

# 第1章 计算机基础知识



## 本章导读

- 计算机概述——计算机的发展、计算机的特点、计算机的应用、计算机的分类、计算机的新技术、计算机的发展趋势、信息技术的发展
- 数制与编码——数制的基本概念、二进制数、八进制数、十进制数和十六进制数、数制间的转换、西文字符的编码、汉字的编码
- 计算机指令和程序设计语言——计算机指令、程序设计语言
- 计算机系统组成——计算机的硬件组成、计算机软件系统、计算机的结构
- 微型计算机的硬件系统——微型计算机的主要性能指标、微型计算机的基本配置、微机的硬件系统
- 多媒体技术简介——多媒体技术简介、多媒体个人计算机、媒体的数字化、多媒体数据压缩
- 计算机病毒与防治——计算机病毒的特点、计算机感染病毒后的常见症状、病毒的分类、计算机病毒的预防、计算机病毒的清除

## 1.1 计算机概述

### 1.1.1 概述

计算机 (Computer) 俗称电脑, 是一种能够高速运算、具有内部存储能力、由程序控制其操作过程并能够自动进行信息处理的电子设备。目前, 计算机已成为我们学习、工作和生活中使用最广泛的工具之一。

本节主要介绍计算机的发展过程、特点、分类和应用领域等内容。

#### 1. 计算机的发展过程

1946年, 世界上第一台电子数字积分式计算机 ENIAC (Electronic Numerical Integrator And Computer) 在美国宾夕法尼亚大学研制成功。这台计算机结构复杂、体积庞大, 但功能远不及现在一台普通微型计算机。

ENIAC 的诞生宣告了电子计算机时代的到来, 它奠定了计算机发展的基础, 开辟了计算机科学技术的新纪元。从第一台电子计算机诞生到现在, 计算机技术经历了大型机阶段和微型机阶段。在 ENIAC 的研制过程中, 美籍匈牙利科学家冯·诺依曼总结并归纳了以下三个特点。

1) 采用二进制。在计算机内部, 程序和数据采用二进制代码表示。

2) 存储程序控制。程序和数据存放在存储器中, 即程序存储的概念。计算机执行程序时无需人工干预, 能自动、连续地执行程序, 并得到预期的结果。



3) 计算机的五个基本组成部分。计算机具有运算器、控制器、存储器、输入设备和输出设备五个基本组成部分。

通常,计算机的发展过程根据采用电子元件的不同被划分为电子管,晶体管,集成电路以及大规模、超大规模集成电路四个阶段,分别称为第一代至第四代计算机。在这四个阶段的发展过程中,计算机的体积越来越小,功能越来越强,价格越来越低,应用越来越广泛。

#### (1) 第一代计算机(1946~1957)

主要元件是电子管。

运算速度为每秒几千次到几万次,内存容量仅为 1000~4000 字节。

主要用于军事和科学研究。

体积庞大、造价昂贵、运算速度低、存储容量小、可靠性低、不易掌握、维护困难。

代表机型为 UNIVAC-1。

#### (2) 第二代计算机(1958~1964)

主要元件是晶体管。

运算速度提高到每秒几十万次,内存容量扩大到几十万字节。

应用已扩展到数据处理和事务处理。

体积小、重量轻、耗电量少、运算速度快、可靠性高、工作稳定。

代表机型为 IBM-7000 系列机器。

#### (3) 第三代计算机(1965~1970)

主要元件采用小规模集成电路(SSI)和中规模集成电路(MSI)。

主要用于科学计算、数据处理及过程控制。

功耗、体积、价格等进一步下降,运算速度及可靠性相应提高。

代表机型为 IBM-360 系列机器。

#### (4) 第四代计算机(1971 年至今)

主机元件采用大规模集成电路(LSI)和超大规模集成电路(VLSI)。

运算速度可达每秒几百万次至上亿次。

应用领域不断向社会各个方面渗透。

体积、重量、功耗进一步减小。

代表机型有 IBM 4300/3080/3090/9000 系列机器。

熟记第一台计算机的名称、诞生时间和地点;第一到四代计算机的主要元件和代表机型;我国大型计算机的代表。

### 2. 微型计算机的发展

1971 年,世界上第一片 4 位微处理器 4004 在 Intel 公司诞生,标志着计算机进入了微型计算机时代。

以微处理器为核心的微型计算机属于第四代计算机。通常以微处理器为标志来划分微型计算机,如 286、386、Pentium 和 Pentium 4 等。微型计算机的发展史实际上就是微处理器的发展史。微处理器一直遵循摩尔(Moore)定律,其性能以平均每 18 个月提高一位的高速度发展着。Intel 公司的芯片设计和制造工艺一直处于业界前端,在宏观上可划分为 80×86 时代和 Pentium 时代。

