

高等学校电子信息类系列教材

信息论与信息传输

张树京 齐立心 编著



清华大学出版社
<http://www.tup.tsinghua.edu.cn>



北京交通大学出版社
<http://press.bjtu.edu.cn>



高等学校电子信息类系列教材

信息论与信息传输

张树京 齐立心 编著

清华大学出版社
北京交通大学出版社
·北京·

内 容 简 介

本书是按照高等学校电子信息类本科专业与研究生基础课的教学内容编写的。本书的特点是将信息论和信息传输的基本内容紧密地融合为一体,并以信息论的观点和方法来审视信息传输中出现的各种关系和问题,以及介绍与信息传输有关的信息和信号处理方法。全书共分6章,其内容安排如下:第1章信息论基础,第2章信源编码,第3章信道编码,第4章数字信息传输,第5章多媒体信息传输,第6章无线信息传输。本书适用于高等学校电子信息类专业,作为本科、研究生教材或教学参考用书;也可供成人教育和培训班学员参考。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

本书防伪标签采用特殊防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

信息论与信息传输 / 张树京, 齐立心编著. —北京: 清华大学出版社; 北京交通大学出版社, 2005.6

(高等学校电子信息类系列教材)

ISBN 7-81082-382-5

I. 信… II. ①张… ②齐… III. ①信息论-高等学校-教材 ②信息传输-高等学校-教材 IV. ①G201 ②TN919.1

中国版本图书馆 CIP 数据核字 (2005) 第 051158 号

责任编辑: 韩 乐 特邀编辑: 新 正

出版者: 清华大学出版社 邮编: 100084 电话: 010-62776969 <http://www.tup.com.cn>

北京交通大学出版社 邮编: 100044 电话: 010-51686414 <http://press.bjtu.edu.cn>

印刷者: 北方交通大学印刷厂

发 行 者: 新华书店总店北京发行所

开 本: 185×260 印张: 22 字数: 563 千字

版 次: 2005 年 7 月第 1 版 2005 年 7 月第 1 次印刷

书 号: ISBN 7-81082-382-5/TN·36

印 数: 1~4000 册 定价: 29.00 元

本书如有质量问题, 请向北京交通大学出版社质监局反映。对您的意见和批评, 我们表示欢迎和感谢。

投诉电话: 010-51686043, 51686008; 传真: 010-62225406; E-mail: press@center.bjtu.edu.cn。

前 言

在高等学校电子信息类本科专业与研究生基础课教学计划中,通常将信息论和信息传输原理安排为两门独立的课程,分别介绍信息传输理论(或称信息论)和信息传输原理(或称通信原理)方面的内容。编写本书的意图在于期望改进传统的课程体系,将这两门课程的相关内容紧密地融合为一体,并以信息论的观点和方法来审视信息传输中出现的各种关系和问题,以及介绍与信息传输有关的信息和信号处理方法。

本书的特点可以归纳如下。第一是努力将信息论的基本观点、基本概念和基本方法融合到信息传输技术和信息传输系统中去,并尽量以信息传输理论来指导解决信息传输中出现的各种关系和矛盾。例如:用信息论中信源编码定理来指导信源压缩编码技术;用信息论中信道编码定理来指导差错控制编码技术;用信息论中信源与信道匹配的基本概念来指导设计最佳的信息传输系统等。第二是将传统称为通信系统的概念拓展为信息传输系统,这是因为随着数字信息化时代的到来,信息传输的内容已经由单一的语音通信业务扩展为集数据、文字、语音、图像与音频视频为一体的多媒体信息传输,无论在内容与形式上都与传统的报话通信有很大的区别,其结果是带来了信息传输技术的更新换代,并出现了许多应用广泛的新型信息传输系统,例如交通信息系统、金融信息系统、物流信息系统等。第三是将数字信息、多媒体信息和无线信息传输都放在同一个指标体系内进行衡量和比较,从信息论的角度出发,不难发现它们有许多共同之处,同时又有不少的区别。例如:在数字信息传输系统中研究不同调制技术的改进,使得信道利用率尽量提高,接近于理想传输容量;在多媒体信息传输系统中着重研究信源压缩编码技术,以获得逼近于理想的最佳编码方法,使得信源效率最高;在无线信息传输系统中强调充分利用有限的无线频率资源,研究无线信道中的抗干扰措施,并着重考虑多用户通信移动化问题。第四是本书强调信息传输网络的重要性,因为在传统的通信原理教材中介绍信息传输设备(硬件)的功能和技术比较多,但信息传输网络的质量既受传输设备的性能影响,也包含交换(节点)设备的重要因素在内,特别是在网络通信中讲究国际上统一的通信协议(软件),这是网络通信的核心。因此,再沿用通信系统原理的内容已经远远不够全面和难以深入,只有从信息传输系统扩展到信息传输网络的角度,才能覆盖信息传输的全部内涵。

