内 容 简 介

本书内容包括三部分，即第一篇信息论基础，第二篇纠错编码导论，第三篇保密通信。第一篇介绍信息的量度，离散信源和连续信源的信息量，离散信道和连续信道的信道容量，消息在信道上的信息传输速率，信源编码定理和信道编码定理，信息率失真函数等。第二篇介绍纠错编码的代数基础，各种纠错码的编码和译码原理及实施方案等。第三篇介绍保密通信的基础理论和实现通信保密的方法等。本书的特点是把纠错编码和保密学作为信息论的应用编写的，全书着重物理概念的介绍，其目的是欲使读者能够较清楚地了解信息论在通信技术中的具体应用情况。

本书可作为通信与电子系统等专业的研究生教材，以及无线电技术、通信、信息工程、生物医学工程、计算机科学与工程等专业的本科生教材，亦可供从事无线电技术和通信工程的技术人员进修使用。

信息论与编码

吴伯修　祝宗泰　钱霖君　编

东南大学出版社出版

南京四牌楼2号

江苏省新华书店发行　　江苏阜宁印刷厂印刷

开本 787×1092　毫米　1/16　印张19.5　字数500千

1991年6月第1版　1991年6月第1次印刷

印数：1—2000册

ISBN　7—81023—421—8

TN·45　　定价：5.10元

出 版 说 明

根据国务院关于高等学校教材工作分工的规定，我部承担了全国高等学校、中等专业学校工科电子类专业教材的编审、出版的组织工作。由于各有关院校及参与编审工作的广大教师共同努力，有关出版社的紧密配合，从1978年至1985年，已编审、出版了两轮教材，正在陆续供给高等学校和中等专业学校教学使用。

为了使工科电子类专业教材能更好地适应“三个面向”的需要，贯彻“努力提高教材质量，逐步实现教材多样化，增加不同品种、不同层次、不同学术观点、不同风格、不同改革试验的教材”的精神，我部所属的七个高等学校教材编审委员会和两个中等专业学校教材编审委员会，在总结前两轮教材工作的基础上，结合教育形势的发展和教学改革的需要，制订了1986～1990年的“七五”（第三轮）教材编审出版规划。列入规划的教材、实验教材、教学参考书等有近400种选题。这批教材的评选推荐和编写工作由各编委会直接组织进行。

这批教材的书稿，是从通过教学实践、师生反映较好的讲义中经院校推荐，由编审委员会（小组）评选择优产生出来的。广大编审者、各编审委员会和有关出版社为保证教材的出版和提高教材的质量，作出了不懈的努力。

限于水平和经验，这批教材的编审、出版工作还会有缺点和不足之处，希望使用教材的单位，广大教师和同学积极提出批评建议，共同为不断提高工科电子类专业教材的质量而努力。

电子工业部教材办公室

目　录

第一篇　信息论基础

第二篇　纠错码导论

第三篇　保密通信

一内

第一篇　信息论基础

第一章　通信的基础知识

第一节　引　言

众所周知，客观世界有三大基本要素，即物质、能量和信息。人们在生产和生活之中，对于物质和能量早已有所认识，所以这两者早已成为自然科学的研究对象，分别形成了专门的学科，例如物理学、化学、天文学、地学、生物学等。本世纪五十年代后，人们终于认识到信息的客观存在，并建立了研究信息的学科——信息论。信息论是应用统计数学的方法研究信息的传输及有关问题的科学。因为信息论的研究对象是信息，所以有必要了解信息的含义。在当今信息社会中，“信息”这个名词应用已十分广泛，例如报纸上和电视节目中的经济信息、商品信息、人才交流信息等等，但是要给信息下一个完整的、确切的定义迄今尚未实现。有很多学者曾对信息作过各式各样的定义。对信息论曾有杰出贡献的科学家维纳（N.Wiener）说：“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，同外部世界进行交换的内容的名称。”综合信息的各种定义，比较一致的定义为：信息是认识主体（人、生物、机器）所感受的或所表达的事物运动的状态和运动状态变化的方式。信息存在于自然界，也存在于人类社会。可以这样说：哪里有事物，哪里就有事物的运动，同时哪里就产生了信息。进一步，信息的特征为：

（1）接收者在收到信息之前，对它的内容是不知道的，所以信息是新知识、新内容；

（2）信息是能使认识主体对某一事物的未知性或不确定性减少的有用知识；

（3）信息可以产生，也可以消失，同时信息可以被携带、贮存及处理；

（4）信息是可以量度的，信息量有多少的差别。

由于人们迄今对于信息的属性和信息的传输媒介还没有完全弄清楚，譬如信息、物质和能量这三者之间存在怎样的关系？信息的传递是否存在着信息感应和信息辐射这一类情况？人体的特异功能又该怎样解释？凡此种种需要人们深入地研究和探索，以便能在此基础上获得圆满的信息定义。