1981 年,IBM 公司首次推出配置 Intel 8088 芯片的准 16 位 IBM PC(个人计算机),1983 年又推出了 IBM PC/XT,使微型计算机进入了一个迅速发展的时期。

仅仅 20 多年时间,微处理器已发展到 Pentium 4.1/4.2,与最初的 IBM PC 相比,性能已不可同日而语。

### 3. 我国计算机技术的发展概况

我国计算机技术研究起步晚、起点低,但随着改革开放的深入和国家对高新技术的扶持、对



创新能力的倡导，计算机技术的水平正在逐步提高。我国计算机技术的发展历程如下所述。

1956年，开始研制计算机。1958年，研制成功第一台电子管计算机——103计算机。1959年，104计算机研制成功，这是我国第一台大型通用电子数字计算机。1964年，研制成功晶体管计算机。1971年，研制成功以集成电路为主要器件的DJS系列机器。这一时期，在微型计算机方面，我国研制开发了长城、紫金、联想系列微型机。

1983年，我国第一台亿次巨型计算机“银河”诞生。1992年，10亿巨型计算机“银河”诞生。1997年，每秒130亿浮点运算、全系统内存容量为9.15GB的巨型机“银河”研制成功。

1995年，第一套大规模并行机系统“曙光2000-”超级服务器问世，峰值速度可达每秒1117亿次，内存高达50GB。

1999年，“神威”并行计算机研制成功，其技术指标位居世界第48位。

2001年，中科院计算所成功研制我国第一款通用CPU——“龙芯”。

2002年，我国第一台拥有自主知识产权的服务器“龙腾”诞生。

2005年，联想并购IBM PC，一跃成为全球第三大PC制造商。

2008年，我国自主研发制造的百万亿次超级计算机“曙光5000”获得成功。

### 1.1.2 计算机的特点

计算机具有以下特点。

(1) 处理速度快。现在运算速度高达 $10^5$ 亿次/秒的计算机，使过去人工计算需要几年或几十年完成的科学计算能在几小时或更短时间内完成。

(2) 计算精度高。能够处理的字长的增加配合先进的计算技术，计算机的高精度计算能力解决了其他计算工具无法解决的问题。

(3) 存储容量大。主存储器（内存）的容量越来越大；随着大容量的磁盘、光盘等存储器的发展，外部存储器的存储容量也非常可观。

(4) 可靠性高。现在的计算机可靠性达到了很高的水平，很少发生错误。我们通常所说的“计算机错误”，大多是计算机的外设错误和人为造成的错误。

(5) 自动化工作。计算机可以在编制好的程序控制下自动工作，不需要人工干预，实现自动化工作。

(6) 适用范围广。计算机在各行各业中的广泛应用，常常产生显著的经济效益和社会效益，从而引起产业结构、产品结构、经营管理和方式等方面的重大变革。在产业结构中已出现了计算机制造业和计算机服务业，以及知识产业等新的行业。计算机还是人们的学习工具和生活工具。借助家用计算机、个人计算机、互联网、数据库系统和各种终端设备，人们可以学习各种课程，获取各种情报和知识，处理各种生活事务（如订票，购物，存、取款等），甚至可以居家办公。越来越多的人的工作、学习和生活中将与计算机发生直接或间接的联系。

### 1.1.3 计算机的应用

计算机的应用主要分为数值计算和非数值计算两大类。信息处理、计算机辅助设计、计算机辅助教学、过程控制等均属于非数值计算，其应用领域远远大于数值计算。

据统计，目前计算机有5000多种用途，并且以每年300~500种的速度增加。计算机的主要应用领域可以分为以下几个方面。

#### 1. 科学计算（数值计算）

科学计算，主要解决科学研究和工程技术中产生的大量数值计算问题。这是计算机最初的也是最重要的应用领域。



计算机“计算”能力的增强,推进了许多科学研究的进展,如人类基因序列分析计划、人造卫星的轨道测算等。

## 2. 信息处理

信息处理是指对大量数据进行加工处理,如收集、存储、传送、分类、检测、排序、统计和输出等,再筛选出有用信息。信息处理是非数值计算,与科学计算不同,处理的数据虽然量大,但计算方法简单。

## 3. 过程控制

过程控制又称实时控制,是指用计算机实时采集控制对象的数据,加以分析处理后,按系统要求对控制对象进行控制。工业生产领域的过程控制是实现工业自动化的重要手段。利用计算机代替人对生产过程进行监视和控制,可以大大提高劳动生产率。

## 4. 计算机辅助设计和辅助制造

计算机辅助设计简称 CAD (Computer Aided Design)。在 CAD 系统的帮助下,设计人员能够实现最佳的模拟设计,提前做出设计判断,并能很快地制作图纸。

