

高 等 学 校 计 算 机 课 程 规 划 教 材

教育部—微软精品课程建设立项项目

数据结构与算法教程 (C++版)

唐宁九 游洪跃 孙界平 朱宏 杨秋辉 主编
李炳法 伍良富 主审



清华大学出版社



教育部——微软精品课程建设立项项目
高等学校计算机课程规划教材

数据结构与算法教程 (C++版)

唐宁九 游洪跃 孙界平 朱宏 杨秋辉 主编

清华大学出版社
北京

内 容 简 介

本书结合 C++ 面向对象程序设计的特点,构建了数据结构与算法,书中的所有算法都在 Visual C++ 6.0、Visual C++ 2005、Visual C++ 2005 Express、Dev-C++ 和 MinGW Developer Studio 开发环境中进行了严格的测试,而且,在作者个人网页上提供了大量的教学支持内容。

本书共分 11 章,第 1 章是基础知识,介绍基本概念及其术语,讨论实用程序软件包;第 2 章引入线性表;第 3 章介绍栈和队列,用栈实现表达式求值;第 4 章介绍串,详细讨论串的存储结构与模式匹配算法;第 5 章介绍数组和广义表,首次提出了广义表的使用空间表存储结构;第 6 章介绍树结构,应用哈夫曼编码实现压缩软件;第 7 章介绍图结构,实现图的常用存储结构,讨论图的相关应用,并实现相应算法;第 8 章介绍查找,讨论静态查找表、动态查找表与散列表,并实现了所有算法;第 9 章介绍排序,以简洁方式实现各种排序算法;第 10 章介绍文件,讨论各种常用文件结构;第 11 章介绍算法设计技术与算法分析技术。

本书在内容组织上特别考虑了读者的可接受性;在算法实现时,重点考虑了程序的可读性;并且在习题、上机实验或课程设计中进一步实现更强的功能。通过本书学习,读者不但能迅速提高数据结构与算法的水平,还能提高 C++ 程序设计的能力,经过适当的选择,本书可以作为数据结构、数据结构与算法分析、数据结构与算法设计、数据结构与算法等课程的教材,本书可作为高等院校计算机及相关专业的教材,也可供其他从事软件开发工作的读者学习参考使用。

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

数据结构与算法教程: C++ 版/唐宁九等主编. —北京:清华大学出版社,2012.12

高等学校计算机课程规划教材

ISBN 978-7-302-28030-9

I. ①数… II. ①唐… 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. ①TP311.12 ②TP312

中国版本图书馆 CIP 数据核字(2012)第 023032 号

责任编辑:汪汉友 徐跃进

封面设计:傅瑞学

责任校对:李建庄

责任印制:李红英

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:24.5

字 数:615 千字

版 次:2012 年 12 月第 1 版

印 次:2012 年 12 月第 1 次印刷

印 数:1~3000

定 价:39.00 元

产品编号:042628-01

出版说明

信息时代早已显现其诱人魅力,当前几乎每个人都随身携有多个媒体、信息和通信设备,享受其带来的快乐和便捷。

我国高等教育早已进入大众化教育时代,而且计算机技术发展很快,知识更新速度也在快速增长,社会对计算机专业学生的专业能力要求也在不断翻新。这就使得我国目前的计算机教育面临严峻挑战。我们必须更新教育观念——弱化知识培养目的,强化对学生兴趣的培养,加强培养学生理论学习、快速学习的能力,强调培养学生的实践能力、动手能力、研究能力和创新能力。

教育观念的更新,必然导致教材的更新。一流的计算机人才需要一流的名师指导,而一流的名师需要精品教材的辅助,而精品教材也将有助于催生更多一流名师。名师们在长期的一线教学改革实践中,总结出了一整套面向学生的独特的教法、经验、教学内容等。本套丛书的目的就是推广他们的经验,并促使广大教育工作者进一步更新教育观念。

在教育部相关教学指导委员会专家的帮助和指导下,在各大大学计算机院系领导的协助下,清华大学出版社规划并出版了本系列教材,以满足计算机课程群建设和课程教学的需要,并将各重点大学的优势专业学科的教育优势充分发挥出来。

