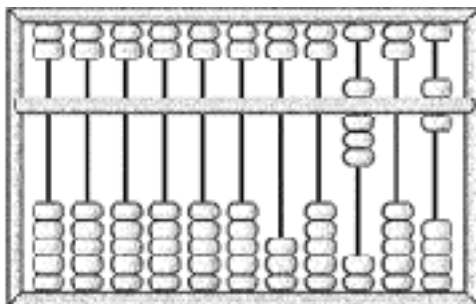


第18章 从算盘到芯片

纵观历史，人类发明了很多灵巧的工具和机器以满足广泛的需求，从而使数学运算变得更容易了些。虽然人类天生就有使用数字的能力，但仍能经常需要帮助。人们常遇到一些自己不能轻易解决的问题。

数字可看成是早期帮助人类记录商品和财富的工具。许多文明，包括古希腊和美洲土著，都借用石子或谷物来计数。在欧洲使用计数板，而在中国则对由框和珠子组成的算盘较为熟悉：



没有人真的喜欢乘法和除法，但却有人为它做过什么，苏格兰数学家 John Napier(1550-1617)就是这少数人中的一个。他发明了对数来简化这些操作，两数之积简化为它们对数的和。因此，如果你想使两数相乘，先在对数表中分别查出它们的值，然后相加，再用相反的方法查对数表就可得到它们的积。

对数表的建立，使得随后 400 年里一些最伟大的思想家一直为此忙碌，而另一些人却在设计使用小装置来代替对数表。一种有对数标尺的滑尺已有很长的历史了，它由 Edmund Gunter (1581-1626) 发明并由 William Oughtred (1574-1660) 修正。1976 年，当 Keuffel & Esser 公司将其公司最后制造的滑尺捐赠给华盛顿特区的 Smithsonian 学院时，滑尺的历史也就宣告结束了，其中的原因是手持计算器的出现。

Napier 也发明了一种乘法辅助器，它由刻在骨头、号角、象牙上的数字条组成，因而这样的辅助器称为 Napier 骨架。1620 年左右，Wilhelm Schickard (1592-1635) 制造出了最早的有点儿自动功能的由 Napier 骨架组成的机械计算器。几乎在同时出现了由互相连结的轮子、齿轮和水平仪组成的另外一种计算器，这种机械计算器的两个最主要的制造者是数学家和哲学家布莱兹·帕斯卡 (1623-1662) 和莱布尼兹 (1646-1716)。

你一定能记得最初的 8 位加法器和能自动进行多于 8 位数的加法计算的计算机中的进位是多么令人讨厌。进位原先似乎只是加法运算中的一个小问题，但在加法机中却成了一个中心问题。即使设计一个能进行除进位外的所有工作的加法机，也不能说工作就算完成了。

进位处理是否成功是评估老式计算机的关键。例如，帕斯卡设计的进位机制禁止减法运算。为了进行减法，必须加上 9 的补码，这在第 13 章中已经讲到。直到 19 世纪后期，才出现了真正可以为人们所使用的机械计算器。

一个奇特的发明对计算的历史产生了深远的影响，就像它对纺织所产生的深远影响一样，这就是约瑟夫·玛丽·杰奎德 (1752-1834) 所发明的自动织布机。杰奎德织布机（大约产生于 1801 年）使用上面已打孔的金属卡片（就像钢琴上的金属卡片）来控制编织物的图案。杰奎德的一大杰作就是用黑白丝线织成的自画像，为此使用了大约 1 万张卡片。

在 18 世纪（甚至直到 20 世纪 40 年代），计算机就像一个以计算数字谋生的人。使用星星进行航海导航经常需要对数表，并且三角函数表也是必需的。如果需要发布新表，则需要许多计算机来工作，然后把结果汇总起来。当然，在这一过程的任何阶段，即从初始化计算到设置类型来打印最后几页都可能会出现错误。

从数学表中消除错误的愿望激发了查尔斯·巴贝芝 (1791—1871)。巴贝芝是一位英国的数学家和经济学家，他和摩尔斯差不多是同一时代的人。

在那时，数学表（以对数表为例）并不是通过计算表中每一项确切的对数值而建立的，因为这得花费很多时间。取而代之的是选择一些数进行对数计算，而介于这些数中间的那些数则采用插补，即称作差分的方法，通过相对简单的计算来得到。

大约在 1820 年，巴贝芝认为可以设计并制造一台机器来自动建立表，甚至可以到自动设置打印类型这一步，这样可以消除错误。他构想差了分机，这是一个很大的机械加法机。通过切换，可使位于 10 个不同位置的轮子来表示各位数的十进制数字，负数用 10 的补码来计算。尽管一些早期的模型可以证明巴贝芝的设计是可行的，并且也从英国政府获得了一些支持，但差分机却从未完成过。巴贝芝于 1833 年放弃了这一工作。



然而，就在那个时候，巴贝芝又有了一个更好的构想，这就是解析机（重复的设计和再设计不断耗费着巴贝芝的生命，直到他死去），解析机是 19 世纪最接近计算机的发明。在巴贝芝的设计中，有一个存储系统（类似于今天存储器的概念）和运算器（算术单元）。乘法由重复加法来实现，除法由重复减法来实现。

解析机最精华的部分在于它可以用卡片来编程，这些卡片是由杰奎德的按图案编织的织布机上的卡片经过改造而制成的。正如艾达·奥古斯塔，即拉弗雷斯女伯爵 (1815-1852) 在她翻译的由一个意大利数学家写的，关于巴贝芝解析机的文章的按语里写的：“我们可以说解析机编织的是代数模型，正如杰奎德织布机编织的是花和叶一样”。

巴贝芝可能是第一个意识到计算机中条件转移的重要性的人。拉弗雷斯女伯爵关于此也曾写道：“操作循环必须理解成一批操作，这些操作可以重复多次。无论是只重复两次还是无穷次，都是一个循环，归根到底重复组成了循环的这些操作。许多情况下，还会出现一个或多个循环的重复，即循环的循环或多个循环的循环”。

尽管差分机最终由 Georg 和 Edvard Scheutz 父子在 1853 年制成，但巴贝芝的机器那时已被遗忘了很久，直到 20 世纪 30 年代人们开始追寻 20 世纪计算机的根源时才再次想起。巴贝芝曾经做的东西已经被后来的技术所超越，除了他对自动化的超前认识外，他并没有为 20 世纪的计算机工程留下什么东西。

