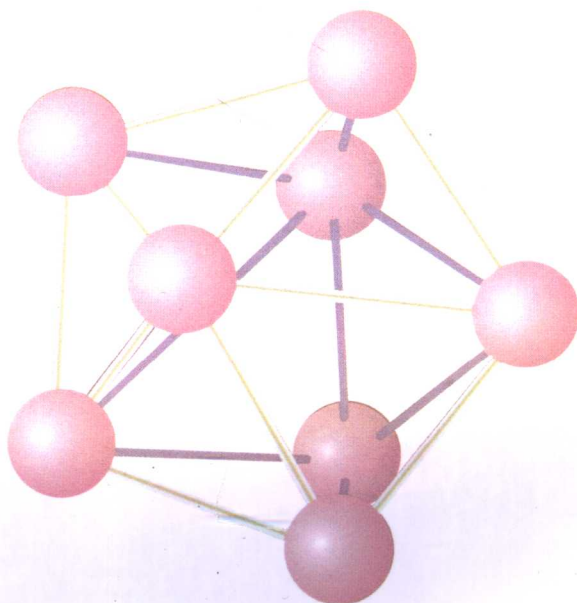
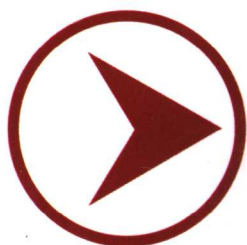




高等职业院校规划教材·计算机应用技术系列

数据结构

张亦辉 李 波 主编



中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

高等职业院校规划教材·计算机应用技术系列

数 据 结 构

主 编 张亦辉 李 波

副主编 周朋红 张克瑜

参 编 李婷婷 牛文琦 李 琳

吴振国 解英杰

内 容 简 介

本书主要阐述数据结构的相关概念及算法分析方法,分别介绍了几种常见的数据结构类型。全书共分9章,主要内容包括顺序表、链式表、栈、队列、串、多维数组、广义表、树、图、查找和内部排序;第9章针对本书各章的内容,提供了9个实验指导。

本书融入山东省“数据结构”精品课程在教学方法、教学内容上的先进思想,做到理论知识的阐述由浅入深、通俗易懂,内容的组织和编排以应用为主线,各章都配有“应用举例分析”(除第1、8、9章),列举并分析很多实用的例子,便于教学中举例,并且有助于学生加深对基础理论知识的理解,培养学生的实际应用能力。

本书可作为高等职业学校计算机类专业学生学习“数据结构”的教材。为方便学生学习,本书的程序部分均采用 Turbo C 语言描述,学生可以很方便地将书中的算法上机测试。

图书在版编目(CIP)数据

数据结构/张亦辉,李波主编. —北京:中国铁道出版社,2007.7

高等职业院校规划教材. 计算机应用技术系列
ISBN 978-7-113-08119-5

I. 数… II. ①张…②李… III. 数据结构—高等学校: 技术学校—教材 IV. TP311.12 ~

中国版本图书馆CIP数据核字(2007)第133575号

书 名: 数据结构

作 者: 张亦辉 李 波 等

出版发行: 中国铁道出版社(100054,北京市宣武区右安门西街8号)

策划编辑: 严晓舟 秦绪好

责任编辑: 翟玉峰

特邀编辑: 薛秋沛

封面设计: 傅 巍

封面制作: 白 雪

责任校对: 贾 星

印 刷: 三河市国英印务有限公司

开 本: 787×1092 1/16 印张: 14.75 字数: 342千

版 本: 2007年8月第1版 2007年8月第1次印刷

印 数: 1~5 000册

书 号: ISBN 978-7-113-08119-5/TP·2473

定 价: 20.00元

版权所有 侵权必究

本书封面贴有中国铁道出版社激光防伪标签,无标签者不得销售

凡购买铁道版的图书,如有缺页、倒页、脱页者,请与本社计算机图书批销部调换。

高等职业院校规划教材

主任：邓泽民（教育部职业技术教育中心研究所）

委员：（按姓氏音序排列）

安志远（北华航天工业学院）
陈焕文（湖南信息职业技术学院）
丁桂芝（天津职业大学）
高 嵩（北京青年政治学院）
贺 平（番禺职业技术学院）
蒋川群（上海第二工业大学）
李 雪（安徽职业技术学院）
刘兴东（深圳职业技术学院）
曲宏山（郑州经济管理干部学院）
宋文官（上海商学院）
王首义（黑龙江司法警官职业学院）
徐 红（山东商业职业技术学院）
宣仲良（苏州市职业大学）
严晓舟（中国铁道出版社）
张 剑（广东纺织职业技术学院）
张晓云（西安航空技术高等专科学校）
张亦辉（济南铁道职业技术学院）

