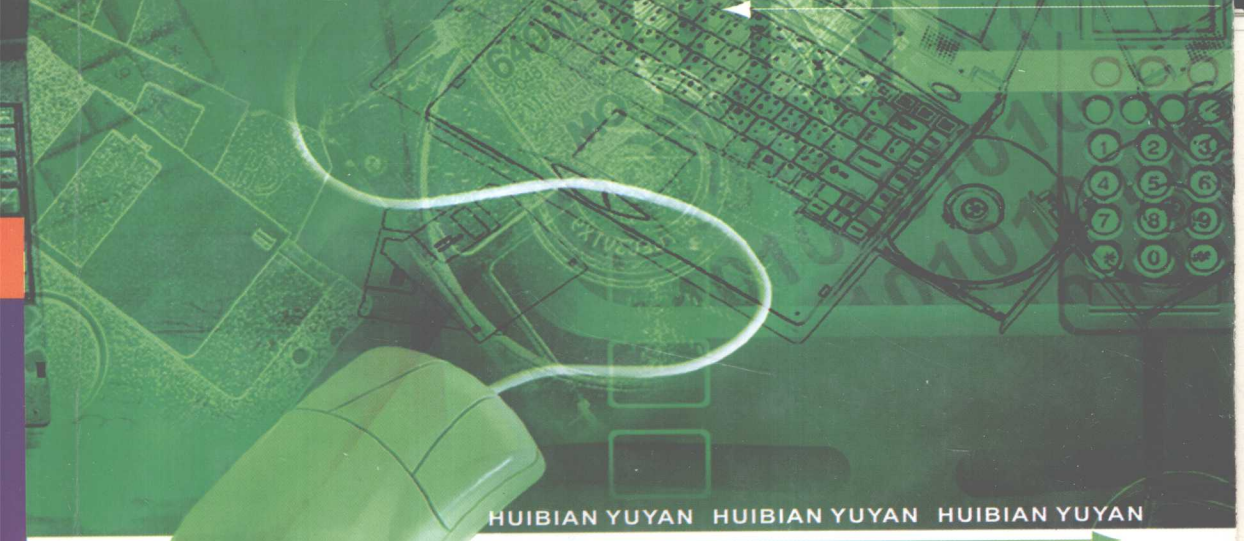


21世纪高等学校计算机科学与技术规划教材

汇编语言

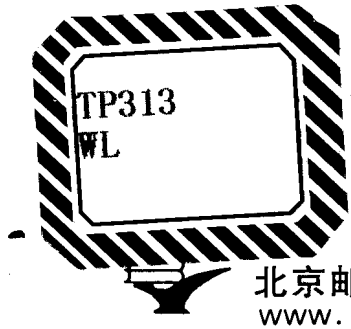
汪 黎 吴庆波 编著

北京邮电大学出版社

21世纪高等学校计算机科学与技术规划教材

汇编语言

汪 黎 吴庆波 编著



北京邮电大学出版社
www.buptpress.com

内 容 简 介

本书根据计算机专业学生学习、应用汇编语言的特点和需要，以32位微处理器80386为主要背景，深入系统地介绍了32位汇编语言程序设计的一般概念、基本技术和常用算法。本书共8章，分为两个层次：基础部分（第1～第5章）主要介绍了汇编语言程序设计的一般概念、基础知识，80386 CPU的寻址方式和指令系统，汇编语言的基本语法，汇编语言程序的基本结构和设计方法；提高部分（第6～第8章）主要介绍了DOS和BIOS系统功能调用、中断程序设计，宏汇编、重复汇编、条件汇编、保护方式下的汇编程序设计，以及如何用汇编语言编写WINDOWS程序。本书内容深入浅出，新颖丰富，实用性强，各章之后均针对性地附有适量的习题。全书反映了汇编语言程序设计的主流技术和发展趋势，体现了基础性、实用性和先进性的统一。

