

目 录

第 1 章	计算机基础知识	1
1.1	计算机概述	1
1.2	计算机中信息的表示与存储单位	6
1.3	计算思维基础	11
第 2 章	计算机系统组成	22
2.1	计算机系统工作原理	22
2.2	计算机系统组成	23
2.3	计算机硬件系统	24
2.4	计算机软件系统	30
第 3 章	操作系统基础	34
3.1	操作系统概述	34
3.2	常见操作系统介绍	47
第 4 章	多媒体技术	57
4.1	多媒体技术概述	57
4.2	常见媒体类型	61
4.3	多媒体工具	68
4.4	超文本与超媒体	73
第 5 章	数据库基础	78
5.1	数据库	78
5.2	数据模型	82
5.3	数据库模型	84
5.4	关系数据库	86
5.5	数据库技术的新发展	88
第 6 章	计算机网络基础	103
6.1	计算机网络基础知识	103
6.2	INTERNET 的基础	118
6.3	INTERNET 的应用	126

6.4	网页设计与制作基础	138
第7章	信息安全与职业道德	151
7.1	信息技术与信息系统	151
7.2	计算机信息系统的安全	154
7.3	计算机病毒与预防	158
7.4	职业道德及相关法规	165
第8章	软件工程基础	171
8.1	软件工程概述	171
8.2	软件生命周期	175
8.3	软件生命周期模型	183
8.4	系统分析与设计方法	185
第9章	算法与数据结构	188
9.1	算法基础	188
9.2	数据结构基础	192
第10章	计算机新技术发展	195
10.1	物联网基础	195
10.2	云计算技术基础	203
10.3	大数据基础	207
参考文献		214

1.1 计算机概述

电子计算机是一种能自动、高速、正确地完成数值运算和数据处理等功能的电子设备,是现代科学技术的一项伟大的成就,已广泛应用于国防、工业、农业、文教等各个领域,随着信息技术和经济的飞速发展,计算机占据越来越重要的地位,成为人类生活中必不可少的基本工具。了解计算机的发展史,熟悉它的运行机制,是学好计算机的必备基础。

1.1.1 计算机的发展

1946 年 2 月 14 日,世界上第一台计算机“埃尼阿克”(英文缩写 ENIAC,即 Electronic Numerical Integrator and Calculator,中文是电子数字积分计算机)在美国宾夕法尼亚大学诞生。该机采用电子管作为基本元件,使用了 18800 只电子管,占地 170 平方米,重量 30 吨,耗电 140~150 千瓦,每秒可进行 5000 次加减运算。如图 1.1 所示。

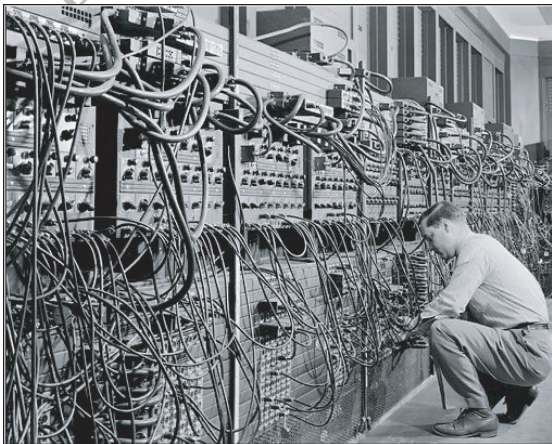


图 1.1 ENIAC 计算机

在 ENIAC 研制过程中,美籍匈牙利数学家冯·诺依曼参与进来,他最新提出“存储程序控制”的思想,这一卓越的思想为电子计算机的逻辑结构设计奠定了基础,已成为计算机设计的基本原则。世界上首次实现存储程序思想的计算机“埃德沙克”(英文缩写 EDSAC)于 1949 年在英国问世,世界上首次设计的存储程序计算机“埃德瓦克”(英文缩写 EDVAC)于 1952 年投入运行,其速度比 ENIAC 提高了 240 倍,这是第一台现代意义的通用计算机。

冯·诺依曼计算机工作原理的主要内容:

(1) 计算机的数制采用二进制

将 ENIAC 采用的十进制改成二进制,因为计算机是由数字电路组成的,数字电路中只有 0 和 1 两种状态。将计算机中的指令和数据均以二进制形式存储,这可以充分发挥电子元件高速运算的优越性。

(2) “存储程序”和“程序控制”结合

程序和数据都放在内存中,这样,不仅可以使计算机的结构大大简化,而且为实现运算控制自动化提供了良好的条件,计算机在程序的控制下自动完成操作。

(3) 采用控制器、运算器、存储器、输入设备和输出设备五个基本部件的结构。

根据冯·诺依曼原理设计制造的计算机被称为冯·诺依曼结构计算机,现代计算机仍然是基于这一原理设计的,也是冯·诺依曼机。由于冯·诺依曼在计算机逻辑结构设计上的伟大贡献,他被誉为“计算机之父”。

ENIAC 的问世标志着计算机时代的到来,之后计算机技术发展异常迅速,根据计算机所采用的主要元器件的不同将计算机的发展分为四个阶段:

第一代(1946 年~1957 年),以电子管为主要部件,内存储器采用汞延迟线、磁芯和磁鼓等,外存储器采用磁带。软件上采用机器语言,后期采用汇编语言,主要用于科学计算。

第二代(1958 年~1964 年),以晶体管为主要部件,内存用磁芯,外存用磁盘。软件上广泛采用高级语言,并出现了早期的操作系统,应用以科学计算和各种事务处理为主,并开始用于工业控制。

第三代(1965 年~1970 年),以中小规模集成电路为主要部件,内存用磁芯、半导体,外存用磁盘。软件上广泛使用操作系统,产生了分时、实时等操作系统和计算机网络。

第四代(1971 年~至今),以大规模、超大规模集成电路为主要部件,以半导体存储器和磁盘为内、外存储器。在软件方法上产生了结构化程序设计和面向对象程序设计的思想。另外,网络操作系统、数据库管理系统得到广泛应用。微处理器和微型计算机也在这一阶段诞生并获得飞速发展。

计算机技术是世界上发展最快的科学技术之一,产品不断升级换代。当前计算机正朝着巨型化、微型化、智能化、网络化、多媒体化等方向发展,计算机本身的性能越来越优越,应用范围也越来越广泛,从而使计算机成为工作、学习和生活中必不可少的工具。

1.1.2 计算机的特点

(1) 运算速度快

现代巨型计算机的运算速度已达百万亿次/秒以上。过去人工需要几年、几十年才能完

成的大量、复杂的科学计算工作,现在使用计算机只需几天或几小时甚至更少的时间就完成了。

(2)运算精度高

电子计算机的计算精度在理论上不受限制,一般的计算机均能达到 15 位有效数字,通过一定的技术手段,可以实现任何精度要求。目前使用计算机计算得到的圆周率的值已达到小数点后的上亿位。

(3)有很强的记忆能力

计算机具有记忆信息的能力,可存储大量的数据和程序,并将处理或计算结果保存起来。这也是计算机区别于其他计算工具的基本特点。

(4)具有逻辑判断

计算机除了能进行数值计算,还可以进行逻辑判断运算。借助于逻辑运算,可以让计算机做出逻辑判断,分析命题是否成立,并可根据命题成立与否做出相应的对策。例如,数学中有个“四色问题”,即不论多么复杂的地图,使相邻区域颜色不同,最多只需四种颜色就够了。100 多年来不少数学家一直想去证明它或者推翻它,却一直没有结果,成了数学中著名的难题。1976 年两位美国数学家终于使用计算机进行了非常复杂的逻辑推理验证了这个著名的猜想。

(5)运行过程自动化

计算机与以前所有计算工具的本质区别在于它能够摆脱人的干预,自动、高速、连续地进行各种操作。计算机从正式操作开始到输出操作结果,整个过程都是在程序控制下自动进行的。

1.1.3 计算机的分类

现在的计算机种类繁多,并表现出各自不同的特点。要对其分类,可以从不同的角度进行,分类依据不同,分类结果也不同。

按计算机的用途不同分为通用计算机和专用计算机。通用计算机广泛适用于一般科学运算、学术研究、工程设计和数据处理等,市场上销售的计算机多属于通用计算机。专用计算机是为适应某种特殊需要而设计的计算机。

按计算机信息的表示形式和对信息的处理方式不同分为数字计算机、模拟计算机和混合计算机。

按计算机的性能不同分为巨型机、大中型机、小型机、微型机、工作站与服务器等。

(1)巨型机

又称超级计算机,是指运算速度超过每秒 1 亿次的高性能计算机,它是目前功能最强、速度最快、软硬件配套齐备、价格最贵的计算机,主要用于解决诸如气象、太空、能源、医药等尖端科学研究和战略武器研制中的复杂计算。它们安装在国家高级研究机关中,可供几百个用户同时使用。运算速度快是巨型机最突出的特点。如美国 Cray 公司研制的 Cray 系列

机中,Cray-Y-MP 运算速度为每秒 20~40 亿次,我国自主生产研制的银河Ⅲ巨型机(如图 1.2 所示)为每秒 100 亿次,IBM 公司的 GF-11 可达每秒 115 亿次,日本富士通研制了每秒可进行 3000 亿次科技运算的计算机。最近我国研制的曙光 4000A 运算速度可达每秒 10 万亿次。



图 1.2 银河Ⅲ巨型机

(2)大中型计算机

这类计算机也有很高的运算速度和存储容量,具有极强的综合处理能力。一般用于大中型企事业单位的中央主机,是事务处理、商业处理、信息管理、大型数据库和数据通信的主要支柱。

这类计算机使用范围相对巨型机较普遍,大中型机通常都像一个家族一样形成系列,如 IBM370 系列、DEC 公司生产的 VAX8000 系列、日本富士通公司的 M-780 系列。

(3)小型机

小型机具有体积小、价格低、性能价格比高等优点,适合中小企业、事业单位用于工业控制、数据采集、分析计算、企业管理以及科学计算等,也可做巨型机或大中型机的辅助机。典型的小型机是美国 DEC 公司的 PDP 系列计算机、IBM 公司的 AS/400 系列计算机,我国的 DJS-130 计算机等。

(4)微型计算机

简称微机,是当今使用最普及、产量最大的一类计算机,体积小、功耗低、成本少、灵活性强,性能价格比明显地优于其他类型计算机,已广泛应用于办公自动化、数据库管理、图像识别、多媒体技术等领域,并且开始成为老百姓家庭中的一种常规电器。

