

第1章

计算机基础知识

本章主要内容

- 📖 计算机的发展与多核处理器。
- 📖 计算机的类型及应用领域。
- 📖 微型计算机系统的主要技术指标。
- 📖 进位计数制的概念及几种进位计数制之间的转换。
- 📖 数据的存储单位（位、字节、字）。
- 📖 ASCII 字符集及其编码，汉字及其编码（国标码）的基本概念，包括 GB2312 编码、BIG5 编码等。

1.1 计算机的发展与多核处理器

1.1.1 计算机的发展概况

1. 计算机发展的几个阶段

(1) 第一代计算机（1946—1957 年）——电子管数字计算机时代

世界上第一台计算机于 1946 年 2 月 15 日在美国诞生，命名为电子数字积分计算机（Electronic Number Integrator And Calculator, ENIAC）。ENIAC 采用的电子元件是电子管（真空管），并使用机器语言编程，主要应用于军事目的和科学研究。

(2) 第二代计算机（1958—1964 年）——晶体管数字计算机时代

随着晶体管的发明，用晶体管取代电子管作为计算机的逻辑元件，使得计算机进入了第二代，即晶体管数字计算机时代。第二代计算机和第一代计算机相比，具有体积小、重量轻、耗电少、运算速度快等特点。由于使用了操作系统，采用高级语言编程，应用领域广泛。

(3) 第三代计算机（1965—1970 年）——中、小规模集成电路数字计算机时代

第三代计算机的电子元件采用了中、小规模集成电路（MSI、SSI），用半导体存储器取

代磁芯储存器, 计算机的体积更小、耗电更少、可靠性更高、功能更强、运算速度更快。

(4) 第四代计算机 (1971 年至今) ——大规模、超大规模集成电路数字计算机时代

随着大规模集成 (LSI) 和超大规模集成 (VLSI) 技术的发展, 在一块半导体芯片上可以集成几千个甚至数十亿个晶体管, 产生了微处理器及微型计算机, 产生了当代的超级计算机。

2. 微处理器及微机的发展阶段

1971 年 Intel 公司推出了由大规模集成电路组成的具有控制器和运算器功能的中央处理器 (Central Processor Unit, CPU), 通常称为微处理器 (Microprocessor, MP)。以微处理器为核心, 配上由大规模集成电路制作的存储器、输入/输出接口电路及系统总线等所组成的计算机, 称为微型计算机 (Microcomputer), 简称微机。

微机的发展取决于微处理器的发展, 发展先后大体上分为如下五代。

(1) 第一代微型计算机 (1971—1973 年)

美国 Intel 公司于 1971 年推出了以 Intel 4004 为微处理器的型号为 MCS-4 的计算机, 它是世界上第一台微型计算机。

(2) 第二代微型计算机 (1974—1978 年)

1974 年, Intel 公司推出了第二代微处理器 Intel 8080, 1975—1976 年相继出现了集成度更高、功能更强的微处理器, 如 Motorola 公司的 6800 和 Zilog 公司的 Z80 等处理器芯片。

(3) 第三代微型计算机 (1978—1984 年)

Intel 公司于 1978 年推出了 16 位的 8086 微处理器, 它属于第三代微处理器, 1981 年, 以 8088 微处理器为核心首次组成了 IBM 微型计算机, 开创了微型计算机的新时代。IBM 是国际商业机器公司 (International Business Machines Corporation) 的简称。

(4) 第四代微型计算机 (1985—1992 年)

随着超大规模集成技术的发展, 出现了 32 位微处理器, 即第四代微处理器, 相应地有 386 微机和 486 微机。

1985 年 Intel 公司推出了 80386 微处理器后, 确定了 80386 芯片的指令集结构 (Instruction Set Architecture) 作为以后开发 80X86 系列处理器的标准, 称其为 Intel 32 位结构 (Intel Architecture-32, IA-32), 后来的 80486、Pentium 等微处理器统称为 IA-32 处理器, 或称 32 位 80X86 处理器。

(5) 第五代微型计算机 (1993—1995 年)

1993 年 3 月, Intel 公司推出了第五代微处理器 Pentium (译名为“奔腾”) 586, 简称 P5, 相应推出了 Pentium 系列微机。

3. 中国计算机发展简况

1958 年, 我国成功研制第一台小型电子管通用计算机 103 机。

1974 年, 清华大学等单位采用集成电路联合研制成功 DJS-130 小型计算机。

1983 年, 研制成功运算速度每秒上亿次的银河-I 巨型机, 并于 1997 年由国防科技大学成功研制银河-III 百亿次并行巨型计算机系统。

2001 年, 中科院计算机所研制成功我国第一款通用 CPU——“龙芯”芯片, 2002 年, 曙光公司推出了采用“龙芯—1”CPU 设计制造的“龙腾”服务器。

2003 年 12 月 9 日, 联想承担的国家网格主节点“深腾 6800”超级计算机正式研制成功, 其实际运算速度达到每秒 4.183 万亿次, 全球排名第 14 位。

2005 年 5 月 1 日, 联想完成并购 IBM PC。联想正式宣布完成对 IBM PC 全球业务的收购,

联想以合并后年收入约 130 亿美元、个人计算机年销售量约 1400 万台的成绩，一跃成为全球第三大 PC 制造商。

2010 年，我国“天河 1 号”落户天津滨海新区的国家超级计算天津中心。2010 年 11 月 17 日，国际超级计算机 TOP500 组织正式发布第 36 届世界超级计算机 500 强排名榜，安装在国家超级计算天津中心的“天河 1 号”超级计算机系统，以峰值速度 4700 万亿次、持续速度 2566 万亿次每秒浮点运算的优异性能位居世界第一，实现了中国自主研发超级计算机综合技术水平进入世界领先行列的历史性突破。

2013 年下半年，TOP500 组织公布了的全球计算机 TOP500 排名，上半年我国研制的“天河 2 号”成功卫冕，蝉联第一。

2016 年 6 月 20 日德国法兰克福国际超级计算机大会（ISC）公布了新一期世界计算机 500 强榜单，我国最新的超级计算机“神威·太湖之光”登顶。“神威·太湖之光”击败的是霸占世界榜首 3 年的“天河 2 号”。

“天河 2 号”采用的核处理器是英特尔 Xeon E5-2692 v2 12，“神威·太湖之光”实现了核心处理器的全国产化。首次采用国产核心处理器“申威 26010”，“申威 26010”大小约 25 平方厘米，集成了 260 个运算核心，内置数十亿个晶体管。每块“申威 26010”运算能力为每秒 3 万多亿次。“神威·太湖之光”运算速度达到 93Petaflop（每秒运算一千万亿次），约为“天河 2 号”的两倍多。