总之,本书是本着与时俱进的精神,力求在内容上推陈出新,留出更多的篇幅介绍最新的信息传输技术和信息传输系统,以及信息传输的组网技术。本书删除了模拟信息传输的内容,因为目前它已经退出了信息传输技术的舞台,取而代之的是无线移动通信技术和宽带IP网络技术,正是它们吹响了21世纪信息传输技术向前迈进的号角,正在不断地创新和发展。

根据上述编写原则,本书共有6章。第1章是信息论基础,着重介绍信息熵的概念和计算方法,以及信源模型、信道模型和信息率失真理论。第2章是信源编码,着重介绍各类信源的不同编码技术,特别是视频类信源(图像)编码在多媒体信息传输中比较重要。第3章是信道编码(也称差错控制编码),重点介绍分组码和卷积码,尤其是近年来用得较多的循环码和交织码。第4章是数字信息传输,介绍基带数字信息传输和频带(调制后)数字信息

传输两大类，其中多元调制数字信息传输技术更为常用；另外，还介绍数字信息传输网络的基本内容，其中包括电信网和计算机网，前者的代表是分组数据传输网络，而后者的代表是局域网。第5章是多媒体信息传输，介绍多媒体信息业务和多媒体信息终端，重点在于国际上制定的多媒体信息编码标准，其中包含了各种信源编码技术的综合运用；另外，着重介绍多媒体信息传输网络，特别是宽带综合数字业务网（B-ISDN）和互联网技术。第6章是无线信息传输，除了介绍一般的天线和电波传播基本内容外，着重介绍扩大移动通信系统容量和提高抗干扰能力的码分多址（CDMA）扩频技术；同时，以蜂窝和集群两类移动通信系统为代表，介绍无线信息传输网络的组网技术和系统性能。

在编写本书时力求做到内容前后协调，使用的符号和单位一致，注意理论联系实际，适当地引入数学推导，但着重阐明其物理意义和应用场合。各章内容还可以根据实际教学情况加以选用。但因受篇幅限制，仅适当将各章小结和习题纳入本书范围，读者如有必要可以参考相关教材。有关随机过程和噪声分析的内容放在本书附录内作参考。

本书适用于高等学校电子信息类专业，作为本科、研究生教材或教学参考用书，也可供通信工程技术人员或成人教育和培训班学员学习参考。

作 者

2005年5月

目 录

第 1 章 信息论基础	1
1.1 信息量	1
1.1.1 信息的定义	1
1.1.2 信息的量度	3
1.1.3 信息熵	4
1.1.4 连续消息的信息量	9
1.2 信源	12
1.2.1 信源模型	12
1.2.2 离散信源的熵	17
1.2.3 连续信源的熵	20
1.2.4 信源编码定理	21
1.3 信道	22
1.3.1 信道模型	22
1.3.2 离散信道容量	25
1.3.3 连续信道容量	29
1.3.4 信道编码定理	32
1.4 信息率失真理论	33
1.4.1 失真函数	33
1.4.2 信息率失真函数	34
1.4.3 离散信源的 $R(D)$ 函数	36
1.4.4 连续信源的 $R(D)$ 函数	37
小结	39
习题	41
第 2 章 信源编码	43
2.1 信源编码器	43
2.1.1 信源编码器模型	43
2.1.2 信源冗余度	46
2.1.3 信源效率	46
2.1.4 信源编码方法	48
2.2 字符类信源编码	49
2.2.1 最佳编码(熵编码)	49
2.2.2 算术编码	55
2.3 音频类信源编码	58
2.3.1 量化编码	58
2.3.2 差值编码	65

2.3.3 预测编码声码器	70
2.4 视频类信源编码	72
2.4.1 正交函数变换编码	72
2.4.2 多值正交函数集	75
2.4.3 离散矩阵变换编码	77
小结	80
习题	81
第3章 信道编码	83
3.1 差错控制	83
3.1.1 差错控制方式	83
3.1.2 差错控制编码	84
3.1.3 有限域算术	88
3.2 线性分组码	89
3.2.1 简单的分组码	89
3.2.2 一致监督矩阵	92
3.2.3 生成矩阵	95
3.2.4 伴随式	96
3.2.5 汉明码	97
3.3 循环码	99
3.3.1 码多项式	99
3.3.2 循环码的编译码	103
3.3.3 BCH 码	105
3.3.4 RS 码	108
3.4 卷积码	109
3.4.1 卷积码的编码过程	110
3.4.2 最大似然译码	113
3.4.3 维特比译码	114
3.5 级连码	116
3.5.1 乘积码	118
3.5.2 交错码	119
小结	122
习题	124
第4章 数字信息传输	125
4.1 基带数字信号传输	125
4.1.1 基带数字信号码型	125
4.1.2 基带数字信号功率谱	127
4.1.3 基带数字信号传输	131
4.1.4 基带数字信号的误码性能	137
4.2 二元调制数字信号传输	141

