

走
实
践
应
用
案
例
教
学
之
路
·
培
养
技
能
型
紧
缺
人
才

21 世纪高职高专计算机教育规划教材



数 据 结 构

科海策划

乔国荣 主编

- ◆ 知识讲解
- ◆ 课堂习题
- ◆ 上机实训
- ◆ 课程设计
- ◆ 网络资源

 科 学 出 版 社

数 据 结 构

乔国荣 主编

科 学 出 版 社

内 容 简 介

本书是高职高专计算机专业的基础教材。

本书以 C 语言为程序设计语言。第 1~6 章,采用循序渐进的方式,引导读者逐步掌握线性表、栈和队列、串和数组、树与二叉树、图等数据结构;第 7~8 章,系统地讲解了查找和排序的各种实现方法;第 9 章是课程设计指导,给出了设计题目、设计思想和设计实例,这非常有助于教师指导学生完成课程设计任务,也是本书的重要特色。讲解过程中,每一章除了介绍学生必须掌握的知识点,还给出了教学目标、课堂习题及实训题,这使学生不但了解重点掌握的内容,也可巩固所学知识,提高其应用能力。

本书内容通俗易懂,概念清楚,简明实用,既可作为高职高专和应用型本科院校计算机应用及相关专业的教材,还可供计算机专业相关人员使用。

图书在版编目(CIP)数据

数据结构/乔国荣编著.

—北京:科学出版社,2006

ISBN 7-03-017876-9

I. 数… II. 乔… 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(2006)第 098404 号

责任编辑:刘秀青 / 责任校对:刘雪莲
责任印刷:科海 / 封面设计:林陶

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>

北京市艺辉印刷有限公司印刷

科学出版社发行 各地新华书店经销

2006 年 9 月第一版

开本:16 开

2006 年 9 月第一次印刷

印张:16

印数:1-4000

字数:389 千字

定价:25.00 元

(如有印装质量问题,我社负责调换)

本书编委会

主 编：乔国荣

副主编：吕振凯 毕万新 袁仁水

李文锋 宫明明

参 编：翟 慧 张 震 刘 佳

赵海琳 苗专生 刘 燕

前 言

《数据结构》是计算机专业一门重要的专业基础课，是计算机学科的核心课程。通过数据结构课程的学习，应使学生能运用数据结构的知识和技巧更好地进行算法和程序的设计，也为后继课程的学习打下良好的基础。

全书共分为9章。第1章介绍了数据结构和算法的基本概念和常用术语；第2~6章介绍了基本的数据结构，如线性表、队列、栈、串、数组、广义表、树和图结构，分别讨论了这几种类型数据的逻辑结构和存储结构，以及相应的算法；第7~8章介绍了几种常用的查找和排序方法；最后一章是本书的特色，增加了课程设计指导的内容，使学生在学完基本知识的同时，能够综合利用所学知识完成一些实际课题的设计与制作。书中的程序大部分都可以直接上机运行，其余小部分程序只须稍加改动即可上机运行。另外，为了便于教学，每章后面还配有习题和实训题，并给出了所有习题及实训题的相应参考答案。全书概念表述清楚、简洁，内容由浅入深，强调实践环节，利于教学和自学。

本书可作为高等职业学校、高等专科学校、成人高等学校及本科院校举办的二级职业技术学院计算机类专业或信息类相关专业的教材，也可作为非计算机专业学生选修课或全国计算机等级二级考试理论辅修课的教材，还可作为计算机相关人员的自学参考书。

本书第1、2、9章由毕万新编写；第3、4章由吕振凯编写；第5、6、7、8章由乔国荣编写。乔国荣统编全稿。

编写过程中，作者得到了所在单位大连轻工业学院职业技术学院的领导和同行的支持，另外，袁仁水、李文锋、宫明明、翟慧、张震、刘佳、赵海琳、苗专生、刘燕等也提供了宝贵的意见，在此一并表示衷心的感谢。

由于编者水平有限，书有难免有一些不足之处，请读者批评指正。

编 者

2006年6月

丛书序

当前,我国就业和经济发展正面临着两个大的变化——社会劳动力就业需要加强技能培训,产业结构优化升级需要培养更多的高级技工。温家宝总理在全国职业教育工作会议上指出:要把发展职业教育纳入国民经济和社会发展“十一五”规划,统筹安排,加大扶持力度,把基础教育、职业教育和高等教育放在同等重要位置,创办具有中国特色的职业教育。

