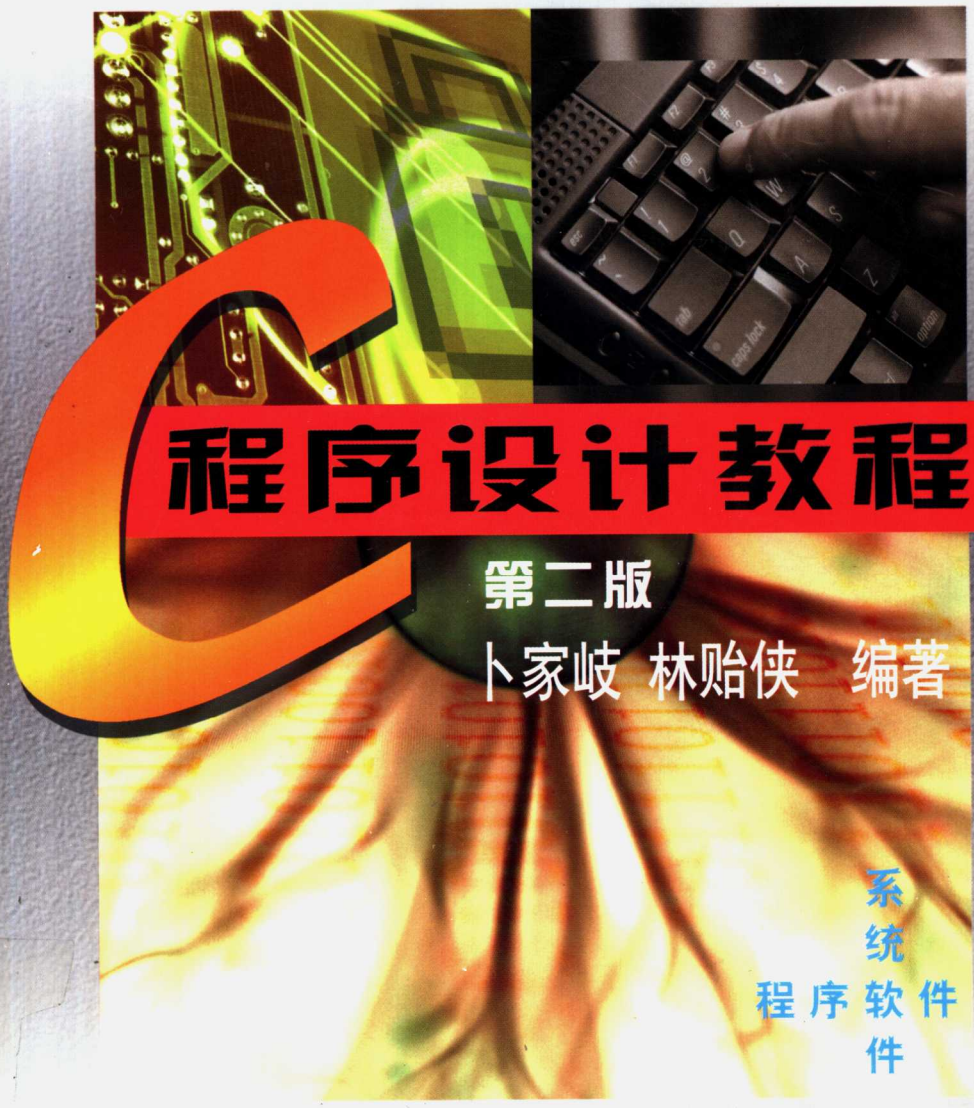
计算机教育丛书

谭浩强　主审



C程序设计教程

第二版

卜家岐 林贻侠　编著

系统软件
程序软件

中国科学技术出版社

计算机教育丛书

谭浩强　主审

C 程序设计教程

第二版

卜家岐　林贻侠　编著

中国科学技术出版社

• 北 京 •

图书在版编目（CIP）数据

C 程序设计教程/卜家岐，林贻侠编著.—2 版. —北京：
中国科学技术出版社，2002.6
(计算机教育丛书)
ISBN 7-5046-3309-7

Ⅰ.C… Ⅱ.①卜… ②林… Ⅲ.C 语言-程序设计-
高等学校-教材 Ⅳ.TP312

中国版本图书馆 CIP 数据核字（2002）第 039433 号

中国科学技术出版社出版
北京市海淀区中关村南大街 16 号 邮政编码：100081
电话：62103202
新华书店北京发行所发行 各地新华书店经售
北京市卫顺印刷厂印刷
*
开本：787 毫米×1 092 毫米 1/16 印张：17.25 字数：450 千字
2002 年 6 月第 2 版 2002 年 6 月第 1 次印刷
印数：1—8000 册 定价：25.00 元

序

重视和加强计算机教学是上海大学钱伟长校长一贯教导的。几年来,在这个领域中我校的教学工作不断加强,重视改革,成绩是显著的,尤其是在占学生人数90%以上的非计算机专业中,从教学内容和方法上,年年有所变革,有所提高。该教材的第一版1998年获上海市优秀教材奖。

计算机科学的发展日新月异,在教学内容上要跟得上时代的发展,从教学方法上必须不断创新。此书作为教学改革的尝试,特别是对C语言中“指针”的教学作了较大幅度的改革,作者的用意是使这个C语言中的难点分散到每一章中进行教学,便于学生掌握,这比传统教材中把指针集中在一章中进行教学更合理、更有效。