本书可作为高等院校计算机或相关专业的汇编语言课程教材，也可供科研和软件开发人员学习或自学参考。

图书在版编目(CIP)数据

汇编语言/汪黎，吴庆波编著. —北京：北京邮电大学出版社，2005

ISBN 7－5635－0839－2

Ⅰ.汇... Ⅱ.①汪...②吴... Ⅲ.汇编语言—程序设计 Ⅳ.TP313

中国版本图书馆CIP数据核字(2005)第027034号

书　　名	汇编语言
编　　著	汪　黎　吴庆波
责任编辑	陈露晓
出版发行	北京邮电大学出版社
社　　址	北京市海淀区西土城路10号(100876)
电话传真	010－62282185(发行部)　010－62283578(传真)
	E-mail　sanwen99@mail.edu.cn
经　　销	各地新华书店
印　　刷	国防科技大学印刷厂印刷
开　　本	787mm×960mm　1/16
印　　张	20
字　　数	351千字
版　　次	2005年4月第1版　2005年4月第1次印刷

ISBN 7－5635－0839－2/TP·111　　定价　29.50元

如有质量问题请与发行部联系

版权所有　侵权必究

序

自20世纪80年代以来，高等学校计算机教育发展迅速，计算机教育的内容不断扩展、程度不断加深。特别是近十年来，计算机向高度集成化、网络化和多媒体化发展的速度一日千里；社会信息化不断向纵深发展，各行各业的信息化进程不断加速；计算机应用技术与其他专业的教学、科研工作的结合更加紧密；各学科与以计算机技术为核心的信息技术的融合，促进了计算机学科的发展，各专业对学生的计算机应用能力也有更高和更加具体的要求。

基于近年来计算机学科的发展，以及国家教育部关于计算机基础教学改革的指导思路，我们确立了这套“21世纪高等学校计算机科学与技术规划教材”的编写思想与编写计划。教材是教学过程中的“一剧之本”，是高校计算机教学的首要问题。该套系列教材编写计划的制定凝聚了编委会和作者的心血，是大家多年来计算机学科教学和研究成果的体现，并得到了陈火旺院士的亲自指导与充分肯定。

这套系列教材由北京邮电大学出版社三文工作室精心策划和组织。编写过程中，充分考虑了计算机学科的发展与《计算机学科教学计划》中内容和模块的调整，使得整套教材更具科学性和实用性。整套系列教材体系结构按课程设置进行划分。每册教材均涵盖了相应课程教学大纲所要求的内容，既具备学科设置的合理性，又符合计算机学科发展的需要。从结构上遵循教学认知规律，基本上能够满足不同层次院校、不同教学计划的要求。

各册教材的作者均为多年来从事教学、研究的专家和学者，他们有丰富的教学实践经验，所编写的教材结构严谨、内容充实、层次清晰、概念准确、论理充分、理论联系实际、深入浅出、通俗易懂。

教材建设是一项长期艰巨的系统工程，尤其是计算机科学技术发展迅速、内容更新快，为使教材更新能跟上科学技术的发展，我们将密切关注计算机科学技术的发展新动向，以使我们的教材编写在内容上不断推陈出新、体系上不断发展完善，以适应高校计算机教学的需要。

21世纪高等学校计算机科学与技术规划教材编委会

2005年3月

前　言

汇编语言是功能强大的语言，它是直接面向机器的。汇编语言能最大限度地利用机器硬件提供的支持，直接控制硬件，充分发挥硬件的性能。早期的操作系统都是用汇编语言开发的，而且只能用汇编语言开发。随着高级语言和开发工具的发展，虽然现代操作系统可以用C等高级语言来开发，但其核心部分，例如操作硬件，保护方式切换等仍只能用汇编语言开发。使用汇编语言还有一大优势，那就是汇编语言的运行效率是最高的，因而在特别要求程序的运行时间和空间的场合，汇编语言是首选。另外，汇编语言在单片机，可编程控制器等方面应用也很广泛。

“汇编语言程序设计”是计算机专业必修的一门重要的核心课程。读者只有在较为深刻地理解了汇编语言的原理和运行环境，能够熟练地应用汇编语言后，才能更好地学习操作系统、微机原理等课程。另外，学习汇编语言程序设计也可为读者打下程序设计基本技术和基本方法的基础，对以后学习别的程序设计语言大有裨益。

