

第一章 计算机系统概论

1.1 计算机与语言

科学技术的高度发展,导致了计算机的诞生。在现代化社会中,计算机已深入到人类工作与生活的各个角落。计算机与其它机器一样,是人类和自然作斗争以及从事各项社会活动的工具。由于它具有计算、模拟、分析问题、操纵机器、处理事务等能力,所以被看作是人脑的延伸,是一种有“思维”能力的机器,从这点出发,计算机又被称为“电脑”。但是一切机器,包括计算机在内,都是人类智慧的结晶,都是人创造的,同时又受人的操纵与控制。

人们经常用语言(或文字)来表达思想、交流经验、互通信息。其中汉语、英语、法语等是使用人数最多的语种。人类相互交流信息所用的语言称为自然语言,但是当前的计算机还不具备理解自然语言的能力,于是人们希望找到一种和自然语言接近,并能为计算机接受的语言,这种语言被称为计算机的高级语言。从计算机的发展历史来看,最初在计算机中使用的不是高级语言,由于它难以理解,使用困难,因而需要改进,这样才导致了高级语言的诞生。

常用于科学计算和数据处理的高级语言有: BASIC、FORTRAN、ALGOL、PASCAL、COBOL 和 Ada 等。常用于人工智能的语言有 LISP 和 PROLOG 等。用这些语言编写的程序是由英文字母、数字、运算符号等按照一定的语法规则组成的。然而目前的通用计算机不会直接执行用高级语言编写的程序,而是先将其翻译成机器能执行的语言,称为机器语言(由二进制代码表示的指令组成),再在机器上运行,因此解题的过程可归结为

- (1) 程序员用高级语言编写程序;
- (2) 将程序与数据输入计算机,并由计算机将程序翻译成机器语言程序,保存在计算机的存储器中;
- (3) 运行程序,输出结果。

存储器是计算机中用以存放原始数据、程序以及中间运算结果的设备,最后的处理结果也往往先暂时存放在存储器中,然后再输送出去。存储器分成一个个单元,每个单元有自己的编号,称为该单元的地址。数据或指令以二进制代码形式存放在存储器中。

在计算机中能执行的程序是由指令组成的,因此计算机执行程序的过程,实际上就是按照给定次序执行一组指令的过程。

一条指令通常分成两部分:

- (1) 操作码 规定该指令执行什么样的运算(或操作),因此被命名为操作码。
- (2) 地址码 规定对哪些数据进行运算,通常表示的是数据地址,因此被称为地址码。当前,计算机指令类型很多,各条指令的功能差异很大,并不局限于对数据进行运算,甚至有的指令不需要地址,因此地址码的含义是灵活多变的。我们将在第五章指令系统专门讨论这个问题。

由于二进制码不易辨认,因此往往用符号来表示一条指令,例如加法运算指令可用符号表示如下:

ADD A,B

其中 ADD 为指令的操作码,A、B 为两个操作数的地址码,并隐含指定将运算结果送到地址 A 或 B 中。假如 A 中已存放有十进制数 2,E 中为 3,并默认运算结果送 A,那么执行本条指令以后,A 中的内容将更换成 5,B 中的内容保持不变,仍为 3。

用机器语言编写程序,比用高级语言麻烦得多,那是因为一条机器指令的功能比一条高级语言的语句功能弱很多而造成的,例如用 BASIC 语言编写的程序,执行语句

LET d=b * b-4 * a * c

即可得到 $d=b^2-4ac$ 的值,而用机器语言则需要五条指令才能实现。当用符号来表示指令时,其程序如下:

程序	注释(运算结果)
1. MUL B B	b^2 送入 B
2. MUL A E	$4a$ 送入 A
3. MUL A C	$4ac$ 送入 A
4. SUB B A	b^2-4ac 送入 B
5. MOV D B	b^2-4ac 从 B 传送到 D
A	a
B	b
C	c
D	d
E	4

其中 1~5 为指令,MUL 为乘法指令的操作码,SUB 为减法指令,MOV 为传送指令。A、B、C、D、E 分别表示存储数 a、b、c、d 及常数 4 的地址,上述这些指令统称为算述逻辑运算指令。

指令前面的序号有示指令的执行顺序,也表示该指令在存储器中的相对位置,必须按此顺序将指令存放在相邻的存储单元中。

例如,第 1 条指令存放在地址为 n 的存储单元中,则其后继的指令依次存放在 $n+1$ 、 $n+2$ 、 $n+3$ 和 $n+4$ 的存储单元中。编制程序时还需考虑求得 d 值后机器如何运行问题。如果此时已不再需要进行其它工作,则在 $n+5$ 可安排一条停机指令或动态停机指令(等待指令)。动态停机指令不完成任何有效的具体操作,仅使计算机处于“空转”状态,待有某些特定信号(如中断信号,见本书第十章)来到时,才转到相应的程序入口继续运行。如果此时还需要进行其它工作,则从 $n+5$ 开始继续编制程序,或者安排一条转移指令,将程序转到需执行处。

数据地址 A、B、C、D 和 E,从原则上讲,相互之间不受约束,即可存放在主存储器任何有空闲的地方。但习惯上经常也是顺序安放的。于是可将程序改写如下:

n	MUL	$n+7$	$n+7$	$n+6$	a
$n+1$	MUL	$n+6$	$n+10$	$n+7$	b
$n+2$	MUL	$n+6$	$n+8$	$n+8$	c
$n+3$	SUB	$n+7$	$n+6$	$n+9$	d
$n+4$	MOV	$n+9$	$n+7$	$n+10$	4
$n+5$	HLT				

HLT 为停机指令。

1.2 计算机的硬件

组成计算机的基本部件有中央处理器 CPU(运算器和控制器)、存储器和输入/输出设备。

输入设备用来输入原始数据和处理这些数据的程序。输入的信息有数字符、字母和控制符等,人们经常用 8 位二进制码来表示一个数字符(0~9)、一个字母(A、B、C、 \dots , X, Y, Z)或其它符号,当前通用的是 ASCII 码,它用七位二进制码来表示一个字符,最高的一位可用于奇偶校验或作其它用处。在计算机中,一般把 8 位二进制码称为一个字节。在我国使用的计算机,一般有处理汉字的能力,在本书第九章作进一步说明。

