



近年来,计算机及其应用技术得到了迅猛的发展,已渗透到生产、科研、教学、企业管理乃至家庭用户等各个领域。计算机应用技术的高速发展也极大地促进了信息技术革命的到来,使社会步入信息时代。信息获取、分析处理、传递交流和开发应用是现代入必须具备的基本素质。

1.1 计算机的产生与发展

在人类文明发展的历史长河中,计算工具经历了从简单到复杂、从低级到高级的发展过程,如算筹、算盘、计算尺、手摇机械计算机、电动机械计算机、电子计算机等,它们在不同的历史时期发挥了各自的作用,而且也孕育了电子计算机的设计思想和雏形。计算机技术是信息技术的基础,在人类生活中起着极其重要的作用。

1.1.1 计算机的产生与发展过程

1. 计算机的产生

第二次世界大战期间,美国宾夕法尼亚大学的 John Mauchly (莫希利) 博士和他的学生 J. Presper Eckert 应美国军方的要求构思和设计了 ENIAC (Electronic Numerical Integrator And Calculator, 埃尼阿克), 它于 1946 年 2 月 15 日完成, 为美国陆军的弹道研究实验室 (BRL) 所使用, 用于计算火炮的火力表, 研制和开发新型大炮和导弹。这台计算机共用了 18 000 个电子管, 1 500 多个继电器, 质量达 30 t, 占地 170 m², 耗电 150 kW, 运算速度为每秒 5 000 次加、减运算。

电子计算机的问世, 最重要的奠基人是英国科学家艾兰·图灵 (Alan Turing) 和美籍匈牙利科学家冯·诺依曼 (John von Neuman)。图灵的贡献是建立了图灵机的理论模型, 奠定了人工智能的基础。而冯·诺依曼则是首先提出了计算机体系结构的设想。

冯·诺依曼理论的要点是: 数字计算机的数制采用二进制, 计算机应该按照程序顺序执行。人们把冯·诺依曼的这个理论称为冯·诺依曼体系结构, 包含以下三个要点。

- (1) 计算机由运算器、控制器、存储器、输入设备、输出设备五大基本部件组成。
- (2) 程序和数据均存放在存储器中, 并能自动依次执行指令。
- (3) 所有的数据和程序均用二进制的 0、1 代码表示。

半个多世纪以来, 计算机制造技术发生了巨大变化, 但冯·诺依曼体系结构仍然沿用至今, 人们总是把冯·诺依曼称为“计算机鼻祖”、“现代电子计算机之父”。

2. 计算机的发展

随着电子技术的不断发展, 计算机先后以电子管、晶体管、集成电路、大规模和超大规模集成电

路为主要元器件，共经历了四代变革，如表 1.1 所示。每一代的变革在技术上都是一次新的突破，在性能上都是一次质的飞跃。

表 1.1 各代计算机主要特点比较

| 代别 | 起止年份 | 硬件特征 | 软件发展状况 | 应用领域 |
|-----|-----------|--------------|---|----------------------------|
| 第一代 | 1946—1958 | 电子管 | 机器语言和汇编语言 | 科学计算 |
| 第二代 | 1959—1964 | 晶体管 | 高级语言（编译程序），简单的操作系统 | 科学计算、数据处理和事务管理 |
| 第三代 | 1965—1970 | 小规模集成电路 | 功能较强的操作系统，高级语言，结构化，模块化的程序设计 | 科学计算、数据处理、事务管理和过程控制 |
| 第四代 | 1971 年至今 | 大规模、超大规模集成电路 | 操作系统进一步完善，数据库系统、网络软件得到发展，软件工程标准化，面向对象的软件设计方法与技术广泛应用 | 网络分布式计算、人工智能，迅速推广并普及到社会各领域 |

目前使用的计算机都属于第四代计算机。从 20 世纪 80 年代开始，发达国家开始研制第五代计算机，研究的目的是能够打破以往计算机固有的体系结构，使计算机能够具有像人一样的思维、推理和判断能力，向智能化发展，实现接近人的思考方式。

1956 年，周恩来总理亲自主持制定了我国《12 年科学技术发展规划》，选定了“计算机、电子学、半导体、自动化”作为“发展规划”的四项内容，并制定了计算机科研、生产、教育发展计划，我国由此开始了计算机研制的起步。

1958 年，中国研制出第一台电子计算机。

1964 年，中国研制出第二代晶体管计算机。

1971 年，中国研制出第三代集成电路计算机。

1977 年，中国研制出第一台微型计算机 DJS-050。

1983 年，中国研制成功“深腾 1800”计算机，运算速度超过 1 万次/s，同时“银河 1 号”巨型计算机研制成功，运算速度达 1 亿次/s。

2010 年 10 月，中国国防科技大学研制出“天河一号”，中国首台千万亿次超级计算机。

1.1.2 计算机的发展趋势

1. 电子计算机的发展方向

目前，科学家们正在使计算机朝着巨型化、微型化、网络化、智能化和多功能化的方向发展。巨型机的研制、开发和利用，代表着一个国家的经济实力和科学水平；微型机的研制、开发和广泛应用，则标志着一个国家科学普及的程度。

（1）向巨型化和微型化两极方向发展。巨型化是指要研制运算速度极高、存储容量极大、整体功能极强，以及外设完备的计算机系统（巨型机），巨型机主要用于尖端科学技术及军事、国防系统；而微型化是随着大规模集成电路技术的不断发展和微处理器芯片的产生，以及进一步扩大计算机的应用领域而研制的高性价比的通用微型计算机，这种微型机操作简单，使用方便，所配软件丰富。

（2）智能化是未来计算机发展的总趋势。智能化就是要求计算机能够模拟人的逻辑思维功能和感官，能够自动识别文本、声音、图形、图像等多媒体信息，具有逻辑推理和判断功能。其中最具代表性的领域是专家系统和智能机器人。

（3）非冯·诺依曼体系结构是提高现代计算机性能的另一个研究焦点。冯·诺依曼型计算机工作原理的核心是存储程序和程序控制，整个计算机的工作都是在程序设计人员设计的程序控制下工作的，计算机不具备智能功能。因此，要想真正实现计算机的智能化，就必须打破目前的冯·诺依曼体系结构，研制新型的非冯·诺依曼型计算机。

