

项目一 计算机基础知识

项目概述：本项目通过 9 个具体任务的完成使读者掌握计算机的基础知识，包括信息及信息技术，计算机的发展历史、特点、分类，计算机内数据的编码方法、不同进制间的转换，计算机系统的组成，当前主流计算机配置，汉字输入方法，多媒体技术和计算机病毒的有关知识。

任务一 信息与信息技术

任务描述：随着科学技术的发展，信息技术也在突飞猛进地发展。21 世纪，人类已全面进入了信息时代，信息技术已广泛应用于社会的各个领域。信息技术的发展与应用已成为衡量一个国家综合国力的关键因素。那么什么是信息和信息技术呢？通过本任务的学习，读者可以系统地了解信息与信息的概念以及信息技术的应用和发展。

任务分析：想认识 and 了解信息与信息技术，必须从基本概念出发，掌握其特征和发展。

知识点：信息 信息技术 信息的特征、应用和发展

1.1.1 信息的基本概念

“信息”一词有着非常悠久的历史，早在几千年前，即有“信”字出现。“信”常可作消息来理解。作为日常用语，“信息”经常是指“音讯、消息”的意思。如我国周朝时期用烽火台传递边关警报，古罗马用悬灯来报告敌人进攻的消息，近代发明的电报、电话也是为了传递信息，人们每天都在利用语言、文字、声音、信件、网络等来传递信息。信息已是现代社会中广泛使用的一个概念，它存在于人们日常生活的各个环节之中。人们对信息的认识和理解也越来越深入，但至今信息还没有一个公认的定义，不同的学科，从不同的角度对信息的概念有不同的解释。一般认为，信息是自然界、人类社会和人类思维活动中普遍存在的一切物质和事物的属性，或者说信息就是指以声音、语言、文字、图像、动画、气味等方式所表示的实际内容。

“信息”是有价值的，就像不能没有空气和水一样，人类也离不开信息。因此人们常说，物质、能量和信息是构成世界的三大要素。所以说，信息的传播是极其重要的。

1.1.2 信息的特征

信息是人们认识世界、改造世界，取之不尽、用之不竭的宝贵资源，它具有区别于其他事物本质的特征，信息必须依赖于载体而存在，具有客观性、普遍性、时效性、价值性、依存性、共享性、传递性和可加工性等。

1. 客观性

客观性是指信息的内容必须真实可靠。事实是信息的中心价值，不符合事实的信息不仅没有价值，而且可能引到负面价值上。

2. 普遍性

普遍性是指信息无处不在、无时不有。信息普遍存在于自然界、人类社会中，也存在于人类的思维或精神领域中。它是事物运动的状态和方式，只要有事物存在，就会有其运行的状

态和方式,就存在着信息。

3. 时效性

时效性是信息的重要特征,是指信息从发出、接收到利用的时间间隔及其效率。信息的时效性与信息的价值性密不可分,任何有价值的信息都是在一定的条件下起作用的,如时间、地点、事件等,离开一定的条件,信息就会失去一定的价值。

4. 依存性

依存性是指信息的存储、传递和交流必须依附在一定的载体之上。信息本身是看不见、摸不着的,它必须依附于一定的物质形式,不可能脱离物质而单独存在。

5. 共享性

共享性是指信息可由不同个体或群体在同一时间或不同时间共同享用。正是由于信息可被共享的特点,信息资源才能够发挥最大效用,生生不息。

6. 传递性

传递性是指信息可以通过一定的传输工具和载体进行空间上和时间上的传递。一个完整的信息传递过程必须具备信源(信息发送方)、信宿(信息接收方)、信道(媒介、载体)和信息4个基本要素。

7. 加工性

信息可以被分析、综合、扩大等,也就是人们可以对信息进行加工处理,把信息从一种形式变换成另一种形式。

1.1.3 信息技术

信息技术是指同获取、传递、再生和利用信息有关的技术。或者说凡是能扩展人的信息功能的技术都是信息技术。它主要是指利用计算机和现代通信手段实现获取信息、传递信息、存储信息、处理信息、显示信息、分配信息等的相关技术,主要包括感测与识别技术、信息传递技术、信息处理与再生技术、信息实用技术等几个方面。它是利用科学的原理、方法及先进的工具和手段有效地开发和利用信息资源的技术体系。

从应用的角度来看,信息技术经历了数值处理、数据处理、知识处理、智能处理、网络处理5个阶段,目前正在向网络处理阶段过渡。

目前,以网络化、数字化、多媒体和智能化为代表的现代信息技术正在改变着人们传统的生活、学习、工作和思维方式,影响着教育的内容与方法。信息技术的发展使人类社会从工业社会步入了信息时代,信息技术教育成为全世界教育课程改革的热点研究课题。信息技术的发展使各行各业的应用技术不断地升级换代,推动了产业结构的整体调整,使整个社会处在转型的过程中。

1.1.4 信息化发展

1. 信息化概述

信息化是指培育、发展以智能化工具为代表的新的生产力并使之造福于社会的历史过程。社会信息化过程,就是在经济活动和社会活动中建设和完善信息基础设施,发展信息技术和信息产业。信息化社会中信息成为社会的战略资源,信息产业成为国民经济的主导产业,推动国民经济的快速发展。

2. 信息化建设

“信息高速公路”是国家信息基础设施(NII)的俗称,是美国政府于1993年正式提出的。紧跟着日本、加拿大和欧洲的工业发达国家也都决定加速建设“信息高速公路”。“信息高速公路”是一个可以交流各种信息的大容量、高速率的多媒体通信网络。

我国为加快国民经济信息化建设步伐,也在规划信息高速公路,从“三金”(即金桥、金关、金卡)工程起步建设信息高速公路。目前,已经取得了很大的成就,如已建成的四大互联网络:中国互联网、中国教育科研网、中国科技网和中国金桥网。

3. 我国信息化的发展与应用

我国信息化建设发展很快,由中国互联网络信息中心(CNNIC)在2014年1月16日发布的第33次《中国互联网络发展状况统计报告》(以下简称《报告》)显示:

(1)网民规模增长进入平台期,发展主题从“量变”转向“质变”。

《报告》显示,截至2013年12月,中国网民规模达6.18亿,全年新增网民5358万人。互联网普及率为45.8%,较2012年底提升3.7个百分点。综合近年来网民规模数据及其他相关统计,中国互联网普及率逐渐饱和,互联网发展主题从“数量”向“质量”转变,具有互联网在经济社会中地位提升、与传统经济结合紧密、各类互联网应用对网民生活形态影响力度加深等特点。

(2)手机网民数量持续增长,高流量手机应用成亮点。

截至2013年12月,中国手机网民规模达到5亿,年增长率为19.1%,继续保持稳定增长。网民中使用手机上网的人群比例由2012年底的74.5%提升至81.0%,远高于其他设备上网的网民比例,手机网民规模的持续增长促进了手机端各类应用的发展,手机依然是中国网民增长的主要驱动力,成为2013年中国互联网发展的一大亮点。