信息论是在信息可以量度的基础上，研究有效地和可靠地传递信息的科学。它涉及信息量度、信息特性、信息传输速率、信道容量、干扰对信息传输的影响等方面的知识。通常把上述范围的信息论称为**狭义信息论**，又因为它的创始人是香农（C.E.Shannon），故又称为**香农信息论**。本书第一篇仅介绍香农信息论。香农信息论的研究对象是语法信息，因为语法信息只与事件的出现概率有关，所以它又称为概率信息。还有两种信息是语义信息和语用信息，前者是指这种信息不仅与出现概率有关，而且还与信息的含义有关；后者是指该种信息的大小还与接收者本身的条件有关。**广义信息论**则包含通信的全部统计问题的研究，除了香农信息论之外，它还包括信号设计、噪声理论、信号的检测与估值等。

信息论是一门数学性的通信理论基础知识，蕴藏着大量新的数学问题，因此近数十年来得到许多数学家、科学家和工程师们的重视，发展较快。同时信息论的研究成果也已经在通信、电视、雷达、制导、数据处理、计算机、自动控制、生物学、心理学、医学等领域中得到越来越广泛的应用。

近三十多年来，信息论沿着两条不同途径发展。其一是在维纳的“平稳时间序列的内插、外推和平滑方法”和“控制论”这两本名著的基础上发展起来的一个信息论分支——微弱信号检测理论；其二是在香农的“通信的数学理论”和“噪声中的通信”这两篇经典著作的基础上发展起来的另一个信息论分支——传输消息的设计和编码理论。虽然维纳和香农都认为“携带信息的信号”也可以用随机过程进行描述，但是他们探讨的问题并不相同。维纳模型是要在接收端尽可能逼真地恢复“携带信息的信号”，因此维纳认为只有当信号受到噪声干扰时才需要进行处理。而香农模型在接收端要恢复的是依附在信号上的信息，而不是信号本身，因此香农认为信号在通过有噪声的信道的前后都需要进行处理。尽管有这些差别，但他们的工作却有许多相同之处，其中最重要的是“最佳系统”的设想。在这之前，人们满足于设计或建造一个优良系统，但是该系统好到怎样程度，缺少衡量标准。由于最佳系统的性能是实际系统性能的上界，因此可作为设计优良系统的努力方向。维纳的工作及微弱信号检测理论着重研究干扰作用下信号的最佳接收问题。它是通信、雷达、导航、遥测、遥控以及电子对抗等技术的理论基础。而香农的工作及传输消息的设计和编码理论则着重研究信源和信道的统计特性及编码方法，以提高信息传输的有效性和可靠性。

随着自动控制、计算机、系统工程、人工智能等技术的发展和结合，近十年来出现了一门新兴的信息技术，它包括信息的获取、传输、处理、贮存和分发的技术。简言之，信息技术是以微电子（大规模集成电路）为基础，通信和计算机的紧密结合。与信息技术相应的科学称为信息科学。信息科学的基本理论是信息论和控制论，但它比信息论研究的范围更广阔，涉及的内容更深奥、更复杂，因此需要更好的工具——电子计算机。电子计算机为信息技术提供了运算、贮存和逻辑思维的能力，把信息处理技术提高到一个更高的水平。此外，信息论和仿生电子学、人工智能理论相结合将为信息技术的发展开辟一个广阔的新方向。自然界各种生物都有它们独特的信息识别和信息处理功能。从生物的识别和处理信息的机理中吸取技术思想，用电子技术的各种手段在信息系统中实现，就有可能明显地改善机器的信息处理能力。人脑的思维是高级信息处理系统，若能从人脑的思维活动中吸取技术思想来丰富机器的能力，就能使机器具有逻辑演译、推理和形式逻辑思维能力。具有这种能力的系统称为智能信息系统。它是信息技术研究的主要对象。可以预期，随着信息技术和信息科学的发展，将为人类提供最有效、最可靠的信息传输、处理和控制的手段，它将对现代科学技术的发展产生深远的影响。

第二节　信息的传输

在信息论和通信理论中经常会遇到信息、消息和信号这三个既有联系又有区别的名词。在介绍信息传输的模型之前有必要先介绍这三个名词。

（1）信息：信息的定义在上一节中已介绍过，简言之，信息是指各个事物运动的状态及状态变化的方式。人们从来自对周围世界的观察得到的数据中获得信息。信息是抽象的意识或知

识，它是看不见、摸不到的。人脑的思维活动产生的一种想法，当它仍贮存在脑子中的时候它就是一种信息。

（2）**消息**：一般而言，消息是指包含有信息的语言、文字和图象等，例如我们每天从广播节目、报纸和电视节目中获得各种新闻及其它消息。在通信中，消息是指担负着传送信息任务的单个符号或符号序列。这里的符号包括字母、文字、数字和语言等。单个符号消息的情况，例如我们用x_1表示晴天，x_2表示阴天，x_3表示雨天。符号序列消息的情况，例如一个英语单词“BOY”由三个字母构成；“今天是晴天”这一个消息由五个汉字构成。可见消息是具体的，它载荷信息，但它不是物理性的。

（3）信号：信号是消息的物理体现。为了在信道上传输消息，就必须把消息加载（调制）到具有某物理特征的信号上去。信号是信息的载荷子或载体，它是物理性的。在近代通信中信号往往是电信号和光信号（光纤通信）。

