



电子计算机 情报检索

陕西省科学技术情报学会
陕西省电子技术情报站

电子计算机情报检索

陕西省科学技术情报学会

陕西省电子技术情报站

1980年2月 西安

前 言

要在本世纪末实现四个现代化，速度是个核心问题。用以电子计算机为代表的现代化手段和新技术武装科技情报工作，及时了解和全面掌握世界科学技术的最新成就和动向，对实现高速度地赶超世界先进科技水平，具有重要意义。

为了向科技情报人员普及电子计算机情报检索的知识，交流和学习有关经验和技术，我们于一九七九年四月份在西安举办了陕西省电子计算机情报检索技术交流会，并根据各单位的要求，搜集有关资料选编本书内部出版。

本书由徐连科、施维复、王仲东、周伟等同志编辑。由于水平有限，时间仓促，书中缺点在所难免，敬希读者指正。

(001) 第 一 章 文 献 情 报 学 概 论

目 录

一、前言	
二、电子计算机原理	(1)
三、COBOL程序设计	(46)
四、DJ S—130机的指令系统和程序设计方法	(77)
五、数据库系统	(107)
六、主题词表(典)	(124)
七、《国防科学技术主题词典》及其在情报检索中的应用	(129)
八、计算机情报检索	(146)
九、微型机在情报检索中的应用	(154)
十、中文信息化与文字改革	(160)

电子计算机原理简介

陕西省电子技术研究所 翟高齐

说 明

华主席在五届人大政府工作报告中指出：“现代科学技术，以原子能的利用，电子计算机技术和空间科学技术的发展为主要标志，正在经历着一场伟大的革命，引起一系列新兴工业的诞生，广泛推动生产技术的飞跃发展”。“大力发展新兴科学技术，建设原子能电站，发射多种应用卫星，加强激光理论和应用的研究，重视遗传工程的研究，特别是加速发展集成电路和电子计算机的研究，并使它们广泛应用于各个方面”。

这份“简介”是在陕西省电子计算机情报检索交流会议上所作的科普报告，应会议上同志们的要求，现把讲稿整理出来，供有关同志们参考。由于是科普性报告，所以不能讲得很细，侧重于概念性和逻辑原理、功能方面的介绍。如果有些同志工作需要或有兴趣，可参考北京大学电子仪器厂编著的《电子数字计算机原理》、铁道部电子计算机编写组写的《电子计算机原理》及有关书籍。

第一章 概 述

一、计算机的分类和特点

1. 计算机分类

(1) 按表示数值的方法

A、模拟机：连续式，连续量作用的计算机。例如，计算尺就是利用尺上的有限长度来表示出被计算的数值。模拟式电子计算机，是以连续变化的电压作为输入，而算得的结果又是以连续变化形式的电压输出出来（用图纸或示波器描述出连续变化曲线）。

由于使用元件不同，模拟机又分为机械的、电磁的和电子式的。

在解微分方程，实时控制和机车牵引控制等方面都应用了电子模拟机。

缺点：精确度受限制（百分之0.1—3内）。

B、数字机：断续式，断续量作用的计算机。例如，算盘就是断续作用形式的一种。由单个算珠组成，每一个算珠代表一个数值，对算珠进行的运算，也就是对数字进行运算。手摇或电动计算机是应用齿轮上的牙齿或叫转角来代替数字进行运算的，这类计算机也可称为数字计算机。而对于电子数字计算机来说，它的数值是应用电脉冲个数或电流、电压的跳变形式实现对数值的计算。

在数字机上，问题的解，都是由逐个完成算术运算组成的，准确度高（准确度取决

于表示数的位数，即机器的字长），通用性强。

对于变数来说，模拟机的变数是连续变化的物理量，而数字机的变数则是不连续的数字。

(2) 按操作原理和功能：

A、通用机：实现任何解题的算法，用程序控制。

B、专用机：实现一种或几种一定的解题算法，用固定式程序控制。

在专用计算机中，又可分为处理信息的控制机和处理数据的数据处理机。

由于当前广泛应用的是电子数字计算机，所以，我们通常习惯上所说的“电子计算机”一般都是指“电子数字计算机”而言，严格来说，电子计算机应该包括：电子数字计算机和电子模拟计算机。凡是模拟计算机都加上“模拟”二字。我们下面所谈的，如不特别注明，也都是指“电子数字计算机”。

