

新世纪高等院校精品教材

数据结构与算法分析 (悦语言版)

魏宝刚 陈越 王申康 编著

内 容 简 介

本书描述了各种类型的数据结构,包括线性表、树、堆、图,以及查找、排序等算法。自始至终将数据结构的基本原理与算法分析紧密结合,强调了算法性能的重要性,并介绍了算法设计技术。主要以 悦语言描述算法,列举了大量的实例,便于在计算机上实际运行、分析各种算法。每章之后附有习题以备读者进一步练习。

本书逻辑性强、内容新颖、全面,可以作为大专院校计算机专业的教材和参考书,也可供其它理工科专业学生和计算机工程技术人员参考。

图书在版编目 (CIP) 数据

数据结构与算法分析 : 悦语言版 / 魏宝刚, 陈越, 王
申康编著. — 杭州 : 浙江大学出版社, 2009.12
新世纪高等院校精品教材
ISBN 978-7-309-07111-1

I. ①数... II. ①魏...②陈...③王... III. ①数据
结构 ②算法分析 ③悦语言 ④程序设计
IV. ①TP311.1②TP311.1③TP311.1④TP311.1⑤TP311.1⑥TP311.1⑦TP311.1⑧TP311.1⑨TP311.1⑩TP311.1⑪TP311.1⑫TP311.1⑬TP311.1⑭TP311.1⑮TP311.1⑯TP311.1⑰TP311.1⑱TP311.1⑲TP311.1⑳TP311.1㉑TP311.1㉒TP311.1㉓TP311.1㉔TP311.1㉕TP311.1㉖TP311.1㉗TP311.1㉘TP311.1㉙TP311.1㉚TP311.1㉛TP311.1㉜TP311.1㉝TP311.1㉞TP311.1㉟TP311.1㊱TP311.1㊲TP311.1㊳TP311.1㊴TP311.1㊵TP311.1㊶TP311.1㊷TP311.1㊸TP311.1㊹TP311.1㊺TP311.1㊻TP311.1㊼TP311.1㊽TP311.1㊾TP311.1㊿

中国版本图书馆 CIP 数据核字 (2009) 第 123456 号

责任编辑 杜希武 钱欣平

封面设计 俞亚彤

出版发行 浙江大学出版社

(杭州浙大路 33 号) 邮政编码 310027

(电话: 0571-87959389) 电子邮箱: zjupress@163.com

(网址: www.zjupress.com) 印刷厂: 杭州求是印刷厂

排版 浙江大学出版社电脑排版中心

印刷 德清第二印刷厂

开本 787mm×1092mm 1/16 印张 18.5

印数 00001—00000

字数 350 千字

版次 2009 年 12 月第 1 版 2009 年 12 月第 1 次印刷

印数 00001—00000

书号 ISBN 978-7-309-07111-1

定价 28.00 元

前 言

《数据结构和算法》是计算机科学与技术专业的主要专业基础课,也是信息技术的重要理论基础,它所讨论的知识内容和提倡的技术方法,无论对进一步学习计算机领域的其他课程,还是对从事大型信息工程的开发,都有着枢纽的作用。本书作者在从事多年数据结构教学和大量科研实践的经验中,领悟到数据结构和算法在未来 工作 者中的实际需求,所以本书的编写内容着重于理论和实际应用的紧密结合。

数据结构是计算机存储、组织数据的方式。一般选择合适的数据结构可以带来更高的运行或者存储效率的算法。许多大型系统的构造经验表明,系统实现的困难程度和系统构造的质量都严重地依赖于是否选择了最优的数据结构。所以数据结构的选择是一个基本的设计考虑因素。许多时候,确定了数据结构后,算法就容易得到了。然而,有些时候事情也会反过来,我们根据特定算法来选择数据结构与之适应。所以数据结构和与之相关的算法是不可隔离的。本书的编写在内容的安排上,以交融的方式同时导出数据结构和算法。这样学生在数据结构的学习中能知其所用,在算法设计的学习中也不会枯燥无味。这也可以算是本书的特点吧。

在数据结构表达语言的选择上,面对目前数据结构教程所用的 等程序设计语言,我们采用了 悦程序设计语言。这是合乎我国目前教育的国情,我国学生一年级的程序设计语言课可能没有或极个别是讲 的, 的教学更要滞后,对计算机专业的学生在继后的操作系统、数据库等学习中可能更需要用 悦程序设计语言,所以我们选择了 悦程序设计语言,毕竟数据结构的真谛不在于程序设计语言。

