



教育部高职高专规划教材  
Jiaoyubu Gaozhi Gaozhuan Guihua Jiaocai

高职高专计算机专业系列教材

# 数据结构

张世和 编著 黄国兴 审



学(北京)  
711.12-43  
03



清华大学出版社

<http://www.tup.tsinghua.edu.cn>



教育部高职高专规划教材  
高职高专计算机专业系列教材

# 数 据 结 构

张世和 编著  
黄国兴 审

清华大学出版社

(京)新登字158号

## 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。主要包括:数据结构的基本概念;算法描述和算法分析初步;线性表、堆栈、队列、串、数组、树、图等结构;排序和查找的各种方法;另外还用一章的篇幅详细介绍了链式存储结构以加深读者的理解。每一章后面均列举了典型应用实例,并配有算法和程序以供教学和实践使用。

本书作为“高职高专计算机专业系列教材”之一,主要面向高职高专院校计算机类专业的学生,也可以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自学参考书。

**版权所有,翻印必究。**

**本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。**

## 图书在版编目(CIP)数据

数据结构/张世和编著. —北京:清华大学出版社,2000  
教育部高职高专规划教材. 高职高专计算机专业系列教材  
ISBN 7-302-04065-6

I. 数... II. 张... III. 数据结构-高等学校:技术学校-教材 IV. TP311.12

中国版本图书馆 CIP 数据核字(2000)第 75134 号

出版者:清华大学出版社(北京清华大学学研大厦,邮编 100084)

<http://www.tup.tsinghua.edu.cn>

印刷者:北京密云胶印厂

发行者:新华书店总店北京发行所

开 本: 787×1092 1/16 印张: 12 字数: 280 千字

版 次: 2000 年 12 月第 1 版 2000 年 12 月第 1 次印刷

书 号: ISBN 7-302-04065-6/TP·2394

印 数: 0001~6000

定 价: 16.00 元

# 出版说明

---

教材建设工作是整个高职高专教育教学工作中的重要组成部分。改革开放以来,在各级教育行政部门、学校和有关出版社的共同努力下,各地已出版了一批高职高专教育教材。但从整体上看,具有高职高专教育特色的教材极其匮乏,不少院校尚在借用本科或中专教材,教材建设仍落后于高职高专教育的发展需要。为此,1999年教育部组织制定了《高职高专教育基础课程教学基本要求》(以下简称《基本要求》)和《高职高专教育专业人才培养目标及规格》(以下简称《培养规格》),通过推荐、招标及遴选,组织了一批学术水平高、教学经验丰富、实践能力强的教师,成立了“教育部高职高专规划教材”编写队伍,并在有关出版社的积极配合下,推出一批“教育部高职高专规划教材”。

“教育部高职高专规划教材”计划出版500种,用5年左右时间完成。出版后的教材将覆盖高职高专教育的基础课程和主干专业课程。计划先用2~3年的时间,在继承原有高职、高专和成人高等学校教材建设成果的基础上,充分汲取近几年来各类学校在探索培养技术应用性专门人才方面取得的成功经验,解决好新形势下高职高专教育教材的有无问题;然后再用2~3年的时间,在《新世纪高职高专教育人才培养模式和教学内容体系改革与建设项目计划》立项研究的基础上,通过研究、改革和建设,推出一大批教育部高职高专教育教材,从而形成优化配套的高职高专教育教材体系。

“教育部高职高专规划教材”是按照《基本要求》和《培养规格》的要求,充分汲取高职、高专和成人高等学校在探索培养技术应用性专门人才方面取得的成功经验和教学成果编写而成的,适用于高等职业学校、高等专科学校、成人高校及本科院校举办的二级职业技术学院和民办高校使用。

