

信 息

什么是信息？就是进行传递或交流的一组语言、文字、符号或图象。（参阅封三）。在早期，信息不过是消息的同义语，其深邃的内涵却鲜为人知。

在人类社会里，应该说，信息与物质、能量一样，有其重要的地位，是人类赖以生存发展的基本要素。现代社会，信息的地位日趋重要。因此，了解信息，掌握信息，懂得如何充分有效地利用信息也就变得非常迫切了。当然，给信息一个明白无误的定义是非常困难的。信息所涉及的范围十分广泛，不仅包括所有的知识，还包括通过我们五官感觉到的一切。例如，新的科学技术成果，报纸上的新闻，市场行情，天气预报以至一幅画、一张照片，无不属于信息的范畴。可以这样说，信息是人类社会不可缺少的部分，我们是生活在信息的汪洋大海之中。信息科学也横跨了许多不同的学科，物理学、电子学、通讯工程、计算机科学、生物学、语言学、密码学等。

很明显，信息要以相互联系为前提，没有联系也就无所谓信息，任何事物都可以作为信息源，事物的特征和状态是潜在的信息，信息的储存不过是延迟了的传输。这就是说，信息是一种相对的概念：它自身不能单独存在，必须依附于一定的载体，而且也还要和接收者以及它所要达到的目的相联系，这才开始成为信息。正如维纳(N. Wiener)所说：“信息就是信息，不是物质，也不是能量。不承认这一点的唯物论今天就不能存在下去。”当然信息离不开物质载体，对它进行处理、传输或操作，必然要消耗能量。

信息具有多种多样的载体，这是信息的重要特征。例如，人类通过语言、符号等来传递信息，而生物体内的信息则是通过电化学的变化，经过神经系统来传递；信息本身与载体，载带信息的物理现象之间

是有区别的，但要完全分开，有时也会有困难。譬如，图书、杂志、唱片、电视等，都是载带信息的媒介物，可是人们又往往把它们看成是信息的本身。信息在传输过程中，由于不可避免的噪声干扰或译码错误，往往会发生信息的减损，最理想的传输过程在于保真，即将信息一成不变地传输过去。但另一方面，信息



熵与信息

——麦克斯韦妖的启示



冯端 · 冯步云

· 基础物理知识讲座 ·

还有一个重要特征是不但不会在使用中消耗掉，而且还可以复制、散布，也就是说它跟物质和能量不同，不会越用越少。就象书，它可供千万人阅读产生不可估量的影响，举一个例子来说明：马克思与恩格斯合写的“共产党宣言”这本小册子，经过多次印刷、翻译和传播，在社会发展和人类历史上产生了天翻地覆的作用，而信息本身依然存在，毫无减损。

显而易见，信息既有量上的差别，又有质的不同。一段文字、字数的多少反映了量的差别，而其蕴含的意义则反映了质的不同。在日

常生活中，我们对于信息的量与质有相当深的体验，量的差别固然重要，质的不同更不容忽视。同样是20个字，李白的“床前明月光，疑是地上霜；举头望明月，低头思故乡”情景交融，蕴藉隽永，余韵袅袅，成为千古传诵的名诗；而一份电报写道“我因为生了病不能及时赶来参加会议非常抱歉”只是直截了当地说明了一件事；另外还可以随便凑上20个字，也许是一派胡言乱语，毫无意义。就信息的量而言，三者并无差别，发电报的话，收费完全一样；而就其含意或价值而言，却有天壤之别。当然，信息的价值和意义却又是相对的，是随着接收信息者的条件和状态不同而变的。例如封三 b、c 中所显示的二段文字，对于不通晓这些语言的读者来说，无异于天书，毫无意义。对于相应文字的专家来说，就可以传达一定的意义；而复活节岛上发现的铭文(封三 a)却至今尚未破译，因此虽则它蕴含了意义没有人能够理解；至于乐谱(封三 e)与数学公式(封三 f)，更是如此，只有对于这方面有素养的人，方有意义。

