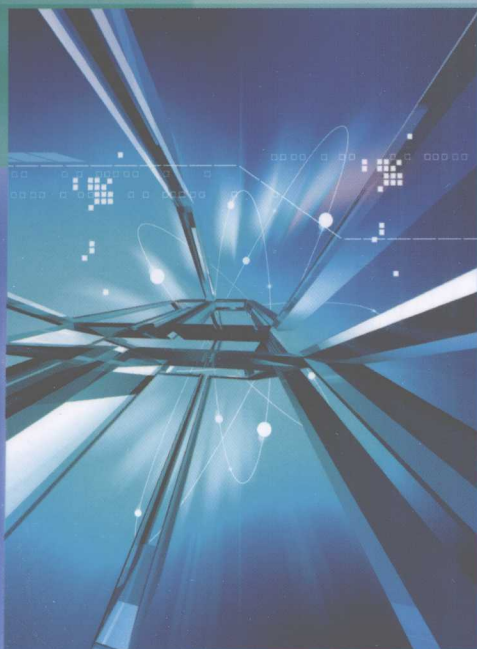


信息管理与信息系统专业规划教材

# 数据结构与算法

—— C语言实现

张千帆 主编



科学出版社

[www.sciencep.com](http://www.sciencep.com)

内容简介

信息管理 with 信息系统专业规划教材

# 数据结构与算法

## ——C语言实现

张千帆 主编

清华大学出版社

ISBN 7-302-10327-2

定价：35.00元

ISBN 7-302-10327-2

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

清华大学出版社

# 科学出版社

北京

清华大学出版社

## 内 容 简 介

本书是为“数据结构”课程编写的教材,主要介绍各种基本类型的数据结构及其算法实现。本书所有算法都有算法功能说明、算法思想分析、详尽的实例描述、C语言编写并可编译执行的完整程序及运行结果图示,典型算法附有算法分析。

本书是数据结构的入门书籍,结构严谨,条理清晰,按照线性数据结构、层次数据结构和网状数据结构的顺序,由易到难介绍主要抽象数据类型及其应用,最后介绍各种查找和排序方法。抽象的数据结构原理与算法实现紧密结合的写作特点使读者能够快速而卓有成效地掌握数据结构原理和经典算法,以加深读者对数据结构和算法的理解,从而提高编程能力。

本书可以作为高等院校信息管理类专业的本科和专科教材,也可以作为其他理工科专业的选修教材或实验指导教材。

### 图书在版编目(CIP)数据

数据结构与算法: C语言实现/张千帆主编. —北京: 科学出版社, 2009  
(信息管理 with 信息系统专业规划教材)

ISBN 978-7-03-024517-5

I. 数… II. 张… III. ①数据结构-高等学校-教材 ②算法分析-高等学校-教材 ③C语言-程序设计-高等学校-教材 IV. TP311.12 TP312

中国版本图书馆 CIP 数据核字 (2009) 第 065874 号

责任编辑: 陈晓萍/责任校对: 耿 耘

责任印制: 吕春珉/封面设计: 耕者设计工作室

科学出版社出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

双青印刷厂印刷

科学出版社发行 各地新华书店经销

\*

2009 年 6 月第 一 版 开本: B5 (720×1000)

2009 年 6 月第一次印刷 印张: 23 1/2

印数: 1—3 000 字数: 471 000

定价: 35.00 元

(如有印装质量问题, 我社负责调换〈双青〉)

销售部电话 010-62134988 编辑部电话 010-62138978-8003

版权所有, 侵权必究

举报电话: 010-64030229; 010-64034315; 13501151303

## 编 委 会

### 顾问

马费成 薛华成 侯炳辉

### 主任

张金隆

### 编委

(以姓氏笔画为序)

