



21世纪高校计算机应用技术系列规划教材

谭浩强 主编

实用数据结构基础

陈元春 张亮 王勇 编著

★本书以“数据结构实验演示系统”

为主干线来组织编写教材，

采用“以应用为目的，以够用为度”的原则，

从实际应用的需要出发，

大胆取舍，

注重实用性。

★论述简明扼要且条理清楚，

集教学和实验指导于一体。

★习题难易适当，

题型丰富；

注重算法和实现过程，

适合于大学本科和高职高专计算机专业的学生。



中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE



谭浩强 主编

21 世纪高校计算机应用技术系列规划教材

实用数据结构基础

陈元春 张亮 王勇 编著

中国铁道出版社

2003·北京

(京)新登字 063 号

内 容 简 介

数据结构是计算机及其相关专业的一门重要的专业基础课程。它不仅是计算机程序设计的理论基础,而且是学习计算机操作系统、编译原理、数据库原理等课程的重要基础。本书以“数据结构实验演示系统”(共 50 多个算法)为主线来系统地全面地介绍了线性表、栈、队列、串、树、图、查找和排序。每章后面都有一个相对独立的子系统实验以巩固各章的重点算法。本书最后还简单介绍了在做软件开发时所应注意的系统设计的要求、文件的包含处理及主控模块的设计等知识,为如何完成一个完整的数据结构实验系统和以后学习软件工程打下基础。

本书对概念和原理的阐述通俗易懂,例子翔实;习题难易适当,题型丰富;注重算法分析和实现过程。适合于作为大学本科和高职高专计算机专业数据结构学科的教材,也可作为计算机应用的工程技术人员的参考书。

图书在版编目(CIP)数据

实用数据结构基础/陈元春,张亮,王勇编著. —北京:中国铁道出版社,2003.8

(21世纪高校计算机应用技术系列规划教材)

ISBN 7-113-05456-0

I. 实… II. ①陈… ②张… ③王… III. 数据结构-高等学校-教材 IV. TP311.12

中国版本图书馆 CIP 数据核字(2003)第 075743 号

书 名:实用数据结构基础

作 者:陈元春 张 亮 王 勇

出版发行:中国铁道出版社(100054,北京市宣武区右安门西街8号)

策划编辑:严晓舟 魏 春

责任编辑:苏 茜 夏华香 王占清

封面设计:孙天昭

印 刷:北京兴顺印刷厂

开 本:787×1092 1/16 印张:15.75 字数:368千

版 本:2003年9月第1版 2003年9月第1次印刷

印 数:1~5000册

书 号:ISBN 7-113-05456-0/TP·1020

定 价:24.00元

版权所有 侵权必究

凡购买铁道版的图书,如有缺页、倒页、脱页者,请与本社计算机图书批销部调换。

21 世纪高校计算机应用技术系列规划教材

编委会名单

主 任： 谭浩强

副主任： 陈维兴 严晓舟

委 员： （以下排名按姓氏字母的先后顺序为序）

安淑芝 安志远 侯冬梅 李雁翎 吕凤翥

秦建中 宋 红 宋金珂 孙中胜 魏 春

魏善沛 熊伟建 薛淑斌 赵乃真 訾秀玲

丛 书 序 言

21 世纪是信息技术高度发展并且得到广泛应用的时代,信息技术深刻地改变了人类的生活、工作和思维方式。每一个人都应当学习信息技术、应用信息技术。人们平常习惯说的计算机教育其内涵实际上已经发展为信息技术教育,内容主要包括计算机和网络的基本知识和应用。

对大多数人来说,学习计算机的目的是为了利用计算机这个现代化工具去处理工作和面临的各种问题,使自己能够跟上时代前进的步伐,同时要在学习的过程中努力培养自己的信息素养,使自己具有信息时代所要求的科学素质,站在信息技术发展和应用的前列,推动我国信息技术的发展。