4.2.1 ASK 数字信号传输	141
4.2.2 FSK 数字信号传输	146
4.2.3 PSK 数字信号传输	152
4.3 多元调制数字信号传输	158
4.3.1 MASK 数字信号传输	158
4.3.2 MPSK 数字信号传输	159
4.3.3 MSK 数字信号传输	161
4.3.4 MPSK 数字信号传输	168
4.3.5 MQAM 数字信号传输	172
4.4 网格编码调制(TCM)	176
4.5 数字信息传输网络	180
4.5.1 电路交换数据网(CDN)	180
4.5.2 分组交换数据网(PDN)	181
4.5.3 数字数据网(DDN)	186
4.5.4 帧中继(FR)	188
4.5.5 局域网(LAN)	190
小结	195
习题	198
第 5 章 多媒体信息传输	199
5.1 多媒体信息传输业务	199
5.1.1 多媒体信息特征	199
5.1.2 多媒体信息的业务配置	200
5.1.3 多媒体业务的服务质量(QoS)	203
5.2 多媒体信息编码	205
5.2.1 JPEG 标准	207
5.2.2 MPEG 标准	208
5.2.3 H.261 标准	213
5.2.4 H.263 标准	215
5.2.5 音频压缩编码标准	215
5.3 多媒体信息终端	216
5.3.1 基于 H.320 标准的会议电视终端	218
5.3.2 基于 H.324 标准的可视电话终端	219
5.3.3 基于 H.322 标准的可视电话终端	220
5.3.4 基于 H.323 标准的会议电视终端	221
5.3.5 基于 H.321 标准的会议电视终端	224
5.3.6 基于 H.331 标准的会议电视终端	224
5.4 多媒体信息传输网络	226
5.4.1 宽带综合业务数字网(B-ISDN)	227
5.4.2 异步转换模式(ATM)网	231

5.4.3 互联网	236
5.4.4 因特网(Internet)	240
5.4.5 高速城域网(MAN)	245
5.4.6 宽带接入网(B-AN)	247
小结.....	257
习题.....	258
第 6 章 无线信息传输	260
6.1 无线电波的辐射和接收	260
6.2 天线及其特性参数	268
6.3 电波传播及场强计算	275
6.3.1 自由空间中传播	275
6.3.2 地球曲率的影响	279
6.3.3 大气不均匀性的影响	280
6.3.4 衰落与多径效应	282
6.4 多址技术	284
6.4.1 频分多址(FDMA)	284
6.4.2 时分多址(TDMA)	286
6.4.3 码分多址(CDMA)	288
6.5 扩频技术	291
6.5.1 直接序列扩频(DS-SS)	293
6.5.2 跳频扩频(FH-SS)	299
6.6 无线信息传输网络	303
6.6.1 无线数据网	303
6.6.2 公众移动通信网(无线蜂窝系统)	309
6.6.3 专用移动通信网(无线集群系统)	319
小结.....	328
习题.....	330
附录 A 概率积分表	332
附录 B 误差函数表	333
附录 C 随机噪声分析	335
C.1 高斯型白噪声	335
C.1.1 理想白噪声	335
C.1.2 限带白噪声	336
C.2 窄带高斯噪声	338
C.3 余弦信号加窄带高斯噪声	341
参考文献	344

第 1 章 信息论基础

本章重点介绍香农信息论的基本内容,它包括对信息的科学定义,各种信息量的计算方法,引入信息熵的概念,以及分析信源模型和信道模型,计算不同信源的信息传输速率(或称时间熵),不同信道的信道容量,特别是香农信道容量公式成为衡量信息传输极限的基础,因此本章内容是信息传输理论的核心,是实现信息传输系统的理论根据。

1.1 信息量

1.1.1 信息的定义

自从 1928 年美国学者哈特莱(Hartly)提出信息的概念以来,信息这个名词就与社会生活和经济发展产生了解之缘。有人断言,物质、能源和信息是人类生存和社会进步的必要条件,可见信息在现代社会中的作用和地位是何等重要。

信息是一切事物运行状态和运动方式的表征,它来源于物质的运动,是物质的一种固有属性。可以说没有物质,没有物质的运动,就没有信息。但是信息不等同于物质,信息不是物质本身,它只是反映物质的运动状态和方式。同时,信息与能量也有着密切的关系,没有能量,事物就不能运动,当然就没有信息。但是信息也不等于能量,能量是事物运动的原因,信息则是事物运动的结果。

