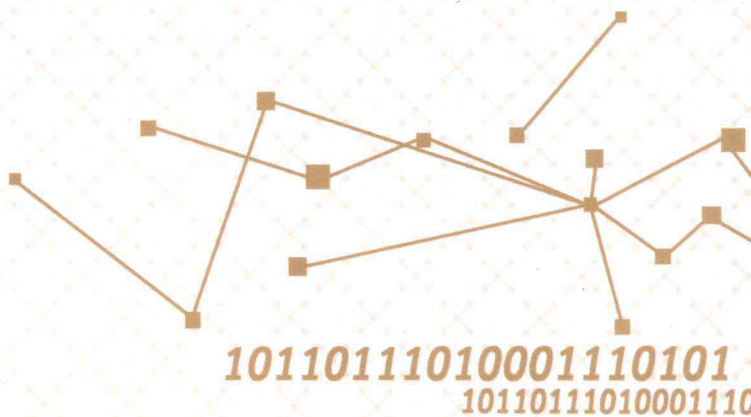




“十二五”国家重点图书出版规划项目
中国科学技术大学精品教材



袁平波 顾为兵 尹东 朱明 / 编著

数据结构及 应用算法

第2版

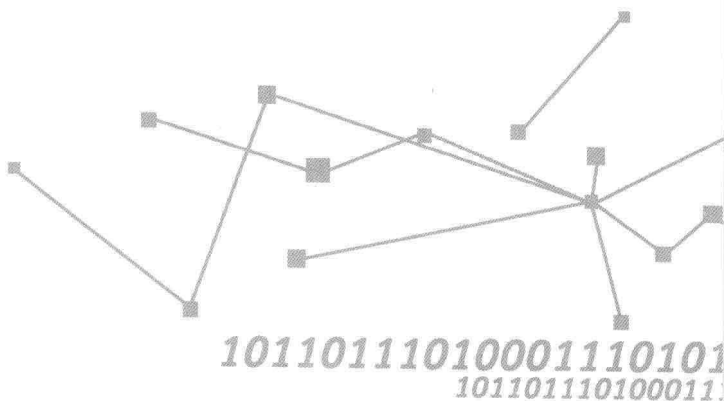
Data Structure and Algorithm

中国科学技术大学出版社



“十二五”国家重点图书出版规划项目

中国科学技术大学精品教材



袁平波 顾为兵 尹东 朱明 / 编著

Data Structure and Algorithm

数据结构及 应用算法

第2版

中国科学技术大学出版社

内 容 简 介

计算机程序设计主要包括数据对象的定义表示和数据对象的处理算法两大部分。本书从数据对象的类型、表示方法及其常用处理方法入手,分别介绍了三种类型数据结构——线性结构、树状结构和图状结构的常用表示方法,以及基于这些数据结构的基本操作函数的实现。此外,本书还介绍了查找和排序算法。同时为方便读者更好地学习掌握数据结构知识,本书还介绍了C程序设计预备知识和计算机主要算法的设计策略等内容,并为大部分章节安排了习题。

本书内容全面丰富,概念阐述清晰,不仅适合作为普通高校信息技术类专业的本科生教材,也适合作为信息技术相关工科专业的“数据结构”或“软件工程”课程的本科教材。对于从事信息技术方面学习和工作的科技人员,本书也是一本很好的参考书。

图书在版编目(CIP)数据

数据结构及应用算法/袁平波等编著. —2版.—合肥:中国科学技术大学出版社,2013.9

(中国科学技术大学精品教材)

“十二五”国家重点图书出版规划项目

ISBN 978-7-312-03320-9

I. 数… II. 袁… III. ① 数据结构—高等学校—教材 ② 算法分析—高等学校—教材 IV. TP311.12

中国版本图书馆 CIP 数据核字(2013)第 195907 号

中国科学技术大学出版社出版发行

安徽省合肥市金寨路 96 号,230026

<http://press.ustc.edu.cn>

安徽省瑞隆印务有限公司印刷

全国新华书店经销

开本:710 mm×960 mm 1/16 印张:24.5 插页:2 字数:471 千

2008 年 9 月第 1 版 2013 年 9 月第 2 版 2013 年 9 月第 2 次印刷

印数:3001—6000 册

定价:45.00 元

总 序

2008年,为庆祝中国科学技术大学建校五十周年,反映建校以来的办学理念 and 特色,集中展示教材建设的成果,学校决定组织编写出版代表中国科学技术大学教学水平的精品教材系列。在各方的共同努力下,共组织选题281种,经过多轮、严格的评审,最后确定50种入选精品教材系列。