(4) 多媒体计算机仍然是计算机研究和开发的热点。多媒体技术是集文字、声音、图形、图像和计算机于一体的综合技术,它以计算机技术为基础,包括数字化信息技术、音/视频技术、图像技术、通信技术、人工智能技术、模式识别技术等,是一门多学科、多领域的高新技术。多媒体技术虽然已经取得很大的发展,但高质量的多媒体设备和相关技术仍需要进一步研制,主要包括视频数据的压缩、解压缩技术,多媒体数据的通信,以及各种接口的实现方案等。因此,多媒体计算机仍然是计算机研究和开发的热点。

(5) 网络化是今后计算机应用的主流。计算机网络技术是在计算机技术和通信技术的基础上发展起来的一种新型技术。目前世界上最大的计算机网络就是被广大用户所使用的 Internet。

2. 未来新一代的计算机

(1) 模糊计算机。1956年,英国人查德创立了模糊信息理论。依照模糊理论,判断问题不是以是、非两种绝对的值或0与1两种数码来表示,而是取许多值,如接近、几乎、差不多及差得远等模糊值来表示。用这种模糊的、不确切的判断进行工程处理的计算机就是模糊计算机,或称模糊电脑。模糊电脑是建立在模糊数学基础上的电脑。模糊电脑除具有一般电脑的功能外,还具有学习、思考、判断和对话的能力,可以立即辨识外界物体的形状和特征,甚至可帮助人们从事复杂的脑力劳动。用模糊逻辑芯片和电路组合在一起,就能制成模糊计算机。

日本科学家把模糊计算机应用在地铁管理上:日本东京以北320 km的仙台市的地铁列车,在模糊计算机控制下,自1986年以来,一直安全、平稳地行驶着。车上的乘客可以不必攀扶拉手吊带,因为在列车行进中,模糊逻辑“司机”判断行车情况的错误,几乎比人类司机要少70%。1990年,日本松下公司把模糊计算机装在洗衣机里,能根据衣服的肮脏程度、衣服的材质调节洗衣程序。我国有些品牌的洗衣机也装上了模糊逻辑芯片。人们又把模糊计算机装在吸尘器里,可以根据灰尘量以及地毯的厚实程度调整吸尘器功率。模糊计算机还能用于地震灾情判断、疾病医疗诊断、发酵工程控制、海空导航巡视等方面。

(2) 生物计算机。生物计算机(Biological Computer)又称仿生计算机(Bionic Computer),是以生物芯片取代在半导体硅片上集成数以万计的晶体管制成的计算机。涉及计算机科学、脑科学、神经生物学、分子生物学、生物物理、生物工程、电子工程、物理学和化学等有关学科。1986年日本开始研究生物芯片,研究有关大脑和神经网络结构的信息处理、加工原理,以及建立全新的生物计算机原理,探讨适于制作芯片的生物大分子的结构和功能,以及如何通过生物工程(利用脱氧核糖核酸重组技术和蛋白质工程)来组装这些生物分子功能元件。

(3) 光子计算机。光子计算机是一种由光信号进行数字运算、逻辑操作、信息存储和处理的新型计算机。它由激光器、光学反射镜、透镜、滤波器等光学元件和设备构成,靠激光束进入反射镜和透镜组成的阵列进行信息处理,以光子代替电子,光运算代替电运算。光的并行、高速,天然地决定了光子计算机的并行处理能力很强,具有超高运算速度。光子计算机还具有与人脑相似的容错性,系统中某一元件损坏或出错时,并不影响最终的计算结果。光子在光介质中传输所造成的信息畸变和失真极小,光传输、转换时能量消耗和散发热量极低,对环境条件的要求比电子计算机低得多。

(4) 超导计算机。超导计算机是利用超导技术生产的计算机及其部件,其性能是目前电子计算机无法比拟的。目前制成的超导开关器件的开关速度,已达到几皮秒(10^{-12} s)的高水平。这是当今所有电子、半导体、光电器件都无法比拟的,比集成电路要快几百倍。超导计算机的运算速度比现在的电子计算机快100倍,而电能消耗仅是电子计算机的千分之一,如果目前一台大中型计算机每小时耗电10 kW,那么同样的一台超导计算机只需一节干电池就可以工作了。

(5) 量子计算机。量子计算机(Quantum Computer)是一类遵循量子力学规律进行高速数学和逻辑运算、存储及处理量子信息的物理装置。当某个装置处理和计算的是量子信息,运行的是量子算法

时，它就是量子计算机。量子计算机的概念源于对可逆计算机的研究。研究可逆计算机的目的是为了解决计算机中的能耗问题。

1.2 计算机的特点、应用及分类

计算机自诞生以来，其发展速度非常惊人，应用范围不断扩大，目前已渗透到人类生活的各个方面。

1.2.1 计算机的特点

计算机技术是信息化社会的基础、信息技术的核心，这是由计算机的特点决定的。概括地说，电子计算机和过去的计算工具相比具有以下几个方面的特点。

(1) 运算速度快。计算机的运算速度是其他任何一种工具无法比拟的。现在一台微型计算机的运算速度可以达到每秒处理千万条指令。目前，世界上速度最快的巨型计算机的运算速度可达每秒数十万亿次以上。正是有了这样的运算速度，使得过去不可能完成的计算任务可以完成，如天气预报、地震预报等。

(2) 计算精度高。计算机进行数值计算时所获得的精度可达到小数点后几十位、几百位甚至上万位。1981年，日本筑波大学利用计算机计算，将 π 值精确到小数点后200万bit。

(3) 具有超强的“记忆”能力和逻辑判断能力。“记忆”功能是指计算机能存储大量的信息，供用户随时检索和查询。逻辑判断功能是指计算机不仅能够进行算术运算，还能进行逻辑运算、实践推理和证明。记忆功能、算术运算和逻辑判断功能相结合，使得计算机能模仿人类的智能活动，成为人类脑力延伸的重要工具，所以计算机又称为“电脑”。