在这样的一个大市场环境下,通过市场调研,许多高职高专的一线教师反映,目前仍然缺乏真正满足行业需求、培养学生技能、满足学生就业需要的好教材。针对这一现状,我们适时推出这套切合当前教育改革和社会需要的面向应用技能型人才培养的系列教材。殷切希望这套教材助力推动办学体制和运作机制的改革,提高高等职业教育的整体水平,而且有助于加快改进职业教育的办学模式、课程体系和教学方法,形成具有多元化特色的教育体系。

作为最早开发计算机科普、教育图书的出版单位,原科海培训中心积极地参与到高职高专计算机教育教材改革中。我们针对内蒙、山东、辽宁、吉林、河南等地的高职高专院校进行了大面积的调研,邀请全国高职高专院校计算机相关专业的专家与名师,共同探讨相关教材的出版问题,经过多次研讨,我们确定了教材编写指导方针:

- 突出职业教育特色,以操作技能为本,重视学生实践能力的培养。
- 理论学习与技能训练并重,基本技能培养与主流技术相结合,以市场对人才的需求为依据,从应用和工程实践的角度安排教材内容。
- 力求让教材涵盖国家有关资格考试的知识和技能要求,让学生在获得能力的同时获得相应从业资质。

在此指导方针下,我们根据教育部的指导思想,按照教学大纲的要求,结合社会对各类人才的技能需求,充分考虑教师的授课特点和授课条件,对以往深受广大师生欢迎的部分教材进行重新审核、重新编写,对新出现的社会职业需求进行研究,邀请资深专家和教师共同编写了这套《21世纪高职高专计算机教育规划教材》。本丛书宗旨是,走实践应用案例教学之路,培养技能型紧缺人才。

丛书特色

本套教材采用“课本 + 网络教学服务”的形式为师生提供各类服务,使教材建设具有实用性和前瞻性,与就业市场紧密结合。

(1) 课本

由“理论、理论理解(或应用)辅助示例(课堂练习)、阶段性理论综合应用中典型案例(上机实验)、习题、大型实践性案例(课程设计)”五大部分组成。

- 理论讲解以“够用”为原则。
- 讲解基础知识时，以“易学易懂易用”为原则，先对知识点作简要介绍，然后通过小实例来演示知识点及其对实际工作的影响，专注于解决问题的经典方法，保证读者看得懂，学得会，以最快速度融入到这个领域中来。
- 阶段性练习，则用于培养学生综合应用所学内容解决实际问题的能力。
- 课程设计实践部分以“贴近实际工作需要为原则”，让学生了解社会对从业人员的真正需求，为就业铺平道路。

(2) 网络教学服务

- 向师生提供电子教案、教材配套素材、教学辅助素材、教学辅助案例等教学资源。
- 开设教学论坛，供师生及专业人士互动交流。

丛书组成

本套教材涵盖计算机基础、程序设计、数据库开发、网络技术、多媒体技术、计算机辅助设计及毕业设计和就业指导等诸多领域，后面将陆续推出：

- | | |
|----------------------|---------------------------|
| • 软件工程理论与实践教程 | • 计算机专业毕业设计指导教程 |
| • 计算机文化基础教程 | • 网页设计与制作教程（MX 2004 版） |
| • Authorware 多媒体制作教程 | • AutoCAD 机械制图教程 |
| • AutoCAD 建筑制图教程 | • SQL Server 2000 数据库开发教程 |
| • Access 数据库开发教程 | • C#程序设计教程 |
| • VB.NET 程序设计教程 | • C++程序设计教程 |
| • Java 程序设计教程 | • ASP 技术应用教程 |
| • 计算机网络管理与维护教程 | • 计算机网络安全教程 |
| | |

编者寄语

如果说科学技术的飞速发展是 21 世纪的一个重要特征的话，那么教学改革将是 21 世纪教育工作不变的主题。要紧跟教学改革，不断创新，真正编写出满足新形势下教学需求的教材，还需要我们不断地努力实践、探索和完善。本丛书虽然经过细致的编写与校订，仍难免有疏漏和不足，需要不断地补充、修订和完善。我们热情欢迎使用本丛书的教师、学生和读者朋友提出宝贵意见和建议，使之更臻成熟。

我社网站（<http://www.khp.com.cn>）免费提供本套丛书相关教材的素材文件、电子教案及相关教学资源。在教材使用中，老师们有什么意见或建议，请直接与我们联系，联系电话是：(010)82896448，电子邮件地址是：hnlycyq2003@vip.sina.com，chj01080108@126.com。

