

前 言

当今信息技术日新月异,融入了社会生活的方方面面,深刻地改变了人类的思维、生产、生活和学习方式。人工智能、5G 移动通信、区块链、物联网、云计算及大数据等新概念和新技术的出现,在社会经济、人文科学、自然科学的许多领域引发了一系列革命性的突破。

计算机组成原理是计算机专业的核心课程,既具有很强的技术性,又具有很强的应用性、时代性,课程所涉及的内容反映了信息技术及其应用发展状况。通过本课程的学习,学生不仅能够掌握计算机硬件系统的内部结构和工作原理,而且还能利用所学知识、方法来分析和解决实际问题,为后续其他专业课程的学习打下良好的基础。

本书作为一本教材,编者希望将先进的教学方法和理念融入其中。BOPPPS 是一种以教学目标为导向、以学生为中心的有效教学结构和教学模式,它由导言(bridge-in)、学习目标(objective)、前测(pre-assessment)、参与式学习(participatory learning)、后测(post-assessment)和总结(summary)6个教学环节构成,“BOPPPS”由这6个教学环节英文单词的首字母组成。编者在教学中采用这种教学模式,由浅入深,从学生身边能够看到的科技常识入手,通过恰当的实例“导入”,让枯燥的硬件系统原理更形象化、具象化,而每章的习题会围绕前测和后测环节展开,在这些习题的设计上注重培养学生运用课程知识解释或解决实际问题的能力。本书提供每章1~2节课程的BOPPPS教学案例设计稿,感兴趣的教师请到北京邮电大学出版社官方网站下载,欢迎交流。

本书还有机地融入了思政元素,例如书中体现的科学世界观和方法论,以及我国科技工作者对当代计算机发展的贡献等。在BOPPPS教学案例设计稿中,还提供了更多与课程内容有机融合的思政案例。

计算机系统与其他任何系统一样,由一组相互关联的部件组成。说明系统最好的方法是描述其结构和功能,结构是各个部件之间具有层次的互连方式,而功能是各个部件的操作。一般地,每个部件还可以进一步分解为若干个子部件,通过描述每个子部件的结构和功能,反过来就能揭示这个部件的结构和功能。本书章节的安排是按照这种自顶向下的顺序逐层展开的,从总体描述到子部件描述,从结构介绍到功能原理阐述,结构和功能两者密不可分。例如,计算机系统的主要部件是处理器、内存和输入/输出系统,而处理器的主要部件是控制单元、寄存器、ALU和执行单元,其中控制单元为处理器所有部件的操作提供控制信号,控制单元传统上是基于微程序控制器原理工作的,为了描述这种设计思想,需要进一步分析控制存储器、微命令、微指令序列逻辑和寄存器。

因此,在学习本书时,建议读者搞清楚前后章节的联系,并熟悉自顶向下的描述方式。

本课程内容与多门专业基础课和专业课的内容有紧密的联系,例如数字电路、微机原理与接口技术、操作系统、各种高级编程语言、编译原理等。其中,数字电路是本课程的先

修课程,本课程中所讲述的功能部件,如加法器、算术逻辑单元、寄存器、存储器、缓冲器等各类基本硬件电路都属于数字电路课程的内容,但不同点在于数字电路主要介绍门级电路,而本课程主要介绍寄存器级(Register Transfer Level,RTL)以上层次的电路。

本课程介绍计算机系统通用知识框架,着重介绍的是其基本概念、基本原理和结构,而微机原理与接口技术课程的主要讲授对象是计算机发展历程中的第四代机型——微型计算机,介绍其CPU、存储器、汇编语言、指令系统以及外围接口电路,可以看成本课程的实际应用案例。因此,在本课程中的一些较为抽象的原理如寻址方式、存储器扩展,在后者中会有更具体的实现方式。

操作系统主要用于管理计算机资源,包括本课程介绍的处理器资源、内存资源、虚拟存储器资源、外设和磁盘空间等,因此,在这些内容上两门课程具有紧密的联系。

高级编程语言通常设计有多种数据类型,如整型、浮点型、字符串型、字节型等,在计算机内部表示和存储时都会涉及编码方式,这与本课程中的数据表示方法联系紧密;另外,高级编程语言经过编译系统的编译之后,最终会变成机器指令,并在本课程的计算机硬件上执行。

在学习本课程时,建议读者主动探究课程联系,这对读者深入了解并掌握计算机专业学科系统的知识将具有很大的帮助。

本书按照计算机组成原理大纲的要求进行编写,全书共分为7章。第1章为计算机系统概述,主要介绍计算机发展历程,计算机的层级结构、组成,以及计算机性能指标。第2章为运算方法和运算器,从信息表示开始,系统地介绍数制与编码、定点数的表示和运算、浮点数的表示和运算、数据校验方法,以及算术逻辑单元的结构和功能。第3章为存储系统,介绍存储系统的概念、基本结构、工作过程、各类存储器的特点和工作原理,包括主存储器、高速缓冲存储器、双端口存储器及虚拟存储器等。第4章为指令系统,介绍指令系统的基本概念、指令格式、指令种类及寻址方式,并给出两种基于RISC技术的指令系统实例。第5章为中央处理器,从中央处理器的基本结构和功能开始,接续上一章指令系统,具体说明指令执行过程、数据通路的结构和功能,然后介绍控制器的结构和工作原理、流水线的概念和实现方式。第6章为总线系统,介绍总线的概念、总线结构、总线控制与通信方式,并给出微型计算机的总线系统实例。第7章为输入/输出系统,介绍I/O系统的基本概念、外部设备的分类和特点、I/O接口控制器,以及I/O控制方式。

本书中使用的符号:

- **【注意】**:在学习概念和工作原理时需要注意的问题。
- **【课程关联】**:本课程与其他课程的联系点。

本书的编写得到了全国高等院校计算机基础教育研究会计算机基础教育教学研究项目的支持,在此表示衷心感谢。

由于编者水平有限,书中难免存在不足之处,恳请广大读者及同行予以批评和指正,谢谢!