一、通信系统的模型

通信系统是用来传递消息的系统。我们熟知的电报、市内电话、长途电话、短波通信、卫星通信和移动通信等都属于通信的范畴。尽管广播、电视、雷达、声纳、导航、遥控和遥测等系统在结构上和所传输的消息上各不相同，但它们传递消息的方式则与通信系统是相似的，故也把它们列入广义通信的范畴。单向传递消息的广义通信系统的简化物理模型可用图1-1来表明。在该图上，我们设信源发出的是单个符号的消息u，信宿收到的也是单个符号

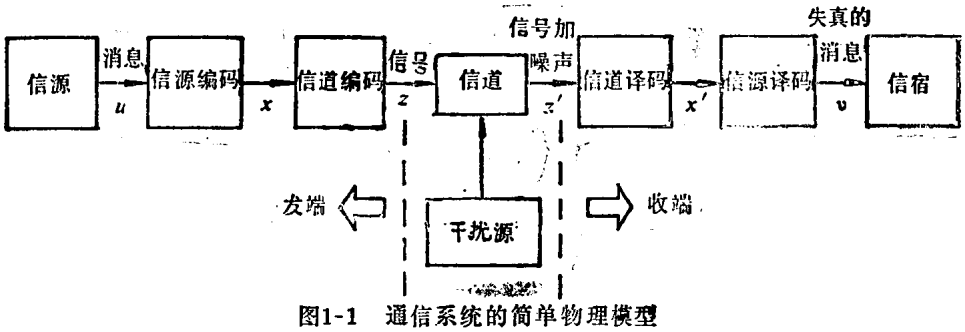

图1-1　通信系统的简单物理模型

的消息v。由于信道上存在干扰，所以信宿收到的是有失真的消息。图1-2也是通信系统的

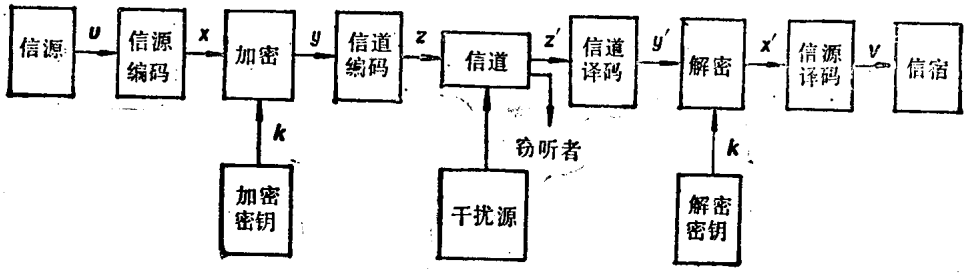

图1-2　通信系统的物理模型

物理模型。它是在图1-1所示模型的基础上增加了保密通信的功能，并设信源发出的是符号序列的消息u，信宿收到的也是符号序列v。下面我们利用图1-2的通信模型来介绍通信系统在传递消息过程中带有普遍性的规律。

信源是向通信系统提供消息u的人或机器。信源本身是十分复杂的，在信息论中我们仅研究信源的输出。信源输出的是以符号形式出现的具体消息，它载荷信息。信源输出的消息

可以有多种形式，但可归纳成两类：（1）**离散消息**，例如由字母、文字、数字等符号组成的符号序列，或者单个符号。（2）**连续消息**，例如话音、图象、在时间上连续变化的电参数等。因为通信系统的接收者（信宿）在收到消息之前并不知道信源所发出消息的内容，所以一般地说信源发出的是随机性的消息。但因信源发出的消息都携带着信息，可见消息的变化是具有一定规律性的，因此严格地说信源发出消息并不是完全随机性的。

信宿是消息传递的对象，即接收消息的人或机器。根据实际需要，信宿接收的消息 v 其形式可以与信源发出的消息 u 相同，也可以不相同。当两者形式不相同时，v 是 u 的一个映射。

信道是传递消息的通道，又是传送物理信号的设施。信道可以是一对导线、一条同轴电缆、传输电磁波的空间、一条光导纤维等传输信号的媒质。在图1-1和图1-2上的信道实际上是由发信机（调制器和发射机）、收信机（接收机和解调器）以及天线、传输信号的媒质这四个部分组成的。这种信道的输入信号和输出信号都是基带信号。收、发信机又称为信道机，发信机是把基带信号变成便于在信道上传输的射频大功率信号，而收信机是把射频信号恢复成所需要的基带信号。

信源编码器的作用是把信源发出的消息变换成由二进制码元（或多进制码元）组成的代码组，这种代码组就是基带信号。同时通过信源编码可以压缩信源的冗余度（即多余度），以提高通信系统传输消息的效率。**信道编码器**的作用是在信源编码器输出的代码组上有目的地增加一些监督码元，使之具有检错或纠错的能力。从信道编码器输出的也是代码组形式的基带信号。**信道译码器**具有检错或纠错的功能，它能将落在其检错或纠错范围内的错传码元检错或纠错，以提高传输消息的可靠性。**信源译码器**的作用是把信道译码器输出的代码组变换成信宿所需要的消息形式，它的作用相当于信源编码器的逆过程。

