



21st CENTURY
实用规划教材

21世纪全国高职高专
计算机系列实用规划教材

数据结构(C语言版)

主 编 夏 燕 张兴科
副主编 李笑雪 蔡 芸

内容特点:

- 每章都有要重点掌握的、在C语言环境下调试通过的上机实训题,章末配有大量标准化习题和上机操作题,并且提供参考操作步骤
- 内容详实,通俗易懂,理论的讲述够用为度,注重实践
- 本书可作为高职高专院校、技校、职高及社会办学相关课程的教材,也非常适合计算机相关专业初学者学习使用,特别是要参加专升本考试的读者,是很好的学习参考书



北京大学出版社
PEKING UNIVERSITY PRESS

21 世纪全国高职高专计算机系列实用规划教材

数据结构(C 语言版)

主 编	夏 燕	张兴科
副主编	李笑雪	蔡 芸
参 编	张顺利	李琳琳



北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

本书系统地介绍了较为常用的数据结构,主要包括线性表、栈、队列、串、数组、树和二叉树、图、查找表和排序,并按照高职高专计算机专业本课程大纲要求,对每种数据结构算法实现进行阐述,并对各种算法的时间和空间性能作了简要分析。

本书内容详实,通俗易懂,理论的讲述够用为度,注重实践。每章都有要重点掌握的、在C语言环境下调试通过的上机实训题,章末有大量标准化习题和上机操作题,并且上机操作题还配有参考操作步骤,使读者对每章的学习从理论到实践都能得到进一步巩固。

本书可作为高职高专院校、技校、职高及社会办学相关课程的教材,也非常适合计算机相关专业初学者学习使用,特别是要参加专升本考试的读者,是很好的学习参考书。

图书在版编目(CIP)数据

数据结构(C语言版)/夏燕,张兴科主编. —北京:北京大学出版社,2007.7

(21世纪全国高职高专计算机系列实用规划教材)

ISBN 978-7-301-12409-3

I. 数… II. ①夏… ②张… III. ①数据结构—高等学校:技术学校—教材②C语言—程序设计—高等学校:技术学校—教材 IV. TP311.12 TP312

中国版本图书馆CIP数据核字(2007)第083359号

书 名: 数据结构(C语言版)

著作责任者: 夏 燕 张兴科 主编

责任编辑: 李彦红

标准书号: ISBN 978-7-301-12409-3/TP·0904

出 版 者: 北京大学出版社

地 址: 北京市海淀区成府路205号 100871

网 址: <http://www.pup.cn> <http://www.pup6.com>

电 话: 邮购部 62752015 发行部 62750672 编辑部 62750667 出版部 62754962

电子邮箱: pup_6@126.com

印 刷 者: 河北滦县鑫华书刊印刷厂

发 行 者: 北京大学出版社

经 销 者: 新华书店

787毫米×1092毫米 16开本 20.75印张 471千字

2007年7月第1版 2007年7月第1次印刷

定 价: 28.00元

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究

举报电话: 010-62752024

电子邮箱: fd@pup.pku.edu.cn

21 世纪全国高职高专计算机系列实用规划教材

专家编审委员会

主 任 刘瑞挺

副主任 (按拼音顺序排名)

陈玉国 崔锁镇 高文志 韩希义

黄晓敏 魏 峥 谢一风 张文学

委 员 (按拼音顺序排名)

安志远 丁亚明 杜兆将 高爱国 高春玲 郭鲜凤

韩最蛟 郝金镇 黄贻彬 季昌武 姜 力 李晓桓

连卫民 刘德军 刘德仁 刘辉珞 栾昌海 罗 毅

慕东周 彭 勇 齐彦力 沈凤池 陶 洪 王春红

闻红军 武凤翔 武俊生 徐 红 徐洪祥 徐受容

许文宪 严仲兴 杨 武 易永红 于巧娥 袁体芳

张 昕 赵 敬 赵润林 周朋红 訾 波

信息技术的职业化教育

(代丛书序)

刘瑞挺/文

北京大学出版社第六事业部组编了一套《21 世纪全国高职高专计算机系列实用规划教材》。为此,制订了详细的编写目的、丛书特色、内容要求和风格规范。在内容上强调面向职业、项目驱动、注重实例、培养能力;在风格上力求文字精练、图表丰富、脉络清晰、版式明快。

