



电子·教育

高职高专计算机系列规划教材

中国计算机学会高职高专教育学组推荐出版

数据结构 (C语言版)

李益明 邓文华 主 编



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

高职高专计算机系列规划教材

数 据 结 构 (C 语言版)

李益明 邓文华 主 编

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书对常用的数据结构做了系统的介绍,力求概念清晰,注重实际应用。全书共分9章,依次介绍了数据结构的基本概念、线性表、栈、队列和数组、树结构和图结构,以及查找和排序等基本运算。此外,第9章给出了10套模拟试卷及相应的试题分析,供师生参考。全书用C语言作为算法描述语言并且每章后面均列举了典型应用实例,并配有PowerPoint演示文稿,以供教师教学使用。

本书主要面向高职高专院校计算机专业的学生,也可作为大学非计算机专业学生的选修课教材和计算机应用技术人员的自学参考用书。

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。

图书在版编目(CIP)数据

数据结构: C语言版 / 李益明, 邓文华主编. —北京: 电子工业出版社, 2004.9

(高职高专计算机系列规划教材)

ISBN 7-5053-9994-2

I. 数… II. ①李…②邓… 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. ①TP311.12②TP312

中国版本图书馆 CIP 数据核字 (2004) 第 093560 号



责任编辑: 张云怡

印 刷: 北京天宇星印刷厂

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

经 销: 各地新华书店

开 本: 787×1092 1/16 印张: 13 字数: 333 千字

印 次: 2005 年 8 月第 3 次印刷

印 数: 3000 册 定价: 19.00 元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系。联系电话:(010)68279077。质量投诉请发邮件至 zlts@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

出版说明

高职高专的计算机专业面临着两方面的巨大变化,一方面是计算机技术的飞速发展,另一方面是高职高专教育本身的改革和重组。

当前,计算机技术正经历着高速度、多媒体、网络化的发展,计算机教育特别是计算机专业的教材建设必须适应这种日新月异的形势,才能培养出不同层次的、合格的计算机技术专业人才。为了适应这种变化,国内外都在对计算机教育进行深入的研究和改革。美国 IEEE 和 ACM 在推出了《Computing Curricula 2000》之后,立即又推出了《Computing Curricula 2001》。全国高校计算机专业教学指导委员会和中国计算机学会教育委员会在 1999 年 9 月也提出了高等院校《计算机学科教学计划 2000》(征求意见稿)。目前,国内许多院校的老师、专家正在研究《Computing Curricula 2001》,着手 21 世纪的中国计算机教育的改革。

高专层次和本科层次的计算机教育既有联系又有区别,高专层次的计算机教育旨在培养应用型人才。自 20 世纪 70 年代末高等专科学校计算机专业相继成立以来,高等专科学校积极探索具有自己特色的教学计划和配套教材。1985 年,在原电子工业部的支持下,由全国数十所高等专科学校参加成立了中国计算机学会教育委员会大专教育学组,之后又成立了大专计算机教材编委会。1986 年~1999 年,在各校老师的共同努力下,已相继完成了三轮高等专科学校计算机教材的规划与出版工作,共出版了 78 种必修课、选修课、实验课教材,较好地解决了高专层次计算机专业的教材需求。

为了适应计算机技术的飞速发展,以及高职高专计算机教育形势发展的需要,中国计算机学会教育委员会高职高专教育学组和高职高专计算机教材编委会于 2000 年 7 月开始,又组织了一批本科高校、高等专科学校、高等职业技术学院和成人教育高等院校的有教学经验的老师,学习研究参考了高等院校《计算机学科教学计划 2000》(征求意见稿),提出了按照新的计算机教育计划和教学改革的要求,编写高专、高职、成人高等教育三教统筹的第四轮教材。