输出设备用来输出计算机的处理结果,可以是数字、字母、表格、图形等。最常用的输入输出设备是显示终端和打印机,终端设备采用键盘作为输入工具,处理结果显示在屏幕上,而打印机则将结果打印在纸上,除此以外,为了监视人工输入信息的正确性,在用键盘输入信息时,将刚输入的信息显示在屏幕上,如有错误,可及时纠正。

存储器用来存放程序和数据,是计算机各种信息的存储和交流中心。存储器可与 CPU、输入输出设备交换信息,起存储、缓冲、传递信息的作用,在这里,我们要注意把存储单元的地址和存储单元里存放的内容(数据或指令)区分开。

存储器又有主存储器和辅助存储器之分。当前在计算机上运行的程序和数据是存放在主存储器中的。

中央处理器又叫 CPU,在早期的计算机中分成运算器和控制器两部分,由于电路集成度的提高,现在已把它们集成在一个芯片中。

运算器是对信息或数据进行处理和运算的部件。经常进行的是算术运算和逻辑运算,所以在其内部有一个算术及逻辑运算部件(ALU)。算术运算是按照算术规则进行的运算,例如加、减、乘、除、求绝对值、求负值等。逻辑运算一般是指非算术性质的运算,例如比较大小、移位、逻辑乘、逻辑加等。在计算机中,一些复杂的运算往往被分解成一系列算术运算和逻辑运算。

当 CPU 处理的数据局限于整数时,这个 CPU 有时被称为整数运算部件 IU。为了快速而有效地对实数进行处理,在某些计算机中专门设置了浮点运算部件。

控制器主要用来实现计算机本身运行过程的自动化,即实现程序的自动执行。在控制器控制之下,从输入设备输入程序和数据,并自动存放在存储器中,然后由控制器指挥各部件(运算器、存储器 \dots)协同工作以执行程序,最后将结果打印输出。作为控制用的计

算机则直接控制对象。

在计算机中,各部件间来往的信号可分成三种类型,它们是:地址、数据和控制信号。通常这些信号是通过总线传送的,如图 1.1 所示。CPU 发出的控制信号,经控制总线送到存储器和输入输出设备,控制这些部件完成指定的操作。与此同时,CPU(或其它设备)经地址总线向存储器或输入输出设备发送地址,使得计算机各个部件中的数据能根据需要互相传送。输入输出设备和存储器有时也向 CPU 送回一些信号,CPU 可根据这些信号来调整本身发出的控制信号。现代计算机还允许输入输出设备直接向存储器提出读写要求,控制数据传送。

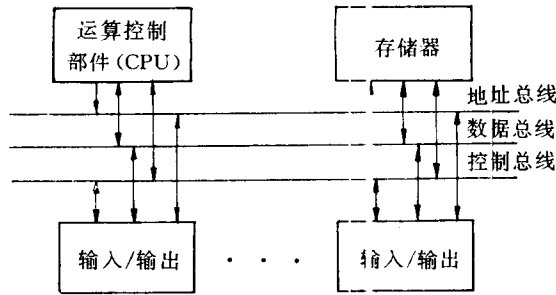


图 1.1 以总线连接的计算机框图

1.3 计算机系统的层次结构

现代计算机解题的一般过程:用户用高级语言编写程序,连同数据一起送入计算机(用户程序一般称为源程序),然后由计算机将其翻译成机器语言程序(称为目标程序),在计算机上运行后输出结果,其过程如图 1.2 所示。

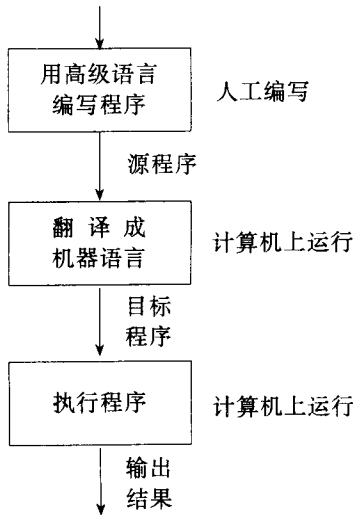


图 1.2 计算机的解题过程

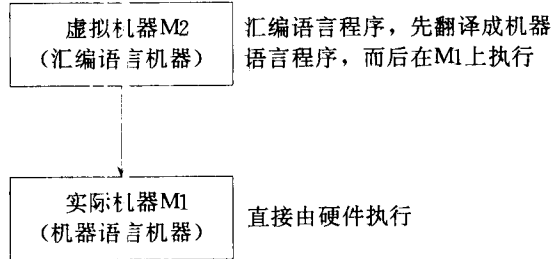


图 1.3 汇编语言虚拟机的层次结构

早期的计算机只有机器语言,用户必须用二进制码表示的机器语言编写程序(用八进制或十六进制书写),因此工作量大,容易出错,而且对程序员的要求很高,要求他们计算机的硬件和指令系统有正确和深入的理解,并有熟练的编程技巧,只有少数专家才能达到此要求,于是在50年代,出现了符号式程序设计语言,称为汇编语言,对此,程序员可用ADD, SUB, MUL, DIV等符号分别表示加法、减法、乘法、除法的操作码,并用符号来表示指令和数据的地址。汇编语言程序的大部分语句是和机器指令一一对应的。用户用汇编语言编写程序后,依靠计算机将它翻译成机器语言(二进制代码),然后再在计算机上运行。这个翻译过程是由汇编程序实现的。

我们可以把一台具有汇编程序的计算机看作是在实际机器级(硬件)之上出现的一台虚拟机器,该机允许使用汇编语言编程。我们称它为虚拟机器的原因是因为它依靠了软件(汇编程序)才存在。图1.3表示这台机器的层次结构,对用户来讲,他所看到的机器已不是实际机器M1,而是虚拟机器M2。

由于汇编语言的语法、语义结构仍然和机器语言基本一样,而与人的传统解题方法相差甚远,因而经过了人们的努力又出现了面向题目的高级语言。随同研制出来的是这些语言的翻译程序,因此我们可以设想在汇编语言级之上又出现了高级语言级,它的实现是先把高级语言程序翻译成汇编语言程序或中间语言程序,尔后再翻译成机器语言程序。有的计算机则将高级语言程序直接翻译成机器语言程序。图1.4(a)和(b)分别表示这两种情况。