一、组编过程

2004 年 10 月,第六事业部开始策划这套丛书,分派编辑深入各地职业院校,了解教学第一线的情况,物色经验丰富的作者。2005 年 1 月 15 日在济南召开了“北大出版社高职高专计算机规划教材研讨会”。来自 13 个省、41 所院校的 70 多位教师汇聚一堂,共同商讨未来高职高专计算机教材建设的思路和方法,并对规划教材进行了讨论与分工。2005 年 6 月 13 日在苏州又召开了“高职高专计算机教材大纲和初稿审定会”。编审委员会委员和 45 个选题的主、参编,共 52 位教师参加了会议。审稿会分为公共基础课、计算机软技术专业、计算机网络技术专业、计算机应用技术专业 4 个小组对稿件逐一进行审核。力争编写出一套高质量的、符合职业教育特点的精品教材。

二、知识结构

职业生涯的成功与人们的知识结构有关。以著名侦探福尔摩斯为例,作家柯南道尔在“血字的研究”中,对其知识结构描述如下:

- ◆ 文学知识——无;
- ◆ 哲学知识——无;
- ◆ 政治学知识——浅薄;
- ◆ 植物学知识——不全面。对于药物制剂和鸦片却知之甚详。对毒剂有一般了解,而对于实用园艺却一无所知;
- ◆ 化学知识——精深;
- ◆ 地质学知识——偏于应用,但也有限。他一眼就能分辨出不同的土质。根据裤子上泥点的颜色和坚实程度就能说明是在伦敦什么地方溅上的;
- ◆ 解剖学知识——准确,却不系统;
- ◆ 惊险小说知识——很渊博。似乎对近一个世纪发生的一切恐怖事件都深知底细;
- ◆ 法律知识——熟悉英国法律,并能充分实用;
- ◆ 其他——提琴拉得很好,精于拳术、剑术。

事实上,我国唐朝名臣狄仁杰,大宋提刑官宋慈,都有类似的知识结构。审视我们自己,每人的知识结构都是按自己的职业而建构的。因此,我们必须面向职场需要来设计教材。

三、职业门类

我国的职业门类分为 18 个大类：农林牧渔、交通运输、生化与制药、地矿与测绘、材料与能源、土建水利、制造、电气信息、环保与安全、轻纺与食品、财经、医药卫生、旅游、公共事业、文化教育、艺术设计传媒、公安、法律。

每个职业大类又分为二级类，例如电气信息大类又分为 5 个二级类：计算机、电子信息、通信、智能控制、电气技术。因此，18 个大类共有 75 个二级类。

在二级类的下面，又有不同的专业。75 个二级类共有 590 种专业。俗话说：“三百六十行，行行出状元”，现代职业仍在不断涌现。

四、IT 能力领域

通常信息技术分为 11 个能力领域：规划的能力、分析与设计 IT 解决方案的能力、构建 IT 方案的能力、测试 IT 方案的能力、实施 IT 方案的能力、支持 IT 方案的能力、应用 IT 方案的能力、团队合作能力、文档编写能力、项目管理能力以及其他能力。

每个能力领域下面又包含若干个能力单元，11 个能力领域共有 328 个能力单元。例如，应用 IT 方案能力领域就包括 12 个能力单元。它们是操作计算机硬件的能力、操作计算机软件包的能力、维护设备与耗材的能力、使用计算机软件包设计机构文档的能力、集成商务计算机软件包的能力、操作文字处理软件的能力、操作电子表格应用软件的能力、操作数据库应用软件的能力、连接到互联网的能力、制作多媒体网页的能力、应用基本的计算机技术处理数据的能力、使用特定的企业系统以满足用户需求的能力。

显然，不同的职业对 IT 能力有不同的要求。

五、规划梦想

于是我们建立了一个职业门类与信息技术的平面图，以职业门类为横坐标、以信息技术为纵坐标。每个点都是一个函数，即 IT(Professional)，而不是 IT+Professional 单纯的相加。针对不同的职业，编写它所需要的信息技术教材，这是我们永恒的主题。

这样组合起来，就会有 $IT((328) * (Pro(590)))$ ，这将是一个非常庞大的数字。组织这么多的特色教材，真的只能是一个梦想，而且过犹不及。能做到 $IT((11) * (Pro(75)))$ 也就很不容易了。

因此，我们既要在宏观上把握职业门类的大而全，也要在微观上选择信息技术的少而精。

六、精选内容

在计算机科学中，有一个统计规律，称为 90/10 局部性原理(Locality Rule)：即程序执行的 90% 代码，只用了 10% 的指令。这就是说，频繁使用的指令只有 10%，它们足以完成 90% 的日常任务。

事实上，我们经常使用的语言文字也只有总量的 10%，却可以完成 90% 的交流任务。同理，我们只要掌握了信息技术中 10% 频繁使用的内容，就能处理 90% 的职业化任务。