(5)工作站

工作站(workstation)是介于 PC 和小型机之间的高档微型计算机,通常配备有大屏幕显示器和大容量存储器,具有较高的运算速度和较强的网络通信能力,有大型机或小型机的多任务和多用户功能,同时兼有微型计算机操作便利和人机界面友好的特点。工作站的独到之处是具有很强的图形交互能力,因此在工程设计领域得到广泛使用。SUN、HP、SGI 等

公司都是著名的工作站生产厂家。

(6) 服务器

服务器是一种可供网络用户共享的高性能计算机,服务器一般具有大容量的存储设备和丰富的外部接口,运行网络操作系统,要求有较高的运行速度。服务器常用于存放各类资源,为网络用户提供丰富的资源共享服务。常见的资源服务器有 DNS(Domain Name System,域名解析)服务器、E-mail(电子邮件)服务器、Web(网页)服务器、BBS(Bulletin Board System,电子公告板)服务器等。

1.1.4 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如下。

(1) 科学计算(或数值计算)

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。例如,建筑设计中为了确定构件尺寸,通过弹性力学导出一系列复杂方程,长期以来由于计算方法跟不上而一直无法求解。而计算机不但能求解这类方程,并且引起弹性理论上的一次突破,出现了有限单元法。

(2) 数据处理(或信息处理)

信息处理包括对信息的收集、转换、存储、分类、排序、统计、查询、传输等一系列加工处理工作,其结果是获得有用的信息,为管理和决策提供依据。据统计,80%以上的计算机主要用于信息处理,这类工作量大面广,决定了计算机应用的主导方向。目前信息处理已广泛应用于办公自动化、经济领域的各种生产经营管理、人口统计、档案管理、飞机订票等各个方面。

(3) 过程控制

过程控制又称实时控制。过程控制是利用计算机及时采集、检测、分析控制对象的有关数据,按最优值迅速地对控制对象进行自动调节或自动控制。采用计算机进行过程控制,不仅可以大大提高控制的自动化水平,而且可以提高控制的及时性和准确性,从而改善劳动条件、提高产品质量及合格率。因此,计算机过程控制已在机械、冶金、石油、化工、纺织、水电、航天等部门得到广泛的应用。例如,在汽车工业方面,利用计算机控制机床、控制整个装配流水线,不仅可以实现精度要求高、形状复杂的零件加工自动化,而且可以使整个车间或工厂实现自动化。

(4) 计算机辅助工程

利用计算机的高速计算能力、逻辑判断功能、大容量存储和图形处理功能来部分地代替或帮助人完成各种工作,称之为计算机辅助工程。如计算机辅助设计(CAD)、计算机辅助制造(CAM)、计算机辅助测试(CAT)、计算机辅助教学(CAI),计算机辅助出版(CAP)等。

(5) 人工智能

人工智能(Artificial Intelligence)是计算机模拟人类的智能活动,诸如感知、判断、理解、学习、问题求解和图像识别等。现在人工智能的研究已取得不少成果,有些已开始走向实用阶段。例如,能模拟高水平医学专家进行疾病诊疗的专家系统,具有一定思维能力的智能机器人等等。

(6) 计算机网络

计算机网络是计算机技术与现代通信技术结合的产物,因特网的发展是信息技术领域时代的里程碑。网络彻底改变了人们获取信息的方式,必将对人们的生产和生活产生革命性的影响。

(7) 电子商务

电子商务是利用计算机系统和互联网所进行的商业活动。它包括电子邮件、电子数据交换、电子资金转账、快速响应系统、电子表单和信用卡交易等电子商务的一系列应用。

1.2 计算机中信息的表示与存储单位

在计算机中,信息的表示依赖于计算机内的物理器件的状态,信息用什么表示形式直接影响计算机的结构和性能。无论是指令、数据、图形还是声音,在计算机中都以二进制表示。采用二进制的主要原因有以下几点。

(1) 技术实现简单,计算机是由逻辑电路组成,逻辑电路通常只有两个状态,开关的接通与断开,这两种状态正好可以用“1”和“0”表示。

(2) 简化运算规则:两个二进制数和、积运算组合各有三种,运算规则简单,有利于简化计算机内部结构,提高运算速度。

(3) 适合逻辑运算:逻辑代数是逻辑运算的理论依据,二进制只有两个数码,正好与逻辑代数中的“真”和“假”相吻合。

(4) 易于进行转换,二进制与十进制数易于互相转换。

(5) 用二进制表示数据具有抗干扰能力强,可靠性高等优点。因为每位数据只有高低两个状态,当受到一定程度的干扰时,仍能可靠地分辨出它是高还是低。

1.2.1 进位计数制

所谓进位计数制就是使用一组固定的数字和一套统一的规则来表示数的方法。在日常生活中最常用的是十进制数。十进制数是一种进位计数制,它有 0~9 这十个数字,进位的规则是“逢十进一”。我们把 0~9 这十个数字符号称为数码,数码的总个数称为基数,(如十进制的基数是 10,二进制的基数是 2,依次类推,八进制的基数是 8),不同的位置有各自的位权,通常称某个固定位置上的计数单位为位权。例如,在十进制计数中,十位数位置上的位

权为 10^1 , 百位数位置上的位权为 10^2 , 千位数位置上的位权为 10^3 , 而在小数点后第 1 位上的位权为 10^{-1} 。在计算机中, 我们常用到的数制有二进制、八进制、十进制和十六进制, 如表 1.1 所示。

进制数的书写规则有两种: 在进制数后面加英文标识或在括号外面加数字下标。

(1) 在数字后面加英文标识

B(Binary): 表示二进制数。如, 二进制数 500 可写成 500B。

O(Octonary): 表示八进制数。如, 八进制数 500 可写成 500O。

D(Decimal): 表示十进制数。如, 十进制数 500 可写成 500D。一般约定 D 可省去不写, 即无后缀的数字为十进制数。

H(Hexadecimal): 表示十六进制数。如, 十六进制数 500 可写成 500H。

(2) 在括号外面加数字下标

$(1001)_2$: 表示二进制数 1001。

$(3423)_8$: 表示八进制数 3423。

$(5679)_{10}$: 表示十进制数 5679。

$(3FE5)_{16}$: 表示十六进制数 3FE5。

表 1.1 常用进制数

名称	表示符号	基本数字符号	进位规则
十进制	D	0~9	逢十进一
二进制	B	0, 1	逢二进一
八进制	O	0~7	逢八进一
十六进制	H	0~9, A, B, C, D, E, F	逢十六进一

1.2.2 进位计数制之间的转换

(1) 非十进制数转换成十进制数

转换方法: 将要转换的非十进制数的各位数字与它的位权相乘, 其积相加, 和数就是十进制数。例:

$$\begin{aligned}
 (101101.11)_2 &= 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 1 \times 2^{-2} \\
 &= 32 + 0 + 8 + 4 + 0 + 1 + 0.5 + 0.25 \\
 &= (45.75)_{10}
 \end{aligned}$$

$$(123.4)_8 = 1 \times 8^2 + 2 \times 8^1 + 3 \times 8^0 + 4 \times 8^{-1} = 64 + 16 + 3 + 0.5 = (83.5)_{10}$$

$$(5F.A)_{16} = 5 \times 16^1 + 15 \times 16^0 + 10 \times 16^{-1} = 80 + 15 + 0.0625 = (95.0625)_{10}$$

(2) 十进制数转换成非十进制数

转换方法: 将十进制数转换为其他进制数时, 可将此数分成整数与小数两部分分别转换, 然后再拼接起来即可。

整数部分转换:将十进制整数连续除以非十进制数的基数,并将所得余数保留下来,直到商为0,然后用“倒数”的方式(第一次相除所得余数为最低位,最后一次相除所得余数为最高位),将各次相除所得余数组合起来即为所要求的结果。此法称为“除以基数倒取余法”。

小数部分转换:将十进制小数连续乘以非十进制数的基数,并将每次相乘后所得的整数部分保留下来,直到小数部分为0或已满足精确度要求为止,然后将每次相乘所得的整数部分按先后顺序(第一次相乘所得整数部分为最高值,最后一次相乘所得的整数部分为最低值)组合起来。

说明:

1)十进制纯小数转换时,若遇到转换过程无穷尽时,应根据精度的要求确定保留几位小数,以得到一个近似值。

2)十进制与八进制、十六进制的转换方法和十进制与二进制之间的转换方法相同,这里不再举例。

(3)二、八、十六进制数的相互转换

1)二进制数与八进制数之间的转换,由于一位八进制数对应三位二进制数,因此转换方法如下:

二进制数转换为八进制数:将二进制数以小数点为界,分别向左、向右每三位分为一组,不足三位时用0补足(整数在高位补0,小数在低位补0),然后将每组三位二进制数转换成对应的八进制数。

例:将 $(1011010.1)_2$ 转换成八进制数

001 011 010. 100

1 3 2 4

$(1011010.1)_2 = (132.4)_8$

八进制数转换为二进制数:按原数位的顺序,将每位八进制数等值转换成三位二进制数。

例:将八进制数 $(756.3)_8$ 转换成二进制数

7 5 6. 3

111 101 110 011

$(756.3)_8 = (111101110.011)_2$

2)二进制数与十六进制数之间的转换:由于一位十六进制数对应四位二进制数,因而转换方法如下:

二进制数转换为十六进制数:将二进制数以小数点为界,分别向左、向右每四位分为一组,不足四位时用0补足(整数在高位补0,小数在低位补0),然后将每组的四位二进制数等值转换成对应的十六进制数。

例:将二进制数 $(1100111001.001011)_2$ 转换成十六进制数。

0011 0011 1001. 0010 1100

3 3 9 2 C

$(1100111001.001011)_2 = (339.2C)_{16}$

十六进制数转换为二进制数:按原数位的顺序,将每位十六进制数等值转换成四位二进

制数。

例:将 $(AB3.57)_{16}$ 转换成二进制数

A B 3. 5 7
 $1010\ 1011\ 0011\ 0101\ 0111\ (AB3.57)_{16}=(101010110011.01010111)_2$
八进制转换为二进制数的方法与十六进制转换为二进制的方法相同。

1.2.3 信息的存储单位

b(位,bit):是二进制数的最小单位,即一位二进制数,有 0 和 1 两个值,在计算机网络通信中,常用 bps(位/秒)来衡量数据传输速率的快慢。

B(字节,Byte):由八位二进制数构成,即:1 Byte=8 bit。字节是衡量存储器容量的基本单位,常用的单位还有:KB、MB、GB、TB 等,换算关系如下:

$1KB=2^{10}B$
 $1MB=2^{10}KB$
 $1GB=2^{10}MB$
 $1TB=2^{10}GB$

1.2.4 非数值信息的表示

一般将数据分为数值型数据和非数值型数据。数值型数据用于衡量量的大小;非数值型数据用于表示各类信息,如文字、声音、图形、图像等。关于非数值型数据的编码,各个国家的文字表示有相应国家标准,声音、图形、图像有相应的行业标准和国际标准。下面介绍几种常见的文字编码标准。

(1)字符型信息的编码(ASCII 码)

ASCII(American Standard Code for Information Interchange)码,是“美国信息交换标准代码”的简称,是英文文字系统的编码标准。编码包括 0~9 十个数码符号、52 个大小写英文字母,32 个标点符号和运算符,34 个控制符,共计 128 个字符。这需要用 7 位二进制编码表示,所以采用一个字节,最高位置 0,传输时最高位可以作为奇偶校验。7 位 ASCII 码字符编码表如表 1.2 所示。

表 1.2 ASCII 码表

高 3 位 低 4 位	000	001	010	011	100	101	110	111
0000	NUL	DLE	SP	0	@	P	,	p
0001	SOH	DC1	!	1	A	Q	a	q
0010	STX	DC2	“	2	B	R	b	r
0011	ETX	DC3	#	3	C	S	c	s
0100	EOT	DC4	\$	4	D	T	d	t

续表

高 3 位 低 4 位	000	001	010	011	100	101	110	111
0101	ENQ	NAK	%	5	E	U	e	u
0110	ACK	SYN	&	6	F	V	f	v
0111	BEL	ETB	•	7	G	W	g	w
1000	BS	CAN	(	8	H	X	h	x
1001	HT	EM	)	9	I	Y	i	y
1010	LF	SUB	*	:	J	Z	j	z
1011	VT	ESC	+	;	K	[	k	{
1100	FF	FS	,	<	L	\	l	
1101	CR	GS	—	=	M	]	m	}
1110	SO	RS	.	>	N	^	n	~
1111	SI	US	/	?	O	_	o	DEL

例如:字符 A 的 ASCII 码是 1000001。若用十进制数可表示为 65D,用十六进制数可表示为 41H。

(2)汉字编码

汉字是象形文字,由于汉字自身的特点,汉字没法像英文一样通过简单元素(如字母)来表示。因此,汉字的编码采用一字一码的方式。汉字编码的输入、处理和显示方式都和英文不同,包含了用于输入的输入码,用于交换的国标码,用于内部处理的内码和用于打印显示的字形码。

1)汉字输入码。汉字输入码是为了将汉字通过键盘输入计算机而设计的代码,是代表某个汉字的一组键盘符号,也称外码。外码多数为 4 个字节。每一个汉字对应一个外码,不同的输入方法,其汉字的外码是不同的。输入法必须将键盘所输入的字符序列转换成机器内部表示的内码存储和处理。输入码和机内码之间的转换通过键盘管理程序实现。

2)汉字交换码。汉字交换码是指在对汉字进行传递和交换时使用的编码,也称国标码。1981 年,国家标准局颁布了《信息交换用汉字编码字符集(基本集)》,简称 GB2312—80,代号国标码,是在汉字信息处理过程中使用的代码的依据。GB2312—80 共收集汉字、字母、图形等字符 7445 个,其中汉字 6763 个(常用的一级汉字 3755 个,按汉语拼音字母顺序排列;二级汉字 3008 个,按部首顺序排列),此外,还包括一般符号、数字、拉丁字母、希腊字母、汉语拼音字母等。在该标准集中,每个汉字或图形符号均采用双字节表示,每个字节只用低 7 位;将汉字或图形符号分为 94 个区,每个区分为 94 个位,高字节表示区号,低字节表示位号。国标码一般用十六进制表示,在一个汉字的区号和位号上分别加十六进制 20H,即构成该汉字的国标码。例如,汉字“啊”位于 16 区 01 位,其区位码为十进制数 1601D(即十六进制数 1001H),对应的国标码为十六进制数 3021H。

3)汉字机内码。汉字机内码是供计算机内部进行存储、处理、传输汉字用的代码,也称

内码。在计算机内,一个汉字或图形符号的内码一般用2个字节表示,每个字节一般只用7位,最高位均置1。无论汉字的输入方式如何,但是对于同一个汉字来说,它的内码是相同的。

汉字国标码作为一种国家标准,是所有汉字都必须遵循的统一标准,但由于国标码每个字节的最高位都是“0”,与国际通用的ASCⅡ码无法区别,必须经过某种变换才能在计算机中使用,英文字符的机内代码是7位的ASCⅡ码,最高位为“0”,而将汉字机内代码两个字节的最高位设置为“1”,这就形成汉字的内码。

4) 汉字字形码。汉字字形码是表示汉字字形信息的编码,也称汉字输出码。目前在汉字信息处理系统中大多以点阵方式形成汉字,所以汉字字形码就是确定一个汉字字形点阵的代码,点阵字形中的每一点用一个二进制位来表示,随着字形点阵的不同,它们所需要的二进制位数也不同,例 24×24 的字形点阵,每字需要72字节; 32×32 的字形点阵,每字共需128字节,与每个汉字对应的这一串字节,就是汉字的字形码。

1.3 计算思维基础

1.3.1 计算思维基本特征

长期以来,计算机被社会看作一种高科技工具,计算机学科也被视为一门专业性很强的工具学科。专业性意味着认知门槛高,工具意味着它是一种辅助性学科,这种狭隘的认知是计算机向各行各业渗透的最大障碍。计算机大师迪克斯彻(Edsger Wybe Dijkstra, 1930—2002, 荷兰)说过:“我们使用的工具影响着我们的思维方式和思维习惯,从而也将深刻影响着我们的思维方式”。

“计算思维”是卡内基梅隆大学(CMU)计算机科学系主任、美国国家自然科学基金会会计学与信息科学工程部助理部长周以真(Jeannette M. Wing)教授提出的一种理论。周以真教授提出了以下计算思维的观点和方法。

1. 计算思维的广泛性

计算思维是每个大学生必须掌握的基本技能,它不仅属于计算机科学家。教育应使每个孩子培养解析能力时,不仅掌握阅读、写作和算术(3R),还要具有计算思维。

应当为大学新生开设一门“怎么像计算机科学家一样思维”的课程,面向所有专业学生:应当传播计算机科学的快乐、崇高和力量,致力于使计算思维成为常识。

一个人可以主修计算机科学而从事任何行业。例如,一个人可以主修英语或者数学,而从事各行各业的职业。一个人可以主修计算机科学,而从事医学、法律、商业、政治以及任何类型的科学和工程,甚至艺术工作。

2. 计算思维的基本原则

计算思维是一种递归思维,它是并行处理的思维方式。

计算思维是概念化,不是程序化。计算思维不光是编程,还要求在抽象层次上进行思维。

计算思维是人的,不是计算机的思维方式。计算思维是人类求解问题的思维方法,而不是让人类像计算机那样思考。计算机枯燥而且沉闷,人类聪明且富有想象力,是人类赋予了计算机激情。

计算思维是数学思维和工程思维的相互融合。计算机科学本质上来源于数学思维,但是受到计算设备的限制,迫使计算机科学家必须进行工程思考,不能只是数学思考。图灵和冯·诺依曼是“数学思维”和“工程思维”的典型代表人物。

3. 利用计算思维解决问题的方法

计算思维建立在计算过程的能力和限制之上。需要考虑哪些事情人类比计算机做得好?哪些事情计算机比人类做得好?最根本的问题是:什么是可计算的?

当人们求解一个待定的问题时,首先会问:解决这个问题有多困难?有什么最佳的解决方法?表述问题的难度取决于人们对计算机理解的程度。

为了有效求解一个问题,可能要进一步提问:一个近似解就足够了吗(如 Excel 中的计算精度是否需要 16 位以上)?是否允许漏报和误报(如视频播放时的数据丢失)?计算思维就是通过简化、转换和仿真等方法,把一个看起来困难的问题,重新阐释成一个人们知道怎样解决的问题。

计算思维采用抽象和分解的方法,将一个庞杂的任务设计分解成一个适合于计算机处理的系统。计算思维是选择合适的方式对问题进行建模,使它易于处理。在人们不必理解每一个细节的情况下,就能够安全的使用或调整一个大型的复杂系统。

上述周以真教授关于计算思维的主要观点,是指导人们利用计算机解决问题的思维方式。

4. 解决复杂问题的方法

大问题的复杂性随着问题规模的增长,问题的复杂性呈非线性增加的反应。在计算机领域,问题的大小与计算的复杂性直接相关。

例如,对矩阵乘法这类问题,2 阶矩阵相乘是个小问题,而 1000 阶矩阵相乘是个大问题。

例如,在 CPU 等大型集成电路芯片中,继承了数亿个晶体管,因为在设计中变得非常复杂,稍有不慎就会造成产品瑕疵。

例如,在软件工程中,随着问题规模的增大,系统的复杂性也在增大。每个新的信息项、功能或限制都有可能影响整个系统的其他元素。因此,随着问题复杂性的增加,系统分析的任务将呈现几何级数增长。在研究一个大系统时,控制和降低系统的复杂性,是区分和选择各种计算机级数的重要因素。

为了解决大问题的复杂性,哲学家和数学家笛卡儿在《方法论》一书中提出一些基本方

法,笛卡儿的主要思想如下:

第一,绝不把任何我没有明确地认识为真的东西,当作真的加以接受。也就是说,小心避免仓促的判断和偏见,只把那些十分清楚地呈现在我的心智之前,是我根本无法怀疑的东西放进我的判断之中。

第二,把我考察的每一个难题,都尽可能地分成细小的部分,直到可以而且适于加以圆满解决的程度为止。

第三,按照次序引导我的思想,以便从最简单、最容易认识的对象开始,一点一点上升到对复杂对象的认识,即便是那些彼此之间并没有自然先后次序的对象,我也给它们设定一个次序。