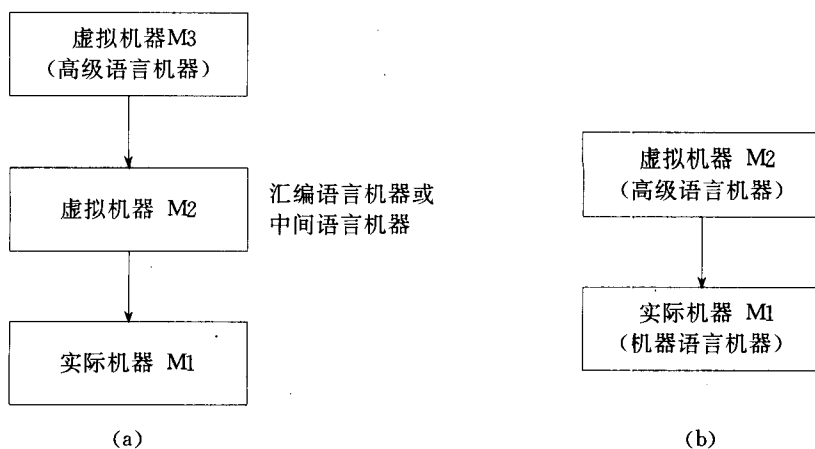


图 1.4 高级语言虚拟机器的层次结构

翻译程序有编译程序和解释程序两种。

编译程序是将用户编写的源程序中全部语句翻译成机器语言程序后,再执行机器语言程序。假如一个题目需要重复计算几遍,那末一旦翻译以后,只要源程序不变,不需要再次进行翻译。但源程序若有任何修改,都要重新经过编译。

解释程序则是在将源程序的一条语句翻译成机器语言以后立即执行它(而且不再保存刚执行完的机器语言程序),然后再翻译执行下一条语句。如此重复,直到程序结束。它的特点是翻译一次只能执行一次,当第二次重复执行该语句时,又要重新翻译,因而效率

较低。当前,ALGOL、FORTRAN、PASCAL 等语言是用编译程序进行翻译的,而 BASIC 语言则有解释和编译两种。

在上述虚拟机 M2 与实际机器 M1 之间还存在一种称为操作系统的软件。操作系统是从管理程序发展而来的,它提供了实际机器所没有的但在汇编语言和高级语言的使用和实现过程中所需的某些基本操作和数据结构。

操作系统的功能(命令)是通过操作系统的指令系统(或称为控制语言)实现的,在一般计算机上操作系统的指令是经过用机器语言编写的解释程序实现的,因此操作系统可看作实际机器的扩充,在计算机系统的多级层次结构中应占有一个席位,它的位置应在实际机器之上、汇编语言机器级之下。图 1.5 是计算机系统的多级层次结构。

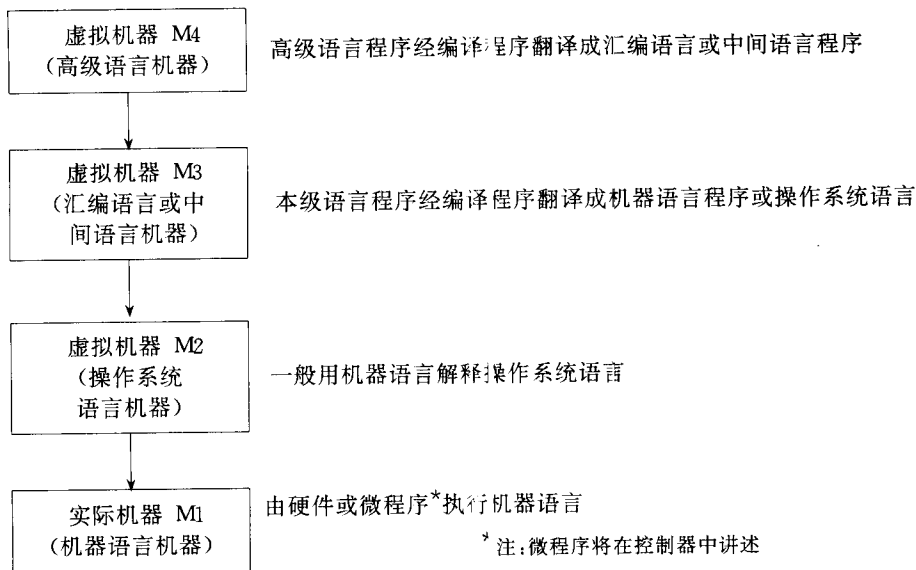


图 1.5 计算机系统的多级层次结构

不同计算机系统之间的多级层次结构的划分与实现方法是有差别的,一般来说,相邻级语言的语法结构的差别不要太大,这样才便于编译或解释,但最后总是要翻译成能在机器上执行的机器语言程序。

1.4 电子计算机的发展简史

电子计算机的发展,如果从第一台计算机的问世算起,到现在才 40 余年,在人类科技史上还没有一种学科可以与电子计算机的发展之快相提并论。

20 世纪四十年代,无线电技术和无线电工业的发展为电子计算机的研制准备了物质基础,1943 年—1946 年美国宾夕法尼亚大学研制的电子数字积分和计算机 ENIAC (Electronic Numerical Integrator And Computer) 是世界上第一台电子计算机。当时第二次世界大战正在进行,为了进行新武器的弹道问题中许多复杂的计算,在美国陆军部的资助下开展了这项研究工作,ENIAC 计算机于 1945 年年底完成,1946 年 2 月正式交付使

用,因为它是最早问世的一台电子数字计算机,所以一般人认为它是现代计算机的始祖。

ENIAC 计算机共用 18000 多个电子管,1500 个继电器,重达 30 吨,占地 170 平方米,耗电 140kW,每秒钟能计算 5000 次加法,领导研制的是埃克特(J. P. Eckert)和莫克利(J. W. Mauchly)。ENIAC 计算机存在两个主要缺点,一是存储容量太小,只能存 20 个字长为 10 位的十进制数,二是用线路连接的方法来编排程序,因此每次解题都要依靠人工改接连线,准备时间大大超过实际计算时间。

与 ENIAC 计算机研制的同时,冯·诺依曼(Von Neumann)与莫尔小组合作研制 EDVAC 计算机,采用了存储程序方案,其后开发的计算机都采用这种方式,称为冯·诺依曼计算机。一般认为冯·诺依曼机具有如下基本特点:

(1) 计算机由运算器、控制器、存储器和输入设备和输出设备五部分组成。

(2) 采用存储程序的方式,程序和数据放在同一个存储器中,指令和数据一样可以送到运算器运算,即由指令组成的程序是可以修改的

(3) 数据以二进制码表示。

(4) 指令由操作码和地址码组成。

(5) 指令在存储器中按执行顺序存放,由指令计数器(即程序计数器 PC)指明要执行的指令所在的单元地址,一般按顺序递增,但可按运算结果或外界条件而改变。

(6) 机器以运算器为中心,输入输出设备与存储器间的数据传送都通过运算器

四十多年来,随着技术的发展和 new 应用领域的开拓,对冯·诺依曼机作了很多改革,使计算机系统结构有了很大新发展,如某些机器程序与数据分开存放在不同的存储器中,程序不允许修改,机器不再以运算器为中心,而是以存储器为中心等等,虽然有以上这些突破,但原则变化不大,习惯上称之为冯·诺依曼机。

四十年来,根据电子计算机所采用的物理器件的发展,一般把电子计算机的发展分成四个阶段,习惯上称为四代。两代计算机之间时间上有重叠。

第一代:电子管计算机时代(从 1946 年第一台计算机研制成功到五十年代后期),其主要特点是采用电子管作为基本器件。在这一时期,主要为军事与国防尖端技术的需要而研制计算机,并进行有关的研究工作,为计算机技术的发展奠定了基础,其研究成果扩展到民用,又转为工业产品,形成了计算机工业。

五十年代中期,美国 IBM 公司在计算机行业中崛起,1954 年 12 月推出的 IBM650(小型机)是第一代计算机中畅销最广的机器,销售量超过一千台。1958 年 11 月问世的 IBM709(大型机)是 IBM 公司性能最高的最后一个电子管计算机产品。

第二代:晶体管计算机时代(从五十年代中期到六十年代后期),这时期计算机的主要器件逐步由电子管改为晶体管,因而缩小了体积,降低了功耗,提高了速度和可靠性。而且价格不断下降。后来又采用了磁心存储器,使速度得到进一步提高。不仅使计算机在军事与尖端技术上的应用范围进一步扩大,而且在气象、工程设计、数据处理以及其它科学研究等领域内也应用起来。在这一时期开始重视计算机产品的继承性,形成了适应一定应用范围的计算机“族”,这是系列化思想的萌芽。从而缩短了新机器的研制周期,降低了生产成本,实现了程序兼容,方便了新机器的使用。

1960 年控制数据公司(CDC)研制高速大型计算机系统 CDC6600,于 1964 年完成,取

得了巨大成功,深受美国和西欧各原子能、航空、宇航、气象研究机构和大学的欢迎,使该公司在研究和生产科学计算高速大型机方面处于领先地位。1969年1月,水平更高的超大型机 CDC7600 研制成功,平均速度达到每秒千万次浮点运算,成为六十年代末、七十年代初性能最高的计算机。

第三代:集成电路计算机时代(从六十年代中期到七十年代前期),这时期的计算机采用集成电路作为基本器件,因此功耗、体积、价格等进一步下降,而速度及可靠性相应地提高,这就促使了计算机的应用范围进一步扩大。正是由于集成电路成本的迅速下降,产生了成本低而功能不是太强的小型计算机供应市场。占领了许多数据处理的应用领域。

IBM360 系统是最早采用集成电路的通用计算机,也是影响最大的第三代计算机。在1964年宣布 IBM360 系统时就有大、中、小型等6个计算机型号,平均运算速度从每秒几千次到一百万次,它的主要特点是通用化、系列化、标准化。

通用化:指令系统丰富,兼顾科学计算、数据处理、实时控制三个方面。

系列化:IBM360 各档机器采用相同的系统结构,即在指令系统、数据格式、字符编码、中断系统、控制方式、输入/输出操作方式等方面保持统一,从而保证了程序兼容,当用户更新机器时原来在低档机上编写的程序可以不作修改就使用在高档机中。IBM360 系统后来陆续增加的几种型号仍保持与前面的产品兼容。后来,西欧与日本的一些通用计算机也保持与 IBM360 系统兼容。苏联和东欧国家联合制造的“统一系统”也是与 IBM360 系统兼容的。

标准化:采用标准的输入/输出接口,因而各个机型的外部设备是通用的。采用积木式结构设计,除了各个型号的 CPU 独立设计以外,存储器、外部设备都采用标准部件组装。

第四代:大规模集成电路计算机时代,七十年代初,半导体存储器问世,迅速取代了磁心存储器,并不断向大容量、高速度发展,此后,存储器芯片集成度大体上每三年翻两番(1971年每片1K位,到1984年达到每片256K位,1992年16M位动态随机存储器芯片上市),价格平均每年下降30%。逻辑电路也得到相应的发展。

随着大规模集成电路的迅速发展,计算机进入大发展时期,通用机、巨型机、小型机、微型机以及工作站都得到了发展。

(1) 通用机

通用机是计算机工业中价值比重最大的产品,其中以 IBM370 系统影响最大,它在与 IBM360 系统兼容的前提下进行了改进。IBM 公司为开发 360 系统的软件耗费了巨大的人力和财力,据估算,IBM 用户在应用程序、培训等方面耗费了两千亿美元,是硬件投资的三至五倍,如此丰富的软件不能抛弃、只能继承,这已成为用户与计算机厂家共同遵守的原则,但也成了计算机发展的制约。继 IBM370 以后,IBM303X 大型机系列仍与 IBM370 系统兼容,但具有更强的科学计算处理能力;IBM4300 系列取代了 IBM370 系统的低档机,但仍与 IBM370 兼容。1982 年宣布的 IBM3084K 大型通用机速度达到每秒 2500 万次,主存容量为 64 兆字节。

其它计算机厂家在发展新机种时也遵循兼容的原则。某些计算机厂家走上与 IBM 计算机兼容的道路,称之为 PCM: Plug Compatible Mainframe(插接兼容主机硬件完全兼容)或 Program Compatible Mainframe(程序兼容主机——软件兼容),制造与 IBM 兼容