《数据结构和算法》作为一门课程,主要目的是使学生较全面地理解算法和数据结构的概念,掌握各种数据结构与算法的实现方式,比较不同数据结构和算法的特点。通过学习,使学生能够提高用计算机解决实际问题的能力。

本书适合于为本科生开设的数据结构与算法课程。学生应该具有中等程度的程序设计知识,包括像指针和递归这样一些内容,还要具有离散数学的某些知识。同时本书也不失为 工作者的良师益友。

第 章介绍数据结构和算法的基础知识,并提供 悦语言程序设计的一些复习材料。第 章阐述算法分析方法,以例子说明渐进分析及其主要特点,介绍了时间复杂度的测试方法。第 章除讨论线性表的一般表示形式外,还重点讨论堆栈和队列两种重要的表结构。第 章重点在搜索树,包括 和离散集合的表示与应用。第 章主要讨论静态查找和动态查找,以及哈希映射技术等经典的查找算法。第 章是关于优先队列的内容,主要阐述各种二叉堆。第 章讲排序,重点对八种排序算法详细地进行了分析 插入排序、希尔排序、堆排序、快速排序、归并排序、基数排序和表排序。另外对外排序也进行了描述。第 章讲授图的表示和算法。以实践应用为背景阐述算法效率对数据结构的依赖性。第 章通过讨论实际应用问题的求解技巧介绍算法设计技术,重点讨论五个经典算法和相应的分析。

读者在阅读中可以根据自身的基础合适地选择相应的章节,对熟悉 悦语言的读者完全可

以跳过第一章的有关 悦语言基础一节,对目录中带 * 号的章节一般读者可以先放到一边,作为扩充内容,以后再学习。

全书由浙江大学计算机科学与技术专业博士生导师王申康教授组织编著和审定,第 员圆猿猿源愿章由魏宝刚副教授编写,第 缘苑怨章由陈越教授编写。朱文浩硕士参与第 员章 悦语言基础部分编写。

由于作者水平所限,时间仓促,书中不当之处在所难免,敬请读者批评指正。

作 者

圆园园源年 缘月 圆日

目 录

第 1 章 基础知识	1
1.1 数据结构与算法	1
1.2 抽象数据类型	1
1.3 悦语言程序设计基础	1
1.3.1 数组	1
1.3.2 指针	1
1.3.3 结构体和共用体	1
1.3.4 函数与参数	1
1.3.5 递归函数	1
1.3.6 局部变量和全局变量	1
习题 1	1
第 2 章 算法分析	2
2.1 算法的定义	2
2.2 空间复杂度	2
2.3 时间复杂度	2
2.3.1 程序步	2
2.3.2 最好、最差和平均性能	2
2.3.3 近似方法 (Ω, Θ)	2
2.4 时间复杂度的测试	2
习题 2	2
第 3 章 线性表、堆栈和队列	3
3.1 线性表	3
3.1.1 线性表的定义	3
3.1.2 线性表的数组表示	3
3.1.3 线性表的链表表示	3
3.1.4 稀疏矩阵与多重表	3
3.2 堆栈	3
3.2.1 迷宫问题	3
3.2.2 堆栈的定义	3
3.2.3 堆栈的实现	3
3.3 队列	3

