



软件 技术 基础教程

范 锐 主编



石油工业出版社

软件技术基础教程

范 锐 主编

石油工业出版社

内 容 提 要

本书以计算机软件应用基础为主线,简明扼要地讲解了计算机软件中的一些重要概念、基本知识和方法,内容包括数据结构、操作系统、软件工程、数据库技术和网络技术五个领域的基础知识和应用技术。全书分为基础篇和提高篇,基础篇对应计算机等级考试(二级)的软件基础知识,提高篇对应计算机等级考试(三级B)的软件应用技术内容。

本书是为大学非计算机专业的学生学习计算机软件技术基础知识编写的。它符合国家教委高教司颁发的“工科非计算机专业计算机基础教学指南”中有关软件技术基础课程的教学要求,基本覆盖了计算机等级考试(二级)和(三级B)中有关软件技术的考核内容。本书可作为非计算机专业各层次学生的教材,也可作为计算机应用软件人员的培训教材或工程技术人员的参考教材。

图书在版编目(CIP)数据

软件技术基础教程/范锐 主编.

北京:石油工业出版社,2001.9

ISBN 7-5021-3454-9

I. 软…

II. 范…

III. 软件-教材

IV. TP31

中国版本图书馆 CIP 数据核字(2001)第 045285 号

石油工业出版社出版

(100011 北京安定门外安华里二区一号楼)

西南石油学院印刷厂排版印刷

新华书店北京发行所发行

*

787×1092 毫米 16 开本 15.5 印张 397 千字 印 1—1500

2001 年 9 月北京第 1 版 2001 年 9 月四川第 1 次印刷

ISBN 7-5021-3454-9/TP·64

定价:25.00 元

前 言

计算机应用已渗透到各个领域。作为计算机应用核心的软件技术,是每一个要把计算机很好地应用到自身专业领域的人必须学习并掌握的基本技术。

本教材是为大学非计算机专业学生学习计算机软件基础技术知识而编写的。它符合国家教委高教司颁发的“工科非计算机专业计算机基础教学指南”中有关软件技术基础课程的教学要求,也基本覆盖了计算机等级考试(二级)和(三级B)中有关软件技术的考核内容。

本教材以计算机软件应用基础为主线,简明扼要地讲解了计算机软件中的一些重要概念、基本知识和方法,内容包括数据结构、操作系统、软件工程、数据库技术和网络技术五个领域的基础知识和应用技术。全书分为基础篇和提高篇,基础篇对应计算机等级考试(二级)的软件基础知识,提高篇对应计算机等级考试(三级B)的软件应用技术内容。我们的目的是要读者通过本教材的学习,基本掌握计算机软件技术的概念、知识和方法,今后会利用所学的软件知识与技术去解决所从事的专业领域的计算机应用问题。

本教材由姜瑞京编写第一章和第六章,胡卫东编写第二章和第七章,范锐编写第三章和第八章,黄诚编写第四章和第九章,李淮编写第五章和第十章。全书由范锐统稿并整理,全书由黎明教授主审。

本教材的知识点较多,信息量较大。各章节内容可根据不同专业、不同层次读者的具体情况酌情取舍或安排自学。建议基础篇学时数为48学时,提高篇另增加32学时。

本教材在已连续使用三届的院内自编教材基础上,作了较大的改动,许多章节全部重写,并增加了计算机等级考试(三级B)的绝大部分软件技术内容。其目的在于使有志从事计算机应用软件开发开发的读者,通过深入的学习,掌握更多的软件技术和知识,更好地把计算机应用到自己从事的专业领域中。