有人把它改为 80/20 局部性原理，似乎适应的范围更广些。这个规律为编写符合职业教育需要的精品教材指明了方向：坚持少而精，反对多而杂。

七、职业本领

以计算机为核心、贴近职场需要的信息技术已经成为大多数人就业的关键本领。职业教育的目标之一就是培养学生过硬的 IT 从业本领，而且这个本领必须上升到职业化的高度。

职场需要的信息技术不仅是会使用键盘、录入汉字，而且还要提高效率、改善质量、降低成本。例如，两位学生都会用 Office 软件，但他们的工作效率、完成质量、消耗成本可能有天壤之别。领导喜欢谁？这是不言而喻的。因此，除了道德品质、工作态度外，必须通过严格的行业规范和个人行为规范，进行职业化训练才能养成正确的职业习惯。

我们肩负着艰巨的历史使命。我国人口众多，劳动力供大于求的矛盾将长期存在。发展和改革职业教育，是我国全面建设小康社会进程中一项艰巨而光荣的任务，关系到千家万户人民群众的切身利益。职业教育和高技能人才在社会主义现代化建设中有特殊的作用。我们一定要兢兢业业、不辱使命，把这套高职高专教材编写好，为我国职业教育的发展贡献一份力量。

刘瑞挺教授 曾任中国计算机学会教育培训委员会副主任、教育部理科计算机科学教学指导委员会委员、全国计算机等级考试委员会委员。目前担任的社会职务有：全国高等院校计算机基础教育研究会副会长、全国计算机应用技术证书考试委员会副主任、北京市计算机教育培训中心副理事长。

本系列教材编写目的和教学服务

本系列教材在遍布全国的各位编写老师的共同辛勤努力下，在编委会主任刘瑞挺教授和其他编审委员会成员的指导下，在北京大学出版社第六事业部的各位编辑刻苦努力下，本系列教材终于与广大师生们见面了。

教材编写目的

近几年来，职业技术教育事业得以蓬勃的发展，全国各地的高等职业院校以及高等专科学校无论是从招生人数还是学校的软、硬件设施上都达到了相当规模。随着我国经济的高速发展，尽快提高职业技术教育的水平显得越来越重要。教育部提出：职业教育就是就业教育，也就是说教学要直接面对就业，强调实践。不但要介绍技术，更要介绍具体应用，注重技术与应用的结合。本套教材的主要编写思想如下。

1. 与发达国家相比，我国职业技术教育教材的发展比较缓慢并且滞后，远远跟不上职业技术教育发展的需求。我们常常提倡职业教育的实用性，但在课堂教学中仍然使用理论性和技术性教材进行职业实践教学。针对这种现状，急需推出一系列切合当前教育改革需要的高质量的优秀职业技术实训型教材。

2. 本套教材总结了目前优秀计算机职业教育专家的教学思想与经验，与广大职业教育一线老师共同探讨，最终落实到本套教材中，开发出一套适合于我国职业教育教学目标和教学要求的教材，它是一套能切实提高学生专业动手实践能力和职业技术素质的教材。

3. 社会对学生的职业能力的要求不断提高，从而催化出了许多新型的课程结构和教学模式。新型教学模式是必须以工作为基础的模仿学习，它是将学生置于一种逼真的模拟环境中，呈现给学生的是具有挑战性、真实性和复杂性的问题，使学生得到较真实的锻炼。

4. 教材的结构必须按照职业能力的要求创建并组织实施新的教学模式。教学以专项能力的培养展开，以综合能力的形成为目标。能力的培养既是教学目标，又是评估的依据和标准。

5. 本套的重点是先让学生实践，从实践中领悟、总结理论，然后再学习必要的理论，用理论指导实践。从这一个循环的教学过程中，学生的职业能力将得到极大的提高。

教学服务

1. 提供电子教案

本系列教材绝大多数都是教程与实训二合一，每一本书都有配套的电子教案，以降低任课老师的备课强度，此课件可以在我们网站上随时下载。

2. 提供教学资源下载

本系列教材中涉及到的实例(习题)的原始图片和其他素材或者是源代码、原始数据等文件，都可以在我们网站上下载。

3. 提供多媒体课件和教师培训

针对某些重点课程，我们配有相应的多媒体课件。对大批量使用本套教材的学校，我们会免费提供多媒体课件，另外还将免费提供教师培训名额，组织使用本套教材的教师进行相应的培训。

北京大学出版社第六事业部(<http://www.pup6.com>)

前 言

数据结构是计算机相关专业重要的专业基础课,是计算机程序设计的重要理论技术基础。通过该课程的学习,不仅可以使学生掌握数据结构的基本内容、典型算法和使用方法,而且能够训练学生应用数据结构和算法进行具体应用问题的程序设计能力,为学生后续课程的学习打下良好的基础。

