

硬件篇：计算机的硬件组成

第 1 章 计算机体系结构

本章是关于计算机基本原理的介绍,从冯·诺依曼体制入手,引出了计算机体系结构的一些基本概念。本章主要介绍计算机的简单发展历史、信息的数字化表示、数字的编码、信息的储存及处理,以及信息的输入和输出问题。

1.1 计算机与信息技术

国际标准化组织(ISO)对信息的定义是:信息(information)是对人有用的,能够影响人们行为的数据。

信息具有如下特性:可传输性、可存储性、可处理性(加工)、可共享性、时滞性等。信息同物质和能源一样,是人们赖以生存与发展的重要资源。人类通过信息认识各种事物,借助信息的交流沟通人与人之间的联系,互相协作,从而推动社会前进。

信息技术(information technology, IT)是指信息存储、加工、传输和使用的理论和方法,以及相关设备设施的设计、制造、运行、工艺和技术。信息技术主要包括计算机技术、微电子技术和通信技术。

计算机技术是关于数字电子信息处理自动化的学问,它解决了数字化的信息如何存储、如何运输和如何加工运算、管理的问题。计算机技术是信息技术中的核心部分,计算机技术的迅猛发展引发了信息革命。

电子计算机是信息的处理机,它是人脑功能的延长,能帮助人更好地存储信息、检索信息、加工信息和再生信息。

计算机是一种能按照事先存储的程序,自动、高速地进行大量数值计算和各种信息处理的现代化智能电子装置。

1.2 存储程序与冯·诺依曼体制

计算机的运行采取编制程序、存储程序、自动连续运行程序的工作方式,称为存储程

序方式。对此作出重大贡献的是出生于匈牙利的美国数学家冯·诺依曼(1903—1957)。绝大多数人认为,1946年制成的ENIAC是世界上第一台电子数字计算机。但是ENIAC基本上是十进制而不是二进制,程序和数字分开存储,程序的进入与修改需通过人工设置开关和插拔导线来设置,被称为台外程序式。

1945年,冯·诺依曼通过一篇著名的论文概括了数字计算机的设计思想,被后人称为冯·诺依曼思想。这是计算机发展史中的一个里程碑。几十年来,计算机体系结构发生了许多演变,但存储程序的概念仍是普遍采用的结构原则。冯·诺依曼体制中仍广泛采用的要点可归纳如下。

(1) 采用二进制形式表示数据和指令

数据和指令在代码的外形上并无区别,都是由0和1组成的代码序列,只是各自约定的含义不同而已。采用二进制,使信息数字化容易实现,可以用二值逻辑工具进行处理。程序信息本身也可以作为被处理的对象,进行加工处理,例如对源程序进行编译,就是源程序被当做加工处理的对象。

(2) 采用存储程序方式

这是冯·诺依曼思想的核心内容。如前所述,它意味着实现编制程序,事先将程序(包含指令和数据)存入主存储器中,计算机在运行中程序就能自动地、连续地从存储器中依次取出指令并且执行。这是计算机高速自动运行的基础。计算机的工作体现为执行程序,计算功能的扩展在很大的程度上体现为所存储程序的扩展。计算机的许多具体工作方式也是由此派生的。诺依曼机的这种工作方式,可称为控制流(指令流)驱动方式。即按照指令的执行序列,依次读取指令;根据指令所含的控制信息,调用数据进行处理。因此在执行程序的过程中,始终以控制信息流为驱动工作的因素,而数据信息流则是被动地被调用处理。

为了控制指令序列的执行顺序,人们设置了一个程序(指令)计数器(program counter,PC),让它存放当前指令所在的存储单元的地址。如果程序现在是顺序执行的,每取出一条指令后PC内容加1,指示下一条指令该从何处取得。如果程序将转移到某处,就将转移后的地址送入PC,以便按新地址去读取后继指令。所以,PC就像一个指针,一直指示着程序的执行进程,这也就是指示控制流的形成。程序与数据都采用二进制代码,可按照PC的内容作为地址读取指令,再按照指令给出的操作数地址去读取数据。由于在多数情况下程序是顺序执行的,所以大多数指令需要依次紧挨着存放。除了个别即将使用的数据可以紧挨着指令存放外,一般将指令和数据分别存放在该程序区的不同区域。

