“十三五”普通高等教育规划教材

信息与编码理论

张长森　郭 辉　主编

提供电子教案、习题解答

http://www.cmpedu.com

机械工业出版社
CHINA MACHINE PRESS

“十三五”普通高等教育规划教材

信息与编码理论

张长森　郭　辉　主编
邓　超　李宝平　王小旗　参编

机械工业出版社

本书重点介绍信息论的基本知识，以及常用信源编码和信道编码技术的原理与实现方法。全书共6章，主要内容包括绪论、信息论基础、信源编码、线性分组码、BCH码和RS码，以及卷积码。本书面向工程应用的需要，在介绍了信息论基本概念和必要数学知识的基础上，重点讲解常见的信源编码和信道编码技术的基本原理、物理意义和实现方法，内容全面，例题与图表丰富，便于教学与自学。

本书可用作高等院校通信工程、电子信息工程、计算机科学与技术等专业的本科生或研究生教材和参考书，也可供相关专业的科研和工程技术人员参考。

图书在版编目（CIP）数据

信息与编码理论/张长森，郭辉主编. —北京：机械工业出版社，2019.5

“十三五”普通高等教育规划教材

ISBN 978-7-111-62845-3

Ⅰ.①信… Ⅱ.①张…②郭… Ⅲ.①信息论-高等学校-教材②信源编码-高等学校-教材 Ⅳ.①TN911.2

中国版本图书馆CIP数据核字（2019）第101194号

机械工业出版社（北京市百万庄大街22号 邮政编码100037）

策划编辑：李馨馨 责任编辑：李馨馨

责任校对：陈 越 郑 婕 责任印制：张 博

三河市国英印务有限公司印刷

2019年7月第1版第1次印刷

184mm×260mm · 12.75印张 · 309千字

0 001—2 500册

标准书号：ISBN 978-7-111-62845-3

定价：39.80元

电话服务 网络服务

客服电话：010-88361066 机 工 官 网：www.cmpbook.com

010-88379833 机 工 官 博：weibo.com/cmp1952

010-68326294 金 书 网：www.golden-book.com

封底无防伪标均为盗版 机工教育服务网：www.cmpedu.com

前　言

1948 年，C. E. Shannon 发表了他的开创性论文《通信的数学理论》，宣告了信息论学科的诞生。信息论是研究信息传输和信息处理的科学，是现代信息与通信技术的理论基础。信息论及其衍生的编码理论和技术既是科学理论又是工程应用知识，对实际通信系统的设计与实现产生着深刻的影响，并已经渗透到其他领域。

编码理论和技术在本质上是高度数学化的，对其深入理解需要掌握丰富的通信理论、概率论和近世代数的背景知识。为了帮助读者尽快理解和掌握常用的重要编码和译码技术，本书在讲解信息论的基本概念之后，使用了最少的数学基础知识，重点对常用的信源和信道编译码理论和技术进行了深入的讲解。本书内容兼顾知识性和实用性，联系工程实践，强调物理意义，结构合理、概念清晰、示例丰富准确，可用作通信工程、电子信息工程、计算机科学与技术等专业的本科生或研究生教材和参考书，也可供相关专业的科研和工程技术人员参考。

全书共分 6 章。第 1 章简要介绍了数字通信系统模型、信源编码问题和信道编码问题；第 2 章介绍了信息论的基础知识和一些重要结论；第 3 章讲解了常见的无失真和限失真信源编码方法；第 4 章全面讲解了线性分组码和循环码的相关知识；第 5 章讲解了有限域的基本知识以及 BCH 码和 RS 码的编译码原理；第 6 章讲解了卷积码的编码原理、维特比译码算法和 Turbo 码的基本概念。

本书由河南理工大学张长森、郭辉、邓超、李宝平、王小旗编写。张长森编写了第 1 章，邓超编写了第 2 章，王小旗编写了第 3 章，郭辉编写了第 4、5 章，李宝平编写了第 6 章。在本书的编写过程中，得到了河南理工大学研究生院和机械工业出版社的大力支持，在此表示衷心的感谢。