但是应该指出，有关信息内容的问题，实际上涉及对价值的评估，显然超出了自然科学的范围。目前对此尚无法做出具有客观性的、大家都接受的论断，因而不得已只能舍此求其次：采用电报局的办法，只计字数不问内容，回避了涉及信息内容这一重大而有争议的问题，而单在信息量的问题上上下功夫，这正是香农建立信息论这一门科学的出发点。

信息的统计理论

在信息量的问题上上下功夫，关键在于如何给出具有普遍意义的信息量的定义。由于字数和所采用的语言文字或符号系统密切相关，而各种语言文字和符号系统的情况又千差万别，例如，一段中文，系由许多汉字所组成，每个汉字又是在上万个汉字中挑选出来，其几率约为

1/10⁴; 拼音出来,它就变成一组字母(包括空白),每一字母则是在 26 个拉丁字母和一空白间抉择的结果,其几率为 1/27; 翻译成莫斯电码,每一电码,只有二种可能性,一划或一点,二者必居其一,其几率为 1/2; 因而要定义信息量,必须摆脱具体的语言或符号系统的限制,从根本上来考虑,正是基于这个考虑,香农提出了信息的统计理论。

首先考虑存在有 P 种可能性,其几率是均等的。例如,一个莫斯电码 $P = 2$; 一个拉丁字母 $P = 27$; 一旦在 P 种可能性之中选定其一,我们就取得了信息, P 愈大,相应地作出了选择之后的信息量也愈大,这样,信息 I 被定义为:

$$I = K \ln P$$

这里的 K 为比例常数。由于相互独立的选择可能性(或几率)是相乘的,对应的信息量按此定义就具有相加性。如果考虑一个信息量是一连串几个相互独立的选择的结果,其中每一个选择都是在 0 或 1 之间作出的,因而总的 P 值应为 $P = 2^n$, 于是 $I = K \log_2 P = nK \log_2 2$ 。如果令 I 与 n 等同,则

$$K = 1/\log_2 e = \log_2 e$$

这样定出的信息量的单位就是在计算机科学中普遍使用的比特(bit); 如果令 K 等于玻耳兹曼常数 k , 那末信息量就用熵的单位来度量。

上述的例子中,终态都是唯一的,很显然可以将 I 的定义推广到终态还存在有多种可能性的情况,就需要分别知道始态的可能性 P_0 和终态的可能性 P_1 , 这样

$$I = K \log_2 P_0/P_1 = K \log_2 P_0 - K \log_2 P_1$$

例如,考虑掷骰子所获得的信息,在未掷之前, $P_0 = 6$, 掷出某一确定数字的信息 ($P = 1$) 等于 $K \log_2 6$, 这样,掷出偶数的信息 ($P_1 = 3$) 就等于 $K \log_2 2$ 。

熵与信息

按照布里渊的思想,信息之不同的可能性可以和状态相貌数联系

起来,从而获得信息与熵的关系。考虑某一系统,始态时,信息 $I_0 = 0$ 相貌数为 P_0 , 熵就等于

$$S_0 = k \log_2 P_0,$$

而终态时,信息 $I_1 \neq 0$, 相貌数 $P_1 < P_0$, 熵 $S_1 = k \log_2 P_1$; 显然,在所考虑的情况,系统并非孤立的,当信息获得后,使相貌数降低,而导致熵的减少,而这信息必须由外界机构提供,它的熵增加了,这样

$$I_1 = K(\log_2 P_0 - \log_2 P_1) = S_0 - S_1$$

即信息相等于物理系统中总熵中的一个负值的量:

$$\text{信息} = \text{熵 } S \text{ 的减少} \\ = \text{负熵 } N \text{ 的增加}$$

(定义负熵 $N = -S$)。就是说,信息可以转换为负熵,反之亦然——这就是信息的负熵原理。这一思想萌生于西拉德早先的论文中,但未被学术界注意,熵与信息的关系重新为香农所发现,但他所定义的熵与热力学的熵之标准定义差一正负号,因而香农定义的信息在传递的不可逆过程中,由于噪音的干涉,传递的差错,只减不增,呈现了负熵的特征。