另外,物质和能量受到空间和时间的限制,但是信息原则上却可以延伸和开拓到无限的空间和时间。物质和能源只存在于客观世界,但信息除了客观存在外还接受主观世界的影响,例如信息的内涵就与感知者的理解、思维能力和爱好、欣赏水平等有关。有人将信息分为客观信息和有效信息两类。前者以统计概率作基础,用信息的量度、传输速率、信息容量等来衡量,又称为狭义信息。后者则是在统计信息的基础上,进一步考虑信息的逻辑含义和实效内容,并用平均效用度来衡量,可称为广义信息。本书则主要研究狭义信息,又称香农信息论,有关广义信息论的内容可在信息科学的论著内找到。

1948 年美国著名学者香农(Shannon)连续发表论文,阐述通信的数学理论,引入信息熵的概念,提出信源和信道编码定理,解决了通信系统中传输有效性和可靠性的矛盾,找到了理想信道容量的极限值,从此形成了完整的狭义信息论,故又称为香农信息论,它奠定了信息传输理论的基础。从 20 世纪 60 年代开始研究的广义信息论,将信息的概念渗透到生物医学、仿生学、心理学、语言学及社会经济学等各个领域,并将信息论与电子学、计算机、自动控制 and 系统工程学科相结合,形成了一门广泛的综合学科,称之为信息科学。

现有各种信息传输技术和信息传输系统均与信息传输理论有关,后者可以说是香农信息论的核心,它研究信息的量度方法、信息传输模型,以及解决信息传输系统中有效性和可靠性这对矛盾的关键技术,例如调制技术、编码技术及抗干扰技术等。同时,还对多媒体信息传输、

无线信息传输和多用户信息传输网络中的理论问题提供了数学基础,这些内容正是本课程的学习重点。

1. 消息和信息

这里还必须澄清一些基本概念。首先是信息和消息有何区别?可以这样说,消息是信息传输的具体对象,而信息是抽象化的消息。比如说在电报通信中电文是消息,在电话通信中话音是消息,在电视中画面图像是消息,在雷达中目标距离、高度、方位等参量是消息,在遥测和遥控系统中一些测量数据和指令也是消息。很明显,各种消息在物理特征上极不相同,各种消息的组成亦不可能相同。所有消息都可以归纳成两大类:一类是离散消息,如电报的电文,它是由有限个字符所组成的符号序列;另一类是连续消息,如电话的话音,它是随时间连续变化的过程。无论是连续消息还是离散消息,它们都有一个共同的性质,这就是消息的随机性。发信端在发送消息时可以随心所欲地发出这样或那样的序列或过程,收信端在收到消息之前是无法预测的,这就是说消息的出现是随机的。

在信息传输系统内传递的消息本身无法量度,但其不确定性是可以量度的。因此,人们又引用信息的概念来测度消息的不确定性。信息的含义要比消息更广泛,而且比较抽象,例如在遗传工程中要研究细胞核中的信息,在生物工程中要研究包含在生物体内的各种信息,以及资源信息、地震信息、气象信息等。甚至在社会科学中也存在着各种信息,例如经济信息、生活信息、艺术信息等。因此,信息的内容是多种多样的,其中有些是比较具体的,但也有不少是比较抽象的。可以说,消息就是信息范畴中的一种具体内容,信息传输系统也就包含在信息系统的领域之中。既然消息是随机的,那么信息也是随机的,并且根据哈特莱关于信息的科学定义可以量度其随机性,在香农信息论中还给出了衡量不同信息的量度方法,这样使信息的概念更加明确和科学化。

传递信息是通信的目的,每一个消息都包含着一定的信息量。那么信息量如何确定,它与哪些因素有关呢?哈特莱给出的信息定义是信息量大小与消息发生的概率有密切关系。例如,某人告诉我们一件非常可能发生的消息,比起告诉我们一件不太可能发生的消息来说,它所传递的信息量就比较少。因此,消息发生的概率可以作为人们预期程度的量度单位,它和信息量大小有关。换句话说,在信息量度中最基本的就是不确定性的概念,亦即我们对消息的内容越是不能确定,则消息所包含的信息量就越大。如果我们能够预测给定消息的内容是什么,那就没有传送什么新的消息,因此信息量就很少,甚至等于零。

2. 信息的科学定义

既然消息中包含信息的量度是基于消息出现的概率,并且信息量的大小和消息发生的概率有着相反的关系。如果消息是确定的(即概率为1),那么它所包含的信息量为零;如果消息是完全不可能的(即概率为零),那么它包含着无穷的信息量。这就提示我们,信息量可用消息发生概率的倒数来表示。另外,如果我们得到的不只是一个消息,而是若干个独立的消息之和,那么总的信息量应该是每个消息的信息量之和,这就是说信息量的定义应该满足相加性的条件。因此,在信息论中信息的科学定义为

$$I = \log \frac{1}{P} = -\log P \quad (1-1)$$

式中 P 为消息发生的概率,也称先验概率; I 为从消息发生中能够得到的信息量; $\log$ 是对数。

对数的底决定着量度信息的单位。若取2为底,则 I 的单位为二进制单位,即比特(bit);