缘猿猿 猿文猿对	猿猿猿
缘猿猿 猿伸展树 猿猿猿猿猿	猿猿猿
猿猿猿 猿哈映射	猿猿猿
缘猿猿 猿概述	猿猿猿
缘猿猿 猿哈希表	猿猿猿
缘猿猿 猿哈希函数	猿猿猿
缘猿猿 猿冲突处理	猿猿猿
猿猿题 缘	猿猿猿
第 猿章 猿堆 (优先队列)	猿猿猿
猿猿猿 猿堆的定义和表示	猿猿猿
猿猿猿 猿最大堆	猿猿猿
猿猿猿 猿最大堆的插入	猿猿猿
猿猿猿 猿最大堆的删除	猿猿猿
猿猿猿 猿最大堆的建立	猿猿猿
猿猿猿 猿最小 猿最大堆 (配猿猿猿猿猿)	猿猿猿
猿猿猿 猿配猿猿猿猿猿的插入	猿猿猿
猿猿猿 猿配猿猿猿猿猿的删除	猿猿猿
猿猿猿 猿左右堆 (猿猿猿猿)	猿猿猿
猿猿猿 猿猿猿猿的定义	猿猿猿
猿猿猿 猿猿猿猿的插入	猿猿猿
猿猿猿 猿猿猿猿中最大元素的删除	猿猿猿
猿猿猿 猿左高堆 (猿猿猿猿猿猿猿)	猿猿猿
猿猿猿 猿左高堆的性质	猿猿猿
猿猿猿 猿左高堆的操作	猿猿猿
猿猿题 猿	猿猿猿
第 猿章 猿排序	猿猿猿
猿猿猿 猿插入排序	猿猿猿
猿猿猿 猿算法	猿猿猿
猿猿猿 猿效率分析	猿猿猿
猿猿猿 猿其他改进	猿猿猿
猿猿猿 猿希尔排序	猿猿猿
猿猿猿 猿算法	猿猿猿
猿猿猿 猿效率分析	猿猿猿
猿猿猿 猿堆排序	猿猿猿
猿猿猿 猿算法	猿猿猿
猿猿猿 猿效率分析	猿猿猿
猿猿猿 猿快速排序	猿猿猿
猿猿猿 猿算法	猿猿猿

第 1 章 绪论	1
1.1 数据结构与算法分析	1
1.2 算法效率分析	1
1.3 渐近符号	1
1.4 递归与分治策略	1
1.5 本章小结	1
习题 1	1
第 2 章 线性表	2
2.1 线性表的定义	2
2.2 线性表的表示	2
2.3 线性表的实现	2
2.4 线性表的应用	2
2.5 本章小结	2
习题 2	2
第 3 章 栈和队列	3
3.1 栈的定义	3
3.2 栈的表示	3
3.3 栈的应用	3
3.4 队列的定义	3
3.5 队列的表示	3
3.6 队列的应用	3
3.7 本章小结	3
习题 3	3
第 4 章 树和二叉树	4
4.1 树的定义	4
4.2 树的表示	4
4.3 二叉树的遍历	4
4.4 二叉树的应用	4
4.5 本章小结	4
习题 4	4
第 5 章 图	5
5.1 图的定义	5
5.2 图的表示	5
5.3 图的遍历	5
5.4 最短路径问题	5
5.5 最小生成树	5
5.6 网络流	5
5.7 本章小结	5
习题 5	5
第 6 章 排序	6
6.1 插入排序	6
6.2 希尔排序	6
6.3 选择排序	6
6.4 归并排序	6
6.5 快速排序	6
6.6 基数排序	6
6.7 本章小结	6
习题 6	6
第 7 章 查找	7
7.1 顺序查找	7
7.2 二分查找	7
7.3 哈希查找	7
7.4 二叉查找树	7
7.5 平衡二叉树	7
7.6 本章小结	7
习题 7	7
第 8 章 算法设计技术	8
8.1 贪心法	8
8.2 动态规划	8
8.3 回溯法	8
8.4 分支限界法	8
8.5 本章小结	8
习题 8	8

第 1 章 绪论	1
1.1 问题的提出	1
1.2 问题的分类	2
1.3 问题的求解方法	3
1.4 问题的求解步骤	4
1.5 问题的求解结果	5
1.6 问题的求解评价	6
1.7 问题的求解展望	7
第 2 章 问题的求解方法	8
2.1 问题的求解方法概述	8
2.2 问题的求解方法分类	9
2.3 问题的求解方法选择	10
2.4 问题的求解方法应用	11
2.5 问题的求解方法评价	12
2.6 问题的求解方法展望	13
第 3 章 问题的求解结果	14
3.1 问题的求解结果概述	14
3.2 问题的求解结果分类	15
3.3 问题的求解结果选择	16
3.4 问题的求解结果应用	17
3.5 问题的求解结果评价	18
3.6 问题的求解结果展望	19
第 4 章 问题的求解评价	20
4.1 问题的求解评价概述	20
4.2 问题的求解评价分类	21
4.3 问题的求解评价选择	22
4.4 问题的求解评价应用	23
4.5 问题的求解评价评价	24
4.6 问题的求解评价展望	25
第 5 章 问题的求解展望	26
5.1 问题的求解展望概述	26
5.2 问题的求解展望分类	27
5.3 问题的求解展望选择	28
5.4 问题的求解展望应用	29
5.5 问题的求解展望评价	30
5.6 问题的求解展望展望	31
参考文献	32