本教材受到西南石油学院教材委员会的出版资助,在此,我们表示衷心的感谢。并对一直关心和支持本教材写作和出版的各位同事表示由衷的谢意。

由于编者水平有限,加之时间匆忙,本书难免存在缺点和错误,敬请广大读者批评指正。

编 者

2001年5月于西南石油学院

目 录

基 础 篇

1 数据结构基础	3
1.1 绪 论	3
1.1.1 什么是数据结构	3
1.1.2 数据结构中的基本概念	4
1.2 线性结构	8
1.2.1 线性表	8
1.2.2 栈与队列	18
1.2.3 串	21
1.3 非线性结构	23
1.3.1 树	23
1.3.2 二叉树	24
1.4 查找与排序	30
1.4.1 查找	30
1.4.2 排序	32
习题一	35
2 操作系统基础	37
2.1 绪论	37
2.1.1 操作系统的概念(Operating System)	37
2.1.2 常见的操作系统	37
2.1.3 操作系统的功能	37
2.2 进程管理基础	38
2.2.1 进程的基本概念	38
2.2.2 进程状态	38
2.2.3 进程管理	39
2.2.4 进程间的关系	39
2.2.5 进程调度概述	39
2.2.6 进程通信简介	40
2.2.7 死锁的基础知识	40
2.3 作业管理基础	41
2.3.1 作业的基本概念	41
2.3.2 作业的基本状态	42
2.3.3 作业调度算法	42
2.4 存储管理基础	43
2.4.1 存储管理的功能	43

2.4.2 单一连续区分配概述	45
2.4.3 分区存储管理方案概述	45
2.4.4 分区管理的扩充	45
2.4.5 虚拟存储器概念	46
2.4.6 分页存储管理概述	46
2.4.7 段式存储管理方案概述	47
2.4.8 段页式存储管理方案概述	47
2.5 设备管理基础	47
2.5.1 设备管理概述	47
2.5.2 数据传送方式	48
2.5.3 缓冲技术的概念	49
2.5.4 设备分配的内容	49
2.6 文件管理基础	49
2.6.1 文件系统	50
2.6.2 文件的组织	50
2.6.3 文件目录	50
2.6.4 文件的共享及存取控制	51
2.6.5 辅存空间管理	51
2.6.6 文件的使用	51
习题二	51
3 软件工程基础	53
3.1 概论	53
3.1.1 软件危机	53
3.1.2 软件工程及软件工程学	53
3.2 软件生存周期	54
3.2.1 软件定义	54
3.2.2 软件生存周期	55
3.3 需求分析	56
3.3.1 需求分析的任务	56
3.3.2 结构分析方法的步骤	56
3.3.3 结构分析方法的基本工具	57
3.4 软件设计	58
3.4.1 软件设计的任务	58
3.4.2 模块独立性准则	59
3.4.3 软件总体设计方法	60
3.4.4 详细设计方法	60
3.5 软件编程	61
3.5.1 编程风格	61
3.5.2 程序设计语言	63
3.6 软件测试	65

3.6.1 软件测试的目的	65
3.6.2 常用的测试方法	65
3.7 软件维护	66
3.7.1 软件维护的必要性	66
3.7.2 软件维护的副作用	67
习题三	67
4 数据库基础知识	68
4.1 数据库管理概述	68
4.1.1 数据和信息处理	68
4.1.2 数据管理技术的发展	69
4.1.3 数据库系统	72
4.1.4 数据库的体系结构	74
4.2 数据模型	76
4.2.1 数据和数据联系的描述	76
4.2.2 数据模型	77
4.3 数据库设计基础	83
4.3.1 数据库设计的内容和方法	83
4.3.2 数据库设计步骤	84
习题四	86
5 计算机网络基础	87
5.1 计算机网络的形成和发展	87
5.1.1 计算机网络的产生	87
5.1.2 计算机网络的发展	88
5.2 计算机网络的定义与分类	88
5.2.1 计算机网络的定义	88
5.2.2 计算机网络的组成	89
5.2.3 计算机网络的分类	89
5.2.4 计算机网络拓扑结构	90
5.3 数据交换技术	91
5.3.1 线路交换	91
5.3.2 报文交换	92
5.3.3 分组交换	92
5.3.4 ATM 交换	93
5.4 网络体系结构	93
5.4.1 协议的分层结构	94
5.4.2 ISO/OSI 参考模型	94
5.5 TCP/IP 参考模型与协议	97
5.5.1 TCP/IP 的基本概念	97
5.5.2 TCP/IP 协议的分层结构	98
5.6 局域网技术	99