这里顺便再介绍两个名词：硬件和软件。

“硬件”是指电子计算机的机器系统，通常称为硬件或叫硬设备。硬设备是整个计算机系统的物理装置。

“软件”是指电子计算机的程序系统，通常称为软件或叫软设备。也就是管理和运用机器的程序，也可以说是计算机所具有的各种程序的总和。

程序就是机器解题的工作顺序，每一步计算和操作都是一道指令，而指令就是用来指定实现某种控制或运算的代码，每一条指令应该指出操作功能和操作对象等内容。

从电子数字计算机要讲的问题来看，就是概括为硬件和软件两大部分。

2. 计算机的特点：

(1) 计算速度快

由于电子计算机主要部件大部分采用快速的电子元件组成，因此它的工作速度远远超过了机械式或机电式的计算工具。当前已经制成了一些大型、高速的计算机，运算速度达到每秒钟几百万至几千万次，甚至最高达到一亿五千万次。

(2) 精确度高

采用二进制，使数据的表示，保存和运算都能以很高的精确度进行，从千分之几到百万分之几甚至更高。

(3) 具有“记忆”能力和逻辑判断能力

计算机的存储器保存着大量的解题程序数据等信息。计算机还可以进行各种逻辑判断，如判断数据的大小、性质等，并能根据判断的情况自动决定计算机应执行什么命令。这种功能使计算机能自动工作。

(4) 自动地工作

计算机能在程序控制下自动地进行工作，不需要人参与计算过程。

随着计算机的发展，计算机自动工作的程度愈来愈高。特别是“软件”的发展，使用编译程序和管理程序，只要向计算机输入一定的语言、符号、文字、图形，计算机系统就能自动地进行工作，并可以按照人们的要求打印出数据、绘出图形、表格等。

(5) 通用性强

A、科研与工程设计方面

利用计算机帮助人们进行各种研究和设计工作（设计样品性能模拟和方案比较等）取得比过去用实验方法和粗略估算方法精确得多的结果，从而大大减少研究、试制的时间和费用。

B、数据处理方面

企事业单位的管理、统计工作，农业、生物、医学大量实验数据的分析，情报检索等等都可以利用电子计算机，它能提高工作效率，避免差错，即时预报各种统计结果。

C、自动控制方面

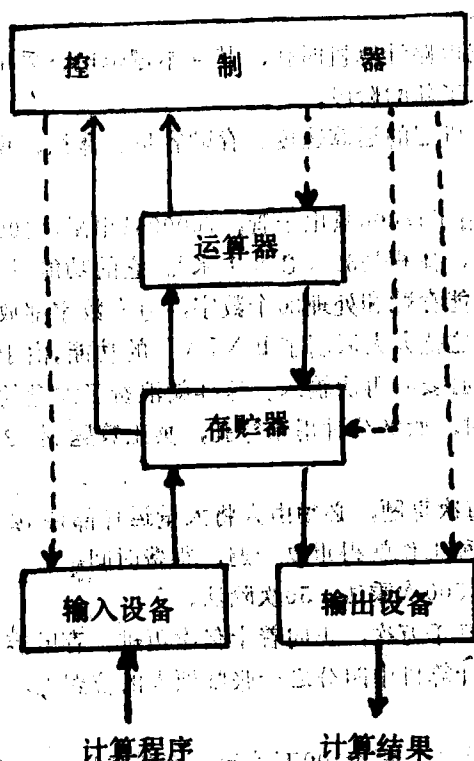
电子计算机与自动仪表配合在一起，能够使生产过程实现闭环的最优控制，实现生产过程的全盘自动化。

计算机在实时自动控制方面应用也很广泛，如对飞机、导弹、人造卫星、火箭等实际目标的跟踪、观测与控制等。

在一些对人体有严重危害的工作中，也离不开电子计算机的控制，如在原子能工业中，在一些高温下的工作中，在一些易燃易爆腐蚀性比较强的工作中，人不能直接参与生产过程的操作，这就要求使用机器人来操作，而机器人又离不开电子计算机的控制。

二、电子数字计算机的逻辑结构框图

电子数字计算机主要由以下几个部分组成（见下图）：



计算程序
原始数据

计算结果

1. 运算器——对数据快速地进行加、减、乘、除等算术运算及其他一些常见的逻辑运算（如比较两数的大小等）。

在运算过程中，运算器不断地得到由存储器提供的数据，并能把求得的结果送回存储器保存起来（少量运算结果可以暂时存在运算器内，如中间结果，下次运算中要用，以后也没有保存的价值，所以就不存入存储器）。运算器的整个运算过程是在控制器的统一指挥下，按程序中编排的先后次序有规律地进行的。