本系列教材行文注重趣味性,立足课程改革和教材创新,广纳全国高校计算机专业一线优秀名师参与,从中精选出佳作予以出版。

本系列教材具有以下特点。

1. 有的放矢

针对计算机专业学生并站在计算机课程群建设、技术市场需求、创新人才培养的高度,规划相关课程群内各门课程的教学关系,以达到教学内容互相衔接、补充、相互贯穿和相互促进的目的。各门课程功能定位明确,并去掉课程中相互重复的部分,使学生既能够掌握这些课程的实质部分,又能节约一些课时,为开设社会需求的新技术课程准备条件。

2. 内容趣味性强

按照教学需求组织教学材料,注重教学内容的趣味性,在培养学习观念、学习兴趣的同时,注重创新教育,加强“创新思维”和“创新能力”的培养、训练,强调实践,案例选题注重实际和兴趣度,大部分课程各模块的内容分为基本、加深和拓宽内容3个层次。

3. 名师精品多

广罗名师参与,对于名师精品,予以重点扶持,教辅、教参、教案、PPT、实验大纲和实

验指导等配套齐全,资源丰富。同一门课程,不同名师分出多个版本,方便选用。

4. 一线教师亲力

专家咨询指导,一线教师亲力;内容组织以教学需求为线索;注重理论知识学习,注重学习能力培养,强调案例分析,注重工程技术能力锻炼。

经济要发展,国力要增强,教育必须先行。教育要靠教师和教材,因此建立一支高水平的教材编写队伍是社会发展的需要,特希望有志于教材建设的教师能够加入到本团队。通过本系列教材的辐射,培养一批热心为读者奉献的编写教师团队。

清华大学出版社

前 言

数据结构与算法内容丰富,技巧性强,包含了计算机科学与技术的许多重要方面,对学生的计算机软件素质的培养作用明显。

本书采用 C++ 面向对象的观点介绍数据结构与算法,并使用模板程序设计技术,与传统采用面向过程的观点相比优势较大,使所设计的程序更容易实现代码重用,在提供通用性和灵活性的同时,又保证了效率。本书将面向对象程序设计的思想融合到数据结构与算法中,读者通过学习可进一步提高面向对象程序设计的能力。

全书共分为 11 章,第 1 章是基础知识,介绍基本概念及其术语、抽象数据类型的实现,还讨论算法的概念和算法分析的简单方法。作为预备知识,读者应具有一定的 C++ 程序设计的基础。但为了降低读者的门槛,本章还介绍了要用到的 C++ 的主要知识点,并介绍了实用程序软件包。

第 2 章引入线性表,详细讨论线性表的顺序存储结构与链式存储结构。在讨论链式存储结构时,首先仿照传统方法实现线性表,在此基础之上,在链表结构中保存当前位置和元素个数,这样在难度增加不大的情况下提高了算法效率,使学生逐步体会改进算法的途径与方法。

第 3 章介绍栈和队列,讨论栈和队列的顺序存储结构与链式存储结构,用栈实现表达式求值,通过学习能掌握各种栈和队列的实现与使用方法,对后继课程(如操作系统原理和编译原理)的学习打下良好基础。本章还讨论了优先队列,使队列应用更加广泛。

第 4 章介绍串,详细讨论串的存储结构与模式匹配算法,为开发串应用软件(如实现文本编辑软件)打下坚实的基础。

第 5 章介绍数组和广义表,详细讨论数组、特殊矩阵、稀疏矩阵和广义表的存储结构及实现方法,首次提出了广义表的使用空间表存储结构,并使用广义表实现 m 元多项式的表示。

第 6 章介绍树结构,讨论了二叉树、线索二叉树、树、森林及其哈夫曼树的结构及其实现,还用哈夫曼编码实现了压缩软件。

第 7 章介绍图结构,实现图的常用存储结构,并讨论图的相关应用,实现相应算法(如求最小生成树的 Prim 算法与 Kruskal 算法,求最短路径的 Dijkstra 算法与 Floyd 算法)。

第 8 章介绍查找,讨论静态查找表、动态查找表与散列表,还讨论二叉排序树、二叉平衡树与 B 树,并实现所有算法。

第9章介绍排序,以简洁方式实现各种排序算法,还测试了各种排序算法的实际运行时间。

第10章介绍文件,讨论主存储器和辅助存储器,以及各种常用文件结构,还特别介绍在数据库中经常采用的VSAM文件,对读者研究与学习数据库有一定的帮助。

第11章介绍算法设计技术与算法分析技术,详细讨论各种算法设计技术的使用方法并实现了各种算法,并对算法分析技术进行深入浅出的讨论。对读者的算法设计和算法分析的理论和实践都有极大的帮助。