编者
2022年11月

目 录

第 1 章 计算机系统概述	1
1.1 计算机的发展历程	1
1.1.1 计算机的起源	1
1.1.2 计算机的发展简史	2
1.1.3 计算机硬件发展的特点和规律	3
1.1.4 计算机的形态和分类	4
1.1.5 丘奇-图灵论题	6
1.2 计算机的基本组成	6
1.2.1 冯·诺依曼计算机原型	6
1.2.2 计算机硬件组成和工作原理	8
1.2.3 个人计算机的物理结构	12
1.2.4 计算机软件分类	13
1.3 计算机系统的层级结构	13
1.3.1 计算机系统的层级结构概述	13
1.3.2 计算机体系结构和计算机组成	14
1.3.3 计算机硬件和软件的逻辑等价性	15
1.4 计算机系统的性能指标	16
1.5 本章小结	18
习题	19
第 2 章 运算方法和运算器	21
2.1 数据信息和二进制编码	21
2.1.1 数据信息	21
2.1.2 进位计数制	22
2.2 数据的编码表示	25
2.2.1 符号位的表示	25
2.2.2 定点数表示	29
2.2.3 浮点数表示	31
2.2.4 十进制数的表示	35
2.2.5 非数值数据的表示	36
2.2.6 数据的宽度和存储顺序	41
2.3 数据的校验	43

2.3.1	奇偶校验	43
2.3.2	海明校验*	45
2.3.3	循环冗余校验	46
2.4	定点运算和定点运算器	47
2.4.1	定点数的加、减法运算	48
2.4.2	定点数的乘、除法运算	52
2.4.3	逻辑运算	53
2.4.4	定点运算器	55
2.5	浮点运算和浮点运算器	60
2.5.1	浮点数的加、减法运算	60
2.5.2	浮点数的乘、除法运算	66
2.5.3	浮点运算器	66
2.5.4	浮点运算流水线	67
2.6	本章小结	68
	习题	70
第3章	存储系统	73
3.1	存储器概述	73
3.1.1	存储器容量单位	73
3.1.2	存储器的分类	74
3.1.3	存储器的多级结构	76
3.2	半导体存储器	77
3.2.1	随机读写存储器	77
3.2.2	微型计算机内存	79
3.2.3	只读存储器	81
3.3	主存储器	83
3.3.1	主存储器技术指标	83
3.3.2	主存储器的基本组成	84
3.3.3	主存储器扩展	86
3.4	高速存储器	92
3.4.1	双端口存储器	92
3.4.2	单体多字存储器	93
3.4.3	多模块存储器	94
3.4.4	相联存储器	97
3.5	高速缓冲存储器	98
3.5.1	cache的基本原理	98
3.5.2	cache-主存地址映射方式	102
3.5.3	替换策略	107
3.5.4	写操作策略	107
3.6	虚拟存储器	108

3.6.1	虚拟存储器的基本原理	108
3.6.2	页式虚拟存储器	109
3.6.3	段式虚拟存储器	110
3.6.4	段页式虚拟存储器	111
3.6.5	快表	111
3.6.6	内页表和外页表	112
3.6.7	CPU 访存过程总结	112
3.6.8	替换算法	113
3.7	本章小结	114
	习题	115
第 4 章	指令系统	119
4.1	指令系统概述	119
4.1.1	机器指令和指令系统	119
4.1.2	指令系统的设计要求	119
4.2	指令格式	120
4.2.1	指令的组成	120
4.2.2	操作码	121
4.2.3	地址码	121
4.2.4	二地址指令类型	123
4.3	指令字长和指令助记符	124
4.3.1	指令字长	124
4.3.2	定长指令和不定长指令	125
4.3.3	指令助记符	125
4.4	指令种类	126
4.4.1	指令操作种类	126
4.4.2	操作数风格分类	127
4.5	寻址方式	128
4.5.1	指令寻址方式	128
4.5.2	操作数寻址方式	129
4.6	RISC 技术	136
4.7	指令系统实例	138
4.7.1	ARM 指令系统	138
4.7.2	RISC-V 指令系统	141
4.8	本章小结	145
	习题	146
第 5 章	中央处理器	148
5.1	CPU 的功能和组成	148
5.1.1	CPU 的功能	148

5.1.2 CPU 的基本组成	149
5.1.3 CPU 中的主要寄存器	151
5.1.4 操作控制器和时序发生器	153
5.2 CPU 的工作过程	153
5.2.1 指令的执行过程	153
5.2.2 指令周期	154
5.2.3 数据通路	159
5.2.4 时序信号和时序发生器	165
5.2.5 控制方式	166
5.3 操作控制器	167
5.3.1 组合逻辑控制器	167
5.3.2 微程序控制器	169
5.3.3 硬布线逻辑控制器与微程序控制器的比较	174
5.4 流水线技术	174
5.4.1 并行处理技术概述	174
5.4.2 流水线技术概述	175
5.4.3 流水线中的冒险处理	180
5.5 本章小结	183
习题	184
第 6 章 总线系统	187
6.1 总线系统概述	187
6.1.1 总线的基本概念	187
6.1.2 总线的特性及其标准化	189
6.1.3 总线性能指标	190
6.2 总线结构	191
6.2.1 总线内部结构	191
6.2.2 总线接口	193
6.2.3 总线的连接方式	193
6.3 总线控制与通信	197
6.3.1 总线控制方式	197
6.3.2 总线通信方式	199
6.3.3 信息传送方式	200
6.4 总线系统实例	201
6.4.1 微型计算机多总线结构	202
6.4.2 微型计算机总线介绍	203
6.5 本章小结	205
习题	206