高等职业院校规划教材·计算机应用技术系列

主任: 石 冰

副主任: 徐 红 张亦辉

委 员: (按姓氏音序排列)

安丰彩	崔凤磊	黄丽民	李 敏	梁国浚
刘 学	毛红旗	彭丽英	曲桂东	孙学农
王明晶	王兴宝	王秀红	王宜贵	王 玉
徐新艳	燕居怀	杨 云	张序政	赵吉兴
赵 敬				

出版说明

自 2002 年全国职业教育工作会议以来, 全国各地区、各部门认真贯彻《国务院关于大力推进职业教育改革与发展的决定》(国发[2002] 16 号), 加强了对职业教育工作的领导和支持, 以就业为导向改革与发展职业教育逐步成为社会共识。2005 年, 在北京召开的全国职业教育工作会议上, 国务院总理温家宝提出, 在今后一个时期, “教育结构调整总的方向是, 普及和巩固义务教育, 大力发展职业教育, 提高高等教育质量”, “把基础教育、职业教育和高等教育放在同等重要位置”。此次讲话精神将职业教育的地位提到了一个新的高度, 为大力发展职业教育奠定了思想基础, 指明了方向。

作为高等职业教育的重要组成部分, 计算机教育和教学也面临着“以就业为导向”的重要转变和改革。为顺应高等职业教育改革和发展的趋势, 配合高等院校的教学改革和教材建设, 中国铁道出版社联合全国知名职业教育专家和各大职业院校推出了《高等职业院校规划教材》系列丛书。

本套系列教材编写的主要指导思想:

- (1) 定位明确。整套教材贯穿了“以就业为导向”的思想, 面向就业, 突出实际应用。
- (2) 内容先进。教材合理安排经典知识和实际应用的内容, 补充了新知识、新技术和新设备。
- (3) 取舍合理。以高等职业教育的培养目标为依据, 注重教材的科学性、实用性和通用性, 尽量满足同类专业院校的需求。
- (4) 体系得当。以岗位职业标准为依据设计教材的体系, 体现岗位技能要求, 紧密结合生产实际, 强化实践环节, 培养创新精神。
- (5) 风格优良。在编写方式和配套建设中体现建设“立体化”精品教材体系的宗旨。为主要课程配套了电子教案、教学大纲、学习指导、习题解答、素材库、案例库、试题库等相关教学资源。

本套教材在编写过程中参考了《中国高职院校计算机教育课程体系 2007》(英文简称为 CVC 2007) 中各专业课程体系的参考方案, 并根据专业类别划分系列, 分为计算机应用技术系列、信息管理技术系列、多媒体技术系列、网络技术系列、软件技术系列、电子商务系列等若干子系列。在本系列丛书的编写和出版过程中, 得到了各专业领域知名职业教育专家以及全国各大高等职业院校的大力支持, 在此表示衷心感谢。希望本系列丛书的出版能为我国高等职业院校计算机教育改革起到良好的推动作用, 欢迎使用本系列教材的老师和同学提出意见和建议, 书中如有不妥之处, 敬请批评指正。



“数据结构”是计算机程序设计的重要理论知识，是计算机相关专业必修的一门重要基础课程和核心课程。本书是中国铁道出版社“高等职业院校规划教材”之一，也是山东省“数据结构”精品课程的立体化教材之一，主要面向高职高专和应用型本科院校计算机相关专业的学生。

本书秉承“以应用为主体、培养实践能力”的思想，在强调理论知识的理解和运用的同时，更加注重结合高职高专和应用型本科教学的要求，体现对学生应用技能的培养。

全书共分9章。第1章主要阐述数据结构的相关概念及算法分析方法；第2章讨论线性结构中的顺序表和链式表的表示、存储和操作；第3章讨论栈和队列的原理及应用；第4章讨论其他的线性数据结构，重点介绍串、多维数组和广义表的相关概念；第5章讨论树的原理和应用；第6章讨论图的原理和基本操作；第7、8章讨论查找和内部排序的概念和基本算法；第9章针对本书各章的内容，提供了9个实验指导。

