

〔日〕 小沢一雅博士著

情报理论基础

国外情报科学编辑部

G35

X 101

情报理论基础

(日) 小泽一雅 著

王秀芝 靖继鹏 译

谢邦同 校



国外情报科学编辑部

1985年

478655

前 言

近年来日本大学，特别是私立学校平均一个教员所教学生人数逐渐增加。这固然说明大学教育取得了成绩，同时也反映还存在许多困难问题。因为这样一来教授一名学生所用的时间必然减少。为了在有限时间内提高教学效果，不得不寄希望于大力提高学生的自学热情。因此，教科书就理所当然地承担起提高自学积极性的重要任务。

本书，是著者在给大阪电讯大学工程系三年级学生讲授情报理论课的讲稿基础上，加工整理形成的。为了有助于学生的理解配有适当例题，并给以简要解答。对重要的公式，用(公式)这一标记加以表示，以引起注意。为便于掌握本书的内容，没有作一般化的讨论，而注重直观性的理解，并避免乱用数学公式。

尽管著者希望此书能增强学生的自学意志，但由于著者水平有限恐难完全达到目的，承蒙指正将是非常荣幸。

本书的出版，得到大阪电讯大学工程系
宫越一雄教授的帮助，在此深表谢意。另外向
参加原稿校对的以浅野哲夫付教授为首的各
位同事再次致以谢意。

著 者

1979年10月

译者前言

本书是日本大阪电讯大学小泽一雅教授，在为工程系三年级学生讲授情报理论课讲义的基础上，经整理、提高而撰写成的。

本书以香农通信理论为依据，在运用概率、统计、积分变换方法与熵概念的基础上，对情报的发生与传递给予了定量的解释与处理。并详细阐述了关于情报量、情报源、信道、符号化及连续信号等重要理论。

著者在阐述过程中注意对概念的直观理解，数学运用也恰到好处，为便于掌握还配有较多的例题。

在目前缺少情报基本理论书籍的情况下，译者相信本书的译出，对于从事情报理论的研究工作者以及对情报理论有兴趣的读者来说，将是一本很有价值的参考书。更可作为情报理论基本教材，供大专院校情报专业学生、研究生和教师使用。

本书在绘图、印刷出版过程中张纯保、

王鸿翱两同志做了不少工作,在此一并表示感谢。

由于译者水平有限,加上时间仓促,译文中的错误与不妥之处在所难免,在此,恳切希望读者给予批评指正。

译 者

1985年12月于吉林工业大学

目 录

1. 绪 论	(1)
2. 情报量	(4)
2.1 事象和记号	(4)
2.2 概率模型	(6)
2.2.1 事象和概率	(6)
2.2.2 条件概率	(9)
2.2.3 完全事象系	(13)
2.3 情报的定量化	(14)
2.3.1 情报量	(14)
2.3.2 情报量的意义	(17)
2.4 熵	(20)
2.4.1 完全事象系的熵	(20)
2.4.2 熵的性质	(23)
2.4.3 有序和无序	(29)
2.4.4 二个完全事象系	(30)
2.4.5 各种熵	(36)
3. 情报的产生和传递	(40)
3.1 情报的种类	(40)
3.2 情报源	(42)
3.2.1 情报源的模型化	(42)
3.2.2 约束记号系列	(45)
3.2.3 英文的统计性质	(51)
3.3 马尔克夫情报源	(56)