计算机历史上另一个里程碑来源于美国宪法第二部分的第一篇。这一部分里除其他事情外还要求每 10 年进行一次人口普查。1880 年人口普查的时候，人口信息按年龄、性别及祖籍

来收集，数据的收集差不多花了七年的时间来进行。

由于担心1890年的人口普查可能需要超过10年的时间，人口普查局寻求使该系统工作自动化的可能性并选用了赫曼·霍勒瑞斯(1860—1929)开发的机器，此人是1880年人口普查的统计员。

霍勒瑞斯的想法是采用马尼拉穿孔卡片，大小是 $6\frac{5}{8} \times 3\frac{1}{4}$ 。(虽然霍勒瑞斯不可能知道巴贝芝如何使用卡片在他的解析机上编程，但他却很熟悉杰奎德织布机上卡片的使用。)卡片上的孔组成24列，每列12个，这样共有288个位置，这些位置表示某个人在人口普查记录中的某些特征。普查员通过在卡片的适当位置上打 $1/4$ 英寸的方孔来标识这些特征。

本书可能使得人们习惯于用二进制码的概念来思考问题，因此，你可能马上会想到卡片上的288个穿孔点可以存储288位信息。但是，这些卡片并不是这样用的。



例如，在纯二进制系统中，人口普查卡片会有一个位置来表示性别，可以用打孔表示男性，未打孔表示女性(或者相反)。但是，霍勒瑞斯的卡片用两个位置表示性别，一个位置打孔表示男性，另一个位置打孔表示女性。同样，用两个穿孔表示年龄，一个穿孔指明一个5年的年龄范围：0~4、5~9、10~14等等，另一个孔用5个位置中的一个来表明在该范围内的确切年龄。年龄编码需要卡片上总共28个位置。而纯二进制系统只需要7个位置就可编码0~127的任何年龄。

我们应该原谅霍勒瑞斯在记录人口普查信息时没有采用二进制系统，对于1890年的人口普查员来说，把年龄转换成二进制数要求太高了一些。还有一个实际原因来解释穿孔卡片系统为什么不能全部是二进制的是因为二进制系统可能出现这样一种情形，即所有的孔都被打孔，使得卡片很脆弱，结构不牢固。

人口普查的数据收集可进行统计或制成表格。你可能想知道每一个区域内有多少人生活，当然，你也可能对人口的年龄分布统计信息感兴趣。正因为如此，霍勒瑞斯制造了一个制表机，它能结合手工操作和自动操作。操作员把一个有288个弹簧针的板子压到每一个卡片上，对应于卡片上每一个穿孔的针接触水银池形成电路，电路触发电磁铁使十进制计数器计数。

霍勒瑞斯在分类卡片的机器上也用了电磁铁。例如，如果需要统计所记录的每一个职业的年龄资料，首先需要按职业对卡片分类，然后对每一个职业统计年龄资料。分类机与制表机一样用手压，但分类机使用电磁铁打开一个开口，对应于26个分隔区域中的一个，操作员把卡片放入分隔区域，然后用手关上开口。

这项实验在自动进行1890年的人口普查工作中取得了巨大成功，处理了超过6200万张的卡片，包含的数据是1880年人口普查的2倍，而数据处理只花了大约1880年人口普查所花时间的1/3。霍勒瑞斯和他的发明享誉全球。1895年，他甚至到了莫斯科并成功地卖出了他的设备，该设备在1897年第一次用于俄罗斯的人口普查。

霍勒瑞斯开始进行各种活动。1896年，他创立了制表机公司，出租和出售穿孔卡片设备。1911年，经过合并，该公司成为计算-制表-记录(computing-Tabulating-Recording)公司，即C-T-R公司。到1915年，C-T-R的主席是Thomas J. Watson(1874-1956)，他在1924年把公司的名字改为国际商用机器公司，即IBM。

1928年,原先的1890年的人口普查卡片已经演化成为著名的“不会卷曲、折叠、翘页”的IBM卡片,有80列12行。它们用了50多年,即使在最后几年也被称作霍勒瑞斯卡片。在第20、21和24章将要讲到这些卡片的影响。

在把目光移到20世纪之前,不要对19世纪那个年代有太多的偏见。显然,本书主要着眼于数字系统的发明,这些发明包括电报、布莱叶盲文、巴贝芝机器和霍勒瑞斯卡片。当与数字概念和数字设备一起工作时,很容易会把整个世界都想像成数字世界。但是,19世纪的特征更多体现在那些不是数字的发明及发现上。的确,我们感受到的自然世界只有很小一部分是数字的,它更接近于是连续的而不那么容易被量化。

尽管霍勒瑞斯在他的卡片制表机和卡片分类机上用了继电器,但是直到20世纪30年代中期,人们才真正开始用继电器来制造计算机(后来叫机电式计算机)。这些机器上用的继电器通常不是电报继电器,而是那些用来在电话系统中控制呼叫路由的继电器。

早期的继电器计算机并不像我们在上一章中制造的继电器计算机(将会看到,后者计算机的设计基础是基于从20世纪70年代开始的微处理器)。特别地,今天对我们来说计算机内采用二进制数是显然的,但那时并不是这样。

我们所设计的继电器计算机与早期的继电器计算机之间的另一个区别是在20世纪30年代,没有人会狂热到用继电器构造524 288位的存储器!所需的造价、空间及功耗使得这样大的存储器不可能实现。可用的很小的存储器只是用来存储中间结果,而程序本身存储在像带有穿孔的纸带这样的物理媒体上。的确,把代码和数据放入存储器的处理方式是一个很现代化的概念。

按年代排列,第一个继电器计算机似乎是由Conrad Zuse(1910-1995)建造的。1935年当他还是一个工程系的学生的时候,他就开始在他父母位于柏林的住所里制造计算机。他采用的是二进制数,但却是早期版本,且用的是机械存储器而不是继电器。Zuse在老式的35毫米的电影胶片上穿孔来进行计算机编程。

1937年,贝尔电话实验室的George Stibitz(1904-1995)把一对电话继电器安装在家里,并且又连接了一个1位加法器到餐桌上,后来他的夫人称它为“K机器”(K表示kitchen,厨房)。该实验导致在1939年产生了贝尔实验室的复数计算机。

与此同时,哈佛大学的研究生Howard Aiken(1900-1973)因需要某种方法来做大量重复计算,从而使得哈佛大学和IBM合作,制造出了最终称为Harvard Mark I的自动顺序控制计算机(ASCC: automated sequence controlled calculator),此项工作在1943年完成。这是第一台打制表格的数字计算机,它终于实现了巴贝芝的梦想。Mark II是以巨大的继电器为基础的机器,使用了13 000个继电器。由Aiken领导的哈佛计算实验室讲授了计算机科学的第1课。

