

現代应用数学丛书

信 息 论

〔日〕喜安善市 室賀三郎 著

上海科学技术出版社

現代应用数学丛书

信 息 论

〔日〕市善安喜 著
室賀三郎

李 文 清 譯
陆 志 剛 校

上海科学技术出版社

5

內 容 提 要

本书是日本岩波书店出版的现代应用数学丛书之一的中译本,介绍 Shannon 的信息传递理论。全书共七章,第一、二两章介绍信息的概念及信息量的数学表达式,第三章讨论离散的信息源,四、五、六章分别讨论无噪音及有噪音的离散信道及编码方法,最后一章讨论了连续的信息源。本书可供高等学校有关专业作为教学参考书,也可供工程师及研究人员参考。

现代应用数学丛书

信 息 论

原书名 情报理论

原著者 [日] 喜安善市
室賀三郎

原出版者 岩波书店 1957

译者 李文清

校者 陆志刚

上海科学技术出版社出版

(上海瑞金二路450号)

上海市书刊出版业营业许可证出 093 号

新华书店上海发行所发行,各地新华书店经售

商务印书馆上海厂印刷

*

开本 850×1168 1/32 印张 3 22/32 字数 85,000

1962年10月第1版 1962年10月第1次印刷

印数 1-4,500

统一书号: 13119·480

定 价: (十四) 0.64 元

出版說明

这一套书是根据日本岩波书店出版的“現代应用数学讲座”翻譯而成。日文原书共 15 卷 60 册,分成 A、B 两组,各編有序号。現在把原来同一題目分成两册或三册的加以合并,整理成 42 种,不另分組編号,陸續翻譯出版。

这套书涉及的面很广,其内容都和現代科学技术密切有关,有一定参考价值。每一本书收集的資料都比較丰富,而叙述扼要,篇幅不多,有利于讀者以較短時間掌握有关学科的主要内容。虽然,这套书的某些观点不尽适合于我国的情况,但其方法可供参考。因此,翻譯出版这一套书,对我国学术界是有所助益的。

由于日文原书是 1957 年起以讲座形式陸續出版的,写作時間和篇幅的限制不可避免地会影响原作者对内容的处理,为了尽可能地减少这种影响,我們在每一譯本中,特請譯者或校閱者撰写序或后記,以介紹有关学科的最近发展状况,并对全书內容作一些評价,提出一些看法,結合我国情况补充一些資料文献,在文內过于簡略或不足的地方添加了必要的注釋和改正原书中存在的一些錯誤。希望这些工作能对讀者有所帮助。

承担翻譯和校閱的同志,为提高书籍的质量付出了巨大劳动,在此特致以誠摯的謝意。

欢迎讀者对本书提出批評和意見。

上海科学技术出版社

现代应用数学丛书

书 名	原作者	译者
代 数 学	张永昌吉等	熊全淹
几 何 学*	矢野健太郎	孙泽瀛
复 变 函 数	功力金二郎	刘书琴
集合·拓扑·测度*	河田敬义	赖英华
泛 函 分 析*	吉田耕作	程其襄
广 义 函 数*	岩 村 联	杨永芳
常微分方程	福原满洲雄	张庆芳
偏微分方程*	南云道夫	钱端壮
特 殊 函 数*	小谷正雄等	钱端壮
差 分 方 程	福田武雄	穆鸿基
富里哀变换与拉普拉斯变换*	河田龙夫	钱端壮
变分法及其应用*	加藤敏夫	周怀生
李 群 论*	岩堀长庆	孙泽瀛
随 机 过 程*	伊 藤 清	刘璋温
回轉群与对称群的应	山内恭彦等	张质贤
結晶統計与代数	伏見康治	孙泽瀛
偏微分方程的应	犬井鉄郎等	杨永芳
微分方程的近似解法	加藤敏夫等	王占瀛
数值計算法	森口繁一等	閻昌齡
量子力学中数学方法	胡永振一郎	周民强
工程力学系統*	近藤一夫等	刘亦奇