第 7 章 输入/输出系统	208
7.1 外围设备的分类和特点	208
7.1.1 外围设备的基本结构	208
7.1.2 外围设备的分类	209
7.1.3 外围设备的特点	209
7.1.4 常见的输入/输出设备	209
7.1.5 外部存储器	212
7.2 I/O 接口	217
7.2.1 I/O 接口的功能和基本结构	217
7.2.2 I/O 编址方式	219
7.3 输入/输出控制方式	220
7.4 程序中断方式	222
7.4.1 中断的基本概念	222
7.4.2 I/O 中断的产生和执行	222
7.4.3 中断源和中断分级	223
7.4.4 程序中断的接口电路	224
7.5 DMA 方式	227
7.5.1 DMA 的基本概念	227
7.5.2 基本的 DMA 控制器	229
7.5.3 选择型 DMA 控制器和多路型 DMA 控制器	232
7.6 本章小结	234
习题	235
参考文献	237

第 1 章 计算机系统概述

📖 本章学习目标

本章为计算机系统概述,从计算机的发展历程讲起,由冯·诺依曼计算机原型引入计算机硬件的主要组成部件、工作原理,最后介绍计算机系统的层级结构与性能指标。

- ① 计算机的起源与发展简史、计算机硬件发展的特点和规律,以及计算机的形态和分类。
- ② 计算机硬件的组成部件、工作原理。
- ③ 计算机软件分类。
- ④ 计算机系统的层级结构。
- ⑤ 计算机系统的性能指标。

1.1 计算机的发展历程

辩证唯物主义认为,事物发展的根本原因在于其内部的矛盾性,人类的认识活动是在实践和认知的矛盾运动中展开的,科学技术作为人类特殊的认识与实践活动,其发展也有其自身复杂的矛盾运动,这些矛盾运动构成了推动科学技术发展的内在动力。从科学技术发展的历程看,科学技术的内容和方向,连同其发展速度,都是社会需求发展的结果。

计算机的发明是对人脑智力的继承和延伸,计算机产生的动力是人们想发明一种能够进行科学计算的机器。它一诞生就立即成了先进生产力的代表,掀起了工业革命之后的又一场科学技术革命,在诞生之后的几十年中,计算机已经广泛应用到社会的各个领域,随着计算机日益智能化的发展,人们干脆将微型计算机称为“电脑”。

按所处理变量的性质,计算机可划分为数字计算机和模拟计算机。20 世纪初,随着电子管的诞生,计算机开始进入电子数字计算机(electronic digital computer)时代,电子数字计算机简称为计算机(computer)。从信息处理角度而言,计算机是指能够按一系列指令对数据进行自动化处理的机器,具体地说,是一种能够接收信息、存储信息并按照存储在其内部的程序对输入信息进行加工、处理,得到期望处理结果,并将处理结果输出的自动化机器。

1.1.1 计算机的起源

在第二次世界大战中,为了解决军事上的科学计算难题,人们开始研究将电子管作为“开关”来提高计算机的运算速度。1946 年,美国宾夕法尼亚大学学者 J. W. 莫利奇(J. W. Mauchly)和 J. P. 艾克特(J. P. Eckert)等人研制了一台“电子数值积分计算机”(Electronic Numerical Integrator and Calculator, ENIAC)。这台叫作 ENIAC(埃尼阿克)的计算机占地面积为 170 m²,使用了 18 000 多个电子管、6 000 多个开关,每秒能够实现 5 000 次加法运算。它内部采用十进制运算,人们需要通过重新连线(拨动开关和插拔线缆)为它编写不同的程序。

这台计算机的问世标志着电子数字计算机时代的开始。

为了满足情报战中“密码编制和破译”的科学计算问题,1943年,战时服务机构技术人员采用英国数学家艾伦·麦席森·图灵(Alan Mathison Turing)提出的概念研制出了CO-LOSSUS(巨人机),其能执行计数运算、二进制算术运算及布尔代数逻辑运算,后来人们共生产了10台CO-LOSSUS。在CO-LOSSUS的辅助下,军方出色地完成了密码破译工作,由于此项工作严格保密,所以直到20世纪70年代其才被披露,CO-LOSSUS的发明年代虽然早于ENIAC,但目前世界公认的第一台计算机还是ENIAC。

1945年,数学家约翰·冯·诺依曼(John von Neumann)在“离散变量自动电子计算机方案”(Electronic Discrete Variable Automatic Computer,EDVAC)中第一次提出了程序存储的概念,几乎同时,数学家图灵也提出了这个观点。1946年,冯·诺依曼和他在普林斯顿高等研究院(Institute for Advance Study,IAS)的同事,开始按照“程序存储”的概念来设计计算机,直到1952年才设计完成。这台计算机被称为IAS计算机,IAS计算机的体系结构成为后续各种数字计算机遵循的典范。

1.1.2 计算机的发展简史

从1946年世界第一台电子数字计算机诞生至今,计算机的发展经历了五代。

1. 第一代计算机(1946—1957年):电子管计算机

第一代计算机使用电子管作为数字逻辑元件,在此期间人们确定了计算机的基本体系结构,使用机器语言编写程序。第一代计算机的主要应用领域是科学计算和商业数据处理。

第一代商用计算机的主要代表有UNIVAC公司的UNIVACII、UNIVAC100系列计算机,以及IBM公司的700/7000系列计算机。

2. 第二代计算机(1958—1964年):晶体管计算机

第二代计算机使用晶体管作为数字逻辑元件,并且开始支持用高级编程语言编写的程序,在此期间出现了系统软件,其负责加载程序、将数据传输到外设,以及执行公共计算等,在应用领域出现了“工业控制机”。

第二代计算机的典型代表为IBM 7904,它与IAS计算机相比有两个显著的特点:一是使用了数据通道,数据通道是具有专门处理器与指令集的独立I/O模块;二是使用了多路复用器,用作数据通道、CPU和主存储器的中央节点。