由于编者水平有限，书中疏漏和不当之处在所难免，敬请读者批评指正。

目　　录

第1章 绪 论

1.1 数字通信系统的模型

各种数字通信系统（包括存储系统），如电报、电话、图像、计算机、导航、雷达等，都可以归结成如图 1-1 所示的模型。

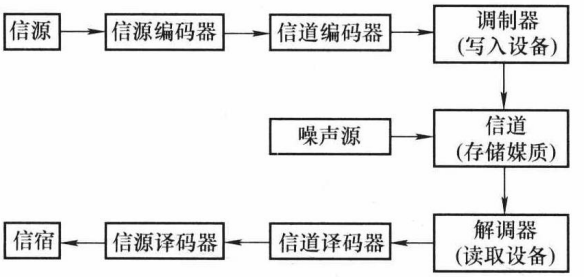

图 1-1 数字通信系统模型

图中信源编码器把信源发出的语音、图像或文字等消息转换成二进制（或多进制）形式的信息序列。有时为了提高传输有效性，还会去除一些与信息传输无关的冗余度来实现数据压缩。

信道编码器则为了抵抗传输过程中的各种干扰，改善误码率性能，往往会在传输的信息序列中人为地增加一些冗余度，使其具有自动检错或纠错的功能。

调制器的功能是把信源编码器输出的信息序列变换成适合信道传输的电信号，然后送入信道传输。解调器则执行与调制器相反的功能，将接收到的电信号还原为信息序列。由于信号在信道传输过程中会受到信道特性、噪声和干扰信号的不利影响，因此在解调器输出的序列中会出现误码的情况。

接下来，解调器的输出序列会送入信道译码器。信道译码器会对接收到的序列进行检错或纠错。然后通过信源译码器恢复成消息送给信宿。

从图 1-1 可以看出，通信的目的是要把对方不知道的消息可靠地传送过去，而消息中真正有意义的部分是信息，因此通信的本质是要实现信息的传输。信息论是研究信息的传输、存储和处理的科学，可以称为通信的数学理论。信息论研究的主要问题包括信源编码和信道编码问题，而由这两个理论问题又延伸出很多实用的编码和译码算法。

1948 年，贝尔实验室的科学家香农（Shonnon）在《通信的数学理论》一文中讨论了通信的基本问题，得出了几个重要的结论。其核心是：在通信系统中采用适当的编码后能够实现高效率和高可靠性的信息传输，并得出了信源编码定理和信道编码定理。从数学角度看，这些定理是最优编码的存在定理；但从工程角度看，这些定理不是构造性的，不能从定理结果直接得出实现最优编码的方法。然而，这些定理给出了编码的性能限，并阐明了通信系统中各个因素之间的关系，为寻找最佳通信系统提供了重要理论依据。

1.2 信源编码问题

1. 无失真信源编码

对于离散信源，当已知信源符号的概率分布时可以计算信源的熵，用它可以表示每个信源符号平均承载的信息量。信源编码定理表明必然存在一种编码方法，使得代码的平均长度可以任意接近但不能低于信源的熵，而且还阐明为了达到这一目标，应该使得概率与码长匹配。

从无失真信源编码定理出发，1948 年香农在论文中提出并给出了一种简单的编码方法，即香农编码。1952 年，R. M. Fano 提出了费诺码。同年，D. A. Huffman 提出了一种编码方法并证明了它是最佳码，称为霍夫曼码。霍夫曼码是有限长度的块码中最好的码，其代码总长度最短。但是，霍夫曼码在实际应用中存在一些块码和变长码所具有的缺点：首先，信源的概率分布必须精确地测量，如果略有变化，就需要更换码表；其次，对于二进制信源，常需要多个符号合起来编码才能取得好的效果。

针对霍夫曼码在实用中的局限，出现了一种称为算术码的非块码，它是从整个序列的概率匹配角度来进行编码的。这种概念也是由香农首先提出的，后经许多学者改进逐渐进入实用阶段。1968 年前后，P. Elias 发展了香农—费诺码，提出了算术编码的初步思想。1976 年，J. J. Rissanen 给出和发展了算术编码，1982 年他和 G. G. Langdon 一起将算术编码系统化，省去了乘法运算，使其更为简单，从而易于实现。

如果离散信源符号的概率分布未知，或是对于不确知的信源进行有效编码时，上述方法就无能为力了，因此人们希望能有一种适用于各类概率特性信源的编码方法。通用编码就是在信源统计特性未知时可以进行编码且编码效率很高的一种码。1977 年，A. Lempel 和 J. Ziv 提出了一种语法解析码，称为 LZ 码。1978 年，他们又提出了改进算法 LZ77 和 LZ78。1984 年，T. A. Welch 以 LZ 编码中的 LZ78 算法为基础设计了一种实用的算法，称为 LZW 算法。LZW 算法保留了 LZ78 算法的自适应性能，压缩效果大致相同，并且逻辑性更强，易于硬件实现，价格低廉，运算速度快，目前作为一种通用压缩方法广泛应用于二进制数据的