3.3.1	定 义.....	(56)
3.3.2	马尔克夫情报源的各种性质.....	(61)
3.3.3	马尔克夫情报源的熵.....	(68)
3.4	信 道.....	(72)
3.4.1	情报传递的模型化.....	(72)
3.4.2	信道容量.....	(79)
3.4.3	无噪声的信道.....	(81)
3.4.4	有噪声的信道.....	(88)
4.	符号化.....	(95)
4.1	效率与冗余度.....	(95)
4.1.1	记号系列的效率和冗余度.....	(95)
4.1.2	符号化和效率.....	(97)
4.2	符号化和信道.....	(101)
4.2.1	情报传送和符号化.....	(101)
4.2.2	香农第1定理.....	(103)
4.2.3	香农第2定理.....	(104)
4.3	高效率符号化法.....	(111)
4.3.1	香农一法奴符号化法.....	(111)
4.3.2	哈夫曼符号化法.....	(115)
4.4	具有冗余度的符号化法.....	(120)
4.4.1	具有冗余度的符号.....	(120)
4.4.2	汉明距离.....	(122)
4.4.3	奇偶检验.....	(127)
4.4.4	线形符号.....	(129)
5.	连续信号.....	(136)
5.1	熵.....	(136)
5.1.1	概率变数和概率密度.....	(136)

5.1.2	连续信号的熵.....	(141)
5.1.3	最大熵.....	(146)
5.1.4	信道容量.....	(150)
5.2	频谱.....	(155)
5.2.1	付立叶级数.....	(155)
5.2.2	付立叶变换.....	(163)
5.2.3	不确定性原理.....	(171)
5.3	典型化定理.....	(177)
5.3.1	频带限制.....	(177)
5.3.2	典型化定理.....	(179)
5.3.3	信号空间.....	(186)
练习题.....		(193)
参考文献.....		(197)

1 绪 论

现在已经进入在社会生活的各种场合中,经常使用“情报”这一语言的时代。产生这种状况的重要原因之一,就是计算机在社会体系中具有了重要而且是决定性的作用。我们正面临着工业革命后出现的新的历史变革。“情报”语言正在变成政治经济、文化、技术革新或社会变革的现时代的主要而又关键的词汇。进入情报时代之后,情报工程给旧工程添加了新的内容,迫使人们改变对事物的看法和想法。自从把“情报”作为科学研究对象,并在理论体系上采用香农(Shannon)通信理论*以后,情报科学已超出了通信工程范围,而且逐渐地突破其他学科领域,开始走上独立发展的道路,并正在取得巨大的进展。

以前的自然科学,像经常所说的那样是有关物质和能量的科学。然而与此相反,情报科学却是关于“情报”这一与人类有密切关系的科学。可以说物质和能量是以与人类无关的方式存在着,但是,“情报”就不能认为是脱离人类而存在。情报科学不但对自然科学,而且对人文、社会科学也都具有很大影响,这说明情报科学并不是与人类没有关系的。所谓意志决定,被认为是缩小选择自由度的行为。香农把情报量看做是选择的自由度,因此,如果把获得情报量当成是逐渐缩小选择自由度的那种所谓意志决定过程时,就能认清

*参考文献 13)

情报科学的对象。

情报时代的社会就是情报社会，创建情报社会光明未来的一种手段是情报工程学。把人类从陈旧的、单纯的劳动中解放出来，使他们能够真正从事社会活动，从而建立了有相互信赖关系的巨大的共同体，这就是情报社会的光明未来。

有人类的地方就有“情报”，使一切情报高效率地顺利地进行“交流”，是情报时代所要解决的基本问题。在任何情况下都能完满的解决这个问题，对多数人来说都是必要的。现今把掌握这种新技术的人员称之为情报工作者。情报工作者，必须具有深刻理解“情报”本质，确切弄清问题根源的能力。情报工作者并不一定限于工程技术人员，但是特别要求从事技术工作的情报工作者具备从工程观点解决这类问题的能力。

情报工程学，是关于情报的产生、传递、搜集、存贮、处理的学问，其内容正在大范围地扩展与加深。香农的通信理论，是以处理情报的产生与传递为主，现在人们所说的“情报理论”的内容大部分以香农通信理论为依据。为此，也有“香农派情报理论”这种叫法。

情报理论，从何谓“情报”这一根本问题出发，给出关于它的产生与传递的系统看法。因此，情报工作者应考虑到有必要首先学习情报理论。而从事技术工作的人员，进行情报工程的各种技术、理论学习也是必要的。作为已进入其他领域的工作者，把掌握的各领域有用知识以“情报”这一新观点加以整理也是需要的。

