高等院校电子信息科学与工程类

通信工程专业教材

信息论基础

赵蓉　叶茵　编著

北京邮电大学出版社
www.buptpress.com

高等院校电子信息科学与工程类＊通信工程专业教材

信息论基础

赵　蓉　叶　茵　编著

北京邮电大学出版社
·北京·

内 容 简 介

本书系统论述了香农信息论，以香农三个编码定理为中心，重点讲述了相关的基本概念、基本原理和基本方法。本书针对本科教学的特点，力求深入浅出，在保持一定理论深度的基础上尽可能地简化数学分析过程，以通俗、生动的语言强化物理概念的描述。为了便于教学和读者自学，部分章节配有习题。

本书可作为高等院校通信、电子、计算机专业本科生的教材，也可作为相关专业人员的参考书。

图书在版编目(CIP)数据

信息论基础/赵蓉，叶茵编著. --北京：北京邮电大学出版社，2011.1

ISBN 978-7-5635-2541-6

Ⅰ. ①信… Ⅱ. ①赵… ②叶… Ⅲ. ①信息论 Ⅳ. ①G201

中国版本图书馆 CIP 数据核字(2010)第 256267 号

书　　名：信息论基础

作　　者：赵　蓉　叶　茵

责任编辑：刘春棠

出版发行：北京邮电大学出版社

社　　址：北京市海淀区西土城路 10 号(邮编：100876)

发 行 部：电话：010-62282185　传真：010-62283578

E-mail：publish@bupt.edu.cn

经　　销：各地新华书店

印　　刷：北京源海印刷有限责任公司

开　　本：787 mm×1 092 mm　1/16

印　　张：13

字　　数：315 千字

印　　数：1—3 000 册

版　　次：2011 年 1 月第 1 版　2011 年 1 月第 1 次印刷

ISBN 978-7-5635-2541-6　　定　价：24.00 元

前　言

信息论是利用概率论、随机过程和数理统计等数学方法来研究信息传输和信息处理过程中一般规律的重要学科，是数学知识与通信技术相结合的边缘学科。

信息论是一门既具有严密的逻辑演绎推理系统，又具有丰富生动的时代气息的科学理论，在通信领域中发挥着越来越重要的作用。信息论所蕴涵的独特的、新颖的解决问题的思路和方法已渗透到生物工程、医学工程、自然科学、社会科学、管理科学等领域。

本书主要介绍香农信息论的基本内容及应用，包括信息论的基本概念、三个编码定理及信源和信道编码。本书注重基本概念、基本理论和基本分析方法的论述，力求概念清晰、结构严谨、内容由浅入深、章节循序渐进。

本书共分7章。第1章为概论，介绍了信息的基本概念和定义、信息论研究的内容；第2章为信源和信源熵，详细地讨论了信息的度量方法，介绍熵的概念、性质、定理和信源冗余度的问题；第3章为信道和信道容量，描述和分析了各种不同类型信道的模型和特性，并介绍了几种特殊信道信道容量的计算方法；第4章为信源编码，核心内容是香农的无失真信源编码定理，介绍了无失真编码的基本概念和定理的基本内容；第5章为信道编码，介绍了信道编码的基本概念及香农的有噪信道编码定理；第6章为信息率失真函数，介绍了信息率失真函数的概念、计算、应用及保真度准则下的信源编码定理；第7章为信息理论在现代通信中的应用，介绍了熵概念在密码学中的应用和MIMO信道的容量。

本书第1章、第3章、第4章、第6章和第7章由赵蓉编写，第2章和第5章由叶茵编写。

本书针对本科教学的特点，力求深入浅出，在保持一定理论深度的基础上尽可能地简化数学分析过程，以通俗、生动的语言强化物理概念的描述。

鉴于编者水平有限，难免存在错误、遗漏和不妥之处，恳请读者指正。

编　者

目　录

第1章 概 论

信息科学是研究信息的获取、传输、处理和利用的一门科学。信息论是人们在长期通信工程的实践中，由通信技术与概率论、随机过程和数理统计相结合而逐步发展起来的一门学科。当代伟大的数学家、美国贝尔实验室杰出的科学家香农(C. E. Shannon)在1948年发表的论文《通信的数学理论》为信息论奠定了理论基础。