本书具有以下特点：

(1) 融入山东省“数据结构”精品课程在教学方法、教学内容上的先进思想，做到理论知识的阐述由浅入深、通俗易懂，内容的组织和编排以应用为主线，略去了一些理论推导和数学证明的过程，淡化算法的设计分析和复杂的时空分析，仅给出分析的结果。我们还将山东省“数据结构”精品课程的相关资源放在网上 (<http://www.jnvp.edu.cn>) 供广大读者学习参考。

(2) 各章（除第1、8、9章）都配有“应用举例分析”，列举并分析很多实用的例子，便于教学中举例，并且有助于学生加深对基础理论知识的理解，培养学生的实际应用能力。

(3) 各章均指出本章要点，并配有本章小结和大量的课后习题，提供了部分习题的参考答案，方便学生课后复习、巩固。

(4) 各章（除第1、7、9章），均配有1~2个上机实验指导，并且贯彻分层次教学的思想，在实验指导中增加了难度较大的“拓展训练”内容，供能力较强的学生练习、提高。

(5) 为方便学生学习，本书的程序部分均采用 Turbo C 语言描述，学生可以很方便地将书中的算法上机测试。

本教材的教学时数以60~80为宜，其中上机时数安排在20课时左右，教师可以根据实际的教学时数、学生情况等自行调整进度和教学内容。

本书由济南铁道职业技术学院张亦辉、李波主编，济南铁道职业技术学院周朋红、德州职业技术学院张克瑜副主编，参加编写的济南铁道职业技术学院李琳老师编写了第4章，济南铁道职业技术学院李婷婷老师编写了第5章，济南铁道职业技术学院牛文琦老师编写了第2章，济南铁道职业技术学院李波老师编写了第1、6章，周朋红老师编写了第3、8章，德州职业技术学院张克瑜老师、潍坊教育学院吴振国老师、山东技师学院解英杰老师编写了第9章及部分程序的上机调试等工作，张亦辉老师编写了第7章和统一定稿。中国铁道出版社、济南铁道职业技术学院信息工程系以及兄弟院校的领导和老师在本书编写过程中给予了大力支持，在此表示衷心的感谢。

由于时间紧迫，加之编者水平有限，书中难免存在疏漏和不妥之处，恳请读者批评指正。

编者

2007年5月

第 1 章 绪论	1
1.1 引言	1
1.2 基本概念和术语	3
1.3 算法描述	5
1.3.1 算法的重要特性	5
1.3.2 数据结构的基本操作	5
1.3.3 算法的描述方法	5
1.4 算法分析	6
1.4.1 算法设计的要求	6
1.4.2 算法时间效率的度量分析	7
本章小结	8
习题一	9
第 2 章 线性表	11
2.1 线性表的概念和基本操作	11
2.2 线性表的顺序存储结构	12
2.2.1 顺序表的定义	12
2.2.2 顺序表中元素的插入	14
2.2.3 顺序表中元素的删除	15
2.2.4 顺序表中元素的定位	16
2.3 线性表的链式存储结构	16
2.3.1 单链表的定义和操作实现	16
2.3.2 循环链表的定义和操作实现	23
2.3.3 双向链表的定义和操作实现	24
2.4 线性表两种存储方式的比较	25
2.5 应用举例分析	26
本章小结	31
习题二	31
第 3 章 栈和队列	34
3.1 栈	34
3.1.1 栈的定义和基本操作	34
3.1.2 栈的顺序存储结构和操作实现	35
3.1.3 栈的链式存储结构和操作实现	37
3.1.4 递归和栈	38
3.1.5 算术表达式的计算	41
3.2 队列	48
3.2.1 队列的定义和基本操作	48