学习计算机课程,有两种不同的方法,一是从理论入手;一是从实际应用入手。不同的人有不同的学习内容和学习方法。大学生中的多数人将来是各行各业中的计算机应用人才。对他们来说,不仅需要解决**知道什么**,更重要的是**会做什么**。因此要以应用为目的,注重培养应用能力,大力加强实践环节,激励创新意识。

根据实际教学的需要,我们组织编写这套“**21 世纪高校计算机应用技术系列规划教材**”。顾名思义,这套丛书的特点是突出应用技术,面向实际应用。在选材上,根据实际应用的需要决定内容的取舍,坚决舍弃那些现在用不到、将来也用不到的内容。在叙述方法上,采取“**提出问题——介绍解决问题的方法——归纳结论和概念**”的三部曲,这种从实际到理论、从具体到抽象、从个别到一般的方法,符合人们的认识规律,实践证明已取得了很好的效果。

本丛书采取模块化的结构,根据需要确定一批书目,也就是提供一个课程菜单供各校选用,以后根据信息技术的发展和教学的需要,不断地补充和调整。只要教学有需要,我们就组织编写新的教材,不受任何框框的限制。我们的指导思想是面向实际,面向应用,面向对象。这样比较灵活,能满足不同学校、不同专业的需要。希望各校的老师把你们的要求反映给我们,我们将会尽最大努力满足大家的要求。

本丛书可以作为大学计算机应用技术课程教材以及高职高专、成人高校和面向社会的培训班的教材,也可作为学习计算机的自学教材。

参加本丛书策划和编写工作的专家和老师有:谭浩强、陈维兴、严晓舟、薛淑斌、秦建中、安淑芝、安志远、赵乃真、吕凤翥、李雁翎、宋红、周永恒、熊伟建、宋金珂、陈元春、冯继生、姚怡、沈洪、沈添、李尊朝、王晓敏、侯冬梅、瞿秀玲、魏善沛、孙中胜、王丙义、程爱民、史秀璋、李振银、刘涛、李宁等。此外参加本丛书编辑和其他工作的还有:魏春、秦绪好、张艳芳、戴薇、郭晓溪、马建、姜淑静、姜天鹏、杨东晓、于静等。对于他们的智慧、奉献和劳动表示深切的谢意。中国铁道出版社以很高的热情和效率组织了丛书的出版工作。在组织编写出版的过程中,得到全国高等院校计算机基础教育研究会和各高等院校老师的热情鼓励和支持,对此谨表衷心的感谢。

本丛书如有不足之处,请各位专家、老师和广大读者不吝指正。

谭浩强谨识

2003 年 2 月于清华园

前 言

数据结构是计算机及相关专业的一门重要的专业基础课程。它不仅是计算机程序设计的理论基础，而且是学习计算机操作系统、编译原理、数据库原理等课程的重要基础。

数据结构的主要任务是讨论数据的各种逻辑结构和数据在计算机中的存储表示，以及各种非数值运算的算法的实现。通过数据结构课程的学习，使学生能使用数据结构的基本分析方法来提高编写程序的能力和应用计算机解决实际问题的能力。

本书编写的定位是大学本科和高职、高专的计算机专业的学生，采用“以应用为目的，以够用为度”的原则，从实际应用的需要出发，大胆取舍，注重实用性。

从体系结构而言，本书是用一个“数据结构实验演示系统”为主干线来组织教材的编写的。每一章的主要算法构成一个相对独立的子系统（即子模块），子系统既是各章教学的重点内容，也是上机实验的主要算法。各个子系统可以通过菜单的选择对本章的基本算法进行实验和演示，也可以用它来检验相关习题的正确性。而系统又是开放式的，对于学有余力的同学，可以将数据结构的其它算法扩充到整个实验演示系统中去。