五十周年校庆精品教材系列于2008年9月纪念建校五十周年之际陆续出版,共出书50种,在学生、教师、校友以及高校同行中引起了很好的反响,并整体进入国家新闻出版总署的“十一五”国家重点图书出版规划。为继续鼓励教师积极开展教学研究与教学建设,结合自己的教学与科研积累编写高水平的教材,学校决定,将精品教材出版作为常规工作,以《中国科学技术大学精品教材》系列的形式长期出版,并设立专项基金给予支持。国家新闻出版总署也将该精品教材系列继续列入“十二五”国家重点图书出版规划。

1958年学校成立之时,教员大部分来自中国科学院的各个研究所。作为各个研究所的科研人员,他们到学校后保持了教学的同时又作研究的传统。同时,根据“全院办校,所系结合”的原则,科学院各个研究所在科研第一线工作的杰出科学家也参与学校的教学,为本科生授课,将最新的科研成果融入到教学中。虽然现在外界环境和内在条件都发生了很大变化,但学校以教学为主、教学与科研相结合的方针没有变。正因为坚持了科学与技术相结合、理论与实践相结合、教学与科研相结合的方针,并形成了优良的传统,才培养出了一批又一批高质量的人才。

学校非常重视基础课和专业基础课教学的传统,也是她特别成功的原因之一。当今社会,科技发展突飞猛进、科技成果日新月异,没有扎实的基础知识,很难在科学技术研究中作出重大贡献。建校之初,华罗庚、吴有训、严济慈等老一辈科学家、教育家就身体力行,亲自为本科生讲授基础课。他们以渊博的学识、精湛的讲课艺术、高尚的师德,带出一批又一批杰出的年轻教员,培养

再 版 前 言

应广大读者的要求,《数据结构及应用算法》在出版5年后决定进行重要修订。在这次修订中,更正了上一版中出现的一些文字编辑错误,调整了部分章节的内容,强化了对一些概念的描述,增补了一些算法代码,并在每章的后面增加了本章小结和配套习题,使本书更加适合作为普通高校信息技术类专业的教材。

在本次修订中,我们秉承了第1版的风格,帮助学生学会从问题入手,分析和研究数据结构的特性,使学生学会在解决问题时用正确的逻辑结构描述数据、用合理的存储结构表示数据和用有效的操作方法处理数据,并初步掌握算法的性能分析技术。本版教材强化了数据结构中抽象数据类型(ADT)的概念,使ADT成为贯穿本教材的主线。

本书可用来作为普通高校电子信息类本科生60~80学时的数据结构课程教材,配有PPT教学课件和实验教材《数据结构实验指导》,以方便教学。本书的第1章、第2章和第11章由朱明老师编写,第3章、第4章和第5章由袁平波老师编写,第6章、第7章由顾为兵老师编写,第8章、第9章和第10章由尹东老师编写。

感谢广大读者一直以来对本教材的支持。

袁平波 顾为兵 尹东 朱明

中国科学技术大学信息科学技术学院

2013年8月

前 言

作为计算机程序设计的重要理论技术基础,“数据结构”是计算机专业的核心基础课程。它涉及数据在计算机中的表示、组织与处理,以及相应结构上的算法设计和初步的算法性能分析技术。它所包含的知识与提倡的技术方法,无论是对学生进一步学习信息技术领域的其他课程,还是对今后从事信息技术方面的理论研究、应用开发及技术管理工作都发挥着重要的作用。

目前,作为最基本的计算机学科核心课程,“数据结构”已发展成为信息技术乃至工科的必修课程。本教材正是针对这一背景和需求而编写的。在本书的内容组织方面,充分考虑了信息技术专业甚至是工科等非计算机专业学生选修学习此课程的实际需要。

帮助学生学会从问题入手,分析和研究计算机加工的数据结构的特性,使其能够为应用中的数据选择适当的逻辑结构、存储结构及其相应的操作算法,并初步掌握算法的性能分析技术是“数据结构”课程的基本教学要求。这同时也是对学生进行复杂程序设计训练并培养学生数据抽象能力的过程;这种能力在软件系统开发过程中体现为创造性思维活动以及算法构造性思维方法。因此,让学生理解、习惯和熟悉这一套算法构造性思维方法,以适应学科迅速发展和知识更新的需要,正是这门课程教学的目的和宗旨。