若取 e 为底,则 I 的单位为自然单位,即奈特(nit)。在实用中通常都取 2 为底,它的单位是比特。例如,一个以等概率出现的二进制码元(0 或 1)所包含的信息量为 1 比特,即当 $P(0) = P(1) = 1/2$ 时它们各自的信息量分别为

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

实际上这里已经引入随机事件的不确定性(或称不肯定性)概念。为简化起见,以后用 lb 来代替 $\log_2$,信息量单位是比特。

1.1.2 信息的量度

根据上述有关信息的科学定义,在香农信息论中可以针对不同情况给出下列各种不同的信息度量方法。

1. 自信息量

如上所述,任意随机事件的自信息量定义为其出现概率的对数的负值。若单个随机字符 X 出现的概率为 $P(X)$,那么它的自信息量 $I(X)$ 定义为

$$I(X) = -\log P(X) \quad (1-2)$$

如果这里对数取 2 为底,则单位应是比特。例如,随机字符 X 出现的概率为 $P(X) = 1/8$,则它所包含的自信息量就是 3 个比特。

若在一个信源内有两个字符 X 和 Y ,则它们的联合自信息量可定义为

$$I(X, Y) = -\log P(XY) \quad (1-3)$$

其中 $P(X, Y)$ 为同时出现 X 和 Y 的联合概率,故 $I(X, Y)$ 可称为联合自信息量。当 X 和 Y 是两个独立的字符,则它们的联合自信息量就是各自的自信息量之和,即 $I(XY) = I(X) + I(Y)$ 。

2. 条件自信息量

条件自信息量的定义是用其条件概率对数的负值来量度。若随机字符 X 有在另一字符 Y 给定条件下的条件概率为 $P(X|Y)$,那么它的条件自信息量 $I(X|Y)$ 可以定义为

$$I(X|Y) = -\log P(X|Y) \quad (1-4)$$

因此,随机字符的条件自信息量可以理解为在限定条件下唯一地能确定该字符所必须提供的信息量。由于每一个随机事件的概率必定在 $(0 \sim 1)$ 范围内,所以无论自信息量还是条件自信息量都是非负值,这是自信息的属性。

3. 互信息量

假定随机字符 X 和 Y 来自不同的两个信源,或者 X 代表发送的字符,而 Y 代表接收的字符,则互信息量的定义可定义为

$$I(X, Y) = \log \frac{P(X|Y)}{P(X)} \quad (1-5)$$

其中 $P(X)$ 是字符 X 的先验概率,而 $P(X|Y)$ 是在给定字符 Y 后的条件概率,故也可称为后验概率。因此,互信息量 $I(X, Y)$ 实际上是后验概率与先验概率比值的对数。当对数的底均为 2 时互信息量的单位也是比特。

互信息量 $I(X, Y)$ 和自信息量 $I(XY)$ 或条件自信息量 $I(X|Y)$ 的属性有所不同,它们表现出以下特点。

(1) 互易性

互信息量具有互易性,即

$$I(X, Y) = I(Y, X) \quad (1-6)$$

也就是说,在信息传输系统中发出信息的发端与接收信息的收端是可以互换的,这也体现了双向通信的原则。

(2) 统计相关性

根据概率论可知,联合概率等于先验概率与条件概率的乘积,即

$$P(XY) = P(X)P(Y|X) = P(Y)P(X|Y)$$

或者

$$P(X|Y) = \frac{P(XY)}{P(Y)}$$

将它代入式(1-5)后得

$$I(X, Y) = \log \frac{P(XY)}{P(X)P(Y)} \quad (1-7)$$

但是在 X 与 Y 保持统计独立的条件下 $P(XY) = P(X)P(Y)$, 即其联合概率就等于各自先验概率的乘积,或者说此时条件概率就演变为先验概率了。此时, X 与 Y 的互信息量 $I(X, Y) = 0$, 意即在两个统计独立的字符之间不存在互信息量。反过来说,只要 X 和 Y 之间存在相关性,则其互信息量 $I(X, Y)$ 就不可能等于零。

(3) 互信息量可正可负

前面提到,自信息量 $I(X)$ 具有非负性,但是互信息量则不一定。由式(1-5)可见,当后验概率大于先验概率,即 $P(X|Y) > P(X)$ 时互信息量 $I(X, Y)$ 是正值;但当后验概率小于先验概率时互信息量就是负的。

这里还要强调一下,自信息量和互信息量的对应关系,即自信息量相当于后验概率为 1 时的互信息量。

4. 条件互信息量

条件互信息量 $I(X, Y|Z)$ 的定义是指在给定 Z 条件下 X 与 Y 之间的互信息量。如用数学公式表示,则条件互信息量的定义为

$$I(X, Y|Z) = \log \frac{P(X, Y|Z)}{P(X|Z)} \quad (1-8)$$

其中 $P(X, Y|Z)$ 是在给定 Z 条件下对 X 和 Y 的条件概率,而 $P(X|Z)$ 是在给定 Z 条件下仅对 X 的条件概率。条件互信息量的属性与互信息量相同。