(4) 能自动运行并支持人机交互。所谓“自动运行”就是人们把需要计算机处理的问题编成程序，存入计算机中，当发出运行指令后，计算机便在该程序控制下依次逐条执行，不再需要人工干预。“人机交互”则是在人想要干预时，采用人机之间“一问一答”的形式，有针对性地解决问题。

(5) 网络与通信功能。目前最大、应用范围最广的国际互联网连接了全世界200多个国家和地区数亿台的各种计算机。在网上的所有计算机用户都可共享网上资料、交流信息、互相学习，将世界变成了地球村。

1.2.2 计算机的应用

计算机问世之初，主要用于数值计算，“计算机”也因此得名。如今的计算机几乎和所有学科相结合，使得计算机的应用渗透到社会的各个领域，如科学技术、国民经济、国防建设及家庭生活等。计算机的应用大致可分为如下几个领域。

(1) 科学计算，也称数值计算。科学计算是计算机应用最早的也是最成熟的应用领域。主要使用计算机进行数学方法的实现和应用。今天，计算机“计算”能力的提高推进了许多科学研究的进展，如著名的人类基因序列分析计划、人造卫星的轨道测算等。还有航天飞机、人造卫星、宇宙飞船、原子反应堆、气象预报、大型桥梁、地震测级、地质勘探和机械设计等都离不开计算机的科学计算。如果没有计算机，如此巨大、繁多的计算单靠人类自身是绝对无法完成的。

(2) 过程检测与控制。过程控制也称实时控制，在工业生产、国防建设和现代化战争中都有广泛的应用。例如，工业生产自动化方面的巡回检测、自动记录、监测报警、自动启停、自动调控等；在交通运输方面的红绿灯控制、行车调度；在国防建设方面的导弹发射中，实施控制其飞行的方向、速度、位置等。

(3) 数据/信息处理，也称非数值计算。现代社会是信息化的社会。随着社会的不断进步，信息量也在急剧增加；现在，信息已和能源、物资一起构成人类社会活动的基本要素。计算机最广泛的应用就是信息处理，有关资料表明，世界上80%左右的计算机主要用于信息处理。信息处理的特点是处理

的数据量较大,但不涉及复杂的数学运算;有大量的逻辑判断和输入/输出,时间性较强,如生产管理、财务管理、人事管理、情报检索、办公自动化、票务管理等。

(4) 计算机辅助系统。当前用计算机辅助工作的系统越来越多,如计算机辅助设计(Computer Aided Design, CAD)、计算机辅助制造(Computer Aided Manufacturing, CAM)、计算机辅助教学(Computer Assisted Instruction, CAI)、计算机辅助测试(Computer Aided Testing, CAT)、计算机辅助工程(Computer Aided Engineering, CAE)、计算机集成制造系统(Computer Integrated Manufacturing System, CIMS)等。

(5) 人工智能(Artificial Intelligence, AI)是研究使计算机来模拟人的某些思维过程和智能行为(如学习、推理、思考、规划等)的学科,主要包括计算机实现智能的原理、制造类似于人脑智能的计算机,使计算机能实现更高层次的应用。人工智能涉及计算机科学、心理学、哲学和语言学等学科。可以说几乎是自然科学和社会科学的所有学科,其范围已远远超出了计算机科学的范畴,人工智能与思维科学的关系是实践和理论的关系,人工智能是处于思维科学的技术应用层次,是它的一个应用分支。从思维观点看,人工智能不仅限于逻辑思维,要考虑形象思维、灵感思维才能促进人工智能的突破性发展,数学常被认为是多种学科的基础科学,数学也进入语言、思维领域,人工智能学科也必须借用数学工具,数学不仅在标准逻辑、模糊数学等范围发挥作用,数学也进入人工智能学科,它们将互相促进而更快地发展。应用领域也不断扩大,可以设想,未来人工智能带来的科技产品,将会是人类智慧的“容器”,也可能超过人的智能。

(6) 网络应用。计算机网络是微电子技术、计算机技术和现代通信技术的结合。计算机网络的建立解决了一个单位、一个地区、一个国家,乃至全世界范围内的计算机与计算机之间的相互通信及各种硬件资源、软件资源和信息资源的共享。目前,世界各国都相继建立了自己的网络系统,并分别与Internet相连。我国已建和在建的信息网络共有9个,并先后启动了政府上网和企业上网工程。网络技术的发展和应用已成为人们谈论的热门话题。

(7) 多媒体应用。多媒体包括文本、图形、图像、音频、视频、动画等多种信息类型。多媒体技术是指人和计算机交互地进行上述多种媒介信息的捕捉、传输、转换、编辑、存储、管理,并由计算机综合处理为表格、文字、图形、动画、音/视频等视听信息有机结合的表现形式。多媒体技术拓宽了计算机的应用领域,使计算机广泛应用于商业、服务业、教育、广告宣传、文化娱乐、家庭等方面。同时,多媒体技术与人工智能技术的有机结合还促进了虚拟现实、虚拟制造技术的发展,使人们可以在计算机上感受真实的场景,通过计算机仿真制造零件和产品,感受产品各方面的功能和性能。

(8) 嵌入式系统。并不是所有计算机都是通用的。有许多特殊的计算机用于不同的设备中,包括大量的消费电子产品和工业制造系统,都是把处理器芯片嵌入其中,完成特定的处理任务。这些系统称为嵌入式系统。如数码相机、数码摄像机以及高档电动玩具等都使用了不同功能的处理器。

(9) VR(Virtual Reality, 虚拟现实, 或称灵境技术),它是一种可创建和体验虚拟世界的计算机系统。虚拟现实的内涵实际是综合利用计算机图形系统和各种现实及控制等接口设备,在计算机上生成的、可交互的三维环境中提供沉浸感觉的技术,类似于3D技术。