第四,把一切情形尽量完全地列举出来,尽量普遍地加以审视,使我确信毫无遗漏。

1.3.2 计算机如何解决问题

人们使用计算机解决问题时,必须人为地规定计算机的操作步骤,告诉计算机“干什么”或“怎么干”,也就是按照任务的要求写出一系列计算机指令,这些指令的集合体成为程序。当然,指令必须是计算机能够识别和执行的指令。计算机的指令是有限的,用它们可以编制出解决各种不同任务的程序。计算机能够自动连续地执行程序中指令规定的操作,是因为计算机是按照“存储程序”原理进行工作的。其实,九三级硬件设备只能进行基本的数学运算与逻辑比较,计算机的各种功能是在计算机程序指挥下完成的。简单地说,有多少种程序,计算机就有多少种功能。

计算机解决一个问题时,大致要经历以下步骤:

1. 问题的可计算性

计算机的优势在于数值计算,很多非数值问题(如文字识别,图像处理等)都可以转化成为数值问题交给计算机处理。但是,使用计算机解决的问题应该在有限步骤内被解决,例如“哥德巴赫猜想”就属于不可计算的问题,因为计算机没有办法给出数学意义上的证明。因此,没有任何理由期待计算机解决世界上所有的问题。

2. 对象的离散化

计算机处理的对象一部分本身就是离散化,例如数字、字母、符号等;但是还有很多处理对象是连续型的,例如色彩、声音、交流电压等。凡是“可计算”的问题,处理对象都是离散型的,因为计算机建立在离散型数字计算的基础上。所有连续型问题必须转化为离散型问题后(数字化),才能被计算机处理。例如,在计算机屏幕上显示一张图片时,计算机必须将图片在水平和垂直方向分解成一定分辨率的像素点(离散化),然后将每个像素点再分解成红绿蓝(RGB)3种基本颜色;每种颜色的变化分解为0~250(1字节)个色彩等级。这样计算机就会得到一大批有特定规律的离散型数字,就能任意处理这张图片了,例如图片的放大、缩小、旋转、变形、变换颜色等操作。

3. 解决问题的算法

通俗地讲,算法就是完成任务的步骤。例如,Windows 系统如何在硬盘中找到用户指定的文件;媒体播放器如何将 MP3 文件转换成动听的音乐;搜索引擎如何在因特网中找到用户需要的网页等,这些问题都需要利用算法来解决。解决一个具体问题时,必须给出所求问题的已知条件,列举所要求的结果,分析已知条件和要求结果之间的联系,然后建立相对应的求解方法,即确定解决问题的算法步骤。然后将算法步骤表达为数学建模、系统结构图、程序流程图等形式。

4. 程序设计

将解决问题的算法步骤转变为程序设计,然后编辑、调试和测试程序代码,直到输出所要求的结果。

图灵在《计算机与智能》论文中指出:“如果一个人想让机器模仿计算机员执行复杂的操作,他必须告诉计算机要做什么,并把结果翻译成某种形式的指令表。这种构造指令表的行为通常称为编程”。

1.3.3 计算机不能解决的问题

大多数机器只能做某一类事情,例如,电视机只能看电视冰箱只能储存食物等。为什么计算机能做多种不同的事呢?是不是计算机什么事都能做?其实,计算机的根本问题是“可计算”问题,可计算性是指一个实际问题是否能够使用计算机来解决。在实际工作中,不可计算的问题有二义性、复杂性、可行性、组合爆炸等诸多问题。

1. 二义性问题

利用计算机解决问题时,必须有充分的已知条件,而且对结果有明确的要求。简单来说,就是问题没有二义性。例如,“为我做一个西红柿炒鸡蛋”这样的问题,计算机是无法解决的。因为这个问题存在太多的二义性,如:西红柿要多大?成熟到什么程度?是否要全红色的?用什么品种?“一个”的含义是多少(1 个西红柿还是 1 斤西红柿)?“炒”的含义又是什么,需要火炒吗?用什么火(柴火、煤气、液化气火、电磁炉等)?炒多长时间?等等。

二义性是一种语法的不完善说明,在计算机应用中应当避免这种情况出现。有两个解决二义性的基本方法。一种方法是设置一些规则,该规则可出现二义性的情况下指出哪一个是正确的;另一种方法是将问题改变成一个强制正确的格式,这样就可以进行计算机编程处理了。

2. 复杂性问题

克拉默(Friedrich Crarner,1923—2003,德国)在经典著作《混沌与秩序——生物系统的复杂结构》中,给出了几个简单例子,用于分析程序的复杂性。

【例 1】 序列 aaaaaa ···

这是一个简单系统,相当的程序为:在每一个 a 后续写 a。这个短程序使这个序列得以随意复制,不管多长都可以办到。

【例 2】 序列 aabaabaabaab ···

与例 1 相比,该例要复杂一些,是一个准复杂系统,但仍可以很容易地写出程序:在两个 a 后续写 b 并重复这一操作。

【例 3】 aabaababbaabaababb ···

与例 2 相似,也可以用很短的程序来描述:在两个 a 后续写 b 并重复。每当三次重写 b 时,将第二个 a 替换为 b。这样的序列具有可定义的结构,有对应的程序来表示。

【例 4】 aababbababbbabaaababbab ···

这个例子中信息的排列毫无规律,如果希望编程解决,就必须将字符串全部列出。这就可以得出一个结论:一旦程序的大小与试图描述的形同大小相提并论时,便无法编程。或者说,当系统的结构你能被描述,或描述它的最小算法与系统本身具有相同的信息比较时,则称该系统为根本复杂系统。在达到根本复杂之前,人们仍可以编写出能够执行的程序,负责计算机不能进行处理。

3. 梵天塔问题

相传印度教的天神梵天(Hanoi)在创作地球时,建了一座神庙。神庙里竖有 3 根宝石柱子,柱子由一个铜座支撑。梵天将 64 个直径大小不一的金盘子,按照从大到小的顺序依次套在第一根柱子上,形成一座金塔,即所谓的梵天塔。天神让庙里的僧侣们将第一根柱子上的 64 个盘子借助第二根柱子全部移到第三根柱子上,即将整个塔迁移,同时定下了 3 条规则。一是每次只能移动一个盘子;二是盘子只能在 3 根柱子上来回移动,不能放在他处;三是在移动过程中,3 根柱子上的盘子必须始终保持大盘在下,小盘在上。

3 个盘子的搬移次数为: $2^3 - 1 = 7$ 。梵天塔有 64 个盘子,移动盘子的次数为: $2^{64} - 1 = 18\ 446\ 744\ 073\ 709\ 551\ 615$ 。假设僧侣们移动依次盘子需要 1 秒,一年有 31 536 000s,则僧侣们一刻不停地来回移动,也需要大约 5 849 亿年的时间。

假设将梵天塔的 3 个柱子命名为 A、B、C,盘子按 1~64 进行编号,然后利用计算机将盘子的搬动顺序全部列举出来。如果计算机以每秒 1 000 万个盘子的速度进行计算,则需要花费大约 58 490 年的时间才能完成计算,这还不包括将结果打印出来的时间。从梵天塔问题可以看出,理论上可以计算的问题,在实际中并不一定能行。

4. 组合爆炸问题

根据笛卡儿积规则,当 2 个元素和 2 个状态具有相互作用时,会有 4 种不同的组合。例如,二进制有 2 个元素(0 和 1),如果采用 2 位二进制,则有 $2^2 = 4$ 中不同的组合。例如,分析计算机日常使用的密码问题,如果规定密码只能采用 0~9 这十个数字,而且只用 6 位密码,则有: $10^6 = 1\ 000\ 000$ 种密码组合;如果密码允许采用 0~9 这十个数字和 26 个英文字母的任意组合,还是采用 6 位密码,则有: $(10 + 26)^6 = 2\ 176\ 782\ 336$ 种密码组合;如何在以上组合中不断增加密码的位数,并将所有可能的密码组合计算出来,计算机必将遇到“组合爆炸”问题,对于“组合爆炸”问题,利用计算机来破解密码是无法处理的。

计算的复杂性包括空间和时间两方面的复杂性。组合爆炸问题体现了时间的复杂性。

从以上分析可以知道,并不是所有问题都是可计算的,即使可计算的问题,也要考虑在实际中是否能行。

1.3.4 人工智能应用

1. 图灵测试

1947年,图灵在一次计算机会议上做了题为“智能机器”的报告(1956年报告以《机器能够思维吗?》为题重新发表),详细地阐述了他关于思维机器的思想,第一次从科学的角度指出:“与人脑的活动方式极为相似的机器是可以制造出来的”。1950年,图灵发表了著名的论文《计算机器与智能》,他逐条反驳了各种机器不能思维的论点,并做出了肯定的回答。他对人工智能问题从行为主义的角度给出了定义,在论文中提出了著名的“图灵测试”。

图灵测试由3人来完成:一个男人(A);一个女人(B);一个性别不限的提问者(C)。提问者呆在与其他两个测试者相隔离的房间里。测试的目标是:让提问者通过对其他两人的提问,来鉴别回答问题的对象是男性还是女性。为了避免提问者通过声音、语调轻易地做出判断,规定提问者和测试者之间只能通过电传打字机进行沟通。

提问者在不知道A、B性别的情况下,可以提问:“请A回答,你的头发长度?”A为了给提问者造成错觉,他可以回答:“我的头发很长,大约有20cm”;如果提问者对B提问:“A说的是真的吗?”那么B可能会回答:“不要相信那个人”之类的回答。反之,A也同样可以做出类似的回答。

如果将上面测试中的男人(A)换成机器,提问者在机器和女人的回答中做出判断。如果机器足够“聪明”,就能够给出类似于人类思考后得出的答案,而且在5min的交谈时间内,如果没被人类裁判识破,那么这台机器就算通过了图灵测试,便可以判定这台机器能够思维。

2. 图灵测试和人工智能

图灵在《计算机器与智能》论文中,逐一反驳了科学、哲学、社会、工程等各个领域对“机器思维”的反对观点。论文发表后,引起了轩然大波,图灵也遭到了来自不同领域的猛烈攻击。

图灵测试不要求接受测试的机器在内部构造上与人脑一样,它只能从功能的角度来判定机器是否能够思维,也就是从行为主义的角度对“机器思维”进行定义。尽管图灵对“机器思维”的定义不够严谨,但他关于“机器思维”定义的开创性工作对后人的研究具有重要意义。因此,专家学者认为,图灵的论文标志着计算机人工智能问题研究的开始,而计算机的终极目标也是达到机器的人工智能。

按照图灵的语言,到2000年左右将会出现足够好的计算机。在长达5min的交谈中,人类裁判在图灵测试中的准确率会下降到70%或者更低(或机器欺骗的成功率达到30%)。到目前为止,虽然没有一台计算机能够完全通过图灵测试,但是在2012年的国际比赛中,计算机的测试中“骗”过人类裁判的成功率已达到了29.2%,已非常接近图灵的预测了。