5.6.1 局域网参考模型与协议	100
5.6.2 随机接入:ALOHA	101
5.6.3 IEEE802.3 标准:CSMA/CD	102
5.6.4 IEEE802.5 标准:Token Ring	103
5.6.5 IEEE802.4 标准:Token Bus	104
5.6.6 交换局域网	104
5.7 Internet 技术	106
5.7.1 IP 地址	106
5.7.2 子网和子网掩码	107
5.7.3 域名与域名系统	109
5.7.4 Internet 的基本服务功能	111
5.7.5 WWW 服务	112
5.7.6 Internet 用户接入方式	113
习题五	113

提 高 篇

6 数据结构	117
6.1 图	117
6.1.1 图的基本概念	117
6.1.2 图的存储结构	119
6.1.3 图的遍历	121
6.1.4 最小生成树	122
6.1.5 AOV 网与拓扑排序	124
6.2 散列法	126
6.2.1 散列法	126
6.2.2 构造散列函数	127
6.2.3 解决冲突的方法	127
6.3 快速排序	129
习题六	130
7 操作系统	132
7.1 进程管理	132
7.1.1 进程的概念与特征	132
7.1.2 进程控制块及进程的状态	133
7.1.3 进程控制	134
7.1.4 进程互斥与同步	135
7.1.5 进程调度(Processes Scheduling)	137
7.1.6 进程通信	138
7.1.7 死锁	138
7.2 作业管理	139

7.2.1 用户和操作系统之间的接口	139
7.2.2 作业的概念、组成	140
7.2.3 作业的建立过程	140
7.2.4 作业的分类和控制	141
7.2.5 作业的状态及其转换	141
7.2.6 作业调度	142
7.3 存储管理	143
7.3.1 单一连续区分配	144
7.3.2 分区存储管理	144
7.3.3 覆盖与交换技术	146
7.3.4 虚拟存储器	146
7.3.5 分页存储管理	147
7.3.6 段式存储管理	151
7.3.7 段页式存储管理	153
7.4 设备管理	153
7.4.1 数据传送方式	154
7.4.2 缓冲技术	155
7.4.3 设备的绝对号、类型号、相对号和符号名	156
7.4.4 设备分配	156
7.4.5 虚拟设备管理与 SPOOLING 技术	158
7.5 文件管理	158
7.5.1 文件的组织	158
7.5.2 文件目录	160
7.5.3 文件的保护和保密	161
7.5.4 文件存储空间管理	162
7.6 UNIX 操作系统	162
7.6.1 UNIX 操作系统的特点	162
7.6.2 UNIX 系统的结构	163
7.6.3 UNIX 的文件系统	163
7.6.4 UNIX 系统的使用	164
习题七	166
8 软件工程	168
8.1 软件开发概述	168
8.1.1 软件的特点	168
8.1.2 软件开发的基本原理和目标	168
8.2 项目计划与可行性研究	169
8.2.1 项目计划与系统目标	169
8.2.2 可行性研究	170
8.3 需求分析方法	171
8.3.1 结构化方法的特点	171