基础知识

1.1 数据结构与算法

信息表示是计算机科学的基础,尽管如今的计算机能够处理视频、语音等各种各样的多媒体信息,但归根结底,这些信息都必须首先转换成计算机能够“理解”的数据。因此,研究和掌握适合于信息处理的计算机数据处理方法和技术是十分重要的。从实用角度讲,研究数据结构是为了支持高效的数据处理。

数据结构是计算机科学与技术领域广泛使用的术语。人们往往会认为数据结构是信息的一种组织方式,它用来反映数据的内部构成,即一个数据由哪些成分构成,以什么方式构成,呈什么结构。实际上这反映出了数据结构的一个方面。那么到底什么是数据结构?它的作用如何?这些都是首先要明确的问题。

在数据结构中,往往涉及数据类型与数据对象的概念,它们之间是存在差异的。一般来说,数据类型是指一种程序设计语言中,变量所具有的数据种类,如 C 语言中的整型、实型、字符型等基本数据类型,还包括用户利用基本数据类型定义的自定义数据类型,如 C 语言提供的结构体、共同体等。而数据对象则是指某种数据类型元素的集合,如整数的数据对象是 $\{ \dots, -1, 0, 1, \dots \}$, 字符类型的数据对象是 $\{ 'a', 'b', 'c', \dots, 'z' \}$ 。数据对象可以是有限的,也可以是无限的。

数据结构不同于数据类型,也不同于数据对象,它不仅涉及数据类型和数据对象,而且要描述数据对象各元素之间的相互关系,比如需要描述数据对象元素之间的运算。因此,比较全面地讲,数据结构包括下面三个方面的内容:

① 数据的逻辑结构——数据之间的逻辑关系;

② 数据的存储结构——数据元素及其关系在计算机存储器中的表示;

③ 数据的运算——对数据对象施加的操作。

数据的逻辑结构是从逻辑关系上描述数据,它包括线性结构和非线性结构两大类,且与数据的存储无关,是独立于计算机的。线性结构的特征是数据元素之间存在直接前驱和后继联系,如本书中将要讲述的线性表。非线性结构的特征是数据元素之间存在多个直接前驱和后继联系,如本书中将要讲述的树、图等。

数据的存储结构是逻辑结构在计算机内的表示,包含数据元素以及它们之间的关联。数据的存储结构主要有顺序存储、链式存储、索引存储和散列存储四种基本方式。

耘晖与刘景韶等在《云外世有真廣學房藏名兩難觀釋卷一》一书中,从数据抽象、算法说明和定义,以及性能分析和度量等方面介绍数据结构,目的是用系统论的方法在整个软件生命周期内定位数据结构。现在更多的作者更是将数据结构和算法分析融为一体,比如 悦蘭與劉景韶著的《粵互難難答問題是林廣學房藏名兩難觀釋卷二》和 悦蘭與劉景韶著的《粵互難難答問題是林廣學房藏名兩難觀釋卷三》。

人们可能误认为现代计算机硬件的性能已经很好了,如何组织数据以及程序的效率问题已不像从前那么重要了。实际上从软件的发展来看,正是由于硬件水平的提高,以前认为计算机难以有效求解的问题现在已被征服,计算机的应用也不断地向更复杂的领域扩展,随着求解问题复杂度的增加,高性能的算法仍然是计算机程序设计追求的目标。

抽象数据类型 (Abstract Data Type, ADT) 是数据描述的一种形式, 它将数据逻辑结构和一组作用于其上的操作或方法结合起来了, 并提倡将对象和操作的定义与实现相互分离。这种描述方法是与计算机无关的, 或者说与具体的程序设计语言无关, 因此, 可以采用灵活的方式, 例如文字、符号、伪代码等来描述数据和对数据的操作。

数据类型名 用来标识一个具体的抽象数据类型,通常可以取名为所要定义的数据结构的英语单词。比如,要定义二叉树,可以命名为 `BinaryTree`,树状图可以命名为 `TreeNode` 等。