1.1 信息的一般概念

信息的概念十分广泛，到目前为止，国内外已有百余种流行的说法，从不同的侧面和不同的层次揭示了信息的本质。例如，“信息是事物之间的差异”，“信息是事物联系的普遍形式”，“信息是物质和能量在时间和空间中分布的不均匀性”，“信息是物质的普遍属性”，“信息是收信者事先所不知道的报道”，“信息是用以消除随机不确定性的东西”，“信息是负熵”，“信息是作用于人类感觉器官的东西”，“信息是通信传输的内容”，“信息是加工知识的原材料”，“信息是控制的指令”，“信息就是数据”，“信息就是情报”，“信息就是知识”……

数学家认为“信息是使概率分布发生改变的东西”，哲学家认为“信息是物质成分的意识成分按完全特殊的方式融合起来的产物”……

1928年，美国数学家哈特莱(Hartley)在《贝尔系统电话杂志》上发表了一篇题为《信息传输》的论文，首先提出“信息”这一概念。他认为，发信者所发出的信息就是其在通信符号表中选择符号的具体方式，并主张用所选择的自由度来度量信息。发信者选择的自由度越大，所能发出的信息量就越大。此外，哈特莱还注意到，选择的具体物理内容是无关紧要的，重要的是选择方式。也就是说，无论符号代表的意义是什么，只要符号表的符号数目一定，“字”的长度一定，发信者所发出的信息的数量就被限定了。所以哈特莱认为“信息是选择的自由度”。

哈特莱的这种理解在一定程度上能够解释通信工程中的一些信息问题，但它存在着严重的局限性。首先，他所定义的信息不涉及信息的价值和具体内容，而只考虑选择的方式。其次，即使考虑选择的方式，也没有考虑各种可能选择方式的统计特性。正是这些缺陷严重地限制了它的适用范围。

1948年，控制论的创始人之一、美国科学家维纳(N. Wiener)在《控制论》一书中是这样来论述信息的，“信息是信息，不是物质，也不是能量”。这就是说，信息就是信息自己，它不是其他什么东西的替代物。它是与“物质”、“能量”同等重要的基本概念。正是维纳第一次将“信息”与“物质”、“能量”相提并论，即信息是独立于物质和能量之外的存在于客

观世界的第三要素。

但后来，维纳在《人有人的用处》一书中提出："信息是人们在适应外部世界并且使这种适应反作用于外部世界的过程中，与外部世界进行相互交换的内容的名称。"他还说："接收信息和使用信息的过程，就是我们适应外部世界环境的偶然性变化的过程，也是我们在这个环境中有效地生活的过程。""要有效地生活，就必须有足够的信息。"的确，信息对人类的生存是很重要的；但是，信息不仅与人类有关，不仅是人与外部世界交换的内容。在自然界中，一切生物体都在与外部世界进行着信息交换。一切生物体都有它们独特的接收信息和交换信息的方式。人们发现，动物之间可以利用气味、声音、不同的运动姿态，乃至超声波、电磁场等多种方式来传递信息。另外，信息的确是人们与外部世界互相交换的内容，但人们在与外部世界相互作用的过程中，还进行着物质与能量的交换。这样，就又把信息与物质、能量混同起来。所以，维纳关于信息的定义是不确切的。

关于信息的定义，意大利学者朗格(G. Longe)在 1975 年出版的《信息论：新的趋势与未决问题》一书的序言中提出："信息是反映事物的形式、关系和差别的东西。信息是包含于客体间的差别中，而不是在客体本身中。""在通信中仅仅差别关系是重要的。"也就是说，他定义信息是客体之间的相互差异。的确，宇宙中到处存在着差异，差异的存在使人们存在着"疑问"和"不确定性"。从这个角度看，差异的确是信息。但是，并不能说没有差异就没有信息。所以，这样定义的信息也是不全面的、不确切的。