王 虎	王学东	刘腾红	孙细明	孙茜雯
何 浩	张唯佳	李 纲	杨新年	肖 华
郑双怡	夏火松	徐绪松	聂规划	曾庆伟
鲁耀斌	雷春明	蔡淑琴	潘开灵	

## 序

国家教育部于1998年7月6日公布了新的《普通高等学校本科专业目录》，将原来的经济信息管理、信息学、科技信息管理、林业信息管理和**管理信息系统**等专业合并为**管理学科门类中的信息管理与信息系统专业**。可以认为，这次合并既是学科相融的必然，也是国家信息化发展的需要。据有关资料介绍，到目前为止，全国已有超过200所高校开设了**信息管理与信息系统专业**。

自20世纪40年代以来，信息技术经过60余年的高速发展，它对人类社会各个领域的**影响越来越广泛和深入**，其影响最大、受益最多的当属**管理和经济领域**。信息作为最主要的经济资源，已经被人们所接受，并且愈来愈受到重视。信息技术的普及和推广，信息资源的组织、开发和利用，促进了企业的发展和产业结构的调整。当前所实施的**电子商务、电子政务和数字图书馆等工程**直接加速了生产力的发展和促进了社会的进步。我国政府提出的“以信息化带动工业化”的战略举措，必将有力提升我国的综合国力，同时也为**信息管理与信息系统专业**带来极大的发展机遇和发展空间。

信息管理与信息系统是一门交叉学科，它不是信息技术和管理科学的简单组合，而需要融合**管理学、经济学、系统科学、运筹学和计算机科学**于一体，因而，必须要有一套具有本专业特点的知识结构体系和适合本专业需要的教材体系。

信息管理与信息系统专业从1998年设立至今的10多年来，许多专家学者在专业建设和教材建设方面倾注了大量的心血，有力地促进了专业和学科的发展。但是，由于该专业具有**跨度大、内容新和变化快**等特点，如何培养适应现代信息技术高速发展需要的、具有创新能力的、既懂信息技术又懂管理的复合型人才，对广大教育工作者而言是一个巨大的挑战。

在科学出版社的直接推动下，在我国信息管理领域的知名学者薛华成教授、侯炳辉教授和马费成教授的指导下，在湖北省信息产业厅和**经济贸易委员会**及相关企业的支持下，**武汉地区**包括华中科技大学、武汉大学、华中师范大学、中南财经政法大学和**武汉理工大学**等20余所高校联合编写了这套针对本科生的**信息管理与信息系统专业规划教材**。

这套教材共22本，除了数学基础类的《运筹学》外，大致可以归为以下3类。

计算机技术类（8本）：《数据库技术》、《计算机网络技术》、《数据结构与算法——C语言实现》、《面向对象的开发方法》、《数据仓库与数据挖掘技术》、《操作系统》、《多媒体信息管理技术基础》和《实用软件工具》。

信息系统类（6本）：《信息系统分析与设计》、《信息系统案例分析》、

《项目管理》、《管理信息系统》、《信息系统原理》和《决策支持系统》。

信息管理类(7本):《信息管理学基础》、《信息资源管理》、《信息经济学》、《信息政策与法规》、《信息组织学》、《信息检索》和《信息安全》。

这套教材具有以下特点:

(1) 内容新。正如前面所指出的一样,这套教材并不是简单地分门别类讲解信息技术和管理科学知识,而是站在信息管理与信息系统专业这个全新的角度上,力求全面、及时反映信息管理与信息系统领域国内外的最新发展和研究成果。

(2) 体系全。为保证本系列教材体系的完整性和内容的系统性,编委会曾多次开会讨论并广泛征求国内信息管理与信息系统领域的有关专家的意见,该套教材主要集中于专业基础课和专业课方面,并考虑了这些课程之间的相互衔接和整体上的协调。

(3) 注重基础。本系列教材从选题到编写均充分考虑到当前我国本科生的知识结构和知识背景及其后续发展的需要,着重于讲解信息管理与信息系统专业的基础知识,注意培养学生的能力。