在第二版的教材中,除保留第一版的主要特点外,在章节组织上亦作了较大幅度的调整,在内容上作了进一步的充实。同时为了适应程序设计的最新潮流,还介绍了C++语言概述。全国高校计算机基础教学研究会理事长谭浩强教授审阅了全书,并提出了不少宝贵意见,这是对我校计算机教学改革的关心和支持,在此表示由衷的感谢。

世界计算机教育会议(WCCE)认为,在信息社会里,计算机基础知识和技能的教育与读、写、算等的传统文化教育同样重要,阅读和编制计算机程序的能力被称为人类的“第二文化”。我们在高等教育中如何加强计算机的基础教学,并把它渗透到各个学科中去,这对提高人才素质和学科水平是极其重要的。本教材的再版希望能起积极的促进作用。

上海大学党委书记兼常务副校长

2002年3月28日

前　言

学习和应用 C 语言的人很多，这是因为 C 语言与其他高级语言相比有着它的独特优点。同时学好 C 语言也为进一步学习 C++语言打下一个坚实的基础。

C 语言是一种理想的结构化语言，它具有高级语言的优点，又具有低级语言的特点，因此它既能用来编写高效简洁、风格优美的应用程序，又能用来编写计算机系统软件。C 语言表达简洁、功能丰富、使用灵活、应用面广、移植性好，能实现汇编语言的某些功能,目标代码质量高，程序运行的速度快。

许多计算机专业和非计算机专业的学生很希望学习 C 语言,有的学校已经把 C 语言作为大学本科“第一计算机语言”进行教学，即上课时间安排在大学一年级或二年级，同时在计算机的后续课程，如“数据结构”、“软件技术基础”中，都以 C 语言作为背景语言进行教学，这样可以巩固、深化 C 程序设计，使得进入高年级以后具有参加实际软件开发的能力。本书有以下几个特点:

1．以大学本科“第一计算机语言”的角度来编写全书，重新安排 C 程序设计内容,力求入门容易，通俗易懂。本书重概念,重实践,通过大量实例（包含 110 多个精心设计的实例）较全面地介绍 C 程序设计的基本概念和方法。