香农在 1948 年发表了一篇著名的论文——《通信的数学理论》。他从研究通信系统传输的实质出发，对信息作了科学的定义，并给出了定性和定量的描述。

香农将各种通信系统概括成如图 1.1.1 所示的框图。在各种通信系统中，其传输的形式是消息。但消息传递过程中一个最基本、最普通又不引人注意的特点是：收信者在收到消息以前是不知道消息具体内容的。在收到消息以前，收信者无法判断发信者将会发来描述何种事物运动状态的具体消息，也无法判断描述的是这种状态还是那种状态。另外，即使收到消息，由于干扰的存在，也不能判断所得到的消息是否正确和可靠。总之，收信者存在"不知"、"不确定"或"疑问"。通过消息的传递，收信者知道了消息的具体内容，原先的"不知"、"不确定"或"疑问"消除或部分消除了。因此，对收信者来说，消息的传递过程是一个从不知到知的过程，或是从不确定到确定的过程，或是从知之甚少到知之甚多的过程。如果不具备这样一个特点，那就不需要通信系统了。

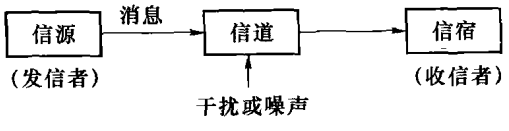


图 1.1.1　通信系统框图

可见，通信过程是一种消除不确定性的过程。消除了不确定性，就获得了信息。原先的不确定性消除得越多，获得的信息量就越多。如果原先的不确定性全部消除，就获得了全部信息；如果消除了部分不确定性，就获得了部分信息；如果不确定性没有任何消除，就没有获得任何信息。由此可见，香农关于信息的定义是：信息是事物运动状态或存在方式的不确定性的描述。

由以上分析可知，在通信系统中形式上传输的是消息，但实质上传输的是信息。消息只是表达信息的工具，是承载信息的客体。显然，在通信中被利用的(即携带信息的)实际客体是不重要的，而信息是重要的。信息是抽象的，而消息是具体的，但不一定是物理性的。通信的结果是消除或部分消除不确定性，从而获得信息。

1.2　信息论研究的对象、目的和内容

信息论研究的是广义的通信系统，它把所有的信息传输系统都抽象成一个统一的模型，如图 1.2.1 所示。信息论研究的对象正是利用这种统一模型来研究信息传输和处理的共同规律。

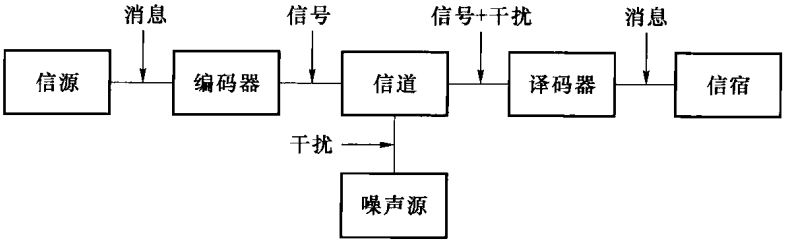


图 1.2.1　通信系统模型

这种模型不仅包括电话、电报、传真、电视、雷达等狭义通信系统，还包括生物有机体的遗传系统、神经系统、视觉系统，甚至人类社会的管理系统。信息以消息的形式在这个通信系统中传递，人们通过研究通信系统中消息的传输和处理来得到信息传输和处理的规律。

该通信系统模型主要包括以下 5 个部分。

(1) 信源

顾名思义，信源是产生消息和消息序列的源。信源可以是人、生物、机器或其他事物。信源的输出是消息，消息是具体的，但它不是信息本身。消息携带着信息，消息是信息的表达者。

信源输出的消息有各种不同的形式，可以是文字、符号、语言、图片、气味等，消息以能被通信双方所理解的形式通过通信进行传输和交换。信源输出的消息是随机的、不确定的，但又有一定的规律性。因此可用随机变量或随机矢量等数学模型表示信源。

(2) 编码器

编码是把消息变换成信号的措施，而译码就是编码的反变换。编码器输出的是适合在信道中传输的信号，信号携带着消息，是消息的承载者。

编码器可分为两种，即信源编码器和信道编码器。信源编码是对信源输出的消息进行适当的变换和处理，目的是为了提高信息传输的效率而压缩信源的冗余度。而信道编码是对信源编码器输出的代码组按照一定的规律添加一些监督码元，使其具有检错、纠错能力，目的是为了提高信息传输的可靠性。