本书主要是针对高职高专院校、职高和技校学生撰写的。教材的编写以高职高专计算机专业本课程教学大纲为主要依据,结合专升本考试大纲的要求,讨论了数据的逻辑结构与存储结构,以及相关算法,还讨论了常用的几种查找方法和排序算法,帮助读者轻松掌握数据结构的基础知识与算法实现。

全书共 10 章,内容包括数据结构概述、线性表、栈、队列、串、数组、树和二叉树、图、查找表和排序。每种数据结构都有较完整的数据类型定义、在计算机内的存储表示和算法实现,并运用大量实例进行讲解,每章都配有要求重点掌握、在 C 语言环境下调试通过的上机实训题,使读者通过上机实训来加深对数据结构课程的理解和掌握。

本书每章后面配有大量标准化习题和上机操作题,并且上机操作题还配有参考操作步骤,通过练习,使读者对每章的学习从理论到实践都能得到进一步巩固。同时还可以到北京大学出版社第六事业部网站(www.pup6.com)下载本书所有习题参考答案和上机操作题 C 语言程序。

本书由昆明冶金高等专科学校的夏燕、山东信息职业技术学院的张兴科担任主编,河南财经学院的李笑雪、北京交通职业技术学院的蔡芸担任副主编,郑州纺织专科学校的张顺利、淄博职业技术学院的李琳琳担任参编。夏燕编写了第 6 章、第 9 章,张兴科编写了第 7 章,李笑雪编写了第 2 章、第 3 章和第 4 章,蔡芸编写了第 10 章,张顺利编写了第 8 章,李琳琳编写了第 1 章和第 5 章。