(4) 结合实际,多采用案例教学。本系列教材的作者都是从事一线教学工作的教师,了解本科生的特点和需求,大多数作者又有从事信息系统开发和信息资源管理的经验,了解实际工作对本专业的需求。因此,在编写过程中作者们能注意理论与实践相结合,通过引入适当的案例和实验,加深学生对理论知识的理解和掌握。

我们希望,这套教材的成功出版,能为推动我国信息管理与信息系统专业教育的发展、促进信息化人才的培养起到积极的作用。

这套教材是我们不同类型的学校,不同专业背景、但同属信息管理与信息系统专业教师合作的一种尝试。我们欢迎信息管理和信息系统及相关专业的教师、学生和科研工作者以及有关人士提出宝贵的意见和建议,以便进一步提高我们的教材质量。

本套规划教材编委会主任

华中科技大学管理学院院长

管理信息研究所所长

张金隆 教授

# 前 言

数据结构是计算机科学中最重要的课程之一，是计算机程序设计的重要理论和实践基础。

对于初学者而言，数据结构的学习有两大难题，一是理解并掌握抽象的数据结构原理与方法，二是把抽象的原理与方法实现出来，以解决实际问题。针对一般数据结构书籍使用伪码，不提供完整算法，给初学者带来不便，本书所有算法都有算法功能说明、算法思想分析、详尽的实例描述、C语言编写的完整程序及运行结果图示，典型算法附有算法分析。本书是数据结构的入门书籍，结构合理，覆盖面广，概念清晰，层次分明，图文并茂。数据结构原理与算法实现紧密结合的写作特点使读者能够快速而卓有成效地掌握数据结构原理和经典算法，以加深读者对数据结构和算法的理解，从而提高编程能力。

第1章为数据结构概述，内容包括数据结构和抽象数据类型等概念及算法复杂度的度量。第2章至第7章介绍线性表、栈、队列、串、数组、广义表、树和二叉树以及图等基本类型的数据结构及其应用。第8章介绍查找的基本概念，各种查找方法及其算法实现与算法分析。第9章介绍排序的概念及主要排序方法的排序原理、算法实现与算法分析。第10章介绍文件的基本概念及常用的文件结构。

本书可以作为高等院校信息管理类专业的本科和专科教材，也可以作为其他理工科专业的选修教材。教师可以根据本学校的专业特点、学生情况和教学学时，选讲部分章节的内容。

在确定教材体系和主要内容过程中，得到了韦司滢老师的帮助。方凌云老师参与了第2、3章的编写工作，程海芳老师参与了第10章的编写工作，漆鹏飞参与了算法实现工作，参与本书编写的还有万元元、毛弘毅、李雄、杨浩、蒋臻和袁乾，在此向他们深表谢意。

由于编者水平有限，本书难免存在缺点和错误，敬请各位专家和读者谅解并批评指正。

张千帆  
2009年3月

# 目 录

第 1 章 绪论	1
1.1 数据结构与程序设计	1
1.1.1 学习数据结构的意义	1
1.1.2 数据与数据结构	3
1.1.3 数据结构类型	4
1.2 抽象数据类型	6
1.2.1 C 语言中的数据类型	6
1.2.2 抽象数据类型	7
1.3 算法分析	8
1.3.1 问题、算法与程序	8
1.3.2 算法效率的度量	9
本章小结	13
思考与练习题	14
第 2 章 线性表	15
2.1 线性表的基本概念	15
2.1.1 线性表的定义与特点	15
2.1.2 线性表的两类存储结构	16
2.2 顺序表的算法实现	18
2.2.1 顺序表的创建	18
2.2.2 顺序表内结点的插入	21
2.2.3 顺序表内结点的查找	25
2.2.4 顺序表内元素的删除	29
2.3 单链表的算法实现	33
2.3.1 单链表的结点结构和一般形式	33
2.3.2 单链表的创建	33
2.3.3 单链表内元素的插入	36
2.3.4 单链表内元素的查找	39
2.3.5 单链表内元素的删除	44
2.3.6 两个单链表的合并	47
2.4 双向链表的算法实现	52
2.4.1 双向链表的结点结构和一般形式	52



