

高等学校教材

数据结构 数据库与编程

冯博琴 刘路放 徐宏喆

西安交通大学出版社

高等学校教材

数据结构 数据库与编程

冯博琴 刘路放 徐宏喆

西安交通大学出版社

内 容 简 介

本书按照《无线电技术与信号系统》教材编审委员会电路与系统教材编审组审定的编写要求编写的，经编审组评审选定，推荐出版。

本书是一本计算机软件基础教材，介绍数据结构和数据库的主要概念和技术以及编程的某些规范，可以丰富具有一门高级程序设计语言基础的读者的软件知识，适合于非计算机类学生和工程人员阅读，内容精炼、实用，深入浅出，例题丰富。本书共分三部分，第一部分讲授几种最主要的数据结构和运算；第二部分介绍一般数据库概念和理论基础、dBASE III. 的基本概念、操作命令和应用程序的编制；第三部分概述了软件工程概念和结构化程序设计。

高等学校教材

数据结构 数据库与编程

冯博琴 刘路放 徐宏喆

责任编辑 陆 薇

*

西安交通大学出版社出版

(邮政编码 710049)

西安电子科技大学出版社印刷厂印装

陕西省新华书店经销

*

开本787×1092 1/32 印张12.125 字数：254千字

1990年12月第1版 1990年12月第1次印刷

印数：1—3000

ISBN 7-5605-0407-8/TP·44 定价：2.40元

出版说明

根据国务院关于高等学校教材工作分工的规定，我部承担了全国高等学校工科电子类专业教材的编审、出版的组织工作。由于各有关院校及参与编审工作的广大教师共同努力，有关出版社的紧密配合，从1978年至1985年，已编审、出版了两轮教材，正在陆续供给高等学校和中等专业学校教学使用。

为了使工科电子类专业教材能更好地适应“三个面向”的需要，贯彻“努力提高教材质量，逐步实现教材多样化，增加不同品种、不同层次、不同学术观点、不同风格、不同改革试验的教材”的精神，我部所属的六个高等学校教材编审委员会和两个中等专业学校教材编审委员会，在总结前两轮教材工作的基础上，结合教育形势的发展和教学改革的需要制订了1986—1990年的“七五”（第三轮）教材编审出版规划。列入规划的教材、实验教材、教学参考书等共400余种选题。这批教材的评选推荐和编写工作由各编委会直接组织进行。

这批教材的书稿，是从通过教学实践、师生反映较好的讲义中经院校推荐，由编审委员会（小组）评选择优产生出来的。广大编审者、各编审委员会和有关出版社为保证教材的出版和提高教材的质量，作出了不懈的努力。

限于水平和经验，这批教材的编审、出版工作还会有缺点和不足之处，希望使用教材的单位，广大教师和同学积极

提出批评建议，共同为不断提高工科电子类专业教材的质量而努力。

机械电子工业部电子类教材办公室

前 言

本书是为非计算机类专业的高年级学生和研究生编写的，旨在用较少的学时介绍软件的基本内容：数据结构、数据库和编程技术。

本书包括三部分，第一部分是数据结构，系统地讲解计算机中数据的基本结构和实现方法，把原有的变量、数组的概念进一步扩大到栈、队列、链表、记录和文件，且对加在这些结构上的运算也给于充分的注意；第二部分是数据库，它以关系数据库为重点，介绍了数据库的概念，以目前最为常用的数据库 dBASE III+ 为例，介绍了数学基础、设计理论和应用数据库系统的设计；第三部分简单介绍软件工程的一些概念以及结构化程序设计方法。

使用本书的读者仅需熟悉一种程序设计语言。在编写中，考虑到如何以较少的篇幅来介绍这些内容，这里简单介绍一下我们的思路和处理手法，希望得到同行的指正，同时也作为使用本书的参考。

讲授第一部分数据结构的困难是读者不易入门，因为他们以前只知道简单变量、数组等少数几种数据类型，对诸如栈、队列、记录、文件、链和树等等感到抽象、生疏，今后又未必有机会学习象编译原理、操作系统等软件课程来扩充这方面的概念，因此本书特别注意引出各种数据结构的背景、用法，以及运算的举例、相互的比较。内容的选取上，主要