3.2.2 队列的顺序存储结构和操作实现	49
3.2.3 队列的链式存储结构和操作实现	52
3.3 应用举例分析	55
本章小结	58
习题三	58
第4章 其他线性数据结构	61
4.1 串	61
4.1.1 串的定义和基本操作	61
4.1.2 串的实现	62
4.1.3 串的基本操作的实现	63
4.2 多维数组	66
4.2.1 多维数组的定义和存储结构	66
4.2.2 稀疏矩阵的压缩	68
4.2.3 特殊矩阵的压缩	71
4.3 广义表	73
4.3.1 广义表的定义和基本运算	73
4.3.2 广义表的存储结构	74
4.4 应用举例分析	76
本章小结	77
习题四	78
第5章 树和二叉树	81
5.1 树的基本定义	81
5.1.1 树的定义	81
5.1.2 基本术语	82
5.2 二叉树	83
5.2.1 二叉树的定义和基本操作	83
5.2.2 二叉树的性质	84
5.2.3 二叉树的存储结构	86
5.2.4 二叉树的遍历	89
5.3 树和森林	95
5.3.1 树的存储结构	95
5.3.2 树、森林和二叉树的转换	98
5.3.3 树和森林的遍历	100
5.4 树的应用	101
5.4.1 哈夫曼树和哈夫曼编码	101
5.4.2 线索二叉树	104
5.5 应用举例分析	107
本章小结	111
习题五	111

第 6 章 图	116
6.1 图的定义和术语	116
6.2 图的存储结构	119
6.2.1 邻接矩阵	119
6.2.2 邻接链表和逆邻接链表	122
6.2.3 十字链表	125
6.2.4 邻接多重表	125
6.3 图的遍历	126
6.3.1 深度优先遍历	127
6.3.2 广度优先遍历	128
6.4 图的应用	130
6.4.1 生成树和最小生成树	130
6.4.2 拓扑排序	133
6.4.3 关键路径	136
6.4.4 最短路径	138
6.5 应用举例分析	141
本章小结	146
习题六	147
第 7 章 查找	151
7.1 查找的基本概念	151
7.2 线性表查找	153
7.2.1 顺序查找	153
7.2.2 二分法查找	154
7.3 索引顺序查找	156
7.4 动态查找	158
7.4.1 二叉排序树的概念和操作	158
7.4.2 平衡二叉树	161
7.4.3 B_树和 B ⁺ 树	165
7.5 散列表	167
7.5.1 散列表和散列函数	167
7.5.2 散列函数的构造	168
7.5.3 解决冲突的主要方法	169
7.5.4 散列表的查找和性能分析	171
7.6 应用举例分析	173
本章小结	179
习题七	180
第 8 章 内部排序	183
8.1 排序的基本概念	183
8.2 简单排序	184
8.2.1 直接插入排序	184

8.2.2 冒泡排序	185
8.2.3 选择排序	187
8.3 高级排序	188
8.3.1 希尔排序	188
8.3.2 快速排序	190
8.3.3 堆排序	192
8.3.4 归并排序	196
8.3.5 基数排序	198
本章小结	201
习题八	201
第 9 章 实验指导	207
实验一 顺序表的建立与基本操作	207
实验二 链表的操作	208
实验三 栈与队列的基本操作	210
实验四 串的操作与稀疏矩阵的压缩	211
实验五 二叉树的基本操作	212
实验六 树的应用	214
实验七 图的建立与遍历	215
实验八 图的应用	216
实验九 简单内部排序	217
参考文献	221
附录 A 各章习题答案	222

第 1 章 绪 论

电子计算机是 20 世纪科学技术最卓越的成就之一。自 1946 年第一台电子计算机问世以来, 计算机产业和应用的发展远远超出了人们的预料。如今, 计算机应用已不再局限于科学计算, 而更多地用于企业管理、工业过程控制、管理信息系统和实时控制等非数值计算的各个方面。与此相对应, 计算机处理的对象也由纯粹的数值数据发展到诸如字符、图像、声音和信号等各种各样的具有一定结构的数据。为了有效地组织和管理这些数据, 编制出高质量的程序, 高效率地使用计算机, 必须深入研究这些数据的特性以及它们之间的相互关系。这就是“数据结构”这门学科形成和发展的背景。