书 名	原作者	译者
非綫性振動論	古 屋 茂	呂紹明
力 学 系 与 映 射 理 論*	岩 田 义 一	孙泽瀛
平 面 彈 性 論*	森 口 繁 一	刘亦奇
有限变位彈性論	山 本 善 之 夫	刘亦奇
变 形 几 何 学*	近 藤 一 夫	刘亦奇
塑 性 論*	鷺津文一郎	刘亦奇
粘性流体理論*	谷 一 郎	刘亦奇
可壓縮流体理論*	河 村 龙 馬	刘亦奇
网 絡 理 論	喜安善市等	陆志刚
自 动 控 制 論	喜安善市等	翟立林
回 路 拓 扑 学	近 藤 一 夫	張鳴鏞
信 息 論	喜安善市等	李文清
推 断 統 計 理 論	北 川 敏 男	李賢平
統 計 分 析*	森 口 繁 一	刘璋温
实 驗 設 計	増山元三郎	刘璋温
群 体 遺 傳 学 的 数 学 理 論*	木 村 資 生	刘祖洞
博 奕 論	官 澤 光 一	張毓春
綫 性 規 划	森 口 繁 一	刘源强
經 济 理 論 中 的 数 学 方 法	安井琢磨等	談祥柏
随机过程的应用*	河 田 龙 夫	刘璋温
計 算 技 术	高 桥 秀 俊	姚 晋
穿孔卡计算机	森 口 繁 一	刘源强

注：有*者已在1962年7月以前出版。

目 录

出版說明

第1章 序論	1
§ 1 历史背景	1
§ 2 信息是什么?	4
§ 3 如何进行通信	6
第2章 信息量与熵	9
§ 4 信息量的数学表达式	9
§ 5 离散随机变数的熵的性质	14
§ 6 連續随机变数的熵的性质	17
第3章 离散的信息源	21
§ 7 Markoff 过程及 Shannon 綫图	21
§ 8 作为信息源的自然語	27
§ 9 信息源的熵	30
§ 10 信息源的冗长度	33
第4章 无噪音的离散信道	37
§ 11 作为变换器的发射机及接收机的功能	37
§ 12 无噪音离散信道的傳輸速率和通信容量	39
§ 13 編碼基本定理	43
§ 14 最佳編碼法	45
第5章 有噪音的离散信道	50
§ 15 曖昧度, 散布度及傳輸速率	50
§ 16 关于通信容量的基本定理	53
§ 17 通信容量的表現	60
第6章 有噪音的离散信道的編碼	68
§ 18 編碼的概念	68
§ 19 无多余的組織碼組	73
§ 20 非无多余的組織碼組	76
§ 21 組織碼的譯碼法及小长度組織碼組的例	81

§ 22 非群碼組	87
第 7 章 連續的通信系統	93
§ 23 連續信息及噪音的表示	93
§ 24 連續信道的傳輸速率及通信容量	96
§ 25 具有相加性噪音的連續信道	98
参考文献	104
校后記	108

第1章 序 論

本章首先叙述信息論的历史与发展过程，然后說明信息論中的信息、通信等基本概念。

§1 历史背景

人类和其他动物的明显区别是具有一种本能以外的通信手段。通信手段不限于声音，还有文字、繪画、雕刻、印刷等方面。其他动物不是沒有通信手段，但不象人类通信发展到那样的高度。准确、无誤和有效的通信是人类社会不可缺少的东西。虽然通信技术如光通信，电通信已經非常发达，但是信息通信的本质，以及妨害通信的噪音，在概率論的基础上加以理論的探討，还是到了廿世紀的中叶才开始的。

自然，現在完善的信息論，是人类多年关于通信本质研究的結果，是逐漸形成的理論，与通信技术密切相关，它的迅速发展則是由电通信的出現所促成的^[1]。举一个熟知的例子，26个英文字母中，E是出現最頻繁的字母，Morse 电报就分配它短的碼，这样就縮短了英文信息傳送的平均時間（但是，妙的是用 Morse 碼对日文來說，就沒有这样的考虑）。其次，我們回顾一下自从电通信发展以来信息論的形成。

現在，家庭用的收音机使用着振幅調制^①，在1922年 Bell 电话研究所的 J. R. Carson 就指出，它是具有一定的頻帶寬度，并且开始明确了边帶的概念。其后，发明了現在长途电话使用单边