2．循序渐进、难点分散。大家知道，指针是 C 语言的精华,指针的概念比较复杂,使用较为灵活，容易搞错，所以指针是学习 C 语言的难点,我们把这个难点分散到每一章中去讲解。首先在第一章计算机的一般知识中，加入内存单元的地址即指针的内容,使初学者提前认识指针概念。而后在第二章简单数据类型中介绍变量时，加入“指针变量”的内容，通过简单的举例，对指针变量有一初步认识。由于提前介绍指针概念,为后续各章引入指针应用创造条件，因此在本教程的每一章中，都会看到指针应用的实例。采取“少吃多餐”的办法,便于读者“消化吸收”，因此本教材中没有为指针特设一章。

3. 本书内容遵循 ANSI C (美国国家标准 C 语言)为基础，按新标准定义和

声明函数；按新标准对数组和结构体变量进行初始化；并建议用缓冲文件系统而不用非缓冲文件系统等。另外，本书用应用广泛的传统流程图和 N-S 结构化流程图描述算法，程序写成锯齿形格式，这些都会养成读者良好的程序设计习惯，编写出可读性好,质量高的应用程序。

4. 本书每章最后一节是“本章的 C++知识”，介绍与本章 C 内容相关的 C++知识。以便读者对照学习，并在第 8 章较详细介绍 C++语言的基本概念，这是 C 语言的最新发展,是程序设计的一个新潮流。读者有了以上基础知识后，会较容易的进入 C++编程环境，去完成中大型软件的设计。这部分内容在安排教学计划时，按任课教师的具体情况可有可无，可多可少，不是必须的。但对于那些有意学习 C++面向对象程序设计的读者，这部分内容就显得十分重要。

5. 本书中的程序使用 Turbo C 2.0 版本（C 部分）和 Visual C++6.0(C++部分)为背景，全部例题都在 PC 机上调试通过，每个例题一般都有运行结果，以便读者学习验证。第 9 章还专门介绍了 Turbo C 2.0 和 Visual C++6.0 两个开发环境，读者可以根据实际情况选用。

本书的再版内容全部由卜家岐教授和林贻侠教授重新编写。并请多年从事本课程教学的教授们通读全稿，提出了很多宝贵意见,最后请全国高等学校计算机基础教学研究会理事长谭浩强教授审阅全书，并得到了他的热忱指导和帮助，为此对他们表示衷心感谢。

由于作者的水平有限，加上计算机科学技术日新月异、发展迅速，本书会有不少缺点或错误，敬请专家和读者指正。

编　者

2002 年 3 月 28 日于上海

目　录

第 1 章　计算机、程序设计和 C 语言概述

为了全面认识和理解 C 程序设计思想，了解一下计算机的工作原理，数据在内存中的存储形式，以及程序设计的发展趋势是很有必要的。学习本章内容后，读者会对计算机、程序设计和 C 语言有个整体上的概念。

1.1　计算机概述

回顾一下计算机技术蓬勃发展的简史，有助于我们从第一计算机语言的角度掌握本课程的内容，提高计算机文化的素质。20 世纪 40 年代，许多新的科技领域，如核反应堆的控制，导弹飞行轨迹的控制等，都要求在较短的时间内完成较复杂的计算工作，靠人工计算是不可能的，必须想办法依靠某种形式的计算机。美籍数学家冯·诺依曼（Von. Neuman）教授指出，如果用二进制而不用十进制进行数值运算，就可以利用电来进行运算，利用电路的开和关两种状态来表示二进制中的 0 和 1，用重叠组合的开关电路就可制成一种计算工具，于是在 1946 年诞生了世界上第一台电子计算机 ENIAC（Electronic Numerical Integrator And Computer）。它用二进制代替十进制完成了复杂的数值运算。但是运算指令和数据还是通过重新连接电路和设定开关来完成的。通过实践，冯·诺依曼又提出存储程序（Stored Program）的概念，即将运算指令和数据用数字的方式存放在电子计算机的存储器中，也就是首先提出了用软件来控制计算机的操作。这就形成了著名的冯·诺依曼原理——“二进制和程序存储控制”的计算机结构思想，为以后计算机的发展奠定了理论基础。