2. 控制器——是整个计算机的中枢，是最重要的一部分，统一指挥整个机器的工作。控制器从存储器顺序地取出指令，并向各部件发出相应的命令，使它们一步步地执行程序所规定的任务。控制器还根据各“执行部件”（如运算器、存储器，输入—输出部件等）向它发回的“反馈信息”，作为下一步发出那些命令的工作条件。如运算器向控制器发回“报告”计算

结果是否超出预定界限的“反馈信息”；控制器将决定下一步是按顺序执行指令还是“中断”（暂停）处理。

3. 存储器——保存大量的信息，即存储数据、程序、资料等。在使用时，根据实际的需要可以把原来记录的内容抹去而重新记录新内容，或者把原记录的内容取出但不破坏原有的记录。

4. 输入—输出设备——是计算机和外界联系的桥梁和通道。输入数据和程序，输出计算结果。

三、计算机发展概况

世界上最早的计算工具是我国的算盘。

1945年世界上第一台电子计算机（ENIAC）在美国诞生，46年正式投入使用，虽然到现在仅30多年的时间，组成计算机的电子元件的性能不断发展，计算机也飞跃发展，大体经历了三个年代跨入第四年代。

第一代：1945—55年，电子管计算机时代，其基本逻辑电路由电子管构成。运算速度约为每秒钟几十次至几千次。

第二代：1956—64年，晶体管计算机时代（分立元件），其基本逻辑电路由晶体管构成。运算速度约为每秒钟几万至几十万次。

第三代：1965—69年，小面积集成电路计算机时代，其基本逻辑电路由中、小规模集成电路构成。速度约为百万次至千万次/秒。

第四代：1970年开始，进入了大面积集成电路计算机时代，其基本逻辑电路采用了大规模集成电路，电子数字计算机的发展进入了新的阶段。

计算机跨越了四个时代后，使用的器件、机器的运算速度、存储容量、体积、重量和成本等都发生了巨大的变化。

世界上第一台电子计算机（ENIAC）用了18000只电子管，70000只电阻，10000只电容，6000个开关，耗电140瓩，重达130吨，体积是 $30 \times 3 \times 1$ 米³。它的功能十分简陋，没有真正称得上内存贮器的东西，它只能存储和处理20个数字，每个数字都放在一个寄存器中。但寄存器的体积太，造价高，这就大大限制了ENIAC的功能。由于计算的中间结果没有地方放，只好写在卡片上，必要时再次输入。这样就消费了大量的时间，例如有人用ENIAC算一个题花了半小时，而20分钟用笔算完，真正算题才2分钟。

ENIAC没有真正称得上运算的东西，每次算题，必须由人将大量运算部件像积木一样搭配成各种解题的布局，每换一题，这种工作就得重复一遍，很费时间。

ENIAC的运算速度是每秒5000次加法，500次乘法，50次除法。

现在的计算机运算速度已经高达每秒1亿5千万次，正向着十亿次迈进。存储容量也相当大，整张报纸的全部信息，可以存储在计算机中四分之一张邮票大的位置上，一台机器存储几亿份资料，也不困难。

第一代美国机器IBM-701型需占几间房子，售价100万美元，而现在与它功能相当的微型计算机只要30—40美元，大小可以放在手掌上。甚至可以把整套计算机的运

辑部件做在一块硅片上，即所谓单片计算机。

与计算机硬件飞速发展的同时，计算机软件也有很大的发展，并经历了三个阶段，即50年代，60年代，和70年代。

第一阶段：（50年代），直接使用机器语言，即用机器指令系统来编程序——手编程序。

第二阶段：（60年代），出现了接近数学语言的所谓高级语言，如ALGOL，FORTRAN等，这些语言比较易学易用，一般人很快就能掌握。

在60年代，软件的理论工作取得了重大的进展，绝大多数的概念都已建立起来了。这些工作包括形式语言和自动控制理论的许多基本成果及其在分析和编译上的应用；发展运算语义和数学语义的理论，语言定义技术，关于编译进程与执行进程模式化的几种结构架子；提出了程序正确性和程序检验的基本思想等。