教育部高等教育司

2000年4月3日

# 高职高专计算机专业 系列教材编审委员会

主任：白英彩

副主任：刘煜海 焦金生

秘书长：高建华 蒋川群

委员（按姓氏笔画顺序）

丁桂芝 白英彩 孙德文 边善裕 石磊

刘煜海 李瑜祥 陈海强 余苏宁 吴永明

林家骏 陶霖 高建华 黄国兴 蒋川群

焦金生 鲍有文

## 序

1999年10月,教育部高教司主持召开了全国高职高专教材工作会议,会议要求尽快组织规划和编写一批高质量的、具有高职高专特色的基础和 专业教材。根据会议精神,在清华大学出版社的支持下,于2000年1月在上海召开了由来自全国各地的部分高职、高专、成人教育及本科院校的代表参加的“高职高专计算机专业培养目标和课程设置体系研讨会”。与会的专家和教师一致认为,在当前教材建设严重滞后同高职教育迅速发展的矛盾十分突出的情况下,编写一套适应高等职业教育培养技术应用性人才要求的、真正具有高职特色的、体系完整的计算机专业系列教材十分必要而且迫切。会议成立了高职高专计算机专业系列教材编审委员会,明确了高职计算机专业的培养目标,即掌握计算机专业有关的基本理论、基本知识和基本技能,尤其要求具有对应用系统的操作使用、维护维修、管理和初步开发的能力。

根据上述目标,编委会拟定了本套教材的编写原则。在教材内容安排上,以培养计算机应用能力为主线,构造该专业的课程设置体系和教学内容体系;从计算机应用需求出发进行理论教学,强调理论教学与实验实训密切结合,尤其突出实践体系与技术应用能力的实训环节的教学;教材编写力求内容新颖、结构合理、概念清楚、实用性强、通俗易懂、前后相关课程有较好的衔接。与本科教材相比,本套教材在培养学生的应用技能上更有特色。

根据目前各高职高专院校计算机专业的课程设置情况,编委会确定了首批出版的十几本教材。这些教材的作者多是在高职高专院校或本科院校的职业技术学院任教的、具有多年教学经验的教师,每本书均由计算机专业的资深教授或专家主审把关。我们还将在在此基础上,陆续征集出版第二、三批教材,力争在3到5年内完成一套完整的高职高专计算机专业教材。

应当说明的是,凡是高等职业教育、高等专科教育和成人高等院校的计算机及其相关专业均可使用本套教材。各学校可以根据实际需要,在教学中适当增删一些内容、实训项目和练习题,从而更有针对性地帮助学生掌握计算机专业知识,并形成相关的应用能力。

由于各地区各学校在教学水平、培养目标理解等方面均有所不同,加上这套教材编写时间的仓促,难免会出现这样或那样的错误,敬请各学校在使用

用过程中及时将错误信息或好的建议返回给教材编审委员会,以便我们即时修订、改版,使该系列教材日趋完善。

我们恳切地希望高职高专院校任课的专业教师和专家对后续教材的编写提出建设性的意见,并真诚地希望各位老师参与我们的工作。

高职高专计算机专业  
系列教材编审委员会  
2000年5月

# 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“高职高专计算机专业系列教材”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”,强调理论知识的理解和运用,实现专科教学以实践体系为主及以技术应用能力培养为主的目标。

本书共分9章。第1章阐述数据、数据结构和算法等基本概念。第2章至第7章分别讨论了线性表、栈、队列、串、数组、树和二叉树以及图等基本数据结构及其应用,其中,第3章专门总结了链式存储结构的基本概念和应用,为学好后面各类数据结构打好扎实的基础。第8,9章讨论了查找和排序的各种实现方法及实用分析。

本教材有以下特点:

(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1,9章)都配有“应用举例和分析”一节,列举分析了很多实用的例子,希望便于教学上的举例,并有助于学生加深对基础理论知识的理解,培养实际应用的能力。

(3) 考虑到此课程的先导课程是“C语言程序设计”,书中所有算法和程序的描述都采用可在计算机上调试运行的C语言函数和程序。这样,降低了算法设计的难度,使学生能更方便地在计算机上验证这些算法。