计算机辅助制造简称 CAM (Computer Aided Manufacturing)。CAM 利用 CAD 输出的信息控制、指挥作业。

将 CAD、CAM 和数据库技术集成在一起,形成 CIMS (计算机集成制造系统) 技术,可实现设计、制造和管理自动化。

## 5. 网络通信

网络通信是指通过电话交换网等方式将计算机连接起来,实现资源共享和信息交流。计算机网络通信的应用主要有以下几个方面。

- 网络互联技术。
- 路由技术。
- 数据通信技术。
- 信息浏览技术。
- 网络技术。

## 6. 人工智能

人工智能是指模拟人类的学习过程和探索过程。人工智能的应用主要有以下几个方面。

- 自然语言理解。
- 专家系统。
- 机器人。

## 7. 数字娱乐

运用计算机网络可以为人们带来丰富多彩的娱乐活动,例如丰富的电影、电视资源,网络游戏等。另外,数字电视的发展也使传统电视的单向播放模式转变为交互模式。

## 8. 嵌入式系统

把处理器芯片嵌入计算机设备中完成特定处理任务的系统称为嵌入式系统。嵌入式应用主要有以下两个方面。

- 消费电子产品。
- 工业制造系统。

### 1.1.4 计算机的分类

依照不同的标准,计算机有多种分类方法,常见的分类有以下几种。



### 1. 按处理数据的类型分类

按处理数据的类型不同,可分为数字计算机、模拟计算机和混合计算机。

- 1) 数字计算机,处理以0、1表示的二进制数字。运算精度高,存储量大,通用性好。
- 2) 模拟计算机,处理的数据是连续的,运算速度快,但精度低,通用性差。
- 3) 混合计算机,具有以上两种计算机的特点。

### 2. 按使用范围分类

按使用范围大小,计算机可以分为专用计算机和通用计算机。

- 1) 专用计算机,专门为某种需求而研制,不能作其他用途使用。工作效率高,运算精度高、速度快。
- 2) 通用计算机,适用于一般应用领域,即通常所说的计算机。

### 3. 按性能分类

依据计算机的主要性能(字长、存储容量、运算速度、外部设备)进行分类,可分为超级计算机、大型计算机、小型计算机、微型计算机、工作站和服务器六类。这也是常用的分类方法。

(1) 超级计算机。用于气象、太空、能源和医药等领域与战略武器研制中的复杂计算,如中国的“银河”、“曙光”和“神威”等。

(2) 大型计算机。用于大型软件企业、商业管理和大型数据库,也可用作大型计算机网络的主机,如IBM 4300、IBM 9000系列。

(3) 小型计算机。价格低廉,适合中小型单位使用,如DEC公司的VAX系列、IBM公司的AS/4000系列。

(4) 微型计算机。小巧、灵活、便宜,一次只能一个用户使用,又称个人计算机(PC),如台式机、笔记本电脑、便携机,掌上电脑等。

(5) 工作站。应用于图像处理、计算机辅助设计以及计算机网络领域。

(6) 服务器。通过网络对外提供服务。相对于PC来说,其稳定性、安全性、性能等方面的要求更高。

## 1.1.5 计算机的新技术

计算机的应用尽管已经渗透到科学技术的各个领域,并扩展到工业、农业、军事、商业以及家庭生活之中,但随着科学的飞速发展和全球范围内的新技术革命的不断兴起,现有的计算机性能已经满足不了社会的需要。于是相应产生了许多计算机的新技术,主要有以下几种。

### 1. 嵌入式技术

嵌入式技术是指将计算机作为一个系统处理部件嵌入到应用系统的一种技术。它的本质是将软件固化集成到硬件系统中,使硬件与软件系统一体化。嵌入式系统几乎涉及了生活中的所有电器设备,如移动计算机设备、电视机顶盒、手机、数字电视、多媒体、汽车、微波炉、数码照相机等,具有软件代码小、高度自动化和响应速度快等特点。同时,嵌入式系统要求具有实时性。

### 2. 网络计算

随着科学的进步,世界上每时每刻都在产生着海量的信息。面对这样海量的计算量,高性能计算机也是束手无策的。于是人们把目光投向了大约数亿台在大部分时间里处于闲置状态的计算机。假如发明一种技术,自动搜索到这些计算机,并将它们连接起来,所形成的计算能力肯定会超过许多高性能计算机。网络计算的出现就来源于这种思想。

网络计算专门针对复杂科学计算的新型计算模式。这种计算模式是利用互联网把分散在不同地理位置的电脑组织成一个虚拟的超级计算机,其中每一台参与计算的计算机就是一个“节点”,



而整个计算是由成千上万个“节点”组成的“一张网格”。网格计算的优势有两个：数据处理能力超强和能充分利用互联网上的闲置处理能力。

网格计算的特点如下。

- 能够提供资源共享，实现应用程序的互相连接。
- 协同工作，共同处理一个项目。
- 基于国际的开放技术标准。
- 可以提供动态服务，适应变化。

### 3. 中间件技术

中间件是介于应用软件和操作系统之间的系统软件。中间件的特点如下。

- 满足大量应用的需要。
- 运行于多种硬件和 OS (操作系统) 平台。
- 支持分布计算，提供跨网络、硬件和 OS 平台的透明性应用或服务的交互。
- 支持标准的协议。
- 支持标准的接口。