第四轮教材的编写工作采取了以招标的方式征求每门课程的编写大纲和主编,要求投标老师详细说明课程改革的思路、本课程和相关课程的联系、重点和难点的处理等。在第四轮教材的编写过程中,编委会强调加强实践环节、强调三教统筹、强调理论够用为度的原则,要求教学计划、教学内容适应高等教育发展的新形势。本套教材的编者均为各院校具有丰富教学实践经验的教师。因此,第四轮教材的特点是体系结构比较合理、内容新颖、概念清晰、通俗易懂、理论联系实际、实用性强。

竭诚希望广大师生对本套教材提出批评建议。

中国计算机学会教育委员会高职高专教育学组
2001 年 1 月

部分学组成员单位名单

安徽淮南联合大学	河北工业职业技术学院
安徽职业技术学院	河北师范大学
保定职业技术学院	河南大学
北方工业大学	河南机电高等专科学校
北京船舶工业管理干部学院	河南新乡平原大学
北京电子信息职业技术学院	河南职业技术学院
北京科技大学职业技术学院	黑龙江大学职业技术学院
北京师范大学信息科学学院	湖北沙市大学
北京市机械局职工大学	湖南财经高等专科学校
北京信息工程学院	湖南城市学院
常州工学院	湖南大学
成都电子机械高等专科学校	湖南环境生物职业技术学院
成都航空职业技术学院	湖南计算机高等专科学校
成都师范高等专科学校	湖南民政职业技术学院
成都信息工程学院	湖南税务高等专科学校
承德石油高等专科学校	湖南铁道职业技术学院
重庆电子职业技术学院	湖州职业技术学院
重庆工业职业技术学院	淮安信息职业技术学院
佛山科技学院	淮海工学院
福州大学职业技术学院	黄石高等专科学校
广东女子职业技术学院	吉林大学
广东轻工职业技术学院	吉林交通职业技术学院
广西水利电力职业技术学院	吉林职业师范学院工程学院
广西职业技术学院	济源职业技术学院
广州大学科技贸易技术学院	江汉大学
广州航海高等专科学校	江苏常州机电职业技术学院
广州市财贸管理干部学院	金陵职业大学
桂林电子工业学院	军械工程学院
哈尔滨师范大学	空军后勤学院
哈尔滨学院	兰州师范专科学校
海淀走读大学信息学院	兰州石化职业技术学院
海口经济职业技术学院	连云港化工高等专科学校
海南职业技术学院	辽东学院
杭州经贸职业技术学院	辽宁交通高等专科学校
杭州商学院	辽阳高等职业技术学院
河北沧州职业技术学院	柳州职业技术学院
河北大学	洛阳大学

漯河职业技术学院
南京工程学院
南京建筑工程学院
南京农业专科学校
南京师范大学
南京钟山学院
南宁职业技术学院
宁波高等专科学校
青岛化工学院
青岛科技大学
青岛职业技术学院
山西大同职业技术学院
山西工业职业技术学院
山西师范大学
陕西工业职业技术学院
上海第二工业大学
上海电机技术高等专科学校
上海交通大学应用技术学院
上海理工大学
上海旅游高等专科学校
上海商业职业技术学院
上海托普职业技术学院
上海应用技术学院
韶关大学
邵阳高等专科学校
深圳职业技术学院
沈阳电力高等专科学校
四川师范学院
四川托普信息职业技术学院

苏州市职工大学
苏州铁路机械学校
苏州职业大学
台州职业技术学院
泰州职业技术学院
天津滨海职业学院
天津渤海职业技术学院
天津大学高职学院
天津电子信息职业技术学院
天津轻工业学院
天津师范大学计算机与信息学院
潍坊高等专科学校
温州大学
无锡职业技术学院
武汉职业技术学院
西安电子科技大学
兖州矿区职业大学
云南财贸学院
浙江大学
浙江工贸职业技术学院
浙江育英学院
郑州工业高等专科学校
郑州经济管理干部学院
郑州经济管理学院
中国保险管理干部学院
中国地质大学
中国人民大学成人教育学院
中州大学

前 言

“数据结构”是计算机程序设计的重要理论基础，是计算机及其应用专业的一门重要基础课程和核心课程。它不仅是计算机软件专业课程的先导，而且也逐渐为其他工科类专业所重视。