“神威·太湖之光”由 40 个运算机柜和 8 个网络机柜组成，每个柜门内有 4 块超节点，每块超节点由 32 块运算插件组成，每个插件由 4 个运算节点板组成，每个运算节点板又含两块“申威 26010”处理器，整台“神威·太湖之光”共有 40960 块“申威 26010”处理器，叠加之下，“神威·太湖之光”一分钟的计算能力相当于全球 72 亿人口用计算器不间断计算 32 年。

“神威·太湖之光”超级计算机由国家并行计算机工程技术研究中心研制，安装在国家超级计算无锡中心，如图 1-1 所示。



图 1-1 “神威·太湖之光”计算机

1.1.2 多核处理器

多核（Multi-core）技术是将多个处理器核心集成在一个半导体芯片上，各处理器核心耦合紧密，构成一个多处理器（Multiprocessor）系统。换言之，多核处理器就是在单个半导体的一个处理器上拥有多个功能相同的处理器核心，这一多处理器系统中的多个处理器核心能够有效地并行执行多个进程或线程，可以同时共享系统总线、内存等资源。

1. 两种主流多核处理器

2005年4月, Intel 仓促推出简单封装双核的奔腾 D 和奔腾四至尊版 840。AMD 在之后也发布了双核皓龙 (Opteron) 和速龙 (Athlon) 64 X2 处理器, 但真正的“双核元年”则被认为是 2006 年。

现在, 微机 (包括笔记本电脑) 中的微处理器一般是多核处理器。多核处理器是在一个集成电路芯片上制作了两个或多个处理器执行核心的芯片, 其特点是提升了 IA-32 处理器硬件的多线程能力。市场上流行的多核处理器主要有 AMD 羿龙系列和 Intel 公司的 Core i3、Core i5、Core i7 系列, 如图 1-2 和图 1-3 所示。



图 1-2 AMD 羿龙微处理器



图 1-3 Core i7 微处理器

2. 微结构

Intel 基于微结构实现了多核技术, 具有两个物理处理器核心的 Intel Pentium 至尊版处理器是第一个引入多核技术的 IA-32 系列处理器, 每个处理器核心都包含超线程技术, 共计支持 4 个逻辑处理器。在图 1-4 中, 可以看到四个核心的 Nehalem 基本构成, 它有超大容量的 L3 高速缓存、输入/输出 (I/O) 控制单元、Intel 的快速通道互联 QPI (Quick Path Interconnect) 总线及内存控制器电路 (Memory Controller)。不同级别的 Nehalem 处理器将会有不同条数的 QPI 连接, 普通桌面处理器通常只有一条 QPI 连接, 工作站以上级别的多核处理器将会有多条 QPI 连接。

3. Intel 酷睿 (Core) i3、i5、i7 处理器

自从 Intel 推出酷睿系列处理器以来, Core i3、Core i5、Core i7 系列处理器成为家喻户晓的品牌, 如图 1-5 所示。

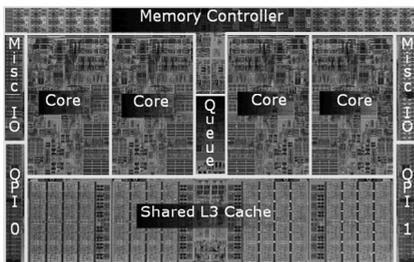


图 1-4 4 个核心的 Nehalem 基本构成

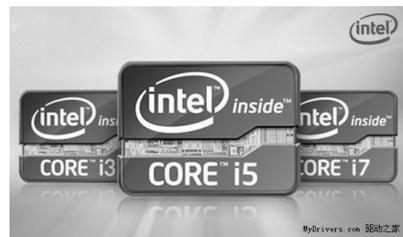


图 1-5 Core i3、Core i5、Core i7 处理器

(1) Intel 酷睿 i3 处理器

Intel 酷睿 i3 是一款基于 Nehalem 架构的双核处理器, 其依旧采用整合内存控制器, 核心线程数是 2 核心 4 线程数, 二级缓存 L2 容量是 $2 \times 256\text{KB}$, 三级缓存 L3 的容量达到 8MB, 功耗 65W。

酷睿 i3 可看作酷睿 i5 的精简版，Core i3 最大的特点是整合 GPU（图形处理器），也就是说 Core i3 将由 CPU+GPU（图形处理器）两个核心封装而成。由于整合的 GPU 性能有限，用户想获得更好的 3D 性能，可以外加显卡。

（2）Intel 酷睿 i5 处理器

Intel 酷睿 i5 核心线程数是 4 核心 4 线程数，二级缓存 L2 容量是 4×256KB，三级缓存 L3 容量达到 8MB，功耗 95W。

Core i3 和 Core i5 的最大区别是，Core i5 支持睿频，Core i3 不支持睿频；Core i3 只有双核，而 Core i5 有双核和 4 核两种。

（3）Intel 酷睿 i7 处理器

Intel 酷睿 i7 核心线程数是 4 核心 8 线程数，功耗大，性能强。

Core i5 和 Core i3 的性能差距比较小，而 Core i7 和 Core i5 的性能差距则比较大。

Core i5 是一款基于 Nehalem 架构的四核处理器，采用整合内存控制器，三级缓存模式，L3 达到 8MB，支持 Turbo Boost 等技术的新处理器。它和 Core i7 的主要区别在于总线不采用 QPI，采用的是成熟的 DMI（Direct Media Interface），并且只支持双通道的 DDR3 内存。结构上它用的是 LGA1156 接口，Core i7 用的是 LGA1366 接口。

（4）Intel 酷睿 i3 系列、i5 系列、i7 系列处理器的型号

随着处理器芯片的发展，其种类繁多，系列处理器还有台式机应用和笔记本电脑应用之分。

酷睿的 i 系列处理器分为几代产品，这与发布时间的先后有关系，这一系列的处理器采用一种流行的级别定位命名方式，即数字越大，代表相应产品价格越贵、性能越好。

例如：

i5-3xxx 属于酷睿 i5 系列三代产品，i5-4xxx 属于酷睿 i5 系列四代产品。

i7-4xxx 属于酷睿 i7 系列四代产品，i7-5xxx 属于酷睿 i7 系列五代产品。

第一代 i3 处理器，型号三位数字，如 i3 380、i3 M380 等。

第二代 i3 处理器，型号为 i3 2XXX 这类，如 i3 2130、i3 2330M 等。

第三代 i3 处理器，型号为 i3 3XXX 这类，如 i3 3330、i3 3340M 等。

第四代 i3 处理器，型号为 i3 4XXX 这类，如 i3 4150、i3 4010U 等。

第五代 i3 处理器，型号为 i3 5XXX 这类，如 i3 5130 等。

i3 后面加三位数字是第一代产品，如 i3 380，这是台式机用的，i3 M380，这是笔记本用的；二代 i3 的命名是 i3 加上 2 开头的四位数字，末尾带 M 的属于笔记本电脑使用的，第三代 i3 与二代 i3 一样，也是 4 位数，只不过是以 3 开头的 4 位数字，如 i3 3340M，最后带 M，属于笔记本电脑使用的，目前市场上已经在销售第六代和第七代 Intel 酷睿微处理器。