郑州牧业高等专科学校的黄贻彬老师、洛阳大学的张红霞老师在本书编写过程中提供了很大帮助,在此深表感谢。

由于时间仓促,编者水平有限,书中错误和疏漏之处在所难免,恳请广大读者批评指正。

编者

2007 年 5 月

目 录

第1章 概论..... 1

1.1 数据结构的概念..... 1

1.1.1 什么是数据结构..... 1

1.1.2 基本术语..... 3

1.2 数据类型..... 6

1.3 算法..... 7

1.3.1 算法的描述..... 7

1.3.2 算法性能分析..... 9

1.4 本章小结..... 12

1.5 上机实训..... 12

1.5.1 实训目的..... 12

1.5.2 实训例题..... 12

1.6 习题与上机操作..... 17

第2章 线性表..... 19

2.1 线性表的逻辑结构..... 19

2.1.1 线性表的类型定义..... 19

2.1.2 线性表的基本操作..... 20

2.2 线性表的顺序存储..... 21

2.2.1 顺序表..... 21

2.2.2 顺序表的基本运算..... 23

2.2.3 顺序表的应用..... 28

2.3 线性表的链式存储..... 30

2.3.1 线性链表..... 30

2.3.2 动态内存分配..... 32

2.3.3 线性链表的基本运算..... 33

2.3.4 循环链表及运算..... 37

2.3.5 双向链表及运算..... 39

2.4 顺序表与链表的比较及应用举例..... 42

2.4.1 顺序表与链表的比较..... 42

2.4.2 线性表应用举例..... 43

2.5 本章小结..... 48

2.6 上机实训..... 49

2.6.1 实训目的..... 49

2.6.2 实训例题..... 49

2.7 习题与上机操作..... 61

第3章 栈..... 64

3.1 栈的定义和基本运算..... 64

3.1.1 栈的定义..... 64

3.1.2 栈的基本运算..... 65

3.1.3 栈的顺序存储结构..... 65

3.1.4 栈的链式存储结构..... 68

3.2 栈的应用举例..... 71

3.2.1 数制的转换..... 71

3.2.2 表达式求值..... 72

3.2.3 子程序调用问题..... 74

3.2.4 迷宫问题..... 77

3.3 本章小结..... 81

3.4 上机实训..... 81

3.4.1 实训目的..... 81

3.4.2 实训例题..... 81

3.5 习题与上机操作..... 92

第4章 队列..... 94

4.1 队列的定义和基本运算..... 94

4.1.1 队列的定义..... 94

4.1.2 队列的基本运算..... 95

4.1.3 队列的顺序存储结构..... 95

4.1.4 队列的链式存储结构..... 100

4.2 队列的应用..... 103

4.2.1 利用队列打印杨辉三角形..... 103

4.2.2 求迷宫的最短路径..... 105

4.3 本章小结..... 107

4.4 上机实训..... 107

4.4.1 实训目的..... 107

4.4.2 实训例题..... 107

4.5 习题与上机操作	116	7.1.3 树的表示方法	153
第 5 章 串	118	7.2 二叉树	154
5.1 串的基本概念和存储结构	118	7.2.1 二叉树的定义	154
5.1.1 串的基本概念	118	7.2.2 二叉树的性质	154
5.1.2 串的静态存储结构	119	7.2.3 二叉树的存储结构	156
5.1.3 串的动态存储结构	120	7.2.4 二叉树的遍历	161
5.2 串的运算	122	7.2.5 二叉树遍历的应用	165
5.2.1 串的基本运算	122	7.3 线索二叉树	169
5.2.2 实现串的基本运算的算法	123	7.3.1 线索二叉树的定义及结构	169
5.3 正文模式匹配	125	7.3.2 线索二叉树的基本运算	170
5.3.1 模式匹配	125	7.4 树、森林和二叉树的关系	174
5.3.2 简单模式匹配算法	125	7.4.1 树的存储结构	174
5.4 串操作应用举例	127	7.4.2 树、森林与二叉树的转换	177
5.5 本章小结	128	7.4.3 树和森林的遍历	179
5.6 上机实训	128	7.4.4 树的应用	180
5.6.1 实训目的	128	7.5 哈夫曼树及其应用	181
5.6.2 实训例题	129	7.5.1 哈夫曼树的定义	181
5.7 习题与上机操作	133	7.5.2 哈夫曼树的构造	183
第 6 章 数组	135	7.5.3 哈夫曼编码	186
6.1 数组的基本概念	135	7.6 本章小结	189
6.1.1 数组的定义及逻辑结构	135	7.7 上机实训	189
6.1.2 数组的顺序存储结构	136	7.7.1 实训目的	189
6.2 矩阵的压缩存储	138	7.7.2 实训例题	189
6.2.1 对称矩阵的压缩存储	138	7.8 习题与上机操作	196
6.2.2 三角矩阵的压缩存储	139	第 8 章 图	200
6.3 稀疏矩阵	139	8.1 图的基本概念	200
6.3.1 稀疏矩阵的概念	139	8.1.1 基本概念	200
6.3.2 稀疏矩阵的转置	141	8.1.2 基本术语	201
6.4 本章小结	142	8.1.3 基本操作	203
6.5 上机实训	142	8.2 图的存储结构	204
6.5.1 实训目的	142	8.2.1 邻接矩阵表示法	204
6.5.2 实训例题	142	8.2.2 邻接表表示法	206
6.6 习题与上机操作	148	8.3 图的遍历	210
第 7 章 树和二叉树	151	8.3.1 图的深度优先搜索遍历	210
7.1 树的定义和基本操作	151	8.3.2 图的广度优先搜索遍历	212
7.1.1 树的定义	151	8.4 连通网的最小生成树	214
7.1.2 树的常用术语	152	8.4.1 求无向图的连通分量	214
		8.4.2 生成树的概念	214

8.4.3 最小生成树	215	9.6 上机实训	266
8.5 最短路径	219	9.6.1 实训目的	266
8.5.1 求某一顶点到其他各项 点的最短路径——迪杰 斯特算法	219	9.6.2 实训例题	267
8.5.2 求任意一对顶点间的 最短路径	221	9.7 习题与上机操作	274
8.6 拓扑排序	223	第 10 章 排序	277
8.6.1 AOV 网	223	10.1 基本概念	277
8.6.2 拓扑排序的过程	224	10.1.1 概述	277
8.7 图的应用举例	225	10.1.2 基本术语	278
8.7.1 图的建立及遍历实用程序	225	10.2 插入排序	279
8.7.2 图的拓扑排序实用程序	228	10.2.1 直接插入排序	279
8.8 本章小结	230	10.2.2 希尔排序	281
8.9 上机实训	230	10.3 交换排序	283
8.9.1 实训目的	230	10.3.1 冒泡排序	283
8.9.2 实训例题	231	10.3.2 快速排序	285
8.10 习题与上机操作	239	10.4 选择排序	288
第 9 章 查找表	244	10.4.1 直接选择排序	288
9.1 查找表的基本概念	244	10.4.2 堆排序	289
9.1.1 基本概念	244	10.5 归并排序	293
9.1.2 术语	244	10.5.1 归并排序的概念	293
9.2 静态查找表	246	10.5.2 二路归并排序	294
9.2.1 顺序表查找	247	10.5.3 归并排序的算法	294
9.2.2 折半查找	248	10.5.4 算法分析	295
9.2.3 索引顺序查找	250	10.6 基数排序	295
9.3 动态查找表	251	10.6.1 多关键词排序	295
9.3.1 二叉查找树的定义	251	10.6.2 基数排序	296
9.3.2 二叉查找树的查找过程	252	10.7 外部排序	297
9.3.3 二叉查找树的插入操作	253	10.7.1 问题的提出	297
9.3.4 二叉查找树的删除操作	254	10.7.2 外部排序的基本过程	299
9.3.5 二叉查找树的性能分析	255	10.8 各种内部排序方法的比较	301
9.4 哈希表查找	256	10.9 排序的应用举例	302
9.4.1 哈希表的基本概念	256	10.9.1 希尔排序的实用程序	302
9.4.2 构造哈希表的方法	257	10.9.2 快速排序的实用程序	304
9.4.3 处理冲突的方法	260	10.10 本章小结	305
9.4.4 哈希表的查找性能	264	10.11 上机实训	306
9.5 本章小结	266	10.11.1 实训目的	306
		10.11.2 实训例题	306
		10.12 习题与上机操作	311
		参考文献	315

