

计算机 软件技术基础

夏清国 编



西北工业大学出版社

计算机软件技术基础

夏清国 编

西北工业大学出版社

【内容简介】 本书是根据工科非计算机专业计算机基础教学的教学计划和教学实践组织编写的。全书主要内容包括数据结构、软件工程、面向对象的设计方法、操作系统、数据库技术基础、信息系统、网络通信技术基础,并在每章最后配有相应的习题。

本书注重介绍计算机软件基础领域的基本理论和设计方法,内容实用、易懂,适合作为非计算机专业本、专科学生的教材,也可作为自学参考书。

图书在版编目 (CIP) 数据

计算机软件技术基础/夏清国编. —西安:西北工业大学出版社,2003.8
ISBN 7-5612-1675-0

I. 计… II. 夏… III. 软件—高等学校—教材 IV. TP31

中国版本图书馆 CIP 数据核字 (2003) 第 062164 号

出版发行:西北工业大学出版社

通信地址:西安市友谊西路 127 号 邮编:710072 电话:(029)8493844

网 址:www.nwpup.com

印 刷 者:陕西宝石兰印务有限责任公司

开 本:787 mm×1 092 mm 1/16

印 张:18.25

字 数:440 千字

版 次:2003 年 8 月第 1 版 2003 年 8 月第 1 次印刷

印 数:1~5 000 册

定 价:23.00 元

前 言

计算机软件技术基础是在学生了解计算机基础知识的基础上,为提高学生对软件本质的理解开设的一门必修课程,通过对计算机软件基础领域的基本原理、方法和思想的学习,来提高学生对软件开发工具和环境的适应能力。

目前计算机软件技术基础课程的教学内容在全国还没有统一的规定。作者在近 10 年从事计算机软件技术基础课程的教学实践过程中,经过不断总结,认为计算机软件技术基础的内容应包括数据结构、软件工程、操作系统基本原理、常见的操作系统的使用、数据库系统基础知识(其中包括有关 SQL 语言的相关知识)、信息系统基础知识、计算机网络通信技术基础以及程序设计方法、语言等。

对于非计算机专业的学生来说,开设该门课程的主旨是使其掌握计算机软件中的一些基本原理、方法、思想,提高自身的软件素质,增强其对软件的悟性和奠定一定的软件设计基础。因此在编写本书的过程中,作者根据工科非计算机专业计算机基础教学的教学计划、教学实践的基本要求及课程教学的特点,对原有计算机软件技术基础课程的内容进行了适当的取舍,并在知识的深浅程度上作了一定的调整。

本书共分为 7 章。第 1 章为数据结构,主要介绍了常见的数据结构(线性表、栈、队列、数组、树、图)以及查找、排序等算法,其中的算法是基于 C 语言实现的;第 2 章为软件工程,主要介绍生命周期法和原型法的设计思想和过程;第 3 章为面向对象的设计方法,主要介绍面向对象的基本概念、理论,并从 OOA, OOD, OOP 到面向对象的测试完整地叙述了面向对象的设计方法;第 4 章为操作系统,主要介绍操作系统的基本功能,并对常见的操作系统进行了介绍;第 5 章为数据库技术基础,主要介绍数据库系统的基本概念、关系数据库的基本理论、数据库的设计以及关系数据库语言 SQL 的使用;第 6 章为信息系统,主要介绍信息系统的基本概念,并对事务处理系统、管理信息系统、决策支持系统和办公信息系统等方面的应用进行了介绍;第 7 章为网络通信技术基础,主要介绍 TCP/IP 网络技术、客户机/服务器模型、Web 的工作原理,并通过一个事例介绍了基于 C/S 计算模式的软件设计中的算法。