8.3.2 数据流分析设计	172
8.3.3 数据字典设计	173
8.3.4 加工说明的几种形式	175
8.4 软件设计	177
8.4.1 设计的基本原则	177
8.4.2 变换型分析设计方法	177
8.4.3 事务型分析设计方法	179
8.4.4 详细设计表示法	180
8.5 软件测试与软件质量	182
8.5.1 软件测试步骤与原则	182
8.5.2 白盒法	183
8.5.3 黑盒法	184
8.5.4 软件调试	185
8.5.5 软件质量保证	185
8.6 软件维护与软件文档	186
8.6.1 影响软件维护的因素	186
8.6.2 软件维护过程	187
8.6.3 软件版本控制	187
8.6.4 软件工程标准	187
习题八	189
9 关系数据库系统	191
9.1 关系数据模型	191
9.1.1 关系模型的数据结构	191
9.1.2 关系数据库的数据完整性	192
9.1.3 关系运算	193
9.2 关系数据库设计理论	194
9.2.1 数据库模式的设计	194
9.2.2 函数依赖	195
9.2.3 关系模式的规范化理论简述	196
9.3 结构化查询语言——SQL	196
9.3.1 SQL 数据库的体系结构	196
9.3.2 SQL 数据定义	197
9.3.3 SQL 数据查询	199
9.3.4 SQL 数据操纵	206
9.3.5 SQL 数据控制	207
9.3.6 嵌入式 SQL	207
习题九	208
10 数据通信与计算机网络	210
10.1 数据通信基础	210
10.1.1 数据通信的基本概念	210

10.1.2 数据通信的基本方式	211
10.1.3 多路复用技术	213
10.1.4 数据编码技术	214
10.1.5 差错控制方法	218
10.2 网络操作系统	219
10.2.1 局域网操作系统的发展与分类	220
10.2.2 局域网操作系统的基本服务功能	221
10.2.3 典型局域网操作系统:Windows NT Server	222
10.2.4 典型局域网操作系统:NetWare	224
10.2.5 UNIX 操作系统的网络功能	225
10.3 网络互联技术	226
10.3.1 网络互联的层次	226
10.3.2 中继器和集线器	226
10.3.3 网桥	227
10.3.4 路由器	230
10.3.5 网关	232
习题十	233

基 础 篇

1 数据结构基础

1.1 绪 论

要点:

- (1) 数据、结构和数据结构。
- (2) 数据结构中的基本概念。

随着计算机科学和硬件技术的飞速发展,计算机的应用已从最初的科学计算逐步发展到人类活动的各个领域。为此,计算机处理的对象已不只是简单的数值和字符,而是带有不同结构的各种数据。而要设计出好的软件,只掌握所用的计算机语言还不够,还应该研究各种数据的特性和数据之间存在的关系。“数据结构”这门学科就是在这样的背景下形成和发展起来的。

1.1.1 什么是数据结构

数据通常指所有能输入到计算机中,并能被计算机程序处理的符号(信息)。结构指数据之间彼此存在的相互关系。数据结构即指数据之间抽象化的相互关系(不涉及数据元素的具体内容,数据元素是数据的基本单位,元素或结点是数据元素在机器中的映象,但有时也把数据元素称元素,把图示法中的数据元素称为结点),而这种关系体现在三个方面:

- (1) 数据之间的逻辑关系(即数据的逻辑结构);
- (2) 数据在计算机内的存储方式(即数据的物理结构或存储结构);
- (3) 数据在计算机内的运算(即数据在某种存储结构上进行的操作)。

例如处理学生成绩表如图 1.1。

学号	姓名	C 语言	英语	数学	物理
001	赵甲	98	90	88	89
002	钱乙	65	71	84	76
003	孙丙	80	77	82	82
004	李丁	78	90	76	73
⋮	⋮	⋮	⋮	⋮	⋮

图 1.1 学生成绩表

在这张学生成绩表中的每一行为一个数据元素，由学号、姓名、各科成绩等数据项（数据项是数据不可分割的最小单位）组成。表中的第一个元素（第一个学生信息）没有前驱元素，最后一个元素（最后一个学生信息）没有后继元素，中间任何一个元素（中间学生信息）有且只有一个前驱和有且只有一个后继。即这张表呈现出一种线性结构关系（这种线性结构关系即为学生成绩表的逻辑结构）；把学生成绩表按元素存放在计算机内有一种存储方法（这种存储方法即为学生成绩表的存储结构）；用程序显示存储内容及对表中元素（某个学生信息）进行查找、插入和删除等算法操作（这种操作即为对学生成绩表进行的运算）。以上三方面的描述，就是这一学生成绩表的数据结构。