第三阶段：在语言方面，一方面出现了一些针对性较强的专用语言，另一方面，企图建立通用性较强的语言，使其能适用于科学计算、数据处理和信息加工等多种用途的所谓汇集性语言。此外，在管理机器资源的操作系统和编译技术方面都有一些进展。

计算机软件是一个复杂的系统，研制一种软件所占用的人力和时间也是可观的。如国产DJS—21的汇编语言是一个低级语言，它的汇编程序需要一个人年（即一个人工作一年）才能完成，它的BASIC语言的编译程序，需要3个人年，更复杂的系统，则需要占用更多的人年。国外一般是45—105个人年。搞一个软件，软件的投资也是很大的。如美国IBM—360操作系统的投资是2亿5千万美元。

在我国，1958年第一台电子管计算机投入运行。1964年，研制成功第二代晶体管计算机。1971年，又生产出第一台集成电路计算机。

与此同时，台式、袖珍式电子计算机也有迅速的发展。现在，我国计算机的研究和生产正朝着系列化的方向前进。

四、电子计算机的发展趋向：

电子计算机目前正向着巨型、微型，计算机网络和智能模拟的方向发展。

1. 提高运算速度：即提高晶体管、磁心的开关性能，研制新的快速电路和提高整机结构。

随着科学技术的发展，对计算机的速度和容量提出了越来越高的要求。如“四色定理”在计算机的帮助下，已经得到了证明，这个定理的证明进行了上百亿次的逻辑判定，所需计算时间多达1200小时，在解决这样的问题上，计算机还显得“慢”。就用一个普通的 n 阶行列式求值问题来说，一个 n 阶行列式展开有 $n!$ 项，每项是 n 个数目相乘，计算它的值就得做 $(n-1) \times n!$ 次乘法，计算一个30阶行列式就要做不少于 2×10^{10} 次乘法。如果在一台每秒一百万次的计算机上算，大约要用3个小时。在高能物理实验中，一次收集的数据多达七千盘磁带，（在DJS—6机上，一盘带存60多万全字长48位二进制的信息）。这样浩大的数据量，等待着目前的计算机去处理，又要白白浪费掉科学家多少宝贵的时间，因此要求计算机向超高速、巨型的方向发展。目前每秒一亿五千万次的计算机已经问世，每秒十亿次、百亿次的计算机正在研制之中。

2. 向微型化发展

由于巨型机虽然速度快，容量大，功能完善，但相应地带来了造价高昂，设备庞大、操作使用和维修十分复杂，不便于普及和推广使用等缺点。因此，具有性能好，造价低，体积小，用途广，方便灵活，稳定可靠等优点的微型计算机就产生并被大力发展，微型机从功能方面来讲是通用机，从应用方面来讲则是专用机，它既可作为专用系统的中央处理机，又可作为分时系统的终端机和计算机网络的接口处理机，它又能用于过程控制，又能进行科学计算和数据处理，所以，它正日益受到人们的密切注意。1979年美国的微型机生产翻了一番，现在几乎每家半导体厂商都在生产十种到几十种微型机。像美国、日本这样计算机数量比较多的国家里，大型机的数量只占计算机总数的5%以下，而微型小型机却占了计算机总数的一半以上。

3. 计算机网络化

把高速运行的计算机与缩短了时间和空间距离的信息传输线结合起来，便形成了一个高效率的计算机网络。参加计算机网络的计算机和若干终端设备可以分散在各地。一个网络至少有计算机、接口设备、数据传输线和终端设备所组成。它能高效率地解决生产活动和科学技术活动中产生的大量数据和信息需要传输、交换和处理的要求。这样就便于利用计算中心，迅速而又方便地解决各个用户（包括小单位的工业、企业、机关、学校等）的科学计算和数据处理，每个用户即一个终端。一台计算机可接几十个甚至几千个终端。例如美国国立医学图书馆1975年采用了IBM370/158型计算机多机系统，联接有500个终端，1976年终端设备数目增加至1000个。

用户只要学会简单指令，就可像打电话一样方便地使用终端设备，通过通信电路直接与远距离的中央计算机对话，迅速进行计算或情报检索。

4. 进行“智能模拟”

智能模拟是探索和模拟人的感觉和思维过程规律的科学，这是一门在控制论、计算机科学、仿生学、心理学的基础上发展起来的一门边缘学科。其主要内容包括：感觉和思维模型的建立，用机器进行图象和物体识别，用计算机求解问题、证明定理、研究学习、探索、联想、启发等活动的过程和机制，理解人的语言等等。