在实际通信系统中，有效性和可靠性是相互矛盾的，去掉信源符号的冗余部分是提高

有效性的必要手段，但这会导致通信系统可靠性的下降；而要提高通信系统的可靠性则需要增加监督码元，这又导致了有效性的降低。所以在很多情况下，为了兼顾有效性，不一定在接收端准确地再现原来的消息，而是允许有一定的误差或失真，近似地再现原来的消息即可。

(3) 信道

信道是指通信系统把承载消息的信号从甲地传输到乙地的媒介或通道。在狭义的通信系统中，实际信道有明线、电缆、波导、光纤、无线电波传播空间等，这些都是属于传输电磁波能量的信道。当然，对广义的通信系统来说，信道还可以是其他的传输媒介。

信道除了传播信号以外，还有存储信号的作用。

在信道中引入噪声和干扰，这是一种简化的表达方式。为了分析方便，常把在系统其他部分产生的干扰和噪声都等效地折合成信道干扰，看成是由一个噪声源产生的，它将作用于所传输的信号上。这样，信道输出的是已叠加了干扰的信号。而噪声源的统计特性又是划分信道的依据，并且决定信道的传输能力。又由于干扰或噪声具有随机性，所以信道常用输入和输出之间的条件概率分布来描述。

(4) 译码器

译码就是把信道输出的已叠加了干扰的编码信号进行反变换，变成信宿能够理解的消息。译码器也可分为信源译码器和信道译码器。译码器应尽可能准确地再现信源输出的消息。

(5) 信宿

信宿是消息传送的对象，即接收消息的人、机器或其他事物。

图 1.2.1 的模型只适用于收发两端单向通信的情况。它只有一个信源和一个信宿，信息传输也是单向的。更一般的情况是：有多个信源和信宿，即信道有多个输入和多个输出，信息传输也可以双向进行。例如，广播通信是单输入、多输出的单向传输的通信系统；而卫星通信网则是多输入、多输出的多向传输的通信系统。要研究这些通信系统，就需要对两端单向通信系统模型作适当的修正，得出多用户通信系统的模型，把两端单向通信的信息理论发展成为多用户通信的信息理论。

信息论研究的具体内容是有过争议的。某些数学家认为，信息论只是概率论的一个分支；有些物理学家认为，信息论只是熵的理论，他们对“熵”特别感兴趣。当然，熵的概念确实是香农信息论的基本概念之一，但信息论的全部内容要比熵的概念广泛得多。