1.1.2 数据结构中的基本概念

1.1.2.1 数据结构的内容

上面提到数据结构涉及到：逻辑结构、存储结构和运算，下面看看它们的基本含义。

(1) 数据的逻辑结构。

这种结构指的是数据间的关系，可用二元组表示如下：

$$B = (K, R), K = \{k_i \mid 1 \leq i \leq n, n \geq 0\}, R = \{r_j \mid 1 \leq j \leq m, m \geq 1\}.$$

其中， B 是一种数据结构， K 表示数据元素的有穷集合， R 表示各数据元素的关系集合， k_i 表示第 i 个数据元素， r_j 表示第 j 个关系。

数据结构还能够利用图形形象地表示出来，图形中的每个结点对应一个数据元素，连线表示数据元素间的关系。

一种数据结构 $linearity = (K, R)$ ，其中

$$K = \{k_1, k_2, k_3, k_4, k_5\},$$

$$R = \{r\},$$

$$r = \{ \langle k_1, k_2 \rangle, \langle k_2, k_3 \rangle, \langle k_3, k_4 \rangle, \langle k_4, k_5 \rangle \}.$$

数据结构所对应的图示法如图 1.2 所示。

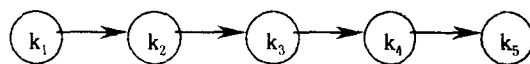


图 1.2 线性结构图示示例

1) 结点的类型。

结点的类型分为两类：初等类型和组合类型。每个结点的基本表示是由 n 个数据项表示的，则

$$k_i = (w_{1i}, w_{2i}, \dots, w_{ni}) \geq 1$$

一般情况： k_i 是初等类型时， $n=1$ （即只有一个数据项）； k_i 是组合类型时， $n>1$ （即有多个数据项）。

2) 结构的分类。

结构的类型也分为两类：线性结构和非线性结构。

线性结构只有一个终端和一个开始结点，并且所有结点都最多只有一个前驱和后继。这种数据结构的特点是数据元素之间的 1:1 联系。图 1.2 是一个典型的线性结构。

非线性结构指不是线性结构的结构类型。它又分有树型结构和图型结构。

树型结构除根结点外，每个结点有且只有一个前驱结点，但可以有任意多个后继结点（树叶可以看作具有零个后继结点）。这种数据结构的特点是数据元素之间的1:N的联系（ $N \geq 0$ ）。

一种数据结构 $tree = (K, R)$ ，其中

$K = \{k_1, k_2, k_3, k_4, k_5\}$,

$R = \{r\}$,

$r = \{\langle k_1, k_2 \rangle, \langle k_1, k_3 \rangle, \langle k_1, k_4 \rangle, \langle k_3, k_5 \rangle\}$ 。

所对应的图示法如图 1.3 所示。

图型结构每个结点可以有任意多个前驱结点和任意多个后继结点。这种数据结构的特点是数据元素之间的M:N的联系（ $M \geq 0, N \geq 0$ ）。

一种数据结构 $graph = (K, R)$ ，其中

$K = \{k_1, k_2, k_3, k_4, k_5\}$,

$R = \{r\}$,

$r_1 = \{\langle k_1, k_2 \rangle, \langle k_1, k_3 \rangle, \langle k_2, k_4 \rangle, \langle k_4, k_3 \rangle, \langle k_4, k_5 \rangle, \langle k_3, k_5 \rangle\}$ （为有向图），

$r_2 = \{(k_1, k_2), (k_1, k_3), (k_1, k_4), (k_2, k_3), (k_2, k_4), (k_3, k_4), (k_3, k_5), (k_4, k_5)\}$ （为无向图）。