干扰源是整个通信系统中各个干扰的集中反映，用以表示消息在信道中传输时遭受干扰的情况。对于任何通信系统而言，干扰的性质、大小是影响系统性能的重要因素。

密钥源是产生密钥 k 的源。信道编码器输出信号 x 经过 k 的加密运算后，就把明文 x 变换为密文 y。若窃听者未掌握发端采用的密钥 k，则他就很难从窃听到的信号 z' 解出明文。而收端的信宿则因知道事先已约定好的密钥 k，因此能从收到的信号 z' 解出明文。对于二进制的代码而言，加密相当于 $y=z\oplus p$ 运算（其中序列 p 通常是受密钥控制的伪随机序列），而解密则相当于 $x'=y'\oplus p$ 运算。这里 x'、y'、z' 之所以不同于发端的 x、y、z 是考虑到信号 z 在信道中传输时所受到的干扰影响。但在正常通信条件下，总会有 $x'\approx x$，$y'\approx y$，$z'\approx z$ 的结果。

顺便指出：不是所有的通信系统都采用如图 1-2 所示的那样全面的技术。例如点对点的有线电话，只要有一对电话机和一条电话线路（铜线）就够了，话音基带信号通过电话机变成相应的电信号（模拟信号），就能在电话线上传送，收端的电话机再把电信号恢复成人耳能听得清的话音。如果是点对点的无线电话，则在发端需要一台发信机，把模拟信号调制到射频上，再用大功率发射机经天线发射出去，然后在无线信道中传输，在收端则应使用收信机把收到的调制射频信号解调恢复为发端的原始话音。若在这样的系统中增加了加密和解密装置，就构成无线保密通信系统。在干扰大、信道容量有限的通信系统中，就需要采用信源编码和信道编码技术，以提高传输消息的有效性和可靠性。

二、通信的目的

通信是一种交流信息的过程或活动，通信构成一种提供服务或业务的功能。通信的目的

是使信源发出的由消息携带的信息能被信宿接受。消息在具有保密装置的通信系统中传输时，还可以防止消息被敌方人员或无关人员所窃听。

三、模拟传输和数字传输

通信的传输方式按照在信道（包括收、发信机和天线的信道）上传输的基带信号的性质可分为模拟传输方式和数字传输方式两类。模拟传输方式是指传输的基带信号是连续信号，即它在时间上的变化是连续的，在幅度上的取值也是连续的。因为这种连续信号能模拟原消息的整个变化过程，故称为**模拟信号**。另一类是数字传输方式，这是指传输的基带信号是二进制码元（即数字信号），它属于**离散信号**。连续信号变换成数字信号的过程为：首先按照抽样定理对连续信号进行时间上的取样，也就是把连续信号变成时间离散的脉冲序列信号，然后再对各个脉冲幅度进行量化，使其幅度值变成所处量化等级的量化值，再把这量化值变成二进制数，从而得到与量化值相对应的由多位二进制码元构成的码元序列。这样连续信号就被变成数字信号了。

与模拟传输方式相比，数字传输方式有较多优点：①抗干扰能力强；②在远距离中继通信中干扰的影响不积累；③便于使用纠错技术来提高通信的可靠性；④便于与计算机通信网相连接等。因此近年来数字传输方式的通信有迅速的发展。但数字传输方式也有频带利用率低的缺点。例如数字电话的频带常为模拟电话的八倍。随着通信系统载波频率的提高，特别是光纤通信的发展，通信系统可资利用的带宽正在增加，可以满足快速数字传输方式通信的需要，因此建立综合业务数字网已在计划之中。

四、信息论研究的基本问题

信息论研究的基本问题是有关信源的输出、信道的统计特性、信源编码和信道编码等问题。

关于信源，信息论着重研究信源包含的平均信息量（熵）和单位时间内信源发出的平均信息量（时间熵）。关于信道，信息论着重研究消息在无干扰信道和有干扰信道上传输的信息传输速率，以及信道传输信息量的能力（信道容量）。关于信源编码，信息论指出：把出现概率大的消息用较短的代码组，而把出现概率小的消息用较长的代码组，就能提高信道传输消息的效率。无失真的信源编码定理指出：当代码组的长度增大时，符号序列消息中每个符号所对应的代码组长度以单符号信源熵为下界，也就是每个符号所对应的代码组长度不能小于单符号信源熵，否则就不能实现信源编码。至于信道编码，信息论的有扰离散信道编码定理指出：当信息传输速率小于信道容量时，通过增大代码组的长度，可以以很小的误码率来传输消息，也就是在误码率可任意小的条件下，信息传输速率的上界为信道容量。这实际上是增加传输信号的多余度来提高它的抗信道干扰的能力。由于所增加的多余度是根据信道干扰的性质设计的，所以它的抗信道干扰的能力远大于信源消息固有多余度的抗干扰能力。因此我们才有可能通过信源编码和信道编码来提高信道传输消息的有效性和可靠性。此外，采用脉冲压缩技术来减小信源消息的多余度也能提高信道的传输效率。顺便指出：压缩信源消息的多余度，对于保密通信而言也是防止窃听者破译的措施之一。