中间件抽象了典型的应用模式，应用软件制造者可以基于标准的中间件进行再开发。中间件可分为多种，如交易中间件、消息中间件、专有系统中间件、面向对象中间件、数据访问中间件、远程过程调用中间件、Web 服务器中间件、安全中间件等。

## 1.1.6 未来计算机的发展趋势

21 世纪是网络的时代，是超高速信息公路建设取得实质性进展并进入应用的年代。那么，计算机的发展趋势是什么？未来新一代计算机的类型是什么？

### 1. 计算机的发展趋势

#### (1) 巨型化

巨型化是指高速运算、大存储容量和强功能的巨型计算机。它的运算能力一般在每秒百亿次以上、内存容量在几百吉字节以上。巨型计算机的发展集中体现了计算机科学技术的发展水平，推动了计算机系统结构、硬件和软件的理论和技术、计算数学以及计算机应用等多个科学分支的发展，主要用于尖端科学技术和军事国防系统的研究开发。

#### (2) 微型化

大规模、超大规模集成电路的出现，使计算机迅速向微型化方向发展。微型化是指体积更小、功能更强、可靠性更高、携带更方便、价格更便宜、适用范围更广的计算机系统，因为微型计算机可以渗透到仪表、家电、导弹弹头等微机无法进入的领域，所以 20 世纪 80 年代以来发展异常迅速。

#### (3) 网络化

计算机网络是指计算机技术发展的又一重要分支，是现代通信技术与计算机技术相结合的产物。网络化就是利用现代通信技术和计算机技术，将分布在不同地点的计算机连接起来，按照网络协议互相通信，共享软件、硬件和数据资源。

#### (4) 智能化

第五代计算机要实现的目标是智能化，让计算机来模拟人的感觉、行为、思维过程，使计算机具有视觉、听觉、语言、推理、思维、学习的能力，成为智能型计算机。

### 2. 未来新一代的计算机

#### (1) 模糊计算机

实际生活中，人们使用着大量的模糊概念，比如“走快一些”、“再来一点”、“休息片刻”



中的“一些”、“一点”、“片刻”等都是含糊、不清的说法，人们往往需要处理大量的模糊信息。目前的计算机只能进行精确运算而不能处理模糊信息，而模糊计算机除具有一般计算机的功能外，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，甚至可以帮助人从事复杂的脑力劳动。

1990年，日本松下公司把模糊计算机装在洗衣机中，它可以根据衣服的材质调节洗衣程序。我国有些品牌的洗衣机也安装了模糊逻辑芯片。此外，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厚实程度调整吸尘器的功率。模糊计算机还能用于地震、火情判断、疾病医疗诊断、发酵工程控制、海空导航巡视等多个方面。

### (2) 生物计算机

微电子技术和生物工程两项高科技的互相渗透为研制生物计算机提供了可能。利用DNA化学反应，通过和酶的相互作用可以使某基因代码通过生物化学的反应转变为另一种基因代码，转变前的基因代码可以作为输入数据，反应后的基因代码可以作为运算结果。利用这一过程可以制成新型的生物计算机。科学家们认为，生物计算机的发展可能要经历一个较长的过程。

### (3) 光子计算机

光子计算机是一种用光信号进行数字运算、信息存储和处理的新型计算机，运用集成电路技术，把光开关、光存储器等集成在一块芯片上，再用光导纤维连接成计算机。1990年1月底，贝尔实验室研制成第一台光子计算机。光子计算机的关键技术为光存储、光互联技术、光集成器件。除了贝尔实验室外，日本和德国的其他公司都投入巨资研制光子计算机，笔者预计未来将出现更先进的光子计算机。

### (4) 超导计算机

超导技术的发展使科学家们想到用超导材料来替代半导体制造计算机。超导计算机具有超导逻辑电路和超导存储器，运算速度是传统计算机无法比拟的，美国科学家已经成功地将5000个超导单元装置在小于 $10\text{cm}^3$ 的主机内，组成一个简单的超导计算机，每秒能执行2.5亿条指令。研制超导计算机的关键之一是要有一套维持超低温的设备。

### (5) 量子计算机

量子计算机中的数据用量子位存储。由于量子有叠加效应，一个量子位可以是0或1，也可以既存储0又存储1，因此一个量子位可以存储2个数据。同样数量的存储位，量子计算机的存储量比传统的电子计算机大许多。传统计算机与量子计算机之间的区别是，传统计算机遵循着众所周知的经典物理规律，而量子计算机则遵循独一无二的量子动力学规律，是一种信息处理的新模式。