本书力求概念清楚、内容新颖、简明扼要、重点突出、易于理解、实用性强。书中的错误和不妥之处，恳请读者批评指正。

编 者

2003 年 5 月

目 录

第 1 章 数据结构及其算法	1
1.1 数据结构的引入	1
1.2 数据结构的基本概念	3
1.2.1 什么是数据结构	3
1.2.2 数据结构的基本概念和术语	3
1.2.3 算法描述及分析	6
1.3 线性表	8
1.3.1 线性表的定义及运算	8
1.3.2 线性表的顺序存储结构	9
1.3.3 线性表的链式存储结构	12
1.3.4 顺序表与链表的比较	18
1.4 栈和队列	18
1.4.1 栈	18
1.4.2 队列	25
1.5 数组	35
1.5.1 数组的定义与运算	35
1.5.2 数组的顺序存储结构	36
1.5.3 几种特殊矩阵的存储方式	37
1.5.4 数组的链式存储结构	39
1.6 树与二叉树	40
1.6.1 树的基本概念	40
1.6.2 树的存储结构	42
1.6.3 二叉树的基本概念	43
1.6.4 几种特殊形式的二叉树	44
1.6.5 二叉树的存储结构	46
1.6.6 二叉树的遍历	47
1.6.7 二叉树的应用	51

1.7 图	57
1.7.1 图的基本概念	57
1.7.2 图的存储结构	59
1.7.3 图的遍历	61
1.7.4 生成树与最小生成树	64
1.7.5 最短路径	65
1.8 查找	68
1.8.1 查找的基本概念	68
1.8.2 线性查找	69
1.8.3 折半查找	69
1.8.4 分块查找	71
1.8.5 二叉排序树查找	72
1.8.6 哈希(Hash)查找	73
1.9 排序	77
1.9.1 排序的基本概念	77
1.9.2 直接插入排序	77
1.9.3 交换排序	78
1.9.4 选择排序	80
1.10 排序方法的比较和选择	83
习题一	84

第2章 软件工程技术基础 87

2.1 概述	87
2.1.1 软件危机	87
2.1.2 软件工程的基本原则	88
2.1.3 软件工程的基本目标	90
2.1.4 传统软件工程模式	90
2.1.5 现代软件工程模式	91
2.2 软件开发过程	92
2.2.1 软件开发过程模型	92
2.2.2 软件生存周期	93
2.2.3 问题定义	93
2.2.4 可行性分析	94
2.2.5 需求分析	95
2.2.6 总体设计	95
2.2.7 详细设计	97
2.2.8 软件编码	100
2.2.9 软件测试	101
2.2.10 软件维护	104

2.3 软件开发方法	105
2.3.1 结构化系统方法	105
2.3.2 原型法	111
2.4 软件工程管理	112
2.4.1 软件工程项目管理	112
2.4.2 软件工程计划管理	114
2.4.3 软件人员组织与管理	114
2.4.4 软件工程文档管理	115
习题二	116
第3章 面向对象的软件开发技术	118
3.1 面向对象方法概述	118
3.2 面向对象的基本概念	121
3.2.1 实体和对象	121
3.2.2 服务和消息	123
3.2.3 类	125
3.2.4 面向对象方法的特性	128
3.3 面向对象分析方法	131
3.3.1 确定对象	132
3.3.2 确定属性	134
3.3.3 定义操作(服务)	134
3.3.4 对象间通信	135
3.3.5 完成对象定义	135
3.4 面向对象的设计	139
3.4.1 面向对象设计概念	140
3.4.2 面向对象设计方法	143
3.5 面向对象编程及面向对象测试	144
3.5.1 面向对象编程	144
3.5.2 面向对象测试	144
3.5.3 面向对象的软件维护	145
习题三	145
第4章 操作系统	146
4.1 操作系统概述	146
4.1.1 操作系统的概念	146
4.1.2 操作系统分类	146
4.1.3 操作系统的作用	148
4.1.4 操作系统的层次模型	149
4.1.5 操作系统的特性	150

4.2 处理器管理	151
4.2.1 作业调度	151
4.2.2 进程调度	153
4.2.3 多道程序并发运行的问题	158
4.3 存储管理	160
4.3.1 存储管理的功能	160
4.3.2 单一连续区存储管理	162
4.3.3 分区存储管理	164
4.3.4 页式存储管理	168
4.3.5 段式存储管理	172
4.4 设备管理	175
4.4.1 设备管理的功能及基本概念	175
4.4.2 设备分配	177
4.4.3 设备处理	178
4.4.4 假脱机技术	178
4.5 文件管理	179
4.5.1 文件和文件系统	179
4.5.2 文件的组织	181
4.5.3 文件的存取方式及存储设备	183
4.5.4 文件目录	183
4.5.5 文件的共享与安全性	188
4.6 几种常见的操作系统	189
4.6.1 DOS 操作系统	189
4.6.2 UNIX 操作系统	192
4.6.3 Windows 操作系统	194
习题四	195