压缩。

2. 限失真信源编码

无失真信源编码只适用于离散信源，对输出模拟信号的连续信源并不适用。因为连续信源输出信号的每个样值所载荷的信息量是无限大的，所以用有限长度的信息序列进行编码时必然导致失真。不过，作为信宿的人或机器都存在一定的灵敏度和分辨力，超过灵敏度和分辨力所传送的信息是毫无价值的，也是完全没有必要的，故而当失真在某一限度以下时是不影响正常通信的。例如，语音信源当量化分层超过256级时人耳就很难分辨，所以没有必要在量化时超过256级。对图像信源亦是如此，人们看电影时，可以充分利用人眼的视觉暂留效应，当放映速度达24张/秒时，人眼就能将离散的照片在人脑内反映成连续画面，因此大大超过24张/秒的放映速度是没有意义的。

限失真信源编码的研究相对于信道编码和无失真信源编码落后约十年左右。1948年，香农的论文已体现了关于率失真函数的思想。1959年，他发表了《保真度准则下的离散信源编码定理》，首先提出了率失真函数及率失真信源编码定理。1971年，T. Berger的《信息率失真理论》是一本较全面的论述有关率失真理论的专著。率失真信源编码理论是信源编码的核心问题，是频带压缩、数据压缩的理论基础，直到今天它仍是信息论研究的课题。

连续信源的信号编成代码后就不能无失真地再恢复成原来的连续值，此时只能根据率失真理论进行限失真编码，因此限失真编码实际上就是最佳量化问题。最佳标量量化常不能达到率失真函数所规定的$R(D)$值，因此人们后来又提出了向量量化的概念，将多个信源符号合成一个向量并对它进行编码。从理论上讲，某些条件下用向量量化来编码可以达到上述的$R(D)$值，但在实现上还是非常困难的，有待进一步改进。1955年，P. Elias提出了预测编码方法，利用前几个符号来预测后一个符号的值，然后将预测值与实际值之差即预测误差作为待编码的符号，这样得到的符号间的相关性就大为减弱，从而可提高压缩比。另一种限失真信源编码方式是变换编码，该方法通过样值空间的变换，例如从时域变到频域，在某些情况下可以减弱符号间的相关性，从而取得良好的压缩比。

1.3 信道编码问题

信道编码的历史仍然可以追溯到香农于1948年出版的那篇著名论文。香农发现，任何通信信道（或存储信道）都有自己的信道容量C，其单位是比特/秒（bit/s），其物理意义为：当一个通信（或存储）系统的信息速率R小于信道容量C时，通过使用信道编码技术，是有可能使得系统输出的错误概率任意小的。但是，香农并没有指出如何找到合适的码，他的贡献主要是证明了这些码的存在并定义了它们的作用。其后在整个20世纪50年

代，人们一直在努力寻找能够得到任意小错误概率且具有明确构造方法的编码方案，但是进展缓慢。20 世纪 60 年代，人们不再痴迷于这个宏伟的目标，信道编码的研究方向开始确定下来，并逐渐分成了两个方向。

1. 线性分组码

第一个研究方向是线性分组码，该类码具有严格的代数结构，并且主要采用分组码的形式。历史上第一种线性分组码是 1950 年 R. W. Hamming 发明的可以纠正 1 个错误的汉明码。其后不久，Muller 提出了一种可以纠正多个错误的编码方法（1954 年），紧接着 Reed 给出了该码的一种译码算法（1954 年）。无论是汉明码还是 Reed - Muller 码，其性能距离香农给出的好码都非常大。之后，学者们付出了大量的辛勤研究工作，但是一直没有找到更好的码，直到十年之后，Bose、Ray—Chaudhuri（1960 年）和 Hocquenghem（1959 年）才提出了一类可以纠正多个错误的编码方法，即 BCH 码。接着，Reed 和 Solomon（1960 年）以及 Arimoto（1961 年）分别独立地发现了一类适用于非二进制信道的编码方法，即 RS 码。

BCH 码的发现引起了一系列关于实用方法的研究工作，人们纷纷通过设计软件或硬件来实现编码器和译码器。第一种较好的译码算法是 Peterson 于 1960 年发现的方法，随后 Berlekamp（1968 年）和 Massey（1969 年）发现了一种更为有效的译码算法，并且随着数字电子技术的进步，该算法的实现已成为可能。此外，面向不同的应用和不同的编码需求，该算法也逐渐出现了很多变种。