讲线性表、链表、树和文件，而且对查找、排序两种运算也有较多的叙述。第二部分数据库的处理颇费脑筋，因为概念多、语句和命令繁杂，要在25~30学时内作全面介绍有困难。由于本部分的要求是以关系数据库dBASE III+为例，掌握数据库的基本概念、理论、操作和设计，因此在第七章介绍了一般数据库的基础概念，如实例、数据、数据模型、数据联系和数据类型等。在第八章介绍关系数据库的数学基础：关系定义、运算、投影、选择和联结，还概述了关系数据库设计理论(关系的规范化)，后面三章详细介绍dBASE III+的基本概念、操作、命令和应用程序的编制。为了避免陷入dBASE III+的细节中去，正文仅对dBASE III+的一般原理、主要命令和功能做了说明，对于大量的开关命令和细节只作了一般性介绍。附录中提供了一切细节，以便查阅。第十二章简单介绍了另外二种数据库类型：网络型和层次数据库的实例。第三部分编程的内容十分广泛，怎样编好程序，如何测试程序的正确与否，评价质量的高低等，都是读者关心的问题，限于篇幅，这里只选了两个基本内容。其一是关于软件工程的常识，使读者了解软件的生命期，区分软件与程序，并建立起正确的软件质量观点。其二是由于编写程序容易陷入细节，忽视程序的整体，重视个别语句含义，忽视数据结构和程序结构。这里介绍了结构化程序设计，使读者在设计和编程时有规范可循。

考虑到读者的特点，在介绍概念和运算时，尽量给出例子。在第九、十、十一章中，每一章都有一节“设计一个数据库系统”，分别介绍了用dBASE III+编写的小型图书情报管理系统的几个数据库的建库、规范化的基本操作和应用表

格，以及建立管理系统的步骤。

这类课程进行上机实习是至关重要的，不动手做一做是不行的。因此每章都配了习题，大部分可以上机练习。

本书是我们近年来为非计算机类学生讲授《数据结构》、《dBASE III》、《软件开发原理及方法》等课程的讲稿修改而成。参加编写的人员有徐宏喆(一、二、三、五章)、刘路放(四章、六至十二章)、冯博琴(一至六章、十三章)，由冯博琴统编全稿。教材编审组的责任编委邹理和教授给作者许多启示和帮助，主审郝克刚教授对本书提出许多宝贵意见，在此谨表示诚挚的感谢。由于编者水平有限，书中难免存在一些缺点和错误，殷切希望广大读者批评指正。

编 者

1988. 9

目 录

第一部分 数据结构

第一章 绪论

§ 1—1 数据与数据结构	(1)
§ 1—2 算法使用的语言	(8)
习题	(13)

第二章 线性表

§ 2—1 线性表及其存贮结构	(15)
§ 2—2 栈和队列	(21)
§ 2—3 数组	(34)
习题	(42)

第三章 链表和串

§ 3—1 链表	(47)
§ 3—2 串	(70)
习题	(75)

第四章 树、二叉树与图

§ 4—1 树的基本概念	(77)
§ 4—2 二叉树	(83)
§ 4—3 遍历二叉树	(88)
§ 4—4 二叉排序树	(93)
习题	(99)

第五章 查找与排序

§ 5—1 顺序查找	(101)
§ 5—2 有序表查找	(103)
§ 5—3 二叉排序树查找	(105)
§ 5—4 Hash技术	(107)
§ 5—5 排序的思想	(112)
§ 5—6 插入排序	(114)
§ 5—7 交换排序	(116)
§ 5—8 二路合并排序	(118)
习题	(123)

第六章 文件

§ 6—1 文件概念及结构	(126)
§ 6—2 顺序文件	(129)
§ 6—3 索引	(133)
§ 6—4 外部排序	(138)
习题	(141)

第二部分 数据库

第七章 数据库系统概述

§ 7—1 引言	(143)
§ 7—2 什么是数据库	(144)
§ 7—3 数据模型	(148)
§ 7—4 数据库的构成	(155)
习题	(158)