目前，关于信息论研究的内容一般有以下 3 种理解。

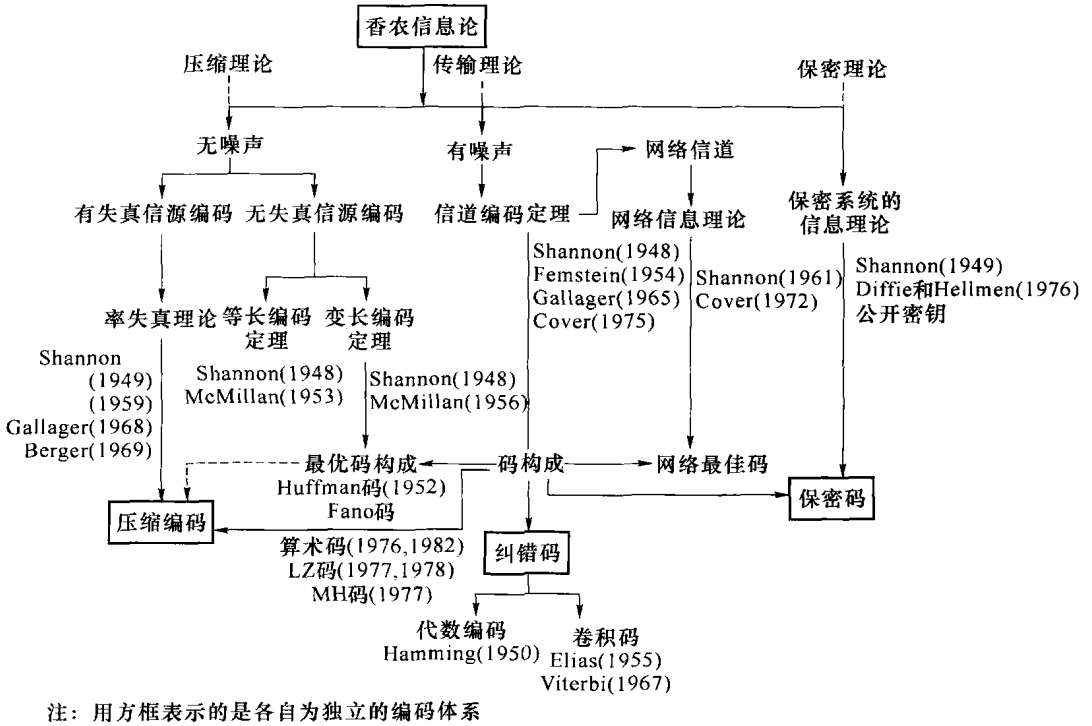
(1) 信息论基础

信息论基础亦称香农信息论或狭义信息论，是以客观概率信息为研究对象，从通信的信息传输问题中总结和开拓出来的理论，主要研究信息的测度、信道容量、信息率失真函数，与这 3 个概念相对应的香农三定理以及信源和信道编码理论。其研究的各部分内容可用图 1.2.2 来描述。

(2) 一般信息论

一般信息论主要也是研究信息传输和处理问题。除了香农基本理论以外，还包括噪声理论、信号滤波和预测、统计检测与估计理论、调制理论、信息处理理论以及保密理论

等。后一部分内容的研究者以美国科学家维纳为代表，其中贡献最大的是维纳和苏联科学家柯尔莫哥洛夫。



注：用方框表示的是各自为独立的编码体系

图 1.2.2　香农信息论的科学体系

虽然维纳和香农等人都是运用概率和统计数学的方法来研究如何准确或近似再现消息的问题，都是通信系统的最优化问题，但他们之间有一个重要的区别。维纳研究的重点是在接收端，研究消息如果在传输过程中被某些因素（如噪声、非线性失真等）干扰后，在接收端如何把消息从干扰中提取出来。在此基础上，他创立了最佳线性滤波理论（维纳滤波器）、统计检测与估计理论、噪声理论等。而香农研究的对象则是从信源到信宿的全过程，是收、发两端联合最优化问题，重点是编码。香农定理指出，只要在传输前后对消息进行适当的编码和译码，就能保证在干扰的情况下，最佳地传送消息，并准确或近似地再现消息。为此，他发展了信息测度理论、信道容量理论和编码理论等。

（3）广义信息论

广义信息论是一门综合性的新兴学科，至今并没有严格的定义。概括起来，凡是能够用广义通信系统模型描述的过程或系统都能用信息基本理论来研究。它不仅包括一般信息论的所有研究内容，而且还包括如医学、生物学、心理学、遗传学、神经生理学、语言学、语义学，甚至社会学和经济管理中有关信息的问题。反之，所有研究信息的识别、控制、提取、变换、传输、处理、存储、显示、价值、作用以及信息量的大小的一般规律以及实现这些原理的技术手段的工程学科，也都属于广义信息论的范畴。

由于信息论研究的内容极为广泛，而各分支又有一定的相对独立性，因此本书只论述信息论的基础理论，即香农信息论。

香农信息论最根本、最本质的问题如下。

(1) 什么是信息？如何度量？

(2) 怎样确定信源的输出含有多少信息量？

(3) 一个信道的信道容量是多少(即最多能传送多少信息量)？

(4) 为了实现无失真地传输信源消息，对信源编码所需的最少的符号数是多少？这是无失真信源编码，即香农第一定理的内容。

(5) 在有噪信道中是否有可能以接近信道容量的信息传输速率传输信息而错误概率几乎为零？这是有噪信道编码，即香农第二编码定理的内容。

(6) 如果对信源编码时允许一定量的失真，所需的最少的码符号数又是多少？这是限失真信源编码，即香农第三编码定理的内容。

目前，在香农信息论方面值得注意的研究动向是信息概念的深化、网络信息理论和多重相关信源编码理论的发展和应用、通信网的一般信息理论研究、磁记录信道的信息理论研究、信息率失真的发展及在数据压缩和图像处理中的应用、信息论在大规模集成电路中的应用等问题。这些领域都是与当前信息工程的前景——光通信、空间通信、计算机互联网、移动通信、多媒体通信、语音和图像的信息处理等密切相关的。