① 調制这一概念，參看文献[2]。对于不熟悉电气通信的讀者可以略去本节以下的內容，不会影响以后的理解。

帶通信方式，討論了調頻，指出了調制指數小的調頻方式對寬頻帶的帶寬點的不實用性。此外，與信息論有關的研究是關於噪音波形的統計研究，這些都是最早的研究。

1924年 H. Nyquist 及 K. Küpfmüller 兩人獨立地指出，電信信號的傳輸速率與信道帶寬有比例關係。這種想法在 1928 年被 R. V. L. Hartley 推廣了。Hartley 的想法也可以說是現在信息論的基礎，他把信息考慮為代碼或是單語的序列，把它所代表的語義當做次要的而不予考慮。於是從 S 個代碼序列中選 N 個碼即構成 S^N 個可能的信息，他指出“信息量 H ”定義為 $H = N \log S$ 是合理的。傳輸一定的信息量時，帶寬及傳輸時間的乘積為常數，他論述了當帶寬愈窄，傳輸時間愈長，帶寬愈大，傳輸時間愈短。但 Hartley 的理論，沒有考慮到噪音和概率統計，現在的理論比他的想法可以說進步多了。

在 1940 年，岡田和藤木兩人曾指出，信息的傳輸與能量的傳輸沒有直接關係，他們曾討論了 Hartley 信息量的各種具體例子。

信息論的進一步發展，是由於各種各樣新通信方式的出現，作為通信的本質即信息的傳輸，要求在更廣泛的基礎上加以研究，又因為統計數學和妨礙通信的噪音理論的發展，信息論只有考慮到噪音及利用了數學統計的處理方法，才形成了理論。從社會環境來說，在第二次大戰期間，為了製造新武器，動員了許許多數學家、物理學家開展了以電子學為中心的軍事科學研究，這也是促使信息論發展的重要原因。

關於噪音的理論，V. D. Landon 在 1936 年及 1941 年曾討論過噪音的波形，指出它的峰值與有效值之比，不論噪音的帶寬如何，常為 3.6 到 4。K. Fränz 在 1940 年討論了射頻接收機的檢波器問題，1943 年 P. A. Mann 及 1947 年瀧保夫對噪音波形作了細緻討論，都是關於噪音波形的理論。1945 年及 1948 年 Bell

研究所的 S. O. Rice 曾討論了随机噪音的統計性质, 并对噪音与正弦波迭加的情况, 作了詳細的数学分析^[3]。

关于通信方式的发展, E. H. Armstrong 在 1936 年最初使用了調頻通信装置, 調制指数很大时, 帶寬虽然变大, 但有抑制噪音的作用。随着在第二次大战时期, 脉冲技术的急速发展, 脉冲幅度調制(PAM), 脉冲宽度調制(PWM), 脉冲頻率調制(PFM), 脉冲位置調制(PPM)等各种通信方式发展起来, 其中特別重要的是对噪音抑制效果最大的脉冲編碼調制(PCM)。早在 1939 年, H. Reeves 曾对他的原理申請专利, 1948 年 Bell 电话研究所克服了技术上的困难完成了制造。此种脉冲編碼調制, 后来对信息論的形成充当了重要角色。

在二次大战时期和战后, 电子計算机和各种武器的自动控制的蓬勃发展, 刺激了英美諸国在这方面的研究, 在信息傳輸形式的广泛基础上, 把噪音也考虑在內的研究才探明了信息通信的本质。

1946 年 D. Gabor 及 1940 年今堀, 討論了信号的长度、頻率的不确定性。其后 Szilard, Weiner, Brillouin 等人討論了信息量与热力学第二定律的关系, 高桥又討論了信息論与統計力学的关系。

Wiener 在战时曾研究抑制噪音的滤波器設計問題^[4]。把信号及噪音看做随机过程, 做出了信号波形与信号上重迭噪音波形之間最小均方誤差的滤波器設計。所化的代价是時間的推迟, 推迟愈多, 錯誤概率愈小。他还研究了預測器, 給出了比輸入信号早一段時間的信号值具有均方誤差的預測器設計。時間愈提前, 錯誤概率愈大。

此后, 李郁荣用电子綫路实现了 Wiener 相关器和預測器, 能把信号噪音比为 -20 分貝的埋藏在噪音中的周期信号用相关器提取出来, 能够改善信号噪音比达 $+20$ 分貝左右^[5]。