1958年,中国科学院计算技术研究所研制成功了我国第一台小型电子管通用计算机103机(八一型)。

3. 第三代计算机(1965—1971年):中小规模集成电路计算机

第三代计算机开始使用中小规模集成电路作为数字逻辑元件,这期间操作系统问世,出现了系列机和小型机,系列机的典型代表是IBM System/360计算机,小型机的典型代表是DEC PDP-8计算机。

1965年,中国科学院计算技术研究所研制成功了我国第一台大型晶体管计算机109乙机,之后又推出了109丙机,109丙机在我国两弹的试验中发挥了重要作用。

第三代计算机之后,在计算机发展阶段的划分上没有统一的意见,下面介绍较为普遍采用的划分方法。

4. 第四代计算机(1972—1990年):大规模集成电路和超大规模集成电路计算机

第四代计算机开始使用大规模和超大规模集成电路作为数字逻辑元件,伴随着微处理器

的问世,微型计算机开始登上历史舞台,典型代表是 IBM PC。

这个阶段是我国计算机迅猛发展的时期。1974年,清华大学等单位联合设计并研制成功了采用集成电路的 **DJS-130 小型计算机**。1983年,国防科技大学研制成功了运算速度每秒上亿次的**银河-I号巨型机**,成为我国高性能计算机发展史上的重要里程碑。1985年,电子工业部计算机管理局研制成功了与 IBM PC 兼容的**长城 0520CH 微型计算机**。

5. 第五代计算机(1991年至今):甚大规模集成电路计算机

这个阶段开始使用甚大规模集成电路作为数字逻辑元件,运算速度提升到每秒十亿次以上。同时,计算机体系结构也出现变革,呈现向多个方向发展的趋势。

① 智能计算机。第五代智能计算机是将信息采集、存储、处理、通信同人工智能结合在一起的计算机系统,它除了能进行数值计算或处理一般的信息外,主要是面向知识处理,具有形式化推理、联想、学习和解释的能力,能够帮助人们进行判断、决策、开拓未知领域和获得新的知识。人机之间可以直接通过自然语言(声音、文字)或图形图像交换信息。

② 生物计算机。生物计算机又称仿生计算机,它是以生物芯片取代在半导体硅片上数以万计的晶体管而制成的计算机。生物计算机已成为当今计算机技术的前沿,生物计算机比硅片计算机在速度、性能上有质的飞跃。

③ 量子计算机。量子计算机是一类遵循量子力学规律进行高速数学和逻辑运算,具有存储和处理信息能力的物理装置,它以量子态为记忆单元和信息储存形式,量子计算机元件尺寸可小到原子或分子量级。不同于传统计算机,量子计算机的突出特征是存储数据的对象是量子比特,量子比特可以同时是 0 和 1,即允许“叠加态”共存,从而拥有更强大的并行能力。量子计算机的另外两个特征是采用量子编码和量子算法来进行数据操作。量子计算技术是颠覆性技术,关系到未来发展的基础计算能力,它成为很多国家竞相争夺的信息时代的制高点。

1.1.3 计算机硬件发展的特点和规律

自从 1946 年第一台计算机诞生以来,计算机的发展表现为运算速度大幅提升,成本大幅降低,体积大幅缩小,产量大幅提升。下面是计算机硬件技术发展过程中表现出来的一些特点和规律。

1. 计算机硬件技术的发展伴随着微电子技术的发展

1947年**晶体管**的问世为微电子技术的发展奠定了基础。微电子技术始于 20 世纪 50 年代,而后出现的集成电路更是革命性的技术,20 世纪 70 年代后期微电子技术进入发展高峰阶段。以计算机中的存储器和 CPU 为例,1970 年仙童半导体公司生产出第一个较大容量的半导体存储器,逐渐代替了存储密度低、体积大的磁芯存储器,从此半导体存储器向着更大容量、更低成本(每位价格)和更短存取时间方向发展,单片存储器的容量从 KB 数量级提升到目前的 GB 数量级。1971 年,Intel 公司开发出第一片微处理器芯片 4004,第一次将 CPU 的所有元件封装在一个芯片内部,从这片 4 位微处理器开始,微处理器位数经历了 8 位、16 位和 32 位的发展,到目前主流微处理器已是 64 位,单个芯片上集成的晶体管数量从几千增加到数以亿计。

2. 摩尔定律

1965 年,戈登·摩尔(Gordon Moore)在整理关于计算机存储器发展趋势的数据时发现了一个惊人的现象:一个集成电路芯片可以容纳的晶体管数目每年便会增加一倍。令很多人包括摩尔吃惊的是:这一现象在之后的几年中得到了证实。到了 20 世纪 70 年代中期,这种趋势

放缓到每 18 个月增加一倍,并在此后一直保持这种发展趋势。这个规律不仅预测了存储器芯片的发展情况,也较为精确地反映了处理机芯片上晶体管及磁盘驱动器存储容量的发展。图 1-1 所示为集成电路上晶体管数量随时间的增长趋势,图中纵坐标采用对数形式,代表每个集成电路芯片上晶体管的数量, m 和 bn 分别代表 10^6 和 10^9 。