操作 是实现数据对象处理的一系列算法,例如整型数据对象的操作包括通常的算术运算——加、减、乘、除、取余等。

例如,一个矩阵可以被看作是由 皂行 灶列共 皂伊灶个元素组成的数值序列,当指定一个具体的行列值 (蚤躁) 时,其中 皂伊蚤躁 皂伊蚤躁 灶矩阵有一个惟一的元素与之对应。对矩阵的操作包括矩阵的转置、矩阵的加法运算、矩阵的乘法运算等。矩阵的这些特性用抽象数据类型定义可以采用如下形式:

粤, 粤, 粤, ..., 粤, 粤, ..., 粤, 皂伊灶个元素构成的约瑟夫增建序列。每个元素是整数, 对于任意一个下标 蚤灶增有惟一的矩阵元素与之对应。蚤灶增是一个有序对 (蚤躁, 蚤和躁是整数, 其大小满足 员 ≤ 蚤 ≤ 皂, 员 ≤ 躁 ≤ 灶)

对于任意三个矩阵 A, B, C 有 $(A+B)C = AC + BC$ 和 $A(B+C) = AB + AC$

函数 `Matrix Create(max_row, max_col)` :: 创建一个最大行为 `max_row` 最大列为 `max_col` 的空矩阵

函数 `Matrix Transpose(a)` :: 若 `a` 是一个 `n` 行 `m` 列矩阵, 返回一个 `m` 行 `n` 列矩阵, 且满足 `Transpose(Transpose(a)) = a`

函数 `Matrix Add(a, b)` :: 若 `a` 和 `b` 的维数相同, 返回一个矩阵 `c` 且 `c[i][j] = a[i][j] + b[i][j]`; 否则返回矩阵维数不相同的错误信息

函数 `Matrix Multiply(a, b)` :: 若 `a` 的列数等于 `b` 的行数, 返回矩阵 `c` 且 `c[i][j] = Σ a[i][k] * b[k][j]`; 否则返回矩阵维数不相同的错误信息

从这个定义可以看出, 抽象数据类型是数据对象以及操作集合的统一体, 通过它简单地描述出客观实体完整的外部特征, 而并未涉及对象是如何存储以及各操作是如何实现的等内在特性。

从对象与操作的集合角度来看, 抽象数据类型描述方法与目前广泛使用的面向对象的思想是相吻合的。按抽象数据类型定义使得客观实体对外部来说是透明的, 外部调用只要满足操作的公共接口定义, 不必关心它们具体是如何实现的。

图 1-1 为矩阵的数据结构的图示表示, 它概括了矩阵的抽象数据类型的定义、矩阵的表示、算法以及外部调用之间的关系。从中可以看出, 抽象数据类型的定义在数据的表示、算法实现和外部调用 (对符合矩阵抽象数据类型的具体实体的使用) 之间架起了一座桥梁。当外部调用要完成矩阵运算时, 只要输入组成矩阵的具体数据, 调用接口提供的 `Matrix Create` 功能, 数据就会以一个特殊的存储形式被存入计算机内存; 调用 `Matrix Transpose` 和 `Matrix Add` 等运算操作就会完成相应的矩阵运算。而矩阵采用的物理存储方式到底是数组还是链表, 以及各操作如何实现、它的效率如何等, 对外部调用来说并不重要。

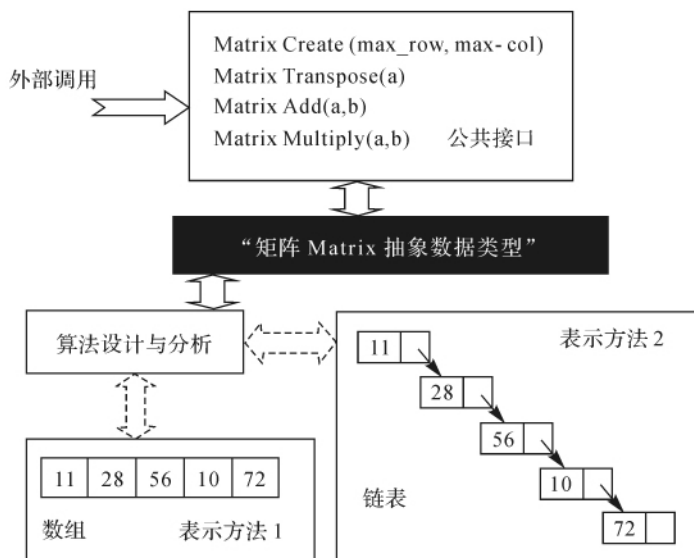


图 1-1 矩阵数据结构

员级悦语言程序设计基础

员媛瑶数组

数组是一个由若干相同类型变量组成的存储结构,它在内存中是一段连续的存储单元,最低地址对应于数组的第一个元素,最高地址对应于最后一个元素。由于数组不同于其他变量类型,在某些算法中,需要用数组来表示一些比较复杂的数据或是优化精简代码。