第八章 关系数据库及其数学基础

§ 8—1 关系及其基本术语	(159)
§ 8—2 关系运算	(165)
§ 8—3 关系模型的数学定义	(172)
§ 8—4 关系代数	(177)
§ 8—5 关系的规范化理论	(181)
习题	(188)

第九章 dBASE III+ 系统总论

§ 9—1 dBASE III+	(191)
§ 9—2 dBASE III+的基本元素	(192)
§ 9—3 数据库的基本操作	(200)
§ 9—4 选择、投影和范围	(209)
§ 9—5 排序和索引	(212)
§ 9—6 多个工作区	(221)
§ 9—7 数据库结构操作命令	(232)
§ 9—8 设计一个数据库系统(一)	(236)
习题	(246)

第十章 向数据库索取更多信息

§ 10—1 列出漂亮的表格	(248)
§ 10—2 计数、求和、平均与汇总	(253)
§ 10—3 宏代换函数与内存变量文件	(258)
§ 10—4 格式文件	(263)
§ 10—5 定义和输出报表及标签	(270)
§ 10—6 怎样和其它程序设计语言交流数据	(274)
§ 10—7 设计一个数据库系统(二)	(278)
习题	(282)

第十一章 应用程序设计

§ 11—1 建立应用程序.....	(284)
§ 11—2 程序控制结构.....	(286)
§ 11—3 过程.....	(291)
§ 11—4 应用程序的编写.....	(293)
§ 11—5 设计一个数据库系统(三).....	(295)
习题	(303)

第十二章 层次和网络数据库管理系统

§ 12—1 DBTG 系统	(304)
§ 12—2 IMS 系统	(321)
习题	(327)

第三部分 软件开发

第十三章 软件开发技术

§ 13—1 软件危机、软件生命期及软件质量...	(329)
§ 13—2 非结构化程序设计.....	(336)
§ 13—3 结构程序设计(SP)	(339)
习题	(352)

附录A dBASE III+命令一览	(353)
--------------------------	-------

附录B dBASE III+函数一览	(370)
--------------------------	-------

附录C dBASE III+全屏幕操作控制键一览	(374)
--------------------------------	-------

参考文献	(375)
------------	-------

第一部分 数据结构

第一章 绪 论

本章有两个内容，一是说明数据和数据结构等一些基本概念，以及数据结构与计算机科学发展的关系，二是简单介绍本书第一部分描述算法用的一种语言。

§1-1 数据与数据结构

计算机的出现，首先是作为一种快速计算工具，用以解算人类无法进行的复杂而耗时的“工程”计算，比如导弹飞行轨道、核反应、桥梁应力等等，这里所涉及的数据主要是实型数据。近 20 年来，计算机的软硬件技术获得了迅速发展。由于计算机在存贮量、速度、配套的外部设备（如终端、打印机、绘图仪等）、软件 and 价格等方面都很优越，使它得到了广泛的应用。近年来，计算机作为一种“纯计算”工具的形象已被所谓的非数值计算所取代。在日常工作和生活中，可以看到计算机在代替人订飞机票、旅馆管理、发工资、检索情报资料 and 进行人口普查等等。的确，现在在大量的计算机中所运行的是称为航空订票系统、信息处理系统、财务系统等等。

据统计,这种非数值计算占计算机总工作量的 90%。在这些系统中,所处理的数据类型及在数据上所进行的运算都比原来的实数及其运算丰富和复杂。计算机科学的发展正是与计算机处理数据的能力的不断提高分不开的,这也是促使人们对数据进行研究的原因。

什么是数据呢?可以说是指计算机能够存取和处理的信息,这些信息在某种意义上就代表了现实世界中客观存在的事物,因此它是对某物的抽象。

以旅馆管理系统和刑事侦破系统为例,前者对于每一个旅客的姓名、年龄、性别、单位、籍贯、从何处来及到何处去感兴趣,而后者对于人的高度、身材、体型、发色、口音、面部特征及衣着等十分关心,因此,每个应用系统仅抽取与解决本题有关的特性,一般是许多项数据。顺便指出,这些数据可能是一些数字,但口音、面部特征和衣着等等,必须利用其他类型的数据来描述。