(10) AR(Augmented Reality, 增强现实),它可以将虚拟信息放在现实中展现,并且让人和虚拟信息进行互动。AR通过技术上的手段能够将现实与虚拟信息进行无缝对接。将在现实中不存在的事物构建一个三维场景予以展现,与现实生活相互衔接。它是一种全新的人机交互技术,利用这样一种技术,可以模拟真实的现场景观,它是交互性和构想为基本特征的计算机高级人机界面。

(11) 数据挖掘(Data mining)是数据库知识发展的一个步骤。数据挖掘一般是指从大量的数据中通过算法搜索隐藏于其中的信息的过程。数据挖掘通常与计算机科学有关,并通过统计、在线分析处理、情报检索、机器学习、专家系统(依靠过去的经验法则)和模式识别等诸多方法来实现上述目标。

1.2.3 计算机的分类

计算机及其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。

1. 传统计算机的分类方法

自计算机诞生以来，先后出现了多种计算机的分类方法，其中较为常见的计算机分类方法主要包括如下几种：

（1）按信息的表示与处理方法不同可分为模拟计算机、数字计算机和混合式计算机。

（2）按计算机的用途不同可分为通用计算机和专用计算机。

（3）按计算机的性能和规模不同可分为个人计算机、工作站、小型计算机、主机、小巨型计算机、巨型计算机。其中，这种分类方法是由美国电气和电子工程师学会（IEEE）于1999年提出的，是目前常用的一种分类方法。

2. 现代计算机的分类方法

随着技术的进步，各种型号的计算机性能指标都在不断地改进和提高，过去一台大型机的性能可能还比不上今天一台微型计算机。按照传统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。现在，通常根据计算机的综合性能指标，并结合计算机应用领域的分布对计算机进行分类。

（1）高性能计算机

高性能计算（High Performance Computing, HPC）俗称超级计算机，通常是指使用多个处理器（作为单个机器的一部分）或者某一集群中组织的几台计算机（作为单个计算资源操作）的计算机系统和环境。目前国际上对高性能计算机的最为权威的评测是世界计算机排名（即TOP500），通过测评的计算机是目前世界上运算速度和处理能力均堪称一流的计算机。我国生产的超级计算机主要有曙光-5000A、深腾7000、神威3000A、天河二号等，2014年6月23日，在德国莱比锡市发布的第43届世界超级计算机500强排行榜上，中国超级计算机系统“天河二号”获得世界超算“三连冠”，其运算速度比位列第二名的美国“泰坦”快了近一倍。

（2）微型计算机

微型计算机简称“微型机”或“微机”，由于其具备人脑的某些功能，所以也称为“微电脑”。微型计算机是由大规模集成电路组成的、体积较小的电子计算机。它是以微处理器为基础，配以内存储器及输入输出（I/O）接口电路和相应的辅助电路而构成的裸机。目前微型计算机已广泛应用于办公、学习、娱乐等社会生活的方方面面，是发展最快、应用最为普及的计算机。我们日常使用的台式计算机、笔记本电脑、掌上型计算机等都属于微型计算机。

（3）工作站

工作站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和外存储器，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。工作站主要是为了满足工程设计、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高性能微型计算机。

（4）服务器

服务器是指在网络环境下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。嵌入式计算机系统是以应用为中心,以计算机技术为基础,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、嵌入式操作系统以及用户的应用程序等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,我们日常生活中使用的电冰箱、全自动洗衣机、空调、电饭煲、数码产品等都采用了嵌入式计算机技术。

1.3 信息技术

信息技术(Information Technology, IT)的飞速发展促进了信息社会的到来。半个世纪以来,人类社会正由工业社会全面进入信息社会,其主要动力就是以计算机技术、通信技术和控制技术为核心的现代信息技术的飞速发展和广泛应用。随着科学技术的飞速发展,各种高新技术层出不穷、日新月异,但是最主要、发展最快的仍然是信息技术。

1.3.1 信息技术概述

1. 信息技术的定义

在现代信息社会中,一切可以用二进制进行编码的东西都可以称为信息。一般来说,信息的采集、加工、存储、传输和利用过程的每一种技术都是信息技术,也就是说,信息技术一般是指一系列与计算机相关的技术,如微电子技术、光电子技术、通信技术、网络技术、感测技术、控制技术、显示技术等。它也常被称为信息和通信技术(Information and Communications Technology, ICT)。

2. 现代信息技术的内容

一般来说,信息技术包含三个层次的内容:信息基础技术、信息系统技术、信息应用技术。

(1) 信息基础技术。信息基础技术是信息技术的基础,包括新材料、新能源、新器件的开发和制造技术。近几十年来,发展最快、应用最广、对信息技术及整个高科技领域的发展影响最大的是微电子技术和光电子技术。

(2) 信息系统技术。信息系统技术是指有关信息的获取、传输、处理、控制的设备和系统技术。感测技术、通信技术、计算机技术、控制技术是它的核心和支撑技术。

(3) 信息应用技术。信息应用技术是针对各种实用目的的,如信息管理、信息控制的信息决策而发展起来的具体技术,如企业生产自动化、办公自动化、家庭自动化、人工智能和互联网技术等。它们是信息技术开发的根本目的所在。信息技术在社会的各个领域得到了广泛的应用,显示出强大的生命力。

3. 现代信息技术的发展趋势

在社会生产力发展、人类认识和实践活动的推动下,信息技术将得到更深、更广、更快的发展,当前信息技术发展的总趋势是以互联网技术的发展和应用为中心,从典型的技术驱动发展模式向技术驱动与应用驱动相结合的模式转变,其发展趋势可以概括为数字化、多媒体化、高速化、网络化、宽带化和智能化等。

1.3.2 信息化和信息产业

1. 信息化

信息化是指以培养、发展以计算机为主的智能化工具为代表的新生产力,并使之造福于社会的历

史过程。信息化是以现代通信、网络、数据库技术为基础，将所研究对象各要素汇总至数据库，供特定人群生活、工作、学习、辅助决策等并和人类息息相关的各种行为相结合的一种技术，使用该技术后，可以极大提高各种行为的效率，为推动人类社会进步提供极大的技术支持。