对于初学者,要完全独立编写数据结构与算法的代码是相当困难的,因此本书讨论的数据结构与算法都加以实现并进行了严格测试,提供了完整的测试程序,读者可参考这些测试程序编写相关算法,但如果只会使用已有的数据结构编写简单的程序也不利于读者对数据结构与算法的深入理解,提高研究新数据结构与算法的能力。因此本书的习题不但包括基本练习题,还包括仿照书中数据结构构造新数据结构的题目,或改造已有算法的题目,这样使读者具有构造新结构及改造或改进算法的能力。

本书各章还提供了实例研究,这些实例研究包含教材基础内容的应用(例如第4章的实例研究——文本编辑,第3章的实例研究——表达式求值等),也包含教材正文内容的补充(例如第6章的实例研究——树与等价关系)。实例研究主要提供给那些精力充沛的学生深入学习与研究,读者通过对实例研究的学习,可提高实际应用数据结构与算法的能力,虽然有一定的难度,但应比读者的想象更易学习与掌握。

为了尽快提高读者的学习能力,本书各章还提供了深入学习导读,包括本书作者实现相关数据结构与算法的最原始思想的资料来源,也包括进一步学习的参考资料,极大地方便了读者与教师查阅资料。

为了加强实践教学的需要,本书的附录中提供了实验题目与课程设计项目,并提供了实验报告格式与课程设计报告格式,以便学生在做数据结构与算法的实验与课程设计时选择,也为教师提供了可供参考的实验与课程设计的素材。

本教材在内容组织上特别考虑了读者的可接受性;在算法实现时,重点考虑程序的可读性,教材基本内容的算法一般选择最简单实现方式,学生容易接受;一般采用启发的方式在习题、上机实验或课程设计中进一步实现更强的功能,这样容易培养起读者的学习兴趣,使读者感到自己具有能发展或改进已有算法的能力,也会使读者感到自己已具有计算机高手的实力。

本书的作者都活跃在教学研究第一线上,有的作者还具有深厚的数学功底,不但完成了所有算法的测试程序,还对算法分析的相关公式进行了严格的数学推理。

本书采用了模板程序设计技术,模板技术已成为现代C++语言的风格基础,C++ 98(1998年标准化的C++)提供的标准程序库中有80%的成分是使用模板机制实现的STL(Standard Template Library,标准模板库)。而国内现阶段教学并未重视C++的模板程序设计,书籍资料也不是很多。作者认为在C++中,只要模仿本书算法,读者会在不知不觉中就掌握了模板程序设计技巧。

现在来讨论一下在国外数据结构与算法教材中上机时喜欢采用的STL。实际上,STL是AT&T贝尔实验室和HP研究实验室的研究人员将模板程序设计和面向对象程

序设计的原理结合起来,创造的一套研究数据结构与算法的统一方法,现在已成为 C++ 标准库的一部分。STL 提供了实现数据结构的新途径。它将(数据)结构(即组织数据的存储结构)抽象为容器。通过使用模板和迭代器,STL 库使程序员能够将广泛的通用算法应用到各种容器类上。通过本书作者的研究与了解,STL 库只覆盖了数据结构中的线性结构和树结构,并没有覆盖图部分,因此,对数据结构来讲,STL 库并不完备,同时,如果读者上机编程都只使用 STL 库解决数据结构的相关算法,可能使读者在数据结构编程方面,只会使用 STL 库,而不能独自设计新数据结构。本书采用模板方法实现了所有书中的数据结构算法,应比 STL 库更完备;同时 STL 库中包含的源代码可读性差,不适合作为教学使用,本书的算法源程序首要强调可读性,使读者容易接受与模仿,并且可进行改进或修改算法实现,因此在某种意义上讲,本书提供的关于数据结构与算法实现的类模板与函数模板是一种 GTL(General Template Library)或 OSGTL(Open Source General Template Library),读者也可从作者个人主页提供的软件包(具体内容见附录 B)来进行实际数据结构与算法方面的软件开发;当然,通过本书的学习,再返回来学习 STL 库的应用,将会事半功倍,读者只要找一本介绍 STL 的书籍或网上找一些介绍使用 STL 库的文档,并用 STL 库试着编程即可完全掌握 STL 库的使用。