本章要点

- 数据结构的基本概念
- 算法的特征
- 算法复杂度分析

1.1 引 言

数据结构主要研究非数值应用问题中数据之间的逻辑关系和对数据的操作, 同时还研究如何将具有逻辑关系的数据按一定的存储方式存放在计算机内。分析数据之间的逻辑关系和确定数据在计算机内的存储结构是进行程序设计前必须完成的两个任务。

处理非数值计算问题和数值计算问题的解决方案不同。例如, 求解梁架结构中应力的数学模型为线性方程组, 预报人口增长情况的数学模型为微分方程。但还有更多的非数值计算问题是无法用数学方程加以描述的。

【例 1.1】某单位职工档案的管理 (见图 1-1)。

工 号	姓 名	性 别	出生年月	婚 否	学 历	进 厂 日 期
0001	张莉莉	女	1981/10/01	否	本科	2006/09/01
0005	王明明	男	1980/02/28	否	本科	2005/04/01
0006	赵涛涛	男	1981/09/10	是	专科	2007/01/01
⋮	⋮	⋮	⋮	⋮	⋮	⋮

图 1-1 某单位职工档案表

图 1-1 中的职工档案表就是一个数据结构。计算机档案管理的主要功能包括查找、浏览、插入、修改、删除和统计等。如果把表中的一行看成一个记录, 并称为一个结点, 则在此表中, 结点和结点之间的关系是一种最简单的线性关系。

【例 1.2】某学校教师的名册。

虽然可以用例 1.1 中的二维表格将全校教师的名单列出, 但是采用如图 1-2 所示的结构更好。它像一棵根在上面倒挂的树, 清晰地描述了教师所在的系和教研组。这样可以从树根沿着某系某教研组很快找到某个教师, 查找的过程就是从树根沿分支到某个叶子的过

程, 类似于树这样的数据结构可以描述家族的家谱、企事业单位中的人事关系, 甚至可以用来反映人机对弈的动态过程等。

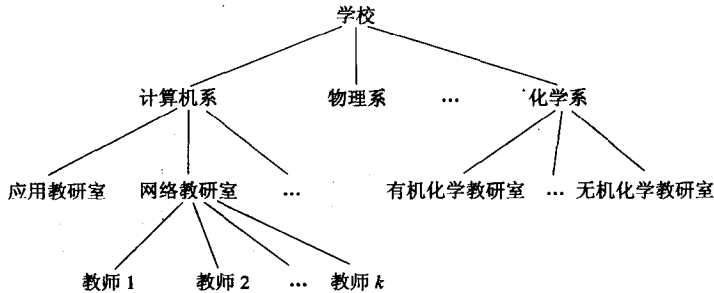
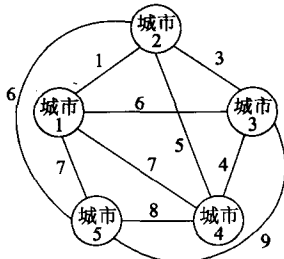


图 1-2 某学校教师名册

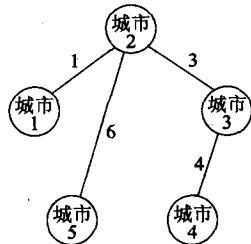
【例 1.3】通信网络设计。

在 n 个城市之间建立通信网络, 要求在其中任意两个城市之间都有直接的或间接的通信线路, 在已知某些城市之间直接通信线路预算造价的情况下, 使网络的造价最低。

当 n 很大时, 这样的问题只能用计算机来求解。可以用如图 1-3 (a) 所示的关系来说明: 图中的小圆圈表示一个城市, 两个圆圈之间的连线表示对应城市之间的通信线路, 连线上的数值表示该通信线路的造价。这一描述的结构称为图状结构, 利用计算机可以求出满足要求的通信网络, 如图 1-3 (b) 所示。



(a) 通信网络中的数据



(b) 最小造价通信网络

图 1-3 用图描述通信网络问题

通过上面三个例子可以看出: 数据结构中元素和元素之间存在着逻辑关系, 而线性表、树和图是三种基本的逻辑结构, 其他各类的数据结构都是由这三种基本结构派生出来的。数据结构就是解决如何分析数据元素之间的关系, 如何确定合适的逻辑结构, 以及如何存储这些数据, 并对为完成数据操作所设计的算法做出时间和空间的分析。“数据结构”在计算机科学中是一门综合性的专业基础课, 它不仅是一般程序设计 (特别是非数值计算的程序设计) 的基础, 而且也是设计和实现编译程序、操作系统、数据库系统及大型应用程序的重要基础。