的计算机,它们按 IBM 系列机的系统结构制造主机,并直接引用 IBM 计算机的软件,因而使产品的性能价格比优于 IBM 原装机,以争夺市场。

(2) 巨型机

现代科学技术,尤其是国防技术的发展,需要有很高运算速度、很大存储容量的计算机,一般的大型通用计算机不能满足要求。集成电路的进展,为制造巨型机提供了条件。从六十年代到七十年代相继完成了一些巨型机,其中取得最高成绩的要推 Cray-1 计算机。针对天气预报、飞行器的设计和核物理研究中存在大量向量运算的特点,Cray-1 计算机的向量运算速度达每秒 8000 万次,并兼顾了一般的标量运算。1983 年研制成功的 Cray X-MP 机向量运算速度达每秒 4 亿次。与此同时,CDC 公司的 CYBER203 和 205 先后完成,CYBER205 每秒可进行 4 亿次浮点运算。这些是八十年代初期的水平最高的巨型机。但是这些成就还不能满足一些复杂问题的需要,所以不少单位开展了性能更高的巨型机的研究工作。近年来微处理器的发展为阵列结构的巨型机发展带来了希望,例如古德伊尔公司为美国宇航局(NASA)研制了一台处理卫星图象的巨型计算机系统 MPP,该机由 16384 个微处理器组成 128×128 方阵。这种采用并行处理技术的多处理器系统是巨型机发展的一个重要方面,称为小巨型机。

日本、英国、苏联、法国也先后开始研制巨型机。

(3) 小型机

小型机规模小、结构简单所以设计试制周期短,便于及时采用先进工艺,生产量大,硬件成本低;同时由于软件比大型机简单,所以软件成本也低。再加上容易操作、容易维护和可靠性高等特点,使得管理机器和编制程序都比较简单,因而得以迅速推广,掀起一个计算机普及应用的浪潮。DEC 公司的 PDP-11 系列是 16 位小型机的代表,到七十年代中期 32 位高档小型机开始兴起,DEC 公司的 VAX11/780 于 1978 年开始生产,应用极为广泛。VAX11 系列与 PDP11 系列是兼容的。

小型机的出现打开了在控制领域应用计算机的局面,许多大型分析仪器、测量仪器、医疗仪器使用小型机进行数据采集、整理、分析、计算等。应用于工业生产上的计算机除了进行上述工作外还可进行自动控制。

(4) 微型机

微型机的出现与发展,掀起计算机大普及的浪潮,利用 4 位微处理器 Intel4004 组成的 MCS-4 是世界上第一台微型机,它于 1971 年问世。Intel8086 是最早开发成功的 16 位微处理器(1978 年),以后开发的 Intel80286、80386 与 8086 兼容。1981 年以后 32 位微处理器相继问世,比较著名的 32 位微处理器有 Intel80386、Motorola 的 68020 和 68030 等。Intel80386 片内集成了 27.6 万个晶体管、Motorola68030 片内集成了 30 万个晶体管。1990 年 Intel80486 和 Motorola68040 推向市场,其集成度达到 120 万个晶体管,与原来的产品相比较,除了提高主频速度外,还将原属片外的有关电路集成到片内。

32 位微处理器采用过去大中型计算机中所采用的技术,因此用它构成的微型机系统的性能可以达到七十年代大中型计算机的水平。

七十年代后期,兴起个人计算机(一种独立微型机系统)热潮,最早出现的是 Apple 公司的 Apple II 型微机(1977 年),此后各种型号的个人计算机纷纷出现。1981 年一向以

生产大中型通用机为主的 IBM 公司推出了 IBM PC 机,后来又推出扩充了性能的 IBM PC/XT、IMB PC/AT 以及 386、486 和 Pentium 等多种机型,由于具有设计先进、软件丰富、功能齐全、价格便宜等特点,很快成为微型机市场的主流,国内外有不少厂家相继生产了与 IBM 兼容的个人计算机。

微型机向小型化发展出现了便携机(膝上型、笔记本型和掌上型),预计在 90 年代会获得迅速发展。

低档的个人计算机可作为家用电脑,国内市场上的家用电脑一般采用 i80286 或 i80386 芯片,用于教育、游戏及家庭管理等。高档的用于经营管理、科学计算以及学校教育等方面。

(5) 工程工作站

工程工作站是八十年兴起的面向广大工程技术人员计算机系统,一般具有高分辨率显示器、交互式的用户界面和功能齐全的图形软件。开始集中应用于各种工程方面的计算机辅助设计,如集成电路设计、机械设计、土木建筑设计等。1980 年成立的 Apollo 公司和 1982 年成立的 Sun 微机系统公司主要从事工作站的研制与生产工作。开始都采用 Motorola 的微处理器芯片,后来改用 RISC(精简指令系统计算机)微处理器。

由于工程工作站出现得比较晚,一般都带有网络接口,并采用开放式系统结构,即将机器的软、硬件接口公开,以鼓励其它厂商、用户围绕工作站开发软、硬件产品。同时尽量遵守国际工业界流行的标准。

(6) 联机系统和计算机网

由于计算机技术和通信技术的迅速发展,为适应高度社会化生产和科技发展的需要,出现了由单个计算中心通过通信线路和若干个远程终端连接起来的联机系统(或称为面向终端的网络),例如库存管理系统、生产管理系统、银行业务系统,飞机订票系统、情报检索系统、气象观测系统等,使分散在各处的信息通过终端能很快集中于计算机中,同时各处的工作人员可通过终端进行查询、获取资料。

在七十年代,能实现计算机之间的通信、并共享资源的计算机网迅速发展。著名的美国 ARPA 网诞生于六十年代末,在七十年代不断扩充网上节点,到 1975 年已连接 60 个以上的节点,一百多台主计算机。地理范围遍布全美并扩展到欧洲。与此同时其它网相继建成。由于这些网络跨越的地理范围比较宽阔,因而称为广域计算机网。一些主要计算机厂家为解决本公司生产的各种计算机之间和计算机与终端设备的联网问题,向用户提供相应的硬件(如通信接口板)和网络软件。