从编写风格而言，本书力求做到简明扼要，条理清楚，并尽量避免抽象的理论论述和复杂的公式推导。本书集教学和实验指导于一体，使得使用本书的学生和自修的读者，一册在手就能方便地进行数据结构课程的学习和实验训练。如果读者在使用本书的过程中需习题答案及“数据结构实验演示系统”，可向中国铁道出版社计算机图书项目中心或任课老师索取。

本书内容共分 10 章，第 1 章绪论，介绍了数据结构与算法的基本概念，并对算法的时间复杂度和空间复杂度做了介绍；第 2 章到第 5 章，介绍了线性表、栈、队列、串等线性结构的逻辑特征，存储方法以及常用算法的实现及基本应用；第 6 章到第 7 章，介绍了树和图两种非线性数据结构的逻辑特征、存储方法以及相关算法的实现和基本应用；第 8 章，主要介绍了顺序查找、二分查找、分块查找和二叉排序树的查找方法以及散列存储的基本方法；第 9 章，介绍了在计算机中广泛使用的各种排序方法，并对各种排序算法的优劣进行了分析和比较。各章内容相对独立，自成体系；每章都有明确的实验目的和实验要求，供学生上机实验之用，在实验参考程序中给出了各章子系统的源代码。书中各章子系统的实验均给出了完整的源程序，并全部在 VC++ 环境中上机运行通过。由于篇幅所限，本书大部分算法都是以单独的函数形式给出的，若读者要运行这些算法，还必须给出一些变量的说明及主函数来调用所给的函数。

本书的第 10 章为“系统的开发”，提出了系统设计的要求，文件包含处理方法及主控模块的设计。实际上它相当于数据结构的一个实训课题，让学生在完成各章子系统的前提下，再设计一个主控模块（即主菜单），来调用各个子系统。通过实训，使学生在原有各章子系统的基础上，组装成一个完整的数据结构实验系统，从而使学生在学好基本算法的基础上，逐步建立起系统的概念。

本书由计春雷副教授、曾宪文副教授和沈学东老师策划，本书第 1 章、第 3 章、第 4 章、

第6章、第10章由陈元春执笔，第2章、第7章、第8章由张亮执笔，第5章、第9章由王勇执笔。实验和习题指导部分由陈元春和张亮合编，并由张亮调试了整个“数据结构实验演示系统”。最后由陈元春完成全书的统稿、修改和定稿工作。王淮亭副教授、刘新铭副教授和郑君华老师审阅了全书的内容，并提出了许多宝贵的修改意见，费宏慧老师绘制了本书的大量图画。另外，还有陈贤淑、陈晓娟、廖康良等参与了本书的编排工作，在此一并表示感谢！

由于作者水平有限，书中的错误或不当之处在所难免，恳请广大专家和读者不吝赐教。同时，我们也会在适当的时间对本书的内容进行修订和补充，并发布在天勤网站：<http://www.tqbooks.net>“图书修订”栏目中。

编者

2003年8月

目 录

第 1 章 绪论	1
1-1 什么是数据结构	2
1-1-1 从数据结构实验演示认识数据结构	2
1-1-2 数据结构研究的内容	3
1-2 数据的逻辑结构	5
1-2-1 基本概念	5
1-2-2 逻辑结构的描述	6
1-3 数据的存储结构	8
1-4 算法和算法分析	9
1-4-1 算法特性	9
1-4-2 算法的效率	10
1-4-3 算法效率的评价	11
小结	12
实验 1	13
习题 1	16
第 2 章 线性表	19
2-1 线性表的定义与运算	20
2-1-1 线性表的定义	20
2-1-2 线性表的基本操作	21
2-2 线性表的顺序存储	22
2-2-1 顺序表	22
2-2-2 顺序表上基本运算的实现	23
2-3 线性表的链式存储	27
2-3-1 线性链表	27
2-3-2 线性链表上基本运算的实现	29
2-3-3 循环链表	35
2-3-4 双向链表	36
小结	38
实验 2 线性表子系统	38
习题 2	43
第 3 章 栈	47
3-1 栈的定义和运算	48
3-1-1 栈 (Stack) 的定义	48