第 2 章　信源和信源熵

香农信息论的基本点是用随机变量或随机矢量来表示信源，运用概率论和随机过程的理论来研究信息。

2.1　信源的描述与分类

在信息论中，信源是信息的来源，它一般以符号的形式发出具体消息。包含信息的符号常带有随机性。如果符号是确定的，而且是预先知道的，那么该消息就无信息可言。因此只有当符号的出现具有随机性，预先无法确定时，该符号的出现才给观察者提供了信息。基于对信源的这种认识，我们通常可用随机变量或随机矢量来描述信源输出的消息，或者说，用概率空间来描述。

2.1.1　离散信源与连续信源

有些信源可能输出的消息数是有限的或可数的，而且每次只输出其中一个消息。例如，掷一颗六面质地均匀的骰子，研究其落下后，朝上一面的点数。每次实验结果必然是 1 点、2 点、3 点、4 点、5 点、6 点中的一个面朝上。其输出消息是“朝上一面是 1 点”、“朝上一面是 2 点”、……、“朝上一面是 6 点”等 6 种不同的消息。每次实验中，出现哪一种消息是随机的，但必定是出现其中的某一种信息。这 6 种不同的消息构成两两互不相容的基本事件集合，用符号 $a_i(i=1,2,\cdots,6)$ 来表示这些消息。另外，大量试验表明，各消息都是等概率出现的，即概率都等于 1/6。因此，可以用随机变量 X 来描述信源输出的消息。把这个信源抽象后得到的数学模型为

$$\begin{bmatrix} X \\ p(x) \end{bmatrix}=\begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix}$$

并且各事件出现的概率满足

$$\sum_{i=1}^{6} p(a_i)=1$$

可见，这个离散信源的概率空间是一个完备的概率空间集，信源输出的消息只能是符号集 $\{a_1,a_2,\cdots,a_6\}$ 中的任何一个，而且每次必定选取其中一个。

在这个典型实例的启发下，我们可以构建一般单符号信源的数学模型。

定义 2.1.1　若信源输出的消息数是有限的或可数的，而且每次只输出符号集中的一个消息，这样的信源称为简单的离散信源。

其数学模型就是离散型的概率空间：

$$\begin{bmatrix} X \\ p(x) \end{bmatrix}=\begin{bmatrix} a_1 & a_2 & \cdots & a_q \\ p(a_1) & p(a_2) & \cdots & p(a_q) \end{bmatrix}$$

显然，$p(a_i)(i=1,2,\cdots,q)$应满足

$$\sum_{i=1}^{q} p(a_i) = 1$$

上式表示信源可能取的消息(符号)只有 q 个：$\{a_1,a_2,\cdots,a_q\}$，而且每次必定取其中一个。

不同的信源对应不同的概率空间。如信源给定，就意味着相应的概率空间已经确定。反之，如信源的概率空间已经给定，就意味着相应的信源已经给定。用概率空间表示信源的数学模型的必要前提就是信源可能发出的各种不同符号的概率先验可知。

在实际情况中，存在着很多这样的信源，如书信文字、计算机的代码、电报符号、阿拉伯数字等。这些信源符号集的取值是有限的或可数的。

定义 2.1.2 若信源的输出是单个符号(代码)的消息，可能出现的消息数是不可数的无限值，即输出消息的取值是连续的，这样的信源称为简单的连续信源。

语音信号及遥控系统中测得的电压、温度、压力等连续数据都是简单的连续信源。我们可用连续型随机变量来描述这些消息，其数学模型为连续型的概率空间：

$$\begin{bmatrix} X \\ p(x) \end{bmatrix}=\begin{bmatrix} (a,b) \\ p(x) \end{bmatrix}\text{或}\begin{bmatrix} \mathbf{R} \\ p(x) \end{bmatrix}$$

并满足

$$\int_a^b p(x)\mathrm{d}x = 1 \text{ 或} \int_{\mathbf{R}} p(x)\mathrm{d}x = 1$$

式中，$\mathbf{R}$ 表示实数集$(-\infty,\infty)$；而 $p(x)$是随机变量 X 的概率密度函数。

2.1.2 离散无记忆信源和离散有记忆信源

上述信源是最简单的情况，因为信源只输出一个消息(符号)，所以可用一维随机变量来描述。然而，很多实际信源输出的信息往往是由一系列符号组成的。例如，常用汉字用 4 位阿拉伯数字($N=4$)来表示，其中每一位都从 0、1、2、3、4、5、6、7、8、9 这 10 种符号($q=10$)中取一种。