在3G网络进一步普及、智能手机和无线网络持续发展的背景下,视频、音乐等高流量手机应用拥有越来越多的用户。截至2013年12月,我国手机端在线收看或下载视频的用户数为2.47亿,与2012年底相比增长了1.12亿,增长率高达83.8%,在手机类应用用户规模增长幅度统计中排名第一。用户上网设备向手机端转移、使用基础环境的改善和上网成本的下降三方面是手机端高流量应用使用率激增的主要原因。

(3)社交类综合平台持续升温,网络游戏终端竞争加剧。

《报告》显示,2013年微博、社交网站、论坛等互联网应用的使用率较2012年有所下降。类似即时通信等以社交元素为基础的平台应用则发展稳定:在2013年,整体即时通信用户规模在移动端的推动下提升至5.32亿,较2012年底增长6440万,使用率达86.2%。与传统即时通信工具、社交网站相比,以社交为基础的综合平台不仅拥有更强的通信功能,还增加了信息分享等社交类应用,并为用户提供了诸如支付、金融等内容的综合服务,最大限度地增加了用户粘性,保证了用户规模的持续增长。

与之相比,2013年中国网络游戏用户增长速度明显放缓。网民使用率从2012年的59.5%降至54.7%。网络游戏用户规模为3.38亿,增长数量仅为234万。与网络游戏市场整体增长乏力现状形成鲜明对比的是,手机网络游戏用户的增长十分迅速:截至2013年12月,我国手机网络游戏用户数为2.15亿,较2012年底增长了7594万,年增长率达到54.5%。传统的PC端网络游戏增长乏力,面临手机网络游戏高速增长的挑战。

(4)网购团购规模增速明显,企业电商应用尚待提升。

《报告》显示,2013年以网络购物、团购为主的商务类应用保持较高的发展速度。2013

年, 中国网络购物用户规模达 3.02 亿, 使用率达到 48.9%, 相比 2012 年增长 6.0 个百分点。在商务类应用中, 团购市场的增长最为迅猛: 2013 年团购用户规模达 1.41 亿, 团购的使用率为 22.8%, 相比 2012 年增长了 8.0 个百分点, 使用率年增速达 54.3%, 成为商务类应用的最大亮点。

对比高速增长的网络购物和团购类商务应用, 企业电子商务应用仍然存在提升空间。2013 年, 中国企业在线采购和在线销售的比例分别为 23.5% 和 26.8%, 利用互联网开展营销推广活动的企业比例为 20.9%。不同行业的电子商务应用普及率差距较大, 其中制造业、批发零售业电子商务应用化较为普遍。在企业电子商务应用的规模方面, 与大中型企业相比, 微型企业对电子商务的应用普及还需要进一步加强。

任务二 认识计算机

任务描述: 对计算机, 特别是微型计算机读者都不陌生了, 那么计算机是谁发明的, 什么时间发明的, 到目前为止, 计算机的发展过程是怎样的? 通过本任务的学习, 读者可以了解计算机的发展历史、特点及我国计算机的研究和发展情况, 从而对计算机产生浓厚的兴趣。

任务分析: 计算机从发明到现在只有几十年的历史, 但已经经历了四代, 正在向第五代发展。认识计算机要从计算机的发展历史、特点讲起。

知识点: 计算机的发展、特点和应用

1.2.1 计算机的产生和发展

20 世纪 40 年代中期, 导弹、火箭、原子弹等现代科技的发展, 迫切需要解决很多复杂的数学计算问题, 原有的计算工具已经满足不了要求; 另一方面电子学和自动控制技术的迅速发展, 也为研制电子计算机提供了技术条件。1946 年, 世界上第一台电子数字计算机诞生了, 它是由美国宾夕法尼亚大学的莫克利 (John Mauchly) 和他的学生埃克特 (J.P.Eckert) 博士共同发明的。这台机器的名字叫“电子数字积分计算机”, 简称“埃尼阿克”——ENIAC (Electronic Numerical Integrator And Calculator), 如图 1-1 所示。它由 18000 多只电子管、1500 多个继电器、10000 多只电容和 7000 多个电阻组成, 重 30 吨, 占地约 170 平方米, 功耗 150 千瓦, 共花费 3 年时间才得以完成。它的计算速度为每秒钟完成 5000 次加法运算, 每天稳定工作时间仅有几个小时。虽然说它的功能比不上当前最普通的一台微型计算机, 但它奠定了电子计算机的发展基础, 开辟了计算机科学技术的新纪元, 被称为人类第三次产业革命开始的标志。

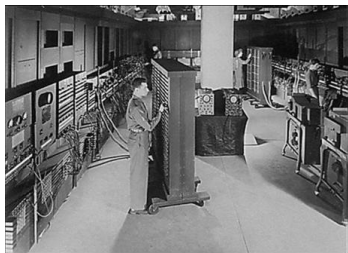


图 1-1 第一台电子管计算机

从第一台电子计算机诞生以来, 短短的半个多世纪里, 计算机发生了翻天覆地的变化。在推动计算机发展的众多因素中, 电子元器件的发展起着决定性的作用。根据电子元器件的发展, 计算机已经经历了四代, 正在走向第五代。

1. 第一代电子计算机 (1946 ~ 1957)

第一代电子计算机又称电子管时代计算机, 采用电子管作为计算机的开关元件, 内存为磁芯, 外存为磁带, 其运算速度只有每秒数千次到数万次; 存储容量小, 只有数千字节; 程序

设计使用机器语言或汇编语言；输入输出设备主要用穿孔的纸带或卡片。由于体积大、功耗大、造价高、使用不便，主要用于军事和科研部门进行数值运算。

2. 第二代电子计算机（1958~1964）

第二代电子计算机是晶体管计算机。其基本特征是用晶体管代替了电子管作为开关元件，大量采用磁芯存储器作内存，磁盘、磁带作外存，大大增加了存储容量，内存容量可达数万字节，运算速度达到每秒数十万次，出现了系统软件（监控程序），提出了操作系统的概念，程序设计不仅可以使使用汇编语言，还可以使用 FORTRAN、COBOL 等高级语言，使编程工作简化方便，具有速度快、寿命长、体积小、重量轻、能耗低等优点。计算机的应用范围从数值计算扩大到数据处理、工业过程控制等领域，并开始进入商业市场。

3. 第三代电子计算机（1965~1974）

第三代电子计算机又称集成电路计算机。其基本特征是使用中小规模集成电路 IC（Integrated Circuit）代替了晶体管。内存为半导体存储器，外存为大容量磁盘，存储容量可达数兆字节，速度达到了每秒几百至上千万次；这个时期，高级程序设计语言有了很大的发展，出现了操作系统和会话语言，推广了结构化程序设计方法，Pascal 语言、C 语言等先后投入使用；计算机制造成本不断降低，为计算机的推广创造了条件，计算机开始广泛应用于各个领域。