2. 卷积码

第二个研究方向是从概率的角度来理解编码和译码的过程，这条道路逐渐产生了序列译码（Sequential Decoding）的概念。序列译码要求引入长度不定的非分组码，这种码可以用树状图来描述，并且可以通过搜索树状图的算法来实现译码。其中最为有用的一种树状码（Tree Codes）是高度结构化的卷积码（Convolutional Codes），这种码可以通过线性移位寄存器电路来生成。到 20 世纪 50 年代末，使用基于序列译码的算法实现了卷积码的成功译码。其后直到 1967 年，A. J. Viterbi 才提出了一种更为简单的译码算法，即维特比算法。对于中等复杂度的卷积码，维特比算法获得了广泛的使用，但是该算法对于强度更大的卷积码是不实用的。

经过二十多年的发展，到了 20 世纪 70 年代，两个研究方向在某些领域开始汇聚并相互渗透。J. L. Massey 和 G. D. Forney 开始研究卷积码的代数理论，开创了卷积码的一种新的研究视角。而在分组码领域，人们开始研究码长较大的好码的构造方案，G. D. Forney 在 1966 年引入了级联码（Concatenated codes）的思想，J. Justesen 使用该思想设计了一种完全可构造且性能很好的长码。同时期的 V. D. Goppa 于 1970 年定义了一类能够确保得到好码的

编码方法，尽管没有给出如何识别好码的方法。

到了20世纪80年代，信道编码器和译码器开始频繁出现在新设计的数字通信系统和存储系统中。例如，在CD中使用了可以纠正两个字节错误的RS码；RS码也常常出现在许多磁带设备、网络调制解调器和数字视频光盘中；而在基于电话线的调制解调器中，代数编码开始被诸如网格编码调制（1982年由G. Ungerboeck提出）的欧式空间编码所取代，这类方法的成功开始引起对基于欧氏距离的非代数编码方法的研究热潮。

到了20世纪90年代，信道编码、信号处理和数字通信之间的界限越来越模糊。Turbo码的出现可以看作是这个阶段的中心事件，用于长二进制码的软判决译码出现了实用的迭代算法，并且实现了香农给出的性能限。1996年，D. MacKay和M. Neal等人对R. G. Gallager在1963年提出的LDPC码重新进行了研究，发现该码具有逼近香农限的优异性能，并且具有译码复杂度低、可并行译码以及译码错误的可检测性等特点，从而成为信道编码理论新的研究热点。

进入21世纪，最大的突破性成果是2008年由土耳其毕尔肯大学Erdal Arikan教授首次提出的Polar码。该码是一种可以实现对称二进制输入离散无记忆信道和二进制擦除信道容量的新型代码构造方法，一出现便在学术界和产业界引起了广泛关注。

第2章　信息论基础

信息是消息的内涵，其与通信问题密切相关。1948 年香农在题为《通信的数学理论》的论文中系统地提出了关于信息的论述，创立了信息论。维纳提出的关于度量信息量的数学公式预示了信息论的广阔应用前景。1951 年，美国无线电工程学会承认了信息论这门学科，此后信息论得到迅速发展。20 世纪 50 年代是信息论向各门学科冲击的时期，而 60 年代信息论属于一个消化、理解的时期，在已有基础上进行重大建设的时期，其研究重点是信息和信源编码问题。70 年代，由于数字计算机的广泛应用，通信系统的能力有了极大提高，如何更有效地利用和处理信息，成为日益迫切的问题。人们越来越认识到信息的重要性，认识到信息可以作为与材料和能源一样的资源加以利用和共享。可以说，信息论是 20 世纪 40 年代后期人们从长期通信实践中总结出来的一门学科，是专门研究信息的有效处理和可靠传输的一般规律的科学，是运用概率论与数理统计的方法研究信息、信息熵、通信系统、数据传输、密码学、数据压缩等问题的应用数学学科。

本章重点学习信息论的基本概念，主要包括信源的数学模型、信息量、信息熵、信道的相关概念以及香农三大定理。通过本章相关概念的学习，有助于理解通信的基本过程，以及信息在传递过程中所涉及的基本理论。本章内容也是后续章节课程学习所必备的知识基础。