(3) 由运算器、存储器、控制器、输入装置和输出装置五大部件组成计算机系统,并规定了这五部分的基本功能

上述概念奠定了现代计算机的基本结构思想,并开创了程序设计的新时代。到目前为止,绝大多数计算机仍沿用这一体制,形成为诺依曼机体制。传统的诺依曼体制从本质上讲是采取串行顺序处理的工作机制,即使有关数据已经准备好,也必须逐条执行指令序列。而提高计算机性能的根本方向之一是并行处理。因此,近年来人们在谋求突破传统的冯·诺依曼体制的束缚,这种努力被称为非诺依曼化。对所谓非诺依曼化的探讨仍在

争议中，一般认为它表现为以下三个方面的突破。

① 在诺依曼体制范畴内对传统诺依曼机进行改造

如采用多个处理部件形成流水处理，依靠时间上的重叠提高处理效率；又如组成阵列机结构，形成单指令流多数据流，提高处理速度。这些方向已比较成熟，成为标准结构。

② 用多个诺依曼机组成多机系统，支持并行算法结构

这方面的研究目前比较活跃。

③ 从根本上改变诺依曼机的控制流驱动方式

例如，采用数据流驱动工作方式的数据流计算机，只要数据已经准备好，有关的指令就可并行地执行。这是真正非诺依曼化的计算机，它为并行处理开辟了新的前景，但由于控制的复杂性，仍处于试验探索之中，不在本书的讨论范围之内。

图 1.1 以框图的形式表示出数字计算机的基本硬件组成。典型的数字计算机硬件由五大部分组成，即运算器、存储器、控制器、输入设备和输出设备。

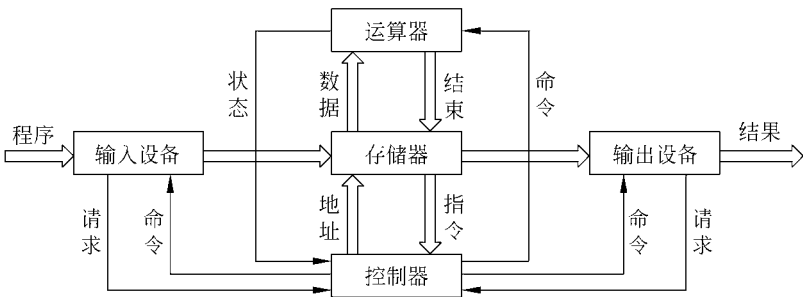


图 1.1 数字计算机的简单框图

在本节中出现的如下概念是下面几节将要阐述的重点：信息的数字化表示、存储器、处理器（运算器和控制器）、输入与输出。

1.3 计算机发展简史

1.3.1 计算机的发展

世界上第一台计算机是 1946 年 2 月由美国的宾夕法尼亚大学研制成功的，该机命名为 ENIAC (Electronic Numerical Integrator And Calculator)，意思是“电子数值积分计算机”。它的诞生在人类文明史上具有划时代的意义，从此开辟了人类使用电子计算工具的新纪元，是人类进入信息时代的里程碑。

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，共经历了四代的变革。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。四代计算机的演变如表 1.1 所示。

1. 电子管计算机

第一代(1946—1957 年)计算机的逻辑元件采用电子管，通常称为电子管计算机。它的内存容量仅有几千个字节，不仅运算速度低，成本也很高。

表 1.1 四代计算机的发展

代次	起止年份	所用电子元器件	数据处理方式	运 算 速 度	应 用 领 域
第一代	1946—1957	电子管	汇 编 语 言、代 码 程序	5000~30000 次/s	国防及高科技
第二代	1958—1964	晶体管	高级程序设计语言	数十万 ~ 几百万 次/s	工程设计、数据 处理
第三代	1965—1970	中、小 规 模 集 成 电路	结构化、模块化程 序设计、实时处理	数百万 ~ 几千万 次/s	工业控制、数据 处理
第四代	1971 至今	大 规 模、超 大 规 模集成电路	分时、实时数据处 理、计算机网络	上亿条指令/s	工业、生活等各 方面

在这个时期,没有系统软件,用机器语言和汇编语言编程。计算机只能在少数尖端领域中得到应用,一般用于科学、军事和财务等方面的计算。尽管存在这些局限性,它却奠定了计算机发展的基础。

2. 晶体管计算机

第二代(1958—1964 年)与第一代相比有很大改进,计算机的逻辑元件采用晶体管,即晶体管计算机。存储器采用磁芯和磁鼓,内存容量扩大到几十 K 字节。晶体管比电子管平均寿命提高 100~1000 倍,耗电却只有电子管的 1/10,体积比电子管小一个数量级,运算速度明显提高,每秒可以执行几万次到几十万次的加法运算,机械强度较高。由于具备这些优点,晶体管计算机很快取代了电子管计算机,并开始成批生产。