本教材作为“高职高专计算机系列规划教材”之一，主要面向高职高专院校计算机类专业的学生。教材内容的构造力求体现“以应用为主体”，强调对理论知识的理解和运用，实现专科教学以实践体系及以技术应用能力培养为主的目标。

本书共分9章，第1章叙述数据、数据结构和算法等基本概念；第2章至第6章分别讨论了线性表、栈和队列、串和数组、树以及图等基本数据结构及其应用；第7,8章分别讨论了查找和排序的各种实现方法及应用。此外，第9章给出了10套模拟试卷及试题分析与参考答案，供学生复习迎考使用。

本教材具有以下特点：

(1) 基础理论知识的阐述由浅入深、通俗易懂。内容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。

(2) 各章（除第1,8章外）都配有“应用举例”，列举分析了很多实用的例子，且大多数算法都直接给出了其相应的C语言程序，以便学生上机练习、实践。

(3) 为了便于学生复习及掌握每章的重点、要点，本书在每章的结束处对该章进行了小结，并列出了本章的重点。

(4) 本教材配有相应的PowerPoint演示文稿，可为教师节省大量的备课时间。

本教材讲课时数为60学时左右，上机时数为20学时左右。教师可根据学时数、专业和学生的实际情况选讲应用举例中的例子。

本书第1,5,6章由李益明副教授编写，第2,3,4章由邓文华副教授编写，第7,8章由张超老师编写，第9章由李益明、邓文华老师共同编写，全书由李益明老师修改定稿。

南京工程学院李平副教授对全书进行了认真仔细的审定，并提出了许多宝贵的意见和建议，作者在此表示衷心的感谢。

编写高职高专计算机专业教材是一项新的尝试，由于时间紧迫，书中难免存在错误和不足之处，恳请广大读者批评指正。

编 者

2004年6月

目 录

第 1 章 绪论	(1)
1.1 从问题到程序	(1)
1.2 有关概念和术语	(4)
1.3 算法及算法分析	(6)
1.3.1 算法特性	(6)
1.3.2 算法描述	(7)
1.3.3 算法分析	(7)
本章小结	(9)
习题 1	(9)
第 2 章 线性表	(11)
2.1 线性表的逻辑结构	(11)
2.1.1 线性表的定义	(11)
2.1.2 线性表的基本操作	(11)
2.2 线性表的顺序存储及操作实现	(12)
2.2.1 顺序表	(12)
2.2.2 顺序表的基本操作实现	(14)
2.2.3 顺序表应用举例	(17)
2.3 线性表的链式存储和运算实现	(19)
2.3.1 单链表	(19)
2.3.2 单链表上基本运算的实现	(20)
2.3.3 循环链表	(26)
2.3.4 双向链表	(27)
2.4 应用举例	(28)
本章小结	(30)
习题 2	(30)
第 3 章 栈和队列	(33)
3.1 栈	(33)
3.1.1 栈的定义及基本运算	(33)
3.1.2 栈的存储实现和运算实现	(34)
3.2 栈的应用举例	(36)
3.3 队列	(40)
3.3.1 队列的定义及基本运算	(40)
3.3.2 队列的存储实现及运算实现	(41)
3.4 队列应用举例	(46)
本章小结	(47)
习题 3	(48)