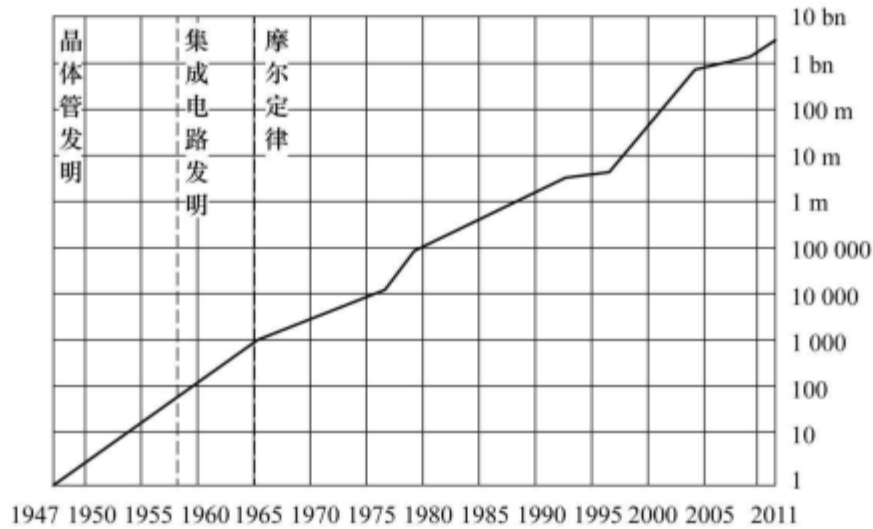


图 1-1 集成电路上晶体管数量随时间的增长趋势

摩尔定律揭示了信息技术进步的速度,也成了行业界对计算机发展趋势的预测基础。需要注意的是,摩尔定律不同于自然科学定律,它只是对信息技术发展速度的预测和估计。

3. 计算机硬件技术的发展也伴随着计算机软件技术的发展

以计算机系统软件为例。在早先的计算机中,人们直接使用机器语言(即机器指令代码)来编写程序,需要熟悉二进制(或十六进制)的机器指令,编程工作枯燥、繁琐,且容易出错,这大大地限制了计算机的使用。后来人们想出了一种办法,将二进制数用有意义的符号和数字代替,这种助记符称为汇编语言,并开发了称为“汇编器”的工具,借助于汇编器将助记符转换为机器指令,从而提高了编写程序的效率。但是助记符仍然是面向机器的,不同型号机器的助记符是不同的。之后,人们又发明了更接近数学符号的算法语言,比如 BASIC、FORTRAN、C、C++ 和 Java 等,使用算法语言编写程序大大地提升了效率。但是,用算法语言编写的程序(称为源程序)无法直接在计算机上运行,还需要使用编译器和链接器,将源程序变成机器可以识别并执行的机器代码(称为可执行文件),并通过加载器将机器代码加载到主存储器中,才能由处理器获取并执行。

随着计算机运算速度的飞速提升,计算机结构和程序的复杂程度越来越高,人们认识到,只有让计算机自己来管理和调度程序的运行,才能摆脱繁琐的手工管理工作,并进一步提高计算机的使用效率,于是出现了一类程序——操作系统。操作系统是一套系统软件,专门负责管理计算机资源(如处理器、内存、外部设备以及各种编译与应用程序)并自动调度用户作业,从而让多个用户能够共享一台计算机系统,操作系统的出现使得计算机的使用效率得到了大大提升。

1.1.4 计算机的形态和分类

计算机的物理形态千差万别,运算速度也差别巨大。图 1-2(a)所示为我国神威·太湖之光超级计算机,其体积庞大,需要放置在专用的机房,运行峰值性能达 125.436 PFlops,持续性

能达 93.015 PFlops。图 1-2(b)所示为常见的个人计算机(Personal Computer, PC),由于在使用时通常被放置在办公桌面上,因此也称为台式计算机,或桌面计算机,它可用于通用计算。图 1-2(c)所示为智能手表,其除了具备计时基本功能外,还具备运动和睡眠情况记录、心率和血氧饱和度测量等专用功能,这类设备功耗很低。



图 1-2 计算机不同的物理形态

按照计算机用途分类,可将其分为通用计算机和专用计算机。

所谓通用计算机是为通用需求和目的而设计的计算机,其通用性强、应用面广。从体积、复杂度、功耗、性能指标、数据存储容量、指令系统规模和价格等维度,又可将通用计算机分为 6 类,即超级计算机(supercomputer)、大型机(mainframe)、服务器(server)、工作站(workstation)、微型计算机(microcomputer)和单片机(single-chip computer),如图 1-3 所示。超级计算机是功能最强、运算速度最快、存储容量最大的一类计算机,通常用于科学计算,以解决经济、科技、国防等领域的复杂计算问题。超级计算机性能的强弱也是一个国家科技发展水平及其综合国力的重要标志之一。而单片机则将计算机主要部件集成到一个芯片上,其体积小,结构简单,性能相对较差,但价格便宜。介于两者之间的大型机、服务器、工作站和微型计算机,它们的结构规模和性能指标依次递减,人们在工作生活中常使用的台式计算机和笔记本计算机(laptop computer)都属于微型计算机。

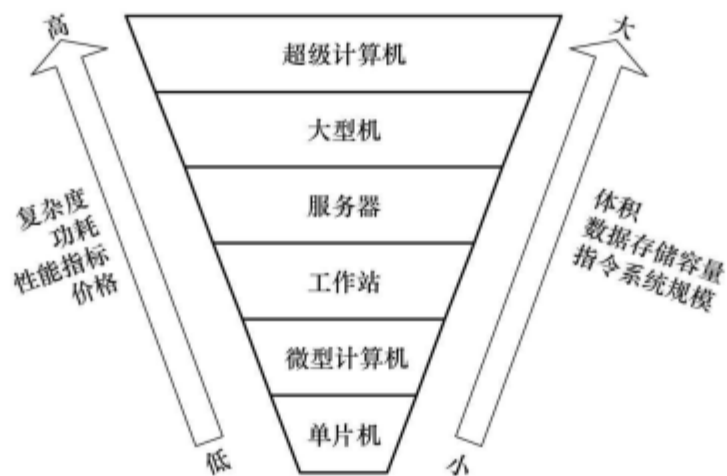


图 1-3 通用计算机分类

所谓专用计算机是针对某类具体应用而设计的计算机,其设计目标追求该应用中最有效、最快和最经济,通常在其他用途方面的适应性较差。嵌入式计算机或嵌入式系统(embedded system)是为特定应用而设计的专用计算机,它以计算机技术为基础,具有软件、硬件可裁剪的特点,以满足不同应用在功能、可靠性、成本、体积、功耗等方面的需求。嵌入式计算机是目前