自 ENIAC 以来，计算机的研制、生产和应用以迅猛的速度发展着，到 20 世纪末计算机已经历了电子管（第一代），晶体管（第二代），集成电路（第三代）和大规模集成电路（第四代）四个阶段。目前应用的计算机，它里面所用的集成电路芯片的集成度越来越高，有的被称为超大规模集成电路。这四个时期的计算机都基于冯·诺依曼原理，所以这四代计算机也称为冯·诺依曼计算机。

现在，计算机已是一种现代化的处理信息的工具。数字信息、文字信息、图像信息、动画信息和声音信息都可以通过计算机来进行存储处理。20 世纪 90 年代以来，许多国家已致力于集文（字）、图（形）、声（音）于一体的多媒体计算机（MPC——Multimedia Personal Computer）的研究，一台多媒体计算机可以具有电视机、录像机、计算机、电话机和录音机等的功能，计算机的应用开始进入每个家庭。加上计算机网络的发展，从小型的局域网到全球的 Internet 网，从传递信息的角度上讲，地球已变得越来越“小”了。与此同时，科学家们还在研制第五代计算机（非冯·诺依曼计算机），它采用完全新的工作原理和体系结构，它的工作接近于人们思考问题的方法，是一种能与人交谈，自我学习，自我推理并有决策能力的智慧型计算机。这是计算机发展的方向，但近期还难以实现。本课程中，计算机指的是目前流行的大规模集成电路计算机。

计算机系统由硬件系统和软件系统两部分组成。

计算机硬件系统指的是计算机的具体设备，即实物。它由主机和外部设备两部分组成。而

计算机主机又由中央处理单元 CPU（Central Processing Unit）和内存储器（简称“内存”）组成。CPU 包含运算器和控制器两部分，外部设备则由输入、输出设备和外存储器组成。

计算机的软件系统指的是程序。为了使计算机正常工作，程序是必不可少的。计算机做任何工作，都是通过存储在内存中的程序来实现的，可以说程序是计算机工作的“灵魂”。根据程序的不同用途，软件系统又可分为系统软件和应用软件两部分。系统软件一般是指生产厂家或公司在出售计算机时提供给用户的，它用来管理计算机的各项工作和为开发运行用户程序配备必要的软件包。如操作系统 OS（Operating System），编译程序（Compiler），解释程序（Interpreter）以及一些集成（多功能）开发环境（软件）等。应用软件包括用户向厂家或公司定制的专用软件包和用户自己建立的各类源程序（高级语言编写的程序）或执行程序等。操作系统用来管理计算机内的所有资源（包括软、硬件资源）。如何充分引用计算机的全部资源和最大限度的发挥计算机各部分的作用是设计操作系统软件的指导思想。在操作系统的支持下，用户才能编辑源程序，并利用编译程序或解释程序翻译源程序成为机器能执行的二进制指令代码。集成开发环境是一个集编辑、编译、运行于一体的系统软件，它为用户编制应用软件提供很大的方便。

1.2 数据在计算机中的存储形式

1.2.1 位、字节、字的概念

在主机内的存储器，称内存储器，简称内存。要运行的程序和数据都存放在内存中。那么内存是由什么组成的呢？它原来是由千千万万个具有二个状态的电子开关组成的。电子开关打开的状态为“1”，闭合时的状态为“0”。每个电子开关用计算机术语“位（bit）”来称呼，它可以代表二进制数（逢二进一）的一个基本单元。计算机内存就是一个庞大的二进制基本单元——电子开关的集合体。

要存放信息，必须把这些电子开关有机地组织起来，一般用若干个二进制“位”组成一个“字节”（byte），多数计算机用 8 位组成一个字节，如图 1.1 所示。

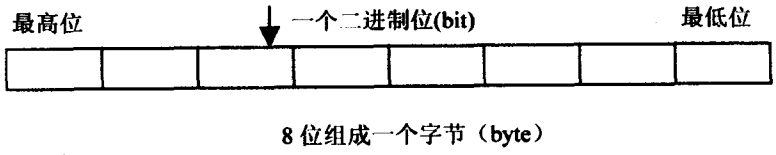


图 1.1 字节示意图

一个字节可以存放一个字符，当然这个字符是用一个二进制数来表示的，这个二进制数就是这个字符的二进制代码。计算机上所应用的全部字符（字母、数字或其他专用字符）都有相应的二进制代码，如字母“a”，它的二进制代码是：

