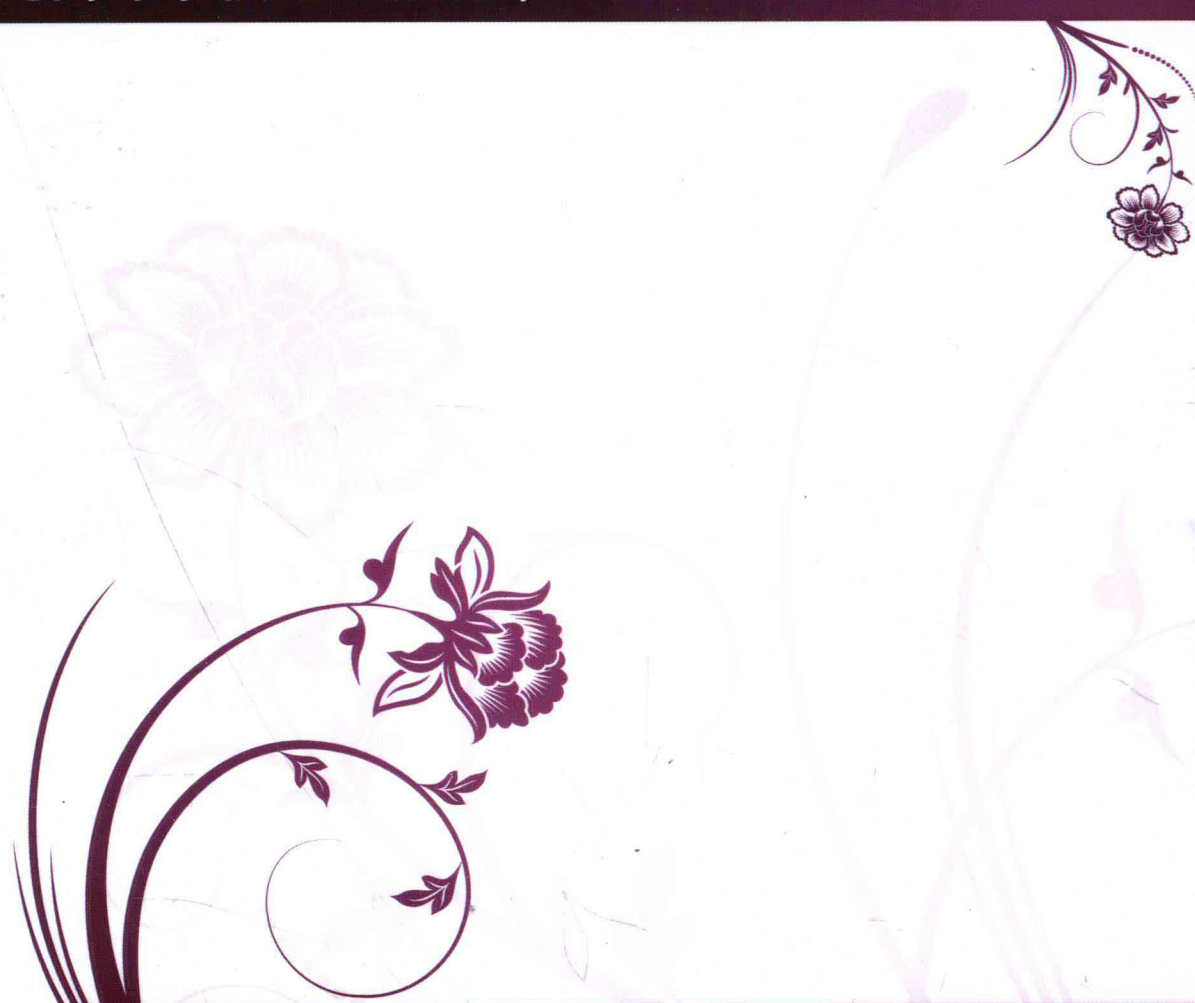




算法与数据结构

(第二版)

陈媛 何波 卢玲 涂飞 等编著



清华大学出版社

21世纪高等学校规划教材 | 计算机科学与技术

算法与数据结构

(第二版)