在这个时期,系统软件出现了监控程序,提出了操作系统概念,出现了高级语言,如 FORTRAN、ALGOL 60 等。

3. 集成电路计算机

第三代(1965—1970 年)计算机的逻辑元件采用集成电路。这种器件把几十个或几百个分立的电子元件集中做在一块几平方毫米的硅片上(称为集成电路芯片),使计算机的体积和耗电大大减小,运算速度却大大提高,每秒钟可以执行几十万次到 100 万次的加法运算,性能和稳定性进一步提高。

在这个时期,系统软件有了很大发展,出现了分时操作系统和会话式语言,采用结构化程序设计方法,为研制复杂的软件提供了技术上的保证。

4. 大规模与超大规模集成电路计算机

第四代(1971 年以后)计算机的逻辑元件采用大规模集成电路(LSI)。在一个 4mm² 的硅片上,至少可以容纳相当于 2000 个晶体管的电子元件。金属氧化物半导体电路(metal oxide silicon,MOS)也在这一时期出现。这两种电路的出现进一步降低了计算机的成本,体积也进一步缩小,存储装置进一步改善,功能和可靠性却进一步得到提高。同时计算机内部的结构也有很大的改进,采取了“模块化”的设计思想,即按执行的功能划分成比较小的处理部件,更加便于维护。

20 世纪 70 年代末期开始出现超大规模集成电路(VLSI),在一个小硅片上容纳相当于几万个到几十万个晶体管的电子元件。这些以超大规模集成电路构成的计算机日益小

型化和微型化,应用和发展的速度更加迅猛,产品覆盖巨型机、大/中型机、小型机、工作站和微型计算机等各种类型。

在这个时期,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。

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,研究的目标是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。

1.3.2 微型计算机的发展

微型计算机(简称微机或 PC 机)是 1971 年出现的,属于第四代计算机。它的一个突出特点是将运算器和控制器做在一块集成电路芯片上,一般称为微处理器(micro-processor unit,MPU)。根据微处理器的集成规模和功能,又形成了微机的不同发展阶段,如 Intel 80486、Pentium、PⅡ、PⅢ、PⅣ、酷睿等。

世界上第一台微机是由美国英特尔公司年轻的工程师马西安·霍夫(M. E. Hoff)于 1971 年研制成功的。它把计算机的全部电路做在四个芯片上:4 位微处理器 Intel 4004、320 位(40 字节)的随机存取存储器、256 字节的只读存储器和 10 位的寄存器,它们通过总线连接起来,于是就组成了世界上第一台 4 位微型电子计算机——MCS-4,从此揭开了微机发展的序幕。

第一代微处理器是在 1972 年由英特尔公司研制的 8 位微处理器 Intel 8008,主要采用工艺简单、速度较低的 P 沟道 MOS 电路,由它装备起来的计算机称为第一代微型计算机。

第二代微处理器是在 1973 年研制的,主要采用速度较快的 N 沟道 MOS 技术的 8 位微处理器。代表产品有英特尔公司的 Intel 8085、摩托罗拉公司的 M6800、Zilog 公司的 Z80 等。第二代微处理器的功能比第一代显著增强,以它为核心的微型计算机及其外部设备都得到相应的发展,由它装备起来的计算机称为第二代微型计算机。

第三代微处理器是在 1978 年研制的,主要采用 H-MOS 新工艺的 16 位微处理器。其典型产品是英特尔公司的 Intel 8086。Intel 8086 比 Intel 8085 在性能上提高了 10 倍。由第三代微处理器装备起来的计算机称为第三代微型计算机。

从 1985 年起采用超大规模集成电路的 32 位微处理器,标志着第四代微处理器的诞生。典型产品有英特尔公司的 Intel 80386、Zilog 公司的 Z80000、惠普公司的 HP-32 等。由第四代微处理器装备起来的计算机称为第四代微型计算机。

1993 年英特尔公司推出第五代 32 位微处理器芯片 Pentium(中文名为奔腾),它的外部数据总线为 64 位,工作频率为 66~200MHz。

1998 年英特尔公司推出 Pentium Ⅱ、Celeron,后来又推出 Pentium Ⅲ。第六代微处理器都是更先进的 32 位高档微处理器,工作频率为 300~860MHz,主要用于高档微机或服务器。