第4章 串和数组	(50)
4.1 串	(50)
4.1.1 串的基本概念	(50)
4.1.2 串的基本运算	(50)
4.1.3 串的存储及基本运算的实现	(51)
4.2 数组	(54)
4.2.1 数组的逻辑结构	(54)
4.2.2 数组的内存映像	(54)
4.2.3 稀疏矩阵	(55)
4.3 应用举例	(58)
本章小结	(60)
习题4	(60)
第5章 树与二叉树	(61)
5.1 树的概念与基本操作	(61)
5.1.1 树的定义及相关术语	(61)
5.1.2 树的基本操作	(63)
5.2 二叉树	(63)
5.2.1 二叉树的基本概念	(63)
5.2.2 二叉树的主要性质	(64)
5.2.3 二叉树的基本操作与存储实现	(65)
5.2.4 二叉树的遍历	(68)
5.3 树与森林	(71)
5.3.1 树的存储	(71)
5.3.2 树、森林与二叉树的转换	(73)
5.3.3 树和森林的遍历	(75)
5.4 最优二叉树——哈夫曼树	(76)
5.4.1 哈夫曼树的基本概念	(76)
5.4.2 哈夫曼树的构造算法	(78)
5.4.3 哈夫曼树在编码问题中的应用	(79)
5.5 应用举例	(80)
本章小结	(82)
习题5	(83)
第6章 图	(85)
6.1 图的基本概念	(85)
6.1.1 图的定义和术语	(85)
6.1.2 图的基本操作	(87)
6.2 图的存储表示	(88)
6.2.1 邻接矩阵	(88)
6.2.2 邻接表	(90)
6.3 图的遍历	(91)

6.3.1 深度优先查找	(91)
6.3.2 广度优先查找	(93)
6.4 图的应用	(94)
6.4.1 最小生成树	(94)
6.4.2 最短路径	(97)
6.4.3 拓扑排序	(100)
6.5 应用举例	(103)
本章小结	(107)
习题 6.....	(107)
第 7 章 查找	(109)
7.1 基本概念与术语	(109)
7.2 静态查找表	(110)
7.2.1 静态查找表结构	(110)
7.2.2 顺序查找	(111)
7.2.3 有序表的折半查找	(112)
7.2.4 分块查找(索引顺序查找)	(114)
7.3 动态查找表	(115)
7.4 哈希表(杂凑或散列表)	(119)
7.4.1 哈希表与哈希方法	(119)
7.4.2 常用的哈希函数构造方法	(120)
7.4.3 处理冲突的方法	(121)
7.4.4 哈希表的查找及其分析	(123)
7.5 应用举例	(124)
本章小结	(130)
习题 7.....	(130)
第 8 章 排序	(132)
8.1 基本概念	(132)
8.2 3 种简单排序方法	(133)
8.2.1 直接插入排序	(133)
8.2.2 冒泡排序(Bubble Sort)	(134)
8.2.3 简单选择排序	(135)
8.3 快速排序	(136)
8.4 堆排序(Heap Sort)	(139)
8.5 归并排序	(141)
8.6 基数排序	(143)
8.6.1 多关键码排序	(143)
8.6.2 链式基数排序	(144)
8.7 各种排序方法的比较与讨论	(146)
本章小结	(146)
习题 8.....	(147)

第9章 模拟试题..... (148)

 模拟试题 1..... (148)

 模拟试题 2..... (150)

 模拟试题 3..... (152)

 模拟试题 4..... (154)

 模拟试题 5..... (156)

 模拟试题 6..... (159)

 模拟试题 7..... (162)

 模拟试题 8..... (165)

 模拟试题 9..... (168)

 模拟试题 10 (171)

附录 部分习题参考答案..... (175)

第1章 绪 论

数据作为计算机加工处理的对象，在计算机中如何表示、存储是计算机科学研究的主要内容之一，更是计算机技术需要解决的关键问题之一。数据是计算机化的信息，它是计算机可以直接处理的最基本和最重要的对象。科学计算、数据处理、过程控制、文件存储、数据库技术等，都是对数据进行加工处理的过程。因此，要设计出一个结构好、效率高的程序，必须研究数据的特性及数据间的相互关系及其对应的存储表示，并利用这些特性和关系设计出相应的算法和程序。

数据之间的逻辑关系和物理实现

1.1 从问题到程序

“数据结构”作为计算机科学与技术专业的专业基础课，是十分重要的核心课程，其主要的研究内容就是数据之间的逻辑关系和物理实现，探索有利的数据组织形式及存取方式。所有的计算机系统软件和应用软件的设计、开发都要用到各种类型的数据结构。因此，要想更好地运用计算机来解决实际问题，仅掌握几种计算机程序设计语言是难以应付众多复杂的课题的。要想有效地使用计算机、充分发挥计算机的性能，还必须学习和掌握数据结构的有关知识。学好“数据结构”这门课程，对学习计算机专业的其他课程，如操作系统、编译原理、数据库管理系统、软件工程、人工智能等都是十分有益的。