第 5 章 数据库技术基础..... 196

5.1 数据库系统基础知识	196
5.1.1 数据、信息、信息处理	196
5.1.2 数据管理技术的发展	197
5.1.3 数据库、数据库管理系统和数据库系统	197
5.1.4 数据描述	199
5.1.5 数据库组织	200
5.2 关系数据库系统	204
5.2.1 关系数据库的特点	204
5.2.2 关系数据库的结构	204
5.2.3 关系代数	205
5.2.4 关系模式的规范化	208

5.3 关系数据库设计	210
5.3.1 数据库设计	210
5.3.2 需求分析	211
5.3.3 概念设计($E-R$ 模型)	212
5.3.4 逻辑设计	213
5.3.5 物理设计	214
5.3.6 构建与维护	215
5.4 SQL	216
5.4.1 SQL 概述	216
5.4.2 SQL 基础	217
5.4.3 SQL 的数据定义功能	218
5.4.4 SQL 的数据操纵	221
5.4.5 SQL 的数据库控制功能	226
5.4.6 嵌入式 SQL	227
5.5 数据库技术的新应用	230
习题五	233
第 6 章 信息系统	235
6.1 信息系统基础	235
6.1.1 信息与信息资源	235
6.1.2 信息的组织与管理	237
6.1.3 信息资源管理	241
6.2 信息系统的结构	242
6.2.1 信息系统的一般结构	242
6.2.2 信息系统的结构类型	243
6.2.3 信息系统的基本组织	244
6.2.4 信息系统的层次结构	244
6.2.5 信息系统的空间结构	245
6.2.6 管理活动的层次结构	245
6.2.7 决策过程的 3 阶段模型	246
6.2.8 决策的类型	247
6.3 事务处理系统	247
6.4 管理信息系统	250
6.4.1 管理信息系统概述	250
6.4.2 管理信息系统的结构	252
6.4.3 建立管理信息系统的基础	256
6.5 决策支持系统	257
6.5.1 决策支持系统概述	257
6.5.2 决策模式	258

6.5.3 决策支持系统的基本结构及信息组织	258
6.6 办公信息系统	259
6.6.1 办公自动化的概念	260
6.6.2 办公信息系统的系统模型	260
6.6.3 办公信息系统的结构	261
习题六	262
第7章 网络通信技术基础	263
7.1 TCP/IP 网络技术简介	263
7.1.1 TCP/IP 的层次模型和功能	263
7.1.2 TCP/IP 的数据链路层	264
7.1.3 TCP/IP 的网络层	265
7.1.4 TCP/IP 的传输层	266
7.1.5 TCP/IP 的应用层	267
7.2 客户机/服务器模式	268
7.3 Web 工作的基本原理	269
7.4 客户机/服务器软件设计中的算法	272
习题七	281
参考文献	282

第 1 章 数据结构及其算法

计算机科学的研究领域,概括来说,就是指在计算机中如何组织信息,如何处理这些信息以及如何利用这些信息,其中如何组织信息是很重要的一个方面。为此,人们提出数据结构这一研究课题,并获得了许多研究成果。

1.1 数据结构的引入

从提出一个实际问题到计算机解出答案,需要经历下列步骤:分析阶段、设计阶段、编码阶段和测试维护阶段等。其中分析阶段就是从实际问题中提取操作对象以及操作对象之间的关系。下面来看几个例子。

【例 1.1】 计算机管理图书目录问题。

利用计算机帮助查询书目,首先必须将书目存入计算机。那么这些书目如何存放呢?我们既希望查询时间短,又要求节省空间。一个简单的办法就是建立一张表,每本书的信息只用一张卡片表示,在表中占一行,如表 1.1 所示。此时计算机操作的对象(数据元素)便是卡片,卡片之间的关系是顺序排列的。计算机对数据的操作是按某个特定要求(如给定书名)进行查询的,继而找到表中满足要求的一行信息。由此可知,从计算机管理图书目录问题抽象出来的模型即是包含图书目录的表和对表进行的查找运算。