在本书的第2章，说明什么是情报，以及情报量和熵的概念，在第3章介绍关于情报的发生和传递，而与它有关的符号化则放在第4章讨论。在第5章介绍与连续信号有关的

情报理论。从“概率”这一观点来处理情报的思想贯穿全书。香农派的情报理论，巧妙地引入了概率概念，建成一个有完美系统的理论体系，这是从学习情报理论中得到的宝贵经验之一。

2 情 报 量

我们来讨论一下平时无意之中常说的“情报”这一语言的实质是什么？特别是应该怎样使“情报量”像电流和电压那样变成一个定量的数量呢。

2.1 事象和记号

当我们听到“明日降雨”，“中国要人来日本”等消息后，就认为已经得到了某种情报。这些消息告诉我们将要发生的某些特定事情和事件，也就是把某些现象即将发生的知识传递给我们。

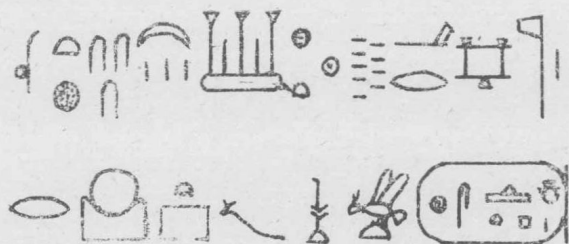


图 2.1 古埃及利用象形文字记录所发生的事件。

于是，当我们得到了某种情报以后，就确知一定会有相应的事情和现象发生。“下雨”，“太阳升起”或者“发生地震”等等这些现象不仅是包罗万象的，而且也是与人类有极其密切的关系，甚至于也包括有待于分类的无数细小的事件和现象。当现象引起了人类的注意，并使之认识到现象和现象与自身有着密切关系时起，就开始出现了“情报”这一

特定词汇。

自人类的古代社会以来，对于引起重视的现象，人们或者用口传，或者以某种形式记录下来，从而得到有关的基本知识。如“国王逝世”这一事象，是用象形文字记录下来的，又“敌人来了”这一事象则曾用“烽火”来传递。我们的常用汉字就其起源来看，则是沿着与象形文字的不同方向产生和发展起来的。如果去掉其包罗万象的形式就很容易在现代汉字中找出其明显的痕迹。事象最初就是以这样的文字、语言，“烽火”或者绘画等方式来表现它与人类具有的密切关系。象这样用于表现事象的文字、语言或者象形文字等，我们把它称之为记号。

图2.2 是把同一事象用绘画，汉字以及英文3种不同的“记号”表示的例子。形态虽然完全不同，可是接受这些记



图 2.2

号的人们却得到了完全相同内容的情报。这样，我们不妨把情报看做是记录、描写事象的记号。反之，不能用记号表示的事象，也可以说它无法给人类带来可选取的任何情报。

图2.3，“阴或晴”是一个表示天气的记号，把这种表征两个不同事象的两个记号组合使用，可用来表现另一个事象。这就是表示几个事象组合成一个新事象的例子。反之，把某个事象分解为几个基本事象的情况也很多。象这种事象



「阴或晴」

的组合与分解，就要用记号来表示它们的各种关系，从数学角度来看可以把上述这些过程看做是某种“演算”。

图 2.3

规定用 E 、 F 表示任意两个事象，按照集合论的定义，关于 E 和 F 的演算如下：

(1) 把「 E 和 F 至少有一个发生」的这一事象叫做 E 和 F 的事象和，用 $E \cup F$ 表示。

(2) 把「 E 和 F 一起发生」的这一事象叫做事象积，用 $E \cap F$ 表示。

(3) 如果发生 E 则一定发生 F 时，写成 $E \subset F$ 。这是 F 包含 E 的一种情况。

(4) 把「不发生 E 」这一事象表示为 $\overline{E}$ ，叫做逆事象。这样，设全部事象的集合为全事象 U 时，由此可写成 $U = E \cup \overline{E}$ 。