机器人的研制就是智能模拟的一个应用方面。目前，世界上已有两千多个不同水平的机器人，在各行各业中大显神通。另外，还可以代替人进行危险、费力的操作。例如，1966年，美国一架军用飞机失事，一颗氢弹掉进了地中海，最后终于依靠一个机器人下海把氢弹打捞了起来。

回顾30多年来的历程，计算机已经获得了辉煌的成果，从天至宇宙空间的探索到小至“基本粒子”的研究，从国防尖端到家庭事务，从文化教育到医药卫生，无论是工业，还是农业，各行各业，各个部门，都需要使用计算机。计算机是一门新兴的技术科学，具有非常广泛的发展前途。

第二章 计算机中数的表示方法

一、二进制的引进

我们通常用到的是十进制数，其数值部分用十个不同的符号0，1，2，3，4，5，6，7，8，9来表示。这些数字符号称为数码。数码处于不同的位置代表的意义是不同的。例如3152.76可写成

$$3152.76 = 3 \times 10^3 + 1 \times 10^2 + 5 \times 10^1 + 2 \times 10^0 + 7 \times 10^{-1} + 6 \times 10^{-2}$$

一般地任意一个十进制数S，可表示为

$$S = k_n (10)^n + k_{n-1} (10)^{n-1} + \dots + k_1 (10)^1 + k_0 (10)^0 + k_{-1} (10)^{-1} + \dots + k_{-m} (10)^{-m} = \sum_{j=n}^{-m} k_j (10)^j \quad (1)$$

其中 k_j 是0，1，2，3，4，5，6，7，8，9中任一个，由S决定
 m, n 为正整数，括号内的10称为计数制的基数。

二进制数的每个数位只可能取两个不同的数码“0”和“1”，在十进制中是“逢十进一”，这是大家都熟悉的。在二进制中是“逢二进一”的。二进制是电子计算机中广泛采用的进位制。

现就几个简单的数字列出十进制与二进制的对照表（见下页）。

十进制数的表达式（1）也适用于二进制数的表达形式，任意一个二进制数S都可以表示成：

$$S = k_n 2^n + k_{n-1} 2^{n-1} + \dots + k_0 2^0 + k_{-1} 2^{-1} + \dots + k_{-m} 2^{-m} = \sum_{i=n}^{-m} k_i 2^i \quad (2)$$

其中 k_i 只能取0或1，它由S决定， m, n 为正整数。

比如 1101011.1101可写成

$$1101011.1101 = 1 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 1 \times 2^{-2} + 0 \times 2^{-3} + 1 \times 2^{-4}$$

对于J进制的数S可表示为

$$S = \sum_{i=n}^{-m} k_i J^i \quad (3)$$

其中 k_i 可以是0，1，2，...，(J-1)中的任一个数码， m, n 为正整数。

若J=10，则（3）式为十进制数的表示形式，J=8，则（3）式为八进制数的表示形式。J=2，则（3）式为二进制数的表示形式。

	十进制	二 进 制			十进制	二 进 制	
整 数	0	0			0.5	0.1	2^{-1}
	1	1	2^0		0.25	0.01	2^{-2}
	2	10	2^1		0.125	0.001	2^{-3}
	3	11			0.0625	0.0001	2^{-4}
	4	100	2^2				
	5	101					
	6	110					
	7	111					
	8	1000	2^3				
	9	1001					
	10	1010					
	11	1011					
	12	1100					
	13	1101					
	14	1110					
	15	1111					
	16	10000	2^4				
	32	100000	2^5				
	128	1000000	2^6				

二、二进制的特点

为什么二进制在计算机中被广泛应用呢？下面我们来看一看二进制的特点，就得到了回答。

1. 二进制只有两个数码 0 和 1。用物理状态表示只需二种状态就可表示，而十进制就需用十种物理状态才能表示。制造二个稳定状态的元件比制造多稳定状态的元件容易得多，我们可以找到很多这样的元件具有这种特性，例如电位的高和低，脉冲的有和无，晶体管的截止和导通等等，只要规定其中的一种状态表示“1”，另一种状态表示“0”，这样就可以表示二进制。