目 录

第1章 绪论.....	1
1.1 数据结构的基本概念.....	1
1.1.1 数据结构的定义.....	1
1.1.2 数据的逻辑结构及存储结构.....	3
1.1.3 数据结构有关概念及术语.....	4
1.2 算法和算法描述.....	5
1.2.1 什么是算法.....	5
1.2.2 算法描述.....	6
1.3 算法分析.....	7
1.3.1 空间复杂度.....	7
1.3.2 时间复杂度.....	7
1.4 本章小结.....	8
习题1.....	8
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.....	12
2.2 线性表的顺序存储结构.....	12
2.2.1 线性表的顺序存储——顺序表.....	13
2.2.2 顺序表基本操作的实现.....	13
2.2.3 顺序表的应用举例.....	17
2.3 线性表的链式存储结构.....	19
2.3.1 线性表的链式存储——链表.....	19
2.3.2 单链表.....	21
2.3.3 循环链表.....	27
2.3.4 双向链表.....	28
2.3.5 单链表应用举例.....	30
2.4 本章小结.....	34
习题2.....	34
实训1.....	37
第3章 栈与队列.....	40
3.1 栈.....	40

3.1.1 栈的定义.....	40
3.1.2 栈的顺序存储及其基本操作的实现.....	41
3.1.3 栈的链式存储及其基本操作的实现.....	46
3.1.4 栈的应用举例.....	49
3.2 队列.....	51
3.2.1 队列的定义.....	51
3.2.2 队列的顺序存储及其基本操作的实现.....	52
3.2.3 队列的链式存储及其基本操作的实现.....	56
3.2.4 队列的应用举例.....	58
3.3 本章小结.....	59
习题 3.....	59
实训 2.....	63
第 4 章 串、数组和广义表.....	70
4.1 串.....	70
4.1.1 串的定义和特性.....	70
4.1.2 串的顺序存储及其基本操作实现.....	71
4.1.3 串的链式存储及其基本操作实现.....	78
4.1.4 串的应用举例.....	78
4.2 数组.....	79
4.2.1 数组的定义和运算.....	79
4.2.2 数组的顺序存储结构.....	80
4.2.3 矩阵的压缩存储.....	80
4.2.4 稀疏矩阵.....	81
4.3 广义表.....	87
4.3.1 广义表的定义和特性.....	87
4.3.2 广义表的存储结构及其基本操作实现.....	88
4.4 本章小结.....	89
习题 4.....	89
实训 3.....	91
第 5 章 树.....	94
5.1 树.....	94
5.1.1 树的定义及基本术语.....	94
5.1.2 树的表示.....	95
5.2 二叉树及其遍历.....	96
5.2.1 二叉树的定义.....	96
5.2.2 二叉树的重要性质.....	96
5.2.3 二叉树的存储结构.....	98

5.2.4 二叉树的遍历	99
5.3 线索二叉树	106
5.3.1 线索二叉树的定义	107
5.3.2 线索二叉树的基本操作	108
5.4 树和森林	109
5.4.1 树的存储结构	109
5.4.2 二叉树与树之间的转换	111
5.4.3 森林与二叉树的转换	112
5.5 二叉树应用实例	112
5.5.1 二叉排序树	112
5.5.2 哈夫曼树	117
5.6 本章小结	119
习题 5	119
实训 4	124
第 6 章 图	129
6.1 图的基本概念	129
6.1.1 图的定义	129
6.1.2 图的基本术语	130
6.2 图的存储结构	132
6.2.1 邻接矩阵	132
6.2.2 邻接表	134
6.3 图的遍历	136
6.3.1 深度优先搜索	136
6.3.2 广度优先搜索	137
6.4 最小生成树	139
6.4.1 普里姆算法	140
6.4.2 克鲁斯卡尔算法	143
6.5 最短路径	146
6.5.1 单源最短路径	147
6.5.2 每对顶点之间的最短路径	150
6.6 拓扑排序	154
6.6.1 AOV 网	154
6.6.2 拓扑 (Topology) 排序的实现	155
6.7 本章小结	158
习题 6	159
实训 5	161
第 7 章 查找	165