2. 信息产业

信息产业属于第四产业范畴，它包括电信、电话、印刷、出版、新闻、广播、电视等传统的信息部门和新兴的电子计算机、激光、光导纤维（简称光纤）、通信卫星等信息部门。主要以电子计算机为基础，从事信息的生产、传递、储存、加工和处理。

1.4 信息的表示及编码基础知识

计算机最主要的功能是信息处理。在计算机内部，各种信息，如数字、文字、图形、视频、声音等都必须采用数字化的编码形式进行存储、处理和传输。由于在计算机内部处理二进制数，所以数字化编码的实质就是用 0 和 1 两个数字进行各种组合，将要处理的信息表示出来。

1.4.1 计算机中的数制

计算机科学中经常使用十进制、二进制、八进制和十六进制。但在计算机内部，不管什么样的数都使用二进制来表示。在具体讨论计算机常用数制之前，首先介绍几个有关数制的基本概念。

1. 进位计数制的概念

在十进制数中，一个数可以用 0~9 这 10 个阿拉伯数字的组合来表示，这 10 个数字再加上数位的概念，就可以表示任何一个十进制数了。例如：

$$2181=2\times 10^3+1\times 10^2+8\times 10^1+1\times 10^0=2000+100+80+1$$

其中，

(1) 0~9 这些数字符号称为数码。

(2) 全部数码的个数称为基数，十进制数的基数为 10。

(3) 用“逢基数进位”的原则进行计数，称为进位计数制。十进制数的基数为 10，所以其计数原则是“逢十进一”。

(4) 所谓权值就是数字在数中所处位置的单位值。

(5) 权值与基数的关系是：权值等于基数的若干次方。

在十进制数中，各个位的权值分别是 10^i (i 为整数)。例如：

$$12\ 345.67=1\times 10^4+2\times 10^3+3\times 10^2+4\times 10^1+5\times 10^0+6\times 10^{-1}+7\times 10^{-2}$$

式中， 10^4 、 10^3 、 10^2 、 10^1 、 10^0 、 10^{-1} 、 10^{-2} 即为各个位的权值，每一位上的数码与该位权值的乘积，就是该位的数值。

(6) 任何一个十进制数 A 都可以用如下形式的展开式表示出来：

设 $A=(a_n a_{n-1} a_{n-2} \cdots a_1 a_0 a_{-1} a_{-2} \cdots a_{-m})_{10}$ ，则

$$\begin{aligned} A &= a_n \times 10^n + a_{n-1} \times 10^{n-1} + \cdots + a_1 \times 10^1 + a_0 \times 10^0 + a_{-1} \times 10^{-1} + a_{-2} \times 10^{-2} + \cdots + a_{-m} \times 10^{-m} \\ &= \sum a_i \times 10^i \end{aligned}$$

式中， a_i 为第 i 位数码，10 为基数， $i=n \sim -m$ ， n 、 m 为正整数。

同样道理，任何一个 R 进制的数 $B=(b_n b_{n-1} b_{n-2} \cdots b_1 b_0 b_{-1} b_{-2} \cdots b_{-m})_R$ 可按一般展开式展开为：

$$\begin{aligned} B &= b_n \times R^n + b_{n-1} \times R^{n-1} + \cdots + b_1 \times R^1 + b_0 \times R^0 + b_{-1} \times R^{-1} + b_{-2} \times R^{-2} + \cdots + b_{-m} \times R^{-m} \\ &= \sum b_i \times R^i \quad (i=n \sim -m) \end{aligned}$$

式中， b_i 为第 i 位数码， R 为基数， $i=n \sim -m$ ， n 、 m 为正整数。

2. 计算机常用的数制

计算机能够直接识别的只是二进制数。这就意味着它所处理的数字、字符、图像、声音等信息，都是1和0组成的二进制数的某种编码。

由于二进制在表达一个数字时，位数太长，不易识别，书写麻烦，因此在编写计算机程序时，经常将它们写成对应的十六进制数或八进制数，也经常采用人们熟悉的十进制数。因此，计算机工作时，往往需要根据情况，在其内部要进行二、八、十、十六进制数的转换。表1.2给出了常用计数制的基数和数码，表1.3给出了各种进制数关系表。

表 1.2 常用计数制的基数和数码

| 数 制 | 基 数 | 数 码 |
|------|-----|---------------------------------|
| 二进制 | 2 | 0、1 |
| 八进制 | 8 | 0、1、2、3、4、5、6、7 |
| 十进制 | 10 | 0、1、2、3、4、5、6、7、8、9 |
| 十六进制 | 16 | 0、1、2、3、4、5、6、7、8、9、A、B、C、D、E、F |

表 1.3 各种进制数关系表

| 十进制 | 二进制 | 八进制 | 十六进制 | 十进制 | 二进制 | 八进制 | 十六进制 |
|-----|------|-----|------|-----|------|-----|------|
| 0 | 0000 | 0 | 0 | 8 | 1000 | 10 | 8 |
| 1 | 0001 | 1 | 1 | 9 | 1001 | 11 | 9 |
| 2 | 0010 | 2 | 2 | 10 | 1010 | 12 | A |
| 3 | 0011 | 3 | 3 | 11 | 1011 | 13 | B |
| 4 | 0100 | 4 | 4 | 12 | 1100 | 14 | C |
| 5 | 0101 | 5 | 5 | 13 | 1101 | 15 | D |
| 6 | 0110 | 6 | 6 | 14 | 1110 | 16 | E |
| 7 | 0111 | 7 | 7 | 15 | 1111 | 17 | F |

3. 书写规则

为了区分各种计数制的数，常采用以下两种方法。

(1) 在括号外面加数字下标。例如：

$(1011)_2$ 表示二进制数 1011

$(1011)_8$ 表示八进制数 1011

$(1234)_{10}$ 表示十进制数 1234

$(23AD)_{16}$ 表示十六进制数 23AD

(2) 在数字后面加写相应的英文字母作为标志。

B (Binary) 表示二进制；O (Octonary) 表示八进制（但为了避免字母 O 与数字 0 相混淆，常用 Q 代替 O）；D (Decimal) 表示十进制（可以省略，默认十进制）；H (Hexadecimal) 表示十六进制。例如：