使用最为广泛的计算机形态之一,例如针对网络应用而设计的路由器,针对通信应用而设计的智能手机,针对音频和视频应用而设计的音视频播放机,针对工业控制而设计的可编程逻辑控制器(Programmable Logic Controller,PLC),医疗设备中的电子计算机断层扫描仪器(CT)、B型超声断面显像仪,家用电器中的智能电视、冰箱、洗衣机和微波炉等。

1.1.5 丘奇-图灵论题

任何类别的计算机,其基本功能都是用来执行计算任务的。根据**丘奇-图灵论题**(Church-Turing thesis),任何一台具有基本功能的计算机,在理论上都能够执行任何其他计算机可以执行的任务。因此,如果不考虑**时间**和**存储容量**,性能和复杂程度相差甚远的各种计算机,大到超级计算机,小到智能手表、手环都能够执行**相同的运算任务**,但由于计算机存储容量不同、执行任务的时间不同而呈现性能差别。在实际应用中,性能高低不同的计算机都有其各自适用的领域。

1.2 计算机的基本组成

一台完整的计算机由**硬件**(hardware)和**软件**(software)两部分组成,其中硬件是有形实体,主要由各种电子器件组成。一台硬件的计算机称为“**实际机器**”或“**裸机**”。软件是无形的,但显示了更高层的逻辑功能,并使硬件最大限度地发挥作用,配备了软件的计算机又称为“**虚拟机器**”。

计算机硬件是组成计算机的所有**电子器件**和**机电装置**的总称,例如,人们熟知的台式计算机和笔记本计算机通常包括集成电路芯片、分离形式的电子元器件,以及散热器、风扇等机电部件。

下面首先介绍计算机的原型机——冯·诺依曼计算机,分析其基本结构和功能,其次介绍现代计算机硬件的组成,最后介绍计算机软件分类。

1.2.1 冯·诺依曼计算机原型

根据冯·诺依曼提出的“**存储程序**”概念而设计的计算机称为**冯·诺依曼计算机**,或**IAS计算机**,或“**存储程序**”计算机,其由5个基本部件组成,如图1-4所示。

- ① **主存储器**:用来存放数据和程序。
- ② **算术逻辑单元(ALU)**:用来对**二进制数据**进行操作和处理。
- ③ **控制单元**:用来解释存储器中的程序,并使其运行起来。
- ④ **输入/输出(I/O)设备**:是由控制单元进行操作的设备。

虽然在冯·诺依曼计算机诞生之后,计算机经历了几代的发展,但目前绝大多数计算机的硬件结构仍然具有冯·诺依曼计算机的特征,因此冯·诺依曼计算机被称为现代计算机的“**原型机**”,冯·诺依曼也因此被称为“**计算机之父**”。

冯·诺依曼计算机在结构上由5个基本部件组成,其基本设计思想可以概括为5个方面。

(1) 采用存储程序的工作方式

采用存储程序的工作方式是后来电子数字计算机的显著特点,一方面通过运行存储的不同程序,即可完成不同的计算任务,而不用更改硬件或重新连线;另一方面不同的程序在后来

被进一步用“软件”来表达,软件成了计算机的重要组成部分。

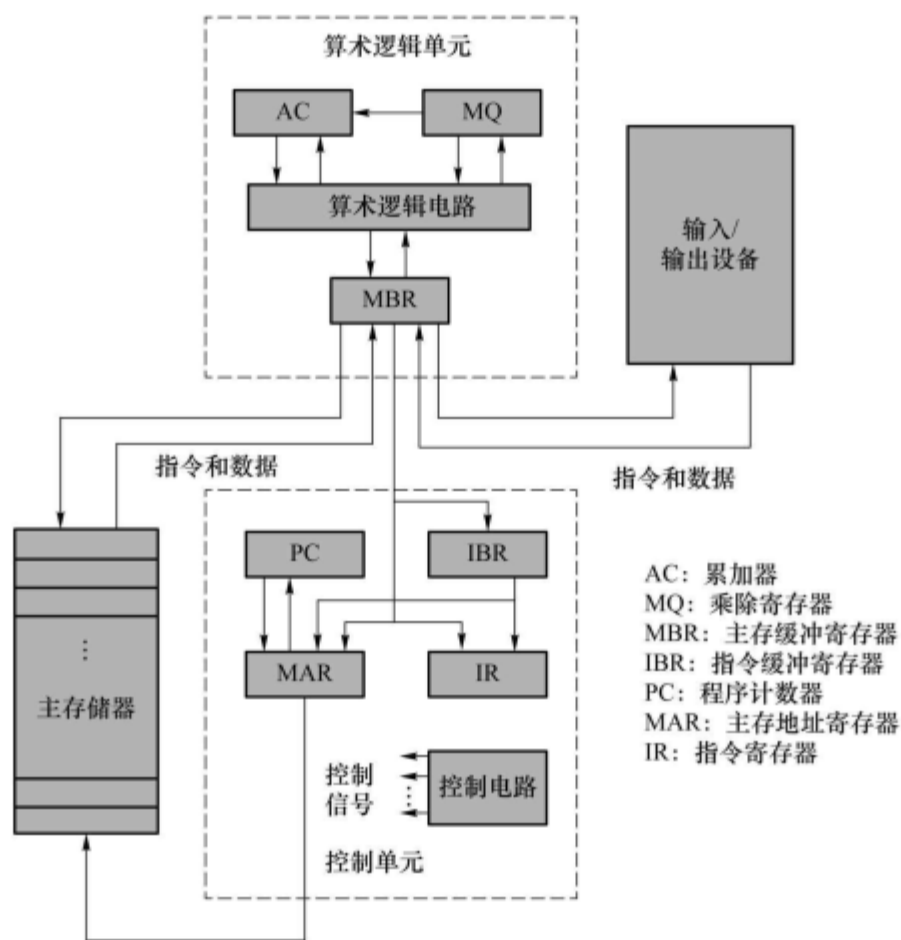


图 1-4 冯·诺依曼计算机的结构示意图

(2) 采用二进制

在计算机中存储的程序、数据都用二进制表示,所有数据的运算都按二进制进行处理,处理结果也表示为二进制数据。

(3) 按地址顺序执行程序

存储器做线性编址,即存储器被划分为多个存储单元,按顺序为各存储单元分配连续的地址码,而程序按地址码顺序被加载到存储器中,然后按地址码顺序从存储器读到运算器和控制器中自动运行。