继电器并不是制造计算机的最好器件,因为它是机械的,工作时需弯曲一个金属簧片,如果超负荷工作,簧片就会折断;如果有一小片污垢或纸片粘在触点之间,继电器就会失效。一个著名的事件发生在1947年,从Harvard Mark II计算机的一个继电器中找到一只蛾子。Grace Murry Hopper(1906—1992)1944年加入Aiken的小组,此人后来在计算机程序设计语言领域非常有名。他在计算机日志中记录了这只蛾子,写道“第一次发现了真正的bug”。

继电器的一种可能的替代品是真空管,真空管由John Ambrose Fleming(1849—1945)和Lee de Forest(1873—1961)发明用来同无线电设备连接。到20世纪40年代,真空管早已用来放大电话信号。事实上,每一家的落地式收音机都装上了用来放大无线电信号的真空管,

以便人们能听见。真空管可以连接成与、或、与非和非门，这一点非常像继电器。

逻辑门是由继电器还是由真空管来制造的并不重要。利用逻辑门可集成加法器、选择器、译码器、触发器和计数器。前面几章讲的基于继电器的器件在当继电器被换成真空管时仍然可用。

不过真空管也有问题，它们昂贵、耗电量大、散发的热量多。然而最大的问题在于它们最终会被烧毁，这也就是它们的寿命问题。有真空管收音机的人就习惯于隔一段时间更换这些管子。电话系统设计成有许多多余的管子，因此损失点儿管子也不是大的问题。（没有人能指望电话系统不出一点儿问题。）然而计算机中的一个管子烧毁以后，并不能很快被检测到，而且，计算机中使用了如此多的真空管，可能每几分钟就会烧毁一个。

使用真空管相对于继电器的最大好处在于它每百万分之一秒（即1微秒）就可以跳变一次。真空管改变状态（开关闭合或断开）的速度比继电器快1000倍，在最好的情况下，继电器状态的变化大约需1毫秒，即千分之一秒。有趣的是，在早期计算机的研究中，速度问题并不是最重要的，这是因为早期计算机总的计算速度与机器从纸带或电影胶片中读取程序的速度密切相关。正是因为计算机是基于这种方式制造的，真空管比继电器速度快多少也就无关紧要了。

在20世纪40年代初，真空管开始在新的计算机中替换继电器。直到1945年，晶体管制成。正如继电器机器称为机电式计算机，真空管则是第一台电子计算机的基础。

在英国，Colossus计算机（1943年开始使用）用于破译德国的“Enigma”代码生成器生成的密码。为这个项目（和英国以后的一些计算机项目）做出贡献的人是艾伦·M·图灵（1912—1954），他当时由于写了两篇很有影响的论文而闻名于世。第一篇论文发表于1937年，其中首先提出了“计算能力”的概念，用以分析计算机可以做到和不能做到的事。他构思出了现在称为图灵机的计算机抽象模型。图灵写的第二篇著名论文的主题是人工智能，他介绍了一个测试机器智能的方法，现在称作图灵测试法。

在摩尔电气工程学校，J.Presper Eckert(1919-1995)和John Mauchly(1907—1980)设计了ENIAC（electronic numerical integrator and computer，电子数字积分器和计算机）。它采用了18 000个真空管，于1945年末完成。纯粹按吨位（大约30吨）计算，ENIAC是曾经制造出来的（也许以后也是）最大的计算机。到1977年，你可以在电器行买到更快的计算机。然而，Eckert和Mauchly的专利却被John V. Atanasoff(1903—1995)给阻挠了。Atanasoff在早期曾设计了一个电子计算机，但它从未很好地工作过。

ENIAC引起了数学家约翰·冯·诺依曼（1903—1957）的兴趣。从1930年开始，匈牙利出生的冯·诺依曼就一直住在美国。他是一个令人瞩目的人物，因能在脑子里构思复杂的算法而享有很高的声誉，他是普林斯顿高级研究院的一名数学教授，研究范围很广，从量子理论到游戏理论的应用再到经济学。

冯·诺依曼帮助设计了ENIAC的后继产品EDVAC（electronic discrete variable automatic computer）。特别是在1946年与Arthur W. Burks和Herman H. Goldstine合写的论文《Preliminary Discussion of the logical Design of an Electronic Computing instrument》中，他描述了有关计算机的几点功能，这些功能使得EDVAC比ENIAC更先进。EDVAC的设计者感觉



到在计算机内部应该使用二进制数，而 ENIAC用的是十进制数；计算机应该具有尽可能大的存储器，当程序执行时，这个存储器可用来存储程序代码和数据；（ENIAC中的情况不是这样，对ENIAC进行编程是通过断开开关和插上电缆来进行的。）指令应该在存储器中顺序存放并用程序计数器来寻址，但也应该允许条件转移。这种设计思想叫作存储程序概念。

这种设计思想是重要的革命化的一步，今天称为冯·诺依曼体系结构，上一章建造的计算机就是典型的冯·诺依曼机器。但冯·诺依曼体系结构也带来冯·诺依曼瓶颈，冯·诺依曼型机器需要花费大量的时间从存储器中取出指令来准备执行。应该还记得第17章最后设计的计算机取指令的时间占整个指令周期的3/4。

在EDVAC时代，用真空管构建存储器是不值得的，因而人们使用一些古怪的方法来解决这个问题。一个成功的方法就是水银延迟线存储器，它使用5英尺长的水银管子。在管子的一端，每隔1微秒向水银发一个小脉冲。这些脉冲需要1毫秒的时间到达管子的另一端（此时，它们像声波一样会被检测到，然后送回到开始的地方），因此每个水银管可存储大约1024位信息。

直到20世纪50年代中期，磁芯存储器才开发出来。这种存储器由大量的环绕着电线的电磁金属环组成，每个小环保存1位信息。在磁芯存储器被其他技术取代后的相当一段时期内，还经常听到老程序员把由处理器访问的存储器称作磁芯。