在计算机技术发展的初期，人们使用计算机的目的主要是处理数值计算问题。当人们使用计算机解决具体问题时，一般需要经过以下几个步骤：首先从该具体问题抽象出一个适当的数学模型，然后设计或选择一个解此数学模型的算法，最后编出程序进行调试、测试，直至得到最终的结果（如图 1.1 所示）。例如，求解梁架结构中应力的数学模型的线性方程组，该方程组可以使用迭代算法来求解。

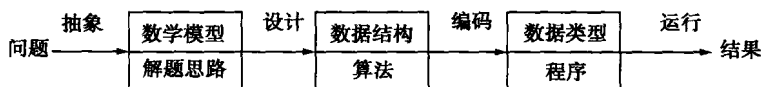


图 1.1 计算机解决问题的一般过程

由于当时所涉及的运算对象是简单的整型、实型或布尔型数据，所以程序设计者的主要精力集中在程序设计的技巧上，而无须重视数据结构。随着计算机应用领域的扩大和软、硬件的发展，非数值计算问题越来越显示出重要性。据统计，当今处理非数值计算性问题占用了 90% 以上的机器时间。这类问题涉及到的处理对象不再是简单的数据类型，其形式更加多样，结构更为复杂，数据元素之间的相互关系一般无法直接用数学方程式加以描述。因此，解决这类问题的关键不再是数学分析和计算方法，而是要设计出合理的数据结构，才能有效地解决问题。下面所列举的就是属于这一类的具体问题。

【例 1.1】 图书信息检索系统。在现代图书馆中，人们往往借助计算机图书检索系统来

查找需要的图书信息；或者直接通过图书馆信息系统进行图书借阅。为此，人们需要将图书信息按不同分类进行编排，建立合适的数据结构进行存储和管理，按照某种算法编写相关程序，就可以实现计算机自动检索。由此，一个简单的信息检索系统就是建立一张按图书分类号和登录号顺序排列的图书信息表，以及分别按作者名、出版社等顺序排列的各类索引表，如图 1.2 所示。由这 3 张表构成的文件就是图书信息检索的数学模型，计算机的主要操作就是按照用户的要求（如给定书名）通过不同的索引表对图书信息进行查询。

图书分类号	登录号	书 名	作者	出 版 社
B259.1	3240	梁启超家书	张品兴	中国文联出版社
C52	5231	探寻语碎	李泽厚	上海文艺出版社
D035.5	6712	市政学	张永桃	高等教育出版社
G206	1422	传播学	邵培仁	高等教育出版社
H319.4	1008	英语阅读策略	李宗宏	兰州大学出版社
K825.4.00	5819	围棋人生	聂卫平	中国文联出版社
P1.00	8810	通向太空之路	邹惠成	科学出版社
TN915	7911	通信与网络技术概论	刘云	中国铁道出版社
TP312	7623	计算机软件技术基础	王宇川	科学出版社
TP393.07	8001	网络管理与应用	张琳	人民邮电出版社
Q3.00	2501	普通遗传学	杨业华	高等教育出版社

(a) 图书信息表

邵培仁	4
李泽厚	2
李宗宏	5
刘云	8
聂卫平	6
王宇川	9
杨业华	11
张琳	10
张品兴	1
张永桃	3
邹惠成	7

(b) 作者名索引表

高等教育出版社	3,4,11
科学出版社	7,9
兰州大学出版社	5
人民邮电出版社	10
上海文艺出版社	2
中国铁道出版社	8
中国文联出版社	1,6

(c) 出版社索引表

图 1.2 图书信息检索系统中的数据结构

诸如此类的还有电话自动查号系统、学生信息查询系统、仓库库存管理系统等。在这类文档管理的数学模型中，计算机处理的对象之间通常存在着一种简单的线性关系，这类数学模型可称为线性的数据结构。