2.4.2	双向链表的创建	53
2.4.3	双向链表内元素的插入	55
2.4.4	双向链表内元素的查找	58
2.4.5	双向链表内元素的删除	58
2.5	循环链表的算法实现	62
2.5.1	循环链表的结点结构和一般形式	62
2.5.2	循环链表的创建	62
本章小结		65
思考与练习题		65
<b>第3章</b>	<b>栈与队列</b>	<b>67</b>
3.1	栈的基本概念	67
3.1.1	栈的定义与特点	67
3.1.2	栈的两类存储结构	67
3.2	顺序栈的算法实现	68
3.2.1	顺序栈的建立和入栈	68
3.2.2	顺序栈出栈	70
3.3	队列的基本概念	73
3.3.1	队列的定义与特点	73
3.3.2	队列的存储结构	73
3.4	顺序队列的算法实现	75
3.4.1	顺序队列建立和入队	75
3.4.2	顺序队列出队	77
3.5	循环队列的算法实现	80
3.5.1	循环队列建立和入队	80
3.5.2	循环队列出队	82
3.6	链队列的算法实现	85
3.6.1	链队列建立和入队	85
3.6.2	链队列出队	87
3.7	栈和队列的应用——算术表达式求值	90
本章小结		101
思考与练习题		101
<b>第4章</b>	<b>串</b>	<b>103</b>
4.1	串的基本概念	103
4.1.1	串的定义与特点	103
4.1.2	串的存储结构	104
4.2	串的算法实现	106

4.2.1	串赋值算法	106
4.2.2	求子串算法	108
4.2.3	串比较算法	110
4.2.4	串联接算法	112
4.3	串的模式匹配算法实现	115
4.3.1	串的朴素模式匹配算法	115
4.3.2	改进的模式匹配算法	118
	本章小结	122
	思考与练习题	122
<b>第5章</b>	<b>数组和广义表</b>	<b>124</b>
5.1	数组的基本概念	124
5.1.1	数组的定义与特点	124
5.1.2	数组的存储结构	125
5.2	特殊矩阵的压缩存储	127
5.3	矩阵的算法实现	130
5.4	广义表的基本概念	136
5.4.1	广义表的定义与图形表示	136
5.4.2	广义表的存储结构	137
5.5	广义表的算法实现	138
	本章小结	143
	思考与练习题	144
<b>第6章</b>	<b>树和二叉树</b>	<b>146</b>
6.1	树的基本概念	146
6.1.1	树的定义与基本术语	146
6.1.2	树的表示形式和存储结构	147
6.2	二叉树的基本概念	151
6.2.1	二叉树的定义与性质	151
6.2.2	二叉树的存储结构	153
6.2.3	树、森林和二叉树的转换	155
6.2.4	二叉树的遍历	157
6.3	二叉树算法实现	159
6.3.1	二叉树的建立	159
6.3.2	递归的二叉树前序遍历	161
6.3.3	非递归的二叉树前序遍历	163
6.3.4	递归的二叉树中序遍历	167
6.3.5	非递归的二叉树中序遍历	169