在某些特定的领域,如博弈领域,图灵测试已取得了成功。1997年,IBM公司的“深蓝”计算机战胜了国际象棋冠军卡斯帕罗夫。“深蓝”有30个处理器,每个处理器的工作频率为120MHz,可以进行并行计算。它还有480个专门用于下棋的集成电路芯片。下棋程序用C语言编写,每秒可以评估2亿个可能的布局。“深蓝”计算机可计算出未来6~8步棋局的形式和最佳走法,最多可以计算出未来20步棋局的形势。

“深蓝”战胜卡斯帕罗夫后,让计算机理解人类的自然语言成了计算机科学家追求的热点。2011年,IBM超级计算机“沃森”(Watson)在美国最受欢迎的智力竞猜电视节目《危险边缘》中,击败了该节目历史上两位最成功的选手。“沃森”计算机存储了海量数据,而且拥有一套逻辑推理程序,可以推理出它认为最正确的答案。“沃森”极富磁性的语言也倾倒了广大观众,机器越来越像人类了。

3. 人工智能的定义

人工智能(Artificial Intelligence, AI)的定义可以分为“人工”和“智能”两个部分。“人工”比较好理解,争议也不大;关于什么是“智能”,问题就多了。“智能”涉及意识、自我、心灵、无意识的精神等问题。目前人们对自身的理解是非常有限的,对构成人类智能的必要元素的了解也是有限的,所以很难定义什么是“人工”制造的“智能”。

人工智能比较流行的定义是美国麻省理工学院麦卡锡(John McCarthy, 1927—2011)教授在1956年提出的:人工智能就是要让机器的行为看起来就像是人所表现出来的智能行为一样。

人工智能的另一个定义是:人造机器所表现出来的智能,通常是指通过计算机实现的智能。

4. 人工智能的研究内容

(1) 搜索与求解

搜索是为了达到某一目标而进行某种操作、运算、推理或计算的过程。搜索是人们求解问题时,在不知道现成方法的情况下所采用的一种方法。搜索可以看作人类和其他生物所具有的一种元知识。人工智能的研究表明,许多问题(包括智力问题和实际工程问题)的求解,都可以描述为对某种图或空间的搜索问题。进一步研究发现,许多智能活动(包括脑智能和群智能)的过程都可以看作或者抽象为一个基于搜索的问题求解过程。因此搜索技术是人工智能最基本的研究内容。

(2) 学习与发现

学习能力是智能行为的一个非常重要的特征。但至今人们对学习的机理尚不清楚。机器学习是研究计算机怎样模拟或实现人类的学习行为,以获取新的知识或技能,并重新组织已有的知识结构之不断改善自身的性能。

学习是一项复杂的智能活动,学习过程与推理过程是紧密相连的,按照学习中使用推理的多少,机器学习所采用的方法大体上可分为4种:机械学习、通过传授学习、类似学习和通过案例学习。学习中所用的推理越多,系统的能力越强。

(3) 知识与推理

培根(Francis Bacon, 1561—1626, 英国)说过“知识就是力量”。在人工智能中,人们进一步领略到这句话的深刻内涵。对智能来说,知识太重要了。发现客观规律是一种智能的表现,而运用知识解决问题也是智能的表现,并且发现规律和运用知识本身还需要知识。因此可以说,知识是智能的基础和源泉。计算机要实现人工智能就必须拥有学习知识和运用知识的能力。为此,就要研究面向机器的知识表示形式和基于各种表示的机器推理技术。

知识的表示要便于计算机的接受、存储、处理和运用。机器的推理方式与知识的表示又息息相关。尤其是一些“默认知识”(一些经常使用却很难通过语言和文字予以表达的知识)很难通过计算机进行处理。例如,“跌倒的人是如何迅速站起来的”“人如何下楼梯”等默认知识,对机器人的跌倒站立研究就有很大的应用价值。

(4) 发明与创造

广泛的发明创造内涵很广泛,如机器、材料、工艺等方面的发明和革新;也包括创新性软件、规划、设计等技术和方法的创新;以及文学、艺术方面的创作;还包括思想、理论、法规的建立和创新等。

发明创造不仅需要知识和推理,还需要想象和灵感。它不仅需要逻辑思维,而且还需要形象思维。所以这个领域是人工智能中最富挑战性的一个研究领域。目前,人们在这一领域已经展开了一些工作,并取得了一些成果,例如,已开发出了计算机辅助创新软件,还尝试用计算机进行文艺创作等。但总体来讲,原创性的机器发明创造进展甚微,甚至还是一片空白。

(5) 感知与交流

感知与交流是指计算机对外部信息的直接感知和人机之间、智能之间的直接信息交流。

机器感知就是计算机像人一样通过“感觉感官”从外界获取信息。例如,机器通过摄像头获取图像信息,通过微型话筒获取声音信息,通过其他传感器获取重量、阻力、温度、光线等外部环境的数据。然后,对这些数据进行处理,感知外部环境的基本情况,做出行动决策(如火星探测车)。

机器交流涉及通信和自然语言处理等技术。自然语言处理又包括自然语言理解和表达。例如,情感和社会技能对一个智能机器人是很重要的。智能机器人通过了解人们的动机和情感状态,能预测别人的行为。这涉及博弈论、决策理论等方面的研究,以及机器人对人类情感感知能力的检测。为了良好的燃机互动,智能机器人也需要表现出情绪来,至少它必须礼貌地和人类打交道。

(6) 记忆与联想

记忆是智能的基本条件,不管是脑智能还是群智能,都以记忆为基础。记忆也是人脑的基本功能之一。在人脑中,伴随着记忆的就是联想,联想是人脑的奥秘之一。

计算机要模拟人脑的思维就必须具有联想功能。要实现联想就需要建立事物之间的联系,在机器世界里就是有关数据、信息或知识之间的联系。建立这种联系的方法很多,如程

序中的指针、函数、链表等。但传统方法实现的联想,只能对那些完整的、确定的(输入)信息,联想起(输出)有关信息。这种“联想”与人脑的联想功能相差甚远,人脑对那些残缺的、失真的、变形的输入信息,仍然可以快速准确地输出联想响应。

例如,人类在回答:“树上有10只鸟,被人用枪打下1只,问树上还剩几只鸟?”这类问题是,就会利用到联想和记忆等方面的能力。而计算机在处理这类问题时,答案往往令人啼笑皆非。

(7) 人工智能的研究方法

人工智能是一门边缘学科,属于自然科学和社会科学的交叉。目前没有统一的原理和范式指导人工智能的研究。在许多问题上研究者都存在争论,几个长久以来没有结论的问题是:是否应从心里或神经方面模拟人工智能?人类生物学与人工智能研究有没有关系?智能行为能否用简单的原则来描述?智能是否可以使用高级符号表达?20世纪50年代计算机研制成功后,研究者开始探索人类智能是否能简化成符号处理。20世纪80年代后,很多人认为符号系统不可能模仿人类所有的认知过程,特别是机器感知、机器学习和模式识别。

5. 人工智能的应用领域

人工智能的应用范围非常广泛,如:难题求解、机器翻译、模式识别、文艺创作、机器博弈、智能机器人、任务规划、资源配置、机器定理证明、自然程序设计、智能控制、智能策略等领域。

(1) 难题求解

这里的难题,主要指那些没有算法解,或虽有算法解,但现有计算机无法完成的困难问题。例如,智力性问题中的梵天塔问题、 n 皇后问题、旅行商问题、博弈问题等。犹如,现实世界中的集体规划、市场预测、股份分析、地震预测、疾病诊断、军事指挥等难题。在这些难题中,有些是数学理论中所称的非确定型多项式(NP)问题或NP问题。NP问题是指那些既不能说明其算法复杂性超出多项式界限,但又没有找到有效算法的一类问题。

研究智力难题的求解具有双重意义:一方面可以找到解决这些难题的途径;另一方面由解决这些难题而发展起来的一些技术和方法,可用于人工智能的领域。

(2) 机器翻译

机器翻译是完全用计算机作为两种语言之间的翻译。机器翻译由来已久,早在计算机问世不久,就有人提出了机器翻译的设想,随后就开始了这方面的研究。当时人们认为只要用一部双向词典及一些语法知识就可以实现两种语言文字间的机器互译,结果遇到了挫折。例如,当把英语句子“The files like an arrow”(光阴时间)翻译成日语,然后再翻译回来的时候,竟变成了“苍蝇喜欢箭”;又如,把英语句子“The spirit is silling but the flesh is weak”(心有余而力不足)翻译成俄语,然后再翻译回来时,竟变成“The wine is good but the meat is spoiled”(酒是好的,肉变质了)。

这些问题出现后,人们发现机器翻译并非想象的那么简单,人们认识到单纯地依靠“查字典”的方法不可能解决机器翻译问题,只有在对语义理解的基础上,才能做到真正的翻译。

因此,机器翻译被认为是人工智能的完整性问题,即为了解决其中一个问题,必须解决全部问题,哪怕是一个简单和特定的任务。例如,在机器翻译组中,要就机器按照作品的观点(推理),知道人物经常谈论些什么(知识),如何忠实地再现作者的意图(情感计算)。因此,机器翻译被认为具有人工智能的完整性。

(3) 模式识别

识别是人和生物的基本职能之一,人们几乎时时刻刻在对周围环境进行识别。模式识别是指计算机对物体进行识别。物体指文字、符号、图形、图像、语言及传感器信息等形式的实体对象,但不包括概念、思想、意识等抽象或虚拟对象。后者的识别属于心理、认知及哲学等学科的范畴。也就是说,模式识别是狭义的模式识别,它是人和生物个体感知能力在计算机上的模拟和扩展。模式识别目前在部分领域取得了很好的应用,如:文字识别(书籍文字的数字化)、车牌识别、指纹识别、人脸识别、视网膜识别、掌纹识别等。