【例 1.2】 人机对弈问题。人机对弈是一个古老的人工智能问题，其解题思想是将对弈的策略事先存入计算机，策略包括对弈过程中所有可能出现的情况以及响应的对策。在决定对策时，根据当前状态，考虑局势发展的趋势做出最有利的选择。由此，计算机操作的对象

(数据元素)是对弈过程中的每一步棋盘状态(格局),数据元素之间的关系由比赛规则决定。通常,这个关系不是线性的,因为从一个格局可以派生出多个格局,因此常用树形结构来表示。如图 1.3 所示是井字棋的对弈树。

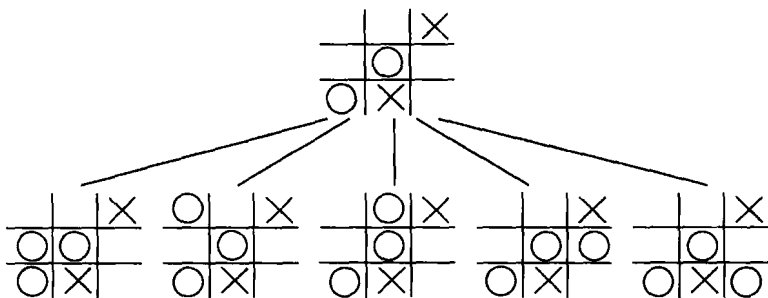
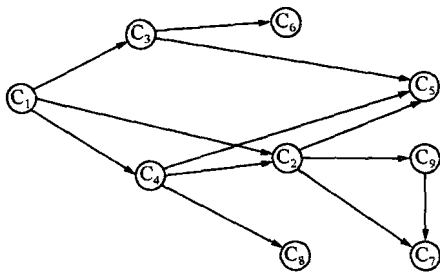


图 1.3 井字棋对弈树

【例 1.3】 教学计划编排问题。一个教学计划包含许多课程,在教学计划包含的许多课程之间,有些课程必须按规定的先后次序进行学习,有些则没有次序要求。即有些课程之间有先修和后续的关系。这种各个课程之间先修和后续的次序关系可用一个称做“图”的数据结构来表示,如图 1.4 所示。有向图中的每个顶点表示一门课程,如果从顶点 C_i 到 C_j 之间存在有向边 $\langle C_i,C_j\rangle$,则表示课程 i 必须先于课程 j 进行。

课程编号	课程名称	先修课程
C_1	计算机导论	无
C_2	数据结构	C_1,C_4
C_3	汇编语言	C_1
C_4	C 程序设计语言	C_1
C_5	计算机图形学	C_2,C_3,C_4
C_6	接口技术	C_3
C_7	数据库原理	C_2,C_9
C_8	编译原理	C_4
C_9	操作系统	C_2

(a) 计算机专业的课程设置



(b) 表示课程之间优先关系的有向图

图 1.4 教学计划编排问题的数据结构

由以上几个例子可见,描述这类非数值计算问题的数学模型不再是数学方程,而是诸如表、树、图之类的数据结构。因此,可以说数据结构课程主要是研究非数值计算的程序设计问题中所出现的计算机处理对象以及它们之间的关系和操作的学科。

学习数据结构的目的是为了了解计算机处理对象的特性,将实际问题中所涉及的处理对象在计算机中表示出来并对它们进行处理。与此同时,通过算法训练提高学生的思维能力;通过程序设计的技能训练促进学生的综合应用能力,提高学生的专业素质。

1.2 有关概念和术语

在系统地学习数据结构知识之前，先掌握一些基本概念和术语的定义。

1. 数据 (Data)

数据是信息的载体，它能够被计算机识别、存储和加工处理，它是计算机程序加工的原料。应用程序可以处理各种各样的数据，它可以是数值数据，也可以是非数值数据。数值数据是一些整数、实数或复数；非数值数据包括字符、文字、图形、图像、语音等。