值得注意,也有人反其道而行之,将统计物理的熵的定义建立在信息论的基础上,即认为熵是对系统无知度(或信息量欠缺)的度量。例如我们对于系统的信息若仅限于若干宏观参量,对其微观状态的细节一无所知,那么我们将预期该系统处于熵为极大值的状态,因为如果处于熵较低的状态,则必会提供更多的信息。这就是杨乃斯(E. T. Jaynes)于 1957 年提出的原理。这种说法不无道理,也有人用它来取代统计物理中熵的定义。例如在非平衡态的气体,我们可能测量出描述其宏观流动的数据,因而提供较高的信息量;而在平衡态仅需要少数几个宏观量来描写其状态。这样,熵的增加就意味着信息的减少。

信息处理消耗的能量极限

下面来讨论处理信息的能量消

耗问题,这是通讯技术,计算机技术和物理测量都十分关心的一个问题。当然一个具体的机器,如一台电子计算机,进行信息操作所消耗的能量取决于计算机的技术水平,随着电子技术的飞速进展,每操作一个比特信息所需的能量随时代在急骤下降,50 年内从 10^{-3} 焦降到 10^{-13} 焦,下降了十个数量级,预期这一趋势仍将继续(图 3)。但显然不可能无限持续下去,物理学的规律必然会对这能量规定一个下限。

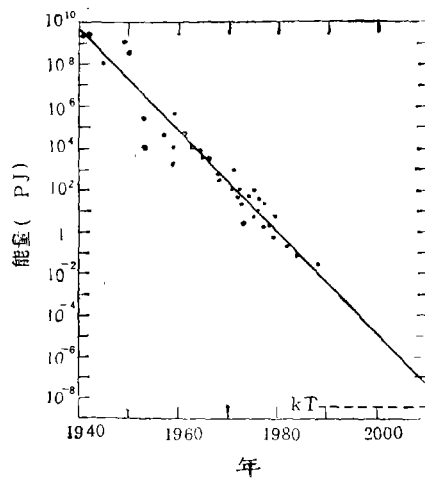


图 3 处理 1 比特信息所消耗的能量

我们知道,热机之效率的上限是由热力学第二定律确定的,不管工艺技术如何改进,这一上限是无从超越的。这里的情况有些相似,有待于作进一步的分析。

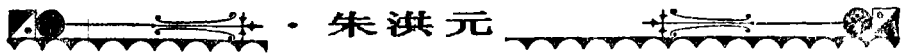
按照能量均分定律,物体每一自由度所分配到的能量约为

$$kT \approx 4 \times 10^{-21} \text{ 焦}$$

(在室温 $T \approx 300 \text{ K}$) 任何进行信息处理的元件总需要实现“开关”或“读写”的功能, kT 的能量相当于噪音的水平,那么“开关”或“读写”的能量至少应为 kT 的好几倍,这是由元件的工作情况所规定限制。

从信息论的角度,1 比特的信息量等于 $\log_2 2$, 对应的熵为 $-k \log_2 2$ 。所给出的热量变化也正好等于 $kT \log_2 2$ 。另外从香农提出的信息通道容量的基本公式也可以

高能物理学发展的回顾与展望(下)



· 朱洪元 ·

(续上期)

标准模型理论显然还需要提高。但今后一段时期的发展不会像 50 年代、60 年代、70 年代那样快。因为没有出现和理论相矛盾的实验结果来指引理论发展的方向,只能根据现有理论研究的长远目标和理论本身所包含的内部矛盾来进行探索。