$(10010001)_2$ 可表示成 10010001B

$(1357)_8$ 可表示成 1357Q

$(1998)_{10}$ 可表示成 1998D 或 1998

$(3DF6)_{16}$ 可表示成 3DF6H

1.4.2 不同进制数之间的转换

不同进制数之间的转换原则是：如果两个有理数相等，则两数的整数部分和小数部分分别相等。因此，进行各计数制之间的转换时，都是把整数部分和小数部分分别进行转换的。

1. 各种进制数转换为十进制数

(1) 二进制数转换成十进制数

根据二进制数的定义，只要将它们按权值展开求和，就可以得到相应的十进制数，将这种方法称为“按权乘基数相加法”。例如：

$$(100110.101)_2 = 1 \times 2^5 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^{-1} + 1 \times 2^{-3} = 32 + 4 + 2 + 0.5 + 0.125 = (38.625)_{10}$$

(2) 八进制数或十六进制数转换成十进制数

八进制数或十六进制数转换成十进制数的方法与二进制数转换成十进制数相同，只是其中的各个数位的权值不同而已。

2. 十进制数转换为二进制数

根据不同计数制之间的转换原则，当要将一个十进制数转换为二进制数时，通常是将其整数部分和小数部分分别进行转换，然后再将转换结果组合在一起。

(1) 整数部分的转换

转换方法：除 2 取余法。

具体做法为：将十进制整数除以 2，得到一个商和一个余数（0 或 1），记下余数，并将所得的商再除以 2，又得到一个新的商和一个新的余数，如此反复进行，直到商为 0 为止，将依次得到的余数反序排列起来，便可得到相应的二进制整数。例如将十进制整数 83 转换成二进制整数，转换过程如图 1.1 所示，转换结果为

$$(83)_{10} = (1010011)_2$$

(2) 小数部分的转换

转换方法：乘 2 取整法。

具体做法为：将给定的十进制纯小数乘以 2，得到一个乘积，将乘积的整数部分取出并记录（0 或 1），将剩余的纯小数部分再乘以 2，又得到一个新的乘积，如此反复进行，直到乘积的小数部分为 0 或满足指定的精度要求为止，将依次得到并记录的各次整数顺序排列起来，便可得到相应的二进制数小数。如将十进制小数 0.687 5 和 0.306 95 转换成二进制小数，转换过程如图 1.2 所示。

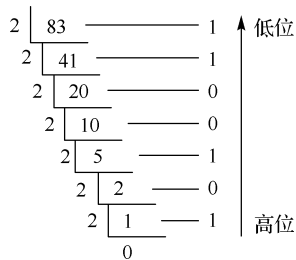


图 1.1 除 2 取余法

| | | | |
|---------------------------|--------|-----------------------------|--------|
| $0.6875 \times 2 = 1.375$ | 取出整数 1 | $0.30695 \times 2 = 0.6139$ | 取出整数 0 |
| $0.375 \times 2 = 0.75$ | 取出整数 0 | $0.6139 \times 2 = 1.2278$ | 取出整数 1 |
| $0.75 \times 2 = 1.50$ | 取出整数 1 | $0.2278 \times 2 = 0.4556$ | 取出整数 0 |
| $0.50 \times 2 = 1.00$ | 取出整数 1 | $0.4556 \times 2 = 0.9112$ | 取出整数 0 |
| | | $0.9112 \times 2 = 1.8224$ | 取出整数 1 |

图 1.2 转换过程

$$(0.6875)_{10} = (0.1011)_2 \quad (0.30695)_{10} = (0.01001)_2$$

注意：多余的位数可以按“0 舍 1 入”的规律取近似值，保留指定的小数位数。

对于包含整数和小数的十进制数，当分别转换为对应的二进制数后，还需将它们组合起来，例如：

$$(83.6875)_{10} = (1010011.1011)_2$$

3. 十进制数转换为其他进制数

十进制数转换为其他进制数的方法与十进制数转换为二进制数的方法相似，也是分为整数部分和

小数部分分别进行转换的，只是每次所要乘除的不是“2”。当把十进制数转换为八进制数或十六进制数时，每次将乘除“8”或“16”。

总之，将十进制数转换为任何进制数时，对于整数部分的转换，所采用的方法都是“除基数取余法”；而对于小数部分的转换，则采用“乘基数取整法”。

4. 二进制数和十六进制数之间的转换

(1) 二进制数转换成十六进制数

由于 $2^4=16$ ，可以用4位二进制数对应于1位十六进制数（0000~1111→0~F），所以将二进制数转换成十六进制数时可以采用“四位一并法”。即从小数点开始向左或向右，每4位1组，不足4位的用0补足，将每4位二进制数用1位与之相对应的十六进制数来代替即可。例如：

$$\begin{array}{ccccccc} (0010 & 1100 & 1010 & 0110 & .1000 & 1110 & 1000)_2 \\ | & | & | & | & | & | & | \\ (2 & C & A & 6 & .8 & E & 8)_{16} \end{array}$$

即

$$(101100101001110.100011101)_2=(2CA6.8E8)_{16}$$

(2) 十六进制数转换成二进制数

其转换是二进制数转换成十六进制数的反过程，可以采用“一分为四法”。例如：

$$\begin{array}{cccccc} (3 & A & 5 & E & .7 & B)_{16} \\ | & | & | & | & | & | \\ (0011 & 1010 & 0101 & 1110 & .0111 & 1011)_2 \end{array}$$

即

$$(3A5E.7B)_{16}=(11101001011110.01111011)_2$$

1.4.3 二进制数的常用单位

在计算机内部，一切数据都用二进制数的编码来表示。为了衡量计算机中数据的量，人们规定了一些二进制数的常用单位，如位、字节、字等。

1. 位 (bit)

位是二进制数中的一个数位，可以是0或1。它是计算机存储信息的最小单位，称为比特 (bit)。

2. 字节 (Byte)