图型结构所对应的图示法如图 1.4 所示。

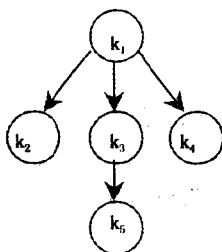
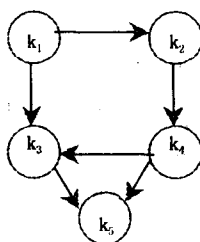
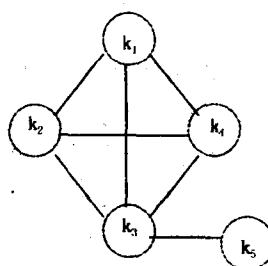


图 1.3 树型结构



(a) 有向图



(b) 无向图

图 1.4 图型结构

3) 结点和结构。

结点和结构是相对而言的。结构是由结点组成的，但是对于一个结点内部，又可以把内部数据项看成结点，本身看成结构。结点内部的组合方式可以是线性和非线性的。

4) 抽象数据类型。

抽象数据类型（ADT）是指一个数学模型以及定义在该模型上的一组操作。抽象数据类型的定义仅取决于它的一组逻辑特性，而与其在计算机内部如何表示和实现无关。即不论其内部结构如何变化，只要它的数学特性不变，都不影响其外部的使用。所定义在数据类型的抽象层次越高，含有该抽象数据类型的模块的复用程度也就越高。

抽象数据类型按其值的不同特性，可细分为三种类型：原子类型，其变量的值是不可分解的；固定聚合类型，其变量的值由确定数目的成分按某种结构组成；可变聚合类型，构成可变聚合类型“值”的成分的数目不确定。

抽象数据类型的定义与使用实现了数据的封装和隐藏，为对象程序设计方法提供了基础，也使过程语言的可维护性大大提高。

目前，在数据结构中描述结点及算法，通常是采用类 PASCAL 语言或类 C 语言来描述。

(2) 数据的存储结构。

数据的逻辑结构是一种数学表示，与计算机本身无关。而数据的存储结构是逻辑结构在计算机存储器中的存储方式，与计算机存储器有密切联系。

逻辑结构在存储器中的存储是一个映象过程。也就是把结点存储在内存中，同时存储的方式又能明显或隐含地反映出逻辑结构的各关系。

数据结构有四种基本的映象方法：

1) 顺序的方法（对应存放）。

这种方法是把逻辑上相邻的结点存储在存储器相邻的单元中。各结点之间的关系是由存储结构相邻关系来体现的。这种方法主要用于线性结构。

2) 链接的方法（用指针来实现结构特征）。

这种方法是将各结点所占的存储单元分为两部分：一部分存放结点本身的信息，称为数据项；另一部分存放结点的后继结点所对应的存储单元的地址，称为指针项。各结点之间的关系由指针项来体现。有些情况，逻辑结构较复杂，可能需要多个指针项。

3) 索引的方法（提高查找速度的方式）。

这种方法是建立索引表，在索引表里指示结点的位置。此方法常应用于结点本身是一个较大（一个单元放不下）的数据项的结构。对于文件结构较为常用。数组的下标就是一种索引方法（数组的下标索引可通过一种计算确定存储的位置）。

4) 散列的方法。

这种方法是用结点的值来确定它存储的地址，确定存储位置的方法是建立一个散列函数：

$$LOC(k) = f(w_i)$$

其中， w_i 是结点的一个关键字段。

一般数据结构的存储映象是以上四种之一，或者是它们的组合。同一个逻辑结构可以用不同的存储映象方法。使用哪一种存储方法与逻辑结构和将来要进行的运算有关。

对于不同的存储映象，其存储的空间利用率是不相同的。存储的空间利用率通常用存储密度来计量。

$$\text{存储密度} = \frac{\text{数据本身所占的存储量（不包括链指针空间）}}{\text{整个结构所占的存储量}}$$

存储密度最大为 1，密度越大空间利用率越高。顺序存储密度为 1，链接存储