7.1 查找的基本概念	165
7.2 顺序查找	166
7.3 二分查找	167
7.4 分块查找	170
7.5 哈希表查找	172
7.5.1 哈希表查找的基本概念	172
7.5.2 构造哈希函数的方法	173
7.5.3 哈希冲突解决方法	175
7.5.4 哈希查找效率的分析	179
7.6 本章小结	180
习题 7	180
实训 6	184
第 8 章 排序	186
8.1 排序的基本概念	186
8.2 插入排序	187
8.2.1 直接插入排序	187
8.2.2 二分法插入排序	189
8.2.3 希尔排序	190
8.3 选择排序	192
8.3.1 简单选择排序	192
8.3.2 堆排序	193
8.4 交换排序	196
8.4.1 冒泡排序	196
8.4.2 快速排序	198
8.5 归并排序	200
8.6 基数排序	202
8.7 本章小结	205
习题 8	206
实训 7	209
第 9 章 课程设计指导	213
9.1 课程设计大纲	213
9.2 课程设计题目及设计要求	214
9.3 飞机售票系统实例	217
9.4 本章小结	222
附录 习题答案	223
参考文献	244

第1章 绪论

本章导读

在深入学习数据结构之前，首先了解学习数据结构的意义、数据结构的基本术语及数据结构的一些相关概念等，这对学习后面的内容将有很大的帮助。本章介绍的数据结构研究对象和有关概念，包括数据、数据元素、数据类型、逻辑结构、存储结构、算法描述（C语言描述）和算法分析等。

教学目标

通过本章学习，要求掌握以下内容：

- 理解和熟悉数据结构中的基本概念。
- 理解和掌握线性结构、树形结构和图形结构的概念。
- 熟悉算法评价的一般规则，算法时间复杂度、空间复杂度的概念和数量级的表示方法。

1.1 数据结构的基本概念

1.1.1 数据结构的定义

计算机是一种数据处理装置。用计算机处理实际问题时，一般要先对具体问题进行抽象化，建立起实际问题的求解模型，然后设计出相应的算法，编写程序并上机调试，直至得到最终结果。

在计算机处理数据过程中，大批量的数据并不是彼此孤立、杂乱无章的，它们之间有着内在的联系。只有利用这些内在的联系，把所有数据按照某种规则有机地组织起来，才能对数据进行有效的处理。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、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。

下面举出几个例子，来说明什么是数据结构。

例 1.1 学生成绩检索问题。见表 1-1。

表 1-1 学生成绩表

学号	姓名	性别	视频技术	C 程序设计	网页设计	多媒体著作工具
2002001	李丽	女	76	88	78	80
2002002	乔丽娜	女	77	82	67	77
2002003	王龙龙	男	84	76	76	67
2002004	李琳	女	85	87	88	78
2002005	朱宏利	男	88	90	78	69

这个学生成绩表是一个二维表格，每一行为一个学生的全部信息，每一列数据的类型相同。整个二维表形成学生成绩的一个线性序列，每个学生的信息按照学号次序存放，各学生之间形成一种线性关系。这是一种典型的数据结构，我们称这种数据结构为线性表。对这种线性表的主要操作，是给出学生的姓名时，能在该表中快速找到学生每门课程的成绩，还有就是对该表如何添加一个新的学生、如何删除已经退学的学生、如何修改表中的数据等操作。这就是数据结构要研究的内容。

例 1.2 某高校的专业设置情况。如图 1-1 所示。

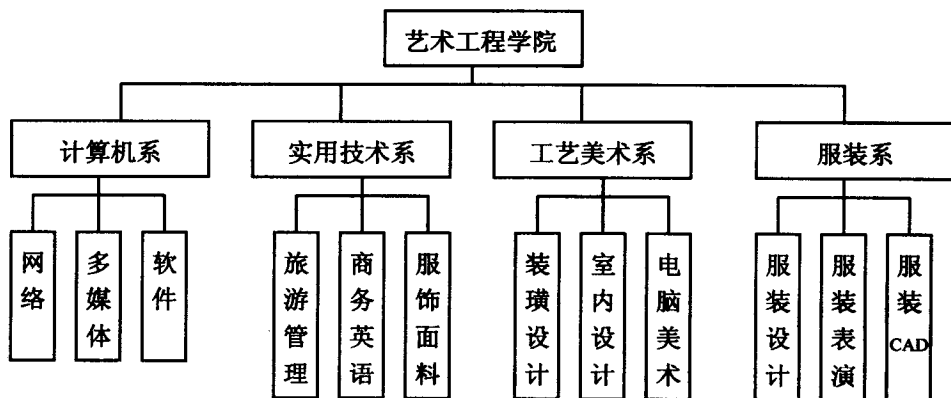


图 1-1 某高校专业设置情况

在上述专业设置情况图中，艺术工程学院分 4 个系，每个系设 3 个专业。在这种数据结构中，数据之间的关系是一对多的非线性关系。这也是常用的一种数据结构，我们称之为树形结构。