（员 一维数组

数组可以是一维的,也可以是多维的。其中最常用的是一维数组,在悦语言中一维数组的声明形式为:

摇摇类型说明符摇摇变量名 [长度值] ;

数组必须显示地说明,以便编译程序为其分配内存空间。类型说明符指明数组的类型,也就是数组中每一个元素的类型。例如,下列语句声明了两个数组:

摇摇float v 随 摇摇/* 数组 v 存放 猿个浮点数,对应的单元是 v 园、v 1 和 v 2

```
摇摇int a[10];    /* 数组 a 含有 10 个整型数据元素 */
```

这里需要注意的是，悦语言中的数组必须是静态的。换言之，数组的大小必须在程序运行前就被确定下来。因此数组声明中的长度值必须是常量，像下面程序段用变量 `len` 定义长度值是错误的：

```
摇摇int n ;
摇摇摇摇scanf ("% d", &n) ;
摇摇摇摇int a [n] ;
```

悦语言允许在说明时用数值表对全局数组和静态局部数组初始化,但不能对非静态局部数组初始化。数值表是一个由逗号分隔的常量表,这些常量的类型与类型说明符相容,第一个常量存入数组的第一个单元,第二个常量存入第二个单元,依次类推,各个常量对号入座,初始化每个数组单元。与其他变量相似,数组初始化的一般形式如下:

摇摇类型说明符摇摇变量名 [长度值] 越{变量值列表} ;

例如,对上面例子中的整型数组 `arr` 的初始化可以表示成如下形式:

摇摇int a [园 越园 员 圆 猿 源 缘 远 苑 愿 怨 ;

一维数组的总字节数由其元素类型以及数组长度决定,可按下式进行计算:

摇摇总字节数 越置集类型) 伊数组长度

若一个浮点数占用 源个字节,一个整数占用 圆个字节的话,前面定义的浮点类型数组 增和整型数组 葬的大小分别为 泽源枣源枣伊园越园以及 泽源枣源枣伊园越园。

对数组中元素的访问可通过下标来实现,下标的取值范围是从 0 至长度值 -1。例如下列函数对整型数组 `arr` 进行赋值:

```

void main ()
{
    int x [10];
    int k;
    for (k = 0; k < 10; k++)
        x[k] = k;
}

```

执行这个程序,编译器会在内存分配一个由 10 个整型单元组成的连续存储空间,并在循环赋值作用下将 0~9 十个整数依次存入各个单元。一维数组在本质上是由同类数据构成的线性表,图 1-1 形象地表示了执行 `main` 函数产生的结果。

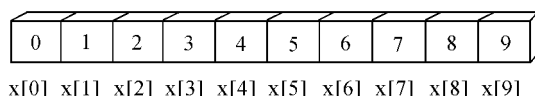


图 1-1 一维数组 `x` 的表示

这里需要注意的是,悦语言并不检验数组边界,因此,数组的两端都有可能越界而使其他变量的数据甚至程序代码被破坏。因此对于悦语言来说,数组的边界检验是程序员的职责。下面的程序可以说明这一点:

```

#include <stdio.h>

void main ()
{
    int x [10];
    int y;
    int k;
    for (k = 0; k < 10; k++)
        x[k] = k;
    y = x[10]; /* 这里 x[10] 的值被覆盖为 0 */
    for (k = 0; k < 10; k++)
        printf ("x [%d] = %d\n", k, x[k]);
}

```

该程序语句输出结果为:

```

x [0] = 0
x [1] = 1
x [2] = 2
x [3] = 3
x [4] = 4
x [5] = 5
x [6] = 6
x [7] = 7
x [8] = 8
x [9] = 9