6.3.6	递归的二叉树后序遍历	172
6.3.7	非递归的二叉树后序遍历	174
6.4	哈夫曼树及其应用	178
6.4.1	哈夫曼树与哈夫曼编码	179
6.4.2	哈夫曼算法实现	180
	本章小结	186
	思考与练习题	186
<b>第7章</b>	<b>图</b>	<b>187</b>
7.1	图的基本概念	187
7.1.1	图的定义和术语	187
7.1.2	图的表示与存储结构	191
7.2	图的构造算法实现	195
7.2.1	构造数组存储的图	195
7.2.2	构造邻接表存储的无向图	199
7.2.3	构造邻接表存储的有向图	202
7.2.4	构造十字链表存储的有向图	205
7.2.5	构造邻接多重表存储的无向图	208
7.3	图的遍历算法实现	212
7.3.1	深度优先遍历算法	212
7.3.2	广度优先遍历算法	217
7.4	最小生成树算法实现	224
7.4.1	普里姆算法	224
7.4.2	克鲁斯卡尔算法	231
7.5	图的应用	237
7.5.1	拓扑排序	237
7.5.2	关键路径	244
7.5.3	最短路径	253
	本章小结	259
	思考与练习题	259
<b>第8章</b>	<b>查找</b>	<b>262</b>
8.1	查找的基本概念	262
8.1.1	相关术语	262
8.1.2	查找表结构	263
8.2	顺序查找算法的实现	263
8.3	折半查找算法的实现	266
8.4	分块查找算法	271

8.4.1	索引表	271
8.4.2	分块查找算法实现	272
8.5	二叉排序树及其算法实现	275
8.5.1	二叉排序树及其查找过程	275
8.5.2	二叉排序树插入结点的过程	276
8.5.3	二叉排序树删除结点的过程	277
8.5.4	二叉排序树的算法实现	278
8.6	平衡二叉树及其算法实现	284
8.6.1	平衡二叉排序树及其构造	284
8.6.2	平衡二叉排序树算法实现	288
8.7	B-树及其算法实现	295
8.7.1	B-树	295
8.7.2	B-树的查找	296
8.7.3	B-树的插入	297
8.7.4	B-树的删除	297
8.7.5	B-树的算法实现	300
8.8	哈希查找的算法实现	307
8.8.1	哈希表	307
8.8.2	哈希函数构造方法	307
8.8.3	哈希冲突的处理方法	308
8.8.4	哈希表的算法实现	310
	本章小结	314
	思考与练习题	315
<b>第9章</b>	<b>排序</b>	<b>317</b>
9.1	排序的基本概念	317
9.1.1	术语介绍	317
9.1.2	常用的内容排序算法简介类型	318
9.2	插入排序的算法实现	320
9.2.1	直接插入排序	320
9.2.2	希尔排序	323
9.3	快速排序的算法实现	327
9.4	选择排序的算法实现	331
9.4.1	直接选择排序	331
9.4.2	堆排序	335
9.5	归并排序的算法实现	339
9.6	基数排序的算法实现	342

151	9.7 各种内部排序方法的比较	349
153	9.7.1 时间性能	349
153	9.7.2 空间性能	349
153	9.7.3 排序方法的稳定性	350
156	9.8 外部排序	350
157	本章小结	350
158	思考与练习题	350
	<b>第 10 章 文件</b>	<b>353</b>
181	10.1 文件的基本概念	353
188	10.1.1 文件记录与文件结构	353
189	10.1.2 文件操作	354
189	10.2 文件的存储结构	354
189	10.2.1 顺序文件	354
197	10.2.2 索引文件	355
197	10.2.3 散列文件	356
199	10.2.4 多关键字文件	357
199	10.2.5 倒排序文件	357
197	本章小结	358
197	思考与练习题	358
	<b>参考文献</b>	<b>359</b>

# 第 1 章 绪 论

本章介绍数据结构的基本概念、数据结构的逻辑结构与物理结构、抽象数据类型的表示与实现、算法和算法设计的要求、算法代价的度量，难点是算法效率的时间渐近度分析方法。

## 1.1 数据结构与程序设计

### 1.1.1 学习数据结构的意义

学习一门程序设计语言，记住了该语言的语法、词法及一些常用的函数，就意味着已经掌握了编写程序的基本工具。不过语言只是语言，光靠语言是写不出好的程序的。就好比写一篇论文，首先确定用哪种语言来写？很显然，如果是写给中国人看，就用中文，写给美国人看，当然用英文。那么确定了语言之后，就能写论文了吗？论文的结构如何设计？写作的方法、技巧和注意事项是什么？如果不知道这些，自然写不出好论文。可见，要写一篇好论文，会语言是远远不够的。同样，要写一个好程序，仅会一门语言也是远远不够的。