简单来说, 数据结构是研究程序设计中数据的逻辑结构、物理结构和对数据的操作 (算法) 的一门课程。数据结构重点分析数据之间的抽象关系, 不涉及数据的具体内容。

1.2 基本概念和术语

数据 (Data) 指所有能输入到计算机中并被计算机程序处理的符号的总称。计算机输入和处理的数据除数值外,还有字符串、表格、图像甚至声音等,它们都属于数字编码范畴。

数据元素 (Data Element) 数据的基本单位,在计算机程序中通常作为一个整体进行考虑和处理。一个数据元素可以由若干个数据项组成,也可以只由一个数据项组成。数据元素又称为元素、结点 (Node) 或记录 (Record)。

数据项 (Data Item) 指数据中的不可分割的、含有独立意义的最小单位,数据项有时也称字段 (Field) 或域。例 1.1 中的职工档案表格是要存放到计算机中进行处理的数据,表中每一行记录了一个职工的档案信息,在数据操作中作为一个整体考虑,对应为一个数据元素,又称为一个记录。这个记录中包含工号、姓名、学历等若干个数据项。操作的基本单位是记录,如职工的插入或删除一定是作用于一个职工的全部信息即一个记录,而不可能是作用于其中的某个数据项。如果删除操作只删除某个职工的姓名或工号,将引起数据不完整等严重后果。每个数据项 (如职工的姓名或工号) 均有独立的含义,但在档案管理这个实际问题中并无完整的意义,而组合在一个记录中构成职工的档案,就具有了完整的实际意义。数据、数据元素和数据项实际上反映了数据组织的三个层次:数据可由若干个数据元素构成,而数据元素又可以由一个或若干个数据项组成。

数据逻辑结构 (Data Logical Structure) 数据结构主要研究数据元素之间的关联方式。数据元素之间存在的一种或多种特定的关系称为数据的逻辑结构。通常有下列四类基本的结构:

(1) 集合结构 在集合结构中,数据元素间的关系是“属于同一个集合”。集合是元素关系极为松散的一种结构。

(2) 线性结构 该结构的数据元素之间存在着一对一的关系。

(3) 树形结构 该结构的数据元素之间存在着一对多的关系。

(4) 图形结构 该结构的数据元素之间存在着多对多的关系,图形结构也称做网状结构。如图 1-4 所示为表示上述四类基本结构的示意图。

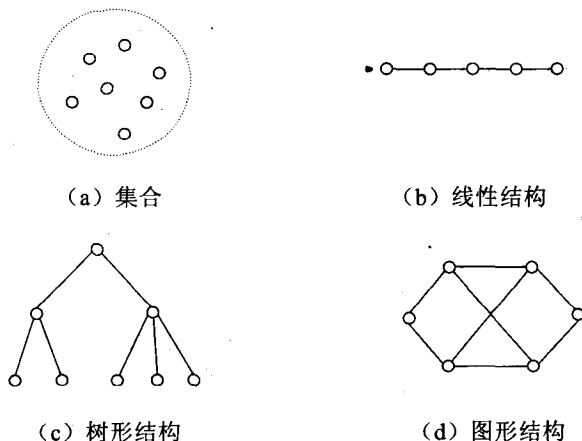


图 1-4 四类基本逻辑结构的示意图

数据物理结构 (Data Physical Structure) 数据在计算机中的存放方式称为数据的物理结构, 又称存储结构。数据的存储结构是逻辑结构在计算机存储器中的实现。数据元素在计算机中有多种不同的存储方法, 最重要的两种是顺序存储结构和链式存储结构。顺序存储的特点是在内存中开辟一组连续的空间 (高级语言中的数组) 来存放数据, 数据元素之间的逻辑关系通过元素在内存中存放的相对位置来确定, 又称向量存储。链式存储的特点是通过指针反映数据元素之间的逻辑关系, 又称动态存储。除此之外, 还有索引和散列两种存储方式。