1.1.3 信息熵

信息熵是香农信息论研究的核心,简单地讲信息熵就是平均信息量。熵是紊乱程度的测度,香农将熵的概念引入信息领域,因此信息熵也是衡量信息紊乱程度的测度。下面介绍在不同情况下信息熵的计算公式。

1. 平均自信息量(信源熵)

前面已讲过,要唯一地确定信源的一个字符 X , 所需要的信息量是它的自信息量 $I(X)$ 。实际信源往往包含着多个字符,而且各个字符的出现概率往往是不同的。如果各个字符的出现为相互独立时,则这种信源称为无记忆信源。无记忆信源的平均自信息量是各字符自信息量的概率加权平均值(统计平均值),即为

$$H(X) = \sum_x P(X) I(X) \quad (1-9)$$

用式(1-2)代入后得

$$H(X) = \sum_x P(X) [-\log P(X)] = - \sum_x P(X) \log P(X)$$

所以平均自信息量 $H(X)$ 也可直接定义为

$$H(X) = - \sum_x P(X) \log P(X) \quad (1-10)$$

上式中的 $H(X)$ 是唯一确定信源 X 中任意一个字符所需要的平均信息量,也称为信源熵,它是信源概率 $P(X)$ 的函数。

由于自信息量 $I(X)$ 为非负量,且 $0 \leq P(X) \leq 1$, 所以信源熵 $H(X)$ 也是非负量。仅当 $P(X)=0$ 和 $P(X)=1$ 时它才等于零。当 $P(X)=0$ 时,说明该字符 X 并不出现,所以它的自信息量为零。当 $P(X)=1$ 时,这说明该信源只有一个字符,且该字符是必然出现的,所以传输这个字符消息并没有提供任何信息量。

信源熵满足不等式

$$H(X) \leq \log M \quad (1-11)$$

式中 M 是信源 X 中的字符数目。当且仅当信源 X 中各字符的出现概率 $P(X)$ 都等于 $\frac{1}{M}$ 时,式(1-11)才取等号。此时, $H(X)$ 为最大值,并取名为最大熵,即

$$H_{\max}(X) = \log M \quad (1-12)$$

例如,信源 X 中只有两个字符,即 $M=2$, 则该信源的最大熵 $H_{\max}(X) = 1$ 比特。假设其中一个消息的出现概率为 P , 另一个消息的出现概率必为 $(1-P)$, 则该信源的熵为

$$H(X) = - [P \log P + (1-P) \log (1-P)] \quad (1-13)$$

$H(X)$ 作为 P 的函数的关系曲线示于图 1-1。由图可见,最大熵出现在 $P = (1-P) = 0.5$ 处,且该曲线以 $P = 0.5$ 为轴对称。式(1-13)可称为最大熵定理。