## 1.1.7 信息技术的发展

信息同物质、能源一样重要，是人类生存和社会发展的三大基本资源之一。数据处理之后产生的结果为信息，信息具有针对性、实时性，是有意义的信息。信息技术是指应用在信息加工和处理中的科学、技术，工程的训练方法与管理技巧；上述方面的技巧应用；计算机与人、机的交互作用；与之相应的社会、经济和文化等多种事物。目前，信息技术主要指一系列与计算机相关的技术。

一般来说，信息技术包括了信息基础技术、信息系统技术和信息应用技术。

### (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。

### (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、控制的设备和系统的技术。感测技术、通信技术、计算机与智能技术和控制技术是其核心和支撑技术。



### (3) 信息应用技术

信息应用技术是针对各种实用目的的技术,如信息管理、信息控制、信息决策等技术。信息技术在社会各个领域得到了广泛的应用,显示出强大的生命力。展望未来,现代信息技术将向数字化、多媒体化、高速度、网络化、宽频带、智能化等方面发展。

## 1.2 数制与编码

计算机要正常工作,就必须存放各种指令及各种类型的数据,这些数据不仅包括数值,还包括各种字符、文字、图形、图像、声音、动画和视频等。

本节前半部分介绍的数制基本知识是本章的重点内容;后半部分讲述计算机内部数据的表示与编码。

### 1.2.1 数制的基本概念

人们在生产实践和日常生活中创造了许多种表示数的方法,如常用的十进制、钟表计时中使用的六十进制等。这些数的表示规则称为数制。

下面以最熟悉和最常用的十进制为例解释数制中的基本要素,并引入  $R$  进制数的概念。数制中的基本要素及其含义如表 1-1 所示。

表 1-1 数制中的基本要素及其含义

基本要素	具体含义
数码	用来表示数据的数字符号
基数	所包含的数码个数。每逢超过或等于此数就要向相邻的高位进一
数位	数码在数中所处的位置。不同的位置代表不同的数值
位权	它是基数为底的指数函数。数码在不同数位时,它所代表的实际值等于数字本身的值乘上一个确定的与数位有关的系数,这个系数称为位权

例如,在十进制中,任何数值都可用 1、2、3、4、5、6、7、8、9、0 中的数字来表示,这十个数字符号就是十进制的数码。十进制的数码中有十个,且逢十向相邻高位进一(简称“逢十进一”),所以十进制的基数  $R = 10$ 。

【例 1-1】十进制数 888.88 中,五个数码都是 8,但每个 8 代表的值却是不同的。

第 1 个数码 8 处于百位,代表 800,它的位权为  $10^2$ 。

第 2 个数码 8 处于十位,代表 80,它的位权为  $10^1$ 。

第 3 个数码 8 处于个位,代表 8,它的位权为  $10^0$ 。

第 4 个数码 8 处于十分位,代表 0.8,它的位权为  $10^{-1}$ 。

第 5 个数码 8 处于百分位,代表 0.08,它的位权为  $10^{-2}$ 。

本例中的五个数码都是“8”,但五个 8 的数位不同,所以其实际值是不一样的。

$$\text{数码的实际值} = \text{数码值} \times \text{位权}$$

“指数”和数的位数有关。整数部分指数的范围“ $0 \sim (\text{整数的位数}-1)$ ”,由 0 开始从小数点向左递增 1;小数部分指数的范围为“ $-1 \sim -\text{小数的位数}$ ”,由 -1 开始从小数点向右递减 1。

如例 1-1 中的 888.88,整数的位数是“3”,所以它的指数范围是“ $0 \sim (3-1)$ ”即“ $0 \sim 2$ ”。从小数点向左排列,其指数就是“0”、“1”、“2”。小数的位数是“2”,所以它的指数范围是“ $-1 \sim -2$ ”,从小数点向右排列,其指数就是“-1”和“-2”。

【例 1-2】十进制数 345.26 的按权展开式为:

$$345.26 = 3 \times 10^2 + 4 \times 10^1 + 5 \times 10^0 + 2 \times 10^{-1} + 6 \times 10^{-2}$$

使用  $R$  个数字符号来表示数据,按  $R$  进位的方法进行记数,称为  $R$  进位记数制,简称  $R$  进制。对于任意具有  $N$  位整数、 $M$  位小数的  $R$  进制数有同样的基数  $R$ 、位权  $R^i$  (其中  $i = -m$  至  $n-1$ ) 和按权展开表示式。

从前面的介绍可知,每个数码的实际值=数码值 $\times$ 位权。而“按权展开”的实际意义就是求整个数的值,即整个数的实际值等于每个数码的实际值相加。

如例 1-1 中,“888.88”按权展开就是 888.88 这个数实际的值,方法就是将 5 个 8 的实际值相加:

$$888.88 = 8 \times 10^2 + 8 \times 10^1 + 8 \times 10^0 + 8 \times 10^{-1} + 8 \times 10^{-2}$$