举例说明，求表 1.1 所示矩阵中所有元素的平均数。

表 1.1 4×4 矩阵

2	0	3	0
0	7	0	0
0	0	1	0
3	0	0	0

第一种设计方法：


```
void main()
{
    int a11,a12,a13,a14;
    int a21,a22,a23,a24;
    int a31,a32,a33,a34;
    int a41,a42,a43,a44; //依次声明矩阵的元素变量
    int sum, average; //声明总和与平均数变量
    a11=2; a12=0; a13=3; a14=0;
    a21=0; a22=7; a23=0; a24=0;
```

```

a31=0; a32=0; a33=1; a34=0;
a41=3; a42=0; a43=0; a44=0; //矩阵赋初值
sum=a11+a12+a13+a14+a21+a22+a23+a24+a31+a32+a33+a34+a41+
    a42+a43+a44;
average=sum/16; //求平均数
printf("矩阵元素的平均数为: %d\n", average); //输出平均数
}

```

程序的运行结果如图 1.1 所示。



```

矩阵平均数为: 1
Press any key to continue

```

图 1.1 程序的运行结果

第二种设计方法:

```

void main()
{
    int a[16]={2,0,3,0,0,7,0,0,0,0,1,0,3,0,0,0};
    //用数组给矩阵赋值
    int sum, average; //声明矩阵总和与平均数变量
    int i; //声明数组下标变量
    sum=0; //设置总和初值
    for(i=0; i<16; i++)
        sum=sum+a[i] ; //计算总和
    average=sum/16; //计算平均数
    printf("矩阵平均数为: %d\n", average); //输出平均数
}

```

程序的运行结果如图 1.2 所示。



```

矩阵平均数为: 1
Press any key to continue

```

图 1.2 程序的运行结果

两种方法都是正确的。但是，第一种方法使用 16 个内存变量存储  $4 \times 4$  矩阵中 16 个位置上的数据，算法的扩充性不是很好。如果不是  $4 \times 4$  的矩阵，而是  $100 \times 100$  的矩阵，这个程序就需要很大的修改。事实上，第二种方法使用数组结构来存储矩阵元素值是一个更好的方法，算法的时空效率都有所提高。

可见，同一个问题的解决方法有多种，采用不同的数据结构，会影响到程序运行的效率和灵活性。当对这些按照一定规律组织的数据进行操作时，理论上存在一种或几种与某种数据结构相匹配的最优（或近似最优）算法，以此来确保操作数据的时空效率。著名的瑞士计算机科学家沃思（N.Wirth）教授曾提出：数据

结构+算法=程序。数据结构和算法设计分析是整个程序设计的基础，是重中之重，这就是学习数据结构这门课程的重要意义。

### 1.1.2 数据与数据结构

计算机是用来处理和存储数据的。数据在计算机中的表示和存储是有规律、有结构的。

#### 1. 数据结构实例

表 1.2 所示的学生信息登记表中，学号、姓名、性别、年龄是数据项 (item)、其取值范围是数据域 (field)；每一行内容是记录 (record)。C 语言将“记录”定义为“结构” (struct)，该学生信息登记表是一个线性表结构。

表 1.2 学生信息登记表

学号	姓名	性别	年龄
2008251001	张平	男	19
2008251002	王晶	男	19
2008251003	李梅	女	20

图 1.3 所示的某高校的院系设置图中，学校、各学院、各系等是结点 (node)；学校与学院，学院与系之间是层次关系或父子关系；该高校的院系设置图是一个树形结构。