```

程序中,访问 `x[10]` 实际上是非法的。事实上从内存的角度来看,假设数组 `x` 的起始地址是 `0x0000` 那么 `x[10]` 与 `x[9]` 都代表同一块内存 (对于悦编译器,上述程序中变量的内存分配是顺序进行的)。从实际效果来看,`x[10]` 可以看作是 `x[9]` 的一个引用。从输出结果中可以看到,虽然在将 `x[10]` 赋值为 0 以后,并没有再对其进行赋值操作,但是由于对 `x[9]` 进行了赋值,因此最后 `x[10]` 的值变为 9。

1.1.1 字符串数组

在悦语言中,字符串通常是以一个空字符 ‘`\0`’ 结尾的字符数组来表示的。这里的空字符

‘\0’在ASCII码表中的值为0它是不能省略的。因此,在说明字符数组时,必须比它要存放的最长字符串多一个字符。对于大多数C语言函数来说,操作一个缺少‘\0’结尾的字符串是非常危险的。表 10.1 是经常用到的一些字符串操作函数。

表 10.1 常用字符串操作函数

函数名称	功能
<code>strcpy(dest, src)</code>	将 <code>src</code> 拷贝到 <code>dest</code>
<code>strcat(dest, src)</code>	将 <code>src</code> 连接到 <code>dest</code> 的末尾
<code>strlen(src)</code>	返回 <code>src</code> 的长度
<code>strcmp(src1, src2)</code>	若 <code>src1</code> 与 <code>src2</code> 相等,返回值为 0;若 <code>src1</code> 小于 <code>src2</code> 返回值小于 0;若 <code>src1</code> 大于 <code>src2</code> 返回值大于 0

下面的程序说明了这些函数的用法以及将‘\0’覆盖后产生的程序错误：

```
#include <stdio.h>

void main ()
{
    char x[100] = "Hello";
    char y[100] = "World";
    char z[100];
    strcpy(z, x); /* 这里将 x 中的内容拷贝到 z 中 */
    printf("After strcpy z is %s\n", z);
    strcat(z, y); /* 这里将 y 中的内容拷贝到 z 中"Hello"的后面 */
    printf("After strcat z is %s\n", z);
    printf("Z is %s than x\n", (strcmp(z, x) < 0 ? "bigger" : "small"));
    printf("Before overwrite, strlen(z) is %d\n", strlen(z));
    z[0] = '!'; /* 这里将 z 中的'\0'覆盖了 */
    printf("After overwrite, strlen(z) is %d\n", strlen(z));
    printf("Z is %s", z);
}

输出结果：
After strcpy z is Hello
After strcat z is HelloWorld
Z is bigger than x
Before overwrite, strlen(z) is 10
After overwrite, strlen(z) is 9
Z is HelloWorld! * * *
```

需要注意的是,在上述程序中,最后一行输出中的‘*’表示乱码。在将 `z` 中的‘\0’ (也就是 `z[0]` 的值) 覆盖后, `z` 的长度的增加会随着运行环境的不同而不同。也就是说最后输出的 `strlen` 不是一定的。事实上,由于 `strlen` 函数中对字符串长度的统计方式是取字符串起始位置到第一个‘\0’之间的长度,于是随着环境的不同, `z` 的下一个‘\0’出现的位置是不可预测的,所

以最后的统计结果也是随机的。对于 `rand()` 函数,会在每次调用后输出一些乱码的原因也是一样的。

1.2 二维及多维数组

多维数组可以看作是一维数组的扩展。一个典型的二维数组可以看作是一个以一维数组为元素的特殊一维数组。当要存储和处理矩阵、图表等数据时,用二维数组更方便。二维数组的定义要指定两个下标范围,其语法形式为:

数据类型说明符 变量名 [长度值 1] [长度值 2];

习惯上,第一个下标称为行坐标,它的取值范围是 0~ 长度值 1-1;第二个下标是列坐标,其取值范围是 0~ 长度值 2-1。例如,下面程序声明了一个二维整型数组,并通过一个二层循环语句给数组赋值:

```
#include <stdio.h>

void main()
{
    int m[2][4];
    int i, j;
    for (i = 0; i < 2; i++)
        for (j = 0; j < 4; j++)
            m[i][j] = i * j;
}
```

二维数组元素在内存中是按行顺序存储的,第一个下标代表行,第二个下标代表列,这意味着按照在内存中的实际存储顺序访问数组元素时,右边的下标比左边的下标变化快一些。图 1-1(a) 是数组 `m` 所表示的矩阵,而图 1-1(b) 则是 `m` 的内存分布表示。