(4) 最后的附录汇总了本书各章中介绍各类数据结构时用到的数据结构类型说明,供学生在上机时参考使用。

本教材讲课时数为50~60学时,上机时数可灵活安排。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。

本书由上海应用技术学院计算机系张世和副教授编写,华东师范大学



计算机系黄国兴教授主审。书中算法由上海应用技术学院计算机系徐继延老师完成,并全部在 PC 机上用 Turbo C 调试通过。

编写高职高专计算机专业系列教材是一项新的尝试,时间紧迫,难免存在疏漏,敬请读者批评指正。

编者

2000年6月

# 目 录

---

<b>第 1 章 绪论</b> .....	1
1.1 引言 .....	1
1.2 基本概念和术语 .....	3
1.3 算法描述 .....	5
1.3.1 算法的重要特性 .....	5
1.3.2 数据结构上的基本操作 .....	5
1.3.3 算法的描述方法 .....	6
1.4 算法分析 .....	7
1.4.1 算法设计的要求 .....	7
1.4.2 算法时间效率的度量分析 .....	7
<b>第 2 章 线性表</b> .....	11
2.1 线性表的定义及逻辑结构 .....	11
2.2 线性表的基本操作 .....	12
2.3 线性表的顺序存储结构 .....	12
2.4 基本操作在顺序表上的实现 .....	13
2.4.1 顺序表上元素的插入 .....	13
2.4.2 顺序表上元素的删除 .....	15
2.4.3 顺序表上元素的定位 .....	16
2.5 应用举例及分析 .....	16
<b>第 3 章 链式存储结构</b> .....	21
3.1 线性表的链式存储结构 .....	21
3.1.1 单链表上的基本运算 .....	23
3.1.2 循环链表 .....	29
3.1.3 双向链表 .....	29
3.2 线性表的顺序和链式存储结构的比较 .....	31
3.3 应用举例及分析 .....	32

<b>第 4 章</b>	<b>栈和队列</b> .....	37
4.1	栈 .....	37
4.1.1	栈的定义及基本操作 .....	37
4.1.2	栈的顺序存储结构 .....	38
4.1.3	栈的链式存储结构 .....	40
4.2	队列 .....	41
4.2.1	队列的定义及基本操作 .....	41
4.2.2	队列的顺序存储结构 .....	42
4.2.3	队列的链式存储结构 .....	46
4.3	应用举例及分析 .....	48
<b>第 5 章</b>	<b>其他线性数据结构</b> .....	53
5.1	串 .....	53
5.1.1	串的定义及基本操作 .....	53
5.1.2	串的存储结构 .....	54
5.1.3	串的基本操作的实现 .....	56
5.2	多维数组 .....	58
5.2.1	二维数组定义及基本操作 .....	58
5.2.2	二维数组的向量存储结构 .....	59
5.2.3	稀疏矩阵的压缩存储 .....	60
5.2.4	稀疏矩阵的转置算法 .....	61
5.3	应用举例及分析 .....	63
<b>第 6 章</b>	<b>树和二叉树</b> .....	67
6.1	树的定义和基本操作 .....	67
6.1.1	树的定义 .....	67
6.1.2	基本术语 .....	68
6.1.3	树的基本操作 .....	69
6.2	二叉树 .....	69
6.2.1	二叉树的定义和基本操作 .....	70
6.2.2	二叉树的性质 .....	71
6.2.3	二叉树的存储结构 .....	73
6.2.4	遍历二叉树 .....	75
6.3	树和森林 .....	78
6.3.1	树的存储结构 .....	78
6.3.2	树、森林与二叉树的转换 .....	81
6.3.3	树和森林的遍历 .....	83
6.4	哈夫曼树和判定树 .....	84