第二次世界大战开始时，Wiener 与当时哈佛大学的医学家 A. Rosenblueth 试图把现代各学科中通信及自动控制的基本问题综合成一门新的学问，他们在很长一段时期内，同各方面的学者举行了学术讨论会，终于在 1947 年命名为控制论 (Cybernetics)^[6]。此种气氛和刺激对于近代信息论的发展有不可否认的影响。在这种气氛中，除 Wiener 的相关器及预测器外，还对信息量提出了新的定义，这就是 R. A. Fisher 的不同定义^[7]，用了统计力学的熵的形式。

1948 年 Bell 电话研究所的 O. E. Shannon 提出了最完善的统一的信息论^[8]。即把 Hartley 的信息量在概率论更广泛的基础上加以定义，同时发展了有噪音时信息传输的理论。比他较早一些时间，W. G. Tuller 曾讨论了有噪音干扰的信道的信息传输量问题^[9]。Shannon 利用熵的形式，导入了通信容量这一新的重要概念。由于 Shannon 划时代的贡献，使信息论很快地形成了雏形。其后，经过许多学者的努力，才使这个理论渐趋丰富与完善。

通常“通信理论”与“信息论”几乎是同义语。重要的区别在于后者与控制论有相同的含义，而前者是控制论中与通信有关的一部分而已。从另一种观点说，所谓“信息论”，即是通信系统中用近代数理统计学的方法探究信息传输问题的科学。关于信息本质的科学从广义来说，除调制问题、传输理论、频谱分析问题外，有人从物理学的角度认为热力学、量子力学也有信息论的本质问题。这本小册子以 Shannon 的研究为中心说明信息理论，可以认为信息本质是最基本的，由这种角度写的信息论专著在国内外还不多。

§2 信息是什么？

“信息” (information)，在辞典（如日文广辞苑）上的意义是所观察事物的知识，“通信”是两人以上互相交换信息或意志。如果

不是交换，就叫做单方传输了。但在定义信息量时，“信息”、“通信”的意义就与通常的涵意不同，必须注意使用了抽象化的概念。“通信”放在下一节讨论，本节只叙述“信息”。

同样一件特定的通信内容，对不同的接收者来说，常引起不同的情感。例如 1956 年埃及纳赛尔总统宣布苏彝士运河国有化时，英法两国感到挫伤，阿拉伯诸国莫不称快。从信息论来说，对不同的人，由于接收的意义不同，不能看做与信息量有关。信息的接受是把事情本身、时间和内容都撇除，使之抽象化，在各种情况下，问题在于代表不同的知识。从这个角度看来，在苏彝士运河国有化问题上，重要的是事情发生以后苏联对中东的影响，英国在地中海霸权的丧失，纳赛尔总统的盛衰等事情发生的可能性。假定可以猜到纳赛尔的盛衰，此外没有其他的可能性，那末苏彝士运河的国有化就算不得什么大信息。若一切的可能性都相等，那就无从猜测国际形势的变化，于是这就成为国际形势中的一个大信息。

收音机中人们喜欢收听的“20 个扉”^①恰好适合信息量的概念。猜谜的人提出 20 个问题获得信息，来判断谜底是什么。后提的问题应和以前提的尽量不同，并且提的问题要最大可能地与谜底有关，这样才能有效地利用这 20 个信息所构成的信息量。

不只是“简单地有多少个可能性”，而更重要的是“这样或那样事件有多少程度的可能性”。Hartley 只是由“有多少个可能性”这一点出发建立信息论，而 Shannon 理论的出发点是“这样或那样的事件有多少程度的可能性”，这里使用了统计概率论，这是近代信息论发展的基础。

信息论最重要的一点是，当通信文已经知道的时候，注重的不在此特定的通信文本身的意义及内容，而是此特定通信文以外可

① “20 个扉”是日本民间流行的一种猜谜形式。猜的人可以向出“谜”的人提出 20 个问题，故称 20 个扉。出谜的人要诚实回答问题。——译者注

能产生多少种通信文，又各通信文实际上有多少出現概率。換言之，重要的不是“我們述說的事物”，而是“我們所能述說事物的大概程度”。

我們用語言交換信息，即进行通信，不論使用自然語或人工語，如果文章的长度相同，則含有不同意义的单字越多的文章，它的信息量愈大。例如用 48 个日文字母写的文章，与夹杂着难解的汉字所写的文章相比較，不論讀和写，前者因为字母少因之很容易。反之，后者的信息量較大。然而夹杂着汉字的文章，汉字数以万計，但大部分的汉字出現的可能性不大，如果限制使用一部分汉字，則不費很大的劳力也能写或讀。