由此可见在计算机中，使用二进制简单、可靠、易于实现。

2. 二进制的数运算简单（二进制数的四则运算）

进行简单的算术运算时，必需记住两个整数的和及乘积的表。对于J进制数要记住 $\frac{J(J+1)}{2}$ 个和与积。对于十进制，需要记住 $\frac{10(10+1)}{2} = 55$ 个和与积（见表）。用

这样的算法设计机器的运算器就很庞大，控制线路也很复杂。

用二进制，则 $J = 2$ ， $\frac{J(J+1)}{2} = \frac{2(2+1)}{2} = 3$

这样的和数表与积数表就特别简单。

加法公式：0 + 0 = 0

1 + 0 = 0 + 1 = 1
1 + 1 = 10

减法公式：

0 - 0 = 0 1 - 0 = 1
10 - 1 = 1 1 - 1 = 0

向高位借1，“借一当二”

乘积公式：0 × 0 = 0

1 × 0 = 0 × 1 = 0
1 × 1 = 1

除法：乘法与减法的混合应用。

下面举例说明二进制正整数的四则运算规则：

111
+ 011

1010

1101
- 110

111

10111
× 101

10111
00000
10111

1110011

11001 — 商
111 / 10110001

111

1000

111

1001

111

10 — 余数

3. 采用二进制，就可以使用逻辑代数（或称开关代数），这就为计算机的逻辑设计提供了便利的工具。

三、八进制数的表示

在J进制数的表示中，J=8即为八进制

八进制有八个不同的数码0, 1, 2, 3, 4, 5, 6, 7，而且是“逢八进一”，由于 $2^3 = 8$ ，三位二进制表示一位八进制，把二进制的三位用一个对应的八进制数码代替，即成八进制表示。

如 6 7 7 2 1 001 101 011 110 111 010 010
 | | | | | | | | | | |
110 111 111 010 001 1 5 3 6 7 2 2

八进制与二进制转换简单，改写即可。

八进制与二进制没有本质上的区别，把二进制用八进制表示书写简单。编写程序一般用八进制书写，指令的书写与穿孔输入一般均用八进制（目前也有用十六进制的）

四、十进制与二进制数之间的转换

1. 十进制数转换为二进制数

A、整数

设给定一个十进制整数725，把它转换为二进制整数。

根据数的表示法，并按数展开则有：

$$(725)_+ = (k_n k_{n-1} \dots k_1 k_0)_- = k_n 2^n + k_{n-1} 2^{n-1} + \dots + k_1 2^1 + k_0 2^0$$

$$= 2 (k_n 2^{n-1} + k_{n-1} 2^{n-2} + \dots + k_1)$$

两边除2，则商为 $k_n 2^{n-1} + k_{n-1} 2^{n-2} + \dots + k_1$

由于等式两边整数与小数对应相等，

$$\therefore \frac{k_0}{2} = \frac{1}{2}, k_0 = 1$$

对商数再除2，则新商（第二次商）为 $k_n 2^{n-2} + k_{n-1} 2^{n-3} + \dots + k_1$ ，新余数（第二次余数）为 k_1 ($k_1 = 0$)

用此方法继续做下去即求得 $k_2, k_3, \dots, k_n$ 。

2 725		余数 = 1 = k_0
2 362		余数 = 0 = k_1
2 181		余数 = 1 = k_2
2 90		余数 = 0 = k_3
2 45		余数 = 1 = k_4
2 22		余数 = 0 = k_5
2 11		余数 = 1 = k_6
2 5		余数 = 1 = k_7
2 2		余数 = 0 = k_8
2 1		余数 = 1 = k_9
0		

$$\therefore (725)_+ = (k_9 k_8 \dots k_1 k_0)_- = (1011010101)_-$$

B、小数

设给定十进制小数0.6875，把它转换成二进制小数。

根据数的表示法，并按数展开则有：

$$(0.6875)_+ = (0.k_{-1} k_{-2} \dots k_{-n})_- = k_{-1} 2^{-1} + k_{-2} 2^{-2} + \dots + k_{-n} 2^{-n}$$