6.4.1	哈夫曼树的定义及构造方法 .....	84
6.4.2	哈夫曼编码 .....	86
6.4.3	分类与判定 .....	87
6.5	应用举例及分析 .....	89
<b>第7章</b>	<b>图 .....</b>	<b>99</b>
7.1	图的定义和术语 .....	99
7.2	图的存储结构 .....	102
7.2.1	邻接矩阵表示法 .....	102
7.2.2	邻接链表表示法 .....	104
7.3	图的遍历 .....	107
7.3.1	深度优先搜索遍历 .....	107
7.3.2	广度优先搜索遍历 .....	108
7.4	图的应用 .....	110
7.4.1	生成树和最小生成树 .....	110
7.4.2	拓扑排序 .....	113
7.4.3	最短路径 .....	116
7.5	应用举例及分析 .....	118
<b>第8章</b>	<b>查找 .....</b>	<b>125</b>
8.1	基本概念 .....	125
8.2	静态查找表 .....	126
8.2.1	顺序表上顺序查找 .....	126
8.2.2	有序表查找 .....	128
8.2.3	索引顺序表查找 .....	130
8.3	动态查找 .....	131
8.3.1	二叉排序树的生成和插入 .....	132
8.3.2	二叉排序树上的查找 .....	134
8.3.3	二叉排序树的删除 .....	135
8.4	散列表 .....	137
8.4.1	散列表与散列函数 .....	137
8.4.2	散列函数的构造方法 .....	138
8.4.3	解决冲突的主要方法 .....	140
8.4.4	散列表的查找及分析 .....	142
8.5	应用举例及分析 .....	144
<b>第9章</b>	<b>内部排序 .....</b>	<b>152</b>
9.1	基本概念 .....	152

9.2	三种简单排序方法 .....	153
9.2.1	直接插入排序 .....	153
9.2.2	冒泡排序 .....	154
9.2.3	简单选择排序 .....	156
9.3	快速排序 .....	157
9.4	堆排序 .....	160
9.5	归并排序 .....	164
9.6	基数排序 .....	166
9.6.1	多关键字的排序 .....	167
9.6.2	链式基数排序 .....	167
9.7	各种内部排序方法的比较与讨论 .....	169
<b>附录 数据存储类型说明 (datastru. h) .....</b>		<b>171</b>

# 第 1 章

# 绪 论

## 1.1 引 言

电子计算机是 20 世纪科学技术最卓越的成就之一。自 1946 年第一台电子计算机问世以来,计算机产业和应用的发展远远超出了人们对它的预料。如今,计算机的应用已不再局限于科学计算,而更多地用于数据处理、信息管理、实时控制等非数值计算的各个方面。用数字计算机解决任何问题都离不开程序设计。为了编制“好”的程序,必须分析程序处理的数据的特性及数据之间的关系,这就是“数据结构”这门学科形成和发展的背景。

数据结构主要研究非数值应用问题中数据之间的逻辑关系和对数据的操作,同时还研究如何将具有逻辑关系的数据按一定的存储方式存放在计算机内。分析数据之间的逻辑关系和确定数据在计算机内的存储结构是程序设计前两个必须完成的任务。

处理非数值计算问题和数值计算问题的解决方案不同。例如,求解梁架结构中应力的数学模型为线性方程组,预报人口增长情况的数学模型为微分方程。但还有更多的非数值计算问题是无法用数学方程加以描述的。

### 例 1.1 某单位职工档案的管理。

图 1.1 中的职工档案表就是一个数据结构。计算机档案管理的主要功能包括:查找、浏览、插入、修改、删除、统计等。如果把表中的一行看成一个记录并称为一个结点,则在此表中,结点和结点之间的关系是一种最简单的线性关系。

工号	姓名	性别	出生年月	婚否	学历	进厂日期
0001	张丽萍	女	08/21/1962	已	大专	09/01/1984
0005	李小明	男	04/06/1972	未	大学	04/10/1996
0006	王冠英	男	06/06/1942	已	高中	03/12/1961
⋮	⋮	⋮	⋮	⋮	⋮	⋮

图 1.1 某单位职工档案表

例 1.2 某学校教师的名册。虽然可以用例 1.1 中的二维表格将全校教师的名单列出,但采用图 1.2 所示的结构更好。它像一棵根在上而倒挂的树,清晰地描述了教师所在的系和教研组,这样一来可以从树根沿着某系某教研组很快找到某个教师,查找的过程就