在20世纪40年代，冯·诺依曼并不是唯一一个对计算机的本质进行概念上思考的人。

克劳德·香农(1916年出生)是另外一个有着重大影响的思想家。第11章曾经提到他1938年的硕士论文，论文中确立了开关、继电器和布尔代数之间的关系。1948年，当他在贝尔电话实验室工作时，他在《Bell System Technical Journal》上发表了一篇题为《A Mathematical Theory of Communication》的论文，其中不仅引入了“位”的概念，而且确立了一个现代称为“信息理论”的研究领域。信息理论涉及在噪声（经常阻碍信息传送）存在的情况下传送数字信息以及如何进行信息补偿等问题。1949年，他写了第一篇关于编写让计算机下棋的程序的文章；1952年他设计了通过继电器控制的机械老鼠，这个老鼠可以在迷宫中记住路径。香农同时也因为他会骑独轮车，玩变戏法而在贝尔实验室很出名。

Norbert Wiener(1894-1964)18岁时就在哈佛大学取得了数学博士学位，因《Cybernetics, or Control and Communication in the Animal and Machine》(1948)一书而闻名于世。他首次使用控制论(Cybernetics)这个词来表示一种把人及动物的生物活动与计算机及机器人的机理联系起来的理论。在现代文化里，广泛使用cyber-前缀表示与计算机相关的东西。更特别的是，成千上万的计算机通过因特网进行的互连称作cyberspace(信息空间)，这个词来自科幻小说作家William Gibson 1984年的小说《Neuromancer》中的词cyberpunk。

1948年，Eckert-Mauchly计算机公司(Remington Rand公司的后继者)开始开发第一台商用计算机UNIVAC(universal automatic computer)，并于1951年完成。第一台被送往人口普查局。UNIVAC的首次网络应用是用于CBS，用来预测1952年的总统选举结果。Walter Cronkite称它为“电脑”。同样是在1952年，IBM发布了它的第一个商用计算机系统，即701。

从此，开始了社团和政府使用计算机的漫长历史。然而，之所以对这段历史感兴趣可能是因为我们要追踪另一段历史轨迹——即降低计算机造价和大小并且使它进入家庭的轨迹，它开始于1947年一场几乎不被人注意的电子技术突破。

贝尔电话实验室许多年里都是这样一个地方：聪明的人可以在此做他感兴趣的任何事。所幸的是，他们之中有人对计算机感兴趣，如已经提到的George Stibitz和Claude Shannon，

他们在贝尔实验室工作的时候都为早期的计算机作出了突出的贡献。后来，在 20 世纪 70 年代，贝尔实验室诞生了很有影响的操作系统 UNIX 和程序设计语言 C 语言，这些在随后的几章里将要讲到。

1925 年 1 月 1 日，美国电话电报公司正式把它的科学和技术研究部分与商业部分分离，另外建立附属机构，这样贝尔实验室诞生了。贝尔实验室的主要目的是为了研究能够提高电话系统性能的技术。幸运的是，它的要求非常含糊，可以包含所有的事情。但在电话系统中，一个明确的长期的目标是：在线路上传输的声音信号能不失真地放大。

从 1912 年起，贝尔电话系统就采用了真空管放大器，大量的研究和工程人员着手提高电话系统使用的真空管的性能。尽管这样，真空管仍然有许多问题。管子体积大，功耗大且最终会烧毁。不过，在当时却是唯一的选择。

1947 年 12 月 16 日，当贝尔实验室的两个物理学家 John Bardeen(1908—1991)和 Walter Brattain(1902—1987)在装配一个不同类型的放大器时，所有的一切都改变了。这种新型放大器由锗片——一种称作半导体的元素——和一条金箔构成。一个星期后，他们给他们的上司 William Shockley (1910—1989) 进行了演示。这就是第一个晶体管，一种被人们称为 20 世纪最伟大的发明的器件。

晶体管不是凭空产生的。8 年前，即 1939 年 12 月 29 日，Shockley 在笔记本写下：“今天我想用半导体而不是真空管做放大器在原理上是可能的。”第一个晶体管被发明以后，随后许多年它继续被完善。1956 年，Shockley、Bardeen 和 Brattain 获得诺贝尔物理学奖——“因为他们半导体上的研究并且发明了晶体管。”

本书的前面谈到了导体和绝缘体。之所以称作导体是因为它们非常容易导电，铜、银和金都是很好的导体。并非巧合，所有这三种元素都在元素周期表的同一列。

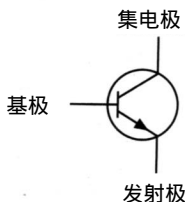
前面讲过，原子中的电子分布围绕在原子核的外层。上述三种导体的特征是只有一个单独的电子在最外层。这个电子很容易与原子的其余部分分离并自由移动形成电流。与导体相对的是绝缘体，像橡皮和塑料，它们几乎不能导电。

锗和硅元素（还有一些化合物）称为半导体，并不是因为它们的导电性是导体的一半，而是因为它们的导电性可以用多种方法来控制。半导体最外层有 4 个电子，是最外层所能容纳电子最大数目的一半。在纯半导体中，原子彼此非常稳固地结合在一起，具有与金刚石相似的晶状结构。这种半导体不是好的导体。

但是半导体可以掺杂，意思是与某种杂质相混合。半导体很容易与其他杂质结合而变得不纯。有一类杂质为原子的结合提供额外的电子，这种半导体叫 N 型半导体（N 表示 negative）；另一种类型的杂质掺杂生成 P 型半导体。

两个 N 型半导体中夹一个 P 型半导体可制成放大器，称作 NPN 晶体管，对应的三部分分别是集电极（Collector）、基极（Base）和发射极（Emitter）。

下面是一个 NPN 晶体管的示意图：



基极上的一个小电压能控制一个很大的电压经过集电极到发射极。若基极上没有电压，它有效截止晶体管。

晶体管通常被封装在一个直径大约为 1/4 英寸的小金属容器里，有三个引脚伸出：



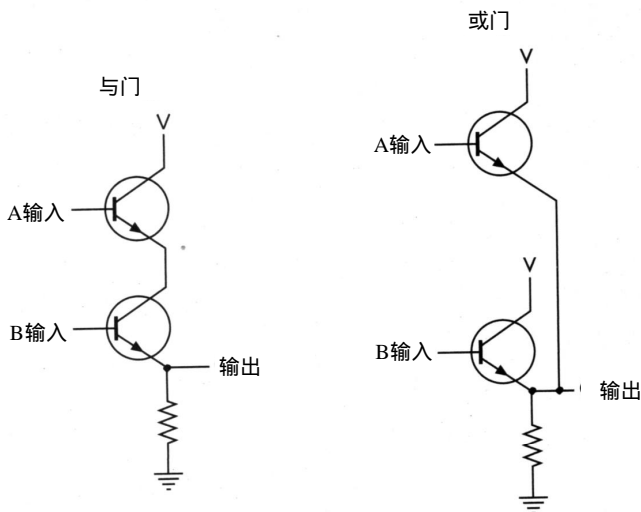
晶体管首创了固态电子器件，意思是晶体管不需要真空，可由固体做成，特别是指由半导体和当今最普遍的硅制成。除了比真空管体积更小外，晶体管功耗小，发热少，并且更耐用。携带一个电子管收音机很不方便。晶体管收音机靠小电池供电，并且不像电子管，它不会变热。对于 1954 年圣诞节早晨收到礼物的幸运者来说携带晶体管收音机已成为可能。那些第一批袖珍晶体管收音机是由德克萨斯仪器公司制造的，该公司是半导体革命中的一个重要公司。