0	1	1	0	0	0	0	1

来表示。这些字符和二进制代码之间的对应关系，很多计算机系统采用 ASC II (American Standard Code for Information Interchange 即“美国标准信息交换码”)代码。这个代码与字符之间的对应关系见附录 I 。

计算机内存就是由很多排列整齐的字节组成，为了管理方便，每个字节都有相应的位置编号，这个编号就是这个字节的“地址”，通过地址可以找到内存中任何一个字节的内容。

一个字节可以存放一个字符，但要存放一个整数或一个实数，一个字节就不够了，有的系统上整数用 2 个字节或 4 个字节来存放，而实数用 4 个字节来存放。由一个或若干个字节组成一个“字”，一个字可以用来存放一个数据或一条指令。

1.2.2 内存单元、内存单元地址和指针的概念

一个内存单元可以用来存放一个数据或一条指令，由于内存中每个字节都有自己的地址，因此每个内存单元也有自己相应的地址，由一个字节组成的内存单元的地址，就是这个字节的地址，而由多个字节组成的内存单元的地址，这个地址指的是组成这个内存单元几个字节中的第一个字节的地址。假定程序中已定义了三个短整型变量 i、j、z，编译时系统分配给它们各一个内存单元，每个内存单元占两个字节，如图 1.2 所示，图中每个长方形框代表两个字节。如果分配给变量 i 的内存单元地址是 5000（这只是假定的地址），那么变量 i 占用地址为 5000 开始的两个字节（即地址为 5000 和 5001 两个字节），5000 是变量 i 的内存单元地址，简称变量 i 的地址，C 语言中用符号“&i”来表示，这里&表示取地址运算符，&i 的值是 5000。若定义变量时 j 是紧跟在变量 i 之后，变量 z 又紧跟在 j 之后，那么系统分配给这些变量的内存单元一般是连续的，即分配给变量 j 的内存单元地址是 5002，分配给 z 的内存单元地址是 5004，所以变量 j 的地址&j 的值是 5002（变量 j 占用地址为 5002 和 5003 两个字节），而变量 z 的地址&z 的值为 5004（变量 z 占用地址为 5004 和 5005 两个字节），所以，当程序定义了某一变量以后，变量名和它占用的内存单元地址有直接的对应关系。

请注意，要区别一个内存单元的地址和这个内存单元的内容（即变量值）这两个概念。对变量值的存取是通过地址来进行的，根据变量名和地址的对应关系，找到变量的地址，然后从这个地址所标识的存储单元中进行存取数据的操作。图 1.2 中，设变量 i，j 的值分别是 5 和 3，要执行一个操作 z = i + j 时，过程是这样的：从变量 i 这个存储单元的地址 5000 开始的两个字节中取出变量 i 的值 5，再从变量 j 这个存储单元的地址 5002 开始的两个字节中取出变量 j 的值 3，通过 CPU 将它们相加后得和 8 送到变量 z 这个内存单元中，即送到地址为 5004 开始的两个字节中。这种按变量地址存取变量的方式称为“直接访问”方式，这里的“访问”指的是在内存单元中存取（或称读写）数据的意思。