第1章 概 论

教学提示：“数据结构”是研究数据在程序设计中操作对象及其之间关系与操作的学科，是介于数学、计算机硬件和计算机软件三者之间的一门核心课程，属于计算机专业的核心基础课程。“数据结构”是数据的组织、存储和运算的总和；是数据按照某种关系组织起来，同时用一定的存储方式存储到计算机中。

教学要求：

- (1) 理解数据结构及相关的概念。
- (2) 掌握各种逻辑结构的特点。
- (3) 掌握算法的定义、特性及用类 C 语言描述算法的规则。
- (4) 掌握评价算法优劣的标准：时间复杂度、空间复杂度的定义及表示。

1.1 数据结构的概念

随着计算机科学技术的快速发展，尤其是计算机的软、硬件技术飞速发展，计算机的应用领域也越来越广泛，从最初的科学计算逐步发展到人类活动的各个领域。

计算机处理的数据对象从单纯的数值计算，扩展到图像、视频、音频等具有不同结构的非数值数据的处理。针对复杂的处理对象，程序设计必须更多的注意到数据和数据之间存在的关系。合理的数据组织形式与良好的程序设计方法相结合是有效地解决问题的前提。通常情况下，精心选择的数据结构可以产生更高的运行效率或者存储效率的算法。这就是“数据结构”这门学科形成和发展的原因。

1.1.1 什么是数据结构

随着计算机应用领域的不断扩大，计算机处理的对象更多的是非数值计算问题，它们的数学模型无法用数学方程来进行描述，此时就必须建立相应的数据结构来进行描述，分析问题中所用到的数据是如何组织的，研究数据之间的关系如何，进而为解决这些问题设计出合适的数据结构。

下面请看三个具体例子。

【例 1.1】 学生档案信息。

将学生信息组织成如表 1-1 所示的花名册。花名册中每个学生的信息由编号、姓名、性别、出生年月、籍贯等项目组成，占表的一行，一行信息称为一个结点，则表中的结点和结点之间是一种简单的线性关系，这就是上述花名册表的线性逻辑结构。当用计算机对上述花名册表中的数据进行运算时，就要考虑那些结点在计算机中的存储表示，即存储结构。另外，还必须考虑如何进行结点的插入、删除、修改、检索或查找，这就涉及数据的运算。

表 1-1 学生档案信息

编 号	姓 名	性 别	出 生 年 月	籍 贯
01	王蕾	女	1981.5	山东
02	万名化	男	1982.1	河北
03	赵力新	男	1981.6	浙江
04	孙仁	女	1983.1	辽宁
05	周立	男	1982.9	江苏
...	...	...	...	...

【例 1.2】 家族族谱。

家庭父子的关系，一父可对应多个子女，就不再如【例 1.1】所示——简单的一对一关系。将其用图示表示，如图 1.1 所示。

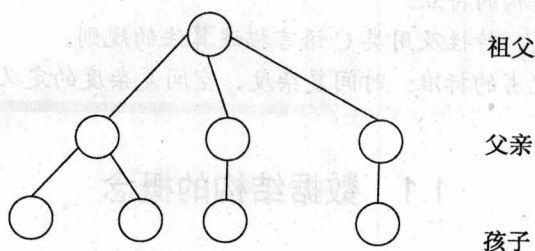


图 1.1 家族父子关系

参照图 1.1，要查找父亲的孩子，就不再是一对一的关系，同一个父亲可以找到多个孩子。同样用计算机进行处理时既要考虑存储问题，又要考虑增加、删除、查询等运算。

【例 1.3】 城市交通图。

众所周知，城市之间的交通联系比前两例要更复杂。也就是说数据元素之间出现了多对多的情况。用顶点表示城市，连线表示通路，如图 1.2 所示。

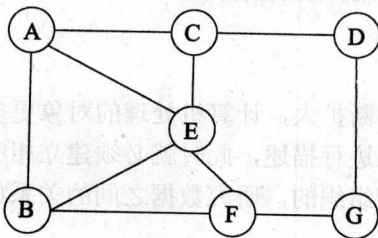


图 1.2 城市交通图

图 1.2 中，为表示城市交通图，可采用一种称为图的结构来表示实际的交通网络，此时，这个交通网络图就表示一个数据结构。在这个数据结构中，结点之间的关系可以是任意的，图中任意两个结点之间都可以相关。同样，必须考虑构造后将这张图存入计算机中，这就要涉及图的存储结构问题，以及图的运算。