1.2 计算机的类型及应用领域

1.2.1 计算机的类型

计算机及相关技术的迅速发展带动了计算机类型的不断分化，形成了各种不同种类的计算机。按用途分，计算机可分为专用计算机和通用计算机。

按字长分，计算机可以分为 8 位机、16 位机、32 位机和 64 位机。

通常按照计算机的运算速度、字长、存储容量等综合性能指标,可将计算机分为超级计算机、大型计算机、小型计算机、微型计算机和 workstation。随着技术的进步,各种型号的计算机性能指标都在不断改进和提高,以致过去一台大型机的性能可能还比不上今天的一台微型计算机。

按照巨型、大型、小型、微型及 workstation 的标准来划分计算机的类型也有其时间的局限性,因此计算机的类别划分很难有一个精确的标准。可以根据计算机的综合性能指标,结合计算机应用领域的分布将其分为 5 大类。

1. 高性能计算机

高性能计算机也就是俗称的超级计算机,或者称巨型机,堪称一流的计算机。例如我国生产的“天河 2 号”和“神威·太湖之光”是世界上顶级的超级计算机。

2. 微型计算机

大规模及超大规模集成电路的发展是微型计算机得以产生的前提。微处理器是微型计算机的核心部件。我们日常使用的台式计算机、笔记本电脑、超薄型笔记本电脑(超级本)等都称为微型计算机。

3. workstation

workstation 是一种高档的微型计算机,通常配有高分辨率的大屏幕显示器及容量很大的内存储器 and 外部存储器,主要面向专业应用领域,具备强大的数据运算与图形、图像处理能力。workstation 是为满足工程设计、动画制作、科学研究、软件开发、金融管理、信息服务、模拟仿真等专业领域而设计开发的同性能微型计算机。

workstation 不同于计算机网络系统中 workstation 的概念,计算机网络系统中的 workstation 仅是网络中的任何一台普通微型机或终端,只是网络中的任一用户节点。

4. 服务器

服务器是指在网络环境中为网上多个用户提供共享信息资源和各种服务的一种高性能计算机,在服务器上需要安装网络操作系统、网络协议和各种网络服务软件。服务器主要为网络用户提供文件、数据库、应用及通信等方面的服务。

5. 嵌入式计算机

嵌入式计算机是指嵌入到对象体系中,实现对象体系智能化控制的专用计算机系统。嵌入式计算机系统以应用为中心,以计算机技术为基础,并且软/硬件可裁剪,适用于应用系统对功能、可靠性、成本、体积、功耗有严格要求的专用计算机系统。它一般由嵌入式微处理器、外围硬件设备、嵌入式操作系统及用户的应用程序 4 部分组成,用于实现对其他设备的控制、监视或管理等功能。例如,日常生活中使用的电冰箱、全自动洗衣机、空调、数码产品等都采用了嵌入式计算机技术。

1.2.2 计算机的应用领域

计算机主要有以下一些应用领域。

1. 科学计算

科学计算也称数值计算,是指利用计算机来完成科学研究和工程技术中提出的数学问题的计算。计算机最早的应用领域是科学计算,在现代科学技术工作中,仍然存在大量的和复杂的科学计算问题。利用计算机的高速计算、大存储容量和连续运算的能力,可以实现人工无法解决的各种科学计算问题。

2. 数据处理

数据处理也称信息处理,是对各种数据进行收集、存储、整理、分类、统计、加工、利用、传播等一系列活动的统称。据统计,80%以上的计算机主要用于数据处理,数据处理从简单到复杂已经历了三个发展阶段。

(1) 电子数据处理(Electronic Data Processing, EDP),它以文件系统为手段,实现一个部门内的单项管理。

(2) 管理信息系统(Management Information System, MIS),它以数据库技术为工具,实现一个部门的全面管理,可以大大提高工作效率。例如,办公自动化(OA)系统、人事档案管理、财务管理软件等都属于计算机信息处理软件。

计算机在管理中的应用首先开始于信息处理。

(3) 决策支持系统(Decision Support System, DSS),它以数据库、模型库和方法库为基础,帮助管理决策者提高决策水平,改善运营策略的正确性与有效性。

目前,数据处理已广泛地应用于办公自动化、企事业计算机辅助管理与决策、情报检索、图书管理、电影电视动画设计、会计电算化等各行各业。

信息正在形成独立的产业,多媒体技术使信息展现在人们面前的不仅是数字和文字,也有声音和图像信息。

3. 辅助技术

计算机辅助技术包括计算机辅助设计(Computer Aided Design, CAD)、计算机辅助制造(Computer Aided Manufacturing, CAM)和计算机辅助教学(Computer Aided Instruction, CAI)等。

(1) 计算机辅助设计

计算机辅助设计是利用计算机系统辅助设计人员进行工程或产品设计,以实现最佳设计效果的一种技术。它已广泛地应用于飞机、汽车、机械、电子、建筑和轻工等领域。

(2) 计算机辅助制造

计算机辅助制造是利用计算机系统对生产设备的管理、控制和操作的过程。例如,在产品的制造过程中,用计算机控制机器的运行,处理生产过程中所需的数据,控制和处理材料的流动及对产品进行检测等。

将CAD和CAM技术集成,实现设计生产自动化,实现无人化工厂,这种技术称为计算机集成制造系统(CIMS)。

4. 计算机辅助教学

计算机辅助教学就是利用计算机系统使用课件来进行教学。

5. 过程控制

过程控制也称实时控制,是利用计算机及时采集数据,按最优值迅速地对控制对象进行自动调节或自动控制。采用计算机进行生产过程控制,不仅可以大大提高控制的自动化水平,而且可以提高控制的及时性和准确性,从而改善劳动条件,提高产品质量及合格率。