(4) 程序与数据一起存储

程序以及程序运行所需要的数据一起存放在存储器中,这种存储结构称为冯·诺依曼体系结构。

与冯·诺依曼体系结构计算机的存储器方式略有不同,哈佛体系结构计算机将程序与数据分别放在不同的存储器中,这源于1944年发明的Harvard Mark I计算机(机电计算机)。将程序和数据分开存储的优势:一方面便于同时进行读取指令和读取数据的操作,从而提高数据吞吐率;另一方面独立存储程序更能保证程序运行的可靠性。这种体系结构适合于一些实时性很强的嵌入式计算机。

(5) 以运算器作为机器中心

输入/输出设备与存储器之间的数据传输通过运算器完成。

1.2.2 计算机硬件组成和工作原理

目前绝大多数计算机都遵循冯·诺依曼体系结构,其硬件结构非常复杂,包含数以百万计的电子元器件,对于这种复杂结构系统,人们通常采用层级(或称分层、分级)方法加以描述。

1. 计算机硬件的层级结构

计算机硬件是一个复杂结构系统,可划分为若干个相互关联的子系统,子系统又称“部件”,而每个部件又可以划分成多个更低层级的子系统,依次逐步向下细分,由此计算机硬件被描述为多层次结构。

层级结构划分对于系统的分析和设计都是非常必要的。通常,系统的分析和设计一次仅在一个特定层级上展开,一个特定层级包括一组部件和它们之间的相互关系,当每个层级上部件的行为特征是从其较低层级中简化和抽象出来时,分析者和设计者只需要关心该层级的组成结构和功能。这里涉及组成结构和功能原理两个概念。

组成结构:指某个层级上相互关联的部件及其连接方式。

功能原理:指某个层级上各个部件的操作。

对计算机硬件的总体描述,即从其顶层结构上看,其包括五大组成部件,即控制器、运算器、存储器、输入设备和输出设备,各个部件采用总线方式连接,图 1-5 所示为计算机顶层结构。

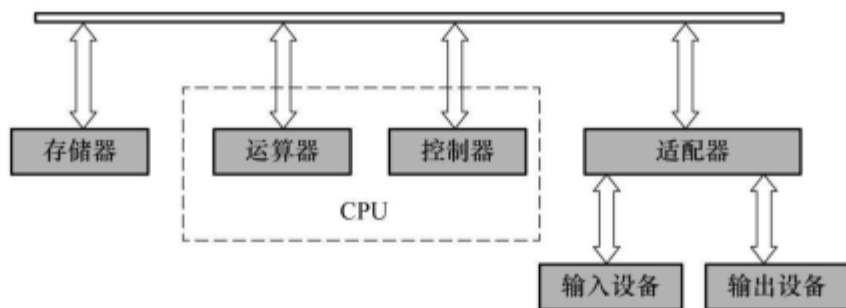


图 1-5 计算机顶层结构示意图

在计算机的部件中,控制器和运算器称为中央处理器(Central Processing Unit,CPU),当代的 CPU 通常被制作在一个集成电路芯片上,称为“微处理器”,例如台式计算机和笔记本电脑使用的 x86 系列的 CPU。当将控制器、运算器、存储器以及输入、输出接口制作在一个芯片上时,该芯片称为“单片机”,又称为“微控制器”,例如 Intel 8051 系列单片机。

从顶层向下逐级分析:顶层结构中的子系统——CPU 部件,其包括控制器、运算器和寄存器,它们之间以内部总线相连接;更底层结构中的子系统——控制器,其包括控制存储器、控制单元寄存器、解码器以及序列逻辑等。本书内容的总体安排就采用这种自顶向下的方法,从整体到细节逐级展开、逐层分析。

图 1-5 所示是计算机的逻辑结构,即部件的逻辑连接结构,为了实现这种逻辑连接,需要利用一些装置将计算机的各个部件分组拼装起来,构成一台完整的物理计算机,称为“计算机的物理结构”。一些逻辑结构图也包含部件之间的位置关系,例如,通常将台式计算机的 CPU 和主存储器放置在一个机箱内部,称为“主机”,而主机外的硬件装置称为“外围设备”,图 1-6 所示为计算机的基本结构。

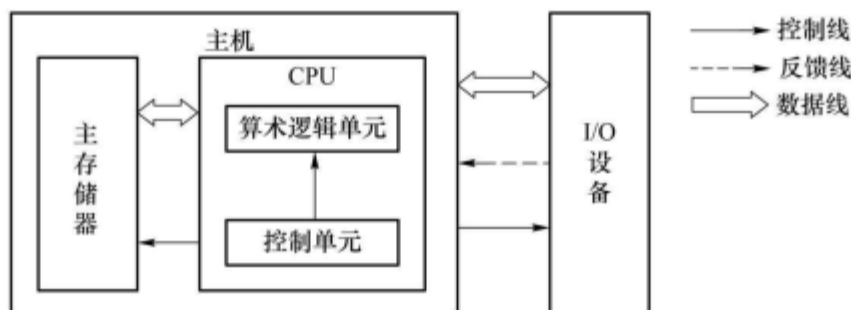


图 1-6 计算机的基本结构图

下面对顶层结构中 5 个部件的主要功能加以简要介绍,并介绍总线,后续章节将会详细阐述。

(1) 存储器

存储器是保存程序和数据的存储介质,也是一个分层级的存储系统,总体上分为主存储器(简称主存,或内存)和辅助存储器(简称辅存,或外存)。其中主存储器是 CPU 能够直接访问的存储器,是计算机实现“存储程序”的核心,而辅助存储器中的信息不能被 CPU 直接访问,必须调入主存后才能被 CPU 访问,它的主要用途是帮助主存记忆更多的信息。