在电报中所用的代碼是一种广义的人工語，亦有相同的情况，单位碼的种类多的通信系統，同一時間所送的信息比較多。例如对于脉冲編碼調制，不管其技术上的难易，用三元碼通信比用二元碼通信，在一定時間內，会送出更多的信息。

§3 如何进行通信

我們如何进行信息通信呢？例如我們給远方的人打电报，把电文写在紙上交給邮电局，电报員把文字变成 Morse 电碼电流（或电傳打字电碼电流），通过綫路傳送出去。在綫路中途有时受到噪音的干扰而产生錯字，收信局再把电碼电流譯成文字，印刷之后，送交收信人。

又如长途電話，打電話人的声音經過炭质送話器变成声音电流送到长途電話局，換成載波电流經過很長的綫路，到达目的地的電話局，再換成声音电流，經過听電話人的電話耳机把声音电流变成声音，再傳到听電話的人。此种情況下，在綫路、交換机、中繼器內可能混入噪音。

信息論是把以上所述各个实际出現的情况加以抽象化、理想

化,成为通信系统模型的数学理论。由图 3.1 看出,这个模型分成下列五部分。

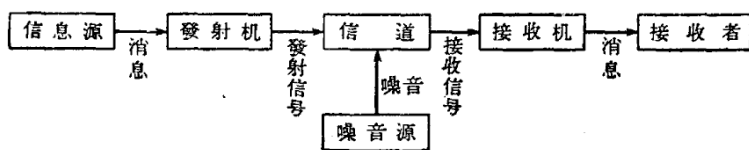


图 3.1

(1) **信息源** (information source) 产生要传输信息的来源称为信息源,从信息源产生的信息叫做消息 (message)。如打电报的人的大脑即是信息源,写在纸上的文字即是消息。在电通信中,电报、电传打字中所用的字母序列 (包括标点在内)、电话或无线电广播的声压变化的时间函数、电视中对应的时间及空间坐标的函数、黑白电视中的二元坐标系所表示的点 (x, y) 处光强度的时间 t 的函数 $f(x, y, t)$, 就都是消息的例子。

(2) **发射机** (transmitter) 把消息变成适合于信道传输的信号装置叫发射机。例如 Morse 电报的情况,发射机就是把电文变成点、划、空白等电码序列所用的电报机,在电话的情况下,发射机是把声压按比例变成电流的炭质送话器。但是,从信息论看来,声压、电流等表示信息的物理要素间的变化,不是重要的事,重要的是在通报时能把字母或电码序列变成别的电码序列,从这种性质着眼,发射机有时亦称为变换器。例如在 Morse 电报情况下,字母变换成电流的技术问题,从信息论的立场看来并不重要,重要的是字母序列的适当限制,例如不使字母空白连续发生两次,不使单语空白与字母空白连续发生的限制,即具有一定时间长度的电码序列的变换是很重要的。

(3) **信道** (channel) 是信号由发射机传输到接收机时所使用的媒介。例如电线、同轴电缆、一定频带宽度的无线电波,在光

通信中的光綫等都是媒介。信号在信道傳送的中途，能受到噪音的干扰，由发射机发出的信号，不一定正确地傳送到接收机。此外，信号一般受到的失真，在接收机中普通是能够还原的，可是，本质上妨碍信息傳送的噪音却与这种失真是不同的。

(4) 接收机 (receiver) 是把信道傳送来的信号与噪音的混合接收下来，把原来的消息加以还原，它进行和发射机相反的操作。此时电碼的变换器的作用非常重要。

(5) 接收者 (destination) 即是接收消息的人或物。

以上的通信的模型，不仅适用于电气通信系統，即使象神經系統等等广义的通信系統也是适用的，它們只有形式上的差別，根本上是相同的。信息交换是由相对方向的两个通信系統組成，各个方向的通信系統即是上述的模型。这模型对观察自然界也是形成一个信息观察系統，并且在某种程度上也适用于社会現象中各种形式的信息傳輸。