然而，晶体管的第一个商业应用是助听器。为了纪念贝尔一生为聋人的贡献，AT&T 公司允许助听器生产厂家不用付任何专利权税就可使用晶体管技术。第一台晶体管电视机出现于 1960 年，而今天，电子管器件几乎已消失了。（然而，并未完全消失，一些高保真发烧友及电子吉他手还是喜欢用电子管放大器来放大声音而不是用晶体管。）

1956 年，Shockley 离开贝尔实验室成立 Shockley 半导体实验室。他迁到加利福尼亚的 Palo Alto——他的成长之地。他的公司是那里成立的第一家从事这种工作的公司。后来，其他半导体及计算机公司也在那儿建立了业务，于是旧金山南部的这个地方也就成为现在非正式地称作硅谷（Silicon Valley）的地方。

真空管原是为放大信号而开发的，但它们也用来作为逻辑门的开关。晶体管也是如此。下面你将会看到一个由晶体管组成的与门在构造上很像用继电器组成的与门。只有当 A 输入为 1 且 B 输入为 1 时，两个晶体管才都导通，输出才为 1。这里所示的电阻用来防止短路。

正如你所看到的，下图连接了两组晶体管，右边一个是或门，左边一个是与门。在与门电路里，上面管子的发射极连接到下面管子的集电极。在或门电路里，两个晶体管的集电极都连到电源，发射极连到一起：



因此，我们所学的有关由继电器构造逻辑门和其他部件的知识也适用于晶体管。继电器、电子管和晶体管最初主要用来放大信号，但也可用同样的方法连接成逻辑门来建造计算机。第一台晶体管计算机制造于1956年，短短几年内，新计算机的设计中就放弃使用电子管了。

这里有一个问题：晶体管当然使计算机更可靠，更小和更省电，但晶体管能使计算机的制造更简单吗？

并非如此。虽然晶体管使得在一个小空间里能安装更多的逻辑门，但你还得担心这些组件的互连。连接晶体管形成逻辑门像连接继电器和真空电子管一样困难，某种程度上，它甚至更困难，因为晶体管较小，不容易掌握。如果想用晶体管建造第17章所建造的计算机及64KB的RAM阵列，设计工作中的一部分将是设法发明某种能容纳所有组件的结构。这些劳动是乏味的，需要在数百万晶体管之间建立起数百万连接。

然而，就像我们所发现的，一些晶体管的连接会重复出现。许多对晶体管几乎都是先连接成门，门再连接成触发器、加法器、选择器或译码器，触发器再组成多位锁存器或RAM阵列。如果晶体管事先能连接成通用的结构，那么组装一台计算机就容易多了。

这种思想应该首先由英国物理学家 Geoffrey Dummer(1909出生)在1952年5月的一次演讲中提出，他说：

“我想预测一下未来。随着晶体的出现及半导体的日益广泛应用，现在似乎可以设想电子设备在一个固体块里而无需接线。这个块由隔离、传导、整理、放大四个层组成，电子功能通过各层的隔离区域直接连接起来。”

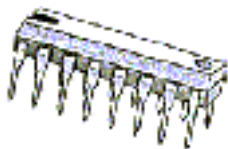
然而，真正可用的产品还不得不再等几年。

在对Dummer的预言毫不知情的情况下，1958年7月，德克萨斯仪器公司的Jack Kilby（生于1923年）想到了在一个硅片上造出许多晶体管、电阻及其他电子组件的方法。6个月以后，即1959年1月，Robert Noyce（1927—1990）也找到了基本上相同的方法。Noyce原先曾为Shockley半导体实验室工作过，但在1957年，他和其他7位科学家离开实验室并成立了Fairchild（仙童）半导体公司。

在技术史上，同时产生的发明很普遍，可能超出了你的想像。尽管Kilby比Noyce早6个月发明了这项技术，并且德克萨斯仪器公司比仙童公司也早申请专利，但Noyce还是首先获得了专利。法律上的争斗随之而来，但仅过了10年，问题就得到了令他们都满意的解决。尽管他们从未在一起工作过，但Kilby和Noyce被认为是集成电路，或称IC，更普遍的是叫芯片的共同发明者。

集成电路要经过复杂的工序才能生产出来，这些工序包括把很薄的硅晶片分层，精确地上涂料，和在不同的区域蚀刻形成微部件。尽管开发一种新的集成电路很昂贵，但收益来自于它的大量生产——生产得越多，就越便宜。

实际的硅晶片薄且很脆弱，为了保护芯片并提供某种方法使芯片中的部件与别的芯片连接，芯片必须安全地封装。集成电路的封装有几种不同的方式，但通常多采用矩形塑料双排直插式封装（或dual inline package，DIP），有14、16或40个管脚从旁边伸出：



这是一个16管脚的芯片。如果你拿着芯片，并使芯片上的凹陷朝左时（如图），从芯片左下角起环绕到右边至末端，管脚依次编号为1~16，第16管脚在左上角。每一边管脚之间相距刚好为1/10英寸。

20世纪60年代，空间项目和军备竞赛刺激了早期的集成电路市场。在民用方面，第一个包含有集成电路的商业产品是1964年由Zenith出售的助听器。1971年，德克萨斯仪器公司开始出售第一批袖珍计算器，同时Pulsar出售了第一块数字表（显然数字手表中集成电路的封装完全不同于刚才图示的例子）。随后出现了其他很多种包含有集成电路的产品。