2.1　信源数学模型

任何信源产生的消息输出都是随机的，可以用随机变量、随机向量或者随机过程来描述，也可以用概率空间来描述信源。信源的输出常常以符号或者代码的形式出现，比如文字、符号、字母等，取值范围是有限或者可数的。例如，投硬币，每次不是正面就是反面；投骰子，每次都是 1 ~6 点中的一个，结果是随机的，并且一定是集合中的某一个。因此，这样在时间和幅度上离散分布的信源称为**离散信源**，数学模型为离散型的概率空间，下面以一维离散随机变量 X 来描述信源输出

$$\begin{pmatrix} X \\ P(x) \end{pmatrix} = \begin{pmatrix} x_1, & \cdots, & x_n \\ P(x_1), & \cdots, & P(x_n) \end{pmatrix} \tag{2-1}$$

式中，$P(x_i)(i=1, \cdots, n)$满足

$$\sum_i^n P(x_i) = 1 \tag{2-2}$$

其中，$P(x_i)(i=1, \cdots, n)$表示随机事件 X 发生某一结果 x_i 的先验概率，该公式表示信源的可能取值范围是有限的，共有 n 个，每次输出必定为其中某一个。这是离散信源中最为基本的一种。

还有一种类型的信源，其取值在时间和幅度上都是连续分布的，同时又是随机的，例如语音信号、图像、电流值和电压值等。这种信源称为**连续信源**，其数学模型为连续型概率空间：

$$\begin{pmatrix} X \\ P \end{pmatrix} = \begin{pmatrix} (a,b) \\ p(x) \end{pmatrix} \tag{2-3}$$

并且

$$\int_a^b p(x)\mathrm{d}x = 1 \tag{2-4}$$

其中，$p(x)$表示随机变量 X 的概率密度函数。

离散信源和连续信源属于最基本的两种信源。下面简单归纳几种常见的信源类型。

离散平稳信源：若信源输出的 N 维随机序列 X 中，每个随机变量 $X_i(i=1, 2, \cdots, N)$都是取值离散的离散型随机变量，并且 X_i 的概率分布都与时间起点无关，即在任意两个不同时刻 x 和 y（x 和 y 为大于 1 的任意整数），对于任意的 $N=0, 1, 2, \cdots$，$X_xX_{x+1}\cdots X_{x+N}$和 $X_yX_{y+1}\cdots X_{y+N}$具有相同的概率分布。这种各维联合概率分布均与时间起点无关的信源称为离散平稳信源。

连续平稳信源：若信源输出的 N 维随机序列 X 中，每个随机变量 $X_i(i=1, 2, \cdots, N)$都是取值连续的连续型随机变量，并且 X_i 的概率分布都与时间起点无关，即在任意两个不同时刻 x 和 y（x 和 y 为大于 1 的任意整数），对于任意的 $N=0, 1, 2, \cdots$，$X_xX_{x+1}\cdots X_{x+N}$和 $X_yX_{y+1}\cdots X_{y+N}$具有相同的概率分布。这种信源称为连续平稳信源。

离散平稳信源分为有记忆和无记忆两种类型。

（1）离散有记忆信源

一般情况下，信源先后发出的符号之间是互相关联的，也就是说信源输出的平稳离散随机序列 X 中，各随机变量之间是有依赖的，例如，在中文字母组成的中文消息中，前后文字之间是有关联的，称这种信源为离散有记忆信源。故在 N 维随机向量的联合概率分布中，引入条件概率分布来描述它们之间的关联。

（2）离散无记忆信源

离散无记忆信源是最简单的离散信源，可以用完备的离散型概率空间来描述，其主要特点是离散和无记忆。离散指的是信源可能输出的消息的种类是有限的或者是可数的，消息的样本空间 R 是一个离散集合。由于信源的每一次输出都是按照消息发生的概率输出 R 中的一种消息，因此信源输出的消息可以用离散随机变量 X 表示。无记忆是指不同的信源

输出消息之间相互独立。

连续平稳信源也可以分为有记忆和无记忆两种类型。

（1）连续平稳无记忆信源

在连续平稳信源情况下，若信源输出的连续型随机向量 $X=(X_1X_2\cdots X_N)$ 中，各随机变量 X_i 之间无相互依赖、统计独立，其联合概率密度函数满足

$$p(x_1x_2\cdots x_N)=p_1(x_1)p_2(x_2)\cdots p_N(x_N) \tag{2-5}$$

因为信源是平稳信源，因此 $p_1(x)=p_2(x)=\cdots=p_N(x)=p(x)$，那么 $p(x_1x_2\cdots x_N)=\prod_{i=1}^{N}p(x_i)$，该信源称为连续平稳无记忆信源。