陈媛 何波 卢玲 涂飞 等编著

清华大学出版社
北京

内 容 简 介

《算法与数据结构(第二版)》是2005年出版的原书的修订版。修订版在保持原书基本框架和特色的基础上,增加了面向对象程序设计C++及实际应用算法及数据结构的实例。

全书系统介绍了算法与数据结构方面的基本知识,重点阐述了基本数据结构及算法在程序开发中的应用方法。通过深入的学习和分析,能够帮助读者极大地提高软件开发和设计能力。本书给出的所有算法和程序均采用C语言描述并已调试通过,部分算法还增加了C++实现代码。本书注重教材的可读性和适用性,书中附有大量的图表、程序,使读者能正确、直观地理解问题。书中每章有学习要点、习题和上机练习题,既便于教学,又便于自学。

本书的读者只要求具有C语言基础,不要求具有面向对象程序设计基础,通过本教材的学习可以帮助读者树立面向对象的编程思想。本书特别适合普通高校本专科学生使用,也可作为其他程序类课程的辅导教材。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

算法与数据结构/陈媛等编著. —2版. —北京:清华大学出版社,2011.9

(21世纪高等学校规划教材·计算机科学与技术)

ISBN 978-7-302-24882-8

I. ①算… II. ①陈… III. ①算法分析—高等学校—教材 ②数据结构—高等学校—教材

IV. ①TP301.6 ②TP311.12

中国版本图书馆CIP数据核字(2011)第033281号

责任编辑:闫红梅 薛 阳

责任校对:梁 毅

责任印制:杨 艳

出版发行:清华大学出版社

地 址:北京清华大学学研大厦A座

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

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62795954, jsjtc@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者:北京富博印刷有限公司

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:185×260 印 张:19 字 数:478千字

版 次:2011年9月第2版 印 次:2011年9月第1次印刷

印 数:1~3000

定 价:29.50元

产品编号:037768-01

编审委员会成员

(按地区排序)

清华大学

周立柱 教授

章 征 教授

王建民 教授

冯建华 教授

刘 强 副教授

北京大学

杨冬青 教授

陈 钟 教授

陈立军 副教授

北京航空航天大学

马殿富 教授

吴超英 副教授

姚淑珍 教授

中国人民大学

王 珊 教授

孟小峰 教授

陈 红 教授

北京师范大学

周明全 教授

北京交通大学

阮秋琦 教授

赵 宏 教授

北京信息工程学院

孟庆昌 教授

北京科技大学

杨炳儒 教授

石油大学

陈 明 教授

天津大学

艾德才 教授

复旦大学

吴立德 教授

吴百锋 教授

杨卫东 副教授

同济大学

苗夺谦 教授

徐 安 教授

华东理工大学

邵志清 教授

华东师范大学

杨宗源 教授

应吉康 教授

东华大学

乐嘉锦 教授

孙 莉 副教授

浙江大学

扬州大学

南京大学

南京航空航天大学

南京理工大学

南京邮电学院

苏州大学

江苏大学

中国矿业大学

武汉大学

华中科技大学

中南财经政法大学

华中师范大学

江汉大学

国防科技大学

中南大学

湖南大学

西安交通大学

长安大学

哈尔滨工业大学

吉林大学

山东大学

中山大学

厦门大学

仰恩大学

云南大学

电子科技大学

成都理工大学

西南交通大学

吴朝晖 教授

李善平 教授

李 云 教授

骆 斌 教授

黄 强 副教授

黄志球 教授

秦小麟 教授

张功萱 教授

朱秀昌 教授

王宜怀 教授

陈建明 副教授

鲍可进 教授

张 艳 教授

何炎祥 教授

刘乐善 教授

刘腾红 教授

叶俊民 教授

郑世珏 教授

陈 利 教授

颜 彬 教授

赵克佳 教授

邹北骥 教授

刘卫国 教授

林亚平 教授

沈钧毅 教授

齐 勇 教授

巨永锋 教授

郭茂祖 教授

徐一平 教授

毕 强 教授

孟祥旭 教授

郝兴伟 教授

潘小轰 教授

冯少荣 教授

张思民 教授

刘惟一 教授

刘乃琦 教授

罗 蕾 教授

蔡 淮 教授

于 春 讲师

曾华桑 教授

出版说明

随着我国改革开放的进一步深化,高等教育也得到了快速发展,各地高校紧密结合地方经济建设发展需要,科学运用市场调节机制,加大了使用信息科学等现代科学技术提升、改造传统学科专业的投入力度,通过教育改革合理调整和配置了教育资源,优化了传统学科专业,积极为地方经济建设输送人才,为我国经济社会的快速、健康和可持续发展以及高等教育自身的改革发展做出了巨大贡献。但是,高等教育质量还需要进一步提高以适应经济社会发展的需要,不少高校的专业设置和结构不尽合理,教师队伍整体素质亟待提高,人才培养模式、教学内容和方法需要进一步转变,学生的实践能力和创新精神亟待加强。