表 1.1 图书信息表

书 名	作 者	登录号	分类号	出版日期	定价/元
C 语言贯通教程	史斌星	347587	TP312188	2001.8	46.00
Delphi 5.0 因特网与数据库程序设计	刘海涛	347545	TP312184	2001.8	36.00
ATM 网络互联	[美]John A. Chiong	344558	TP393-518.7	2000.3	39.00
⋮	⋮	⋮	⋮	⋮	⋮

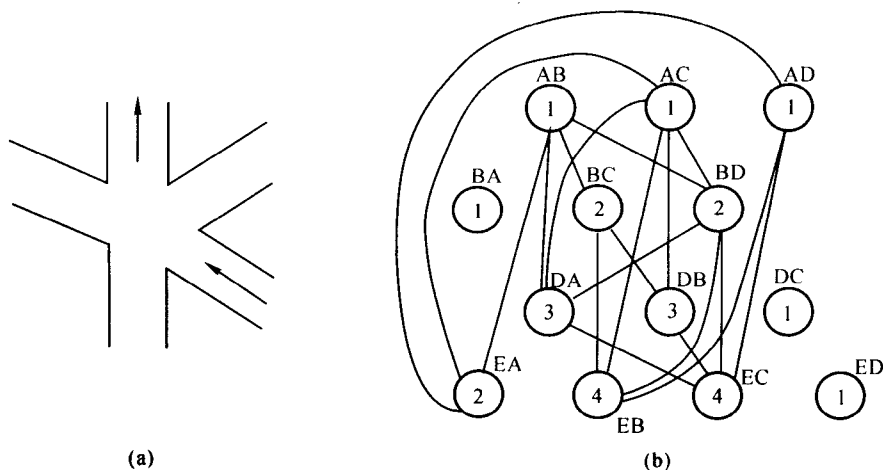


图 1.2 五叉路口交通管理示意图

(a) 五叉路口； (b) 表示通路的图

1.2 数据结构的基本概念

1.2.1 什么是数据结构

数据结构是计算机应用方面的基础知识之一。事实证明,要想有效地使用计算机,仅掌握计算机语言而缺乏数据结构和算法的有关知识是难以应付众多复杂的应用课题的。

早期的计算机主要用于解决数值计算问题,通常是用分析数学的方程式来建立数学模型,以此为加工对象的程序设计称为数值型程序设计。其特点是涉及的操作对象比较简单,一般为整型、实型和布尔型数据。随着计算机应用领域的不断扩大,解决非数值性问题越来越引起人们的关注,如文献检索、金融管理、商业系统数据处理、计算机辅助设计和制造以及以图论为基础的图像模式识别等。这类问题的重点在于数据处理,即对数据集中的各元素以各种方式进行运算,如插入、删除、查找、更新等。在数据处理领域中,数据类型比较复杂,而且数据元素之间具有各种特定的联系,人们最感兴趣的是了解数据集中元素之间的关系以及如何组织和表示这些数据以提高处理效率。数据结构是一门研究非数值运算的程序设计问题中计算机的操作对象以及它们之间的关系和操作等的学科。

1.2.2 数据结构的基本概念和术语

1. 基本术语

(1) 数据:是指反映客观事物的信息的集合,它是数据结构所要描述的东西。如自然数的集合、字符集、汽车部件的集合等都是数据。

(2) 数据元素:是数据的一个个体,它是数据的基本单位。自然数集中的一个自然数,或者字符集中的一个字符均为数据元素。数据元素可以是一个简单的数据项,如字符,也可以是

由几个数据项组合而成的复合数据项。例如,一个班级的学生档案信息数据由一个个学生的信息组成,而每一个学生的信息又包含了姓名、学号和性别等数据项。

(3) 数据对象:是指在数据这个集合中人们感兴趣的一个子集,通常,数据对象中的元素具有某些相同的特性。例如,奇数集和偶数集是两个数据对象,小写字符集 $\{a, b, c, \dots, z\}$ 是字符集中的一个数据对象。