4. 第四代电子计算机（1974 年至今）

第四代电子计算机又称大规模集成电路计算机。其基本特征是逻辑器件使用大规模和超大规模集成电路，内存为高集成度的半导体，外存有磁盘、光盘，运算速度每秒达几亿至几百亿次；微型计算机中微处理器的时钟频率可达每秒数百兆赫兹；主存储器（又称内存）的容量已超过数十亿字节，光盘的容量可达到数百兆字节。速度在加快，容量在加大，体积在下降，性能价格比在提高，已成为目前计算机发展的趋势。操作系统不断完善，应用软件已成为现代工业的一部分，计算机的发展进入了以计算机网络为特征的时代。

5. 新一代计算机

目前，新一代的计算机正在设想和研制阶段。新一代的计算机应以处理知识信息为主，具有推理、联想、学习和应用知识等人工智能方面的能力，如生物计算机、量子计算机、神经网络计算机等。

1.2.2 我国计算机的发展

我国从 1956 年开始研制计算机，1960 年我国第一台自行设计的通用电子计算机 107 机问世。1964 年我国研制成功了大型通用电子计算机 119 机。1978 年我国开始了巨型计算机的研制工作，1983 年 12 月 22 日，我国第一台每秒运算一亿次以上的巨型计算机——“银河-I”（如图 1-2 所示）由国防科技大学计算机研究所在长沙研制成功。“银河-I”的研制成功填补了我国巨型计算机的空白，标志着中国进入了世界研制巨型计算机的行列，成为由中国科技人员自行设计的第一个每秒向量运算 1 亿次的巨型计算机系统。

1992 年研制成功每秒运算十亿次的“银河-II”巨型计算机，如图 1-3 所示。1993 年 6 月 22 日，“银河-II”巨型计算机通过了国家的鉴定验收，从而宣布了我国第一台 10 亿次巨型计算机的问世。

1993 年，我国自行研制的第一台高性能计算机曙光一号诞生。3 天后，发达国家即解除了 10 亿次计算机对中国的禁运。

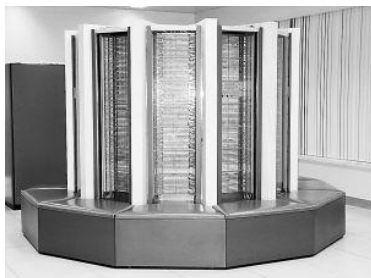


图 1-2 银河-I



图 1-3 银河-II

1997 年,曙光 1000 A 成功落户辽河油田,首次实现了国产高性能计算机商品化。随后,气象、生物制药、化工、航空航天、精密仪器等行业开始大规模应用高性能计算机,并因此大大提升了自己的产出成果。

1997 年 6 月 19 日由国防科技大学计算机研究所研制的“银河-III”百亿次巨型计算机系统通过国家技术鉴定。1999 年 9 月,我国每秒运算 3840 亿次的“神威”高性能计算机问世,它的反应速度在当时世界前 500 台投入商业运行的巨型计算机中排在第 48 位。

2002 年 8 月 9 日,我国联想集团推出具有自主知识产权与核心技术的超级计算机“深腾 LSSC-2”,实际速度达到每秒 1.027 亿次浮点运算。这是我国第一个实际运算速度达到万亿次水平的计算机产品,在每年两次公布的全世界当时安装的实际运算速度最快的前 500 台计算机名单中排在第 24 位。

2005 年 4 月,中国首个拥有自主知识产权的高性能 CPU “龙芯 2 号”正式亮相,从而打破国外在该领域长达数十年的技术垄断。

2007 年 12 月 26 日我国首台采用国产高性能通用处理器芯片“龙芯 2F”和其他国产器件、设备和技术的万亿次高性能计算机“KD-50-I”通过了专家委员会的鉴定,在中国科技大学研制成功。“KD-50-I”万亿次机的研制成功,确定了国产高性能通用处理器在高端并行机应用中的核心地位。

2008 年 9 月 16 号,我国首台超百万亿次超级计算机“曙光 5000A”在天津成功下线,如图 1-4 所示。

“曙光 5000A”共有约三万颗计算核心,浮点运算处理能力可达每秒 230 万亿次,内存 120TB,存储能力超过 700TB。

2009 年 10 月 29 日,我国第一台国产千亿次“天河一号”超级计算机在湖南长沙问世,如图 1-5 所示。



图 1-4 曙光 5000A



图 1-5 天河一号

每秒钟 1206 万亿次的峰值速度和每秒 563.1 万亿次的 Linpack^①实测性能使这台名为“天河一号”的计算机位居同日公布的中国超级计算机前 100 强之首,也使中国成为继美国之后世界上第二个能研制千万亿次超级计算机的国家。

2010 年 6 月 1 日,具有自主知识产权的我国第一台实测性能超千万亿次的“星云”超级计算机系统由曙光公司正式发布。理论峰值 3000 万亿次,实测峰值每秒达 1271 万亿次,是国内第一台、世界第三台实测性能超千万亿次的超级计算机。

2012 年 2 月 24 日,科技部高技术中心组织验收专家组在国家超级计算深圳中心对中国科学院计算技术研究所、曙光信息产业(北京)有限公司等单位承担的“十一五”国家 863 计划“高效能计算机及网络服务环境”重大项目的课题“曙光 6000 千万亿次高效能计算机系统”进行了现场验收,如图 1-6 所示。



图 1-6 曙光 6000 验收现场

对比 20 年前曙光一号亿次计算的量级,曙光 6000 已经到了千万亿次的量级,增长 400 多万倍。而技术与应用结合、科研与产业结合,让曙光的超算产品不只是能比拼速度的机器,更承担起普惠于民的重任。

标志着我国超级服务器技术和产品走向成熟的曙光 2000、使中国超级计算机首次进入世界前十的曙光 4000A、开启中国高性能计算机新纪元的曙光“星云”……不断刷新的成果让世界一次次为“中国速度”惊叹,更推动中国高性能计算产业步入“快车道”。

从助力描绘北京奥运会开幕式上的“中国卷轴”,到帮助汽车厂商降低新车开发成本,从基因测序、药物分子设计,到智能化精细气象预报……曙光在文化创意、高端装备制造、新材料、新能源、节能环保、金融等行业得到广泛应用。

2013 年岁末,嫦娥三号成功落月令世界瞩目。这完美一落离不开一位幕后英雄——曙光高性能计算机,正是曙光对嫦娥轨道路线、飞行姿态等精准、快速的计算,为实现中国探月梦想提供了坚实支撑。

1.2.3 计算机的特点

计算机作为一种通用的信息处理工具,主要有以下特点:

(1) 运算速度快。

计算机的运算速度已经达到了每秒千万亿次,并以每隔几个月提高一个数量级的速度在快速发展。这使得过去需要几年甚至几十年才能完成的工作,现在只要几天、几小时甚至更短的时间就可以完成,极大地提高了工作效率。

(2) 计算精度高。

计算机的计算精度在理论上不受限制。如用计算机计算圆周率 π 的值可以精确到小数点后几亿位。

(3) 超强的记忆能力。

计算机具有一个能存储大量信息的存储器。它类似于人类的大脑,可以“记忆”大量的

^① Linpack 是 Linear system package (线性系统软件包) 的缩写, Linpack 测试是现在最为知名的评判一个系统计算力的方法。

原始数据、中间结果、运算指令等,需要时可以快速调出使用。存储器不但能够存储大量的信息,而且能够快速准确地存入或取出这些信息。

(4) 具有逻辑判断能力。

计算机可以对各种数据或信息进行逻辑推理和判断。它可以根据预先编写的程序自动选择执行程序,还可以进行逻辑推理,具有感知和识别能力以及推理判断能力,进而可以用计算机模仿人的智能活动等。

(5) 具有实时通讯能力。

计算机网络可以消除地理位置差异,使分布在不同地理位置的计算机及其外围设备组成一个互相通讯的系统,可以使数据信息实时地传输和交换。

(6) 自动化能力强。

计算机是由程序控制其操作过程的。只要根据应用的需要编制好程序并输入计算机,计算机就能自动地、连续地工作,完成预定的任务。

1.2.4 计算机的应用

计算机的应用领域已渗透到社会的各行各业,正在改变着传统的工作、学习和生活方式,推动着社会的发展。计算机的主要应用领域包括以下几个方面:

(1) 科学计算。

科学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。在现代科学技术工作中,科学计算问题是大量的和复杂的。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(或信息处理)。

数据处理是指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。计算机具有逻辑判断和数据处理能力,可以存储大量的信息,并进行信息处理。如银行管理系统、财务管理系统、人事管理系统等,节省了大量的人力,大大提高了工作效率。

(3) 过程控制(或实时控制)。

过程控制也叫实时控制,就是用计算机对连续工作的控制对象进行自动控制或自动调节。利用计算机进行过程控制,不仅大大提高了自动化水平,而且也提高了控制的及时性和准确性。

(4) 计算机辅助系统。

计算机辅助系统是指用计算机来辅助人类进行一部分工作,包括计算机辅助设计 CAD (Computer Aided Design)、计算机辅助制造 CAM (Computer Aided Manufacturing)、计算机辅助测试 CAT (Computer Aided Testing)、计算机辅助工程 CAE (Computer Aided Engineering)、计算机辅助教学 CAI (Computer Assisted Instruction) 等。

(5) 人工智能。

人工智能,简称 AI,也是计算机应用的一个重要领域。所谓智能就是利用计算机模拟人类的智能活动,使计算机对知识具有“推理”和“学习”的功能,让计算机可以为人类的决策提供帮助。如专家系统、智能机器人等。人工智能将为计算机的硬件和软件带来革命,最终导致智能计算机的出现。

(6) 网络应用。

计算机技术与现代通信技术的结合构成了计算机网络。计算机网络是利用通信设备和线

路将地理位置不同、功能独立的多个计算机系统互联起来,以实现世界范围内网络中的信息资源、软硬件资源的共享和信息交换。可谓是“一线联五洲”,这是传统通信手段难以达到的。

(7) 多媒体技术。

随着计算机技术的发展,人们已把文本、音频、视频、动画、图形和图像等各种媒体综合起来,构成一种全新的概念——多媒体技术,并且发展迅速,目前已广泛地应用到各行各业。

(8) 电子商务。

电子商务(Electronic Commerce)最早产生于20世纪60年代,发展于20世纪90年代,一般指的是在网络上通过计算机进行业务通信和交易处理,实现商品和服务的买卖以及资金的转账,同时还包括企业公司之间及其内部借助计算机及网络通信技术能够实现的一切商务活动,也就是通过网络进行的生产、营销、销售和流通活动,不仅包括在互联网上的交易,而且也包括利用信息技术来降低商务成本、增加流通价值和创造商业机遇的所有商务活动。

商务活动的核心是信息活动,在正确的时间和正确的地点与正确的人交换正确的信息是电子商务成功的关键。电子商务的显著特点是突破了时间和地点的限制、低成本、高效率、虚拟现实、功能全面、使用更灵活和更加安全有效。

电子商务的运行模式按照电子商务交易主体之间的差异可以有多种不同的模式,其中最典型的运行模式有:商家—商家模式(Business to Business, B2B)、商家—消费者模式(Business to Customer, B2C)、消费者—消费者模式(Customer to Customer, 简称C2C)。

(9) 电子政务。

电子政务就是政府机构运用现代计算机技术和网络技术,将管理和服务的职能转移到网络上去,实现政府组织结构和 workflows 的重组优化,超越时间、空间和部门分隔的制约,向全社会提供高效优质、规范透明和全方位的管理与服务。它开辟了推动社会信息化的新途径,创造了政府实施产业政策的新手段。电子政务的出现有利于政府转变职能,提高运作的效率。

电子政务的特点是转变政府工作方式,提高政府科学决策水平,优化信息资源配置,借助信息技术降低管理和服务成本。

从电子政务服务的对象看,电子政务的主要内容包括:政府—政府电子政务(Government to Government, G2G)、政府—企业电子政务(Government to Business, G2B)、政府—公民电子政务(Government to Citizen, G2C)。

1.2.5 计算机的分类

对计算机分类有很多方法,依据IEEE(美国电气和电子工程师协会)的划分标准,计算机可分为巨型计算机、大型计算机、小型计算机、工作站和微型计算机。

1. 巨型计算机

巨型计算机又称超级计算机,是一种运算速度快、处理信息流量大、可容纳的用户多、价格昂贵的计算机,主要用于国防、科学研究等高技术领域,它是衡量一个国家经济实力和科技水平的重要标志。

2. 大型计算机

大型计算机运算速度低于巨型计算机,具有很强的数据处理和管理能力,速度相对较快,目前主要应用于高等学校、银行和科研机构等。

3. 小型计算机

小型计算机结构相对简单,价格相对大型计算机来说较低,可以适应一般用户的需要。

4. 工作站

工作站是一种高档次的微型计算机, 功能强、速度快, 能够进行较多专业化的工作, 具有较强的联网能力。

5. 微型计算机

微型计算机也称为个人计算机, 它体积小、重量轻、价格便宜、功能齐全、设计先进、更新速度快、使用方便, 广泛应用于个人用户, 具有极强的生命力。

1.2.6 计算机的发展趋势

计算机技术是当今世界发展最快的科学技术之一, 从计算机的研究看其发展, 未来的计算机主要趋向于巨型化、微型化、多媒体化、网络化、智能化和非冯·诺依曼化。

1. 巨型化

巨型化是指发展高速、大存储量和功能强大的超级计算机, 主要应用在高科技领域, 它是一个国家科学技术和工业发展的重要标志。

2. 微型化

微型化是指利用超大规模集成电路, 使计算机的集成度越来越高, 元器件越来越小, 而使计算机速度快、功能强、可靠性提高、耗能量减小、体积小、重量轻。计算机的微型化已成为计算机发展的一个重要方向。