随着计算机的广泛应用,特别是小型机和微型机的普及,一个单位在一幢大楼或一个建筑群内安装多台计算机的情况日益普遍,将这些计算机联接在一起的网络称为局部网。

计算机网的蓬勃发展,加速了社会信息化的进程。

上面讲到,根据所用器件的不同,电子计算机经历了四代的变革,但其基本思想一直遵循冯·诺依曼计算机结构的原理。

计算机的发展促进了人工智能的发展,1981 年日本政府提出了发展第五代计算机的十年计划,突破了冯·诺依曼结构原理。该计划从 1982 年开始执行,现已结束,没有取得预期的结果。日本的第五代计算机的目标是实现智能计算机。美国也有多家公司推出了

智能计算机。一般要求智能计算机具有下列功能：

(1) 智能接口功能 能自动识别自然语言(文字、语音)和图形、图象(视觉)能力。语音识别与理解、机器翻译等均属于自然语言处理的范畴。通过摄像机把图象输入计算机后,计算机能通过图象理解系统和景物感知系统得到信息,并对这些信息进行分析、认识和理解,这是智能机器人应具备的重要功能。军事侦察、地图摄制、字符识别、自动检索等领域内都有计算机视觉的功能。

(2) 理解和推理功能 能根据计算机内存储的信息(知识)进行推理,具有问题求解和学习的功能。

(3) 知识库管理功能 要求能完成知识获取、知识检索和知识更新等功能。

随同计算机硬件发展的还有软件,应该指出,发挥计算机的作用,推广计算机的应用,改进计算机的设计以及简化计算机的操作,使它从只供专家使用转为面向大众,软件工作者起了决定性的作用。

高级程序设计语言在第二代计算机时期趋向成熟并迅速普及,操作系统自动地管理计算机系统中各个设备以及多个程序的高效运行,是第三代计算机时期的重大成就,以上这些软件属于系统软件。

广泛应用计算机的结果,在科学计算、数据处理、商业经营、经济管理、工业控制、工程设计等领域中开发出各自的程序,称为应用软件。计算机厂家向用户提供软件(系统软件和应用软件)时与硬件分别计价,并产生了专门从事软件研制、生产、销售工作的软件公司(例如美国的 Microsoft 公司)。但是软件的发展跟不上需要,软件费用急剧增长,这是因为硬件是工业化生产,价格不断下降,而软件为人工劳动,生产率低。一些科学家提出了软件工程的概念,对软件开发实行工程化管理,以期得到廉价、可靠、有效的软件。软件还具有容易复制的特点,软件成果容易被别人占有,因此影响了软件开发者进行软件开发及将软件投入市场的积极性。为了保护软件不被剽窃,可以采取加密码等技术措施以及低价销售、随硬件提供等经营措施,发挥一定的保护作用,但不能彻底解决问题,因此由国家来制订、实施对软件的保护法律是至关重要的。但是一个国家的法律只适用于国内,而软件很容易在国家之间传播,因此国与国之间相互承担保护对方公民(和法人)软件的义务已成为各国之间经济合作关系的一个重要组成部分。

世界上科学技术发达的国家,十分重视鼓励创造性的脑力劳动,已经研究、制定并贯彻实施专利法、版权法、商标法等一系列保护公民知识产权的法律。很多发展中国家也开始重视这项工作,正纷纷开展保护知识产权的立法工作。1991年5月我国颁布了《计算机软件保护条例》。

1.5 计算机的应用

1.5.1 科学计算

科学计算一直是电子计算机的重要应用领域之一。例如在天文学、量子化学、空气动力学、核物理学等领域中,都需要依靠计算机进行复杂的运算。在军事上,导弹的发射及飞行轨道的计算控制、先进防空系统等现代化军事设施通常都是由计算机控制的大系统,其

中包括雷达、地面设施、海上装备等。现代的航空、航天技术发展,例如超音速飞行器的设计,人造卫星与运载火箭轨道计算更是离不开计算机。

除了国防及尖端科学技术以外,计算机在其它学科和工程设计方面,诸如数学、力学、晶体结构分析、石油勘探、桥梁设计、建筑、土木工程设计等领域内也得到广泛的应用,促进了各门科学技术的发展。

有些系统,要求计算机处理所得的结果立即反过来作用或影响正在被处理的事物本身,例如在控制导弹飞行的系统中,不断测量导弹飞行的参数(包括飞行环境),并及时作出反应,修正导弹飞行的轨迹,这样的系统称为实时处理系统。

科学计算的特点是计算量大和数值变化范围大。

1.5.2 数据处理

当前大部分计算机都用于数据处理。下面以银行系统(储户处理)为例来说明数据处理计算机的工作情况。

用计算机处理储户的存款、取款、直接发薪以及其它诸如信用卡系统、销售点系统等银行业务,这类计算机系统一般配置有多台终端设备,用于将储户的有关信息及数据输入计算机或用于自动支付现金(取款)。

信用卡是一张上面具有一小磁条的卡片,磁条上记有持有者特征,诸如编号等信息,当插入银行终端设备的插口后,磁条上记录的特征信息就读入与终端相连的计算机。假如信用卡持有者是该银行的储户,那末计算机自动将与该储户有关的帐目从计算机中取出,然后进行存款、取款或冲帐等工作,并将结算后的金额重新存入计算机。

信用卡还可用来买东西,凡是有相应银行业务终端的地方,都可用它来付款。

通过计算机系统将工作人员的工资,直接转到银行,自动加到储户的存款金额中去,可以实现自动发放工资。

从上例可见,数据处理系统具有输入/输出数据量大而计算却很简单的特点。为了实现各储蓄所之间的通存通兑以及在商店、酒楼中使用信用卡,计算机需要联网使用。

近年内获得迅速发展的 IC 卡(集成电路卡),将 CPU 和存储器安装在卡内芯片中,用它作为信用卡,具有更高的保密性和安全性。

在企业数据处理领域中,计算机广泛应用于财会统计与经营管理中,如编制生产计划、统计报表、成本核算、销售分析、市场预测、利润评估、采购订货、库存管理、工资管理等。为了适应计算机管理,在报表格式的修改,名词统一编码等多方面要进行大量工作。

1.5.3 计算机控制

在现代化工厂里,计算机普遍用于生产过程的自动控制,例如在化工厂中用计算机来控制配料、温度、阀门的开闭等;在炼钢车间用计算机控制加料、炉温、冶炼时间等;程控机床加工的机械零件具有尺寸精确的特点,而且不需要专用工卡具、模具和熟练技工就可以制造出形状复杂的产品。