2. 数据元素 (Data Element)

数据元素是数据的基本单位，在计算机程序中通常作为一个整体进行考虑和处理。一个数据元素可由若干个数据项 (Data Item) 组成。在不同的条件下，数据元素又可称为元素、结点、顶点、记录等。例如，学生信息检索系统中学生信息表中的一个记录、教学计划编排问题中的一个顶点等，都被称为一个数据元素。

3. 数据项 (Data Item)

数据项指不可分割的、含有独立意义的最小数据单位，数据项有时也称字段(Field)或域。例如，学籍管理系统中学生信息表的每一个数据元素就是一个学生记录。它包括学生的学号、姓名、性别、籍贯、出生年月、成绩等数据项。这些数据项可以分为两种：一种叫做初等项，如学生的性别、籍贯等，这些数据项是在数据处理时不能再分割的最小单位；另一种叫做组合项，如学生的成绩，它可以再划分为数学、物理、化学等更小的项。通常，在解决实际应用问题时是把每个学生记录当做一个基本单位进行访问和处理的。

4. 数据结构 (Data Structure)

数据结构是指互相之间存在着一种或多种关系的数据元素的集合。在任何问题中，数据元素之间都不会是孤立的，在它们之间都存在着这样或那样的关系，这种数据元素之间的关系称为结构。根据数据元素间关系的不同特性，通常有以下 4 类基本结构 (数据的逻辑结构)。

(1) 集合结构：在集合结构中，数据元素间是“属于同一个集合”的关系。数据元素之间除了同属一个集合外，不存在其他关系。

(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) 图状结构：该结构的数据元素之间存在着多对多的关系，图状结构也称网状结构。

图 1.5 为表示上述 4 类基本结构的示意图。

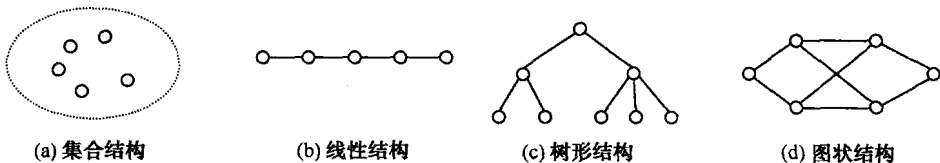


图 1.5 4 类基本结构示意图

由于集合是数据元素之间关系极为松散的一种结构，因此也可用其他结构来表示它，本书不做专门讨论。

从上面所介绍的数据结构的概念中可以知道，一个数据结构有两个要素。一个是数据元素的集合，另一个是数据元素之间关系的集合。在形式上，数据结构通常可以用一个二元组来表示：

$$\text{Data_Structure} = (D, R)$$

其中，D 是数据元素的有限集合，R 是 D 中元素之间关系的有限集合。

数据结构包括数据的逻辑结构和数据的物理结构。数据的逻辑结构可以看做是从具体问题抽象出来的数学模型，它与数据的存储无关。我们研究数据结构的目的是为了在计算机中实现对它的操作，为此还需要研究如何在计算机中表示和存储一个数据结构。数据的逻辑结构在计算机中的表示（又称映像）称为数据的物理结构，或称存储结构。它所研究的是数据结构在计算机中的实现方法，包括数据结构中数据元素的表示及数据元素之间关系的表示。

数据的存储结构可采用顺序存储或链式存储的方法。

(1) 顺序存储方法是通过数据元素在计算机中存储位置上的关系来表示出元素逻辑上的关系，通常把逻辑上相邻的元素存储在物理位置相邻的存储单元中，由此得到的存储表示称为顺序存储结构。顺序存储结构是一种最基本的存储表示方法，通常借助于程序设计语言中的数组来实现。

(2) 链式存储方法对逻辑上相邻的元素不要求其物理位置相邻，元素间的逻辑关系通过附设的指针字段来表示，由此得到的存储表称为链式存储结构，链式存储结构通常借助于程序设计语言中的指针类型来实现。