两边乘以2，得 $2(0.6875)_+ = (k_{-1} + k_{-2} 2^{-1} + \dots + k_{-n} 2^{-n+1})_-$

$$1.3750 = k_{-1} + k_{-2} 2^{-1} + \dots + k_{-n} 2^{-n+1}$$

由于等式两边整数与小数对应相等，

$$\therefore k_{-1} = 1$$

$$0.3750 = k_{-2} 2^{-1} + \dots + k_{-n} 2^{-n+1}$$

等式两边再乘以2，得

$$0.750 = k_{-2} + k_{-3} 2^{-1} + \dots + k_{-n} 2^{-n+2}$$

$$\therefore k_{-2} = 0$$

$$0.750 = k_{-3} 2^{-1} + \dots + k_{-n} 2^{-n+2}$$

如此继续下去, 可得 $k_{-3}, k_{-4}, \dots, k_{-n}$

转换的结果为:

$$(0.6875)_{10} = (0.k_{-1}k_{-2}k_{-3}k_{-4})_2 = (0.1011)_2$$

$$\begin{array}{r} 0.6875 \\ \times 2 \\ \hline 1.3750 \end{array}$$

$$\text{整数部分} = 1 = k_{-1}$$

$$\begin{array}{r} 0.3750 \\ \times 2 \\ \hline 0.7500 \\ \times 2 \\ \hline 1.5000 \end{array}$$

$$\text{整数部分} = 0 = k_{-2}$$

$$\text{整数部分} = 1 = k_{-3}$$

$$\begin{array}{r} 0.5000 \\ \times 2 \\ \hline 1.0000 \end{array}$$

$$\text{整数部分} = 1 = k_{-4}$$

对于十 $\rightarrow$ 二转换的方法, 综述如下:

根据“任意两数相等, 则整数部分与小数部分(分别)对应相等”的原理。当进行整数转换时, 我们用除以 2 的办法, 将整数的最低位孤立出来(每次除 2 后的余数)即得 $k, k_1, \dots, k_n$; 当进行小数转换时, 用乘 2 的办法, 将小数的最高位孤立出来(每次乘 2 后的整数部分)即得 $k_{-1}, k_{-2}, \dots, k_{-n}$ 。

2. 二进制数转换成十进制数

方法很简单, 只需要将二进制数按权相加即可。下面举例说明。

$$\begin{aligned} (101011.1001)_2 &= 2^5 + 0 \times 2^4 + 2^3 + 0 \times 2^2 + 2^1 + 2^0 + 2^{-1} + 0 \times 2^{-2} \\ &\quad + 0 \times 2^{-3} + 2^{-4} \\ &= 32 + 8 + 2 + 1 + 0.5 + 0.0625 \\ &= (43.5625)_{10} \end{aligned}$$

十进制数与八进制数的转换(十 $\rightarrow$ 八)方法同(十 $\rightarrow$ 二), 整数部分除 8, 小数部分乘 8。(八 $\rightarrow$ 十), 按权相加。

五、数的定点与浮点表示

一个十进制数可表示为

$$\text{如 } 67.819 = 10^2 \times 0.67819$$

$$0.000653 = 10^{-3} \times 0.653$$

二进制数也可这样表示: $N = 2^j \times S$

j : 二进制整数, 表示数的整数位数(阶码)

S : 二进制小数, N 的全部有效数字(尾数)

$$\text{如 } 10011.101 = 2^5 \times 0.10011101 = 2^{101} \times 0.10011101$$

一般, 底可以取值 2^k , 则 $N = (2^k)^j \times S$

若 $k=1$, 则底为 $2^1=2$, 若 $k=2$, 则底为 $2^2=4$,

若 $k=3$ 则底为 $2^3=8$ 若 $k=4$ 则底为 $2^4=16$

对于一种机器, k 是固定的。

在 $N = (2^k)^j \times S$ 表达式中,

(1) 若阶码 j 固定 (对任何数) 即小数点位置固定, 则 N 为定点数

(2) 若阶码 j 不固定 (对不同数取不同的值) 则称为浮点数

在定点数中, 若 $j=0$, 只能表示小数

若 $j=n$ } 表示整数
S为 n 位 }

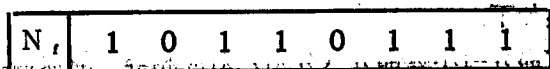
定点机: 计算机中数采用定点表示, 小数点固定

浮点机: 计算机中数采用浮点表示, 小数点浮动

定一浮两用机: 定点、浮点同时采用

下面举例说明数在机器中的表示形式:

(1) 定点数: $N = 0.10110111$

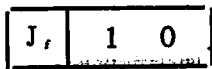


符号

数值部分

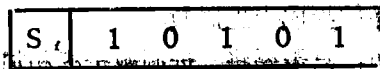
小数点固定在小数最高位和符号位 N_1 之间

(2) 浮点数: $N = 2^{10} \times 0.10101$



阶符

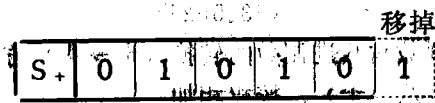
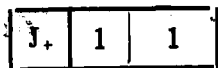
阶码



尾符

尾数

尾数的小数点在尾数最高位的前面, 由于有了阶码, 小数点是可以移动的, 如向左移一位, $N = 2^{10} \times 0.10101$ 便变成 $N = 2^{11} \times 0.010101$ 1移掉



移掉

小数点左移一位, 相当于数码右移一位, 阶码加1。

在浮点计算机中, 通常都采用规格化数的表示法, 这样不丢失有效数字, 提高运算的精确度。

究竟什么是规格化的数呢? 对于二进制来说, 如果尾数的第一位 (即最高位) 是1时, 就算已规格化的数, 反之, 若尾数第一位是0, 就是非规格化的数。也就是说二进制浮点数的尾数 S 满足 $\frac{1}{2} \leq |S| < 1$ 时, 就是已规格化的数。对非规格化的浮点数, 只要移动尾数的小数点就可实现规格化。

“溢出”: 计算机中的数超出了它所能表示的范围, 称为“溢出”。

在定点机中: 初始数据和中间运算结果出现绝对值 ≥ 1 时, 就发生“溢出”, 停机处理。

在浮点机中: 对数进行规格化后发生两种情况:

A. 下溢、当阶码超出机器所能表示的最小数时，称为“下溢”，机器按“0”处理（称为机器零）。

B. “上溢”：阶码超出机器所能表示的最大数时，称为“上溢”，停机处理（中断处理）。

由此可见，在定点机中，只能表示小数，为了防止“溢出”，事先必须对数据按一定的比例进行缩小，这个比例因子的选择比较麻烦，在这点上浮点机比较方便。而数在浮点机中运算比较复杂（增加了对阶和阶码运算），所以浮点机的运算器和控制器比定点机复杂。

例如 $a = 2^{11} \times 0.1001$, $b = 2^{01} \times 0.1100$

$$\begin{aligned} a \pm b &= 2^{11} \times 0.1001 \pm 2^{01} \times 0.1100 = 2^{11} \times 0.1001 \pm 2^{11} \times 0.0011 \\ &= 2^{11} \times (0.1001 \pm 0.0011) = 2^{11} \times 0.11 \\ \text{差} &= 2^{11} \times 0.0110 = 2^{10} \times 0.11 \end{aligned}$$

对阶：让小阶向大阶看齐，阶小的尾数小数点左移，在机器中是尾数数码右移，右移一位，阶码加1，直至阶码相同为止。

$$\begin{aligned} a \times b &= (2^{11} \times 0.1001) \times (2^{01} \times 0.1100) = 2^{11+01} \times (0.1001 \times 0.1100) \\ a \div b &= (2^{11} \times 0.1001) \div (2^{01} \times 0.1100) = 2^{11-01} \times (0.1001 \div 0.1100) \end{aligned}$$

所以，两浮点数相乘（除），只要阶码相加（减），尾数相乘（除）。尾数的乘除同定点数。

六、机器数的原码、补码和反码表示法

1. 原码

将一个数（连同符号）在机器中的表示加以数值化，称为机器数。例如 0.1011，-0.1011 在机器中表示为 0.1011，1.1011，即正数符号位用“0”表示、负数符号位用“1”表示。这种表示法称为原码，记作 $[0.1011]_{\text{原}} = 0.1011$ 及 $[-0.1011]_{\text{原}} = 1.1011$

$$\text{一般定义: } [X]_{\text{原}} = \begin{cases} X & 1 > X \geq 0 \\ 1 - X & 0 \geq X > -1 \end{cases}$$

式中 $[X]_{\text{原}}$ 就是机器数，X 为真值

2. 补码

在做加法运算时，为了判断是同号还是异号，比较哪个数绝对值大，才能确定进行相加还是相减。究竟谁减谁，就要增加机器设备或增加机器的运算时间，为此人们设法将负数转化为正数，将减法转化为加法，这就是采用补码表示法

$$\begin{aligned} [0.1011]_{\text{补}} &= 0.1011 \\ [-0.1011]_{\text{补}} &= 1.0101 \end{aligned}$$