1965年，戈登·E·摩尔（当时在仙童公司，后来是Intel公司的合伙创始人）注意到技术以这样一种方式在发展：1959年后，可以集成到一块芯片上的晶体管的数目每年都翻一番。他预测这种趋势将继续下去。事实上后来这种趋势放慢了些，因此摩尔定律（最终这样命名）修改为预计每18个月在一个芯片上集成的晶体管数量将增长一倍。这仍是一个令人惊异的增长速度，这也就解释了为什么家用电脑只用了短短几年好像就已过时了。一些人认为摩尔定律将继续适用到2015年。

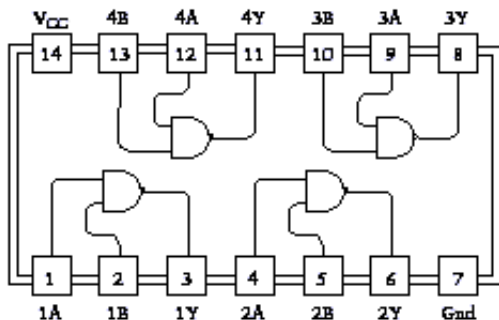
早期，人们习惯于说小规模集成电路，即SSI（small-scale integration），指那些少于10个逻辑门的芯片；中规模集成电路，即MSI（10~100个门）；大规模集成电路，即LSI（100~5000个门）。随后的术语为超大规模集成电路，即VLSI（5000~50000个门）；极大规模集成电路，即SLSI（50000~100000个门）；特大规模集成电路，即超过100000个门。

本章的其余部分和下一章将把时钟停留在20世纪70年代中期，即第一部《星球大战》发行前的年代，那时，VLSI刚刚出现，且好几种技术用来制作构成集成电路的组件，每一种技术有时被称为一个IC家族。到70年代中期，两个“家族”流行开来：TTL和CMOS。

TTL代表晶体管-晶体管逻辑。70年代中期，如果你是一位数字电路设计工程师（即用IC来设计大的电路），一本1.25英寸厚的、德克萨斯仪器公司于1973年出版的名为《The TTL Data Book for Design Engineer》（TTL工程师设计数据手册）的书（以下简称《TTL数据手册》）就会是你书桌上的常客。这是一本关于德克萨斯仪器公司和其他公司出售的TTL集成电路7400系列的完整的参考书，之所以这样称呼是因为该家族的每个集成电路都以开头是74的数字标识。

7400系列中每一个集成电路由许多预先按特定方式连接的逻辑门组成。一些具备简单的预先连接的逻辑门的芯片可以用来建造更大的组件。还有一些芯片提供通用组件，如：触发器、加法器、选择器和译码器。

7400系列中第一个集成电路是号码7400本身，它在《TTL数据手册》里描述为：“四个双输入正与非门”。这意味着这个特定的集成电路包含四个2输入与非门。它们之所以称为“正”与非门是因为有一个电压值与逻辑1对应，没有电压值与逻辑0对应。这个集成电路是一个14管脚的芯片，数据手册中的一个小图展示了管脚对应的输入和输出：



该图是芯片的俯视图（管脚在下面），小的凹陷在左边。

管脚14标为 V_{cc} ，等同于V符号，它表明是电压（顺便说一下，大写字母V的双下标表明是一个电源，下标中的C是指晶体管的集电极输入，它在内部连接到电源）。管脚7标为GND，表示接地。在特定电路中所使用的任何集成电路都必须接电源和公共地。

对于TTL7400系列， V_{cc} 必须在4.75~5.25V之间，换句话说，电源电压必须在 $5V \pm 5\%$ 的范围内。若电压低于4.75V，芯片就无法工作；若超过5.25V，芯片则会被烧坏。通常TTL器件不能使用电池供电，即使你刚好有一个5V的电池，电压也不可能刚好适合于芯片。TTL通常需要从墙上接入电源。

7400芯片中每一个与非门有两个输入和一个输出，它们独立工作。前面曾把输入用1（表明有电压）和0（表明无电压）加以区分，事实上与非门的输入电压可在0（地）~5V（ V_{cc} ）之间变化。TTL中，0~0.8V之间的电压视为逻辑“0”，2~5V之间的电压视为逻辑“1”，应当避免出现0.8~2V之间的输入电压。

TTL的典型输出是0.2V代表逻辑“0”，3.4V代表逻辑“1”。由于电压可能会有点儿变化，集成电路的输入/输出有时称作“高”或“低”，而不是“1”或“0”。此外，低电平也可以表示逻辑“1”，高电平也可以表示逻辑“0”，这种说法称为负逻辑。当7400芯片称作“四个双输入正与非门”时，“正”就表示上面讲到的正逻辑。

如果TTL的典型输出是0.2V代表逻辑“0”，3.4V代表逻辑“1”，则输出确实是在输入范围内，即逻辑“0”在0~0.8V、逻辑“1”在2~5V之间，这也是TTL能隔离噪声的原因。一个“1”输出即使下降1.4V后，电压仍可以高到作为“1”电平输入；一个“0”输出升高0.6V后，电压仍可以低到作为“0”电平输入。

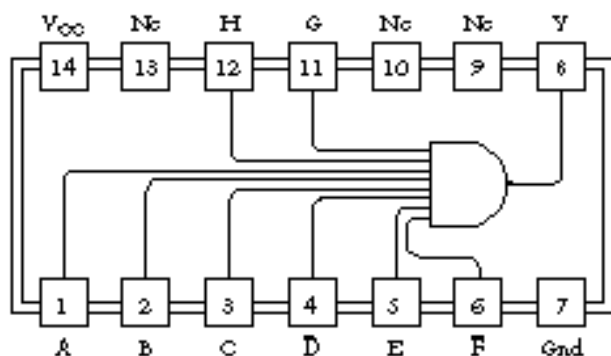
了解一个集成电路最重要的事实可能是它的延迟时间，这指的是输入变化反应到输出所花费的时间。

芯片的延迟通常以纳秒来计算，纳秒缩写为ns。1纳秒非常短暂，千分之一秒为1毫秒，百万分之一秒为1微秒，十亿分之一秒才是1纳秒。7400芯片与非门的延迟时间要保证少于22纳秒，也就是0.000000022秒，或十亿分之22秒。

不能感觉纳秒的长短不仅仅是你一人，地球上任何人除了对纳秒有智力上的理解之外就没有什么了。纳秒比人所经历的任何事物要短暂得多，所以它们永远也不会被理解，任何解释都只会使纳秒变得更难以捉摸。例如，你拿着一本书离你的脸有一英尺远，那么1纳秒就是光从书到你的眼所用的时间。现在，你能说你对纳秒有了更好的认识？