(4) 数据结构:是指相互之间有关联的数据元素的集合。数据结构具有两个要素:数据元素的集合 D 和在 D 上的一个关系(二元关系),关系通常用 R 表示。因此数据结构可以表示为

$$\text{数据结构} = (D, R)$$

2. 数据结构示例

【例 1.4】 学生档案表。

如表 1.2 所示为学生档案表。该表由若干数据元素,即学生信息组成,而每一项学生信息都由几个相同的数据项组合而成,包括学号、姓名、性别、专业、出生年月、籍贯。该表结构按学号排序,前、后数据元素之间的关系由学号决定。除了首、尾元素外,每个元素均只有一个前趋元素和一个后继元素。这就是后面将要介绍的线性结构的一个典型例子。

表 1.2 学生档案表

学 号	姓 名	性 别	专 业	出生年月	籍 贯
990612	刘 华	女	计算机通信	1971.7	河 北
991001	王鲁彬	男	计算机应用	1971.3	山 东
991105	田 军	男	机械制造	1972.1	甘 肃
⋮	⋮	⋮	⋮	⋮	⋮

【例 1.5】 文件目录结构。

文件目录结构如图 1.3 所示。在这种结构中,每一个目录名都是一个数据元素,除了根目录以外,其他目录都有且仅有一个父目录,也就是只有一个前趋。同时,每个目录可以有不止一个子目录,即可以有不止一个后继。这样的结构称为树型结构,后面的章节将会详细讨论这一数据结构。当然,还有一些相对比较复杂的数据结构,例如索引结构就是其中一种,这些都是后面将要讲到的常用数据结构。

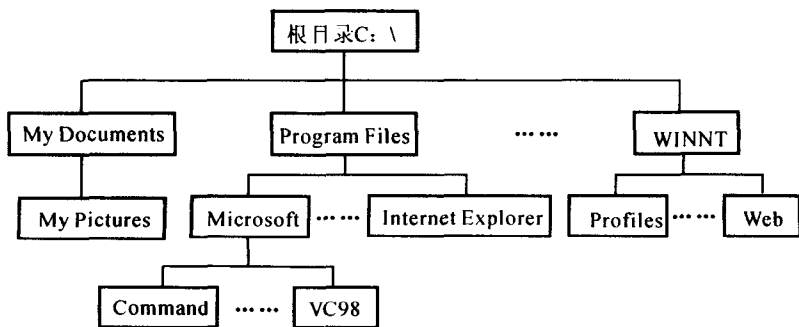


图 1.3 文件目录结构

【例 1.6】 城市交通网。

城市交通网如图 1.4 所示。在这种数据结构中,每一个数据元素为一个城市,它们之间具有多个前趋、后继的关系。这样的结构是最复杂的一种,称为图,这将在后面加以介绍。

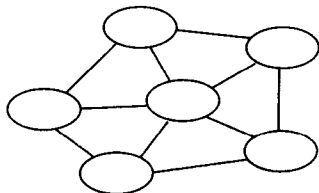


图 1.4 城市交通网

3. 数据结构研究的内容

数据结构研究的内容包括数据的逻辑结构、数据的存储结构和数据的运算。

(1) 数据的逻辑结构：

1) 线性结构:前面例 1.4 学生档案表就是线性结构的例子,从中可以看出,线性结构具有以下特点:

- (i) 非空结构有且仅有一个结点没有前趋元素,称为根结点;
- (ii) 每个结点最多有一个前趋,最多有一个后继;
- (iii) 插入或删除一个结点后仍然满足前两条。

线性数据结构包括线性表、栈、队列、串和数组。

2) 非线性结构:不属于以上线性结构的数据结构通称为非线性结构。因此,非线性结构中的元素具有任意关系,每一个元素的前趋数和后继数是任意的。例如前面提到的树结构和图结构都是非线性结构。

(2) 数据的存储结构:是指数据结构中的各个元素在计算机存储单元中的存放方式以及二元关系 R 的表示方法。比较常用的存储方式有顺序存储和链式存储两种。

1) 顺序存储结构:如果线性数据结构中的每一个数据元素所占用的存储单元数是相等的,则由线性结构的特点可知,二元关系 R 在计算机存储器中可以顺序实现,即可以用数据元素在存储器中存放位置的相邻关系来表示 R ,如图 1.5 所示。