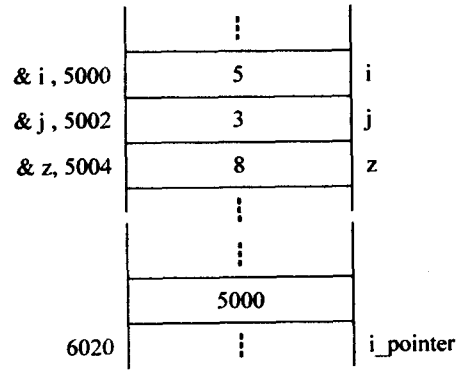


图 1.2 内存单元分配示意图

总之，通过变量的地址能找到变量的内存单元，因此，把变量的地址称作变量的“指针”，如地址 5000 是变量 i 的指针，地址 5002 是变量 j 的指针。通过变量的指针可以找到变量的内存单元来对变量进行读写操作。指针指的是某个内存单元的首地址（或说一个内存单元的入口地址），因此 C 语言中应用的指针是有数据类型的（详见第二章），而每一种数据类型的变量在内存中存放的数据是以内存单元为单位的。

与“直接访问”相对应的另一种访问内存的方式是“间接访问”方式。C 语言中可以定义一种变量专门用来存放地址，假定我们定义一个变量 i_pointer 是用来存放整型变量地址的，系统编译时分配给这个变量的地址假定是 6020，那么可以通过下面赋值语句将变量 i 的地址赋给变量 i_pointer：

```
i_pointer = & i ;
```

变量 i 的地址值是 5000，所以通过以上赋值语句就把地址值 5000 放入地址是 6020 开始的两个字节中，i_pointer 的值就是 5000。这样，我们就可以对变量 i 进行间接访问了，过程是：先从变量 i_pointer 中提取变量 i 的指针 5000，然后到地址是 5000 和 5001 两个字节中存取 i 的值。这个过程，C 语言中用一个存取内容运算符*，即*i_pointer，它可间接存取 i 的值。如

```
*i_pointer=5;          /*间接访问*/
```

等效于

```
i=5;                   /*直接访问*/
```

因为这两个语句的结果都是给变量 i 赋以 5，但它们执行的过程不同，一个是通过间接访问，而另外一个是直接访问。

图 1.3 表示直接访问和间接访问的示意图。

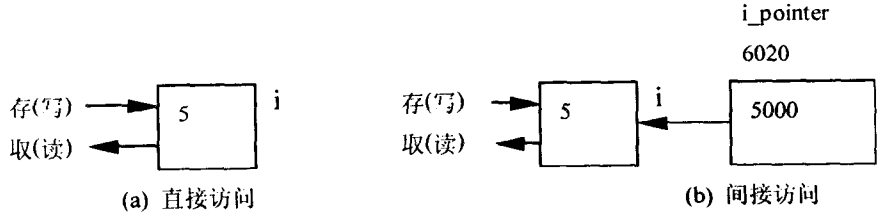


图 1.3　直接访问和间接访问

图 1.3（a）表示对变量 i 的值直接进行读写操作（直接访问）；图 1.3（b）表示对 i_pointer 所"指向"的内存单元进行读写操作（间接访问）。因为 i_pointer 的值是变量 i 的地址，所以称 i_pointer 指向变量 i ，借助于 i_pointer 可以对变量 i 进行读写操作。当 i_pointer 的值用变量 j 的地址&j 重新赋值，那么借助于 i_pointer 可以对变量 j 进行读写操作。因为 i_pointer 的值可以改变，所以称它为"指针变量"。关于指针变量的定义详见第二章，这里只作概述，目的是说明借助于它访问内存的方式（间接访问的方式）。