微机具有体积小、重量轻、功耗小、可靠性高、对环境要求低、价格低廉、易于成批生产等特点。所以,微机一出现,就显示出强大的生命力。

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。

1.3.3 我国计算机的发展

1958年8月1日,中国科学院计算所与北京有线电厂共同研制成我国第一台计算机——103型通用数字电子计算机,运行速度为1500次/s,字长为31位,内存容量为1K字节。同年9月,数字指挥仪901样机问世,这是中国第一台电子管专用数字计算机。

104机是1959年10月1日宣布诞生的我国第一台大型通用数字电子计算机,平均每秒运算1万次,接近当时英国、日本计算机的指标。103机共生产了36台,104机生产了7台,为我国尖端武器的发展作出了重要贡献。

1960年我国第一台大型通用电子计算机——107型通用电子数字计算机研制成功,字长为32位,内存容量为1K字节,有加减乘除等16条指令,主要用于弹道计算。

1963年,中国第一台大型晶体管电子计算机——109机研制成功,它在我国两弹试验中发挥了重要作用,被誉为“功勋机”。

1973年年初,由北京大学、北京有线电厂和燃化部等有关单位共同研制成功中国第一台百万次集成电路电子计算机,字长为48位,存储容量为13KB。

1977年4月,安徽无线电厂、清华大学和四机部六所联合研制成功我国第一台微型计算机DJS-050机。

1983年,“银河Ⅰ号”巨型计算机研制成功,运算速度达1亿次/s。

1987年,第一台国产的286微机——长城286正式推出。

1993年,中国第一台10亿次巨型银河计算机Ⅱ型通过鉴定。1994年,银河计算机Ⅱ型在国家气象局投入正式运行,用于天气中期预报。

1995年,曙光1000大型机通过鉴定,其峰值可达25亿次/s。

1997年,银河Ⅲ并行巨型计算机研制成功,该机采用可扩展分布共享存储并行处理体系结构,由130多个处理结点组成,峰值性能为130亿次/s浮点运算,系统综合技术达到20世纪90年代中期国际先进水平。

1998年,中国微机销量达408万台,国产占有率高达71.9%。

1999年,银河四代巨型机研制成功。

2000年,我国自行研制成功高性能计算机“神威Ⅰ”,其主要技术指标和性能达到国际先进水平。我国成为继美国、日本之后,世界上第三个具备研制高性能计算机能力的国家。

2004年上半年,推出曙光4000A超级服务器,该机处理器总数为2560个,内存总容量为5TB,磁盘总容量为42TB,峰值浮点运算速度为11.2万亿次/s。

2008年9月,曙光5000A集群超级计算机(魔方)在天津成功下线,曙光5000A以峰值速度为230万亿次/s、Linpack测试值为180万亿次/s的成绩再次跻身世界超级计算机前十名,这一成绩让中国成为世界上第二个可以研发生产超百万亿次超级计算机的国家。它拥有30720颗计算核心。

1.3.4 计算机的分类

计算机的种类有很多,从不同角度对计算机有不同的分类方法。下面从计算机处理数据的方式、使用范围、规模和处理能力三个角度进行说明。

1. 按计算机处理数据的方式分类

- 数字计算机

数字计算机处理的是非连续变化的数据,这些数据在时间上是离散的,输入是数字量,输出也是数字量,如职工编号、年龄、工资数据等。基本运算部件是数字逻辑电路,因此其运算精度高、通用性强。

- 模拟计算机

模拟计算机处理和显示的是连续的物理量,所有数据用连续变化的模拟信号表示,其基本运算部件是由运算放大器构成的各类运算电路。模拟信号在时间上是连续的,通常称为模拟量,如电压、电流、温度都是模拟量。一般说来,模拟计算机不如数字计算机精确、通用性不强,但解题速度快,主要用于过程控制和模拟仿真。

- 数模混合计算机

数模混合计算机兼有数字和模拟两种计算机的优点,既能接受、输出和处理模拟量,又能接受、输出和处理数字量。

2. 按计算机使用范围分类

- 通用计算机

通用计算机是指为解决各种问题,具有较强的通用性而设计的计算机。该机适用于一般的科学计算、学术研究、工程设计和数据处理等广泛用途,这类机器本身有较大的适用面。

- 专用计算机

专用计算机是指为适应某种特殊应用而设计的计算机,具有运行效率高、速度快、精度高等特点。该机一般用在过程控制中,如智能仪表、飞机的自动控制、导弹的导航系统等。

3. 按计算机的规模和处理能力分类

- 巨型计算机

巨型计算机是指运算速度快、存储容量大,每秒可达 1 亿次以上浮点运算速度,主存容量高达几百兆字节甚至几百万兆字节,字长可达 32 位的机器。这类机器价格相当昂贵,主要用于复杂、尖端的科学研究领域,特别是军事科学计算。国防科技大学研制的“银河”和国家智能中心研制的“曙光”都属于这类机器。