位权指数:整数部分,由 0 开始自小数点向左递增 1。

## 1.2.2 二进制数、八进制数、十进制数和十六进制数

了解了数制的规律后,下面具体介绍二进制数、八进制数、十进制数和十六进制数,如表 1-2 所示。

表 1-2 常用数制简介

二进制	基数为 2; 两个数码:0~1; 逢二进一; 位权为 $2^i$	十进制	基数为 10; 十个数码:0~9; 逢十进一; 位权为 $10^i$
八进制	基数为 8; 八个数码:0~7; 逢八进一; 位权为 $8^i$	十六进制	基数为 16; 16 个数码:0~9 和 A~F; 逢十六进一; 位权为 $16^i$

以上  $i = -m \sim n-1$ ,  $m$ 、 $n$  为自然数,分别代表数的小数、整数部分的位数。

二进制是计算机中采用的数制,这是因为二进制具有如下特点。

简单可行,容易实现。二进制数仅有两个数码 0 和 1,可以用两种不同的稳定状态(如高电位与低电位)来表示。

运算规则简单。二进制数的计算规则非常简单。以加法为例,二进制加法规则是“逢二进一”。

适合逻辑运算。二进制中的 0 和 1 正好分别表示逻辑代数中的“假”(False)和“真”(True)。二进制数可用于代表逻辑值,容易实现逻辑运算。

但是二进制也有明显的缺点,即数字冗长、书写麻烦且容易出错、便于阅读。所以,在计算机技术文献的书写中,常用十六进制表示。

二进制、十进制、十六进制是学习“数制”最基本的内容,要求读者能做到:在一定的数值范围内直接写出二进制、十进制、十六进制的对应关系。表 1-3 所示为十进制数 0~15 与二进制数、十六进制数的对应关系。

表 1-3 十进制数 0~15 与二进制数、十六进制数的对应关系

十进制	二进制	十六进制	十进制	二进制	十六进制
0	0000	0	8	1000	8
1	0001	1	9	1001	9
2	0010	2	10	1010	A
3	0011	3	11	1011	B
4	0100	4	12	1100	C



续表

十进制	二进制	十六进制	十进制	二进制	十六进制
5	0101	5	13	1101	D
6	0110	6	14	1110	E
7	0111	7	15	1111	F

### 1.2.3 数制间的转换

#### 1. 非十进制数转换成十进制数

非十进制数转换成十进制数的方法是按权展开, 见例 1-3。

【例 1-3】

$$(11010)_2 = 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = (26)_{10}$$

$$(111.01)_2 = 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 0 \times 2^{-1} + 1 \times 2^{-2} = (7.25)_{10}$$

$$(A2B)_{16} = 10 \times 16^2 + 2 \times 16^1 + 11 \times 16^0 = 2560 + 32 + 11 = (2603)_{10}$$

#### 2. 十进制数转换成 R 进制数

将十进制数转换为 R 进制数时, 可将此数分成整数与小数两部分分别转换, 然后拼接起来即可。

十进制整数转换成二进制整数的方法是“除二取余法”, 具体步骤如下。

步骤 1: 把十进制数除以 2 得一个商和余数, 商再除以 2 又得一个商和余数, 依次进行该步骤直到商是零为止。

步骤 2: 以最先除得的余数为最低位, 最后除得的余数为最高位, 从最高位到最低位依次排列, 即为该十进制整数的等值二进制整数了。

十进制小数转换成二进制采用“乘二取整法”, 具体步骤如下。

步骤 1: 把十进制数乘以 2 得一个新数, 若整数部分为 1, 则二进制纯小数相应位为 1; 若整数部分为 0, 则相应位为 0。

步骤 2: 从高位向低位逐次进行, 直到满足精度要求或乘以 2 后的小数部分是 0 为止。

同理, 十进制与八进制相互转换使用“除八取余法”和“乘八取整法”, 十进制转换成十六进制使用“除十六取余法”。

【例 1-4】将十进制数  $(2606)_{10}$  转换成十六进制数。

$$(2606)_{10} = (A2D)_{16}$$

#### 3. 二进制与十六进制间的转换

由于 16 是 2 的 4 次幂, 所以可以用 4 位二进制数来表示 1 位十六进制数。

##### (1) 十六进制数转换成二进制数

对每 1 位十六进制数, 用与其等值的 4 位二进制数代替。

【例 1-5】将十六进制数  $(1AC0.6D)_{16}$  转换成二进制数。

$$(1AC0.6D)_{16} = (110101100000.01101101)_2$$

##### (2) 二进制数转换成十六进制数。

其方法是从小数点开始, 整数部分向左、小数部分向右每 4 位数为 1 节, 整数部分最高位不足 4 位或小数部分最低位不足 4 位时补 0, 然后将每节依次转换成十六进制数, 再把这些二进制数连接起来即为等值十六进制数。