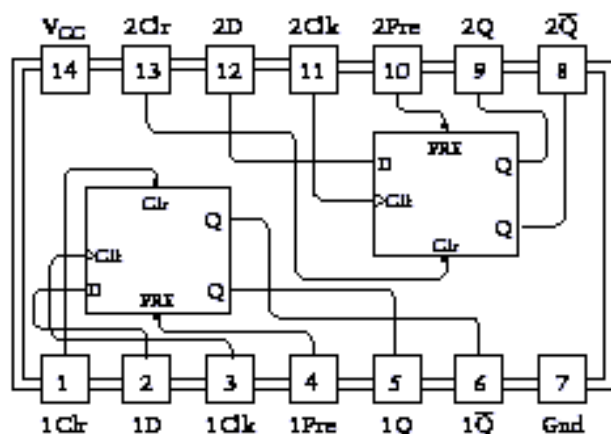
然而，纳秒在计算机中出现是可能的。正如在第17章所见，计算机处理器很笨拙地做着简单的事情——从存储器中取一个字节到寄存器，把一个字节同另一个字节相加，再把结果存回存储器。计算机（不是第17章中的计算机，而是现在使用的计算机）能做大量事情的唯一原因就是那些操作能迅速执行。引用Robert Noyce的话：“当你理解了纳秒之后，在概念上计算机操作就相当简单了。”

我们继续细读《TTL工程师设计数据手册》，就会在书中看到许多熟悉的小条目。7402芯片有四个双输入或非门，7404有六个反相器，7408有四个双输入与门，7432有四个双输入或门，7430有一个8输入与非门：

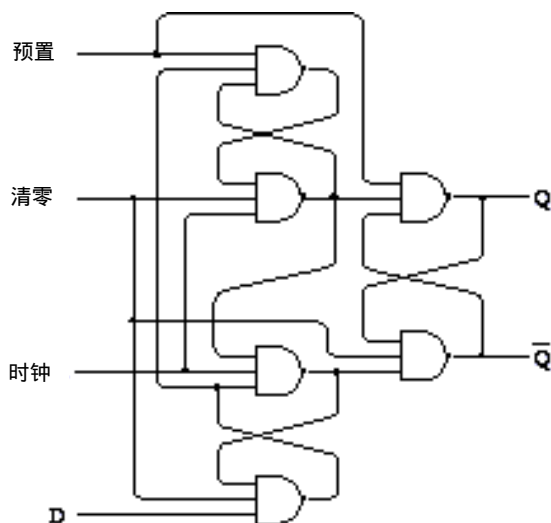


图中缩写NC表示该引脚没有连接到内部电路。

7474芯片是听起来很熟悉的一个芯片，它是“带预置和清零的双D型正边沿触发器”。如下图所示：



TTL数据手册甚至还包含这个芯片的每个触发器的逻辑图：



你会发现这与第14章结尾处的图很相似，只是第14章的图中使用的是或非门。《TTL数据手册》中的逻辑表也有一点点不同：

输入				输出	
Pre	Clr	Clk	D	Q	\bar{Q}
L	H	X	X	H	L
H	L	X	X	L	H
L	L	X	X	H*	H*
H	L	↑	H	H	L
H	H	↑	L	L	H
H	H	L	X	Q ₀	\bar{Q}_0

上表中，“H”代表高电平，“L”代表低电平。你也可以把它们想成是1或0。在上述触发器里，预置与清零输入通常为“0”，在这里它们为“1”。

继续翻阅《TTL数据手册》，你会发现7483芯片是一个4位二进制全加器，74151是一个8-1数据选择器，74154是4-16译码器，74161是同步4位二进制计数器，74175是四个带清除功能的D型触发器。你可以选择这些芯片中的两个做一个8位锁存器。

所以现在你该明白从第11章起使用的各种各样的组件是如何来的了，它们都是从《TTL工程师设计数据手册》中得来的。

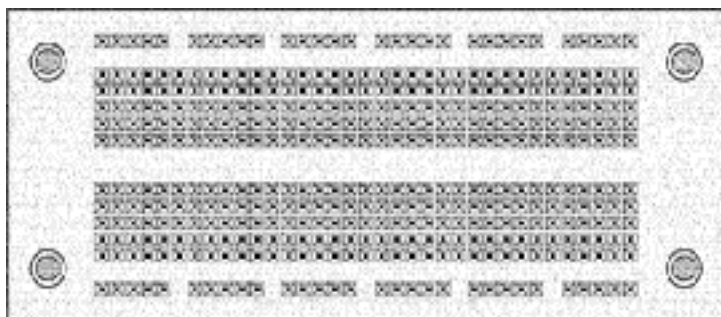
作为一名数字电路设计工程师，需要花费大量的时间去通读《TTL数据手册》，了解要使用的TTL芯片的类型。一旦掌握了你所需要的工具，你就可以用TTL芯片实际组装第17章所示例的计算机。把芯片连接起来要比连接晶体管容易得多，然而你可能不想用TTL组成64KB的RAM阵列。在1973年的《TTL数据手册》中，所列最大容量的RAM芯片才256×1位，需要2048个这种芯片才能组成64KB的RAM！TTL远不是组织存储器的最好技术，第21章将要更多地谈到关于存储器的事情。

你可能也想用好一些的振荡器。可以将TTL反相器的输出端连到输入端，但使用一个事先可预测频率的振荡器要更好一些。构造一个这样的振荡器很容易，就是使用一个石英晶体，在一块小片上引出两条线。这些晶体在特定的频率下产生振荡，通常每秒至少100万周。每秒一百万周称为兆赫，缩写为MHz。如果第17章所示的计算机是用TTL构造的话，那么它在10MHz的时钟频率下可能会运行地很好。每一条指令需要400纳秒的执行时间。当然，这已经比用继电器工作时所能想像的要快多了。

另一个流行的芯片家族是（现在仍然是）CMOS，它代表由金属氧化物填充的半导体。如果你是70年代中期用CMOS集成电路进行电路设计的业余爱好者，你可能会使用一本由National Semiconductor(国家半导体公司)出版的参考书，它在你所在地方的电器行就能见到，书名为《CMOS Databook》。此书包含了CMOS集成电路4000系列的信息。