教育部一直十分重视高等教育质量工作。2007年1月,教育部下发了《关于实施高等学校本科教学质量与教学改革工程的意见》,计划实施“高等学校本科教学质量与教学改革工程”(简称“质量工程”),通过专业结构调整、课程教材建设、实践教学改革、教学团队建设等多项内容,进一步深化高等学校教学改革,提高人才培养的能力和水平,更好地满足经济社会发展对高素质人才的需要。在贯彻和落实教育部“质量工程”的过程中,各地高校发挥师资力量强、办学经验丰富、教学资源充裕等优势,对其特色专业及特色课程(群)加以规划、整理和总结,更新教学内容、改革课程体系,建设了一大批内容新、体系新、方法新、手段新的特色课程。在此基础上,经教育部相关教学指导委员会专家的指导和建议,清华大学出版社在多个领域精选各高校的特色课程,分别规划出版系列教材,以配合“质量工程”的实施,满足各高校教学质量和教学改革的需要。

为了深入贯彻落实教育部《关于加强高等学校本科教学工作,提高教学质量的若干意见》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院、系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上。精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展,顺应并符合21世纪教学发展的规律,代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版

社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。推出的特色精品教材包括:

(1) 21 世纪高等学校规划教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。

(2) 21 世纪高等学校规划教材·计算机科学与技术——高等学校计算机相关专业的教材。

(3) 21 世纪高等学校规划教材·电子信息——高等学校电子信息相关专业的教材。

(4) 21 世纪高等学校规划教材·软件工程——高等学校软件工程相关专业的教材。

(5) 21 世纪高等学校规划教材·信息管理与信息系统。

(6) 21 世纪高等学校规划教材·财经管理与计算机应用。

(7) 21 世纪高等学校规划教材·电子商务。

清华大学出版社经过二十多年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会

联系人:魏江江

E-mail:weijj@tup.tsinghua.edu.cn

前言

1. 关于算法与数据结构

计算机的日益发展,其应用早已不局限于简单的数值运算,而涉及问题的分析、数据结构框架的设计以及插入、删除、排序、查找等复杂的非数值处理和操作。算法与数据结构的学习就是为以后利用计算机高效地开发非数值处理的计算机程序打下坚实的理论、方法和技术基础。

算法与数据结构旨在分析研究计算机加工的数据对象的特性,以便选择适当的数据结构和存储结构,从而使建立在其上的解决问题的算法达到最优。

2. 结构安排

全书共分为 9 章,主要内容如下。

第 1 章:绪论。主要介绍数据结构和算法的基本概念和术语,C 语言的基本数据类型及用 C 语言描述算法的要点,C++ 语言的类与抽象数据类型的关系,C++ 特性及与 C 程序的区别,C++ 语言验证算法的方法。

第 2 章:线性表。主要介绍线性表的逻辑结构,线性表的顺序存储结构和链式存储结构,线性表的应用实例。

第 3 章:栈和队列。主要介绍栈与队列的基本概念及存储结构,栈与队列的应用实例,递归的概念及设计方法,递归实现与栈的关系。

第 4 章:数组和字符串。主要介绍数组存储结构及应用实例,串的基本概念和存储结构,串的应用实例。

第 5 章:树。主要介绍树及二叉树的基本概念及存储结构,二叉树的应用——哈夫曼树及其编码。

第 6 章:图。主要介绍图的基本概念及存储结构,图的遍历,图的生成树和最小生成树,最短路径,有向无环图,图的应用实例。

第 7 章:查找。主要介绍静态查找,动态查找,哈希表查找。

第 8 章:排序。主要介绍插入排序、选择排序、交换排序、归并排序、基数排序及外部排序的简介。

第 9 章:标准模板库。简单介绍标准模板库的组成及使用要点,同时介绍 STL 的应用实例。

本书第 1、6 章由陈媛教授编写,第 2、5 章由何波副教授编写,第 3、4 章由卢玲编写,第 7、8 章由涂飞编写,第 9 章由刘恒洋编写。全书由陈媛教授统稿。

3. 本书特点

本书给出的所有算法和程序均采用 C 语言描述并已调试通过,部分算法还增加了 C++ 实现代码,用 C 和 C++ 两种语言描述算法和数据结构,使数据结构的学习与随后的编程课程紧密结合。本书注重教材的可读性和适用性,书中附有大量的图表、程序,使读者能正确、直观地理解问题。书中每章有学习要点、习题和上机练习题,既便于教学,又便于自学。