(4) 文艺创作

在文艺创作领域,人们也尝试开发和运用人工智能技术。现在计算机创作的诗词、小说、音乐、绘画时有报道。例如下面的两首“古诗”就是计算机创作的。

云松

无题

銮仙玉骨寒,松虬雪友繁。

白沙平舟夜涛声,春日晓露路相逢。

大千收眼底,斯调不同凡。

朱楼寒雨离歌泪,不堪肠断雨乘风。

以上仅仅讨论了人工智能应用的部分领域和部分课题。目前人工智能的研究与实际应用的结合越来越紧密,受应用的驱动越来越明显。

6. 强人工智能和弱人工智能的争论

(1) 强人工智能

强人工智能观点认为有可能制造出真正能推理和解决问题的智能机器,并且,这样的机器将被认为是有知觉的,有自我意识的。强人工智能可以有两类:一种是类人的人工智能,即机器的思考和推理就像人的思维一样;另一类是非人类的人工智能,即机器产生了与人完全不一样的知觉和意识,使用与人完全不一样的推理方式。持这一观点的代表人物是图灵。

(2) 弱人工智能

弱人工智能观点认为不可能制造出能真正推理和解决问题的智能机器,这些机器只不过看起来像是智能的,但是并不是真正拥有智能,也不会有自主意识。持这一观点的代表人物是希尔勒(John Rogers Earle, 1932—, 美国)。

(3) 人工智能的哲学争论

如果一台机器的唯一工作原理就是转换码数据,那么这台机器是不是有思维的? 希尔勒认为这是不可能的。他举了一个“中文房间”的例子进行说明。如果机器仅仅是转换数据,而数据的本身是对某些事情的一种编码,那么在不理解这一编码与实际事情之间对应关系的前提下,机器不可能对其处理的数据有任何理解。基于这一论点,希尔勒认为即使有机体通过了图灵测试,也不一定说明机器就真的像人一样有思维和意识。

也有哲学家持不同的观点。丹尼尔·丹尼特(Daniel Dennett,1942—,美国)在著作《意识的解释》里认为:人也不过是一台有灵魂的机器而已,为什么我们认为:“人可以有智能,而普通机器就不能”呢?他认为像上述数据转换机器的有可能有思维和意识的。

有些哲学家认为,如果弱人工智能是可实现的,那么强人工智能也是可实现的。如西门·布莱克伯恩(Simon Blackburn,英国)在哲学入门教材《思考》里谈到,一个人看起来是“智能”的行动,并不能真正说明这个人就是智能的。我永远不可能知道另一个人是否真的像我一样是智能的,还是说他仅仅是看起来是智能的。基于这个论点,既然弱人工智能认为可以令机器看起来是智能的,那就不能完全否定这机器是真的有智能的。布莱克伯恩认为这是以后主观认定的问题。

需要指出的是,弱人工智能并非和强人工智能完全对立,也就是说,即使强人工智能是不可能的,弱人工智能仍然是有意义的。至少,今天计算机能做的事,在一百多年前是被认为很需要智能的。

吉林大学出版社

2.1 计算机系统工作原理

2.1.1 指令、指令系统与程序

1. 指令

指令是一组能被计算机识别并执行的二进制数据代码,是让计算机完成某个操作的命令。一条指令通常由两个部分组成,前面部分称为操作码,后面部分称为操作数。操作码指明该条指令要完成的操作,如加、减、乘、除、逻辑运算等。操作数是指参加运算的数据或者数据所在的存储单元地址。

2. 指令系统

一台计算机所能执行的全部指令的集合,称为这台计算机的指令系统。指令系统比较充分地说明了计算机对数据进行处理的能力。不同种类的计算机,其指令系统的指令数目与格式也不同。指令系统是根据计算机使用要求设计的。

指令系统是计算机基本功能具体而集中的体现。但无论哪种类型的计算机,指令系统都应具有以下功能的指令:

- (1)数据传送指令:将数据在内存与 CPU 之间进行传送。
- (2)数据处理指令:对数据进行算术或逻辑运算。
- (3)程序控制指令:控制程序中指令的执行顺序。如条件转移、无条件转移、调用子程序、返回、暂停、终止等。
- (4)输入/输出指令:用于实现外部设备与主机之间的数据传输。
- (5)其他指令:对计算机的硬件进行管理。

3. 程序

用户根据解决某项问题所需的步骤,选择适当的指令,将它们一条一条地按照某种顺序进行有序的排列,计算机依次执行这些指令序列,便可完成预定的任务。按照一定要求组织

构成的可完成若干项操作的指令序列就是程序。

2.1.2 计算机的工作原理

1. 冯·诺依曼原理

“存储程序控制”原理是 1946 年由美籍匈牙利数学家冯·诺依曼提出的,所以又称为“冯·诺依曼原理”。该原理确立了现代计算机的基本组成的工作方式,直到现在,计算机的设计与制造依然沿用“冯·诺依曼”体系结构。

2. “存储程序控制”原理的基本内容

(1)采用二进制形式表示数据和指令。

(2)将程序(数据和指令序列)预先存放在主存储器中(程序存储),使计算机在工作时能够自动高速地从存储器中取出指令,并加以执行(程序控制)。

(3)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计算机硬件体系结构。

3. 计算机工作过程(如图 2.1)

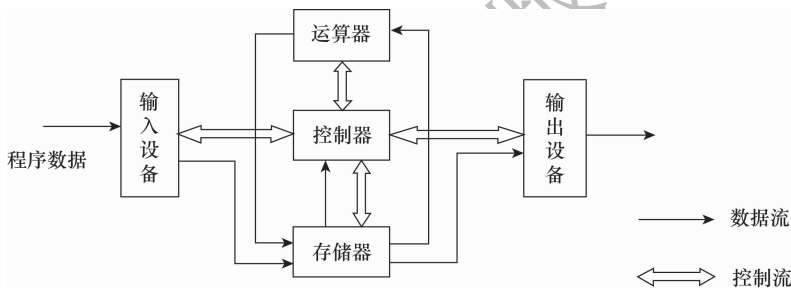


图 2.1 计算机工作原理图

第一步:将程序和数据通过输入设备送入存储器。

第二步:启动运行后,计算机从存储器中取出程序指令送到控制器去识别,分析该指令要做什么事。

第三步:控制器根据指令的含义发出相应的命令(如加法、减法),将存储单元中存放的操作数据取出送往运算器进行运算,再把运算结果送回存储器指定的单元中。

第四步:当运算任务完成后,就可以根据指令将结果通过输出设备输出。

2.2 计算机系统组成

一个完整的计算机系统是由计算机硬件系统和计算机软件系统两部分组成(如图 2.2)。硬件系统是构成计算机系统的各种物理设备的总称。软件系统是运行、管理和维护计算机的各类程序和文档的总和,是计算机的灵魂。

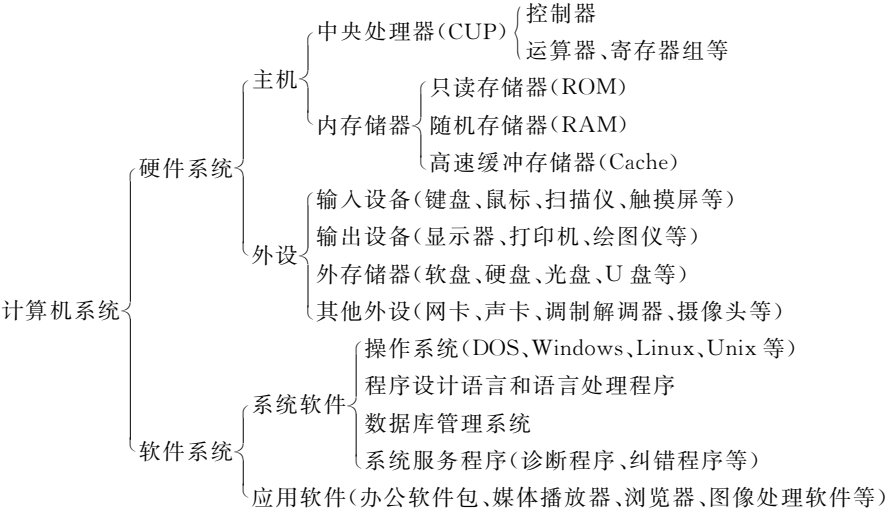


图 2.2 计算机系统的组成

2.3 计算机硬件系统

在计算机系统中,硬件系统是构成计算机系统各个功能部件的物理实体,是计算机能够工作的物质基础,这些部件一般是由电子电路和机械设备构成的。一个计算机系统性能的高低在很大程度上取决于硬件的性能配置。

2.3.1 计算机硬件系统组成

根据冯·诺依曼提出的计算机设计思想,计算机的硬件结构主要由五部分构成。

1. 控制器

控制器是计算机系统的神经中枢和指挥中心,用于控制、指挥计算机系统的各个部分协调工作。其基本功能是从内存中取出指令,对指令进行分析,然后根据该指令的功能向有关部件发出控制命令,以完成该指令所规定的任务。它主要由程序计数器、指令寄存器、指令译码器、操作控制电路和时序控制电路等组成。

2. 运算器

运算器又称算术逻辑单元,是对信息进行加工处理的部件,主要由算术逻辑运算器、累加寄存器、数据缓冲寄存器和状态寄存器等组成。运算器的主要功能是在控制器的控制下,对取自内存或寄存器的二进制数据进行各种加工处理,如加、减、乘、除等算术运算和与、或、非、异或等逻辑运算后,再将运算结果暂存在寄存器或送到内存中保存。

控制器和运算器之间在结构关系上非常紧密,随着半导体工艺的进步,运算器和控制器集成在一个芯片上,形成中央处理器(Central Processing Unit,CPU)。

3. 存储器

存储器是用来存储程序和数据部件,有了存储器,计算机才有记忆功能,才能保证正常工作。按用途,存储器可分内存储器和外存储器两大类。

(1)内存储器。内存又称主存储器(如图2.3),主要用来存放CPU工作时用到的程序和数据以及计算后得到的结果。它的特点是工作速度快、容量较小、价格较高。内部存储器按读写方式又可分为两类。

1)只读存储器(ROM):ROM主要用来存放固定不变的程序、数据,如BIOS程序,这些程序是它们厂商在制造时用特殊方法写入的,断电后其中的信息不会丢失。主板和显卡上的BIOS芯片(存储CPU参数、内存参数、芯片组参数)属于ROM。

2)随机存储器(RAM):RAM是一种读写存储器,其内容可以随时根据需要读出,也可以随时重新写入新的信息,由于信息是通过电信号写入的,因此,在计算机断电后RAM中的信息会丢失。内存芯片、显存芯片,CPU的缓存(Cache)都属于RAM。