从以上例子可见，要描述诸如此类的非数值计算问题的数学模型，则要涉及一些诸如表、图还有树之类的数据结构。

数据结构课程实际上是研究非数值计算问题的程序设计中计算机的操作对象以及它们之间的关系和运算操作等的一门学科。在计算机科学技术中,数据结构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,而且是设计和实现编译程序、操作系统、数据库系统及其他系统程序和大型应用程序的重要基础。

1.1.2 基本术语

下面介绍几个与数据结构相关的基本术语。

1. 数据

数据是对现实世界客观事物的符号化描述。因此,从整数、实数、字符、文字到声音、图形、图像都是数据。在计算机领域中,数据是指能够被计算机识别、存储、处理的一切信息符号。数据是计算机程序加工的“原料”。在程序设计语言中常说由程序处理数据,但程序也是数据。如C语言源程序由编译程序经过编译生成二进制的目标程序,对编译程序来说,源程序就是要处理的数据。

2. 数据元素

数据元素是数据整体中的相对独立的单位,即是数据的基本单位。通常在计算机程序中作为一个整体进行处理。数据元素一般由数据项组成,数据项(item)是具有独立含义的最小标识单位。

如表1-1所示,学生档案信息数据库中,一个学生信息记录可称为一个数据元素,而这个数据元素中的字段:学号、姓名、性别等就是一个个的数据项。

3. 数据对象

数据对象是具有相同特征的数据元素的集合,是数据的一个子集。数据对象与数据间的关系是部分与整体的关系。如整数数据对象是集合 $N=\{0,\pm 1,\pm 2,\dots\}$, 字母字符数据对象是集合 $C=\{'A','B','C',\dots,'Z'\}$, 表1-1所示的学生档案信息表也可看作一个数据对象。

4. 数据结构

数据结构是指相互之间存在一种或多种特定关系的数据元素集合,是带有结构的数据元素的集合,它指的是数据元素之间的相互关系,即数据的组织形式。

数据结构一般包括以下三方面的内容:

(1) 数据元素之间的逻辑关系,有时也称为数据的逻辑结构。

(2) 数据元素及其关系在计算机内存中的表示(又称为映像),称为数据的物理结构,又称为数据的存储结构,它包括数据元素的表示和数据元素之间关系的表示。

(3) 数据的运算及实现,即对数据元素可以施加的操作及其这些操作在相应的存储结构上的实现。

1) 数据的逻辑结构

数据的逻辑结构是指数据元素之间的逻辑关系描述。数据的逻辑结构通常有4类:集合结构、线性结构、树形结构、图形结构或称网状结构。

(1) 集合结构:数据元素之间组织形式松散,其结构为离散型,可以认为是数据元素间无逻辑关系,只是在同一个集合中,如图1.3所示。

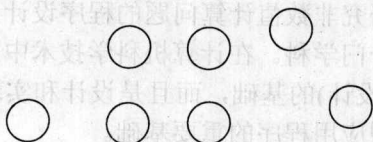


图 1.3 集合结构

(2) 线性结构: 数据元素之间是一一对应的关系。【例 1.1】中, 各学生记录按学号排列之间的关系, 可用图 1.4 来表示。很容易看出, 每个数据元素有且只有一个前驱元素(除第一个元素之外)和一个后继元素(最后一个元素除外)。也就是说, 每个元素都有唯一的前驱和后继(除第一个和最后一个元素外), 元素关系可以记为 1:1 的关系。



图 1.4 线性结构

(3) 树形结构: 数据元素之间是一对多的关系。【例 1.2】中, 父子关系, 可用图 1.5 表示, 像一棵倒挂的树。在树中, 最上面的结点(数据元素)称为根结点, 最下面的结点(数据元素)为叶子结点。可以看出, 每个结点最多只能有一个父亲, 但可以有多个孩子。即每个数据元素(根结点除外)有且只有一个前驱, 但可以有多个后继。元素关系可以记为 1:m 的关系。