【例 1-6】将二进制数  $(10111100101.00011001101)_2$  转换成十六进制数。

$$(10111100101.00011001101)_2 = (5E5.19A)_{16}$$

同理, 8 是 2 的 3 次幂, 所以可以用 3 位二进制数来表示 1 位八进制数。



【例 1-7】将八进制数  $(2731.62)_8$  转换成二进制数。

$$(2731.62)_8 = (010111011.110010)_2$$

## 1.2.4 计算机数据的存储

### 1. 计算机数据的常用单位

由于在计算机内部指令和数据都是用二进制数表示的，因此，计算机系统中信息存储、处理也都是以二进制数为基础的。以下为计算机内数据的常用单位。

位：一个二进制位，称为位（bit），是数据的最小单位，表示为 bit。

字节：8 位二进制数编为一组，称为一个字节（Byte），是信息处理的最基本单位，表示为 B。

字：由若干字节组成的（通常取字节的整数倍）。

现代计算机中存储数据是以字节作为处理单位的，如一个英文大写字母用一个字节表示，而一个汉字和国标图形字符需用两个字节表示。实际使用中这样的单位表示量太小，常用 KB、MB、GB 和 TB 作为数据的存储单位。常用的二进制数的数据单位如表 1-4 所示。

表 1-4 常用的二进制数的数据单位

单 位	名 称	含 义	说 明
bit	位	表示 1 个 0 或 1，称为 bit	最小的数据单位
B	字节	8bit 为 1B	数据处理的基本单位
KB	千字节	1KB=1024B=2 <sup>10</sup> B	适用于文件计量
MB	兆字节	1MB=1024KB=2 <sup>20</sup> B	适用于内存、软盘、光盘计量
GB	吉字节	1GB=1024MB=2 <sup>30</sup> B	适用于硬盘的计量单位
TB	太字节	1TB=1024GB=2 <sup>40</sup> B	适用于硬盘的计量单位

### 2. 计算机数据类型

计算机使用的数据可分为两类：数值型数据和字符型数据（非数值数据）。

在计算机中，不仅数值数据是用二进制数来表示的，字符数据也都用二进制数进行编码。

## 1.2.5 西文字符的编码

所谓编码，就是用二进制数来表示数据的代码。

计算机中常用的字符（西文字符）编码有两种：EBCDIC 码和 ASCII 码。IBM 系列的大型计算机采用 EBCDIC 码，微型计算机采用 ASCII 码。下面主要介绍 ASCII 码。

ASCII 码是美国信息交换标准代码（American Standard Code for Information Interchange）的缩写。该编码被国际标准化组织 ISO 采纳，作为国际通用的信息交换标准代码，是目前在微型机中普遍使用的字符编码。

标准 ASCII 码中通用控制字符有 34 个，阿拉伯数字 10 个，大、小写英文字母 52 个，各种标点符号和运算符 32 个。

比较字符的大小其实就是比较字符 ASCII 值的大小。一般来说，可见控制符的 ASCII 值 < 数字的 ASCII 值 < 大写字母的 ASCII 值 < 小写字母的 ASCII 值。

例如，空格的 ASCII 码是 32，“！”的 ASCII 码是 33，“0”的 ASCII 码是 48，“A”的 ASCII 码是 65，“Z”ASCII 码是 90，“a”的 ASCII 码是 97，“z”的 ASCII 码是 122。

## 1.2.6 汉字的编码

为使计算机可以处理汉字，也需要对汉字进行编码。计算机进行汉字处理的过程，实际上是



各种汉字编码间的转换过程。汉字编码有汉字信息交换码、汉字输入码、汉字内码、汉字字形码和汉字地址码等。下面分别介绍各种汉字编码。

### 1. 汉字信息交换码

汉字信息交换码(国标码)是用于汉字信息处理系统之间或汉字信息处理系统与通信系统之间信息交换的汉字代码。我国于1981年颁布了国家标准的汉字编码集,即《信息交换用汉字编码字符集—基本集》,国家标准号是GB 2312-1980”,简称交换码或国标码。下面介绍国标码的有关知识。

国标码的字符集共收录了7445个图形符号和两级常用汉字,有682个非汉字图形符号和6763个汉字的代码。汉字代码中有一级常用汉字3755个,二级常用汉字3008个。