3. 多媒体化

多媒体化是指对声音、图像的处理具有比现在更强大的能力。

4. 网络化

网络化是指用现代通讯技术和计算机技术把分布在不同地点的计算机互联起来, 组成一个规模更大、功能更强的可以互相通信的网络结构, 从而实现网络中的资源共享。事实证明, 现代网络技术已经成为计算机技术中不可缺少的内容, 计算机网络化是计算机发展的又一个趋势。

5. 智能化

智能化是指使计算机具有模拟人的感觉和思维过程的能力。未来的新一代计算机是一种智能化的计算机, 是一种能思维的计算机。

6. 非冯·诺依曼化

随着计算机应用领域的不断扩大, 采用存储方式进行工作的冯·诺依曼化的计算机逐渐显示出其局限性, 从而出现了新思维, 这就是非冯·诺依曼化计算机的构想。

任务三 计算机中的数制及其转换

任务描述: “ $1+1=10B$ ”、“ $(10)_2+(11)_2=(101)_2$ ”正确吗? 通过本任务的学习你将解除疑问。本任务主要讲解计算机中数据的表示方法及各种数制间的转换、计算方法等。

任务分析: 数制有十进制、二进制、八进制和十六进制等, 人们常用的是十进制, 而计算机中用的是二进制。学习数制及其转换要从数制的概念及转换规则开始。

知识点: 数制的概念 不同进制数据间的转换

1.3.1 数制的基本概念

1. 进位计数制

进位计数制是指用进位的方法进行计数的数制。数有不同的进位计数制，日常生活中使用的多为十进制，而目前使用的计算机是一种电子设备，主要由开关元件构成，故只能识别由“0”和“1”构成的二进制代码。但用二进制数表示一个较大的数时，既冗长又难以记忆，为了阅读和书写方便，或适应某些特殊场合的需要，在计算机中有时也采用十六进制或八进制。

要理解数制，必须先理解基数和位权两个概念。

基数指用该进制表示数时所用到的数字符号的个数，常用“R”表示，称R进制。如十进制数用十个数字来表示大小不同的数，因而基数为10。依此类推，十六进制的基数为16，八进制的基数为8，二进制的基数为2。

位权指数码在不同位置上的权值。每一种进制数中的数字符号所在的位置叫数位，不同数位有不同的“位权”，用一个以基数为底的指数来表示，即 R^i ，R代表基数，i是数位的序号。一般规定整数部分个位为0，十位为1，……，依次增1；小数部分小数点右面的第一位为-1，第二位为-2，……，依次减1。如十进制数123.45，基数为10，1的位权为 10^2 ，2的位权为 10^1 ，3的位权为 10^0 ，4的位权为 10^{-1} ，5的位权为 10^{-2} 。

2. 计算机中的数制

(1) 十进制数。

十进制数是用0、1、2、3、4、5、6、7、8、9共十个数字来表示大小不同的数，基数为10，它的计数规则是“逢十进一，借一当十”，它的权是以10为底的幂。按位权展开的形式是（以1234.56为例）：

$$1234.56=1\times 10^3+2\times 10^2+3\times 10^1+4\times 10^0+5\times 10^{-1}+6\times 10^{-2}$$

(2) 二进制数。

二进制数是用0和1共两个数字表示大小不同的数，基数为2，它的计数规则是“逢二进一，借一当二”，它的权是以2为底的幂，按位权展开的形式是（以(1101.11)₂为例）：

$$(1101.11)_2=1\times 2^3+1\times 2^2+0\times 2^1+1\times 2^0+1\times 2^{-1}+1\times 2^{-2}$$

(3) 八进制数。

八进制数是由0、1、2、3、4、5、6、7共八个数字来表示大小不同的数，基数为8，它的计数规则是“逢八进一，借一当八”，它的权是以8为底的幂，按位权展开的形式是（以(1261.11)₈为例）：

$$(1261.11)_8=1\times 8^3+2\times 8^2+6\times 8^1+1\times 8^0+1\times 8^{-1}+1\times 8^{-2}$$

(4) 十六进制数。

十六进制数是使用0、1、2、3、4、5、6、7、8、9、A、B、C、D、E、F共十六个符号来表示大小不同的数，其中字母A、B、C、D、E、F分别表示10、11、12、13、14、15，基数是16，它计数规则是“逢十六进一，借一当十六”，它的权是以16为底的幂，按位权展开的形式是（以(2D5F.2A)₁₆为例）：

$$(2D5F.2A)_{16}=2\times 16^3+13\times 16^2+5\times 16^1+15\times 16^0+2\times 16^{-1}+10\times 16^{-2}$$

以上介绍的几种数制除了用在括号外面加数字下标的形式表示外，还可以在数字后面加写相应的英文字母作为标识：B（二进制）、O（八进制）、D（十进制）、H（十六进制）。十进制的括号和字母可以省略。

例如, 二进制数 100 可写成 100B 或 $(100)_2$, 八进制数 100 可写成 100O 或 $(100)_8$, 十进制数 100 可写成 100D 或 $(100)_{10}$, 十六进制数 100 可写成 100H 或 $(100)_{16}$ 。

1.3.2 不同数制之间的转换

将数从一种数制转换为另一种数制的过程叫数制间的转换。

1. 二进制数、八进制数、十六进制数转换为十进制数

对于任何一个二进制数、八进制数、十六进制数转换为十进制数, 只需把各数位的值乘以该位位权, 再按十进制加法相加即可。这种方法也叫“位权法”。

【例 1-1】将二进制数 101.11 转换为十进制数。

$$(101.11)_2 = 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 1 \times 2^{-2} = 4 + 1 + 0.5 + 0.25 = 5.75$$

【例 1-2】将八进制数 136.4 转换为十进制数。

$$(136.4)_8 = 1 \times 8^2 + 3 \times 8^1 + 6 \times 8^0 + 4 \times 8^{-1} = 64 + 24 + 6 + 0.5 = 94.5$$

【例 1-3】将十六进制数 2A.C 转换为十进制数。

$$(2A.C)_{16} = 2 \times 16^1 + 10 \times 16^0 + 12 \times 16^{-1} = 32 + 10 + 0.75 = 42.75$$