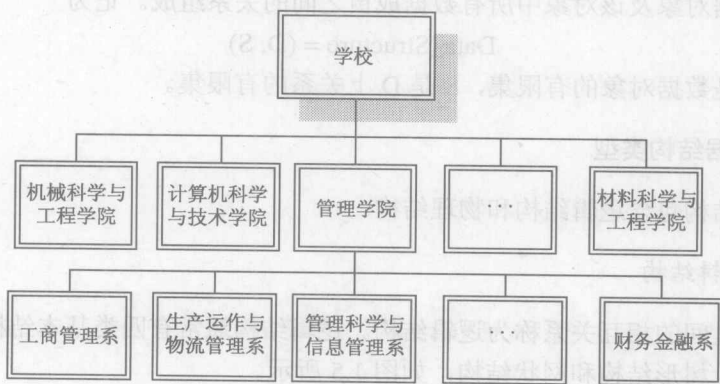


图 1.3 某高校的院系设置图

图 1.4 所示的部分城市之间的飞机直航线路图中，北京、武汉、广州等城市是顶点 (vertex)，两个顶点之间的边 (edge) 表明这一对顶点之间有直航线路，图中任意两个顶点之间都可能有关系，该直航线路图是网状结构。



## 2. 数据结构基本术语

### (1) 数据 (Data)

数据是所有能输入到计算机中并被计算机程序加工、处理的符号的总称。如整数、实数、字符、声音、图像、图形等。

### (2) 数据元素 (Data Element)

数据元素是数据的基本单位。元素、记录、结点、顶点等在计算机程序中通常作为一个整体进行考虑和处理。

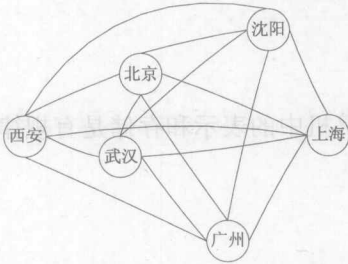


图 1.4 部分城市之间的直航线路图

### (3) 数据项 (Data Item)

数据项是数据不可分割的最小单位。如学号、姓名、性别等。一个数据元素可由一个或多个数据项组成。例如，一个学生的个人信息作为一个数据元素，该数据元素包含学号、姓名、性别等数据项。

### (4) 数据对象 (Data Object)

数据对象是由类型相同的数据元素组成的集合。数据对象是数据的一个子集。例如，由正整数组成的数据对象  $D_1 = \{1, 2, 3, \dots\}$ ，由 26 个字母组成的数据对象  $D_2 = \{A, B, C, \dots, Z\}$ ，其中  $D_1$  是无穷集， $D_2$  是有穷集。

### (5) 数据结构 (Data Structure)

数据结构是相互之间存在一种或多种特定关系的数据元素的集合。数据结构由某一数据对象及该对象中所有数据成员之间的关系组成。记为

$$\text{Data\_Structure} = (D, S)$$

其中， $D$  是数据对象的有限集， $S$  是  $D$  上关系的有限集。

## 1.1.3 数据结构类型

数据结构包含逻辑结构和物理结构。

### 1. 逻辑结构

数据之间的相互关系称为逻辑结构。逻辑结构通常有四类基本结构：集合、线性结构、树形结构和网状结构，如图 1.5 所示。

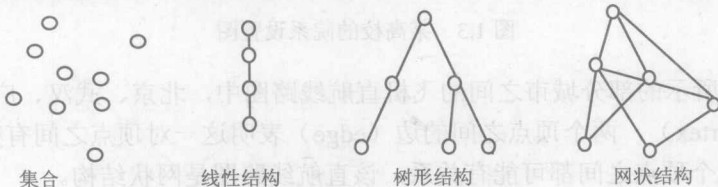


图 1.5 数据结构的四种基本逻辑结构