本书内容和结构体现了教学改革成果。全书由重庆市精品课程“数据结构”重庆理工大学课程组的教师集体编写完成。作者都是长期在高校从事“算法与数据结构”教学的一线教师,有丰富的教学经验和软件开发能力。作者以多年的教学经验和多项教研课题的研究成果,构建了数据结构概念建立和编程思想培养的框架体系,总结提炼了学习本课程的重难点和解决方法,大部分样例都经过整理和组织,以便更好地理解掌握。

为了方便教学,本书配套资料含教师用电子教案和精品课程“数据结构”教学网站材料。另外,我们还开发了练习系统,使用本书的院校可通过邮箱 cy@cqut.edu.cn 向作者索取。

4. 适用对象

本书的读者只要求具有 C 语言基础,不要求具有面向对象程序设计基础,通过本教材的学习可以帮助读者树立面向对象的编程思想。本书可作为计算机专业、信息专业及其他相关专业的本专科教材,也可以作为广大参加自学考试的人员和软件工作者的参考资料。本书既可作为“数据结构和算法”课程的教材,也可作为其他程序类课程的辅导教材。

由于我们的水平有限,本书可能会有不尽人意之处,错漏之处在所难免,敬请读者批评指正,以便我们及时修改。

编 者

2010 年 10 月

目 录

第 1 章 绪论	1
1.1 数据结构的基本概念与学习方法	1
1.1.1 “数据结构”的研究对象	1
1.1.2 数据结构的基本概念和基本术语	2
1.2 算法与数据结构	8
1.2.1 算法的概念	8
1.2.2 描述算法的方法	9
1.2.3 算法的分析	9
1.3 学习算法与数据结构的意义和方法	13
1.4 C 语言的数据类型及其算法描述的要点	13
1.4.1 C 语言的基本数据类型概述	14
1.4.2 C 语言的数组和结构数据类型	14
1.4.3 C 语言的指针类型概述	17
1.4.4 C 语言的函数	22
1.4.5 用 C 语言验证算法的方法	25
1.5 从 C 到 C++	27
1.5.1 C++ 的类和抽象数据类型	27
1.5.2 C++ 语言验证算法的方法	30
1.5.3 C++ 语言与 C 语言程序的区别	31
1.5.4 C++ 的重要特性	31
学习要点	37
习题 1	37
上机练习 1	38
第 2 章 线性表	39
2.1 线性表的逻辑结构	39
2.1.1 线性表的定义	39
2.1.2 线性表的运算	40
2.2 线性表的顺序存储结构——顺序表	41
2.2.1 顺序表	41
2.2.2 顺序存储结构的优缺点	42
2.2.3 顺序表上的基本运算	43

2.3 线性表的链式存储结构——链表	46
2.3.1 单链表	47
2.3.2 循环链表和双向链表	54
2.4 线性表的应用示例	57
2.5 C++中的线性表	62
2.5.1 C++中线性表的抽象数据类型	62
2.5.2 C++中线性表的顺序存储	63
2.5.3 C++中线性表的链式存储	66
学习要点	69
习题2	69
上机练习2	70
第3章 栈和队列	72
3.1 栈	72
3.1.1 栈的基本概念	72
3.1.2 栈的顺序存储结构	74
3.1.3 栈的链式存储结构	77
3.2 栈的应用实例	79
3.2.1 表达式的求值	79
3.2.2 栈与函数的调用	81
3.2.3 栈在回溯法中的应用	81
3.3 队列	82
3.3.1 队列的基本概念	83
3.3.2 队列的顺序存储结构	83
3.3.3 队列的链式存储结构	88
3.4 队列的应用实例	91
3.4.1 舞伴问题	91
3.4.2 模拟打印队列的管理	92
3.5 递归	95
3.5.1 递归的定义及递归模型	95
3.5.2 递归的实现	98
3.5.3 递归的设计	99
3.5.4 递归到非递归的转换	100
3.6 C++中的栈和队列	101
3.6.1 C++中的栈	101
3.6.2 C++中的队列	103
学习要点	104
习题3	105
上机练习3	105