本书的第二篇纠错码导论和第三篇保密通信作为第一篇信息论基础的应用。第二篇着重介绍在有扰离散信道编码定理的基础上发展的差错控制编码理论及其实现技术；第三篇着重介绍保密通信的基础理论和一些实现技术。

可以认为保密通信的加密措施是把消息内容加以变换隐蔽其信息内容，加密的作用对于敌方人员和无关人员来说是增加了传输消息的模糊度，从而增加窃听者获得明文的难度。但我方接收端则因掌握密钥故能方便地从密文解出明文。

密码学是研究如何隐蔽消息中的信息内容，以便它在传输过程中不被窃听，而信息论则是研究如何能在有扰信道中可靠和快速地传输消息。可见它们研究的是消息传输的两个不同的侧面，但是所研究的问题却是密切相关的，这就是香农在四十年代同时研究信息论和密码学的原因。香农于1948年发表的《通信的数学理论》奠定了信息论的基础，此后又在1948年发表《保密系统的通信理论》是一篇用信息论的观点来讨论密码学的论文。正由于这两个领域有着共同的理论基础，而且它们的发展过程又是相辅相成的，因此我们在本书的第三篇介绍保密通信的内容有其可能性和必要性。尤其目前人们已经把密码学看作信息论的一个分支。在保密通信的应用日益普遍的当今，把保密通信的基础知识介绍给读者更是必要，但限于篇幅，本书第三篇只能介绍保密通信的一些基本知识。

第二章 信息的量度

第一节 自信息量和条件自信息量

在第一章第一节中，我们已指出信息论是在信息可以量度的基础上，研究有效地、可靠地传递信息的科学。可见，信息的可量度性是建立信息论的基础。本章将介绍消息和信源的不确定度，以及信息的量度方法。信息的量度又是以不确定度的减少量为基础的。信息可以分为语法信息、语义信息和语用信息三种。本书仅论述狭义信息论，故只研究语法信息的量度，而不考虑概率事件的含义及信宿接收信息的能力。

一、自信息量

一个随机事件的**自信息量**定义为其出现概率对数的负值。若随机事件 x_i 的出现概率为 $P(x_i)$，那么它的自信息量 $I(x_i)$ 为

$$I(x_i) \triangleq -\log P(x_i) = \log \frac{1}{P(x_i)} \tag{2.1}$$

自信息量的单位与所用的对数底有关。在信息论中常用的对数底为2，信息量的单位为**比特**（bit，binary unit的缩写）。在信息论的公式推导中，为方便起见，常取自然对数，即对数底取 e，信息量的单位为**奈特**（nat，nature unit的缩写）。又当概率事件的概率很小时，特别是当 $P(x_i) = 10^{-b}$，b 是一个相当大的正整数时，为了运算方便，可以取10作为对底数，信息量的单位是哈脱莱（Hartley），以纪念科学家哈脱莱首先提出用对数值来量度信息。这三个信息单位之间的转换关系如下：

$$1 \text{ nat} = \log_2 e \simeq 1.433 \quad \text{bit}$$

$$1 \text{ Hartley} = \log_2 10 \simeq 3.322 \quad \text{bit}$$

由式(2.1)可知，一个以等概率出现的二进制码元（0，1）所包含的自信息量为1 bit。其由来为当 $P(0) = P(1) = \frac{1}{2}$ 时，有

$$I(0) = I(1) = -\log_2 \frac{1}{2} = \log_2 2 = 1 \quad \text{bit}$$

这里需要引入随机事件的**不确定度**概念。根据日常知识，各个出现概率不同的随机事件所包含的不确定度是有差别的。一个出现概率接近于1的随机事件，它发生的可能性很大，所以它包含的不确定度就很小。反之，一个出现概率很小的随机事件，很难猜测在某个时刻它能否发生，所以它所包含的不确定度就很大。再则，确定性事件的出现概率为1，那么它包

含的不确定度为0。随机事件的不确定度在数量上等于它的自信息量，所以不确定度已可用(2.1)式计算，两者的单位也相同，但含义并不相同。

例 某地二月份气候的概率分布根据多年气象资料统计出下表：

x_i	晴x_1	阴x_2	雨x_3	雪x_4
$P(x_i)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

这四种气候的不确定度分别为 $I(x_1)=1$ bit，$I(x_2)=2$ bit，$I(x_3)=3$ bit，$I(x_4)=3$ bit。

在二维联合集XY上的元素对x_iy_j的自信息量定义为

$$I(x_iy_j)\triangleq -\log P(x_iy_j)=\log\frac{1}{P(x_iy_j)}, \tag{2.2}$$

式中$P(x_iy_j)$是元素对x_iy_j的联合概率。当x_i和y_j相互独立时，有$P(x_iy_j)=P(x_i)\cdot P(y_j)$，那么就有$I(x_iy_j)=I(x_i)+I(y_j)$。元素对$x_iy_j$的不确定度在数值上也等于它们的自信息量，也可用(2.2)式计算。