6. 人工智能

人工智能(Artificial Intelligence)也称智能模拟,是指利用计算机模拟人类的智能活动,如感知、判断、理解、学习、问题求解和图像识别等。

7. 网络应用

目前,在计算机技术、现代通信技术及其他电子技术相结合的基础上,已经构建并进入

了宽带、高速、综合、广域型数字化的网络时代, 计算机网络的建立, 不仅解决了一个单位、一个地区、一个国家中计算机与计算机之间的通信问题, 各种软/硬件资源的共享也大大促进了国家间的文字、图像、视频和声音等各类数据的传输与处理。

从计算机的应用上看, 计算机将向系统化、网络化和智能化的方向发展, 它将更深入地渗透到社会的各行各业, 推动社会的进一步发展。推动移动互联网、云计算、大数据、物联网等与现代制造业结合, 促进电子商务、工业互联网和互联网金融健康发展。

1.3 微型计算机系统的主要技术指标

早期只用字长、运算速度和存储容量三大指标来衡量和评价计算机的性能, 随着微机技术的不断发展, 现在评价计算机性能的主要指标有下面几项。

1. 字长

字长是 CPU 内部一次能并行处理二进制数码的位数, 字长取决于 CPU 内部寄存器、运算器和数据总线的位数。字长越长, 一个字所能表示数据的精度就越高, 处理速度也越快。为了更灵活地表达和处理信息, 计算机通常以字节 (Byte) 为基本单位, 用大写字母 B 表示, 一个字节等于 8 个二进制位 (bit)。一般机器的字长都是字节的 1 倍、2 倍、4 倍、8 倍等, 目前微型计算机的机器字长有 8 位、16 位、32 位、64 位等几种。

2. 主频

主频即时钟频率, 是指计算机的 CPU 在单位时间内发出的脉冲数目, 它在很大程度上决定了计算机的运行速度。目前随着 CPU 主频的增加, 常用单位由 GHz 取代了 MHz。

在微机的配置中常看到 P4 2.4G 字样, 其中数字 2.4G 就表示处理器的时钟频率是 2.4GHz。

3. 主存容量

主存容量是指一个主存储器所能存储的全部信息量。通常, 我们把以字节数来表示存储容量的计算机称为字节编址的计算机。也有一些计算机是以字为单位编址的, 它们用字数乘以字长来表示容量。主存容量的基本单位是字节, 还可用 KB、MB (兆字节)、GB (吉字节)、TB (太字节) 和 PB (皮字节) 来衡量主存容量。它们之间的关系如表 1-1 所示。

表 1-1 K、M、G、T、P 的定义

单 位	通常比例	实际比例
K (Kilo)	10^3	$2^{10}=1024B=1KB$
M (Mega, 兆)	10^6	$2^{20}=1024KB=1MB$
G (Giga, 吉)	10^9	$2^{30}=1024MB=1GB$
T (Tera, 太)	10^{12}	$2^{40}=1024GB=1TB$
P (Peta, 皮)	10^{15}	$2^{50}=1024TB=1PB$

4. 运算速度

运算速度是一项综合性指标, 它与许多因素有关, 如机器的主频、执行何种操作及主存本身的速度等。对运算速度的衡量有不同的方法, 常用方法有以下几种。

(1) 根据不同类型指令在计算过程中出现的频繁程度, 乘上不同的系数, 求出统计平均值, 这时所指的运算速度是平均运算速度。

(2) 以每条指令执行所需的时钟周期数 (Cycles Per Instruction, CPI) 来衡量。

(3) 以 MIPS 作为计量单位来衡量运算速度。MIPS (Million Instruction Per Second) 表示每秒执行几百万条指令。

5. 外存储器容量

微型计算机一般配有硬盘驱动器和光盘驱动器, 分别可以驱动硬盘和光盘; 还配有 USB 接口, 可以外接 U 盘。这些外存储器中, 硬盘容量最大, 决定了微机能存放系统软件和应用软件的多少。

6. 处理器综合参数

微处理器 (CPU) 性能的高低是影响计算机系统的重要因素, 微处理器档次越高, 计算机系统的性能越好。例如 CPU 主频高, 一般计算机运行速度就快, 核心数量多、线程数多, 计算机数据处理能力增强, 运行速度加快。

例如, Intel 酷睿 i5 4690 采用 22 纳米 (nm) 工艺制程, 采用了 LGA 1150 处理器插槽。i5 4690 原生内置四核心, 四线程, 处理器默认主频高达 3.5GHz, 最高睿频可达 3.9GHz。二级缓存为 1MB, 同时三级高速缓存容量高达 6MB, 这就使得 CPU 在处理数据时提高了命中率, 并且使软件加载时间大大缩短。

Intel 酷睿 i5 4690 处理器的综合参数如下。

包装形式: 盒装。

CPU 主频: 3.5GHz。

插槽类型: LGA 1150。

针脚数目: 1150pin。

核心代号: Haswell。

CPU 架构: Haswell。

核心数量: 四核心。

线程数: 四线程。

制作工艺: 22nm。

热设计功耗 (TDP): 84W。

二级缓存: 1MB。

三级缓存: 6MB。

内存控制器: DDR3 1333/1600MHz。

虚拟化技术: Intel VT。

处理器字长: 64 位。

集成显卡。

显卡基本频率: 350~1100MHz。

1.4 计算机中信息的表示

在应用科学技术领域, 计算机是处理信息的工具。各种形式的信息, 如数字、文字、声音、图像、温度和压力等, 都要转换成为计算机能识别的符号。信息的符号化就是数据, 是计算机所能识别的数据。在现代计算机系统中, 所指的数据都是二进制形式的。

二进制数只需要两个数字符号 (0 和 1) 来组成。在计算机的逻辑电路中, 用两个不同的电信号, 即低电平 (代表逻辑 0) 和高电平 (代表逻辑 1), 就可以方便地表示二进制数, 只需要制造有两个状态的电子器件来表示二进制数中的两个数字符号即可, 物理上容易实现, 且简单可靠, 十分方便。

1.4.1 几种进位计数制

由于阅读和书写二进制数很不方便, 因此在书写 (编程) 和计算机输入、输出时通常使用十进制数十六进制数或八进制数。计算机通过软件及输入/输出接口可将十进制数、八进制数、十六进制数转换成计算机能够接受的二进制数, 并且能将二进制数转换成八进制数、十进制数及十六进制数输出。