- 大/中型计算机

大/中型计算机是指通用性能好、外部设备负载能力强、处理速度快的机器。这类机器运算速度在 100 万次至几千万次/秒,字长为 32~64 位,主存容量在几十兆字节至几百兆字节左右。它有完善的指令系统、丰富的外部设备和功能齐全的软件系统,并允许多个用户同时使用。这类机器主要用于科学计算、数据处理或做网络服务器。

- 小型计算机

小型计算机具有规模较小、结构简单、成本较低、操作简单、易于维护、与外部设备连

接容易等特点,是在 20 世纪 60 年代中期发展起来的一类计算机。当时的小型机字长一般为 16 位,存储容量在 32KB 至 64KB 之间。DEC 公司的 PDP 11/20 到 PDP 11/70 是这类机器的代表。当时微型计算机还未出现,因而小型计算机得以广泛推广应用,许多工业生产自动化控制和事务处理都采用小型机。近期的小型机,像 IBM AS/400,其性能已大大提高,主要用于事务处理。

- 微型计算机

微型计算机(简称微机)是以运算器和控制器为核心,加上由大规模集成电路制作的存储器、输入输出接口和系统总线构成的体积小、结构紧凑、价格低但又具有一定功能的计算机。如果把这种计算机制作在一块印刷线路板上,则称之为单板机。如果在一块芯片上集成了运算器、控制器、存储器和输入输出接口,则称之为单片机。以微机为核心,再配以相应的外部设备(例如,键盘、显示器、鼠标器、打印机)、电源、辅助电路和控制微机工作的软件就构成了一个完整的微型计算机系统。

- 工作站

工作站是指为了某种特殊用途而将高性能的计算机系统、输入输出设备与专用软件结合在一起的系统。它的独到之处是有大容量主存、大屏幕显示器,特别适合计算机辅助工程。例如,图形工作站一般包括主机、数字化仪、扫描仪、鼠标器、图形显示器、绘图仪和图形处理软件等。它可以完成对各种图形与图像的输入、存储、处理和输出等操作。

- 服务器

服务器是在网络环境下为多用户提供服务的共享设备,一般分为文件服务器、打印服务器、计算服务器和通信服务器等。该设备连接在网络上,网络用户在通信软件的支持下远程登录,共享各种服务。

目前,微型计算机与工作站、小型计算机乃至中、大型机之间的界限已经愈来愈模糊。无论按哪一种方法分类,各类计算机之间的主要区别是运算速度、存储容量及机器体积等。

1.3.5 计算机的特点

- 运算速度快

目前最快的巨型机每秒钟能进行数千亿次运算。

- 计算精度高

由于计算机内部采用二进制数进行运算,使数值计算非常精确。一般计算机可以有十几位以上的有效数字。

- 具有“记忆”和逻辑判断的能力

计算机的存储设备可以把原始数据、中间结果、计算结果、程序等信息存储起来以备使用,存储能力取决于所配备的存储设备的容量。

计算机不仅能进行计算,还具有逻辑判断能力,并能根据判断的结果自动决定以后执行的命令,因而能解决各种各样的问题。

- 内部的操作是自动化的

由于程序和数据存储在计算机中,一旦向计算机发出运行指令,计算机就能在程序的

控制下,按事先规定的步骤一步一步执行,直到完成指定的任务为止。这一切都是计算机自动完成的,不需要人工干预。

1.3.6 计算机系统组成

一个完整的计算机系统包括硬件系统和软件系统两部分。组成一台计算机的物理设备的总称是计算机硬件系统,是实实在在的物体,是计算机工作的基础。指挥计算机工作的各种程序的集合称为计算机软件系统,是计算机的灵魂,是控制和操作计算机工作的核心。

• 硬件系统

计算机硬件(computer hardware)或称硬件平台,是指计算机系统所包含的各种机械的、电子的、磁性的装置和设备,如运算器、磁盘、键盘、显示器、打印机等。每个功能部件各司其职、协调工作,缺少其中任何一个就不能成为完整的计算机系统。