第 4 章 数组和字符串	108
4.1 数组	108
4.1.1 数组的定义与操作	108
4.1.2 数组的顺序存储结构	109
4.1.3 矩阵的压缩存储方法	110
4.2 字符串	117
4.2.1 字符串的定义与操作	117
4.2.2 字符串的存储结构	119
4.2.3 字符串基本操作的实现	123
4.2.4 字符串的应用举例	128
4.3 C++ 中的数组和字符串	130
4.3.1 C++ 中的数组	130
4.3.2 C++ 中的字符串	131
学习要点	132
习题 4	132
上机练习 4	133
第 5 章 树	134
5.1 树	134
5.1.1 树的概念	134
5.1.2 树的基本操作	136
5.2 二叉树	136
5.2.1 二叉树的概念	137
5.2.2 二叉树的性质	139
5.2.3 二叉树的存储结构及其实现	140
5.3 二叉树的遍历	142
5.3.1 递归的遍历算法	142
5.3.2 二叉树遍历操作应用举例	146
5.4 线索二叉树	148
5.4.1 线索二叉树的定义	148
5.4.2 线索二叉树的常见运算	150
5.5 一般树的表示和遍历	154
5.5.1 一般树的表示	154
5.5.2 二叉树与树之间的转换	156
5.5.3 一般树的遍历	157
5.6 哈夫曼树及其应用	158
5.6.1 哈夫曼树(最优二叉树)	158
5.6.2 哈夫曼树的应用	161

5.7 C++中的树	165
5.7.1 C++中的二叉树结点类	165
5.7.2 C++中的二叉树类	166
5.7.3 C++中的二叉树非递归遍历	167
学习要点	169
习题 5	169
上机练习 5	172
第 6 章 图	174
6.1 图	174
6.1.1 图的定义	174
6.1.2 图的基本术语	175
6.2 图的存储结构	178
6.2.1 邻接矩阵	178
6.2.2 图的邻接表	179
6.2.3 十字链表	181
6.2.4 边集数组	182
6.3 图的遍历	184
6.3.1 深度优先搜索	184
6.3.2 广度优先搜索	187
6.4 图的连通性	189
6.4.1 无向图的连通分量	189
6.4.2 生成树和最小代价生成树	189
6.5 有向无环图及其应用	195
6.5.1 拓扑排序	196
6.5.2 关键路径	200
6.6 最短路径及其应用	204
6.6.1 单源最短路径	204
6.6.2 每个顶点之间的最短路径	208
6.7 C++中的图	209
6.7.1 C++中的图类	209
6.7.2 图的邻接表的 C++ 程序	211
6.7.3 图的遍历的 C++ 程序	212
6.7.4 图的最小生成树的 C++ 程序	213
学习要点	214
习题 6	214
上机练习 6	216

第 7 章 查找	217
7.1 基本概念与术语	217
7.2 静态查找表	219
7.2.1 静态查找表结构	219
7.2.2 顺序查找	220
7.2.3 有序表的折半查找	221
7.2.4 有序表的插值查找和斐波那契查找	223
7.2.5 分块查找	224
7.3 动态查找表	225
7.3.1 二叉排序树	225
7.3.2 平衡二叉树(AVL 树)	229
7.3.3 B-树和 B+树	234
7.4 哈希表查找	239
7.4.1 哈希表与哈希方法	239
7.4.2 常用的哈希函数	240
7.4.3 处理冲突的方法	242
7.4.4 哈希表的查找分析	244
7.5 C++中的查找	245
7.5.1 静态查找的 C++程序	245
7.5.2 动态查找的 C++程序	246
学习要点	247
习题 7	247
上机练习 7	248
第 8 章 排序	249
8.1 基本概念	249
8.2 插入排序	250
8.2.1 直接插入排序	250
8.2.2 希尔排序	252
8.3 交换排序	253
8.3.1 冒泡排序	253
8.3.2 快速排序	254
8.4 选择排序	257
8.4.1 简单选择排序	257
8.4.2 堆排序	258
8.5 归并排序	262
8.6 基数排序	264
8.7 外部排序简介	268

8.7.1 外存信息的存取·····	269
8.7.2 外部排序的基本方法·····	269
8.8 C++中的排序·····	270
学习要点·····	272
习题 8·····	272
上机练习 8·····	273
第 9 章 标准模板库·····	274
9.1 标准模板库·····	274
9.1.1 容器·····	274
9.1.2 迭代器·····	277
9.1.3 算法·····	281
9.2 STL 应用实例·····	283
9.2.1 双向链表操作的 STL 实现·····	283
9.2.2 STL 测试程序·····	285
学习要点·····	289
习题 9·····	289
参考文献·····	290