回想在本世纪 20 年代,中子还没有被发现,认识到的基本粒子只有两种:电子和质子。当时已经发现的基本相互作用只有两种:万有引力相互作用和电磁相互作用。理论中的基本参数只有六个:

$$\hbar, c, m_e, m_p, e, G_N$$

当时许多理论物理学家期望,假使能建立万有引力相互作用和电磁相互作用、电子和质子的统一理论,从中将三个无量纲常数

$$\frac{e^2}{\hbar c} = \frac{1}{137}$$

$$\frac{m_p}{m_e} = 1836$$

$$\frac{G_N m_e m_p}{\hbar c} = 3.2 \times 10^{-44}$$

推导出来,那末我们就终于达到最终理论(Theory of Everything).

爱因斯坦和爱丁顿在他们的后半生中致力于这种探索,但都没有成功。

从那时到现在又发现了二种基本相互作用:强相互作用和弱相互作用;发现了几百种前所未有的粒子,已知物质的“基本”组元从两种增加到 30 种以上。理

得到类似的结果。近年来,朗道尔(R. Landauer)等仔细分析信息处理中的能量极限,认为在理想化的计算和测量等过程之中,不可避免的能流消耗仅在于将存储的信息抹去,对于一个比特的信息,也正好等于 $kT \log_2$ 。这样,殊途同归,结论基本相同。

我们不妨回头来对图 2 所示的西拉德的理想机器的问题作一定量的分析。当妖精对分子的位置在左还是右作出判断,提供正好一个比

特的信息,最少需要 $kT \log_2$ 的能量。但他利用信息使单分子气体膨胀作功,这相当于气体容积加倍的真空膨胀,提供的功为 $kT \log_2$ 。两者正好得失相抵。

最近有人来探讨妖精是否有采集信息更经济的方法,提出将 N 个西拉德机器耦合起来,能否有利可图。当等到 N 个分子同样都处在左侧时,妖精才来操纵机器作功。这些对外作功等于 $NkT \log_2$, 而清除信息所需能量仍为 $kT \log_2$, 似

论中的基本物理参数从 6 种增加到 21 种。理论上需要解释的无量纲常数从 3 种增加到 18 种。看来即使 20 年代最伟大的物理学家也将宇宙看得太简单了。在 20 年代提出来的问题到今天已经发展为:如何统一地理解轻子、层子、光子、 $W^+ \cdot W^- \cdot Z^0$ 中间玻色子、胶子、引力子、希格斯粒子等一切粒子和它们之间的一切基本相互作用;如何建立一个基本理论,能统一地解释所有的实验结果,又能统一地将 18 个无量纲常数从这个理论本身中推导出来。显然,自从 20 年代以来,我们对自然的认识扩充和深入了很多。在另一方面也使我们进一步认识到宇宙的深广和我们的无知。我想,认识到我们的无知是一大进步。这将激励我们更加努力向更高的目标去探索。

标准模型理论和广义相对论中的一小部分参数来自规范场和万有引力场部分,因此和物理规律的对称性有关。要减少来自这方面的参数,看来得探索物理现象深处是否隐藏着更大的对称性。目前理论探索中相当大一部分就属于这一个方面。如

大统一理论
超对称理论
超引力理论
超弦理论

在 18 个无量纲参数中,绝大部分来自希格斯场部分,因此和对称性的破缺有关,一切粒子的质量都和希格斯场有关。质量不仅是一切粒子的一个基本性质,而且是万有引力场的源。而且不同代的粒子之间的连系也是通过希格斯场实现的。在另一方面,理论中和

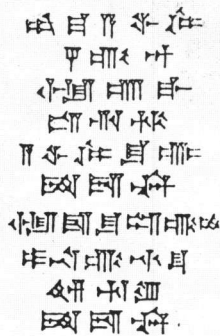
乎有利可图了。但是且慢, N 个粒子都在一侧的几率是非常小的,对应于极其难得的涨落,需要等待很长的时间。类似于斯摩罗柯夫斯基设想的利用布朗粒子来作功的机器,实际上还是行不通的。通过以上的分析,似乎可以得出结论,信息处理所消耗能量的下限还是由热力学第二定律所规定的,否则将导致第二类永动机的问世。到头来,妖精虽然神通广大,还是象孙悟空一样,翻不出如来佛的手掌心。(待续)