例 1.3 城镇之间的公路网。如图 1-2 所示。

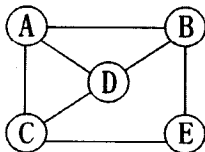


图 1-2 城镇之间的公路网

在城镇公路网中，每个顶点代表一个城镇，边表示城镇之间的道路。在这种数据结构

中，数据之间的关系是多对多的非线性关系，我们称这种数据结构为图形结构。

综合 3 个例子可见，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，是一门研究非数值计算的程序设计中，计算机的操作对象以及它们之间的关系和操作的学科。

1.1.2 数据的逻辑结构及存储结构

1. 数据的逻辑结构

数据的逻辑结构与数据在计算机中的存储无关，因此，数据的逻辑结构可以看成是从具体的问题中抽象出来的数学模型。根据数据元素之间关系的不同特性，数据的逻辑结构可划分为下面 4 种。

(1) 集合

结构中各数据元素之间不存在任何关系。这是数据结构的一种特殊情况，不在本书讨论范围之内。

(2) 线性结构

该数据结构中的数据元素存在着一对一的关系。

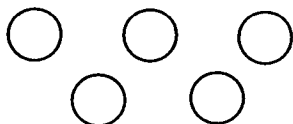
(3) 树形结构

该数据结构中的数据元素存在着一对多的关系。

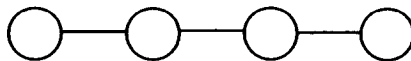
(4) 图形或网状结构

该数据结构中的数据元素存在着多对多的关系。

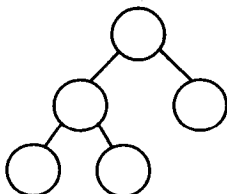
上述 4 类基本结构关系如图 1-3 所示。



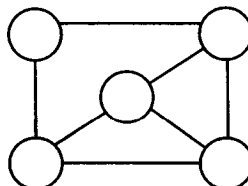
(a) 集合结构示意图



(b) 线性结构示意图



(c) 树形结构示意图



(d) 图形结构示意图

图 1-3 4 种基本数据结构

2. 数据的存储结构

数据的逻辑结构需要用计算机处理,要存入到计算机的存储单元中。数据的逻辑结构在计算机中的表示称为数据的存储结构,又称为数据的物理结构。讨论数据结构时,不但要讨论数据的逻辑结构,还要讨论数据的存储结构。通常,计算机内的数据元素用一组连续的位串来表示,这个位串称为结点。数据元素之间的关系,也称结点之间的关系。在计算机内有以下4种存储数据元素的表示方法。

(1) 顺序存储方法

该方法是将逻辑上相邻的结点存储在物理位置上也相邻的存储单元里,结点之间的逻辑关系由存储单元的邻接关系来表示。用这种方法存储数据元素时,只存储结点的值,不存储结点之间的关系。这种存储表示称为顺序存储结构,主要应用于线性的数据结构;非线性的数据结构也可以通过某种线性化的过程后进行存储。

(2) 链式存储方法

链式存储方法不要求逻辑上相邻的结点在物理位置上也相邻,结点间的关系由附加的指针来表示。通过指针指向结点的邻接结点,将所有结点串联在一起,这称为链式存储结构。也就是说,链式存储方法不仅存储结点的值,而且还存储结点之间的关系。所以,链式存储方法中的结点由两部分组成,一个是存储结点本身的值,称为数据域;另一个是存储该结点的各后继结点的存储单元地址,称为指针域(可包含一个或多个指针)。

(3) 索引存储方法

索引存储方法是在存储结点信息的同时,再建立一个附加的索引表,然后利用索引表中索引项的值来确定结点的实际存储单元地址。索引表中的每一项称为索引项,索引项的一般形式为(关键字,地址),关键字能唯一标识一个结点。

(4) 哈希存储方法

哈希存储方法的基本思想是根据结点的关键字直接计算出结点的存储地址。方法是把结点的关键字作为自变量,通过一个称为哈希函数(Hash)的计算规则,确定出该结点的确切存储单元地址。