为了方便信息技术专业甚至是工科的学生学习掌握本书内容,本书采用大量实例对数据结构的各个知识点进行了详细的解释和说明。每一类数据结构中均给出了多个具有代表性的应用实例,便于读者理解掌握各类数据结构的基本表示方法,基本操作函数的实现方法,以及如何利用这些基本数据结构和操作函数完成复杂问题的计算机求解,从而帮助读者真正熟悉掌握所学习的数据结构内容。

本书的第1章给出学习数据结构所需要掌握的C语言程序设计预备知识,重点复习了数据结构中需要熟练使用的指针、结构体、函数设计和文件操作等内容;第2章介绍了数据、数据结构及其算法等基本概念与基本知识;第3章介绍了线性结构及其基本操作内容,并给出基于指针与数组表示情况下,其基本操作函数的实

目 次

总序	(i)
再版前言	(iii)
前言	(v)
第 1 章 预备知识	(1)
1.1 程序设计概述	(1)
1.2 指针与结构体	(11)
1.3 文件操作	(21)
1.4 函数与模块化程序设计	(27)
1.5 本章小结	(37)
习题	(37)
第 2 章 数据结构导论	(41)
2.1 概念与术语	(41)
2.2 抽象数据类型	(46)
2.3 算法概述	(48)
2.4 算法分析	(51)
2.5 本章小结	(60)
习题	(60)
第 3 章 线性表	(63)
3.1 线性表的基本概念	(63)
3.2 线性表的顺序表示	(67)
3.3 线性表的链式表示	(74)
3.4 线性结构的深入	(83)
3.5 本章小结	(92)
习题	(93)

第 4 章	栈和队列	(95)
4.1	栈的基本概念	(95)
4.2	栈的表示与实现	(97)
4.3	栈的应用	(102)
4.4	队列的基本概念	(111)
4.5	队列表示与实现	(113)
4.6	队列的应用	(119)
4.7	递归及其应用	(122)
4.8	本章小结	(128)
	习题	(129)
第 5 章	串和数组	(132)
5.1	串的基本概念	(132)
5.2	串表示与实现	(134)
5.3	串的应用	(138)
5.4	模式匹配	(139)
5.5	数组	(145)
5.6	矩阵压缩	(150)
5.7	本章小结	(157)
	习题	(158)
第 6 章	树和二叉树	(160)
6.1	树的基本概念	(160)
6.2	二叉树的概念	(165)
6.3	二叉树的遍历及应用	(173)
6.4	线索二叉树	(183)
6.5	树和森林	(188)
6.6	哈夫曼树和哈夫曼编码	(202)
6.7	本章小结	(210)
	习题	(211)
第 7 章	图	(214)
7.1	图的基本概念	(214)
7.2	图的表示与实现	(220)
7.3	图的遍历	(230)
7.4	最小生成树	(236)

7.5	拓扑排序	(242)
7.6	关键路径	(245)
7.7	最短路径	(248)
7.8	本章小结	(253)
	习题	(254)
第 8 章	查找表	(257)
8.1	查找表的基本概念	(257)
8.2	静态查找表	(258)
8.3	动态查找表	(265)
8.4	本章小结	(288)
	习题	(288)
第 9 章	排序	(290)
9.1	排序的基本概念	(290)
9.2	简单排序	(292)
9.3	希尔排序	(298)
9.4	快速排序	(300)
9.5	堆排序	(307)
9.6	归并排序	(310)
9.7	基数排序	(313)
9.8	本章小结	(317)
	习题	(320)
第 10 章	文件	(322)
10.1	文件的基本概念	(322)
10.2	顺序文件	(323)
10.3	索引文件	(325)
10.4	ISAM 文件	(328)
10.5	散列文件	(329)
10.6	本章小结	(330)
	习题	(330)
第 11 章	算法设计策略	(332)
11.1	概述	(332)
11.2	分治策略	(335)
11.3	贪心策略	(343)

11.4	动态规划策略	(349)
11.5	回溯策略	(357)
11.6	分枝定界策略	(372)
11.7	本章小结	(379)
	习题	(380)
	参考文献	(381)

第 1 章 预 备 知 识

一台计算机是由硬件系统和软件系统两大部分构成的。硬件是物质基础,而软件可以说是计算机的灵魂,没有软件,计算机就是一台“裸机”,是什么也不能干的,有了软件,才能运行起来,成为一台真正的“电脑”。所有的软件程序,都是用计算机语言编写的。