硬件是组成计算机系统的物质基础,不同类型的计算机,其硬件组成是不一样的。从计算机的产生发展到今天,各种类型的计算机都是基于冯·诺依曼思想而设计的。这种计算机的硬件系统结构从原理上来说主要由运算器、控制器、存储器、输入设备和输出设备五部分组成。

硬件系统组成如图 1.2 所示。

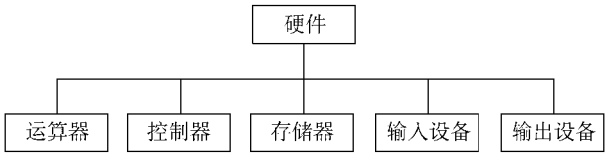


图 1.2 计算机硬件的组成

• 软件系统

计算机软件(computer software)是相对于硬件而言的。它包括计算机运行所需的各种程序、数据及有关资料。脱离软件的计算机硬件称为“裸机”，它是不能做任何有意义的工作的,硬件只是软件赖以运行的物质基础。因此,一个性能优良的计算机硬件系统能否发挥其应有的功能,关键取决于所配置的软件是否完善和丰富。软件不仅提高了机器的效率、扩展了硬件功能,也方便了用户使用。

软件内容丰富、种类繁多,通常根据软件用途可将其分为系统软件和应用软件两类。系统软件是用于管理、控制和维护计算机系统资源的程序集合,如操作系统等。应用软件是在系统软件下二次开发的、为解决特定问题而编制的应用程序或用户程序等。利用应用程序用户可以创建用户文档,如字处理软件、表处理软件等。

软件系统组成如图 1.3 所示。

计算机系统的组成如图 1.4 所示。

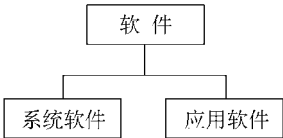


图 1.3 计算机软件分类

一个完整的计算机系统,硬件和软件是按一定的层次关系组织起来的。最内层是硬件,然后是软件中的操作系统,而操作系统的外层为其他软

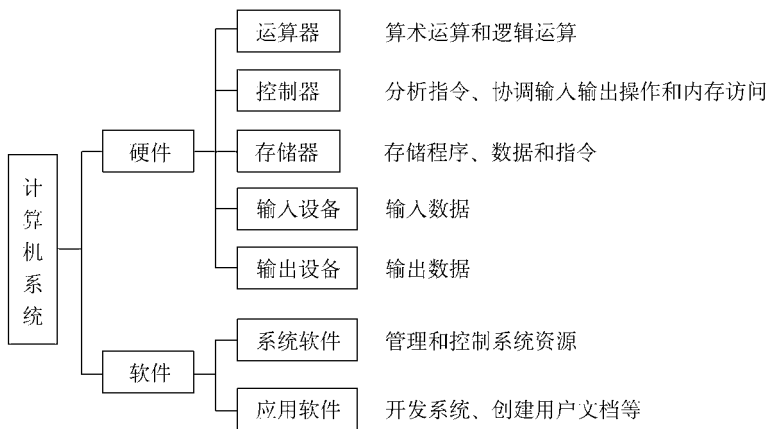


图 1.4 计算机系统组成

件,最外层是用户程序。所以说,操作系统是直接管理和控制硬件的系统软件,自身又是系统软件的核心,同时也是用户与计算机打交道的桥梁——接口软件。

计算机系统的层次结构如图 1.5 所示。

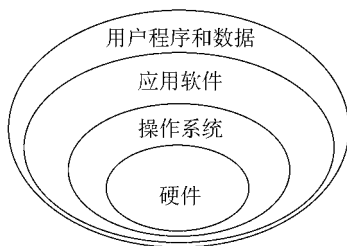


图 1.5 计算机系统的层次结构

1.4 信息的数字化表示和编码

要了解计算机,首先要了解信息在计算机中的表示,这是计算机得以运转的基础。本节先介绍有关数制、编码和如何将信息数字化的问题。

1.4.1 数制

数制是人们对数量计算的一种统计规律。日常生活中最常遇到的进位计数制是十进制,在数字系统中被广泛采用的则是二进制、八进制、十六进制。

一种进位计数包含着两个基本的因素。

(1) 基数

它是计数制中所用到的数码的个数,一般来说,基数为 R 的计数制(简称 R 进制)中,包含的是 $0, 1, \dots, R-1$ 等数码,进位规律是“逢 R 进一”,即每个数位计满 R 就向高位进 1,称为 R 进制计数制。