特别要提一下有关 C++ 编译器的问题。在 C++ 之外的任何编程语言中,编译器都没有受到过如此之重视,这是因为 C++ 是一门非常复杂的语言,以致编译器也难于构造。下面介绍一些常用的优秀 C++ 编译器。

Visual C++ 编译器:由微软开发,现在主要流行 Visual C++ 6.0、Visual C++ 2005 以及 Visual C++ 2005 Express,特点是集成开发环境用户界面友好,信息提示准确,调试方便,对模板支持最完善;Visual C++ 6.0 对硬件环境要求低,现在安装的机器最多,但对标准 C++ 兼容只有 83.43%,Visual C++ 2005 与 Visual C++ 2005 Express 在软件提示信息上做了进一步的优化与改进,并且对标准 C++ 兼容程度达到了 98% 以上,但对硬件的要求较高;还有 Visual C++ 2005 Express 是一种轻量级的 Visual C++ 软件,易于使用。对于编程爱好者、学生和初学者来说是很好的编程工具,微软在 2006 年 4 月 22 日正式宣布 Visual Studio 2005 Express 版永久免费。

GCC 编译器:著名的开源 C++ 编译器。是在 UNIX 操作系统(例如 Linux)下编写 C++ 程序的首选,有非常好的可移植性,可以在非常广泛的平台上使用,也是编写跨平台、嵌入式程序很好的选择。GCC 3.3 与标准 C++ 兼容程度大概能够达到 96.15%。现已有一些移植在 Windows 环境下使用 GCC 编译器的 IDE(集成开发环境),例如 Dev-C++ 与 MinGW Developer Studio,其中 Dev-C++ 是能够让 GCC 在 Windows 下运行的集成开发环境,提供了与专业 IDE 相媲美的语法高亮、代码提示与调试等功能;MinGW Developer Studio 是跨平台下的 GCC 集成开发环境。目前支持 Windows、Linux 和 FreeBSD;根据作者的实际使用,感觉使用 GCC 编译器的 IDE 错误信息提示的智能较低,错误提示不太准确,还有就是对模板支持较差,对语法检查较严格,有些在 Visual C++ 编译器中编译通过的程序可能在 GCC 编译器的 IDE 还会显示有错误信息。

本书所有算法都同时在 Visual C++ 6.0、Visual C++ 2005、Visual C++ 2005 Express、Dev-C++ 和 MinGW Developer Studio 中通过测试。读者可根据实际情况选择

适当的编译器,建议选择 Visual C++ 6.0。

教师可采取多种方式来使用本书作为讲授数据结构、数据结构与算法分析、数据结构与算法设计、数据结构与算法等课程,应该根据学生的背景知识以及课程的学时数来进行内容的取舍。为满足不同层次的教学需求,本教材使用了分层的思想,分层方法如下:没加有星号 * 及双星号 * * 的部分是基本内容,适合所有读者学习;加有星号的部分是适合计算机专业的读者深入学习的选学部分;加有双星号的部分适合于感兴趣的同学研究,尤其适合于那些有志于 ACM 竞赛的读者加以深入研究。

作者为本书提供了全面的教学支持,如果在教学或学习过程中发现与本书有关的任何问题都可以与作者联系:youhongyue168@gmail.com,作者将尽力满足读者的要求,并可能将解答公布在作者的教学网站 <http://teachhelp.changeip.net:9988/>或 <http://teachhelp.3322.org:9988/>上。在教学网站上还将提供如下内容:

(1) 提供书中所有算法在 Visual C++ 6.0、Visual C++ 2005、Visual C++ 2005 Express、Dev-C++ 和 MinGW Developer Studio 开发环境中的测试程序,今后还会提供当时流行的 C++ 开发环境的测试程序,每种开发环境还将提供基本开发过程的文档;还提供本书作者开发的软件包(包含所有本书所讲的数据结构与算法的类模板与函数模板)。