目前国内使用最广泛的PC系列机，都以Intel的80x86，Pentium，Celeron系列微处理器或者与其兼容的微处理器作为CPU。现在主流的微处理器都是32位的，但16位微处理器8086/8088是基础，因此在本书中，我们在讲述16位汇编语言的基础上，以经典的32位微处理器80386为重点，主要讲述32位的汇编语言设计的一般概念、基本技术和常用算法。

本书共8章。第1章讲述汇编语言的基本特点，计算机中的数据表示(二进制，十进制，十六进制表示之间的转换和运算，计算机中数和字符的表示)以及微机的结构组成。第2章介绍寻址方式和指令系统，在介绍8086/8088的基础上，重点讲述80386。第3章介绍汇编语言的基本语法，程序的基本结构和运行的过程。第4、第5章介绍程序设计的基本方法，其中第4章讲解顺序、分支、循环程序，第5章介绍子程序结构和模块化程序设计技术。第6章介绍DOS和BIOS系统功能调用、输入输出、中断程序设计。第7章介绍汇编中较为高级的

试读结束：需要全本请在线购买：　www.ertongbook.com

技术，包括宏汇编、重复汇编、条件汇编等。第 8 章面向较高水平的程序员和有兴趣的读者，讲解保护方式下的编程以及用汇编语言编写 WINDOWS 程序。书中给出了大量的程序实例，每章的后面都有习题，便于读者复习巩固以及检测学习效果。

本书内容深入浅出、新颖丰富、实用性强，前五章是基础部分，可作为一般院校计算机专业的教材，有条件的院校还可讲授第 6、第 7 章。第 8 章可供有兴趣，水平较高的读者参考。

本书由国防科技大学汪黎、吴庆波编著，由于作者水平有限，加之时间仓促，错误和不妥之处在所难免，敬请广大读者和专家批评指正。

编者

2005 年 2 月于国防科技大学

目　录

试读结束：需要全本请在线购买： www.ertongbook.com

第1章 基础知识

1.1 汇编语言概述

不知道读者有没有编程的经验，可能大部分读者都会有这样的问题：汇编语言是什么，跟现在市面流行的 C、Delphi、C++、C++ Builder、Visual C++等编程语言及编程环境有什么区别，为什么要使用汇编语言来编程？本节就从汇编语言的概念开始讲述。

1.1.1 汇编语言的概念

汇编语言，毋庸置疑，也是一种程序设计语言，在这一点上，它和 BASIC、PASCAL、C、C++是一样的。那么，不同的程序设计语言有什么区别，它们主要是语言的语法不一样，另外，具有的功能一般也有或大或小的差别。一般地，我们可以把程序设计语言划分为三大类：机器语言，汇编语言，高级语言。前面所提到的 BASIC、PASCAL、C 等都属于高级语言。读者可能很惊讶，汇编语言居然“独树一帜”。不错，汇编语言确实可以“独当一面”。

下面来讲述 3 类语言的区别。

1. 机器语言

机器语言的概念可能读者听得较少。所谓机器语言就是用机器指令作为程序设计的语言。那么什么是机器指令呢？机器指令就是 CPU 能直接识别和运行的一段二进制编码序列。执行一条机器指令，CPU 可以完成一定的功能。一种 CPU 所能识别和执行的所有机器指令的集合，称为该 CPU 的指令集。指令集是与 CPU 的类型相关的，不同的 CPU，尤其是不同厂家的 CPU，或是不同体系结构的 CPU，它们的指令集可能有很大的差异。这种差异体现在指令的编码、功能等方面。可以类比人类的语言，同是指人，英语和汉语的单词写法显然不同，这就是编码方式上的差异。功能上的差别也很明显，一个英语单词可能同