2. 十进制数转换为非十进制数

将十进制数转换为非十进制数, 方法是: 将整数部分采用“除基取余倒读”法, 小数部分采用“乘基取整正读”法, 再把两部分组合起来, 就可以得到对应的结果。

【例 1-4】将十进制数 75.375 转换为二进制数。

分析: 整数部分转化为二进制数, 应除以 2 倒取余法。

		余数	
2	75	1	最低位
2	37	1	
2	18	0	
2	9	1	
2	4	0	
2	2	0	
2	1	1	最高位
0			

小数部分转化为二进制数, 应乘以 2 取整法。

		所取整数	
×	0.375	2	最高位
	0.750	0	
×	0.5	2	
	1.500	1	
×	1.0	2	最低位
	2.0	0	

所得结果为: $(75.375)_{10} = (1001011.011)_2$

依此类推, 可以完成将 75.375 转化为八进制数和十六进制数。

但是必须注意的是,在有些情况下,十进制小数不能精确地转化为非十进制小数,例如 0.33。在这种情况下,只能根据需要的精度对十进制小数作近似转换。

3. 二进制数与八进制数的相互转换

转换规则是:

(1) 二进制数转换成八进制数:以小数点为中心,分别向左、向右每三位划分成一组,不足三位的分别向高位或低位以 0 补足,每组分别转化为对应的一位八进制数,最后将这些数字从左到右连接起来即可。

(2) 八进制数转换成二进制数:将每一位八进制数转换成对应的三位二进制数,不足三位分别向高位以 0 补足,将这些二进制数从左到右连接起来即可。

【例 1-5】将二进制数 10010011.1011 转换为八进制数,八进制数 672.25 转换为二进制数。

$$\begin{array}{ccccccc} \underline{010} & 010 & 011 & . & 101 & \underline{100} & \\ 2 & 2 & 3 & . & 5 & 4 & \\ (10010011.1011)_2 = (223.54)_8 & & & & & & \end{array} \quad \begin{array}{ccccccc} 6 & 7 & 2 & . & 2 & 5 & \\ 110 & 111 & \underline{010} & . & \underline{010} & 101 & \\ (672.25)_8 = (110111010.010101)_2 & & & & & & \end{array}$$

4. 二进制数与十六进制数的相互转换

二进制数与十六进制数之间的转换规则和二进制数与八进制数之间的转换规则类似。

(1) 二进制数转换成十六进制数:以小数点为中心,分别向左、向右每四位划分成一组,不足四位分别向高位或低位以 0 补足,每组分别转化为对应的一位十六进制数,最后将这些数字从左到右连接起来即可。

(2) 十六进制数转换成二进制数:将每一位十六进制数转换成对应的四位二进制数,不足四位分别向高位以 0 补足,将这些二进制数从左到右连接起来即可。

【例 1-6】将二进制数 11111010011.101101 转换为十六进制数,十六进制数 3B5.6A 转换为二进制数。

$$\begin{array}{ccccccc} \underline{0111} & 1101 & 0011 & . & 1011 & \underline{0100} & \\ 7 & D & 3 & . & B & 4 & \\ (11111010011.101101)_2 = (7D3.B4)_{16} & & & & & & \end{array} \quad \begin{array}{ccccccc} 3 & B & 5 & . & 6 & A & \\ \underline{0011} & 1011 & \underline{0101} & . & \underline{0110} & 1010 & \\ (3B5.6A)_{16} = (1110110101.0110101)_2 & & & & & & \end{array}$$

提示:二进制数与八进制数、十六进制数之间的转换虽然很简单,但一定要注意在转换过程中 0 的补充。

表 1-1 列出了十进制、二进制、八进制和十六进制之间的换算关系。

表 1-1 十进制、二进制、八进制和十六进制的换算关系

十进制	二进制	八进制	十六进制
0	0000	0	0
1	0001	1	1
2	0010	2	2
3	0011	3	3
4	0100	4	4
5	0101	5	5
6	0110	6	6
7	0111	7	7

续表

十进制	二进制	八进制	十六进制
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

1.3.3 二进制数的运算规则

在计算机中, 二进制运算分为算术运算和逻辑运算两类。

1. 算术运算

算术运算包括加、减、乘、除 4 种运算, 运算规则为:

加法规则: $0+0=0$; $0+1=1$; $1+0=1$; $1+1=10$ (向高位有进位)

减法规则: $0-0=0$; $1-0=1$; $1-1=0$; $10-1=1$ (向高位有借位)

乘法规则: $0\times 0=0$; $0\times 1=0$; $1\times 0=0$; $1\times 1=1$

除法规则: $0\div 1=0$; $1\div 1=1$

【例 1-7】 $(1100)_2 + (101)_2 = (10001)_2$

$$\begin{array}{r}
 1 \quad 1 \quad 0 \quad 0 \\
 + \quad \quad 1 \quad 0 \quad 1 \\
 \hline
 1 \quad 0 \quad 0 \quad 0 \quad 1
 \end{array}$$

【例 1-8】 $(1110)_2 - (101)_2 = (1001)_2$

$$\begin{array}{r}
 1 \quad 1 \quad 1 \quad 0 \\
 - \quad \quad 1 \quad 0 \quad 1 \\
 \hline
 1 \quad 0 \quad 0 \quad 1
 \end{array}$$

2. 逻辑运算

逻辑运算包括逻辑与、逻辑或、逻辑非和逻辑异或 4 种, 运算规则如下:

(1) 逻辑与运算 (AND): 逻辑与运算又称逻辑乘法, 运算符可以用 “ $\times$ ” 或者 “ $\wedge$ ”, 运算规则如下:

$$\begin{array}{ll}
 0\times 0=0 & \text{或} \quad 0\wedge 0=0 \\
 0\times 1=0 & 0\wedge 1=0 \\
 1\times 0=0 & 1\wedge 0=0 \\
 1\times 1=1 & 1\wedge 1=1
 \end{array}$$

仅当两数均为 1 时, 结果方为 1, 否则为 0。

(2) 逻辑或运算 (OR): 逻辑或运算又称逻辑加法, 运算符可以用 “+” 或者 “ $\vee$ ”, 运算规则如下:

$$0+0=0 \quad \text{或} \quad 0\vee 0=0$$

$$0+1=1 \quad 0\vee 1=1$$

$$1+0=1 \quad 1\vee 0=1$$

$$1+1=1 \quad 1\vee 1=1$$

仅当两数均为 0 时，结果方为 0，
否则为 1。

(3) 逻辑非运算 (NOT): 逻辑非又叫逻辑否定，运算符是在参加运算的逻辑量上加一横线，如 $\bar{0}$ ，运算规则如下：

$$\bar{0}=1; \quad \bar{1}=0$$

按位取反，0 的非为 1，1 的非为 0。

(4) 逻辑异或运算 (XOR): 逻辑异或运算即实现按位加的功能，运算符可以用 $\odot$ ，运算规则如下：

$$0\odot 0=0; \quad 0\odot 1=1;$$

$$1\odot 0=1; \quad 1\odot 1=0$$

仅当两个逻辑值不相同，结果
才为 1，否则为 0。

任务四 计算机中的数据编码

任务描述：人类最初使用十进制计数系统是因为人们有 10 个手指头，现实的生活、学习和工作中人们也早已习惯于用十进制数，而计算机能够直接识别和处理的只有 0 和 1 两个符号所表示的二进制数。这样就导致了人和计算机之间的通信问题。如何解决这个问题？通过本任务的学习，你将得到正确的答案。