3-1-2 栈的运算	48
3-2 栈的存储和实现	49
3-2-1 顺序栈	49
3-2-2 链栈	52
3-3 栈的应用举例	53
3-3-1 数制转换	53
3-3-2 表达式求值	54
3-3-3 子程序调用 (Subroutine Call)	57
3-3-4 递归调用	58
3-3-5 中断处理和现场保护	59
小结	60
实验 3 栈子系统	60
习题 3	65
第 4 章 队列	67
4-1 队列的定义和基本运算	68
4-1-1 队列 (Queue) 的定义	68
4-1-2 队列的基本运算	68
4-2 队列的存储实现及运算实现	69
4-2-1 顺序队列	69
4-2-2 链队列	73
4-3 队列应用举例	75
小结	77
实验 4 队列子系统	77
习题 4	82
第 5 章 串	85
5-1 串的定义和基本运算	86
5-1-1 串的定义	86
5-1-2 串的输入与输出	87
5-1-3 串的基本运算	87
5-2 串中表示和实现	88
5-2-1 定长顺序存储	89
5-2-2 链接存储	90
5-2-3 串的堆分配存储结构	91
5-3 串的基本运算	92
小结	96
实验 5 串子系统	96
习题 5	101

第 6 章 树和二叉树	103
6-1 树的定义和术语	104
6-1-1 树的定义	104
6-1-2 基本术语	105
6-2 二叉树	106
6-2-1 二叉树的定义	106
6-2-2 二叉树的性质	107
6-2-3 二叉树的存储	108
6-3 遍历二叉树和线索二叉树	112
6-3-1 遍历二叉树	112
6-3-2 恢复二叉树	115
6-3-3 线索二叉树	117
6-4 二叉树的转换	118
6-4-1 一般树转换为二叉树	118
6-4-2 森林转换为二叉树	120
6-4-3 二叉树转换为树和森林	121
6-5 二叉树的应用	122
6-5-1 二叉树的基本应用	122
6-5-2 标识符树与表达式	124
6-6 哈夫曼树及其应用	125
6-6-1 哈夫曼树的引入	125
6-6-2 哈夫曼树的建立	128
6-6-3 哈夫曼编码	130
小结	133
实验 6 树子系统	134
习题 6	142
第 7 章 图	147
7-1 图的定义和术语	148
7-1-1 图的定义	148
7-1-2 图的相关术语	148
7-1-3 图的基本操作	150
7-2 图的存储表示	151
7-2-1 邻接矩阵	151
7-2-2 邻接表	152
7-3 图的遍历	154
7-3-1 深度优先搜索	155
7-3-2 广度优先搜索	156
7-4 图的连通性	157



实用数据结构基础

7-4-1 无向图的连通分量和生成树	157
7-4-2 最小生成树	159
7-5 最短路径	161
小结	162
实验 7 图子系统	163
习题 7	168
第 8 章 查找	171
8-1 查找的基本概念	172
8-2 静态查找表	173
8-2-1 顺序查找	173
8-2-2 二分查找	174
8-2-3 分块查找	178
8-3 动态查找表	178
8-3-1 二叉排序树	178
8-3-2 平衡二叉树 (AVL 树)	184
8-4 哈希表	184
8-4-1 哈希表与哈希方法	184
8-4-2 哈希函数的构造方法	185
8-4-3 处理冲突的方法	186
小结	188
实验 8 查找子系统	189
习题 8	195
第 9 章 排序	199
9-1 概述	200
9-2 插入排序	201
9-2-1 直接插入排序	201
9-2-2 二分插入排序 (Binary Inserting Sort)	203
9-2-3 希尔排序 (Shell's Sort)	204
9-3 快速排序法	205
9-3-1 冒泡排序 (Bubble Sort)	205
9-3-2 快速排序 (Quick Sort)	207
9-4 选择排序	209
9-4-1 简单选择排序	209
9-4-2 树形选择排序	211
9-4-3 堆排序 (Heap Sort)	211
9-5 归并排序	214
9-6 各种排序方法的比较	215
小结	216