3)高速缓冲存储器(Cache):主存由于容量大、寻址系统繁多、读写电路复杂,造成了主存的工作速度大大低于CPU的工作速度,影响了计算机的性能。高速缓冲存储器(Cache)解决了主存速度慢与CPU运算快的矛盾。Cache中存放常用的程序和数据,当CPU访问这些程序和数据时,首先从高速缓存中查找,如果所需程序和数据不在Cache中,则到主存中读取数据,同时将数据写入Cache中。因此采用Cache可以提高系统的运行速度。CPU和内存存储器构成计算机的主机。

(2)外存储器。又称辅助存储器,是内存存储器的补充和后援,主要用于存放计算机当前不处理的程序和大量的数据。保存在外存储器中的程序和数据只在需要时,才会调入到内存中,再由CPU执行或处理。外存储器容量大,保存的程序和数据在断电后也不会丢失,弥补了内存存储器RAM的容量小,断电后丢失数据的缺陷。

1)软盘存储器

软盘由圆形塑料薄片表面镀磁粉,然后外加硬的塑料护套组成,目前基本淡出市场(如图2.4)。计算机通过软盘驱动器对软盘进行读写。常用的3.5英寸(1英寸=2.54cm)软盘,它的存储容量为1.44MB。软盘存储信息的方式是将软盘的两面划分成磁道,磁道内再划分为扇区来存储数据。磁道是由外向内的一个个同心圆,磁道编号从外向内越来越大,每个磁道又等分成若干个扇区。1.44MB软盘片有两面,每面80个磁道,每道18个扇区,每个扇区存储512个字节。(如图2.5)

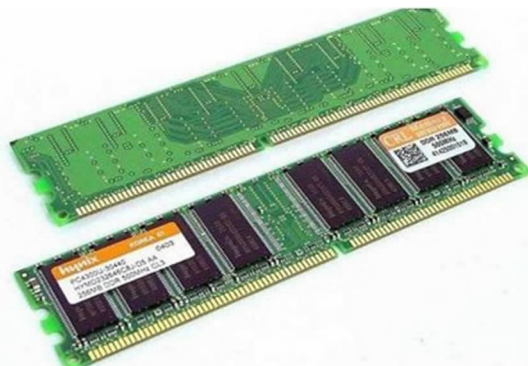


图 2.3 内存条外观



图 2.4 软盘

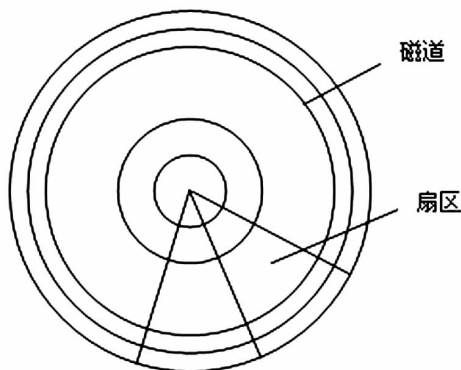


图 2.5 磁道及扇区

2) 硬盘存储器

硬盘是由固定在机箱内硬质的合金材料构成的多张盘片组成,连同驱动器一起密封在壳体中。硬盘多层磁性盘片被逻辑划分为若干同心柱面,每一柱面又被分成若干个等分的扇区,每个扇区的容量也与软盘一样,通常是 512 个字节。

硬盘常常被封装在硬盒内,固定安装在机箱里,难以移动。因此,它不能像软盘那样便于携带,但它比软盘存储信息密度大、容量大、读写速度也比软盘快。所以,人们常用硬盘来存储经常使用的程序和数据。目前微型机配备的硬盘存储容量大多在 120~320GB。(如图 2.6)



图 2.6 硬盘内部结构

3) 光盘存储器

光盘存储器包括光盘驱动器(光驱)和光盘,光驱是采用激光扫描的方法从光盘上读取信息,光盘是利用塑料基片的凹凸来记录信息的。光盘具有存储容量大、记录密度高、读取速度快、使用寿命高的特点,并且携带方便、价格低廉,已经成为存储数据的重要手段。目前市面上的光盘包括 CD-ROM、DVD-ROM 和可擦写型光盘 3 种。

4) 移动存储器

移动存储器是一种可以直接插在通用串行总线 USB 端口上的能读写的外存储器,主要是指移动硬盘和闪存。移动硬盘实质上就是将笔记本硬盘加上移动硬盘盒构成,利用它可以将大量数据随身携带,弥补了电脑硬盘容量的不足。闪存,也就是 U 盘,由于它具有体积小、重量轻、读写速度、保存信息可靠等优点,通常作为软盘的替代品。

4. 输入设备

输入设备是向计算机输入程序、数据和命令的部件。输入设备的种类很多,目前微机上常用的有键盘、鼠标器,有时还用到扫描仪、条形码阅读器、手写输入装置及语音输入装置等。

(1) 键盘

键盘是计算机必备的标准输入设备,用户的程序、数据以及各种对计算机的命令都可以通过键盘输入。常用的键盘有 101 键、104 键和 108 键等,PC 机一般使用 104 键盘。键位一般分为四个区,主键盘(打字)区、功能键区、编辑键区、数字小键盘区,如图 2.7 所示。



图 2.7 键盘

1) 主键区

主键区是键盘的主要使用区,用来输入各种字母、数字、常用运算符、标点和汉字等。主键区常用键的作用如下:

Enter:回车键,换行键。

Caps Lock:大小写字母转换键。

Shift:上档键,常与其他键或鼠标组合使用,主要用于输入键位上方的字符。

Ctrl:控制键,常与其他键或鼠标组合使用。

Alt:变换键,常与其他键组合使用。

Backspace:退格键,按一次,消除光标前边的一个字符。

Tab:制表键,按一次,光标跳到下一个制表位。

2) 功能键区

功能键区的按键又分为功能键(Esc、F1~F12)和控制键。在不同的软件中,可以对功能键 F1~F12 进行重新定义。

3) 编辑键区

编辑是指在整个屏幕范围内,对光标的移动和有关的编辑操作等。该键区的光标移动键位只有在运行具有编辑功能的程序中才起作用。该键区的操作主要有以下几类:

↑、↓、←、→:相应为光标上移一行、光标下移一行、光标左移一列、光标右移一列。

Home、End、Page Up、Page Down:光标移到行头或当前页头、光标移到行尾或当前页尾、光标移到上一页、光标移到下一页。

Delete:删除光标当前位置及后边的一个字符。

Insert:设置改写或插入状态。

4) 小键盘区

是为专门从事录入数据的工作人员提供方便而准备的,包括:10 个数字键以及+、-、×、/等键,共 17 个。

(2) 鼠标

鼠标也是一种输入设备,主要应用于图形界面的系统,可以快速移动选择对象并完成特定操作。目前,微机上常用的鼠标有机械式和光电式 2 种。鼠标有指向、单击、双击、拖动和右键单击 5 种基本操作。

(3) 扫描仪

扫描仪是计算机输入图片使用的主要设备,它内部有一套光电转换系统,可以把各种图片信息转换成计算机图像数据,并传送给计算机,再由计算机进行图像处理、编辑、存储、打印输出或传送给其他设备。

按色彩来分,扫描仪分成单色和彩色两种;按操作方式分,可分为手持式和台式扫描仪。扫描仪的主要技术指标有分辨率、灰度层次、扫描速度等。

5. 输出设备

输出设备就是由计算机向外输出信息的外部设备。常用的输出设备有显示器、打印机、绘图仪等。

(1) 显示器

显示器由监视器和显示适配器两部分组成,是微型计算机不可缺少的输出设备,是实现人机对话的主要工具。它既可显示程序运行的结果,又可显示输入的程序和数据等。显示器按大小来分,有 17 英寸、19 英寸、21 英寸等规格。按显示原理来分,主要有阴极射线管(CRT)显示器(如图 2.8)和液晶(LCD)显示器(如图 2.9)。



图 2.8 阴极射线管显示器



图 2.9 液晶显示器

显示器的主要技术指标有屏幕尺寸、点距、显示分辨率、灰度和颜色深度及刷新频率。

分辨率:是指屏幕上横向、纵向发光点的点数。一个发光点称为一个像素。分辨率越高,显示图像越细腻越清晰。目前常见显示器的分辨率有 640×480 、 800×600 、 1024×768 、 1280×1024 等。

灰度级:每个像素点的亮暗层次级别,或者可以显示的颜色数目,其值越高,图像层次越清楚逼真,若用 8 位来表示一个像素,则可有 256 级灰度或颜色。

刷新频率:指每秒钟内屏幕画面刷新的次数。刷新频率越高,画面闪烁越小,通常是

75Hz~90Hz。

(2) 打印机

打印机是计算机系统中常用的输出设备。可以将电子化的各种文档,如文字、图形、图像输出到纸张上。根据打印机的工作原理,可以将打印机分为针式打印机、喷墨打印机和激光打印机3类。(如图2.10—图2.12)



图 2.10 针式打印机



图 2.11 喷墨打印机



图 2.12 激光打印机

针式打印机:目前主要用于发票、存款凭票等专用票据的打印,在银行、财务部门里应用较广。它在打印时是用带色的钢针击打纸张,所以打出的文字会有明显的色斑,从而具有一定的防伪效果。

喷墨打印机:是通过喷墨头喷出的墨滴来进行打印的。它的价格低廉、打印速度适中且打印质量较好,非常适合普通家庭和学生使用。

激光打印机:打印速度快、质量高,由于它的价格较高,目前主要用于公司或集团用户。

(3) 绘图仪

绘图仪是一种输出图形硬拷贝的输出设备。绘图仪可以绘制各种平面、立体的图形,已成为计算机辅助设计(CAD)中不可缺少的设备。绘图仪按工作原理分为笔式绘图仪和喷墨绘图仪。它主要运用于建筑、服装、机械、电子、地质等行业中。如图2.13所示。



图 2.13 绘图仪

2.3.2 计算机主要性能指标

1. 字长

字长指的 CPU 中寄存器存储单元的长度,即 CPU 一次能够直接处理的二进制数据的

位数。它的长度直接关系到计算机的计算精度、运算速度和功能的强弱,常用于衡量 CPU 的性能。一般情况下,字长越长,计算精度越高,处理能力越强。微处理器的字长已从早期的 4 位、8 位,发展到了 16 位、32 位,目前已到 64 位。