计算机科学是一门研究信息的表示和处理的科学,而信息的表示和组织又直接关系到信息处理程序的效率。由于许多程序规模大、结构复杂、处理对象多为非数值型数据,因此单纯依靠程序设计人员的经验和技巧已不能编制出高效率的处理程序。为了设计出效率高、可靠性强的程序,人们必须对程序设计的方法进行系统的研究。这就要求程序设计人员不但要掌握一般的程序设计技巧,而且要研究计算机程序加工的对象,即研究数据的特性以及数据之间存在的关系,这就是数据(或称信息)的结构。

数据结构是在整个计算机科学与技术领域上被广泛使用的术语。它用来反映一个数据的内部构成,即一个数据由哪些成分数据构成,以什么方式构成,呈什么结构(或关系)。数据结构分为逻辑上的数据结构和物理上的数据结构。逻辑上的数据结构反映成分数据之间的逻辑关系,而物理上的数据结构反映成分数据在计算机内部的存储安排。

1.1 数据结构的基本概念与学习方法

1.1.1 “数据结构”的研究对象

“数据结构”作为一门学科,主要研究数据的各种逻辑结构和存储结构以及对数据的各种操作。它主要有3个方面的内容:数据的逻辑结构;数据的物理存储结构;对数据的操作(或算法或运算)。通常,算法的设计取决于数据的逻辑结构,算法的实现取决于数据的物理存储结构。

用计算机解决一个实际问题时,一般的步骤是,首先得到实际问题的数学模型,然后设计相应的算法,最后编写程序,调试、完善,直至得到问题的答案。当人们用计算机处理数值计算问题时,所用的数学模型是用数学方程描述的:若问题是不变的,要用代数方程描述;若问题是动态的,就要用微分方程来描述。所涉及的运算对象一般是简单的整型、实型和逻辑型数据,因此程序设计者的主要精力集中于程序设计技巧上,而不是数据的存储和组织上。然而,目前计算机应用的更多领域是“非数值计算问题”,它们的数学模型无法用数学方程描述,而是用“数据结构”描述,解决此类问题的关键是设计出合适的数据结构。这是一种什么样的数学模型呢?下面用3个例子对它们进行说明,让读者对这种模型有一个感性的认识。

例 1.1 求一组(n 个)整数中的最大值。

算法:基本操作是“比较两个数的大小”。

模型：由多个整数排成的一个序列(一对一关系)——线性结构。

例 1.2 计算机对弈。

算法：对弈的规则和策略。

模型：由多个格局构成的呈现层次结构的树(一对多关系)——树形结构。

例 1.3 制订教学计划。

算法：教学计划的制订规则和各课程间的先后关系。

模型：各课程构成的复杂的先后关系(多对多关系)——图形结构。

从以上 3 个例子可以看出,描述非数值计算问题的数学模型是用线性表、树、图等结构来描述的,这也就是“数据结构”课程的研究对象。

一般地说,“数据结构”是一门研究非数值计算领域的程序设计问题中计算机的操作对象以及它们之间的关系和操作等的学科。

1.1.2 数据结构的基本概念和基本术语

1. 数据

数据是所有能被输入到计算机中,且被计算机处理的符号的集合。它是计算机程序加工处理的对象。客观事物包括数值、字符、声音、图形、图像等,它们本身并不是数据,只有通过编码转换成能被计算机识别、存储和处理的符号形式后才是数据。例如,例 1.1 中的一批整型数据。

2. 数据元素

数据元素是数据的基本单位,在计算机程序中通常是作为一个整体进行考虑和处理的。例如,例 1.1 中的一个整型数据。

3. 数据项

数据项是数据结构中讨论的最小单位。若数据元素可再分,则每一个独立的处理单元就是数据项,数据元素是数据项的集合;若数据元素不可再分,则数据元素和数据项是同一概念。例如,例 1.1 中的整型数据不可再分,一个整型数据既是数据元素也是数据项。但例 1.3 中的课程(包括课程号、课程名等)这个数据元素可再分,则课程号、课程名就是数据项,不是数据元素了。

4. 数据结构

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。这个概念涉及了两方面内容:一方面是数据元素,另一方面是数据元素之间的相互关系。数据元素不是孤立存在的,在它们之间总是存在某种相互关系。

5. 逻辑结构

逻辑结构是数据元素之间的相互逻辑关系,它与数据的存储无关,是独立于计算机的。

从集合论的观点出发,数据结构是由两个集合构成的一个二元组