TTL的电源要求是在4.75~5.25伏之间，但CMOS则可以是3~18伏之间的任何数值，范围多么大呀！此外，CMOS比TTL功耗要小，这就可以使用电池来驱动小型CMOS电路。CMOS的缺陷是速度慢。例如，CMOS 4008 4位全加器在5伏电压下只能保证750纳秒的延迟。当电源电压升高时，速度加快——10伏时，延迟为250纳秒；15伏时，延迟为190纳秒。但是CMOS设备不能接近于TTL 4位加法器，TTL 4位加法器的延迟为24纳秒。（25年前，在TTL的速度和CMOS的低功耗之间的权衡是很清楚的，今天，也有低功耗的TTL和高速率的CMOS。）

实践上，你可以在塑料面包板上连接这些芯片进行实验：



每一个有5个孔的短行在塑料板下是电导通的。你把芯片插在面包板上，并使芯片跨在中间的长槽上，管脚插入槽两边的孔中。集成电路的每一个管脚都与其他4个孔电连接。你可以将线插入其余孔中来连接芯片。

你也可以使用一种称为线缠绕的技术使芯片的连接更加牢固。每一个芯片插入带有长方形柱子的插座上：



每一个柱子对应于芯片的一个管脚，插座本身也插入打孔的板上。在板的另一边，你会看到特殊的用线缠绕的插槽紧紧包裹着围绕柱子的绝缘线。柱子的方形边缘穿破绝缘层并使它与导线电连接。

如果你实际在使用集成电路来构造一个电路，就要使用一块印刷线路板。以前，这是业余爱好者可以做的事情。板子上有孔，并覆盖一层薄的铜箔。首先，你要在需要保护的区域的铜箔上涂上防酸物，并用酸腐蚀其余部分，然后你把集成电路插座（或集成电路本身）直接焊在板的铜上。由于集成电路之间有许多内部连接，一层铜箔通常是不够用的，商业制造的印刷线路板有许多内部互连的层。

到70年代早期，已可以在一块电路板上用集成电路构造一个完整的计算机处理器。把整个处理器做在一块芯片上已只是时间问题。尽管德克萨斯仪器公司 1971年为单片计算机申请了专利，但实际的制造荣誉却属于 Intel——一家于1968年，由前仙童雇员 Robert Noyce和 Gordon Moore 建立的公司。Intel的第一个主要产品是1970年生产的可存储1024位的存储器芯片，在当时这是可做在一块芯片上的最大存储容量。

Intel在为由日本的 Busicom公司生产的可编程计算器设计芯片时，决定采用不同的方法。正如Intel公司的工程师Ted Hoff写的：“不是想使他们的设备成为一个带有编程能力的计算器，而是想使它作为通常目的计算机可编程为一个计算器”。这就产生了Intel 4004，第一个“芯片上的计算机”或微处理器。1971年11月，4004投入使用，它带有2300个晶体管。（根据摩尔定律，18年后微处理器应该有4000倍数量的晶体管，即大约1000万个，这是相当准确的预计。）

有了晶体管的数量，下面将描述 4004的其他三个很重要的特性。自 4004诞生以来，这三个指标经常用来作为微处理器相互比较的标准。

第一，4004是4位的微处理器。这意味着处理器的数据通路宽度只有 4位，做加、减法运算时，它一次只处理4位。相比较，第17章中开发的计算机有8位数据通路，所以它是8位处理

器。正如我们将看到的4位微处理器很快就被8位微处理器所超越，没有人会停滞不前。70年代后期，16位微处理器出现了。回想一样第17章的内容，以及在8位处理器上进行两个16位数相加所需要的指令代码，你就会欣赏16位处理器带给你的好处。80年代中期，32位微处理器出现了，并从那以后一直作为家用计算机的主流微处理器。

第二，4004有最大每秒108 000周的时钟频率，即108KHz。时钟频率是可连接到微处理器并能运行的振荡器的最大频率。再要快的话，微处理器就可能出错。到1999年，家用计算机微处理器的时钟频率已达到500MHz——大约4004运行速率的5000倍。

第三，4004可寻址的存储器是640个字节。这看起来是一个小得可笑的数量，然而这和当时可用的存储芯片的容量是相匹配的。下一章你就会看到，两年内微处理器可达到64KB的寻址空间，这是第17章所提及的机器的容量。到1999年，Intel的微处理器已达64TB的可寻址空间，尽管大多数人家用电脑的RAM容量还低于256MB。

这三个数字不会影响一台计算机的能力。例如，一个4位处理器要进行32位数的加法，只要简单地按4位一次进行。某种意义上，所有数字计算机都是相同的。如果一个处理器的硬件能做别的处理器做不了的，那么别的处理器可以用软件实现，最终它们能做同样的事情。这也就是图灵1973年的论文里有关计算能力的含义。

然而，处理器根本的不同是在速度上。同时，速度也是我们为什么使用计算机的一个重要原因。

最大时钟频率是影响处理器总体速度的一个显著因素，时钟频率决定了每一条指令的执行速度。处理器的数据宽度也影响其执行速度。虽然一个4位处理器可进行32位数的加法运算，但它的执行速度不可能与32位处理器一样。然而，令人迷惑的是，处理器可寻址的最大存储器容量也是影响速度的一个因素。最初，寻址空间看起来好像与处理器速度无关，而只反映了处理器在执行某些需要大量存储空间的功能时处理器的能力限度。但处理器通过利用存储器地址来控制用于保存或提取信息其他媒体，可避开存储容量的限制。（例如，假设写到某个存储地址的每个字节实际上都是在纸带上穿孔，从存储地址读的每个字节都是从纸带上读。）然而这种做法减慢了整个计算机的处理速度——又是速度问题。

当然，这三个数字都只是粗略地显示了微处理器的运行速度。这些数字没有告诉任何有关微处理器内部体系结构或机器码指令的效率和能力的问题。处理器越来越复杂，许多以前用软件来实现的普通工作现在可以用微处理器来实现。我们在后面的各章中可看到这种趋势的一些例子。

即使所有的数字计算机都具有同等的能力，即使它们只能做与图灵设计的原始计算机一样的工作，处理器的执行速度最终也会影响计算机系统的总体用途。例如，比人类大脑的计算速度还慢的计算机是毫无用处的。当我们在现代计算机的屏幕上看电影时，如果处理器需要花费1分钟的时间来处理每一帧，我们也是无法忍受的。

回到20世纪70年代中期，先不说4004的局限性，但毕竟它是一个开始。1972年4月，Intel发布了8008——一个8位微处理器，时钟频率为200KHz，可寻址16KB的存储空间。（仅用三个数来总结一个处理器是多么容易。）后来，1974年5月期间，Intel和Motorola公司同时发布了对8008进行改进的微处理器，这两种芯片改变了整个世界。