2. 运算速度

运算速度是指计算机每秒钟能够执行的指令数目。常用单位是 MIPS(百万条指令/秒)。

3. 主频

主频指的是计算机的时钟频率,是决定计算机运算速度的一个重要指标。一般来讲,时钟频率越高,运算速度越快。主频的单位是 MHZ 或 GHZ。

4. 内存容量

内存容量通常指的是在内存存储器的 RAM 中能够储存信息的总字节数。它的大小反映了计算机存储程序和处理数据能力的大小,容量越大,运行速度越快。

5. 外部设备的配置

主机所配置的外部设备的多少与好坏,也是衡量计算机综合性能的重要指标。

6. 软件的配置

合理安装与使用丰富的软件可以充分地发挥计算机的作用和效率,方便用户的使用。

2.4 计算机软件系统

2.4.1 计算机软件系统概述

仅有硬件,计算机什么事情也不能干,硬件是计算机的实体,而软件是计算机的“灵魂”,计算机的硬件系统只有与软件系统密切配合,才能正常工作。所谓“软件”是各种程序的总称,不同功能的软件由不同的程序组成,这些程序通常被存储在计算机的外部存储器中,需要使用时装入内存运行。软件的作用是为方便用户使用计算机,充分而有效地发挥计算机的功能。软件系统的好坏会直接地影响计算机的应用。

2.4.2 计算机软件系统分类

计算机软件系统内容丰富,通常将软件分为两大类:系统软件和应用软件。

1. 系统软件

系统软件是为管理、监控和维护计算机资源所设计的软件。它由计算机软件生产厂商研制提供,主要包括操作系统、语言处理程序、数据库管理系统、服务性程序等。

(1) 操作系统(Operating System, OS)

操作系统是最重要的系统软件,是用户和计算机之间的接口。它是最底层的系统软件,是对硬件系统的首次扩充。通常它的主要任务是管理好计算机的全部资源,使用户能充分、

有效地利用这些资源。

目前微机上常用的操作系统有:Windows 系列操作系统、MS-DOS 操作系统、Unix 操作系统和 Linux 操作系统等。

1) 操作系统的功能

计算机系统的资源可分为设备资源和信息资源两大类。设备资源指的是组成计算机的硬件设备,如中央处理器,主存储器,磁盘存储器,打印机,磁带存储器,显示器,键盘输入设备和鼠标等。信息资源指的是存放于计算机内的各种数据,如文件,程序库,知识库,系统软件和应用软件等。因此,从资源管理的观点出发,操作系统的功能可归纳为处理器管理(作业管理和进程管理)、存储器管理、设备管理和文件管理。

处理器管理:是操作系统资源管理功能的一个重要内容。在一个允许多道程序同时执行的系统里,操作系统会根据一定的策略将处理器交替地分配给系统内等待运行的程序。一道等待运行的程序只有在获得了处理器后才能运行。一道程序在运行中若遇到某个事件,例如启动外部设备而暂时不能继续运行下去,或一个外部事件的发生等等,操作系统就要来处理相应的事件,然后将处理器重新分配。

存储管理:就是负责把内存单元分配给需要内存的程序以便让它执行,在程序执行结束后将它占用的内存单元收回以便再使用。对于提供虚拟存储的计算机系统,操作系统还要与硬件配合做好页面调度工作,根据执行程序的要求分配页面,在执行中将页面调入和调出内存以及回收页面等。

设备管理:主要是分配和回收外部设备以及控制外部设备按用户程序的要求进行操作等。对于非存储型外部设备,如打印机、显示器等,它们可以直接作为一个设备分配给一个用户程序,在使用完毕后回收以便给另一个需求的用户使用。对于存储型的外部设备,如磁盘、磁带等,则是提供存储空间给用户,用来存放文件和数据。

文件管理:主要是向用户提供一个文件系统。一般来说,一个文件系统向用户提供创建文件,撤销文件,读写文件,打开和关闭文件等功能。有了文件系统后,用户可按文件名存取数据而无需知道这些数据存放在哪里。这种做法不仅便于用户使用而且还有利于用户共享公共数据。此外,由于文件建立时允许创建者规定使用权限,这就可以保证数据的安全性。

2) 操作系统的分类

随着计算机软件 and 硬件技术的不断发展,已形成了多种不同类型的操作系统,以满足用户各种不同的应用要求,按照系统功能,通常把操作系统的类型分为:

单用户操作系统,系统每次只能支持运行一个用户程序,如在微型计算机上所使用的 MS-DOS。

批处理操作系统,用户将一批算题(或作业)输入到计算机,然后由操作系统来控制作业自动执行,不需要用户对每个作业进行控制。这些若干个计算问题可以同时执行,它们可共享计算机系统的资源。因此提高了整个系统的效率。在多用户小型计算机上主要使用批处理操作系统,如 UNIX 操作系统等。

分时操作系统,硬件上由 1 台主机连接多个用户终端组成,允许多个终端用户同时使用 1 台主机,并保证各用户彼此独立,互不干扰。它主要采用排队及时间片的分时策略,使各终端用户的要求都能及时得到处理。分时操作系统主要侧重于及时性和交互性。

实时操作系统,它能满足计算机系统及时对外部事件的请求进行响应,并在一个较短的时间内尽快完成对该事件的处理,以求实现计算机的实时控制。其及时性要求比分时操作系统的要高。常用于工业控制、导弹发射、飞机航行等领域。

网络操作系统,它除了具有通常操作系统的功能以外还提供了网络通信和网络共享等功能。利用网络“协议”,将许多计算机连成网络,使网络中的计算机系统的资源能共享并进行通信。

分布式操作系统,分布式操作系统也是通过网络将多台计算机连接起来,实现相互通信和资源共享,但它更强调将一个大的算题划分成小任务,分布到不同的计算机上协作完成。

(2) 程序设计语言

要使计算机按照用户的要求去工作,必须使计算机能够接受,并懂得人输送给它的各种命令和数据,而且还应当能够将运算处理后的结果反馈给人。人与计算机之间的这种信息交流同样需要语言。程序设计语言就是人与计算机交流信息的语言工具,提供了让用户按自己的需要编制程序的功能。计算机语言通常分为三大类:机器语言、汇编语言和高级语言。

机器语言:是由二进制代码“0”和“1”组成的一组代码指令,是唯一能被计算机硬件直接识别和执行的语言。机器语言的二进制代码随着 CPU 型号的不同而不同,因此机器语言程序在不同的计算机系统之间不能通用,故将其称之为面向机器的语言。它的优点是占用内存小,执行速度快。缺点是编写程序工作量大、程序可读性差。

汇编语言:也是一种面向机器的程序设计语言,是一种把机器语言符号化的语言。用助记符代替操作码,用地址符代替地址码。如用 ADD 表示加法,用 SUB 表示减法,用 MOV 表示移动,用 LD 表示取数据,D5H 表示两位十六进制的数据等。因为汇编语言的语句和机器指令有对应关系,从而保留了机器语言的优点—执行速度快,所以汇编语言目前仍在使用,主要用于实时控制等对响应速度有极高要求的场合。但编辑复杂,可移植性差。这种程序必须经过翻译(称为汇编)成机器语言程序才能被计算机识别和执行。

高级语言:为了解决机器语言和汇编语言编程技术复杂、编程效率低、通用性差的缺点,20 世纪 50 年代后研制开发了高级语言。高级语言是面向解题过程或者面向对象的语言,按照一定的语法规则编写程序。它们的语句比较接近人类使用的自然语言和数学表达式。用高级语言编写的程序易读、易记、易维护、且通用性强,便于推广和交流。常用的面向过程的高级语言有 Basic、Pascal、Fortran、C 语言等,面向对象的高级语言有 C++、Visual Basic、Delphi、Power builder、JAVA 等。

用高级语言编写的程序(源程序)不能被计算机直接识别、接收和执行,需要用翻译程序将其翻译成机器指令程序(目标程序)才能执行。根据翻译方式的不同,可分为两类:“编译”

方式和“解释”方式。

编译方式是用编译程序(事先编好并装入计算机)将源程序完整地翻译成等价的目标程序后,再执行该目标程序。大部分高级语言都是(或都具有)编译方式,如 Fortran、Pascal、Visual Basic、C/C++、Power builder 等。(如图 2.14)

解释方式是用解释程序将源程序逐句进行翻译,翻译一句执行一句,边翻译边执行,不产生目标程序。如 Basic、FoxBASE,开发阶段的 Power Builder、Visual Basic 等。(如图 2.15)

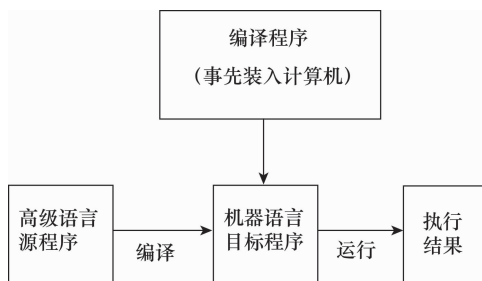


图 2.14 编译过程

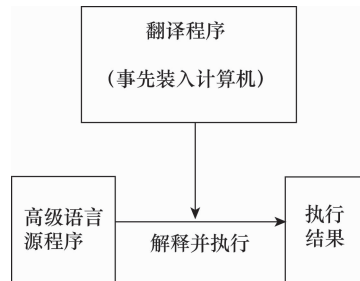


图 2.15 解释过程

(3)数据库管理系统

数据库管理系统向用户提供按照一定的结构组织、管理、加工、处理各类数据的能力,如 Access、Oracle、SQL Server 等。

(4)服务程序

用于调试、检测、诊断、维护计算机软、硬件的程序,如连接程序 Link,编辑程序 Editor 等。

2. 应用软件

应用软件是人们为解决某种问题而专门设计的各种各样的软件,这些软件可以帮助人们提高工作质量和效率。一个计算机系统的应用软件越丰富,越能发挥计算机的作用。目前,微机上广泛使用的应用软件主要有:

(1)文字处理软件 这类软件主要用于办公事务的处理,包括文章的编辑排版、表格的处理、演示文稿的制作等,使用这些软件可实现办公自动化。常用的文字处理软件有 WORD、Excel、PowerPoint、WPS 等。

(2)网络应用软件 这类软件主要用于帮助用户实现网上资源的浏览、远程信息的传送、电子邮件的收发等。常用的网络应用软件有 Internet Explorer、Outlook、Express、Netscape 等。

(3)信息管理软件 信息管理软件主要是使用数据库技术,实现学籍管理、财务管理、人事管理等各类大量数据信息的存储、查询、统计和报表打印。

(4)实用工具软件 微机软件系统中,有许多工具型应用软件,这类软件能帮助用户完成某些专项任务。如多媒体制作软件 Authorware,图形制作软件 Photoshop,辅助设计软件 Auto CAD,数学工具软件 Mathlab。