1. 十进制数

十进制计数有两个特点: 第一, 由十个数字符号 (0、1、2、3、4、5、6、7、8、9) 构成; 第二, 逢十进一。十进制数可以用位权来表示, 位权就是在一个数中同一个数字在不同的位置上代表不同基数的次幂, 任何一个十进制数都可以用它的按位权展开式来表示。

$$N = X_{n-1} \times 10^{n-1} + X_{n-2} \times 10^{n-2} + \dots + X_0 \times 10^0 + X_{-1} \times 10^{-1} + \dots + X_{-m} \times 10^{-m}$$

其中, X 为一个 10 进制的数, 基数是 10, 整数位有 n ($n-1 \sim 0$) 位, 小数位是 m ($-1 \sim -m$) 位。例如:

$$(128.6)_{10} = 1 \times 10^2 + 2 \times 10^1 + 8 \times 10^0 + 6 \times 10^{-1}$$

2. 二进制数

二进制计数也有两个特点: 第一, 由两个数字符号 (0 和 1) 构成; 第二, 逢二进一。二进制数也可以用位权来表示, 任何一个二进制数都可以用它的按位权展开式来表示。

$$N = X_{n-1} \times 2^{n-1} + X_{n-2} \times 2^{n-2} + \dots + X_0 \times 2^0 + X_{-1} \times 2^{-1} + \dots + X_{-m} \times 2^{-m}$$

其中, X 为一个二进制的数, 基数是 2, 整数位有 n ($n-1 \sim 0$) 位, 小数位是 m ($-1 \sim -m$) 位。例如:

$$\begin{aligned} (1011.101)_2 &= 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} \\ &= 8 + 0 + 2 + 1 + 0.5 + 0.00 + 0.125 = (11.625)_{10} \end{aligned}$$

3. 八进制数

八进制数有 0、1、2、3、4、5、6、7 共计 8 个数字符号, 逢八进一, 八进制数也遵守按位权展开的原理, 只不过其中的基数不是 10, 也不是 2, 而是 8。例如:

$$(127.4)_8 = 1 \times 8^2 + 2 \times 8^1 + 7 \times 8^0 + 4 \times 8^{-1} = 64 + 16 + 7 + 0.5 = (87.5)_{10}$$

4. 十六进制数

十六进制数有 0、1、2、3、4、5、6、7、8、9、A、B、C、D、E、F 共计 16 个符号, 逢十六进一, 十六进制数也遵守按位权展开的原理, 只不过其中的基数是 16。例如:

$$(A2C.8)_8 = 10 \times 16^2 + 2 \times 16^1 + 12 \times 16^0 + 8 \times 16^{-1} = 10 \times 256 + 2 \times 16 + 12 + 0.5 = (2604.5)_{10}$$

四种进位制的对照表如表 1-2 所示。

表 1-2 四种进位制的对照表

十进制	二进制	八进制	十六进制
0	0	0	0
1	1	1	1

续表

十进制	二进制	八进制	十六进制
2	10	2	2
3	11	3	3
4	100	4	4
5	101	5	5
6	110	6	6
7	111	7	7
8	1000	10	8
9	1001	11	9
10	1010	12	A
11	1011	13	B
12	1100	14	C
13	1101	15	D
14	1110	16	E
15	1111	17	F
16	10000	20	10

不同进位制的习惯书写单位如下。

- 二进制：Binary、Bin 或 B。
- 八进制：Octal、Oct 或 O。
- 十进制：Decimal、Dec 或 D。
- 十六进制：Hex、Hex 或 H。

在编辑环境下，向计算机输入不同进位制数时，可以分别使用简化符号表示，在十进制数、二进制数、八进制数及十六进制数的末尾分别输入一个大写字母 D、B、O 和 H。

1.4.2 进位计数制之间的相互转换

只要掌握了将十进制数转换成二进制数的方法，就可以按类似的方式将十进制数转换成其他进制数。

将十进制数转换成其他进制的数时，可以先将十进制数转换成二进制数，然后由二进制数转换为八进制数和十六进制数，这种方法经常使用。

将二进制数、八进制数及十六进制数转换成十进制数时，按照按位的权值展开并求和就可以了。

1. 十进制数转换成二进制数

十进制数转换成二进制数包括将十进制整数部分转换成二进制整数及将十进制小数部分转换成二进制小数。

可以采用“除 2 取余法”将十进制整数转换成二进制整数。

具体方法是：通过列竖式将十进制数连续除以 2，将每次除以 2 得到的商和余数分别记录下来，直到商等于 0 为止，所得的余数就是转换成的二进制数，最先得到的一位余数

是转换成二进制数的最低位，按照顺序排列，最后得到的一位余数是转换成的二进制数的最高位。

【例 1-1】 用列竖式的方法把十进制整数 13 转换成二进制数。

解：转换过程如图 1-6 所示。

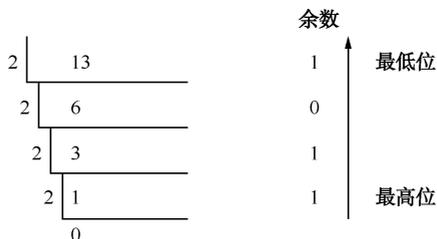


图 1-6 转换过程

转换结果： $(13)_{10}=(1101)_2$ 。

可以采用“乘 2 取整法”将十进制小数转换成二进制小数，具体过程为：用 2 乘以十进制小数，每次得到的一位整数就是转换成的一位二进制小数位，然后，再用 2 去乘以上次乘以 2 后所得结果的小数部分，又可以得到新的一位整数作为转换成的较低一位的二进制小数位。将每次得到的整数部分按先后顺序依次排列，就得到相对应的二进制小数。值得注意的是，先得到的是高位，后得到的是低位。

【例 1-2】 将十进制数 0.8125 转换成二进制小数。

解：

$0.8125 \times 2 = 1.625$	1
$0.625 \times 2 = 1.25$	1
$0.25 \times 2 = 0.5$	0
$0.5 \times 2 = 1.0$	1

转换结果： $(0.8125)_{10}=(0.1101)_2$ 。

以上例子正好转换完成，没有余数，如果不可能全部转换完成，只要保留二进制小数后面几位就可以了。

2. 二进制数转换成八进制数

将二进制数转换成八进制数时，将二进制整数转换成八进制整数并将二进制小数转换成八进制小数。

二进制整数转换成八进制整数的方法是：将二进制整数从右到左，每 3 位一组作为一位八进制数，如果最左边不足 3 位则补 0，同样构成 3 位一组，作为一位八进制数。

二进制小数转换成八进制小数的方法是：将二进制小数从左到右，每 3 位一组作为一位八进制数，如果最右边不足 3 位则补 0，同样构成 3 位一组，作为一位八进制数。