数据的逻辑结构和物理结构是数据结构的密切相关的两个方面，逻辑结构主要为了方便程序员理解数据的形式，而物理结构主要描述数据在计算机中的存储方法。同一逻辑结构可以对应不同的物理存储结构，而不同的逻辑结构也可以对应相同的物理存储结构。算法的设计取决于数据的逻辑结构，而算法的实现依赖于指定的存储结构。

例如, 10 以内的奇数 1、3、5、7、9 用顺序存储结构的方式依次存放在以 300 为起始地址的内存向量中, 并且两个字长存放一个奇数, 如图 1-5 (a) 所示。在顺序存储结构中, 如要读取第三个奇数, 它的地址可以通过起始地址和要读取奇数的位置序号计算得到。若本例中的奇数改用链式存储结构存放, 那么, 第一个奇数存放在地址为 300 的内存单元中, 第二个奇数存放的地址和第一个奇数存放的地址无关, 但第二个奇数所在的地址存放在与第一个奇数相关的指针中, 后面奇数的存放也按此规律。设两个字长存放一个奇数, 两个字长存放一个指针, 如图 1-5 (b) 所示。在链式存储结构中, 如要读取第三个奇数, 只能从第一个奇数所在的地址开始, 通过第一个奇数关联的指针得到第二个奇数存放的地址 105, 再找到第三个奇数存放的地址 400, 才能读出第三个奇数 5。从以上的例子和分析可数据的逻辑结构、相应的存储结构以及完成数据操作的算

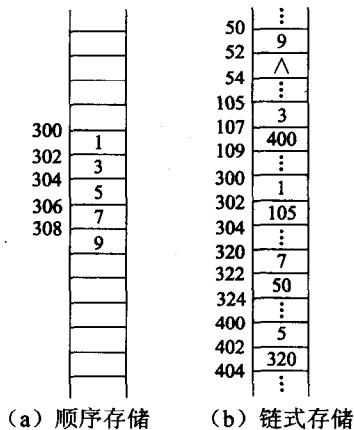


图 1-5 两种存储结构的示意图

数据类型 (Data Type) 和数据结构密切相关的一个概念,在用高级程序设计语言编写的程序中,每个变量、常量或表达式都对应一个确定的数据类型。数据类型可分为两类:一类是非结构的原子类型,如 C 语言中的基本类型(整型、实型、字符型等)、指针类型和空类型;另一类是结构类型,它可以由多个结构类型组成,并可以分解。结构类型的分量可以是非结构的,也可以是结构的。例如数组的值由若干分量组成,每个分量可以是整数,也可以是数组等结构类型。

本书在讨论各种数据结构时,针对其逻辑结构和具体的存储结构给出对应的数据类型,进一步在确定的数据类型上实现各种操作。

算法 (Algorithm) 指解决特定问题的一种方法或一种描述。

1968 年,美国唐·欧·克努特教授开创了数据结构的最初体系,他所著的《计算机程序设计技巧》第一卷《基本算法》是第一部较系统地阐述数据的逻辑结构、存储结构及算法的著作。20 世纪 70 年代初,大型程序出现,软件业飞速发展,结构化程序设计成为程序设计方法学的主要内容,人们越来越重视数据结构,认为程序设计的实质是确定数据的结构,加上设计一个好的算法,也就是人们所说的“程序=数据结构+算法”。

1.3 算法描述

程序设计人员需要对程序处理的问题准确理解,只有准确理解问题后才能研究出解决问题的方法。算法是指解决问题的一种方法或过程描述。如果将问题看作函数,那么算法能把输入转化为输出。解决一个问题可以有多种算法,但一个给定的算法只能解决一个特定的问题。例如对一组数据排序,可给出五种甚至更多种的排序算法。可以用多种算法求解问题的优点在于:根据问题的具体限定条件,可以选用合适的算法求解。例如,有的排序算法适合于元素个数少的序列,有的算法适合于元素个数多的序列,有的算法则适合于定长数值型数据的排序。计算机程序就是用某种程序设计语言去具体地实现的一个算法。本书中主要介绍各种算法,并给出一部分算法对应的C语言程序。当然使用其他的程序设计语言也可以实现算法。综上所述,问题、算法、程序是三个互相关联的不同概念。