术语“数据类型”在这里常要提到,它是指在一种程序设计语言中,变量所具有的数据种类。如 FORTRAN 语言中,有整型、实型、双精度型、复型和布尔型;在 PASCAL 语言中,不仅有整型、实型和布尔型,还有字符型、记录型、指针型和文件型等等。与数据类型紧密联系的又一个术语叫“数据对象”,是指某种数据类型的元素集合。比如整数类型的数据对象是 $\{0, \pm 1, \pm 2, \pm 3, \dots\}$,布尔型的数据对象是 $\{\text{TRUE}, \text{FALSE}\}$,字符类型的数据对象是 $\{A, B, \dots, Z, +, -, *, /, \dots\}$ 。一种程序设计语言所具备的数据类型基本上决定了它所处理问题的复杂程度,如果某种语言的数据类型很贫乏,那么就无法描述某些数据对象,或者说描述得十分别

扭、费劲，因此它就只能适应一个很窄的面。

什么是“数据结构”呢？一般认为是指数据的逻辑结构、数据的存储结构以及数据的运算。下面用例子说明上述概念。

众所周知，进入计算机的数据不仅要求它能代表客观事物的某些特性，而且要经过一定的组织，要精心地选择适当的数据结构来描述它。

例 1-1 图 1-1 是一张价格表，用什么数据结构描述它更为合适要视问题而定。可以把货物名这一列用线性表 A 来表示

$$A : \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix}$$

货物名	价格
A_1	C_1
A_2	C_2
$\vdots$	$\vdots$
A_n	C_n

图 1-1 货物价格表

同样用线性表 C 表示价格

$$C : \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{pmatrix}$$

可以把每一对 (A_i, C_i) 视为一个元素，构成线性表 AC

$$AC : \begin{pmatrix} (A_1, C_1) \\ (A_2, C_2) \\ \vdots \\ (A_n, C_n) \end{pmatrix}$$

线性表 AC 的逻辑结构是指元素 (A_i, C_i) 的序列, 线性表的存储结构是指这 n 个元素在计算机里是怎么存放的, 是顺序存储还是东一个、西一个地靠指针把它们链接起来? 线性表 AC 上的运算是指删去 (A_i, C_i) , 增加 (A_j, C_j) , 以及给定一个货物名, 要查询它的价格, 或修改价格, 或删除具有该货物名的这一对 (A_i, C_i) 等操作。

在不同的数据类型上, 数据的逻辑结构, 存储结构和数据运算不一定相同。

例 1-2 图 1-2 是某大学的行政单位结构表, 我们用树型结构来描述, 如图 1-3 所示:

树型结构用来表示有层次关系的数据结构时特别直观、有效, 第四章要专门研究它。为便于叙述, 图 1-4 给出一般树的有关术语。树型结构的逻辑结构是指这棵树的根结点、叶子、内部结点之间的关系及树的层次和高度。树型结构比线性表更易刻划复杂的事物, 因而它能反映事物的更多侧面, 反映更丰富的特性。树型结构的存储结构描述了每一个结点是顺序存放还是采用链接方式存放, 每一个结点(包括树根和叶子, 下同)的存储格式、结点之间的复杂关系(如父子、兄弟关系)等等。树的运算是指如何查找和增、删结点。不同的数据结构对问题的回答不同。

综上所述, 数据结构是研究在确定的结构中各数据元素之间的逻辑关系, 在计算机内的存储方式, 以及对它允许进行哪些运算。

数据结构的研究工作对于计算机科学发展有着举足轻重作用。学术界有一种颇有影响的观点, 认为计算机科学中具有核心作用的概念是“信息结构的转换”, 计算机科学就是“一

院处、系		科、室	教学小组
研究生院	培养处	硕士生科, 博士生科, 代培科	:
	管理处	学籍科	:
	招生处	:	:
	院办	:	:
管理学院	经济系	:	:
		:	:
		:	:
	企管系	:	:
数学系		代数教研室	:
		几何教研室	:
		微分方程教研室	:
		计算方法教研室	:
计算机系		软件教研室	操作系统组
			网络组
			数据库组
			语言组
		计算机系统教研室	:
		微机教研室	:
		基础及应用教研室	:
教务处		学籍管理科	:
		教研科	:
		生产实习科	:
		成人教育科	:
	:	:	:

图 1-2 学校行政结构表