【例 1-3】 将二进制数 $(1110101001100.1100101)_2$ 转换成八进制数。

解：

$$(1110101001100.1100101)_2 = (001\ 110\ 101\ 001\ 100.110\ 010\ 100)_2$$

$$= (16514.624)_8$$

3. 八进制数转换成二进制数

八进制数转换成二进制数与二进制数转换成八进制数是互逆的过程，但是转换方法是一致的。具体方法是：不管是八进制数的整数位还是小数位，只需要将每一位八进制数用 3 位二进制

数表示就可以了。

【例 1-4】 将八进制数 $(123.456)_8$ 转换成二进制数。

解: $(123.456)_8=(1010011.100101110)_2$

4. 二进制数转换成十六进制数

将二进制数转换成十六进制数时,将二进制整数转换成十六进制整数并将二进制小数转换成十六进制小数。

二进制整数转换成十六进制整数的方法是:将二进制整数从右到左,每4位为一组作为一位十六进制数,如果最左边不足4位则补0,同样构成4位一组,作为一位十六进制数。

二进制小数转换成十六进制小数的方法是:将二进制小数从左到右,每4位为一组作为一位十六进制数,如果最右边不足4位则补0,同样构成4位一组,作为一位十六进制数。

【例 1-5】 将二进制数 $(1110101001100.1100101)_2$ 转换成十六进制数。

解: $(1110101001100.1100101)_2=(0001\ 1101\ 0100\ 1100.1100\ 1010)_2$
 $= (1D4C.CA)_{16}$

5. 十六进制数转换成二进制数

十六进制数转换成二进制数与二进制数转换成十六进制数是互逆的过程,与八进制数转换成二进制数的方法是类似的。具体方法是:不管是十六进制数的整数位还是小数位,将每一位十六进制数用4位二进制数表示就可以了。

【例 1-6】 将十六进制数 $(B2A3.4D)_{16}$ 转换成二进制数。

解: $(B2A3.4D)_{16}=(1011001010100011.01001101)_2$

值得注意的有两点:第一,将二进制数、八进制数及十六进制数转换成十进制数的方法是:按位权展开的方法求其对应的十进制数。第二,八进制数和十六进制数之间的转换都可以先将其转换成二进制数,即借助二进制数再转换成需要的进制数。

1.4.3 计算机中数的表示方法

计算机中的数分为定点数和浮点数两种,相应地有定点运算部件和浮点运算部件,计算机中既可以实现定点数的运算,又可以实现浮点数的运算。

按二进制表示数的方式分,二进制数可分为无符号数与带符号数,带符号的数也称为机器数。

机器数可以区分正数和负数,常用的机器数有原码、反码及补码三种形式,微机中一般选用补码来进行运算和存储。

例如,8位无符号二进制数表示数的范围是 $00000000B\sim 11111111B$,即 $0\sim 255$ 。而8位带符号数表示数的范围则根据原码、反码及补码的不同,其数的范围也不同。

1. 定点数的表示法

在计算机中,约定数据小数点的位置固定在某一位的表示方法即为定点数表示法。原理上讲,小数点的位置固定在哪一位都行,但是通常有两种定点格式,一是将小数点固定在数的最左边(纯小数),二是固定在数的最右边(纯整数)。

假如用宽度为 $n+1$ 位的字来表示定点数 X ,其中 X_0 表示数的符号,一般用1代表负数,0代表正数,其余位代表它的数位,对于任意定点数 $X=X_0X_1X_2\cdots X_n$,在定点计算机中可表示为下列两种。

(1) 如果 X 为纯小数, 小数点固定在 X_0 与 X_1 之间, 则数 X 的表示范围如下。

$$0 \leq |X| \leq 1 - 2^{-n} \quad (1-1)$$

例如, $X=0.1111111$, 表示是一个正的纯小数, 它是二进制形式的, 其十进制数值可以写为 $X=1-2^{-7}$ 。

(2) 如果 X 为纯整数, 小数点固定在 X_n 的右边, 数 X 的表示范围如下。

$$0 \leq |X| \leq 2^n - 1 \quad (1-2)$$

例如, $X=01111111$, 表示是一个正的纯整数, 它是二进制形式的, 其十进制数值可以写为 $X=2^7-1$ 。

2. 浮点数的表示法

任意一个十进制数 N 都可以写成:

$$N=10^E \times M \quad (1-3)$$

同样任意一个二进制数 N 可以写成:

$$N=2^e \times m \quad (1-4)$$

例如, $N=-101.1101=2^{0011} \times -0.1011101$, 其中, m 为浮点数的尾数, 是一个纯小数; e 是比例因子的指数, 称为浮点数的指数, 是一个纯整数; 比例因子的基数是一个常数, 这里取值为 2。

可以看出, 在计算机中存放一个完整的浮点数, 应该包括阶码、阶符、尾数及尾数的符号(数符)共 4 部分, 如表 1-3 所示。

表 1-3 完整的浮点数的组成

E_s	$E_1 E_2 \cdots E_m$	M_s	$M_1 M_2 \cdots M_n$
阶符 (0)	阶码 (011)	数符 (1)	尾数 (1011101)

一般阶符与数符位都只占一位, 且 0 代表正数, 1 代表负数。设阶码占 3 位, 尾数占 7 位, 根据表 1-3 中的数据, 则数 $N=2^{0011} \times -0.1011101$ 。

值得注意的是, 在计算机相同字长的条件下, 浮点数所表示数的范围比定点数所表示数的范围大得多。

Pentium 系列微机的微处理器中包含了改进后的浮点运算部件, 按照 IEEE 754 标准采用了 32 位浮点数和 64 位浮点数两种标准格式。

3. 机器数与真值

计算机中传输与加工处理的信息均为二进制数, 二进制数的逻辑 1 和逻辑 0 分别用于代表高电平和低电平。计算机只能识别 1 和 0 两种状态, 那么如何确定与识别正二进制数和负二进制数呢? 解决的办法是将二进制数最高位作为符号位, 例如, 1 表示负数, 0 表示正数。若字长取 8 位, 10001111B 则可以代表-15, 00001111B 则可以代表+15, 这便构成了计算机能识别的数, 因此, 带符号的二进制数称为机器数, 机器数所代表的值称为真值。在微机中, 机器数有 3 种表示法, 即原码表示法、反码表示法与补码表示法。

4. 原码表示法

若定点整数的原码形式为 $X_0 X_1 X_2 \cdots X_n$, 则原码表示的定义如下。

$$[X]_{\text{原}} = \begin{cases} X & 2^n > X \geq 0 \\ 2^n - X = 2^n + |X| & 0 \geq X > -2^n \end{cases} \quad (1-5)$$