定义 2.1.3 若离散信源输出的消息是由一系列符号组成(设由 N 个符号组成)的，这种信源称为多维信源。

这种信源不能使用一维随机变量来描述，而应使用 N 维随机矢量 $\boldsymbol{X}=(X_1X_2\cdots X_N)$ 来描述。该 N 维随机矢量 $\boldsymbol{X}$ 有时也称为随机序列。

在 N 维随机矢量中，每个随机变量 X_i 的取值 x_i 为

$$x_i \in A=\{a_1,a_2,\cdots,a_q\} \quad (i=1,2,\cdots,N)$$

$$\boldsymbol{X}=(X_1X_2\cdots X_N)$$

则信源的 N 重概率空间为

$$\begin{bmatrix} X^N \\ p(x) \end{bmatrix}=\begin{bmatrix} (a_1a_1\cdots a_1) & (a_1a_1\cdots a_2) & \cdots & (a_qa_q\cdots a_q) \\ p(a_1a_1\cdots a_1) & p(a_1a_1\cdots a_2) & \cdots & p(a_qa_q\cdots a_q) \end{bmatrix}$$

这个空间共有 q^N 个元素。

在某些简单的情况下，信源先后发出的一个个符号彼此是统计独立的，并且它们具有相同的概率分布，则 N 维随机矢量的联合概率分布满足

$$p(\boldsymbol{X}) = \prod_{i=1}^{N} p(x_i = a_{k_i})$$

式中，k_i 可取 $1,2,\cdots,q$，即 N 维随机矢量的联合概率分布可用随机矢量中单个随机变量的概率乘积来表示。这种信源称为离散无记忆信源。

而一般情况下，信源先后发出的符号之间是互相依赖的。例如，在汉字组成的中文序列中，只有根据中文的语法、习惯用语、修辞制约和表达实际意义的制约所构成的中文序列才是有意义的中文句子或文章。所以，在汉字序列中前后文字的出现是有依赖的，不能认为是彼此不相关的。其他如英文、德文等自然语言都是如此。这种信源称为有记忆信源。我们需要在 N 维随机矢量的联合概率分布中引入条件概率分布来说明它们之间的关联。

表述有记忆信源要比表述无记忆信源困难得多。实际上，信源发出的符号往往只与前若干个符号的依赖关系较强，而与更前面的符号依赖关系弱得可忽略不计。为此，可以限制随机序列的记忆长度。

当记忆长度为 $m+1$ 时，称这种有记忆信源为 m 阶马尔可夫信源。也就是信源每次发出的符号只与前 m 个符号有关，与更前面的符号无关。这样可用马尔可夫链来描述信源。这时描述符号之间依赖关系的条件概率为

$$p(x_i \mid x_{i-1}x_{i-2}x_{i-3}\cdots x_{i-m}\cdots) = p(x_i \mid x_{i-1}x_{i-2}x_{i-3}\cdots x_{i-m})$$

如果条件概率与时间起点 i 无关，即信源输出的符号序列可看成为时齐马尔可夫链，则此信源称为时齐马尔可夫信源。

综上所述，信源可分为离散信源和连续信源两种。本章重点讨论离散信源和连续信源的熵。离散信源根据其符号之间是否独立分为无记忆信源和有记忆信源，无记忆信源的理论比较成熟，而有记忆信源的理论还不十分完整，因此只讨论有限记忆信源，即马尔可夫信源的问题。对于连续信源，主要讨论熵和最大熵定理。

2.2　离散信源的信息熵

本节研究最基本的离散信源，即信源输出是单个符号的消息，而且这些消息是两两互不相容的。

对于一般实际输出为单个符号的离散信源，无论它输出的是文字、数字、字母还是其他符号，都可用一维随机变量 X 来描述信源的输出，信源的数学模型统一抽象为

$$\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \cdots & a_q \\ p(a_1) & p(a_2) & \cdots & p(a_q) \end{bmatrix} \tag{2.2.1}$$

其中

$$\sum_{i=1}^{q} p(a_i) = 1. \tag{2.2.2}$$