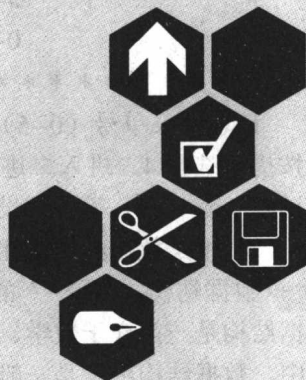
实验 9 排序子系统.....	216
习题 9	224
第 10 章 系统的开发.....	227
10-1 系统设计的要求	228
10-1-1 系统总体结构设计	228
10-1-2 系统详细设计	231
10-1-3 系统调试和维护	232
10-2 文件的包含处理	233
10-2-1 什么是文件包含	233
10-2-2 如何对“数据结构实验系统”进行文件包含	235
10-2-3 数据结构实验系统主控模块	236
10-3 数据结构实验演示系统课程设计任务书.....	237
主要参考文献.....	239

1

绪 论

本章内容概要:

自从世界上第一台电子计算机诞生开始,特别是近 20 年来计算机技术的飞速发展与应用已远远超出人们对它的预料。计算机技术已成为现代化发展的重要支柱和标志,并逐步渗透到人类生活的各个领域。随着计算机硬件的发展,对计算机软件的发展也提出了越来越高的要求。由于软件的核心是算法,而算法实际上是对加工数据过程的描述,所以研究数据结构对提高编程能力和设计高性能的算法是至关重要的。





1-1 什么是数据结构

1-1-1 从数据结构实验演示认识数据结构

我们先来看一个简易的数据结构实验演示系统。

在 Windows 2000 操作系统下运行本书提供的 DS.EXE 文件，就会出现：

数据结构实验演示系统

主菜单

```

*****
*           1-----线   性   表           *
*           2-----   栈                   *
*           3-----队   列                   *
*           4-----   串                   *
*           5-----二   叉   树           *
*           6-----   图                   *
*           7-----查   找                   *
*           8-----排   序                   *
*           0-----退   出                   *
*****

```

请选择菜单号 (0-8):

按提示进行选择，例如选择 2，即进入栈子系统：

栈子系统

```

*****
*           1-----进   栈                   *
*           2-----出   栈                   *
*           3-----显   示                   *
*           4-----数 制 转 换           *
*           5-----逆 波 兰 式           *
*           0-----返   回                   *
*****

```

请选择菜单号 (0-5):

再按提示选择 4，进入二进制—十进制转换的演示……

学过 C (或 C++) 等程序设计的人对结构化程序设计的一些特点应有一定的了解，但是对于数据，特别是数据的结构往往缺乏更深层次的认识。著名的计算机科学家 N.Wirth 提出“算法 + 数据结构 = 程序”的思想，明确地指出了数据结构实际上是程序的主要部分。

数据结构是一门介于数学、计算机硬件和计算机软件三者之间的一门核心课程。在计算机科学中，数据结构不仅是一般非数值计算程序设计的基础，而且是设计和实现汇编语言、编译程序、操作系统、数据库系统，以及其他系统程序和大型应用程序的重要基础。打好“数