计算机程序就是计算机求解问题的指令序列,它包括两个主要部分:对待处理对象数据的定义表示,以及对待处理对象的处理步骤。

如何表示待处理对象数据,就是数据结构所要讲授的内容;而如何通过对处理对象的操作处理,完成问题的求解,则属于算法设计的范畴。

计算机程序是利用计算机编程语言编写而成的,因此为了学好计算机程序编写方法,即数据结构及其算法,首先就需要掌握计算机编程语言,所以本章对目前应用最为普遍的计算机编程语言——C 语言,进行简要回顾,以便为学习好数据结构及其算法奠定基础。

本章首先介绍程序设计的基本知识,具体包括基本控制语句和基本程序设计方法;然后温习 C 语言中较难的部分,具体包括指针、结构体、函数和文件,最后介绍模块化程序设计方法。

1.1 程序设计概述

任何一种程序设计语言都具有特定的语法规则和规定的表达方法。一个程序只有严格按照语言规定的语法和表达方式编写,才能保证编写的程序在计算机中能正确地执行,同时也便于阅读和理解。为了了解 C 语言的基本程序结构,先介绍


```
#include <stdio.h>

main( )                                //主函数
{
    int x,y,z,v;                        //定义整型变量
    scanf("%d,%d,%d",&x,&y,&z); //调用标准函数,从键盘输入 x,y,z 的值
    v = volume(x,y,z);                 //调用 volume 函数,计算体积
    printf("v = %d\n",v);
}

int volume(a,b,c)                      //定义 volume 函数
    int a,b,c;                         //对形参 a,b,c 作类型定义
{
    int p;                             //定义函数内部使用的变量 p
    p = a * b * c;                     //计算体积 p 的值
    return(p);                         //将 p 值返回调用处
}
```

从上面的程序例子,可以看出 C 程序的基本结构。

C 程序为函数模块结构,所有的 C 程序都是由一个或多个函数构成,其中必须只能有一个主函数 `main( )`。程序从主函数开始执行,当执行到调用函数的语句时,程序将控制转移到调用函数中执行,执行结束后,再返回主函数中继续运行,直至程序执行结束。C 程序的函数是由编译系统提供的标准函数(如 `printf`、`scanf` 等)和由用户自己定义的函数(如 `proc`、`func`、`volume` 等)。

1.1.1 程序控制语句

C 程序的组织结构如图 1.1 所示。

C 程序的执行部分是由语句组成的。程序的功能也是由执行语句实现的。C 语句可分为以下五类:表达式语句、函数调用语句、控制语句、复合语句和空语句。

通常的计算机程序总是由若干条语句组成,从执行方式上看,从第一条语句到最后一条语句完全按顺序执行,是简单的顺序结构;若在程序执行过程当中,根据用户的输入或中间结果去执行若干不同的任务则为选择结构;如果在程序的某处,需要根据某项条件重复地执行某项任务若干次或直到满足或不满足某条件为止,这就构成循环结构。大多数情况下,程序都不会是简单的顺序结构,而是顺序、选


```
printf("The result is %d\n",min);
}
```

if 语句只能处理从两者间选择之一,当要实现几种可能之一时,就要用 if... else if,甚至多重的嵌套 if 来实现。当分支较多时,程序变得复杂冗长,可读性降低。C 语言提供了 switch 开关语句专门处理多路分支的情形,使程序变得简洁。

例 1.5 编写一个 C 程序,将键盘输入的一个数转换为英文星期一至星期日并输出。

```
#include <stdio.h>
main( )    //主函数
{
    int a;
    printf("input integer number:");
    scanf("%d",&a);
    switch (a){
        case 1:printf("Monday\n");
        case 2:printf("Tuesday\n");
        case 3:printf("Wednesday\n");
        case 4:printf("Thursday\n");
        case 5:printf("Friday\n");
        case 6:printf("Saturday\n");
        case 7:printf("Sunday\n");
        default:printf("error\n");
    }
}
```

循环控制结构(又称重复结构)是程序中的另一个基本结构。在实际问题中,常常需要进行大量的重复处理,循环结构可以使我们只写很少的语句,而让计算机反复执行,从而完成大量类似的计算。

C 语言提供了 while 语句、do ... while 语句和 for 语句实现循环结构。

例 1.6 编写一个 C 程序,利用格里高利公式求 π : $\pi/4 = 1 - 1/3 + 1/5 - 1/7 + \dots$ 直到最后一项小于等于 10^{-6} 为止。