国际码的存储占用两个字节,编码范围为2121H~7E7E。

区位码,又称国标区位码,是国标码的一种变形。它把全部一级、二级汉字和图形符号排列在一个94行94列的矩阵中,构成一个二维表格,类似于ASCII码表。

“区”:阵中的每一行用区号表示,区号范围是1~94。

“位”:阵中的每一列用位号表示,位号范围是1~94。

区位码:汉字的区号与位号的组合(高两位是区号,低两位的是位号)。

实际上,区位码也是一种汉字输入码,其优点是一字一码,缺点是难以记忆。

### 2. 汉字输入码

汉字输入码是为让用户能够使用英文键盘输入汉字而编制的编码,又称外码。汉字输入码有许多种不同的编码方案,大致分为以下几类。

1) 音码:以汉语拼音字母和数字为汉字编码。例如,全拼输入法和双拼输入法。

2) 形码:根据汉字的字形结构对汉字进行编码。例如,五笔字型输入法。

3) 音形码:以拼音为主,辅以字形字义进行编码。例如,自然码输入法。

4) 数字码:直接用固定位数的数字给汉字编码。例如,区位输入法。

同一个汉字在不同的输入码编码方案中的编码一般也不同。例如,使用全拼输入法输入“嵌”字,就要键入编码“qian”,而用五笔字型的输入码是“mafW”。

### 3. 汉字内码

汉字内码是为在计算机内部对汉字进行处理、存储和传输而编制的汉字编码,应能满足存储、处理和传输的要求。不论用何种输入码,输入的汉字在机器内部都要转换成统一的汉字内码,然后才能在机器内传输、处理。

目前,对应于国标码,一个汉字的内码也用两个字节存储。因为ASCII码是英文的内码,为不使汉字内码与ASCII码发生混淆,将国标码每个字节的最高位置1作为汉字内码。

国标码与内码之间的关系:内码=汉字的国标码+(8080)<sub>16</sub>

### 4. 汉字字形码

汉字字形码是存放汉字字形信息的编码,它与汉字内码一一对应。每个汉字的字形码是预先存放在计算机内的,常称为汉字库。当输出汉字时,计算机根据内码在字库中查到其字形码,得知字形信息,然后就可以显示或打印输出了。

描述汉字字形的方法主要有点阵形和轮廓字形两种。

### 5. 汉字地址码

汉字地址码是指汉字库(这里主要指汉字字形的点阵式字模库)中存储汉字字形信息的逻辑地址码。在汉字库中,字形信息都是按一定顺序(大多数按照标准汉字国标码中汉字的排列顺序)连续存放在存储介质中的,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，从而简化了汉字内码到汉字地址码的转换。

## 1.3 计算机指令和程序设计语言

计算机能按照要求自动完成工作是因为采用了存储程序控制。那么，什么是程序呢？程序又是用何种工具编写的呢？

### 1.3.1 计算机指令

指令就是用户给计算机下达的命令，它告诉计算机要干什么，所要用到的数据在哪里，操作结果又将送往何处。所以，指令包括操作码和地址码。

操作码：指出指令完成操作的类型，如加、减、乘、除、传送等。

地址码（或称操作数）：指出参与操作的数据和操作结果存放的位置。

一条指令只能完成一个简单的操作，而一个比较复杂的操作，就需要由许多简单操作组合而成，这就形成了程序。简单地说，程序就是一组计算机指令序列。

### 1.3.2 程序设计语言

一台计算机可能有多种多样的指令，这些指令的集合称为该计算机的指令系统。

程序设计语言是用来编写计算机程序的，是用户与计算机交互的语言。按其指令代码的类型分为机器语言、汇编语言和高级语言。

#### 1. 机器语言

计算机的指令系统也称为机器语言，具有以下主要特征。

- 1) 它是计算机唯一能识别并且直接执行的语言；
- 2) 每条指令是由 0、1 组成的一串二进制代码，可读性差、不易记忆；
- 3) 用它编写的程序执行速度快，占用内存空间少；
- 4) 程序编写烦琐、难度大，易出错，难以调试和修改；
- 5) 直接依赖于机器；
- 6) 由于不同型号（或系列）计算机的指令系统不完全相同，故可移植性差。

总之，机器语言效率高，但不易掌握和使用。

#### 2. 汇编语言

用比较容易识别和记忆的助记符代替机器语言的二进制代码。这种符号化的机器语言称为汇编语言，又称符号语言，有以下主要特征。

- 1) 指令一般采用相近英语词汇的缩写，如加法运算的指令为 ADD，减法运算的指令为 SUB；
- 2) 在编写程序时，比指令编码容易记忆，出错时也容易修改；
- 3) 汇编语言其实就是用代码表示的机器语言，同机器语言一样，都依赖于具体的机器；
- 4) 计算机不能直接识别和执行汇编语言程序，汇编语言源程序必须经过汇编过程翻译成机器语言程序（目标程序），才能被执行。

#### 3. 高级语言

高级语言是接近于生活语言的计算机语言。常见的高级语言有 BASIC 语言、FORTRAN 语言、C 语言和 Pascal 语言等。和汇编语言程序一样，高级语言程序不能被计算机识别和执行，必须由翻译程序把它翻译成机器语言后才能被执行。

翻译程序按翻译的方法分为解释方式和编译方式两种。