时兼具汉语中多个词的意思，同样，汉语中一个词可能也有英语中多个词的意思。那么，既然机器语言也是一种程序设计语言，那么市面上为什么没有“机器语言程序设计”之类的书呢？那是因为程序员很难掌握和使用机器语言。机器语言是由 CPU 直接执行的，是用二进制表示的，不符合人类书写和使用的习惯，而且没有明显的语义提示。用机器语言编程，非常繁琐困难。机器语言写出的程序，很难维护、调试，更不用说理解、记忆和交流了。为了给读者一个直观的印象，我们来举个例子：假设要将内存中的两个 8 位整数求和，这两个 8 位整数均位于数据段中，它们的内存单元地址偏移分别为 2001H，2002H（这里内存单元和段的概念读者不必深究，本书后面会有详细的讲述），结果赋给另一个 8 位整数，也位于数据段中，偏移为 2003H。完成这个操作需要三条机器指令，换句话说，用机器语言写出的程序如下（用 16 进制表示）：

```
A0 00002001
02 05 00002002
A2 00002003
```

这个程序看来有点“丈二和尚摸不着头脑”，它还是用 16 进制表示的，如果直接用二进制表示，那就是一大串的 01 串。对于以前的计算机，开发人员还真只能用机器语言编程，然后通过在纸带上打孔，把程序“记录”在纸带上，通过纸带读入机把程序输入计算机运行。这样的编程方式，现在恐怕也只有“黑客”才会使用了。

2. 汇编语言

现在来看汇编语言，汇编语言的开发正是为了避免机器语言的上述缺点。人们采用了一些带有明显的语义的单词符号来表示指令的操作码。这些符号被称为指令助记符。指令助记符一般都明显地表示了指令的功能，因而便于理解和应用。另外，对操作数部分，也采用了一些符号来表示，如 CPU 的内部寄存器、存储单元地址等。除此之外，汇编语言中还包含伪指令、宏等不少语言成分。现在，同样是前面的例子，用汇编语言来写，就是：

```
mov al,ds:[2001H]
add al,ds:[2002H]
mov ds:[2003H],al
```

显然看起来要清楚一点。有过高级语言编程经验的读者可能会说，这看起来和一般的高级语言程序还是大不一样。为什么不用高级语言来写呢，多省事呀，只要声明三个 8 位整形变量 a，b，c，然后一条语句：c=a+b 不就可以了吗？对于这个疑问，我们将在后述中解释。这个汇编程序，实际上是与机器语言一一对应的。mov、add 都是指令助记符，al 表示 CPU 的一个内部寄存器，ds:

[2001H]等则表示内存单元的地址。

那么,汇编语言编写的程序如何输入到计算机由计算机运行呢?前面讲过,计算机(确切地说是CPU)能识别的只有机器指令,所以,这里就需要有一个“翻译”,这个“翻译”就是汇编语言编译器,它将汇编语言编译成机器语言,就可以运行了。

3. 高级语言

高级语言就是对汇编语言的进一步抽象,它基本上屏蔽掉了汇编语言中与CPU相关的细节,例如CPU的内部寄存器。也屏蔽掉了具体的内存单元地址的概念,由高级语言的编译器自己来负责内存的分配以及其他的很多问题。这样,程序员就不必关心底层硬件的细节,专注于程序本身的算法、功能设计。前面的例子,用高级语言来编写,可能就是:

```
unsigned char a,b,c;
c=a+b;
```

看到这里,读者可千万不要打退堂鼓,觉得自己为什么不去学高级语言,学汇编有什么好处呢?前面说过,汇编语言是可以“独当一面”的。下面就来讲述汇编语言的优点。

1.1.2 汇编语言的优点

1. 汇编语言的功能强

汇编语言是通过对机器指令逐一加上指令助记符,并对操作数也进行符号化而形成的。也就是说,机器指令与汇编语言指令是一一对应的。换句话说,机器能执行的所有指令都可以通过汇编语言来描述。因此,通过汇编语言,可以非常方便有效地控制机器。以前的操作系统,甚至应用软件都是用汇编语言编写的。而高级语言的设计目的是尽量对程序员屏蔽机器硬件的细节,抽象幅度比较大。这样很多机器指令就可能无法通过高级语言来描述。当需要执行一些特殊的操作时,一般的高级语言就无能为力了。例如,要得到当前进程的堆栈指针的值,用高级语言是没有办法的,只能依靠汇编语言。当编写操作系统或其他系统软件时,汇编语言更是不可或缺的工具。可能有的读者会问,现在不是说可以用高级语言来写操作系统吗?Linux不就是用C语言写的吗?这是不确切的,Linux中有很多与机器硬件打交道的代码或实现核心功能的代码都是用汇编语言写的,这些代码实现的大部分功能用C语言是无法实现的。