用于生产过程自动控制的计算机,一般都是实时控制,它们对计算机的速度要求不高,但可靠性要求很高,否则将生产出不合格的产品,甚至造成重大设备事故或人身事故。

用于控制的计算机,其输入信息往往是电压、温度、机械位置等模拟量,要先将它们转换成数字量,称为模/数转换,然后计算机才能进行处理或计算。当从被控制对象测量到的信息是温度、位置等非电量时,要先将它们转换成电量,然后再转换成数字量。如何测量,用什么仪表测量也是一个很重要的问题。计算机的处理结果是数字量,一般要将它们转换成模拟量去控制对象,称为数/模转换。如有需要,可将结果打印输出或显示在屏幕上,以供观察。提供计算机控制系统的厂家往往已将控制程序(称为应用程序包)编制好,可提供给用户。

1.5.4 计算机辅助设计/计算机辅助制造(CAD/CAM)

由于计算机有快速的数值计算、较强的数据处理以及模拟的能力,因而目前在飞机、船舶、光学仪器、超大规模集成电路 VLSI 等的设计制造过程中,CAD/CAM 占据着越来越重要的地位。

在超大规模集成电路的设计和生产过程中,要经过设计制图、照相制版、光刻、扩散、内部连接等多道复杂工序,是人工难以解决的。

使用已有的计算机辅助设计新的计算机,达到设计自动化或半自动化程度,从而减轻人的劳动强度并提高设计质量,这也是计算机辅助设计的一项重要内容。

由于设计工作与图形分不开,一般供辅助设计用的计算机配备有图形显示、绘图仪等设备以及图形语言、图形软件等。设计人员可借助这些专用软件和输入输出设备把设计要求或方案输入计算机,通过相应的应用程序进行计算处理后把结果显示出来,设计人员可用光笔或鼠标器进行修改或选择,直到满意为止。

1.5.5 人工智能

人类的许多脑力劳动,诸如证明数学定理、进行常识性推理、理解自然语言、诊断疾病、下棋游戏、破译密码等都需要“智能”。

人工智能是将人脑在进行演绎推理的思维过程、规则和所采取的策略、技巧等编成计算机程序,在计算机中存储一些公理和推理规则,然后让机器去自动探索解题的方法,所以这种程序不同于计算机的一般应用程序。

当前人工智能在自然语言的理解、机器视觉和听觉等方面给以极大的重视。自然语言是人类交往所用的语言,计算机理解它是很困难的,因为人们通过各自的生产、生活和社会活动,在大脑中已拥有大量的、高度相似的信息或知识,而计算机则没有,另外讲话的语义经常还跟上下文有关,有时还要对讲话的内容进行推理或者演绎才能得出某些结论。

智能机器人是人工智能各种研究课题的综合产物,有感知和理解周围环境、进行推理和操纵工具的能力,并能通过学习适应周围环境,完成某种动作。在不允许人进入的场所(如高温、有放射性物质等)使用机器人有特殊的意义。

专家系统是指用计算机模拟专家行为的程序。例如模拟我国老中医关幼波对肝病进行辩证治疗的程序,按照关幼波的思维方法,把肝病分成 8 个主型 36 个亚型,根据患者的症状,从 200 多项病状与化验指标以及 170 种药物的综合分析基础上,让计算机从多个实用处方中选出针对病情的处方。本例虽不含有太多的“智能”,但这是一个应用实例。

尽管人工智能的研究已取得一些成果,但与目标(建立真正的智能系统)还相差甚远。

第一章习题

1. 说明高级语言、汇编语言、机器语言三者的差别和联系。
2. 计算机硬件由哪几部分组成? 各部分的作用是什么? 各部分之间是怎样联系的?
3. 计算机系统可分哪几个层次? 说明各层次的特点及其相互联系。
4. 如何划分计算机发展的 4 个阶段(第一代到第四代)? 当前广泛应用的计算机主要采取哪一代的技术?
5. 列出通用机、巨型机、小型机、微型机等计算机的典型机种。这些计算机的运算速度、存储容量、价格和应用范围有哪些主要差别?
6. 计算机能够普及应用的主要原因是什么?
7. 冯·诺依曼结构的特点是什么?
8. 你对知识产权问题和软件保护有什么看法?
9. 试用机器语言编写求 5 个数平均值的程序(自定义指令系统)。

第二章 数字化信息编码与数据表示

2.1 数字化信息编码的概念

计算机的最重要的功能是处理信息,如数值、文字、符号、语言、图形和图象等。在计算机内部,各种信息都必须采用数字化编码的形式被传送、存储和加工。因此掌握信息编码的概念与处理技术是至关重要的。

所谓编码,就是用少量简单的基本的符号,选用一定的组合规则,以表示出大量复杂多样的信息。基本符号的种类和这些符号的组合规则是一切信息编码的两大要素。例如,用 10 个阿拉伯数码表示数字,用 26 个英文字母表示英文词汇等,这就是编码的典型例子。

当基本符号数量很多时,往往还要采取措施,以能使用更量少而简单的基本符号来表示那些量大而复杂的基本符号,这就构成了多重编码。多重编码的典型例子是汉字编码。上万个汉字都作为基本符号编码就太多了。可以用笔形、偏旁部首、拼音或其它方式进一步对它们编码。例如,电报通信中就广泛采用把 0~9 的十个符号作为重编码的基本符号,用 4 位十进制数字对常用的 10000 个以内的汉字进行重编码。

在计算机中,广泛采用的是仅用“0”和“1”两个基本符号组成的基 2 码,或称为二进制码。2 称为码制的基。这是由于:

- 基 2 码在物理上最容易实现。例如,用高、低两个电位表示“1”和“0”,或用脉冲的有无、脉冲的正、负极性表示它,可靠性都较高。

- 基 2 码用来表示二进制数,其编码、计数、加减运算规则简单。可用开关电路实现,简便易行。

- 基 2 码的两个符号“1”和“0”正好与逻辑命题的两个值“是”和“否”,或称“真”和“假”相对应,为计算机中实现逻辑运算和程序中的逻辑判断提供了便利的条件。

基 2 码是进一步对各种基本编码符号进行重编码的媒介。

2.2 常用的信息编码

在这一节,重点讲解计算机内最常用的信息编码方案。包括中西文字符编码、逻辑型数据的表示、数值型数据的表示与编码、不同进制数据间的转换、二进制数据算术与逻辑运算规则。这些内容是数字计算机设计与应用的基本理论之一,要求能熟练掌握并运用自如。