任务分析：计算机中的数据编码有多种，如 BCD 编码、ASCII 字符编码、汉字编码等。学习数据编码，要从数据编码原则开始。

知识点：BCD 码 ASCII 码 数据单位

1.4.1 二进制编码

计算机只认识二进制，而人们习惯于使用十进制，所以在某些情况下也希望计算机能直接处理十进制形式的数据。此外，现代计算机不仅要处理数值领域的问题，还需要处理大量非数值领域的问题，如声音、图形、文字、符号等，这就要求计算机还应能够识别和处理文字、字符和各种符号，如：

字母——26 个英文字母（大小写）：A, B, ..., Z, a, b, ..., z

专用符号——+、-、*、/、\$、%、↑、√等

控制字符——CR（回车）、LF（换行）、BEL（响铃）等

所有这些字符、符号以及十进制数最终都必须转换为二进制格式的代码才能为计算机所处理，即字符和十进制数都必须用若干二进制码来表示，这就是信息和数据的二进制编码。下面以常用的 BCD 码和 ASCII 码为例来讲解各种编码的特点。

1. BCD 数字编码

人们习惯于用十进制来计数，而计算机中则采用二进制，因此，为了解决二进制数表示十进制的问题，引入了 BCD (Binary Coded Decimal) 码，又称为“二—十进制编码”。它的特点是保留了十进制的权，而数字用 0 和 1 的组合编码来表示。

“二—十进制编码”最常用的是 8421 编码，方法是用 4 位二进制数表示一位十进制数，其 4

位二进制编码的每一位都有特定的权值，自左向右分别为： $2^3=8$ 、 $2^2=4$ 、 $2^1=2$ 、 $2^0=1$ ，故称其为 8421 编码。例如 129 用 8421 编码表示的 BCD 码为 0001 0010 1001。它看起来是一个二进制数，实际上它是一个 3 位十进制数。表 1-2 给出了十进制数与 8421BCD 编码的对应关系。

表 1-2 十进制数与 8421BCD 编码的对应关系

十进制数	8421BCD 码	十进制数	8421BCD 码	十进制数	8421BCD 码
0	0000	10	0001 0000	100	0001 0000 0000
1	0001	11	0001 0001	101	0001 0000 0001
2	0010	12	0001 0010	102	0001 0000 0010
3	0011	13	0001 0011	103	0001 0000 0011
4	0100	14	0001 0100	104	0001 0000 0100
5	0101	15	0001 0101	200	0010 0000 0000
6	0110	16	0001 0110	211	0010 0001 0001
7	0111	17	0001 0111	222	0010 0010 0010
8	1000	18	0001 1000	876	1000 0111 0110
9	1001	19	0001 1001	999	1001 1001 1001

BCD 码的计数规律与十进数相同，即“逢十进一”。在书写时，每一个 4 位写在一起，以表示十进制的一位，结尾处加标记 BCD。如(0001 0010 1001)_{BCD}表示十进制数 129。

2. ASCII 字符编码

ASCII 字符编码是“美国标准信息交换代码”(American Standard Code for Information Interchange)的简称，用于给西文字符编码。有 7 位 ASCII 码和 8 位 ASCII 码两种。7 位 ASCII 码称为标准的 ASCII 码，用一个字节表示一个字符，并规定最高位为 0，可以表示 128 个字符，包括 0~9、52 个大小写英文字母，以及各种标点符号、运算符号和控制命令符号等，目前在国际上广泛流行；8 位 ASCII 码称为扩展 ASCII 码，是美国 IBM 公司把标准 ASCII 码的位数增加一位，用 8 位二进制数构成一个字符编码，共有 256 个符号。扩展后的 ASCII 码除了原先的 128 个字符之外，又增加了一些常用的科学符号和表格线条，如表 1-3 和表 1-4 所示。

表 1-3 标准 ASCII 码字符表

ASCII 值		控制 字符	ASCII 值		控制 字符	ASCII 值		控制 字符	ASCII 值		控制 字符
十进制	十六进制		十进制	十六进制		十进制	十六进制		十进制	十六进制	
0	00	NUL	32	20	sp	64	40	@	96	60	`
1	01	SOH	33	21	!	65	41	A	97	61	a
2	02	STX	34	22	”	66	42	B	98	62	b
3	03	ETX	35	23	#	67	43	C	99	63	c
4	04	EOT	36	24	\$	68	44	D	100	64	d
5	05	ENQ	37	25	%	69	45	E	101	65	e
6	06	ACK	38	26	&	70	46	F	102	66	f
7	07	BEL	39	27	,	71	47	G	103	67	g

续表

ASCII 值		控制 字符	ASCII 值		控制 字符	ASCII 值		控制 字符	ASCII 值		控制 字符
十进制	十六进制		十进制	十六进制		十进制	十六进制		十进制	十六进制	
8	08	BS	40	28	(	72	48	H	104	68	h
9	09	HT	41	29	)	73	49	I	105	69	i
10	0A	LF	42	2A	*	74	4A	J	106	6A	j
11	0B	VT	43	2B	+	75	4B	K	107	6B	k
12	0C	FF	44	2C	,	76	4C	L	108	6C	l
13	0D	CR	45	2D	-	77	4D	M	109	6D	m
14	0E	SO	46	2E	.	78	4E	N	110	6E	n
15	0F	SI	47	2F	/	79	4F	O	111	6F	o
16	10	DLE	48	30	0	80	50	P	112	70	p
17	11	DC1	49	31	1	81	51	Q	113	71	q
18	12	DC2	50	32	2	82	52	R	114	72	r
19	13	DC3	51	33	3	83	53	X	115	73	s
20	14	DC4	52	34	4	84	54	T	116	74	t
21	15	NAK	53	35	5	85	55	U	117	75	u
22	16	SYN	54	36	6	86	56	V	118	76	v
23	17	TB	55	37	7	87	57	W	119	77	w
24	18	CAN	56	38	8	88	58	X	120	78	x
25	19	EM	57	39	9	89	59	Y	121	79	y
26	1A	SUB	58	3A	:	90	5A	Z	122	7A	z
27	1B	ESC	59	3B	;	91	5B	[	123	7B	{
28	1C	FS	60	3C	<	92	5C	/	124	7C	
29	1D	GS	61	3D	=	93	5D	]	125	7D	}
30	1E	RS	62	3E	>	94	5E	^	126	7E	~
31	1F	US	63	3F	?	95	5F	—	127	7F	DEL