2. 汇编语言的效率高

汇编语言程序的执行效率比所有高级语言都高。这种高效,体现在时间和空间两个方面。汇编程序的执行速度快,在实现同样功能的情况下,汇编语言程

序生成的二进制可执行程序短。这种优势比较明显，尤其与一些高级语言相比。

这得益于汇编语言与机器语言几乎是一一对应的，而高级语言可能一个语句就对应多条机器指令，而其中某些机器指令可能在当前场合下是不需要的。另外，汇编语言可以充分利用机器硬件的特性。国外很多的编程高手，都非常乐于用汇编语言编程，这很大程度上是从程序的执行性能上考虑的。

除上述优点外，通过学习汇编语言编程，读者将对计算机体系结构，运行原理等有更深刻的认识，这从高级语言是无法学到的。另外，汇编语言编程与高级语言编程有很多地方是相通的，掌握汇编语言程序设计，对学习高级语言编程大有裨益。

1.2 计算机中的数据表示

1.2.1 数制与代码

计算机最基本的功能是对数进行加工处理，进位计数制则是数的表示方法。人们最常用的是十进制计数法。但是，由于数在计算机中是用一个器件的物理状态来表示的，如果计算机中仍用十进制计数制，则为了表示一位数，就要求相应的器件具有 10 种状态。这样的机器制造起来将会非常复杂，而且其稳定性也难以保证，因为这要求 10 种状态之间的差别要非常明显。因此，很自然，计算机采用二进制表示数最为方便。因为一个具有两种不同的稳定状态而且能互相转换的器件就可以用来表示 1 位二进制数，二进制数的表示简单可靠。目前在计算机中，几乎全部采用二进制计数制。

有时候，人们不得不与二进制数打交道，这时，人们常采用十六进制数来表示二进制数。

下面介绍 3 种计数制及它们之间的相互转换。

1. 十进制数(Decimal)

一个任意的十进制数可以表示为：

$$a_n a_{n-1} \cdots a_0 b_1 b_2 \cdots b_n$$

其值为：

$$a_n \times 10^n + a_{n-1} \times 10^{n-1} + \cdots + a_0 \times 10^0 + b_1 \times 10^{-1} + b_2 \times 10^{-2} + \cdots + b_n \times 10^{-n}$$

其中 $a_i(i=0,1,\cdots,n)$，$b_i(i=1,2,\cdots,n)$ 是 0、1、2、3、4、5、6、7、8、9 十个数字之一。

十进制数计数原则是“逢十进一”，所以基数为 10。上式中相应于每位数字

的 10^k 称为该位数字的权，每位数字与其相应的权的乘积的和即为该数的值。

例如：

$$5555.55=5\times10^3+5\times10^2+5\times10^1+5\times10^0+5\times10^{-1}+5\times10^{-2}$$

虽然六个数字都是 5，但它们的权是不同的，最高位的 5 的权是 10^3，最低位的 5 的权是 10^{-2}。

2. 二进制数(Binary)

一个任意的二进制数可以表示为：

$$a_n a_{n-1}\cdots a_0 b_1 b_2\cdots b_n$$

其值为：

$$a_n\times2^n+a_{n-1}\times2^{n-1}+\cdots+a_0\times2^0+b_1\times2^{-1}+b_2\times2^{-2}+\cdots+b_n\times2^{-n}$$