据结构”这门课程的扎实基础，将会对程序设计有进一步的认识，使编程能力更上一个台阶，从而使自己学习和开发应用软件的能力有一个明显的提高。

从我国计算机教学现状来看，数据结构不仅仅是计算机专业教学计划中的核心课程之一，而且已经逐步成为非计算机专业的重要选修课程。

1-1-2 数据结构研究的内容

用计算机解决一个具体问题时，大致需要经过以下几个步骤：

- (1) 从具体问题抽象出适当的数学模型；
- (2) 设计求解数学模型的算法；
- (3) 编制程序、运行并调试程序，直到解决实际问题。

寻求数学模型的实质是分析问题，从中提取操作的对象，并找出这些操作对象之间的关系，然后用数学语言加以描述。

下面请看几个例子：

【例 1-1】学生入学情况登记简表

表 1-1 学生入学情况登记简表

学 号	姓 名	性 别	入学总分
01	丁一	男	440
02	马二	男	435
03	张三	女	438
04	李四	男	430
05	王五	女	445
06	赵六	男	428
07	钱七	女	432
08	孙八	男	437
09	冯九	女	426
10	郑十	女	425

我们把表 1-1 称为一个数据结构，表中的每一行是一个结点（或称为记录），它由学号、姓名、性别、入学总分等数据项组成。在这个表中，第一个记录没有直接前趋，称为开始结点；最后一个记录没有直接后继，称之为终端结点。除了第一个记录和最后一个记录以外，其余的记录，都只有一个直接前趋记录和一个直接后继记录。这些结点之间的关系是“一对一”的关系，它构成了“学生入学情况登记简表”的逻辑结构。

那么，“学生入学情况登记简表”在计算机的外存中又如何存储的呢？表中各元素是存在存储器中连续的存储单元（顺序存储），还是用指针链接，存储在不连续的存储单元（链接存储）呢？它构成了“学生入学情况登记简表”的存储结构。

如何在“学生入学情况登记简表”中查找记录、插入记录、删除记录以及如何对记录进行排序、统计、分析等，构成了数据的运算。

数据结构就是研究数据的逻辑结构、存储结构和运算方法（即算法）的学科。



【例 1-2】井字棋对奕问题

图 1-1 (a) 是井字棋对奕过程中的一个格局，任何一方只要使相同的三个子连成一条直线（可以是一行、一列或一条对角线）即为胜方。如果下一步由“×”方下，可以派生出五个子格局，如图 1-1 (b)；随后由“○”方接着下，每个子格局又可以派生出四个子格局……

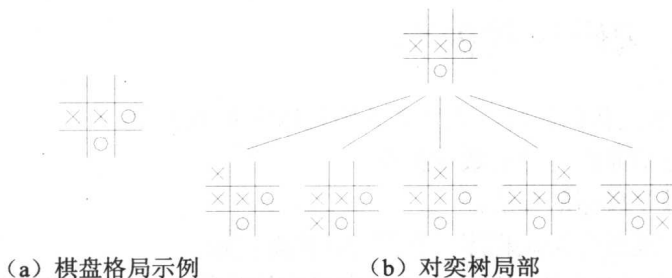


图 1-1 井字棋对奕“树”

若将从对奕开始到结束的过程中所有可能的格局画在一张图上，即形成一棵倒挂的对奕“树”。“树根”是对奕开始时的第一步棋，而所有“叶子”便是可能出现的结局，对奕过程就是从树根沿树叉到某个叶子的过程。在本例中，对奕开始之前的棋盘格局没有直接前趋，称为开始结点（即根），以后每走一步棋，都有多种应对的策略，结点之间存在着“一对多”的关系，它构成了井字棋对奕的逻辑结构。

【例 1-3】七桥问题

位于俄罗斯境内的哥尼斯堡有一条小河叫勒格尔河，河有两条支流，一条叫新河，一条叫老河，它们在市中心汇合，在合流的中间有一座小岛，在小岛和两条支流上建有七座桥。于是哥尼斯堡可分为四块陆地，四周均为河流，四块陆地由七座桥把它们相连。设四块陆地分别为：A、B、C、D，可以用图 1-2 所示的示意图表示。