二、条件自信息量

条件自信息量定义为条件概率对数的负值。设在y_j条件下，随机事件x_i的条件概率为$P(x_i|y_j)$，那么它的条件自信息量$I(x_i|y_j)$定义为

$$I(x_i|y_j)\triangleq -\log P(x_i|y_j)=\log\frac{1}{P(x_i|y_j)} \tag{2.3}$$

在给定y_j条件下，随机事件x_i所包含的不确定度在数值上与条件自信息量相同，即也可用(2.3)式计算，但两者的含义不同。

由于一个随机事件的概率和条件概率总是在闭区间〔0，1〕内，所以自信息量和条件自信息量均为非负值。

第二节 互信息量和条件互信息量

一、互信息量

设有两个离散符号的消息集合X和Y，X是信源发出的符号集合，Y是信宿收到的符号集合。由于信宿事先不知道信源在某一时刻发出的是哪一个符号，所以每个符号消息是一个随机事件。信源发出符号通过有干扰的信道传递给信宿，如图2-1所示。这是最简单的通信系统模型。以信道为基准，信源发出的消息又称为信道输入消息，信宿收到的消息又称为信道输出消息。通常信宿可以预先知道信息X发出的各个符号消息的集合，以及它们的概率

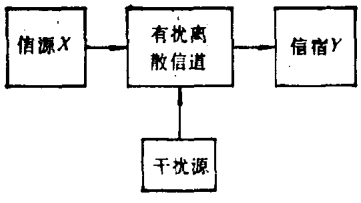

图2-1 最简单的通信系统模型

分布，也就是预先知道信源X的概率空间。对于$X=\{x_1, x_2, \cdots, x_N\}$，其相应的各个元素的概率为$P(x_1)$，$P(x_2)$，…，$P(x_N)$。$x_i$和$P(x_i)(i=1, 2, \cdots, N)$的两个集合就称为**概率空间**，记作$X$，$P(\ )$。各个符号$x_i$的概率称为**先验概率**。当信宿收到一个符号消息y_j后，信宿可以计算信源各消息的条件概率$P(x_i|y_j)$，$i=1, 2, \cdots, N$，这种条件概率又称为**后验概率**。

互信息量定义为后验概率与先验概率比值的对数，即互信息量$I(x_i; y_j)$的定义公式为

$$I(x_i; y_j) \triangleq \log\frac{P(x_i|y_j)}{P(x_i)} \tag{2.4}$$

互信息量的单位与自信息量一样，也取决于对数底。当对数底分别为2、e、10时，互信息量的单位也分别为比特、奈特、哈脱来。

二、互信息量的性质

1．对称性

互信息量的对称性表示为

$$I(x_i; y_j) = I(y_j; x_i) \tag{2.5}$$

上式推导如下：

$$I(x_i; y_j) = \log\frac{P(x_i|y_j)}{P(x_i)} = \log\frac{P(x_i|y_j)\cdot P(y_j)}{P(x_i)\cdot P(y_j)} = \log\frac{P(x_iy_j)/P(x_i)}{P(y_j)}$$

$$= \log\frac{P(y_j|x_i)}{P(y_j)} = I(y_j; x_i)$$

由于互信息量的对称性，它表明了两个随机事件x_i和y_j之间的统计约束程度。当后验概率$P(x_i|y_j)$大于先验概率$P(x_i)$时，互信息量$I(x_i; y_j)$为正值，说明信宿收到的y_j提供了有关x_i的信息。这样，信宿对信源发出的符号消息x_i的不确定度减小了。

2．当x_i和y_j相互独立时，互信息量为零。

当x_i和y_j相互独立时，有$P(x_iy_j)=P(x_i)\cdot P(y_j)$，此时互信息量为

$$I(x_i; y_j) = \log\frac{P(x_iy_j)}{P(x_i)\cdot P(y_j)} = \log\frac{P(x_i)\cdot P(y_j)}{P(x_i)\cdot P(y_j)} = \log 1 = 0$$

这表示x_i与y_j之间不存在统计约束关系。

3．互信息量可为正值或负值。

如前所述，当后验概率大于先验概率时，互信息量为正值。反之，当后验概率小于先验概率时，互信息量就为负值。当后验概率与先验概率相等时，互信息量为零，这就是两个随机事件相互独立的情况。

例 继续讨论第一节的例题，即某地二月份气候的概率空间为

$$X,\ P(\ ) = \begin{cases} x_1(\text{晴}), & x_2(\text{阴}), & x_3(\text{雨}), & x_4(\text{雪}) \\ \frac{1}{2}, & \frac{1}{4}, & \frac{1}{8}, & \frac{1}{8} \end{cases}$$

某一天有人告诉你：“今天不是晴天。”把这句话作为收到的消息x_1'。当收到x_1'后，各种气候的概率变成后验概率了。其中 $P(x_1|x_1')=0$，$P(x_2|x_1')=\frac{1}{2}$，$P(x_3|x_1')=\frac{1}{4}$，