X_0 为符号位, 若 $n=7$, 即计算机的字长是 8 位, 则:

(1) X 取值范围为 $-127 \sim +127$ 。

(2) $[+0]_{\text{原}}=00000000$ 。

(3) $[-0]_{\text{原}}=10000000$ 。

(4) 从原码的定义可以看出, 一个正数的原码等于这个正数的表达形式, 最高位的符号位为 0, 表示正数; 而一个负数的原码等于这个负数的绝对值, 最高位的符号位为 1, 表示负数。如果用得出的这个简便方法求原码, 则十分方便。

例如: $X=01001111$ $[X]_{\text{原}}=01001111$
 $Y=-01001111$ $[Y]_{\text{原}}=11001111$

采用原码表示法简单易懂, 但它最大缺点是加法运算的电路复杂, 不容易实现。

5. 反码表示法

对于定点整数, 反码表示的定义如下。

$$[X]_{\text{反}} = \begin{cases} X & 2^n > X \geq 0 \\ (2^n - 1)X & 0 \geq X > -2^n \end{cases} \quad (1-6)$$

同样 n 取 7, 即字长 8 位, 则:

(1) X 取值范围为 $-127 \sim +127$ 。

(2) $[+0]_{\text{反}}=00000000$ 。

(3) $[-0]_{\text{反}}=11111111$ 。

(4) 从反码的定义可以看出, 一个正数的反码等于这个正数的表达形式, 最高位的符号位为 0, 表示正数; 而一个负数的反码等于这个负数每位取反, 最高位的符号位为 1, 表示负数。

例如: $X=01001111$ $[X]_{\text{反}}=01001111$
 $Y=-01001111$ $[Y]_{\text{反}}=10110000$

6. 补码表示法

对于定点整数, 补码表示的定义如下。

$$[X]_{\text{补}} = \begin{cases} X & 2^n > X \geq 0 \\ 2^{n+1} + X = 2^{n+1} - |X| & 0 \geq X > -2^n \end{cases} \quad (1-7)$$

同样如果 n 取 7, 即字长 8 位, 则:

(1) X 取值范围为 $-128 \sim +127$ 。

(2) $[+0]_{\text{补}}=[-0]_{\text{补}}=00000000$ 。

(3) $[-10000000]_{\text{补}}=10000000$ 。

(4) 从补码的定义可以看出, 一个正数的补码等于这个正数的表达形式, 最高位的符号位为 0, 表示正数; 而一个负数的补码等于这个负数每位取反, 最低位还要加上 1, 最高位的符号位也为 1, 表示负数。

例如: $X=01001111$ $[X]_{\text{补}}=01001111$
 $Y=-01001111$ $[Y]_{\text{补}}=10110001$

综上所述, 一个正数的原码、反码及补码与其真值的表达形式相同, 只是最高位定义为符号位, 用 0 表示; 一个负数的原码、反码及补码的符号位均为 1, 如果是原码, 则其余的位与真值相同, 如果是反码, 则其余的位是真值各位求反所得, 如果是补码, 则在反码的基础上再加 1 便可求得。

【例 1-7】 设机器字长 8 位, $X=00011111\text{B}$, $Y=-00011111$, 用上面总结的方法求 X 和 Y 的原码、反码及补码。

解:

$$[X]_{\text{原}}=[X]_{\text{反}}=[X]_{\text{补}}=00011111\text{B}$$

$$[Y]_{\text{原}}=10011111\text{B}, [Y]_{\text{反}}=11100000\text{B}, [Y]_{\text{补}}=11100001\text{B}$$

1.4.4 计算机中常用的字符编码与汉字编码

计算机中存储的信息都是用二进制数表示的, 而人们在屏幕上看到的英文、汉字等字符是由二进制数转换之后的结果。把不同字符按照一定编排规则编排之后存储在计算机中, 称为“编码”; 反之, 将存储在计算机中的二进制数解析并显示出来, 称为“解码”。

1. 字符集

字符是各种文字和符号的总称, 包括许多国家的文字、标点符号、图形符号和数字等。字符集是指一个系统支持的所有抽象字符的集合。

常见的字符集有 ASCII 字符集、EASCII 字符集、GB 2312 字符集、BIG5 字符集、GB 18030 字符集、Unicode 字符集等。计算机要准确地处理各种字符集文字, 需要进行字符编码, 以便计算机能够识别和存储各种文字。

2. 字符编码

计算机中处理的信息不全是数值, 有时还需要对各种字符或符号进行处理。例如从键盘输入信息或打印输出信息, 这些都要使用字符方式进行输入/输出。因此计算机还必须能够表示字符和符号, 这需要将自然语言的字符转换为计算机可以接受的数字信息, 称为字符代码。各种字符和符号由于表示的内容不同, 编码的规则和编码的长度也不尽相同。

不同的字符集有其相应的字符编码规则, 常用的字符编码有以下 3 种:

- ASCII 码 (包括 EASCII 码);
- 汉字编码 (GB 2312 编码、BIG5 编码);
- Unicode 编码。

3. ASCII 字符集及其编码

ASCII (American Standard Code for Information Interchange, 美国信息交换标准代码) 是基于拉丁字母的一套计算机编码系统, 它主要用于显示现代英语, 是现今最通用的单字节编码系统。

(1) ASCII 字符集

ASCII 字符集主要包括控制字符和可显示字符 (英文大小写字符、阿拉伯数字和西文字符)。

常用的字符包括以下一些。

字母: A、B、…、Z, a、b、…、z。

数字: 0、1、2、3、4、5、6、7、8、9。

专用符号: +、-、*、/、_等。

控制字符 (非打印字符): BEL (响铃)、LF (换行)、CR (回车) 等。

(2) ASCII 编码

ASCII 编码使用 7 位 (bit) 表示一个字符, 即 ASCII 的范围是 00H~7FH, 最高位为 0。7 位编码的字符集只能支持 128 个字符, 如表 1-4 所示。

ASCII 字符集包括 10 个十进制数码 (0~9), 其对应的 ASCII 码是 30H~39H (48~57)。

它还包括 26 个大写的英文字母，大写“A”的 ASCII 码是 41H (65)，大写“B”的 ASCII 码是 42H (66)，即从 41H 开始按英文字母的顺序递增。而小写 26 个英文字母的 ASCII 码则以 61H (“a”的 ASCII 码) 开始，其余也按顺序递增，小写“z”的 ASCII 码是 7AH。