(2) 提供教学用 Power Point 幻灯片 ppt 课件。

(3) 向教师提供所有习题、上机实验题与课程设计项目的解答或参考程序;对学生来讲,将在每学期期末(第 15 周~第 20 周)公布解压码。

(4) 数据结构与算法问答专栏。

(5) 提供至少 8 套数据结构与算法模拟试题及其解答,以供学生期末及考研复习,也可供教师出考题时参考。

(6) 提供数据结构与算法相关的其他资料(例如 Dev-C++ 与 MinGW Developer Studio 软件,流行免费 C++ 编译器的下载网址等)。

希望各位能够抽出宝贵的时间将你对本教材的建议或意见,当然也可以发表对国内外的数据结构与算法课程教学的任何意见寄给作者,你的意见将是我们再版修订教材的重要参考。

张卫华、彭骏、谭斌、李培宇、何凯霖、姜琳、聂清彬、黄维、邹昌文、王文昌、周焯华、胡开文、沈洁、周德华与欧阳等人对本书做了大量的工作,包括提供资料,调试算法,参与了部分内容的编写,在此特向他们表示感谢;作者还要感谢为本书提供直接或间接帮助的每一个朋友,由于你们热情的帮助或鼓励,激发了作者写好本书的信心以及写作热情。

本书的出版要感谢清华大学出版社各位编辑及评审专家,由于他们为本书的出版倾注了大量热情,也由于他们具有前瞻性的眼光才让读者有机会看到本书。

尽管作者有良好而负责任的严格态度,并做出了最大努力,但由于作者水平有限,书中难免有不妥之处,因此,敬请各位读者不吝赐教,以便作者有一个提高的机会,并在再版时尽量采用你们的意见。

编 者

2012 年 10 月

目 录

第 1 章 绪论	1
1.1 数据结构的概念和学习数据结构的必要性	1
1.2 数据结构的基本概念	2
1.2.1 数据	2
1.2.2 数据元素和数据项	2
1.2.3 数据结构	2
1.3 抽象数据类型及其实现	4
1.3.1 数据类型	4
1.3.2 抽象数据类型(Abstract Data Type——ADT)	4
1.3.3 C++ 的类和对象	4
1.3.4 运算符重载	6
1.3.5 有关 C++ 的动态存储分配	9
1.3.6 C++ 的模板(template)	10
1.4 算法和算法分析	11
1.4.1 算法	11
1.4.2 算法分析	12
* 1.5 实用程序软件包	15
1.6 深入学习导读	19
1.7 习题 1	19
第 2 章 线性表	21
2.1 线性表的逻辑结构	21
2.2 线性表的顺序存储结构	23
2.3 线性表的链式存储结构	31
2.3.1 单链表	31
2.3.2 循环链表	39
2.3.3 双向链表	43
* 2.3.4 在链表结构中保存当前位置和元素个数	46

* 2.4 实例研究：一元多项式的表示	55
2.5 深入学习导读	60
2.6 习题 2	60
第 3 章 栈和队列	61
3.1 栈	61
3.1.1 栈的基本概念	61
3.1.2 顺序栈	62
3.1.3 链式栈	67
3.2 队列	74
3.2.1 队列的基本概念	74
3.2.2 链队列	75
3.2.3 循环队列——队列的顺序存储结构	79
* 3.2.4 队列应用——显示二项式 $(a+b)^i$ 的系数	84
* 3.3 实例研究：表达式求值	85
3.4 深入学习导读	88
3.5 习题 3	89
第 4 章 串	90
4.1 串类型的定义	90
4.2 字符串的实现	91
4.3 字符串模式匹配算法	98
4.3.1 简单字符串模式匹配算法	98
** 4.3.2 KMP 字符串模式匹配算法	99
** 4.4 实例研究：文本编辑	104
4.5 深入学习导读	114
4.6 习题 4	114
第 5 章 数组和广义表	115
5.1 数组	115
5.1.1 数组的基本概念	115
5.1.2 数组的顺序表	116
** 5.1.3 数组的类定义	118
5.2 矩阵	122
5.2.1 矩阵的定义和操作	122
5.2.2 特殊矩阵	123
5.2.3 稀疏矩阵	128
5.3 广义表	139