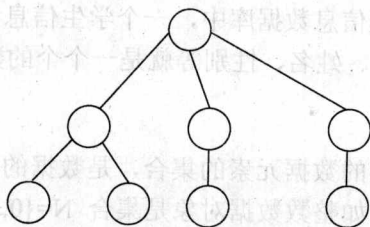


图 1.5 树形结构

(4) 图形结构: 数据元素之间是多对多的关系。如【例 1.3】中的城市交通图, 可以看出, 每个城市有多个前驱或后继。即图形结构特点是每个数据元素可以有多个前驱和后继。元素关系可以记为 $m:n$ 的关系。

2) 数据的存储结构

数据的存储结构是指逻辑结构用计算机语言的实现, 就是数据元素及其关系如何存入计算机的问题。数据的存储结构可按以下四种基本存储方法而得到。

(1) 顺序存储方法。

顺序存储方法是把逻辑上相邻的结点存储在物理位置上相邻的存储单元中, 结点间的逻辑关系由存储单元的相邻关系而体现。由此得到的存储表示称为顺序存储结构, 通常可借助计算机语言的数组来描述。

【例 1.4】顺序存储一棵树。

按照树的结构从上到下, 从左到右给每个结点编号, 在存储器中按编号进行顺序存储, 如图 1.6 所示。

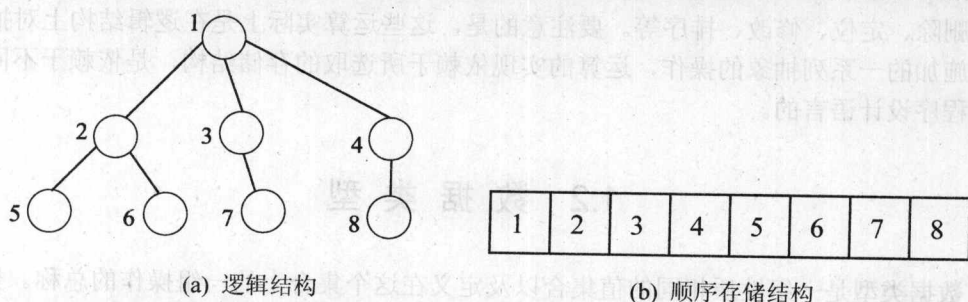


图 1.6 顺序存储结构

(2) 链状存储方式。

不要求逻辑上相邻的结点在物理位置上也相邻,数据元素可以存储在任意位置。为了实现数据元素之间逻辑关系的存储,必须通过一些附加的手段来存储这种相互关系,由此得到的存储表示称为链状存储结构。一般可以用指针来实现,通常可借助计算机程序语言的指针类型或者游标来描述。

【例 1.5】链式存储——线性表(A,B,C)。

如图 1.7 所示,开始元素位置 1006,通过元素 A 的地址(指针)可以找到元素 B,依次类推可以找到下一个元素。

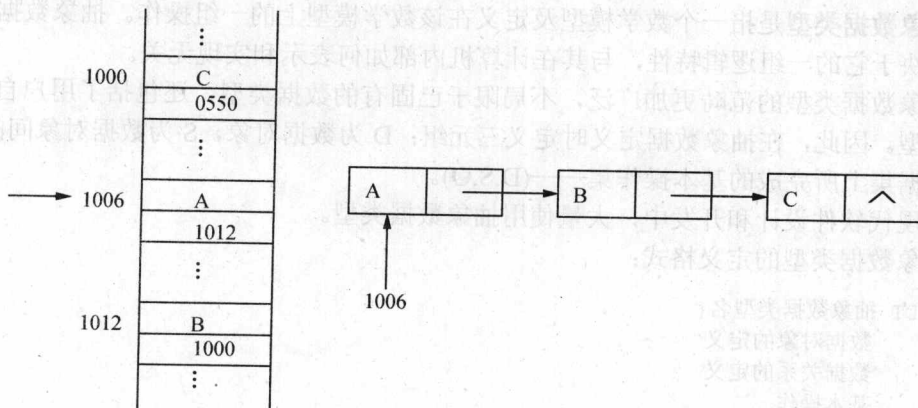


图 1.7 链式存储结构

(3) 索引存储方法。

该方法通常是在存储结点信息的同时,建立附加的索引表。索引表一般有稠密索引和稀疏索引两种。

(4) 散列存储方法。

该方法是依据结点的关键字直接计算出该结点的存储地址,而后将结点按某种方式存入该地址的一种存储方法。

数据的逻辑结构和物理结构是密切相关的,解决不同的问题,面对的数据的逻辑结构是不同的,设计的算法要依据逻辑结构,但算法如何实现又取决于数据的存储结构。

5. 数据的运算

数据的运算是指定义在数据的逻辑结构上的一组操作的集合。例如,检索(查找)、插