表 1-4 标准 ASCII 码字符表的说明

NUL 空	VT 垂直制表	SYN 空转同步
SOH 标题开始	FF 走纸控制	ETB 信息组传送结束
STX 正文开始	CR 回车	CAN 作废
ETX 正文结束	SO 移位输出	EM 纸尽
EOY 传输结束	SI 移位输入	SUB 换置
ENQ 询问字符	DLE 空格	ESC 换码
ACK 承认	DC1 设备控制 1	FS 文字分隔符
BEL 报警	DC2 设备控制 2	GS 组分分隔符
BS 退一格	DC3 设备控制 3	RS 记录分隔符
HT 横向列表	DC4 设备控制 4	US 单元分隔符
LF 换行	NAK 否定	DEL 删除

3. 汉字编码

计算机实现汉字处理的前提条件是对汉字进行编码。汉字个数很多,而且每个汉字要用一个编码,因此汉字编码要比英文字符的编码复杂得多。通过键盘输入的汉字需要使用由英文字符组成的输入码,然后转换成由数字组成的交换码,再由交换码转换为机内码才能在计算机中存储、处理而不会与西文编码冲突;输出汉字时还要再把机内码转换为字形码送到输出设备输出。图 1-7 为计算机中汉字处理原理图。

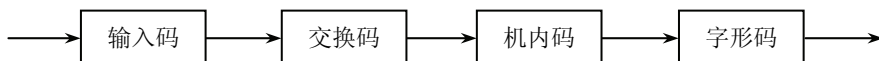


图 1-7 计算机中汉字处理原理图

(1) 汉字交换码。

1981 年,我国颁布了《信息交换用汉字编码字符集·基本集》(GB2312—80)。它是汉字交换码的国家标准,又称为“国标码”。共收集了汉字、图形、符号等 7445 个字符,其中最常用的一级汉字 3755 个,次常用的二级汉字 3008 和 682 个符号。

1995 年 12 月,国家发布了汉字扩展内码规范——GBK 编码方案,该方案共收录了 21003 个汉字以及各类符号共 883 个,并完全兼容 GB2312—80。

国标码规定,每个汉字字符或其他字符由 2 个字节组成,每个字节的最高位为 0,其余 7 位为不同的码值。例如,汉字“北”的编码为 1717,用 BCD 码表示为:0001 0111 0001 0111。

(2) 汉字机内码。

机内码是指一个汉字被计算机内部系统进行存储、处理和传输时使用的编码,为了使汉字处理与 ASCII 处理不发生矛盾,在计算机内部表示汉字时把交换码(国标码)的两个字节的最高位改为 1,称为机内码。这样,当两个字节的最高位都为 1 时,代表一个汉字。当字节的最高位为 0 时,则为 ASCII 码字符。例如,汉字“北”的国标码为 1717,而它的机内码为 9797,用 BCD 码表示为:1001 0111 1001 0111。

因此,汉字的交换码与机内码是不同的,而 ASCII 码的交换码与机内码是相同的。

机内码是计算机内处理汉字信息时所用的汉字代码。尽管各种输入法的输入码不同,但其对应的内码却是相同的。也就是说,一个汉字的机内码是唯一的。汉字机内码与国家标准 GB2312—80 汉字字符集有简明的对应关系。

(3) 汉字输入码。

汉字输入码是为用户通过计算机外部输入汉字而编制的汉字编码,简称外码。汉字输入码面向用户,它的编码原则是简单、易接受,便于学习、记忆,有利于提高输入速度。目前用得较多的有以下 4 类:

- 数字编码(也称流水码):如区位码、电报码、国标码,特点是:整齐、简洁,没有重码,编码和汉字没有直接的对应关系,但难以记忆,一般用于输入一些特殊符号。
- 音码:根据汉字的读音来确定汉字的输入编码,如全拼、智能 ABC、双拼、搜狗拼音、QQ 拼音等。它的特点是:以汉字的拼音为基础,容易掌握和普及,但重码率高,影响输入速度。它是目前使用最为广泛的输入码。
- 形码:根据汉字的字形、结构特征和一定的编码规则来确定汉字的输入编码,如五笔字型、表形码、笔划码、首尾码等。它的特点是:根据汉字的字形编码,重码少,输入速度快,但需要专门的学习才能掌握。

- 音形码：结合汉字的读音和字形而对汉字进行的编码，如自然码。它的特点是：形码部分较简单，记忆量较形码少，重码少，输入速度快，但音形码的掌握需要专门的学习。

(4) 汉字字形码。

字形码又称汉字字模，是用来将汉字显示到屏幕上或打印到纸上所需要的图形数据，是表示汉字字形信息的编码，是汉字的输出形式。目前常用的汉字字形码有点阵码和矢量码。

- 点阵码：输出时采用字符点阵 $M \times N$ 像素阵列来表示汉字，字符点阵不同，一个汉字需要的存储空间也不同。常用的点阵有 16×16 、 24×24 、 32×32 或更高。点阵越大，字形质量越高，所占存储空间也越大，而且点阵字形缩放困难、容易失真。
- 矢量码：用一组数学矢量来记录汉字的外形轮廓，矢量码记录的字体称为矢量字体或轮廓字体。它的特点是节省存储空间、字形美观、容易放大和缩小、字形质量较高等。

1.4.2 计算机中的符号位表示

在计算机中，所有的数据都用二进制的形式来表示。一个二进制数同时包含符号位和数值位两部分，通常规定一个数的最高位作为符号位，用“0”表示正，“1”表示负，余下各位表示数值。把数据符号位数码化后的二进制表示的数称为“机器数”，把在机器外存放的由正负号表示的数称为真值。

1.4.3 计算机中的数据单位

计算机中数据的存储单位通常有位、字节、字、字长等。

1. 位 (bit)

位，是计算机中数据存储的最小单位，指一个二进制位，通常称为比特 (bit)，用“b”表示。一位可表示为“0”或“1”。

2. 字节 (Byte)

字节，是计算机中数据存储的基本单位，规定八位连续的二进制位称为一个字节 (Byte)，通常用“B”表示，即 $1B=8bit$ 。一个汉字占两个字节。由于计算机中存储和处理的信息量很大，因此人们也常用千字节 (KB)、兆字节 (MB)、千兆字节 (GB) 和兆兆字节 (TB) 作为度量单位，它们的换算关系是：

$1KB=1024B$

$1MB=1024KB$

$1GB=1024MB$

$1TB=1024GB$

3. 字 (word)

计算机处理数据时，处理器通过数据总线一次存取、加工和传送的数据称为字。一个字通常由一个字节或若干个字节组成。字是计算机信息交换、处理和存储的基本单元。

4. 字长

字长是处理器能同时处理的二进制数据位数。它直接关系到计算机的精度、功能和速度，是衡量计算机性能的一个重要指标。字长越长，精度越高，处理能力就越强。目前，常见的微机字长为 32 位和 64 位。