5.3.1 基本概念	139
* 5.3.2 广义表的存储结构	141
5.4 深入学习导读	151
5.5 习题 5	152
第 6 章 树和二叉树	153
6.1 树的基本概念	153
6.1.1 树的定义	153
6.1.2 基本术语	154
6.2 二叉树	155
6.2.1 二叉树的定义	155
6.2.2 二叉树的性质	157
6.2.3 二叉树的存储结构	160
6.3 二叉树遍历	168
6.3.1 遍历的定义	168
6.3.2 遍历算法	169
* 6.3.3 二叉树遍历应用举例	175
6.4 线索二叉树	180
6.4.1 线索的概念	180
* 6.4.2 线索二叉树的实现	182
6.5 树和森林	190
6.5.1 树的存储表示	190
* 6.5.2 树的显示	198
6.5.3 森林的存储表示	198
6.5.4 树和森林的遍历	203
6.5.5 树和森林与二叉树的转换	206
6.6 哈夫曼树与哈夫曼编码	208
6.6.1 哈夫曼树的基本概念	209
6.6.2 哈夫曼树构造算法	210
6.6.3 哈夫曼编码	210
* 6.6.4 哈夫曼树的实现	212
** 6.7 树的计数	216
** 6.8 实例研究: 树与等价关系	218
6.9 深入学习导读	222
6.10 习题 6	222
第 7 章 图	224
7.1 图的定义和术语	224

7.2	图的存储表示	228
7.2.1	邻接矩阵	228
7.2.2	邻接表	234
7.3	图的遍历	242
7.3.1	深度优先搜索	242
7.3.2	广度优先搜索	244
7.4	连通无向网的最小代价生成树	246
7.4.1	Prim 算法	246
7.4.2	Kruskal 算法	249
7.5	有向无环图及应用	252
7.5.1	拓扑排序	253
7.5.2	关键路径	256
7.6	最短路径	260
7.6.1	单源点最短路径问题	260
7.6.2	所有顶点之间的最短路径	263
7.7	深入学习导读	265
7.8	习题 7	265
第 8 章	查找	267
8.1	查找的基本概念	267
8.2	静态表的查找	270
8.2.1	顺序查找	270
8.2.2	有序表的查找	271
8.3	动态查找表	274
8.3.1	二叉排序树	274
* 8.3.2	二叉平衡树	285
* 8.3.3	B-树和 B ⁺ -树	310
8.4	散列表	312
8.4.1	散列表的概念	312
8.4.2	构造散列函数的方法	313
8.4.3	处理冲突的方法	313
* 8.4.4	散列表的实现	315
8.5	深入学习导读	320
8.6	习题 8	320
第 9 章	排序	321
9.1	概述	321
9.2	插入排序	322

9.2.1 直接插入排序	322
9.2.2 Shell 排序	324
9.3 交换排序	325
9.3.1 起泡排序	325
9.3.2 快速排序	326
9.4 选择排序	329
9.4.1 简单选择排序	330
9.4.2 堆排序	331
9.5 归并排序	334
* 9.6 基数排序	338
9.6.1 多关键排序	338
9.6.2 基数排序	339
* 9.7 各种内部排序方法讨论	341
* 9.8 外部排序	342
9.8.1 外部排序基础	342
9.8.2 外部排序的方法	343
* 9.9 实例研究: 各种排序算法运行时间测试	344
9.10 深入学习导读	347
9.11 习题 9	347
* 第 10 章 文件	349
10.1 主存储器和辅助存储器	349
10.2 各种常用文件结构	349
10.2.1 顺序文件	349
10.2.2 索引文件	350
10.2.3 散列文件	351
10.2.4 VSAM 文件	351
10.2.5 多关键字文件	352
10.3 深入学习导读	354
10.4 习题 10	354
第 11 章 算法设计与分析	355
11.1 算法设计	355
11.1.1 递归算法	355
11.1.2 分治算法	357
* 11.1.3 回溯算法	358
11.2 算法分析	361
11.2.1 递归分析	361

**11.2.2 利用生成函数进行分析	362
11.3 深入学习导读	363
11.4 习题 11	364
附录 A 调和级数	365
附录 B 课本的软件包	366
附录 C 实验题目	371
附录 D 课程设计项目	372
D1 算术表达式求值	372
D2 简单本文编辑器	372
D3 压缩软件	373
**D4 公园导游系统	373
*D5 专家系统应用——动物游戏	373
*D6 词典变位词检索系统	374
附录 E 实验报告格式	375
附录 F 课程设计报告格式	376
参考文献	377

第 1 章 绪 论

随着计算机功能越来越强大,人们就越要尝试解决更加复杂的问题,而更复杂的问题需要更大的计算量,这使得对程序的运行效率有更高的要求,使得从事软件开发的人员必须学习和具备彻底理解隐藏在程序设计后面的一般原理——数据结构和算法。