其中 $a_i(i=0,1,\cdots,n)$，$b_i(i=1,2,\cdots,n)$ 是 0、1 两个数字之一。

二进制数计数原则是“逢二进一”，所以基数为 2。上式中相应于每位数字的 2^k 称为该位数字的权，所有各位数字的加权和即为该数的值。

例如：

$$(1010.01)_2=1\times2^3+0\times2^2+2^1+0\times2^0+0\times2^{-1}+1\times2^{-2}=(10.75)_{10}$$

其中数的下标表示该数的基数 r，即二进制的 1010.01 与十进制的 10.75 等值。

N 位二进制数可以表示 2^N 个数。例如 4 位二进制数则表示十进制中的 0～15 共 16 个数，如表 1－1 所示。

表 1－1　4 位二进制数对应的十进制

二进制数	0000	0001	0010	0011	0100	0101	0110	0111
相应的十进制数	0	1	2	3	4	5	6	7
二进制数	1000	1001	1010	1011	1100	1101	1110	1111
相应的十进制数	8	9	10	11	12	13	14	15

3. 十六进制数(Hexadecimal)

为了便于人们记忆，经常用八进制或十六进制来表示二进制数。一个任意的十六进制数可表示为：

$$a_n a_{n-1}\cdots a_0 b_1 b_2\cdots b_n$$

其值为：

$$a_n\times16^n+a_{n-1}\times16^{n-1}+\cdots+a_0\times16^0+b_1\times2^{-1}+b_2\times16^{-2}+\cdots+b_n\times16^{-n}$$

其中 $a_i(i=0,1,\cdots,n)$，$b_i(i=1,2,\cdots,n)$ 是 0～9、A、B、C、D、E、F 十六个数字之一。

十六进制数计数是“逢十六进一”，所以基数为 16。上式中相应于每位数字的 16^k 称为该位数字的权，所有数字加权和即为该数的值。

例如：

$(28)_{16}=2\times16^1+8\times16^0=(40)_{10}$

$(18F.4)=1\times16^2+8\times16^1+F\times16^0+4\times16^{-1}=(399.25)_{10}$

读者可能会产生疑问，既然人们习惯于用十进制记数制，而计算机为了制造简单采用二进制记数制，那么，为什么这里又要提到十六进制记数制呢？这是因为二进制和十六进制数之间有一种特殊的关系：$2^4=16$。于是，四位二进制数可以用一位十六进制数来表示，这样便于人们阅读和书写冗长的二进制数，而且它们之间存在直接而又惟一的对应关系，如表 1－2 所示。

表 1－2 二进制、十进制、十六进制数码对照表

十进制	十六进制	二进制	十进制	十六进制	二进制
0	0	0000	8	8	1000
1	1	0001	9	9	1001
2	2	0010	10	A	1010
3	3	0011	11	B	1011
4	4	0100	12	C	1100
5	5	0101	13	D	1101
6	6	0110	14	E	1110
7	7	0111	15	F	1111

4. 任意进制计数

基数为 r 的 r 进制数的值可以表示为：

$$a_n\times r^n+a_{n-1}\times r^{n-1}+\cdots+a_0r^0+b_1r^{-1}+b_2r^{-2}+\cdots+b_3r^{-3}$$

其中 $a_i(i=0,1,\cdots,n)$，$b_i(i=1,2,\cdots,n)$是 0～r－1 中任意一数字，r^k则为各位数相应的权。

使用数制时要考虑计算机的特点，又要顾及到人们的习惯，通常用数字后面跟一个英文字母来表示该数的数制。十进制数一般用 D(Decimal)、二进制数用 B(Binary)、八进制数用 O(Octal)、十六进制数用 H(Hexadecimal)来表示。例如：1101B 表示二进制数的 1101，555D 表示十进制数的 555，3FA2H 则表示十六进制数的 3FA2，……

1.2.2 数制转换

1. 十六进制与二进制数相互转换

每一位十六进制数都可以用相应的四位二进制数表示，反之亦然。这样，要把一个十六进制数转换为对应的二进制数就很简单，只需逐位把十六进制数字

试读结束：需要全本请在线购买： www.ertongbook.com