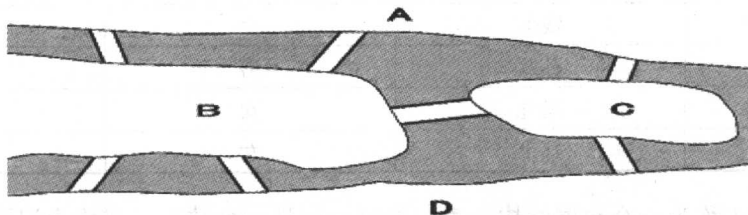


图 1-2 七桥问题

哥尼斯堡的居民有个传统习惯，星期天沿着城市的河岸和小岛散步，同时试图找到一条路线，可经过所有七桥但又不重复经过任意一座桥，这就是著名的“七桥问题”。1736 年，正在哥尼斯堡的瑞士数学家欧拉对“七桥问题”产生了兴趣。他化繁为简，把四块陆地和七座桥抽象为四个点和七条线组成的几何图形，如图 1-3 所示，人们称之为欧拉回路。

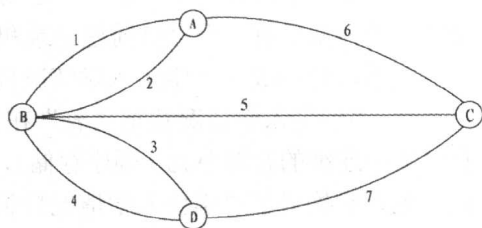


图 1-3 欧拉回路

欧拉对“七桥问题”的结论是：“所有结点的度（一个结点拥有的边数称为度）均为偶

数时，该问题才有解”。反之，“七桥问题”永远无解。在此，我们无意讨论这一数学证明问题，而仅对图 1-3 中，每一个结点都有多个直接前趋和多个直接后继这样一个结果产生兴趣，即那些结点之间存在“多对多”的关系，我们称之为图。

综上所述，非数值计算问题的数学模型不再是数学方程的问题，而是诸如上述的表（如【例 1-1】）、树（如【例 1-2】）、图（如【例 1-3】）之类的数据结构。因此，简单地说，数据结构是一门研究非数值计算的程序设计问题的学科，主要研究数据的逻辑结构、存储结构和算法。本章以下的三个小节就是介绍这三个方面的内容。

1-2 数据的逻辑结构

1-2-1 基本概念

1. 数据（Data）

数据是信息的载体，是对客观事物的符号表示。通俗地说，凡是能被计算机识别、存取和加工处理的符号、字符、图形、图像、声音、视频信号等一切信息都可以称为数据。

在计算机科学中，所谓数据就是计算机加工处理的对象，它可以是数值数据，也可以是非数值数据。数值数据包括：整数、实数、浮点数或复数等，主要用于科学计算、金融、财会和商务处理等；非数值数据则包括：文字、符号、图形、图像、动画、语音、视频信号等。随着多媒体技术的飞速发展，计算机处理的非数值数据已越来越多。

2. 数据元素（Data Element）

数据元素是对现实世界中某个独立个体的数据描述，是数据的基本单位。

数据元素也称为结点（Node），在计算机中，常作为一个整体来处理，如表 1-1 中的某一行，即是一条记录。数据元素在 C（或 C++）语言中可以用结构体来描述，每个数据项则是结构体中的一个分量。

3. 数据项（Data Item）

数据项是数据不可分割的、具有独立意义的最小数据单位，是对数据元素属性的描述。数据项也称为域或字段（Field）。

数据项一般有名称、类型、长度等属性。在 C（或 C++）语言中数据的类型有：整型、实型、浮点型、字符型、指针型等。

数据、数据元素、数据项反映了数据组织的三个层次，即数据可以由若干个数据元素组成，数据元素又由若干个数据项组成。

4. 数据对象（Data Object）

数据对象是性质相同的数据元素的集合，是数据的一个子集。例如在“学生入学情况登记简表”中，数据对象就是全体学生记录的集合。