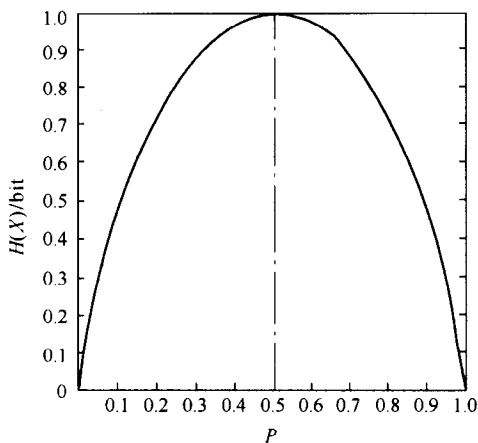


图 1-1 $M=2$ 时信源熵与概率 P 的关系

2. 条件熵

条件熵是联合信源的条件自信息量的概率加权平均值。若给定 X 条件下 Y 的条件自信息量为 $I(Y|X)$, 则它在 XY 集合上的概率加权平均值定义为

$$H(Y|X) = \sum_{XY} P(XY) I(Y|X) \quad (1-14)$$

这里 $H(Y|X)$ 就称为条件熵。根据自信息量的定义,也可把条件熵直接定义为

$$H(Y|X) = - \sum_{XY} P(XY) \log P(Y|X) \quad (1-15)$$

式中取和的范围应包括 XY 集合中的所有值。这里要注意条件熵是用字符 X 和 Y 的联合概率 $P(XY)$, 而不是用条件概率 $P(Y|X)$ 进行加权的。

同理,条件熵 $H(X|Y)$ 可定义为

$$H(X|Y) = - \sum_{XY} P(XY) \log P(X|Y) \quad (1-16)$$

3. 共熵

共熵(又称联合熵)是联合信源 XY 中每个字符对 XY 的自信息量的概率加权平均值,即为

$$H(XY) = \sum_{XY} P(XY) I(XY) \quad (1-17)$$

根据式(1-3),共熵也可直接定义为

$$H(XY) = - \sum_{XY} P(XY) \log P(XY) \quad (1-18)$$

共熵 $H(XY)$ 与信源熵 $H(X)$ 和条件熵 $H(Y|X)$ 的关系为

$$H(XY) = H(X) + H(Y|X) \quad (1-19)$$

上式也可根据

$$I(XY) = I(X) + I(Y|X) \quad (1-20)$$

它们是在 XY 集合上取统计平均值后得出的。

同理,共熵也有下列关系,即为

$$H(XY) = H(Y) + H(X|Y) \quad (1-21)$$

另外,联合集 XY 上的条件熵 $H(Y|X)$ 满足下式

$$H(Y|X) \leq H(Y) \quad (1-22)$$

当且仅当 Y 和 X 是相互独立的,即在 $P(Y|X) = P(Y)$ 的条件下,等号成立。

【例题 1-1】有两个同时输出消息的信源,第一个信源能输出 A 、 B 、 C 这 3 个字符,第二个信源能输出 D 、 E 、 F 、 G 这 4 个字符。若以 i 表示第一个信源, j 表示第二个信源,则第一个信源各消息出现的概率为 $P(i)$,第二个信源各字符出现的条件概率为 $P(j|i)$,它们分别示于例表 1-1 内。每一对字符出现的联合概率 $P(ij) = P(i)P(j|i)$,则列于例表 1-2 内,全部 $P(ij)$ 之和应该是 1。

例表 1-1

符 号		i		
概率		A	B	C
$P(i)$		1/2	1/3	1/6
$P(j i)$	D	1/4	3/10	1/6
	E	1/4	1/5	1/2
	F	1/4	1/5	1/6
	G	1/4	3/10	1/6

例表 1-2

符 号		i		
		A	B	C
$P(ij)$	D	1/8	1/10	1/36
	E	1/8	1/15	1/12
	F	1/8	1/15	1/36
	G	1/8	1/10	1/36

由此可求出第一个信源本身的信息熵为 $H(i) = - \sum_i P(i) \lg P(i) = 1.461$ 比特/符号,条件熵 $H(j|i) = - \sum_i \sum_j P(ij) \lg P(j|i) = 1.956$ 比特/符号,这两个信源的共熵是

$$H(ij) = - \sum_i \sum_j P(ij) \lg P(ij) = 3.416 \text{ 比特/每对符号}$$

可以看出, $H(ij)$ 实际上等于 $H(i)$ 与 $H(j|i)$ 之和。

当上述两个信源没有相关性时,共熵 $H(ij)$ 可达到最大值。此时第二个信源输出符号 D

的概率为

$$P(D) = \sum_i P(i)P(D|i) = P(A)P(D|A) + P(B)P(D|B) + P(C)P(D|C) = \frac{91}{360}$$

同理,可求出 $P(E) = \frac{33}{120}$, $P(F) = \frac{70}{360}$, $P(G) = \frac{91}{360}$, 它们之和应当是 1。因此,第二个信源本身的信息熵为 $H(j) = -\sum_j P(j)\lg P(j) = 1.997$ 比特/符号,于是这两个信源的共熵的最大值为

$$H_{\max}(ij) = H(i) + H(j) = 3.458 \text{ 比特/每对符号}$$

如果用 $P(ji)$ 表示第二个信源输出符号 j , 同时第一个信源输出符号 i 的概率, 则由于 $P(ij) = P(ji)$, 故这两个信源的共熵 $H(ij)$ 和 $H(ji)$ 应该是相同的。

4. 平均互信息量

首先讨论 XY 联合集上的平均条件互信息量, 其定义如下:

$$I(X|Y) = \sum_x P(X|Y) I(X, Y) = \sum_x P(X|Y) \log \frac{P(X|Y)}{P(X)} \quad (1-23)$$

可见, 由 Y 所提供的关于 X 集合的平均条件互信息量就等于由 Y 所提供的互信息量 $I(X, Y)$ 在整个 X 集合中的后验概率加权平均值。显然, $I(X|Y) \geq 0$, 意即在 XY 联合集上的平均条件互信息量也具有非负性。当且仅当 X 集合中的各个字符 X 都与 Y 相互独立, 即 $P(X|Y) = P(X)$ 时, 上式才取等号。

平均互信息量则定义为上述平均条件互信息量 $I(X|Y)$ 在整个 Y 集合上的概率加权平均值, 即为

$$I(X, Y) = \sum_Y P(Y) I(X|Y) \quad (1-24)$$

平均互信息量也可定义为

$$I(X, Y) = \sum_{XY} P(XY) I(X, Y) \quad (1-25)$$

式中 $I(X, Y)$ 是 X 和 Y 的互信息量, $P(XY)$ 是它们的联合概率。当 X 和 Y 相互独立时, 各个符号对的互信息量 $I(X, Y) = 0$, 则它们两个信源的平均互信息量 $I(X, Y) = 0$ 。平均互信息量还具有下列属性。