$P(x_4|x_1')=\frac{1}{4}$。依据(2.4)式,可以计算出x_1'事件与各个气候事件之间的互信息量。对x_1事件，因$P(x_1|x_1')=0$,故不必再考虑x_1与x_1'之间的互信息量。对x_2、x_3、x_4事件，可计算得x_1'与x_2、x_3、x_4的互信息量均为1 bit。这表明x_1'分别对x_2、x_3、x_4提供了1 bit的信息量。再从"今天不是晴天"这句话考虑,因$P(x_1)=\frac{1}{2}$，得$P(x_1')=1-P(x_1)=1-\frac{1}{2}=\frac{1}{2}$，所以$I(x_1')=1$bit。这例题说明信宿收到$x_1'$后,可以使$x_2$、$x_3$、$x_4$的不确定度各减少1 bit。

符号x_i与符号对y_jz_k之间的互信息量定义为

$$I(x_i;\ y_jz_k)\triangleq\log\frac{P(x_i|y_jz_k)}{P(x_i)} \tag{2.6}$$

三、条件互信息量

条件互信息量的含义是在给定z_k条件下，x_i与y_j之间的互信息量。**条件互信息量**$I(x_i;y_j|z_k)$定义为

$$I(x_i;\ y_j|z_k)\triangleq\log\frac{P(x_i|y_jz_k)}{P(x_i|z_k)} \tag{2.7}$$

引用(2.7)式，(2.6)式可写成

$$I(x_i;\ y_jz_k)=I(x_i;\ z_k)+I(x_i;\ y_j|z_k) \tag{2.8}$$

上式推导过程如下:

$$I(x_i;\ y_jz_k)=\log\frac{P(x_i|y_jz_k)}{P(x_i)}=\log\left[\frac{P(x_i|y_jz_k)}{P(x_i|z_k)}\cdot\frac{P(x_i|z_k)}{P(x_i)}\right]$$

$$=\log\frac{P(x_i|z_k)}{P(x_i)}=\log\frac{P(x_i|y_jz_k)}{P(x_i|z_k)}=I(x_i;\ z_k)+I(x_i;\ y_j|z_k)$$

(2.8)式表明：一个联合事件y_jz_k出现后所提供的有关x_i的信息量$I(x_i;\ y_jz_k)$等于z_k事件出现后提供的有关x_i的信息量$I(x_i;\ z_k)$，加上在给定z_k条件下再出现y_j事件后所提供的有关x_i的信息量$I(x_i;\ y_j|z_k)$。

在(2.8)式中，y_j和z_k的位置可以互换，即

$$I(x_i;\ y_jz_k)=I(x_i;\ y_j)+I(x_i;\ z_k|y_j) \tag{2.9}$$

第三节 通信熵

一、平均不确定度(熵)

一个信源总是包含着多个符号消息，各个符号消息又按概率空间的先验概率分布。当各个符号的出现概率相互独立时，这种信源称为**无记忆信源**。无记忆信源的平均不确定度是各个符号的不确定度的数学期望(即概率加权的统计平均值)。无记忆信源的**平均不确定度**$H(X)$定义为

$$H(X)\triangleq E(I(x_i))=\sum_X P(x)\cdot I(x)=\sum_X P(x)\cdot(-\log P(x))=-\sum_X P(x)\log P(x) \tag{2.10a}$$

若信源X中的各符号的概率空间简化表示为

$$\begin{matrix} X \\ P(\ \ ) \end{matrix} = \begin{cases} x_1,\ x_2,\ \cdots,\ x_N \\ P_1,\ P_2,\ \cdots,\ P_N \end{cases}$$

平均不确定度$H(X)$的公式就可写成

$$H(X) \triangleq -\sum_{i=1}^{N} P_i \log P_i \tag{2.10b}$$

因为X中各符号x的不确定度$I(x)$为非负值，$P(x)$也是非负值，且$0 \leqslant P(x) \leqslant 1$，故信源的平均不确定度$H(X)$也是非负量。平均不确定度$H(X)$的定义公式与统计热力学中熵的表示形式相同，故又把$H(X)$称为信源$X$的熵。**信源熵**是表征信源的平均不确定度，平均自信息量是消除信源不确定度时所需要的信息的量度。两者在数值上是相等的，但含义并不相同。$H(X)$的值可作为信源X中任一符号消息所携带的平均信息量，也是唯一地确定信源X中任一符号消息时所需要的最小平均信息量。在信源熵的公式中，当某一符号x_i的概率P_i为零时，$(P_i \log P_i)$在熵公式中无意义，为此规定此时的$(P_i \log P_i)$也为零。当信源X中只含一个符号x时，必有$P(x) = 1$，因此信源熵$H(X)$为零。

二、条件熵

条件熵是在联合符号集合XY上的条件自信息量的联合概率加权统计平均值。在给定y条件下，x的条件自信息量为$I(x|y)$，进一步在给定Y（即各个y）条件下，X集合的**条件熵**$H(X|Y)$定义为

$$H(X|Y) \triangleq \sum_{XY} P(xy) I(x|y) = -\sum_{XY} P(xy) \log P(x|y) \tag{2.11}$$