为了数学上处理的方便，数量通常可分为离散的及連續的两种情况。电报的电文是由离散的信息源发生的离散的消息，Morse 电报的点、划信号是离散的信号，其傳送的信道亦称为离散的。但是电话的消息和信号都是連續的，它的信息源、信道皆称为連續的。但处理的数量，有时同时有离散的及連續的，即所謂混合型，这在数学上很困难，实际上也不重要。

第2章 信息量与熵

关于信息量的数学表达式,根据信息量所应满足的条件,我們自然地导出了熵的形式,然后再阐述熵的性质。

§4 信息量的数学表达式

前章阐述了信息的本质,但信息量的数学公式如何表示呢?以后将要談到,由信息源产生的消息一般可以用 Markoff 过程表示,此处只談构造最简单的信息源,即消息完全不受过去影响的信息源。

为了理解上的便利,現在用概率論的語言加以說明。設有 n 个不同事件(即信息論中所謂消息使用的字母),从中任选一个,选择反复进行构成偶然事体(即信息源),且前一选择与后一选择无关,各事件出現的概率为 $p_1, p_2, \dots, p_n$ 。此种事体的信息量以 $H(p_1, p_2, \dots, p_n)$ 表之,希望有下列性质:

(i) 信息量 $H(p_1, p_2, \dots, p_n)$ 为 $p_1, p_2, \dots, p_n$ 各自变量的連續函数。

(ii) 当 $p_i (i=1, 2, \dots, n)$ 都是 $\frac{1}{n}$ 时, H 是 n 的單調遞增函数。

(iii) 一个事件的选择分成两个步驟时,未分之前的 H 即是既分之后 H 的加权和。

这个(iii)的意义如图 4.1(a) 所示,三个事件的概率为 $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{3}$, $p_3 = \frac{1}{6}$ 。变成图 4.1(b) 时,先取概率为 $\frac{1}{2}$ 的两个事件中的任一个,第二阶段分别选择概率为 $\frac{2}{3}, \frac{1}{3}$ 的事件。选择三个事

件的概率与分两个步骤无关,各个概率最后是同样的,此时希望下式成立:

$$H\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) = H\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{2} H\left(\frac{2}{3}, \frac{1}{3}\right).$$

加权系数 $\frac{1}{2}$ 是由于在第二选择阶段,从时间上說,只成一半了。

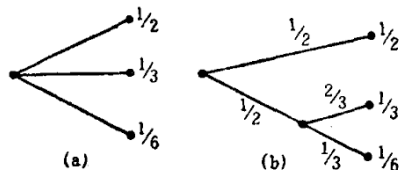


图 4.1

满足上列三个条件的 H , 必然是统计力学的熵的形式。首先考虑 p_i ($i=1, 2, \dots, n$) 都是相等的情况,置

$$H\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = A(n).$$

由条件 (iii), 等概率下发生的 s^α 个事件, 等于概率为 $\frac{1}{s}$ 的 s 个事件分解为 α 阶段重复的选择, 如图 4.2(a), 于是得出

$$A(s^\alpha) = \alpha A(s).$$

同样对于与 s 不同的 t , 得

$$A(t^\beta) = \beta A(t).$$

当 β 适当大时, 选取满足下式的 α :

$$s^\alpha \leq t^\beta < s^{\alpha+1}.$$

取对数, 以 $\beta \log s$ 除它, 得

$$\frac{\alpha}{\beta} \leq \frac{\log t}{\log s} < \frac{\alpha}{\beta} + \frac{1}{\beta}. \quad (4.1)$$

另一方面, 由条件 (ii), 得公式

$$\alpha A(s) \leq \beta A(t) < (\alpha+1) A(s),$$

上式除以 $\beta A(s)$ 得

$$\frac{\alpha}{\beta} \leq \frac{A(t)}{A(s)} < \frac{\alpha}{\beta} + \frac{1}{\beta}.$$

把上式和 (4.1) 式合起来, 设 ε 为任意小正数, α, s, t 相当大时, 取