在计算机中，将8个连续的二进制位称为一个字节 (Byte，简记为B)，即一个字节可以表示8个二进制位。字节是描述计算机存储容量的基本单位，每一个字节可存放一个西文字符的编码，每两个字节存放一个中文字符的编码。随着计算机存储容量的不断扩大，用字节来表示存储容量就显得太小，为此又出现千字节 (KB)、兆字节 (MB)、吉字节 (GB)、太字节 (TB)、拍字节 (PB)、艾字节 (EB) 等单位，它们之间的转换关系如下。

$$1 \text{ KB}=1\,024 \text{ B}=2^{10} \text{ B} \quad 1 \text{ MB}=1\,024 \text{ KB}=2^{20} \text{ B} \quad 1 \text{ GB}=1\,024 \text{ MB}=2^{30} \text{ B}$$

3. 字 (word)

字是计算机一次存取、加工、运算和传输的数据长度。一个字一般由一个或几个字节组成，它是衡量计算机性能的一个重要指标。计算机的字长越长，其运算越快、计算精度越高。计算机的字长通常有8位、16位、32位、64位等。通常我们说多少位的计算机，就是指计算机的字长是多少位。

1.4.4 字符编码

字符是计算机中使用最多的信息形式之一，也是人与计算机通信的重要媒介。将字符变为指定的二进制符号称为编码。在计算机内部，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。

一个编码就是一串二进制位“0”或“1”的组合，二进制数串的位数就决定了符号集的规模。例如，对一个由128个符号构成的符号集进行编码，就需要7位二进制数；对于256个符号的字符集，就需要8位二进制数。

1. ASCII 码

目前，计算机中使用最广泛的符号是ASCII码，即美国国家标准信息交换码（American Standard Code for Information Interchange）。ASCII包括32个通用控制字符、10个十进制数码、52个英文大小写字母和34个专用符号，共128个元素，故需要用7位二进制数进行编码，以区分每个字符。通常使用一个字节（即8个二进制位）表示一个ASCII码字符，规定其最高位总是0。表1.4列出了ASCII码的编码表。

表 1.4 ASCII 码的编码表

| $d_6d_5d_4$ $d_3d_2d_1d_0$ | 000 | 001 | 010 | 011 | 100 | 101 | 110 | 111 | $d_6d_5d_4$ $d_3d_2d_1d_0$ | 000 | 001 | 010 | 011 | 100 | 101 | 110 | 111 |
|-------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| 0000 | NUL | DLE | 空格 | 0 | @ | P | , | p | 1000 | BS | CAN | (| 8 | H | X | h | x |
| 0001 | SOL | DC1 | ! | 1 | A | Q | a | q | 1001 | HT | EM |) | 9 | I | Y | i | y |
| 0010 | STX | DC2 | " | 2 | B | R | b | r | 1010 | LF | SUB | * | : | J | Z | j | z |
| 0011 | ETX | DC3 | # | 3 | C | S | c | s | 1011 | VT | ESC | + | ; | K | [| k | { |
| 0100 | EOT | DC4 | \$ | 4 | D | T | d | t | 1100 | FF | FS | , | < | L | \ | l | |
| 0101 | ENQ | NAK | % | 5 | E | U | e | u | 1101 | CR | GS | - | = | M |] | m | } |
| 0110 | ACK | SYN | & | 6 | F | V | f | v | 1110 | SO | RS | . | > | N | ↑ | n | - |
| 0111 | BEL | ETB | ' | 7 | G | W | g | w | 1111 | SI | US | / | ? | O | - | o | DEL |

例如，分别用二进制数和十六进制数写出“GOOD!”的ASCII编码。

用二进制数表示：01000111B 01001111B 01001111B 01000100B 00100001B

用十六进制数表示：47H 4FH 4FH 44H 21H

2. BCD 码

BCD（Binary Code Decimal）码又称“二-十进制编码”，是用4位二进制数以不同的组合来表示十进制数的一种方法。

由于4位二进制数能表示16个不同的数，现在只需要其中的10个数，因此有6个数是不用的。究竟取其中的哪10个数，有很多方案，也就是说有很多种BCD码，其中最常用的是8421码。8421码是基本BCD码。如果不特殊说明，通常说BCD码就是指8421码。8421码是取16个数中的前10个数，每位的权值分别为8、4、2、1，这样每4位二进制数就对应了1位十进制数。例如，73的8421码为01110011。

1.4.5 汉字编码

ASCII码只对英文字母、数字和标点符号进行编码。为了使计算机能够处理、显示、打印、交换汉字字符，同样也需要对汉字进行编码。

1. 汉字的处理过程

从汉字编码的角度看, 计算机对汉字信息的处理过程实际上是各种汉字编码间的转换过程。这些编码主要包括: 汉字输入码、汉字机内码、汉字地址码、汉字字形码等, 汉字信息处理中的各种编码及流程如图 1.3 所示。

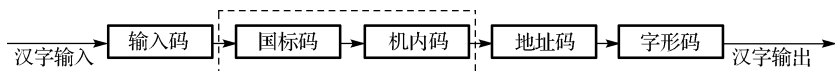


图 1.3 汉字信息处理系统的模型

2. 汉字编码

汉字的编码主要有 4 类: 汉字输入码、汉字交换码、汉字机内码和汉字字形码。

(1) 汉字输入码

汉字输入码是一种用计算机标准键盘上的按键的不同组合输入汉字而编制的编码, 也称汉字外部码, 简称外码。目前已有几百种汉字输入编码法, 但大致可分为如下 4 类。

① 数字编码, 也称顺序码。它用数字代表一个汉字输入, 如区位码、电报码等。

② 字音编码。它是根据汉字的读音进行编码的。由于汉字同音字很多, 输入重码率较高, 输入时一般要对同音字进行选择, 且对不知道读音的字将无法输入, 如拼音码、自然码等。

③ 字形编码。字形码是根据汉字的字形进行编码的。汉字都是由一笔一画组成的, 把汉字的笔画部件用字母或数字进行编码, 按笔画书写的顺序依次输入就能表示一个汉字, 例如五笔字形码、大众码等。