相应地，在给定X（即各个x）条众下，Y集合的**条件熵**$H(Y|X)$为

$$H(Y|X) \triangleq \sum_{XY} P(xy) I(y|x) = -\sum_{XY} P(xy) \log P(y|x) \tag{2.12}$$

下面的推导可以说明求条件熵时要用联合概率加权的理由。

先取在一个y条件下，X集合的条件熵$H(X|y)$为

$$H(X|y) = \sum_{X} P(x|y) \cdot I(x|y) = -\sum_{X} P(x|y) \log P(x|y)$$

进一步把$H(X|y)$在Y集合上取数学期望，就得到条件熵$H(X|Y)$，即

$$H(X|Y) = \sum_{Y} P(y) H(X|y) = -\sum_{XY} P(y) P(x|y) \log P(x|y)$$

$$= -\sum_{XY} P(xy) \log P(x|y)$$

三、共熵

共熵是联合符号集合XY上的每个元素对xy的自信息量的概率加权统计平均值，即**共熵**$H(XY)$定义为

$$H(XY) \triangleq \sum_{XY} P(xy) I(xy) = -\sum_{XY} P(xy) \log P(xy) \tag{2.13}$$

共熵$H(XY)$与熵$H(X)$及条件熵$H(Y|X)$之间存在下列关系：

$$H(XY)=H(X)+H(Y|X) \tag{2.14}$$

上式可根据$I(xy)=I(x)+I(y|x)$在XY联合集合上取统计平均值得出。相应的，共熵还有下列关系：

$$H(XY)=H(Y)+H(X|Y) \tag{2.15}$$

四、各种熵的性质

下面用定理来说明熵的各个性质。

定理2.1 信源X中包含M个不同符号时，信源熵$H(X)$有

$$H(X)\leqslant\log M \tag{2.16}$$

当且仅当X中各个符号的概率全相等时，上式取等号。

证明 自然对数具有性质$\ln\omega\leqslant\omega-1$，$\omega>0$，当且仅当$\omega=1$时，该式取等号。这个性质可用图2.2表示。

$$H(X)-\log M=\sum_X P(x)\log\frac{1}{P(x)}-\sum_X P(x)\log M$$

$$=\sum_X P(x)\log\frac{1}{MP(x)}$$

令$\omega=\frac{1}{MP(x)}$，引用$\ln\omega\leqslant\omega-1$的关系，得

$$H(X)-\log M\leqslant\sum_X\left[\frac{1}{M}-P(x)\right]\log e$$

$$=\left[\sum_X\frac{1}{M}-\sum_X P(x)\right]\log e$$

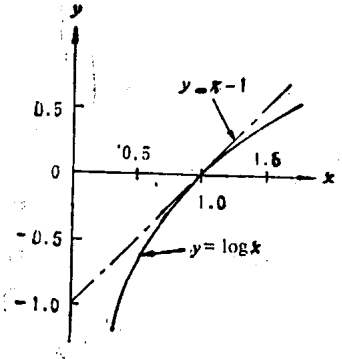

图2-2 自然对数的性质

上式方括号内的第一项和第二项均为1，故得$H(X)-\log M\leqslant 0$，移项后就得(2.16)式的结果。当且仅当$\omega=\frac{1}{MP(x)}=1$，即$P(x)=\frac{1}{M}$时，(2.16)式等号成立。

定理2.1说明：当X集合中各个符号消息以等概率($P=\frac{1}{M}$)出现时，可得到最大信源熵为

$$H(X)_{max}=\log M \tag{2.17}$$

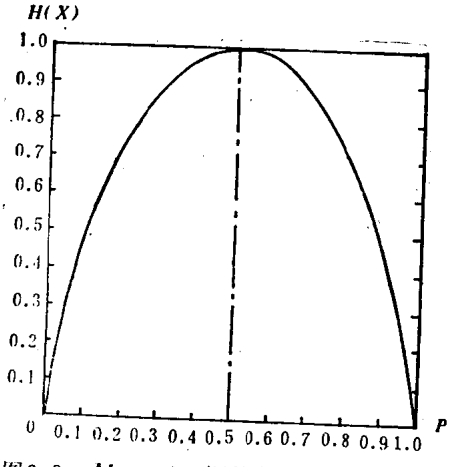

图2-3 $M=2$时信源熵与概率的关系图

若信源只含两个符号消息，即$M=2$，可设一个符号的概率为p，另一个符号的概率为$(1-p)$，该信源熵为

$$H(X)=-\{p\log p+(1-p)\log(1-p)\} \tag{2.18}$$

这个信源熵$H(X)$与概率p的关系示于图2-3。当$p=0.5$时，有最大信源熵$H(X)_{max}$为1 bit。

定理2.2 $H(P_1, P_2, \cdots, P_j, P_{j+1}, \cdots, P_k, \cdots, P_N)=H(P_1, P_2, \cdots, P_j+P_k, P_{j+1}, \cdots, 0, \cdots, P_N)+(P_j+P_k)H\left(\frac{P_j}{P_j+P_k}, \frac{P_k}{P_j+P_k}, 0, \cdots, 0\right)$。 (2.19)

这条定理的证明省略，改用一个例题来说明定理的应用情况。