是从树根沿分支到某个叶子的过程。类似于树这样的数据结构可以描述家族的家谱、企事业单位中的人事关系,甚至可用树来反映人机下棋的动态过程等。

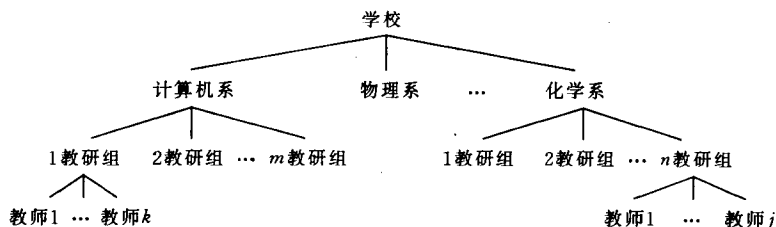


图 1.2 某学校教师名册

**例 1.3** 在  $n$  个城市之间建立通信网络,要求在任意两个城市之间都有直接的或间接的通信线路,在已知某些城市之间直接通信线路预算造价的情况下,使网络的造价最低。

当  $n$  很大时,这样的问题只能用计算机来求解。我们可以用图 1.3(a) 中描述的关系来说明:图中的小圆圈表示一个城市,两个圆圈之间的连线表示对应城市之间的通信线路,连线上的数值表示该通信线路的造价。这一描述的结构为图状结构,利用计算机可以求出满足要求的通信网络,如图 1.3(b) 所示。

通过上面三个例子可以看出:数据结构中元素和元素之间存在着逻辑关系,而线性表,树,图是三种基本的逻辑结构,其他各类的数据结构都是由这三种基本结构派生的。数据结构就是解决如何分析数据元素之间的关系、如何确立合适的逻辑结构、如何存储这些数据,并对为完成数据操作所设计的算法做出时间和空间的分析。“数据结构”在计算机科学中是一门综合性的专业基础课,它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,而且也是设计和实现编译程序、操作系统、数据库系统及大型应用程序的重要基础。

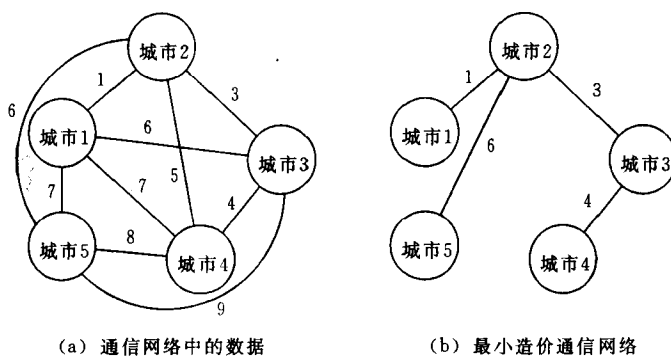


图 1.3 用图描述通信网络问题

简单来说,数据结构是研究程序设计中非数值计算的数据以及它们之间的关系和操作等的一门课程,重点分析数据之间抽象的相互关系,不涉及数据的具体内容。

## 1.2 基本概念和术语

• 数据(data) 指所有能输入到计算机中并被计算机程序处理的符号的总称。计算机输入和处理的数据除数值外,还有字符串、表格、图像甚至声音等,它们都是数字编码范畴。

• 数据元素(data element) 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。一个数据元素可以由若干个数据项组成,也可以只由一个数据项组成。数据元素又被称为元素、结点(node)或记录(record)。

• 数据项(data item) 指数据的不可分割的、含有独立意义的最小单位,数据项有时也称字段(field)或域。上面的职工档案表格是要存放到计算机中进行处理的数据,表中每一行记录了一个职工的档案信息,在数据操作中作为一个整体考虑,对应为一个数据元素,又称为一个记录。这个记录中包含有工号、姓名、学历等若干个数据项。操作的基本单位是记录,如职工的插入或删除一定是作用于一个职工的全部信息即一个记录,而不可能是作用于其中的某个数据项。设想删除操作只删除某个职工的姓名或工号,将引起数据的不完整等严重后果。每个数据项(如职工的姓名或工号)均有独立的含义,但在档案管理这个实际问题中并无完整的意义,而组合在一个记录中构成职工的档案,就具有了完整的实际意义。数据、数据元素、数据项实际上反映了数据组织的三个层次:数据可由若干个数据元素构成,而数据元素又可以由一个或若干个数据项组成。

• 数据逻辑结构(data logical structure) 数据结构主要是研究数据元素之间的关联方式。数据元素之间存在的一种或多种特定的关系被称为数据的逻辑结构。通常有四类基本结构。见图 1.4。