⋮	⋯⋯⋯
Addr1	010101 李 玲 ⋯
Addr1+l	010102 张 宏 ⋯
Addr1+2×l	010103 刘冀平 ⋯
⋮	⋯⋯⋯

图 1.5 顺序存储结构

如图 1.5 所示是学生档案表的逻辑结构。假设每一条学生信息占据长度为 l 的存储区域,该逻辑结构在物理存储器中以一定的顺序占用连续的存储空间。对于这种结构,只需要知道第一个元素的地址和每一个元素所占的存储单元数就可以得到任何一个元素所在的位置。因此,在顺序存储结构中存取任意一个元素所花费的时间相等。

部分特殊的非线性结构在满足一定条件后,也可以采用顺序存储的结构来实现。例如后面将要讲到的完全二叉树就是一例。

2) 链式存储结构:顺序存储结构要求数据结构中的关系 R 是可以顺序实现的,而非线性结构的关系 R 一般不可能顺序实现,因此,非线性结构一般不能用顺序存储的方法实现。非线性结构通常采用链式存储结构。

在链式存储结构中,数据元素被存储在一个称为结点的存储项中,在结点中除了保存数据元素的内容外,还要保存该数据元素的所有后继元素的存储地址。这样,各结点在存储器中的位置可以是任意的,各数据元素之间的关系由存储于各结点中的后继指针来决定。如图 1.6 所示的就是学生档案表的链式存储结构。

链式结构当然也可以用于实现线性结构,例如,用单链表存储一个线性表,就具有插入、删除操作容易,存储空间扩充容易等优点。

此外,数据的存储结构还有索引存储结构和散列存储结构。

(3) 数据的运算:数据的运算是定义在数据逻辑结构上的操作,如检索、插入、删除等。每一种数据结构都有一个运算的集合。

1.2.3 算法描述及分析

1. 算法的概念

算法是由若干条指令组成的有限序列。它必须满足以下性质:

- (1) 输入性:具有零个或多个输入量,即算法开始前对算法给出的初始量。
- (2) 输出性:至少产生一个输出。
- (3) 有穷性:每条指令的执行次数必须是有限的。
- (4) 确定性:每条指令的含义必须明确,无二义性。
- (5) 可行性:每条指令都应在有限的时间内完成。

【例 1.7】 给定两个正整数 m 和 n ,求它们的最大公因子。

求解这个问题通常所用的方法为辗转相除法,下面用 3 个计算步骤描述这个算法。

- (1) 求余数:以 n 除 m ,余数为 r , $0 \leq r \leq n$;
- (2) 判断余数是否等于零:若 $r=0$,输出 n 的当前值,算法结束,否则执行第(3)步;
- (3) 更新被除数和除数: $n \rightarrow m, r \rightarrow n$,执行第(1)步。

上述计算过程给出了 3 个计算步骤,而且每一个步骤都意义明确,切实可行,虽然出现循环,但 m 和 n 都是给定的有限数。每次相除后得到的余数 r 若不为零,也总有 $r < \min(m, n)$,保证了经过有限次循环以后,必会终止,因此上述计算过程是一个算法。

算法的含义与程序十分相似,但二者是有区别的,一个程序不一定满足有穷性。例如,系统程序中的操作系统,只要整个系统不遭到破坏,它就永远不会停止,即使没有作业要处理,它仍处于一个等待循环中,以等待新作业的进入。因此,操作系统就不是一个算法。

另外,程序中的指令必须是机器可执行的,而算法中的指令则无此限制。但是一个算法若用机器可执行的语言来书写,则它就是一个程序。

2. 算法描述语言

算法必须通过算法语言来表达,即选用某一种高级语言编成程序才能在计算机上实现。为了教学需要,在这里采用一种算法描述语言来描述各种算法,它不直接用于计算机,主要为了简单明了地描述算法本身。鉴于读者在学习数据结构前均已具备一定的高级语言基础,因此本算法描述语言基本采用高级语言表达形式,但省略了高级语言形式化的类型说明、变量说明等,代之以较自由的自然语言作非形式化描述,有些部分增加必要的注释(用 $/* \dots */$ 表示),以增加算法的可读性。每一种算法均以函数(过程)的形式表示,即

算法名(参量表)

例如: INSERTLIST(V,i,n,x) $/*$ 顺序表的插入 $*/$

	
A224	010101 李 玲 ... A282
A242	010103 刘冀平 ...
⋮	
A282	010102 张 宏 ... A242
	

图 1.6 链式存储结构