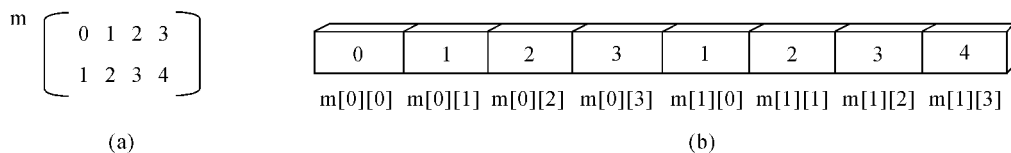


图 1-1 二维数组 `m` 的表示

二维数组的初始化与一维数组相似,为了达到图 1-1(a) 的结果,可以利用下面的语句来初始化上面函数中的 `m`:

```
int m[2][4] = {0, 1, 2, 3, 1, 2, 3, 4};
```

多维数组的情况与二维数组类似,同样可以将其看作是一系列一维数组的嵌套。关于多维数组,需要注意一点:计算机要花大量时间计算数组下标,这意味着存取多维数组中的元素要比存取一维数组的元素花更多的时间。由于这些原因和其他原因,大量的多维数组一般采用 C 语言动态分配函数及指针的方法,每次对数组的一部分动态地分配存储空间。

1.3 指针

1.3.1 指针的基本概念

指针是 C 语言中最为灵活的一个因子,也是最难以掌握的一个概念。简单地讲,指针就

是将内存地址作为值的变量,它与其他变量的不同之处在于,指针的存储空间存放的是地址,而不是一般的数据。

指针的物理意义十分明确,它是 悦语言的精华部分。有了指针技术,我们还可以描述复杂的数据结构,对字符串的处理可以更灵活,对数组的处理更方便,使程序的书写简洁、高效、清爽。通过利用指针,我们能很好地利用内存资源,使其发挥最大的效率。但难点在于它的使用极其灵活,几乎 悦语言的大多语法成分都能应用指针。

指针和前面已介绍的数组有着非常紧密的联系,一个数组一方面可以用其下标进行访问,另一方面也可以用指针来对它进行操作。

和其他变量类型一样,指针必须在使用前进行声明,在 悦语言中指针的声明只要在一般变量声明的变量名之前增加一个星号*,其语法形式为:

摇摇类型说明符 摇 * 变量名;

例如,下面语句声明了两个指针:

```
摇摇char *cp 摇 /* 字符指针变量 cp 可以存放字符变量单元的地址 */
摇摇摇摇int *ip;  /* 整型指针变量 ip 可以存放整型变量单元的地址 */
```

指针的两个基本运算是获取指针所指内存区域的数据和获取变量的地址,对应的运算符是*和取地址符&。下列程序实现了简单的指针运算:

```
摇摇#include <stdio.h>
摇摇void main()
摇摇{
摇摇摇摇char c1='A';
摇摇摇摇char *cp;
摇摇摇摇cp=&c1;
摇摇摇摇char c2;
摇摇摇摇cp=&c2;
摇摇摇摇*cp='B';
摇摇}
```

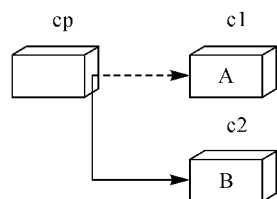


图 员源 指针的表示

执行此函数产生的结果可以用图 员源形象地表示,初始时,变量c1存放的是字符‘A’,指针变量cp存放的是c1单元的地址,用带箭头的线段表示cp所指的存储单元。接着定义另一字符变量c2并将其地址赋给cp,此时cp改变为指向c2存储单元,原指向c1(虚线)实际已不存在,最后一个语句通过运算符*访问指针cp所指的存储单元c2并将其赋值为‘B’。

有了指针,对于变量的访问不仅可以用变量名直接访问,同时也可以利用指针变量提供对变量的间接访问。例如在上面的例子中,最后一行对c2的赋值就是一个间接访问。

对于指针变量可以进行加减操作,但是这里和普通数值变量的加减操作不同,指针变量的加减操作是以所指对象类型大小为单位的,下面的程序说明了这一点:

```
摇摇#include <stdio.h>
摇摇void main()
摇摇{
摇摇摇摇int a=100; 摇摇访问: www.ertongbook.com
```