1.3.1 算法的重要特性

正确性 它必须解决具体的问题,完成所期望的功能,给出正确的输出,这个特性是后面进行算法分析的前提。

确定性 算法执行的每一步和下一步必须确定,不能有二义性。

有限性 一个算法必须由有限步组成。无限步组成的算法无法用计算机程序来实现。因此算法必须可以终止,不能进入死循环。

输入 一个算法有零个或多个输入。

输出 一个算法有一个或多个输出。

1.3.2 数据结构的基本操作

基本操作主要有以下几种:

查找 寻找满足特定条件的数据元素所在的位置。

读取 读出指定位置上数据元素的内容。

插入 在指定位置上添加新的数据元素。

删除 删去指定位置上的数据元素。

更新 修改某个数据元素的值。

根据操作的结果可将操作分为以下两种基本类型:

加工型操作 其操作改变了原逻辑结构的“值”,如数据元素的个数,某数据元素的内容等(一般不考虑改变逻辑结构的类型)。上面基本操作中的后三种操作均为加工型操作。

引用型操作 其操作不改变原逻辑结构的“值”,只是查找或读取。

1.3.3 算法的描述方法

算法的描述方法有很多种,根据描述算法语言的不同,可将算法分为以下四种:

(1) **框图算法描述** 这种描述方法在算法研究的早期曾流行过。它的优点是直观、易懂,但用来描述比较复杂的算法就显得不够方便,也不够清晰简洁。

(2) **非形式算法描述** 用中文语言,同时还使用一些程序设计语言中的语句来描述算法,称为非形式算法描述。

(3) 类语言算法描述 类语言算法又称为伪语言算法。这种算法不能直接在计算机上运行,但专业设计人员经常使用类语言来描述算法,它容易编写、阅读和统一格式。很多语言都有其类语言,比如 C 语言,就有类 C 语言。

(4) 计算机语言编写的程序或函数 这是可在计算机上运行并获得结果的算法,使给定问题能在有限的时间内被解决,通常这种算法也称为程序。

下面以求两个整数 m 和 n ($m \geq n$) 的最大公因子为例来看看不同的算法描述的方法。框图描述如图 1-6 所示。

非形式算法描述:

(1) [求余数] m 除以 n , 并令 r 为余数 ($0 \leq r < n$)。

(2) [余数是零否] 若 $r=0$, 则结束算法, n 就是最大公因子。

(3) [替换并返回 (1)] 若 $r \neq 0$, 则 $m \leftarrow n, n \leftarrow r$ 返回 (1)。

C 语言函数描述:

```
int max_common_factor(int m,int n)
{
    int r;
    r=m%n;
    while(r!=0)
    {
        m=n;
        n=r;
        r=m%n;
    }
    return n;
}
```

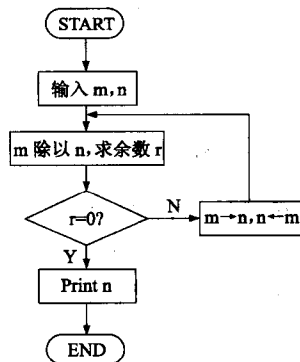


图 1-6 求解最大公因子的框图表示

本书主要介绍算法的思路和实现过程,并且尽可能地将算法对应的 C 语言函数或程序提供给读者阅读或上机运行,以便更好地理解算法。

1.4 算法分析

对于同一个问题,可以用若干种不同的方法来解决,到底采用哪一种方法更为合适呢?这就涉及到算法的比较和选择问题,称之为算法分析,即分析一个算法的时空性能。只有满足算法设计要求的、适合解决问题的算法才是“好”的算法。算法分析主要从时间复杂度和空间复杂度两个方面进行。

1.4.1 算法设计的要求

设计一个“好”的算法应考虑以下几个方面:

易读性 算法应易于阅读和理解,以便于调试、修改和扩充。

健壮性 正确的输入能得到正确的输出是算法必须具有的特性之一。但当遇到非法输入时,算法应能做出反应或处理(如提示信息等),而不会产生不需要的或不正确的结果。

高效率 即达到所需的时空性能。一个算法的时空性能是指该算法的时间性能(时间效率)和空间性能(空间效率)。解决同一个问题,如果有多个算法,执行时间短的算法时间效率高,而存储量和辅助空间量少的算法空间效率高。这两者都和问题的规模有关。