④ 音形编码。根据汉字的读音和字形进行编码。音形码吸收了字音和字形编码的优点, 使编码规则简化, 重码减少, 如双拼码、五十字元码等。

(2) 汉字交换码

汉字交换码是汉字信息处理系统之间或通信系统之间传输信息时所使用的编码, 即国标码。国标码以国家标准局公布的 GB2312—1980《信息交换用汉字编码字符集——基本集》规定的汉字交换码作为标准汉字编码。共收录汉字、字母、图形等符号 7 445 个, 其中汉字 6 763 个(常用的一级汉字 3 755 个, 按汉语拼音字母顺序排列; 二级汉字 3 008 个, 按偏旁部首顺序排列)。国标码规定, 每个字符由一个 2 字节代码组成。每个字节的最高位恒为“0”, 其余 7 位用于组成各种不同的码值。两个字节的代码, 共可表示 $128 \times 128 = 16\,384$ 个符号, 而国标码的基本字符集中, 目前只有 7 445 个字符。GB2312—1980 基本字符集将汉字按规则排列成 94 行、94 列, 形成汉字编码表。汉字输入法中的区位码输入法就是利用了此基本集进行汉字输入的。

2001 年, 我国发布了 GB18030—2000 编码标准, 即《信息交换用汉字编码字符集——基本集的扩充》, 纳入编码的汉字约为 2.7 万个。

(3) 汉字机内码

汉字机内码, 或称汉字内码, 是设备和汉字信息处理系统内部存储、处理、传输汉字而使用的编码。目前西文大多采用 ASCII 编码为内码来设计计算机系统。但汉字字数较多, 一般要用两个字节来存放汉字的内码。为了与 ASCII 编码相区分, 英文字符的机内代码是 7 位 ASCII 码, 最高位为 0, 汉字机内码中两个字节的最高位均为 1, 其余的两个 7 位用对应的国标码, 即用两个 8 位码构成一个汉字内部码。例如, 汉字“补”的国标码是 3239H(0011001000111001B), 它的机内码是 B2B9H(10110010101111001B)。

(4) 汉字字形码

汉字字形码也称输出码, 用于显示或打印汉字时产生字形。该种编码是通过点阵形式产生的。

不论汉字的笔画有多少，都规范在同样大小的范围内书写。把规范的方块再分割成许多小方块来组成一个点阵，这些小方块就是点阵中的一个点，即二进制的位。每个点由“0”和“1”表示“白”和“黑”两种颜色。一个汉字信息系统具有的所有汉字字形码的集合就是该系统的汉字库。根据对输出汉字精美程度的要求不同，汉字点阵的多少也不同，点阵越大输出的字形越精美。简易型汉字为 16×16 点阵，多用于显示；提高型为 24×24 点阵、 32×32 点阵、 48×48 点阵、 64×64 点阵等，多用于打印输出。

（5）汉字地址码

汉字地址码是指汉字库（这里主要是指字形的点阵式字模库）中存储汉字字形信息的逻辑地址码。当需要向输出设备输出汉字时，必须通过地址码。汉字库中，字形信息都是按一定顺序（大多数按标准汉字交换码中汉字的排列顺序）连续存放在存储介质上，所以汉字地址码也大多是连续有序的，而且与汉字内码间有着简单的对应关系，以简化汉字内码到汉字地址码的转换。

3. 汉字编码转换

区位码、国标码与机内码的转换方法

- ① 区位码先转换成十六进制数表示；
- ② $(\text{区位码的十六进制表示}) + 2020\text{H} = \text{国标码}$ ；
- ③ $\text{国标码} + 8080\text{H} = \text{机内码}$ 。

【例 1.1】 已知“大”字的区内码为 2083，试计算对应的国标码和机内码。

解答：“大”字的区号和位号分别为 20 和 83，其对应的十六进制表示为 1453H，因此，“大”字的国标码为 $1453\text{H} + 2020\text{H} = 3473\text{H}$ ，“大”字的机内码为 $3473\text{H} + 8080\text{H} = \text{B4F3H}$ 。

习 题 一

一、选择题

1. 1946 年诞生了世界上第一台电子计算机，它的英文名字是（ ）。
A. UNIVAC-I B. EDVAC C. ENIAC D. Mark-2
2. 现代计算机正朝两极方向发展，即（ ）。
A. 专用机和通用机 B. 微型机和巨型机
C. 模拟机和数字机 D. 个人机和工作站
3. CAD 的中文含义是（ ）。
A. 计算机辅助设计 B. 计算机辅助制造
C. 计算机辅助工程 D. 计算机辅助教学
4. 关于电子计算机的特点，以下论述错误的是（ ）。
A. 运算速度快 B. 运算精度高
C. 具有记忆和逻辑判断能力 D. 运行过程不能自动、连续，需人工干预
5. 数值 10H 是（ ）的一种表示方法。
A. 二进制数 B. 八进制数 C. 十进制数 D. 十六进制数
6. 国标码（GB2312—1980）依据使用频度，把汉字分成（ ）。
A. 简化字和繁体字 B. 一级汉字、二级汉字、三级汉字
C. 常用汉字和图形符号 D. 一级汉字、二级汉字
7. BCD 是专门用二进制数表示（ ）的编码。

- A. 字母符号 B. 数字字符 C. 十进制数 D. 十六进制数
8. 国标码(GB2312—1980)是()的标准编码。
- A. 汉字输入码 B. 汉字字形码 C. 汉字机内码 D. 汉字交换码

二、简答题

1. 世界上第一台电子计算机产生的时间、地点? 被命名为什么?
2. 冯·诺依曼结构计算机的工作原理的核心是什么? 它所具有的三个要点是什么?
3. 计算机的发展经历了哪几个阶段? 各阶段的主要特征是什么?
4. 计算机的发展趋势是什么?
5. 计算机具有哪几个方面的特点?
6. 计算机的主要应用范围是什么?
7. 计算机都有哪些分类方法? 各分为哪几类?
8. 什么是BCD码? 3908的BCD码是什么?
9. 什么是ASCII码? 大写英文字母、小写英文字母与数字三者ASCII码的大小顺序如何?
10. 常用的汉字编码有几种? 它们各自的用途是什么?