(1) 互易性

$$I(X, Y) = I(Y, X) \quad (1-26)$$

意即 X 和 Y 两个信源可以互换, 其平均互信息量不变。

(2) 平均互信息量与信源熵和条件熵的关系

$$I(X, Y) = H(X) - H(X|Y) \leq H(X) \quad (1-27)$$

或者

$$I(X, Y) = H(Y) - H(Y|X) \leq H(Y) \quad (1-28)$$

(3) 平均互信息量与共熵的关系

$$I(X, Y) = H(X) + H(Y) - H(XY) \quad (1-29)$$

(4) 平均互信息量具有非负性

$$I(X, Y) \geq 0$$

当且仅当 X 和 Y 相互独立时等号成立。

在信息传输理论中,平均互信息量具有十分重要的意义。

(1) 有干扰信道的 $I(X, Y)$

有干扰信道上接收消息 Y 所提供的有关传输消息 X 的平均互信息量 $I(X, Y)$, 等于唯一地确定消息 X 时所需要的平均自信息量 $H(X)$ 减去接收 Y 消息后确定消息 X 所需要的条件熵 $H(X|Y)$ 。条件熵 $H(X|Y)$ 可看做是由于受信道噪声而损失掉的平均信息量, 故由于损失掉一部分信息量, 使得唯一地确定信源发出的消息 X 就显得信息量不足。条件熵 $H(X|Y)$ 也可以看做是由于信道噪声所造成的对信源消息的平均不稳定性, 故香农把 $H(X|Y)$ 称为疑义度, 意即接收到的信息应该等于发送的信息减去对接收信息被检验后发送信息仍然存在的不确定性(信息熵)。

用另一种观点来看待, 在有扰信道上所传输的平均互信息量 $I(X, Y)$ 可看做是唯一地确定接收消息 Y 所需要的平均自信息量 $H(Y)$, 减去当信源发出的消息 X 为已知时欲确定接收消息 Y 所需要的条件熵 $H(Y|X)$ 。因此, 条件熵 $H(Y|X)$ 可看做是唯一地确定信道噪声所需要的平均信息量, 通常称它为噪声熵。

(2) 无干扰信道的 $I(X, Y)$

对于无干扰信道的理想情况, 因为信道内没有噪声, 所以不损失任何信息量, 此时 $H(X|Y)$ 和 $H(Y|X)$ 都为零。于是就有

$$I(X, Y) = H(X) = H(Y)$$

这时信源发出的消息 X 和信宿接收到的消息 Y 是一一对应的, 所以不会产生差错。

(3) 噪声极大信道的 $I(X, Y)$

另一个极端情况是信道内的噪声极大, 以致使 $H(X|Y) = H(X)$ 。在这种情况下, 传输的平均互信息量 $I(X, Y)$ 等于零。这说明信宿收到 Y 消息后不能提供有关信源发出消息 X 的任何信息量。对于这种信道, 传输的信息量全部在信道内损失掉了, 故称为全损信道, 这也就是 X 和 Y 相互独立的情况。

图 1-2 示出二元对称有扰信道的一般情况, 其中数字代表传输概率。

【例题 1-2】 已知信源 X 包含 x_0 和 x_1 两种消息, 它们的出现概率相等为 $P(x_0) = P(x_1) = 0.5$, 这两种消息在有干扰的二元对称信道上传输。给定的传输概率分别为 $P(y_0|x_0) = 0.98$, $P(y_1|x_0) = 0.02$, $P(y_0|x_1) = 0.2$ 和 $P(y_1|x_1) = 0.8$, 其中 y_0 和 y_1 是信宿收到的两种消息, 它们均在 Y 集内。首先将信道的传输概率用信道矩阵 $\mathbf{P}(Y|X)$ 来表示, 此时信道矩阵 $\mathbf{P}(Y|X)$ 为

$$\begin{array}{cc} & \begin{matrix} y_0 & y_1 \end{matrix} \\ \begin{matrix} x_0 \\ x_1 \end{matrix} & \begin{bmatrix} 0.98 & 0.02 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$

再根据公式 $P(XY) = P(X)P(Y|X)$ 可计算各个联合概率为

$$P(x_0y_0) = P(x_0)P(y_0|x_0) = 0.5 \times 0.98 = 0.49$$

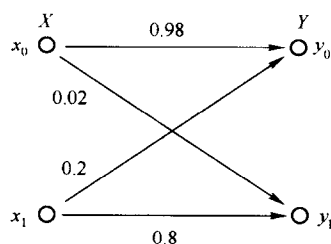


图 1-2 二元对称有扰信道传输特性