（2）连续平稳有记忆信源

有依赖关系的连续平稳信源为连续平稳有记忆信源。

2.2 信息量和信息熵

2.2.1 信息量

所谓信息量就是信息多少的量度。假设信源发送方 A 发送消息给接收方 B，所发出的消息是随机的，在消息收到之前，接收方不能确定会收到什么消息，也就是无法消除消息的不确定性，这种不确定性越大，接收方 B 在收到完整消息时所获得的信息量就越大；如果消息在传输过程中受到干扰，导致接收方 B 对收到消息的不确定性减小甚至不确定性没有改变，那么获得的信息量就较少或者信息量为零。下面举几个简单例子来说明。

“太阳今天早晨竟然从东边升起！”

“吸烟有害健康！”

“天冷要多穿衣服！”

当你看到上面几个消息时，你一定会觉得很无聊，因为我们都知道这些事情，它们发生的概率几乎为 1，所以没有多少信息。

假如你看的新闻标题是：“中国男篮 30 分大胜美国梦之队”“中国足球队 5 比 0 大胜巴西队”，你会立刻产生兴趣，因为这些事情发生的可能性太小。因此，一件事情包含的信息和它发生的概率成负相关关系（即概率越大信息越少，概率趋近于 0，信息趋近于无穷大），信息量的大小跟这种信息不确定的消除有关。

那么可以定义信息量为：

信息量 = 不确定性减少量（消息收到前的不确定性 - 消息收到后的不确定性）

当没有干扰时，收到信息后的不确定性为零，那么此时获得的信息量 = 消息发出前的不确定性，也可以认为是信源消息所含有的信息量。用公式来表示的话，就是某件事发生所含有的信息量等同于该事件发生的先验概率函数

$$I(x_i) = f[P(x_i)] \tag{2-6}$$

式中，$P(x_i)$是事件 x_i 发生的先验概率，而 $I(x_i)$表示事件 x_i 发生所含有的信息量，称之为 x_i 的**自信息量**。

按照刚才的分析，函数$f[P(x_i)]$应该满足如下条件：

（1）$f[P(x_i)]$与概率 $P(x_i)$负相关，即当 $P(x_1) > P(x_2)$时，$f[P(x_1)] < f[P(x_2)]$。

（2）$P(x_1) = 1$，那么$f[P(x_1)] = 0$。

（3）$P(x_1) = 0$，那么$f[P(x_1)] = \infty$。

（4）$I(AB) = I(A) + I(B)$，即两个毫不相关的事件 A 和事件 B 同时发生所提供的信息量等于两个事件各自发生的信息量之和。

1928 年 R. V. L. 哈特莱首先提出信息定量化的初步设想，他将消息数的对数定义为信息量。从数学上可以证明对数形式能够满足刚才所提条件。若信源有 p 种消息，且每个消息是以相等可能产生的，则该信源的信息量可表示为

$$I = \log_2(1/p) \tag{2-7}$$

其中，p 为事件发生的概率，I 为该事件的自信息量。如果对数的底取为 2，那么就是最常见的信息度量，单位为比特（bit），如果取自然对数的底 e，那么单位为奈特（nat），如果底数取为 10，那么单位为哈特（hart），通常情况下，都采用以 2 为底的对数。

由上式可以看出一种特殊情况：若 $p = 1/2$，那么 $I = 1\text{bit}$，也就是二选一时的信息量。这也说明了二元数字所能提供的信息量为 1bit。

自信息量只是表征信源本身的信息特征，但在一般的通信系统里，信源、信道和信宿必须整体进行考虑。假设信源输出为 X，信宿输入为 Y，如图 2-1 所示，那么通常表示某事件在某种条件下出现，这时所带来的信息量就需要在 X、Y 集合中进行考虑，并且要用条件概率来进行描述。

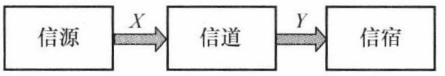

图 2-1　信息系统模型

假设在 y 条件下，随机事件 x 的条件概率为 $P(x|y)$，则 x 的出现所引发的信息量称为**条件自信息量**，如下式所示：

$$I(x|y) = \log_2 P(x|y) \tag{2-8}$$

反之，在 x 条件下，y 所带来的信息量也是条件自信息量：

$$I(y|x) = \log_2 P(y|x) \tag{2-9}$$

对于通信系统而言，该条件概率只与信道特性有关，因此上述条件自信息量可以看作