1.2.3 十、八、十六进制和二进制的相互转换

计算机内部数的存储和传送是用二进制数，而我们习惯上使用的是十进制数。为了表达方便，在一些低级语言中常用八进制或十六进制的数，懂得这些数之间的相互转换是必要的。

大家熟悉的十进制数有两个特点：

(1) 用 0, 1, ..., 9 十个数字来表示；

(2) 逢十进一。

同样原理，二进制数的特点是：

(1) 用 0, 1 二个数字来表示；

(2) 逢二进一。

它是最简单的进位制，也是机器目前所能直接接受的。用二进制组成的代码称机器码。

八进制数的特点是：

(1) 用 0, 1, ..., 7 八个数字来表示；

(2) 逢八进一。

十六进制数的特点是：

(1) 用 0, 1, ..., 9 和 A, B, C, D, E, F 十六个数字或字母来表示 0～15；

(2) 逢十六进一。

由上分析，我们可以把二、十、八、十六进制数之间的对应关系列于表 1.1 所示。

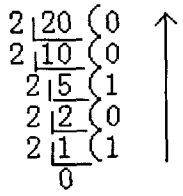
表 1.1　二、十、八、十六进制数之间的对应关系

十进制数	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
十六进制数	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
二进制数	0	1	10	11	100	101	110	111	1000	1001	1010	1011	1100	1101	1110	1111
八进制数	0	1	2	3	4	5	6	7	10	11	12	13	14	15	16	17

有了这些基础，我们就可以来实现它们之间的相互转换了。

1. 十进制和二进制的相互转换

十进制数转换成二进制数可用除 2 取余的长除法。如 20 的二进制求法是：



得 $(20)_{10}=(10100)_2$

二进制数转换成十进制数，可以借助于数的幂级数形式来完成。如有一个十进制数 211，写成以 10 为基数的幂级数形式是：

$$(211)_{10} = 2\times10^2 + 1\times10^1 + 1\times10^0$$

十进制数 211 每位的权是 $10^2, 10^1, 10^0$

同样，二进制数也有它的幂级数形式，如有一个二进制数 1011，可以写成以 2 为基数形式是：

$$(1011)_2 = 1\times2^3 + 0\times2^2 + 1\times2^1 + 1\times2^0$$

二进制数 1011 每位的权是 $2^3, 2^2, 2^1, 2^0$

从二进制的幂级数形式，很容易得到它的十进制数值：

$$(1011)_2 = 8+2+1 = (11)_{10}$$

2. 八进制、十六进制和二进制的相互转换

已知八进制数用：　　0 ——→ 7　八个数字组成

相应的二进制数是：　　000 ——→ 111

即八进制的一位可用二进制的三位来代替。

而十六进制数用：　　0 ——→ F　十六个数字和字母组成

相应的二进制数是：　　0000 ——→ 1111

即十六进制的一位可用二进制的四位来代替。

那么二进制转换成八进制时，只要从二进制数的最右边开始向左数，每三位一组，一直到最左边位数不足时，用 0 代替，每三位二进制数由表 1.1 就可以写出一位八进制数。相反，八进制数要转换成二进制数时，只要把每位八进制数用三位二进制数代替就完成转换工作。

要把二进制数转换成十六进制时，只要把二进制数的最右边开始向左数，每四位一组，一直到最左边位数不足时，用 0 代替，每四位二进制数就可写出一个十六进制的数字或字母，合在一起就是一个十六进制数。相反，要把一个十六进制数转换成二进制数，则只要把每位十六进制数用四位二进制数代替就完成转换工作。例如：

1　0　5　1　2 ——→ 八进制数 10512

001　000　101　001　010 ——→ 二进制数 1000101001010

0001　0001　0100　1010 ——→ 二进制数 1000101001010

1　1　4　A ——→ 十六进制数 114A

那么，十进制和八进制、十六进制之间怎样进行转换呢？从上述可知，通过二进制这个媒介可以把它们连系起来。如十进制要转化成八进制，我们可以先把十进制化成二进制，再把二进制转化成八进制的方法来进行。反之，也同样。当然还有别的方法，在此不一一列举了。

1.2.4 原码、反码和补码的概念

内存中的每一个数都占用一个内存单元，而这个内存单元中的最高位用来表征数的符号，以 0 代表正数，以 1 代表负数，其余各位代表数的绝对值。

一个数的代码有原码、反码和补码三种表达方式。正数的原码、反码和补码的形式都相同，没什么差别，如一个短整数 8，在内存中占两个字节，则+8 的原码、反码和补码都具有如下的