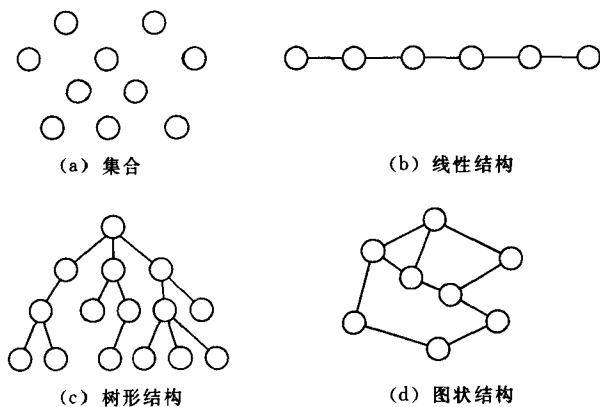


图 1.4 四类基本逻辑结构的示意图

• 数据物理结构(data physical structure) 数据在计算机中的存放方式称为数据的物理结构,又称存储结构。数据的存储结构是逻辑结构在计算机存储器中的实现。



数据元素在计算机中主要有两种不同的存储方法：顺序存储结构和链式存储结构。顺序存储的特点是在内存中开辟一组连续的空间（高级语言中的数组）来存放数据，数据元素之间的逻辑关系通过元素在内存中存放的相对位置来确定，又称向量存储。链式存储的特点是通过指针反映数据元素之间的逻辑关系，又称动态存储。

数据的逻辑结构和物理结构是数据结构的两个密切相关的方面，同一逻辑结构可以对应不同的存储结构。算法的设计取决于数据的逻辑结构，而算法的实现依赖于指定的存储结构。

例如，10 以内的奇数 1,3,5,7,9 用顺序存储结构的方式依次存放在以 300 为起地址的内存向量中，并且两个字长存放一个奇数，如图 1.5(a)所示。在顺序存储结构中，如要读取第三个奇数，它的地址可以通过起地址和要读取奇数的位置序号计算得到。若本例中的奇数改用链式存储结构存放，那么，第一个奇数存放在地址为 300 的内存单元中，第二个奇数存放的地址和第一个奇数存放的地址无关，但第二个奇数所在的地址存放在第一个奇数相关的指针中，后面奇数的存放也按如此的规律。设两个字长存放一个奇数，一个字长存放一个指针，如图 1.5(b)所示。在链式存储结构中，如要读取第三个奇数，只能从第一个奇数所在的地址开始，通过第一个奇数关联的指针得到第二个奇数存放的地址 105，再找到第三个奇数存放的地址 400，才能读出。从以上的例子和分析看，数据结构主要就是研究数据的逻辑结构、相应的存储结构以及完成数据操作的算法设计。

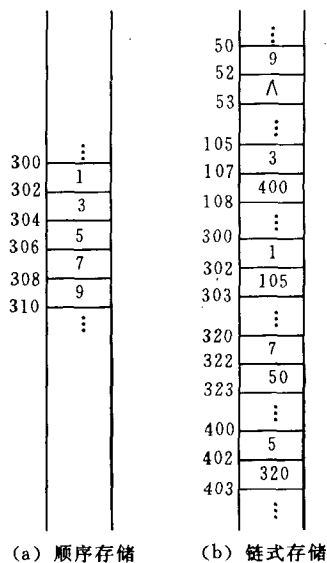


图 1.5 两种存储结构的示意图

- **数据类型 (data type)** 和数据结构密切相关的一个概念，在用高级程序设计语言编写的程序中，每个变量、常量或表达式都对应一个确定的数据类型。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，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对应的数据类型，进一步在确定的数据类型上实现各种操作。

- **算法 (algorithm)** 指解决特定问题的一种方法或一种描述。

1968 年，美国唐·欧·克努特教授开创了数据结构的最初体系，他所著的《计算机程序设计技巧》第一卷《基本算法》是第一部分系统地阐述数据的逻辑结构、存储结构及算法的著作。20 世纪 70 年代初，大型程序出现，软件业飞速发展，结构化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