另外，作为与集合论的空集合概念对应的东西，定义空事象为 $E \cap \overline{E} = \phi$ 。事象积 $E \cap \overline{E}$ 是决不能产生的事象，是作为实体不存在的概念。因此，用特殊记号 ϕ 来表示它。

我们把表征某一事象的记号看作是情报。通过对 E 和 F 这种记号与记号的演算以及构成事象的数学方法的介绍，对理解后边的事象概率概念是很重要的。

2.2 概率模型

2.2.1 事象和概率

我们从报纸、收音机、电视等得到许多情报，其中能引

起注意的信息我们说它是“有信息价值”的。信息是用语言这一记号所记述的各种事象，但信息的价值是怎样确定的呢？这是极其困难的问题。一般情况下，一个事象的信息价值将取决于各种错综复杂的因素很难有准确地评价。其中特别是要掌握主观因素的作用这几乎是不可能的。例如即使是同一个信息，无论怎么看也是对某个人有信息价值，但对另一个人就没有。在情报理论上要排除这种随个人情绪兴趣等而改变的主观因素，而只着眼于客观因素。那么就要考虑一下信息价值的客观因素是什么。

“了解大变革”这一事象不能成为信息，但“接受大变革”这一事象是最好的信息材料。前者是经常遇到的事象，相反的后者是非常珍贵，稀有的事象。在这里，为了简单假设分别用 E 和 F 表示前后两个事象，如果把常有的和稀有的事情，用 E, F 的相应特征值 $p(E), p(F)$

$$p(E) > p(F)$$

这样来表示，就会排除了被接受信息者的主观因素，进而可以抓住信息价值的客观因素。

为满足这种要求，提出了事象概率的概念。即把 $p(E)$ 看作是 E 的概率。这样一来，从决定信息价值许多因素中，去掉决定于每个个人的主观东西，建立了概率这一通用的“尺度”※。

在这个基础上，可进一步把信息价值，简略地表示为
信息价值主要因素 = 概率主要因素 + 非概率主要因素

“毛泽东逝世”这一事象具有很大的信息价值，如果根据这个想法该情况几乎被认为取决于非概率主要因素。因为

※ 即数学上的概率测量

人逝世事件的概率是 1 并不是稀有的事。这种情况下，概率主要因素对决定信息价值是不起作用的。象这样，仅从概率出发的观点，不能作出说明的例子较多。可是为保持理论的客观性，不得不这样规定。

概率是表示产生事象的难易程度的数值，必然会满足如下的克尔莫哥洛夫公理。¹⁾

设 $p(E)$ 为事象 E 的概率

$$(i) \quad 0 \leq p(E) \leq 1, \quad p(\phi) = 0, \quad p(U) = 1$$

$$(ii) \quad \text{如 } E \subset F \text{ 则 } p(E) \leq p(F)$$

$$(iii) \quad \text{如 } E \cap F = \phi \times \text{ 则 } p(E \cup F) = p(E) + p(F)$$

但是， ϕ 是空事象， U 是全事象，当然有 $\phi \subset E \subset U$ 。

首先，考虑某事象是由若干基本事象构成的和事象。例如把物质细分下去，象分到原子那样不能再分割的最小单元。如把某个事象最终分成几个基本单元时，就多半成为简单的事情。这些基本单元叫做基本事象，某事象 A 包含的基本事象数写成 $n(A)$ ，当然，全部基本事象是相互无关的。

假设要具体给出某事象 E 的概率 $p(E)$ 数值时，可以使用统计的方法。即取 r 为产生事象 E 的实例（所谓资料）总数，同样取 s 为 $\bar{E}$ 资料的总数，假设全部资料数为 $n = r + s$ 时，称 r/n 为相对次数。这时，把概率定义为 n 充分大时相对次数的极限值

$$p(E) = \lim_{n \rightarrow \infty} \frac{r}{n}$$

※ $E \cap F = \phi$ 时 E 和 F 是相互无关，或者说是相互否定。另外，(iii) 被称为概率的加法定理。