表 1-4 ASCII 码编码表

低位 LSB		高位 MSB							
		0	1	2	3	4	5	6	7
		000	001	010	011	100	101	110	111
0	0000	NUL	DLE	SP	0	@	P	,	p
1	0001	SOH	DC1	!	1	A	Q	a	q
2	0010	STX	DC2	"	2	B	R	b	r
3	0011	ETX	DC3	#	3	C	S	c	s
4	0100	EOT	DC4	\$	4	D	T	d	t
5	0101	ENQ	NAK	%	5	E	U	e	u
6	0110	ACK	SYN	&	6	F	V	f	v
7	0111	BEL	ETB	,	7	G	W	g	w
8	1000	BS	CAN	(8	H	X	h	x
9	1001	HT	EM)	9	I	Y	i	y
A	1010	LF	SUB	*	:	J	Z	j	z
B	1011	VT	ESC	+	;	K	[k	{
C	1100	FF	FS	'	<	L	\	l	
D	1101	CR	GS	-	=	M]	m	}
E	1110	SO	RS	•	>	N	↑	n	~
F	1111	SI	US	/	?	O	←	o	DEL

在 128 个 ASCII 编码中，有 32 个 ASCII 编码值为 0~31 (00H~1FH)，其不对应任何显示与打印实际字符，它们被用作控制码，控制计算机 I/O 设备的操作及计算机软件的执行，它们代表的含义如表 1-5 所示。

表 1-5 部分 ASCII 编码所代表的含义

编码	含 义	编码	含 义	编码	含 义	编码	含 义
NUL	空	BS	退一格	DLE	数据链换码	CAN	取消
SOH	标题开始	HT	横向列表 (穿孔卡片指令)	DC1	设备控制 1	EM	纸尽
STX	正文开始	LF	换行	DC2	设备控制 2	SUB	取代
ETX	正文结束	VT	垂直制表	DC3	设备控制 3	ESC	换码
EOT	传输结束	FF	走纸控制(换页)	DC4	设备控制 4	FS	文件分隔符
ENQ	询问	CR	回车	NAK	否定应答	GS	组分分隔符
ACK	确认	SO	移位输出	SYN	空转同步	RS	记录分隔符
BEL	响铃	SI	移位输入	ETB	信息组传送结束	US	单元分隔符

在计算机系统的字节传送过程中，往往将最高位用作奇、偶校验位。当最高位恒取 1 时，称为标记校验；当最高位恒取 0 时，称为空格校验；当连同最高位共计 8 位中 1 的个数为奇数个时，则为奇校验；当连同最高位共计 8 位中 1 的个数为偶数个时，则为偶校验。

(3) EASCII 字符集

为了表示更多的欧洲常用字符，对 ASCII 进行了扩展，即形成了 ASCII 扩展字符集 EASCII。

EASCII 使用 8 位 (bit) 表示一个字符, 共 256 字符, 即扩充了 128 个新字符, 如表 1-6 所示。

表 1-6 扩展 ASCII 编码表

高四位 低四位		扩展 ASCII 码字符集															
		1000		1001		1010		1011		1100		1101		1110		1111	
		8		9		A/10		B/16		C/32		D/48		E/64		F/80	
		十进制	字符	十进制	字符	十进制	字符	十进制	字符	十进制	字符	十进制	字符	十进制	字符	十进制	字符
0000	0	128	Ç	144	É	160	á	176	☼	192	⌞	208	⌞	224	α	240	≡
0001	1	129	ü	145	æ	161	í	177	☼	193	⌞	209	⌞	225	β	241	±
0010	2	130	é	146	Æ	162	ó	178	☼	194	⌞	210	⌞	226	Γ	242	≥
0011	3	131	â	147	ô	163	ú	179		195	⌞	211	⌞	227	π	243	≤
0100	4	132	ä	148	ö	164	ñ	180	⌞	196	—	212	⌞	228	Σ	244	∫
0101	5	133	à	149	ò	165	Ñ	181	⌞	197	+	213	⌞	229	σ	245	∫
0110	6	134	â	150	û	166	ª	182	⌞	198	⌞	214	⌞	230	μ	246	÷
0111	7	135	ç	151	ù	167	º	183	⌞	199	⌞	215	+	231	τ	247	≈
1000	8	136	ê	152	ÿ	168	¿	184	⌞	200	⌞	216	+	232	Φ	248	°
1001	9	137	ë	153	ÿ	169	¬	185	⌞	201	⌞	217	⌞	233	Θ	249	•
1010	A	138	è	154	ÿ	170	¬	186		202	⌞	218	⌞	234	Ω	250	·
1011	B	139	ï	155	ç	171	½	187	⌞	203	⌞	219	■	235	δ	251	√
1100	C	140	î	156	£	172	¼	188	⌞	204	⌞	220	■	236	∞	252	n
1101	D	141	ì	157	¥	173	¡	189	⌞	205	—	221	■	237	φ	253	²
1110	E	142	Ä	158	⌞	174	«	190	⌞	206	+	222	■	238	ε	254	■
1111	F	143	Å	159	f	175	»	191	⌞	207	⌞	223	■	239	∩	255	BLANK FF

注: 表中的 ASCII 字符可以用 ALT+“小键盘上的数字键”输入

ASCII 码从显示功能上看, 最大缺点是只能显示 26 个基本拉丁字母、阿拉伯数字和英式标点符号, 因此只能用于显示现代美国英语。在处理英语当中的外来词如 naïve、café、élite 等时, 所有重音符号都不得不去掉, 这样做违反了拼写规则。

EASCII 虽然解决了部分西欧语言的显示问题, 但对更多其他语言依然无能为力。因此现在的苹果电脑已经抛弃 ASCII 而转用 Unicode。

4. 汉字及其编码

(1) 汉字输入和输出的过程

汉字是象形文字, 如何在计算机中实现对汉字的输入和输出呢? 其中包含比较复杂的转换过程。简单地说, 计算机借助二进制数对汉字进行编码、存储、转换、查表等操作, 并由计算机输入/输出设备进行传输。

汉字输入和输出的过程如图 1-7 所示。用户借助汉字的输入码, 如汉字拼音输入码、五笔字型输入码等, 通过键盘输入, 或通过电脑手写输入板输入等, 产生汉字编码, 并输入到计算机内; 由计算机的转换程序找到其对应的机内(编)码, 每个汉字由 2 字节的机内码组成, 机内码存储在计算机的内存储器中; 需要输出时, 再由汉字字形检索程序根据机内码在汉字字形点阵库中找