除了通常采用的顺序存储方法和链式存储方法外，有时为了查找方便还采用索引存储方法和散列表（Hash）存储方法。

5. 数据类型 (Datatype)

数据类型 { 非结构类型
结构类型

数据类型是和数据结构密切相关的一个概念，在高级程序设计语言中用以限制变量取值范围和可能进行的操作的总和称为数据类型。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，它的成分可以由多个结构类型组成，并可以分解。结构类型的成分中可以是非结构的，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等基本类型，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相应的数据类型，进一步在确定的数据类型上实现各种操作。

“数据结构”作为一门独立的课程在国外是从 1968 年才开始的，但在此之前其有关内容已散见于编译原理及操作系统之中。20 世纪 60 年代中期，美国的一些大学开始设立有关课程，但当时的课程名称并不叫数据结构。1968 年美国唐·欧·克努特教授开创了数据结构的最初体系，他编著的《计算机程序设计技巧》第一卷《基本算法》是第一本较系统地阐述数据的逻辑结构和存储结构及其操作的著作。从 20 世纪 60 年代末到 70 年代初，出现了大型程序，软件也相对独立，结构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 算法及算法分析

算法与数据结构的关系紧密，在算法设计时先要确定相应的数据结构，而在讨论某一种数据结构时也必然会涉及相应的算法。下面就从算法特性、算法描述、算法性能分析与度量等3个方面对算法进行介绍。

1.3.1 算法特性

算法 (Algorithm) 是对特定问题求解步骤的一种描述，是指令的有限序列。其中每一条指令表示一个或多个操作。一个算法应该具有以下特性。

- (1) 有穷性：一个算法必须在有穷步之后结束，即必须在有限时间内完成。
- (2) 确定性：算法的每一步必须有确切的定义，无二义性，且在任何条件下算法只有惟一一条执行路径，即对于相同的输入只能得出相同的输出。
- (3) 可行性：算法中的每一步都可以通过已经实现的基本运算的有限次执行得以实现。
- (4) 输入：一个算法具有零个或多个输入，这些输入取自特定的数据对象集合。
- (5) 输出：一个算法具有一个或多个输出，这些输出同输入之间存在某种特定的关系。

算法的含义与程序十分相似，但又有区别。一个程序不一定满足有穷性。例如操作系统，只要整个系统不遭破坏，它将永远不会停止，即使没有作业需要处理，它仍处于动态等待中。因此，操作系统不是一个算法。另一方面，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。算法代表了对问题的解，而程序则是算法在计算机上特定的实现。如果用程序设计语言来描述一个算法，那么它就是一个程序。

算法与数据结构是相辅相成的。解决某一类特定问题的算法可以选定不同的数据结构，而且选择恰当与否直接影响算法的效率。反之，一种数据结构的优劣由各种算法的执行效果来体现。

在算法设计时通常需要考虑以下几个方面的要求。

- (1) 正确性：算法的执行结果应当满足预先规定的功能和性能的要求。正确性要求表明算法必须满足实际需求，达到解决实际问题的目的。
- (2) 可读性：一个算法应当思路清晰、层次分明、简单明了、易读易懂。可读性要求表明算法主要是人与人之间交流解题思路 and 进行软件设计的工具，因此可读性必须强。同时一个可读性强的算法，其程序的可维护性、可扩展性都要好许多，因此，许多时候人们往往在一定程度上牺牲效率来提高可读性。
- (3) 健壮性：当输入不合法数据时，应能做适当处理，不至于引起严重后果。健壮性要求表明算法要全面细致地考虑所有可能出现的边界情况，并对这些边界条件做出完备的处理，尽可能使算法没有意外的情况。
- (4) 高效性：有效使用存储空间和有较好的时间效率。高效性要求主要是指时间效率，即解决相同规模的问题时间应尽可能短。

一般来说，数据结构上的基本操作主要有以下几种。

- (1) 查找：寻找满足特定条件的数据元素所在的位置。
- (2) 读取：读出指定位置上数据元素的内容。