虽然数据结构与算法的原理和方法是独立于具体的计算机语言,然而只能使用具体的某种计算机语言才能在计算机上实现,本书采用目前普遍使用的 C++ 程序设计语言来描述各种数据结构与算法,为使读者更好地理解教材,本章介绍 C++ 的基本结构和语法。

1.1 数据结构的概念和学习数据结构的重要性

对于数值计算问题的解决方法,主要是用数学方程建立数学模型,例如预测人口增长的数学模型为常微分方程,求解这些数学模型的方法是计算数学研究的范畴。

对于非数值计算问题,主要采用数据结构的方法建立数学模型,下面通过实例加以说明。

例 1.1 在人事管理系统中,包含有“员工基本信息”表格,包括了许多员工基本信息记录(例如包含有编号,姓名,性别,籍贯,家庭住址,生日,如表 1.1 所示),将这些记录按照一定的顺序存放在“员工基本信息”表格中,每个员工基本信息记录按顺序排列,形成员工基本信息记录的线性序列,这是一种最简单的线性表结构。

表 1.1 员工基本信息

编 号	姓 名	性 别	籍 贯	家 庭 住 址	生 日
1001	刘倩	女	北京	人民南路 16 号	1985.12.18
1002	朱洪顺	男	成都	一环路北 3 段 56 号	1986.6.28
1003	李世红	男	太原	二环路东 6 段 168 号	1983.10.16
1004	陈冠杰	男	杭州	解放路 18 号	1982.11.29
1005	游倩华	女	苏州	人民西路 98 号	1988.6.8
1006	林键忠	男	青岛	人民西路 69 号	1986.2.28
1007	李代靖	女	太原	一环路东 6 段 16 号	1985.3.19

例 1.2 典型的 UNIX 文件系统结构如图 1.1 所示,属于树结构,是一棵倒置的“树”,此处“树根”代表整个系统,用根目录/表示;下一层表示子系统,如 bin、lib、user 等,

“叶子”就是文件,如 LinkList.h、SqList.h 等。

例 1.3 要在 n 个网站建立通信网络,如图 1.2 所示,这些网站之间形成一种图结构。

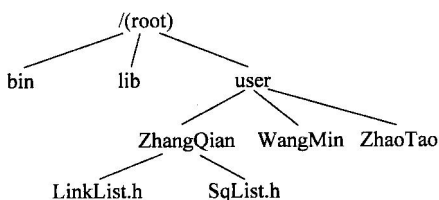


图 1.1 UNIX 文件系统结构图

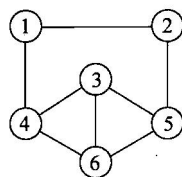


图 1.2 网站之间的“图”状结构

从上面的实例可以看出非数值计算问题的数学模型已不是数学方程,而是线性表、树和图等数据结构,简单地说,数据结构的研究范畴主要是非数值计算问题的操作对象和它们之间的关系以及在计算机中的表示和实现。

在选择数据结构解决特别问题时,只有通过预先分析问题来确定必须达到的性能目标,才有可能挑选出恰当的数据结构,与问题不相称的数据结构,会导致设计的程序效率低,当使用简单的设计就能达到目标时,选择复杂的数据结构来改进程序也没有必要。

1.2 数据结构的基本概念

本节介绍数据结构的基本概念,根据作者的经验,初学者对这些概念在开始时 would 感到非常抽象,难于理解,但随着不断地学习,一定会融会贯通并加以深刻理解。

1.2.1 数据

数据是客观事物的符号表示,是计算机中可以操作的对象,也就是一切能输入到计算机中并能被处理的符号的总称。

数据可以指数值型数据,例如整数、实数和复数等,这些数据主要应用于工程计算,相信读者都比较熟悉;也可以是非数值型数据,例如文字、图形、语音等数据。

1.2.2 数据元素和数据项

数据元素一般在计算机中作为整体进行处理,是数据的基本单位,数据元素也被称为记录,有的数据元素由若干**数据项**所组成,比如在员工基本信息表中,每个员工记录是一个数据元素,而员工的编号、姓名、性别、籍贯、家庭住址和生日等内容为数据项。

1.2.3 数据结构

在现实世界中,不同数据元素之间不是独立的,而是存在着特定的关系,我们将这些