存储器是按二进制码保存信息的,这些二进制信息分为程序和数据,即 CPU 执行的指令和数据。主存储器由很多存储单元组成,通常每个存储单元都是一个字节(8 个二进制位)大小。主存储器做线性编址,按顺序为各存储单元分配连续的地址码,并按地址码对存储器的信息进行访问(包括读取和写入存储器)。

存储单元的总数称为存储器的容量,存储容量越大,计算机能记忆的信息就越多。存储容量通常使用 **KB**(千字节)、**MB**(兆字节)、**GB**(吉字节)、**TB**(太字节)等单位来表示,例如 64 KB、128 MB、256 GB,存储器容量单位之间的换算关系为:1 KB=1 024 B,1 MB=1 024 KB,1 GB=1 024 MB,1 TB=1 024 GB。

图 1-7 所示为一个 64 KB 的存储器,每个格子都表示一个存储单元,格子中的数据就是该存储单元的数据,计算机内部数据是按二进制存储的,为了书写方便,通常写成十六进制形式。图 1-7 中每个存储单元左边的地址码 0000H,0001H,⋯,FFFFH 从上到下按顺序排列,分别表示对应存储单元的地址。

0000H	5AH
0001H	66H
0002H	F0H
⋮	⋮
FFFEH	20H
FFFFH	10H

图 1-7 存储器信息的表示方式

(2) 控制器

现代计算机以存储器为中心,存储器中的指令和数据需要通过控制器的管理和指挥,才能在计算机中实现指令的操作以及数据的流转,并且让程序连续、自动运行。

控制器的基本任务是执行预先存储的指令序列,首先,控制器从存储器中获取一条指令,然后在内部解释该指令,并根据指令意图,协调计算机各执行部件完成指令的操作并实现数据的流转。指令的操作会产生两种信息流,一种是控制流(或称指令流),将操作命令从控制器发送到各个部件;另一种是数据流,在操作命令的控制下,将数据从一个部件发送到另一个部件。

在控制器中如何区分控制流和数据流呢?控制器对指令的操作总体上分为两个周期(两个阶段):取指令周期和执行指令周期。对于存储器中的指令和数据,一般来讲,在取指令周期中从存储器中读出的信息是指令流,它由存储器流向控制器;而在执行指令周期中,从存储器中读出的信息流是数据流,它由存储器流向运算器。因此,控制器可以完全将控制流和数据流

分开处理,如图 1-8 所示。

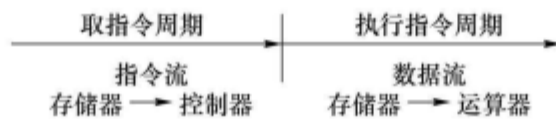


图 1-8 控制器区分指令流和数据流

控制器如何让指令序列从头连续执行呢? 在控制器中设计了一个特殊功能寄存器——**程序计数器**(Program Count, PC),控制器每取出一条指令后,程序计数器会自动加 1,指向下一条指令地址,而且在第一次取指令时,程序计数器已经指向了第一条指令地址,这样就保证了程序从第一条指令开始,按指令顺序连续执行。

(3) 运算器

运算器是一个信息加工部件,完成对数据的算术运算和逻辑运算。运算器通常由算术逻辑单元(Arithmetic Logic Unit, ALU)和一些寄存器组成,如图 1-9 所示,其中 ALU 是完成算术运算和逻辑运算的单元,累加器和寄存器用于存放操作数、中间结果以及最终结果。

一般地,ALU 可以直接完成加法和减法运算,有些 ALU 还可以直接完成乘法、除法运算。由于更为复杂的运算最终都可以分解为上述基本运算,因此理论上 ALU 可以执行任意复杂运算。

ALU 的逻辑运算包括逻辑与(AND)、或(OR)、异或(XOR)、非(NOT)等,它们组合起来可以完成各种复杂的逻辑运算。

有的计算机只包含一个 ALU,有的则包含多个 ALU,例如在超标量计算机中使用多个 ALU,实现多条指令中数据的并行运算。

运算器结构示意图如图 1-9 所示。

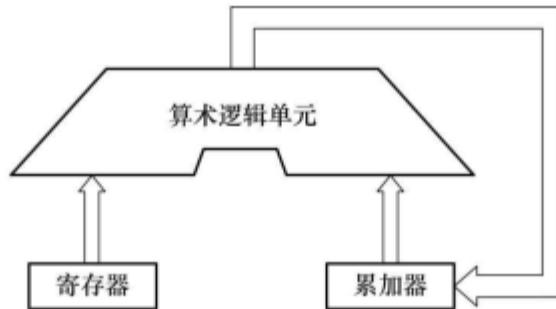


图 1-9 运算器结构示意图

(4) 输入/输出设备

计算机的输入/输出设备是计算机与外界进行信息交换的设备,称为 I/O 设备或者外围设备(peripheral),简称外设。

常见的计算机外设包括键盘、鼠标、写字板等输入设备,以及显示器、打印机等输出设备。有些设备同时具备输入和输出功能,如触摸屏、耳麦(耳机与麦克风的整合体)等。

① 输入设备

输入设备用于将人们熟悉的信息转换为计算机可以接收并识别的信息,例如,当按下键盘上的 A 键时,该键值被翻译成计算机可以识别的 ASCII 码 41H。

② 输出设备

输出设备用于将计算机处理的二进制信息转换为人类或其他设备可以接收或识别的信息,例如显示器,可以将编码的文字或符号的二进制数据,转换为人们可以识别的可读文本。

③ 适配器

外围设备种类非常多,由于它们在速度上有快有慢,在结构上有电子式的,也有机电式的,