上面这4种方法既可以单独使用,也可以组合起来对数据结构进行存储。同一种逻辑结构采用不同的存储方法,可以得到不同的存储结构。选取哪种存储结构来表示相应的逻辑结构,视具体的情况而定,具体要考虑数据的运算是否方便及相应算法的时间复杂度和空间复杂度的要求。

1.1.3 数据结构有关概念及术语

1. 数据(Data)

数据是指能够输入到计算机中,并能被计算机处理的一切对象。对计算机科学而言,数据的含义极为广泛,如整数、实数、字符、文字、图形、图像和声音等都是数据。

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。例如，在表 1-1 中，学生成绩表中的一行就是一个数据元素。数据元素还可以分割成若干个具有不同属性的项（字段）。一般由一个或多个数据项组成。

3. 数据项 (Data Item)

数据项是具有独立意义的最小数据单位，是对数据元素属性的描述。在表 1-1 中，每个数据元素由 7 个数据项组成，其中“学号”数据项描述了顺序，“姓名”数据项描述了成绩所有者的名字，其他几项描述了学生的具体成绩。

4. 数据类型 (Data Type)

数据类型是一组性质相同的值的集合以及定义于这个集合上的一组操作的总称。每个数据项都属于某一确定的基本数据类型。如表 1-1 中，“学号”为数值型，“姓名”为字符型。

5. 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，整数数据对象的集合是 $\{0, \pm 1, \pm 2, \dots\}$ ；字符数据对象的集合是 $\{'A', 'B', 'C', \dots, 'Z'\}$ 。

1.2 算法和算法描述

1.2.1 什么是算法

算法是对某一特定问题求解步骤的一种描述。在计算机系统中，算法是由若干条指令组成的有穷序列，其中每一条指令表示计算机的一个或多个操作。算法满足以下 5 个性质：

- 输入：一个算法可以有零个或多个输入量，在算法执行之前提供给算法。
- 输出：一个算法的执行结果要有一个或多个输出量，它是算法对输入数据处理的结果。
- 有穷性：一个算法必须在执行有穷步骤之后结束，即必须在有限时间内完成。
- 确定性：算法中的每一步骤都有明确的含义，没有二义性。
- 可行性：算法中的每一步都必须是可行的，算法中描述的操作的每一步都能在有限次、有限时间内得以实现。

对于同一个问题，可以有很多种不同的算法，这就需要对算法有一个总的设计要求。一般来说，一个算法必须具有以下几个方面的基本特征。

1. 正确性

正确性是设计一个算法的首要条件，所设计的算法要满足具体问题的要求。在给算法



输入合理的数据后，能在有限的时间内得出正确的结果。

2. 可读性

算法是对特定问题求解步骤的一种描述，它要能转变成计算机可执行的程序，同时必须可以供他人使用。为了使所设计的算法让他人能看懂，在算法或程序中可以增加一些注释来提高可读性。

3. 健壮性

当输入的数据不符合要求时，算法应能判断出数据的非法性，并能进行适当的处理，比如暂停或终止程序的执行、显示错误信息等。算法不应允许产生不可预料的结果。

4. 高效性

算法的效率是指算法执行的时间和占用的存储空间。如果对于同一个问题有多个算法可供选择，应尽可能选择执行时间短、占用空间少的算法。

1.2.2 算法描述

算法的描述方法很多。根据描述算法语言的不同，可将算法描述分为以下 4 种。

1. 框图算法描述

框图算法描述是采用传统流程图或 N-S 图等方式来描述算法，在算法研究的早期曾经很流行。它的优点是直观、易懂，但用来描述比较复杂的算法时就显得不太方便，也不够清晰。

2. 自然语言描述

自然语言描述就是用人类自然语言（如中文，英文等），同时使用一些程序设计语言中的语句来描述算法。

3. 伪语言算法描述

如使用类 C 语言进行算法描述。这种算法不能直接在计算机上运行，但专业设计人员经常使用类 C 语言来描述算法，因为它容易编写、易阅读，有统一格式。

4. 计算机语言描述

这是可以在计算机上运行并获得结果的算法，通常这种算法也称为程序。本书中的大部分算法都是用 C 语言描述的，而且尽可能给出一个完整的 C 语言程序。在本书所有实训的参考答案中，也给出了一个完整的 C 语言程序，以方便学生上机参考。