2.2.1 字符编码

字符是计算机中使用最多的信息形式之一。是人与计算机通信、交互作用的重要媒

介。在计算机中,要为每个字符指定一个确定的编码,作为识别与使用这些字符的依据。这些编码的值,是用一定位数的基 2 码的两个基本符号“1”和“0”进行重编码给出的。

1. ASCII 码和 EBCDIC 码

使用得最多的、最普遍的是 ASCII 字符编码,即 American Standard Code for Information Interchange,如表 2.1 所示。

表 2.1 ASCII 字符编码表

				b ₆ b ₅ b ₄							
				000	001	010	011	100	101	110	111
b ₃	b ₂	b ₁	b ₀								
0	0	0	0	NUL	DLE	SP	0	@	P	,	p
0	0	0	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	STX	DC2	"	2	B	R	b	r
0	0	1	1	ETX	DC3	#	3	C	S	c	s
0	1	0	0	EOT	DC4	\$	4	D	T	d	t
0	1	0	1	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	ACK	SYN	&	6	F	V	f	v
0	1	1	1	BEL	ETB	'	7	G	W	g	w
1	0	0	0	BS	CAN	(8	H	X	h	x
1	0	0	1	HT	EM)	9	I	Y	i	y
1	0	1	0	LF	SUB	*	:	J	Z	j	z
1	0	1	1	VT	ESC	+	;	K	[k	{
1	1	0	0	FF	FS	,	<	L	\	l	
1	1	0	1	CR	GS	-	=	M]	m	}
1	1	1	0	SO	RS	.	>	N	↑	n	~
1	1	1	1	SI	US	/	?	O	-	o	DEL

从表中可以看到:

- 每个字符是用 7 位基 2 码表示的,其排列次序为 b₆b₅b₄b₃b₂b₁b₀,在表中的 b₆b₅b₄ 为高位部分, b₃~b₀ 为低位部分。而一个字符在计算机内实际是用 8 位表示。正常情况下,最高一位 b₇ 为“0”。在需要奇偶校验时,这一位可用于存放奇偶校验的值,此时称这一位为校验位(有关奇偶校验的内容见本章 2.4 节)。

- ASCII 是 128 个字符组成的字符集。其中编码值 0~31 不对应任何可印刷(或称有字形)字符,通常称它们为控制字符,用于通信中的通信控制或对计算机设备的功能控制。编码值为 32 的是空格(或间隔)字符 SP。编码值为 127 的是删除控制 DEL 码。其余的 94 个字符称为可印刷字符,有人把空格也计入可印刷字符时,则称有 95 个可印刷字符。请注意,这种字符编码中有如下两个规律:

(1) 字符 0~9 这 10 个数字字符的高 3 位编码为 011,低 4 位为 0000~1001。当去掉高 3 位的值时,低 4 位正好是二进制形式的 0~9。这既满足正常的排序关系,又有利于完成 ASCII 码与二进制码之间的转换。

(2) 英文字母的编码值满足正常的字母排序关系,且大、小写英文字母编码的对应关系相当简便,差别仅表现在 b_5 一位的值为 0 或 1,有利于大、小写字母之间的编码变换。

另有一种字符编码,是主要用在 IBM 计算机中的 EBCDIC 代码(Extended Binary Coded Decimal Interchange Code)。它采用 8 位码,有 256 个编码状态,但只选用其中一部分。0~9 十个数字的高 4 位编码为 1111,低 4 位仍为 0000~1001。大、小写英文字母的编码均同样满足正常的排序要求,而且有简单的对应关系,即同一个字母其大小写的编码值仅最高的第二位值不同,易于变换与识别。

2. 字符串的表示

随着计算机在文字处理与信息管理中的广泛应用,字符串已成为最常用的数据类型之一,许多计算机中都提供字符串操作功能,一些计算机还给出读写字符串的机器指令。

字符串是指连续的一串字符,通常方式下,它们占用主存中连续的多个字节,每个字节存一个字符。当主存字由 2 个或 4 个字节组成时,在同一个主存字中,既有按从低位字节向高位字节的顺序存放字符串内容的,也有按从高位字节向低位字节的次序顺序存放字符串内容的。这两种存放方式都是常用方式,不同的计算机可以选用其中任何一种。例如,

1F A>B THEN READ (C) 就可以有两种不同的存放方式(图 2.1):

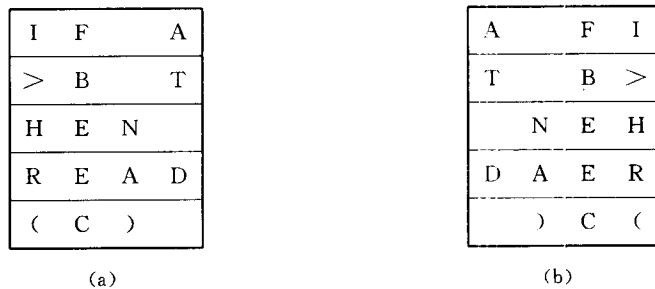


图 2.1 字符串的两种存放方式

假字每个主存字由 4 个字节组成,图(a)是按从高位字节向低位字节的次序存放字符串,图(b)是按从低位字节向高位字节的次序存放字符串。主存中每个字节存的都是相应字符的 ASCII 编码值。例如对图(a)来说,每个字符分别存放的是十进制的 73、70、32、65、62、66、32、84、72、69、78、32、82、69、65、68、40、67 和 41 和 32。

2.2.2 中文的编码表示

我国是个多民族的国家,共有近 60 种民族文字,其中用得最多最广的是汉字,其它还有蒙、藏、朝鲜、僮、苗、哈尼、维吾尔等多种文字。计算机中文处理系统,应尽可能地处理各种中文文字,但其中最关键的技术,是很好地解决汉字的编码方案、存储技术和输入输出方法。有关汉字编码的情况请参阅本书第九章汉字处理设备部分。

2.2.3 逻辑数据的表示

逻辑数据是用来表示二值逻辑中的“是”与“否”、或称“真”与“假”两个状态的数据。很