不同形式的信息载体

(详见本期冯端一文)



(g) 艾契尔 的画



(c) 亚述的楔形文字

வாசிக்குவருகையில், விவொரு வாக்கியத்தையும் அந்நூலிய படத்தோடு சிபிபிட்டுப் பார்த்து, தேராமும் அர்த்தமாகிக்கொண்டு வருகும். மனப்பாடம் பண்ணிக்கொள்ள வேண்டிய அம்சங்கள் மிகக் குறைவு. ஆகையினால் அர்த்தத்துக்கு இணங்க வாக்கிய அமைப்பின் வேறுபாடுகளைக் கண்டறிந்து கொள்வதற்கு வசதிகள் உண்டு. இம்முறை அநுசரித்து ஆங்கிலம் கற்றுக்கொள்வது கஷ்டமில்லாததுமன்றி விளையாட்டுச் சம்பந்தப்பட்டது.

(b) 泰米尔文



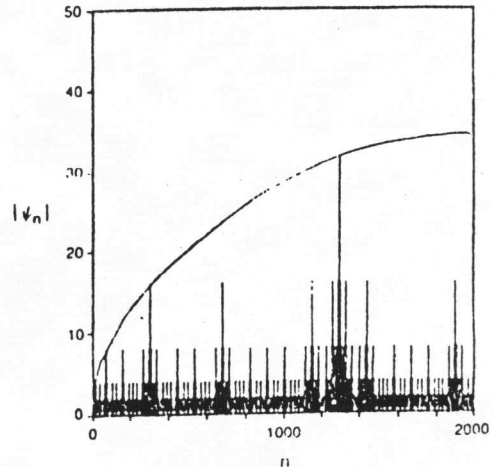
(e) 一段巴哈的乐谱



(a) 复活节岛上发现的铭文

$$\left. \begin{aligned} \frac{\hbar^2}{2m} \frac{d^2 \xi(x)}{dx^2} + (V_a - E_a) \xi(x) &= 0, \\ \frac{\hbar^2}{2m} \frac{d^2 \eta(y)}{dy^2} + (V_v - E_v) \eta(y) &= 0. \end{aligned} \right\}$$

(f) 两维的薛定谔方程组



(d) 准周期结构的波函数

X 光透镜

(详见本期颜一鸣文)

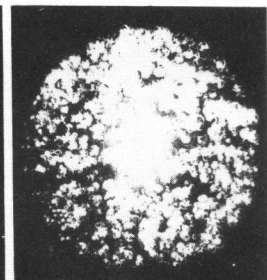
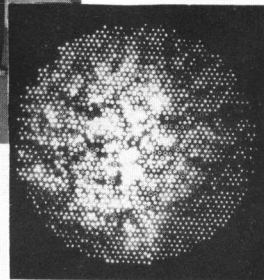
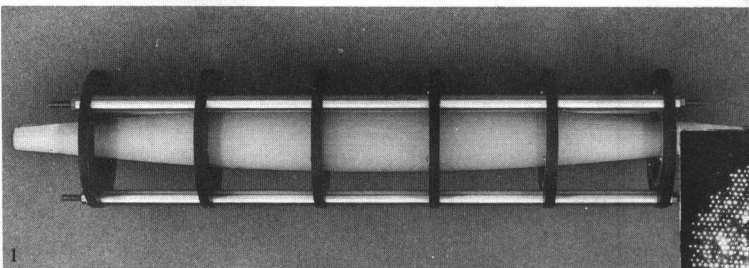


图1 图2：会聚透镜与平行束透镜

图3 图4 图5：X 光束空间分布

