



信息与电子学科硕士研究生学习辅导丛书

信息理论与编码 学习辅导及精选题解

傅祖芸 编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

信息与电子学科硕士研究生学习辅导丛书

信息理论与编码

学习辅导及精选题解

傅祖芸 编著

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书是学习信息理论与编码课程的教辅用书,旨在为硕士研究生的课程学习提供指导和帮助。并且它是本科生的信息论与编码习题解答一书的姊妹篇。

本书以香农信息理论为核心内容,同时还包含信源压缩编码、纠错码以及网络信息论等内容。全书共分8章,每章由基本内容和学习要点、习题详解和精选习题解答三部分组成。基本内容和学习要点部分是对各章节知识点的高度概括。另两部分是对习题做了详细推演和解答。其中精选习题是直接选自于国内外同类的经典教材。

全书概念清晰,解题详细,一题多解,便于自学。书中习题属中度难题和难题,它具有广泛性和代表性,适合研究生的学习。因此,本书不只局限于某本教科书,它可以与不同版本同类教材配套使用。

本书可供高等学校通信、信息、电子类及相关专业的教师和硕士研究生参考使用,也可供有关科技人员学习专业基础理论时参考。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

信息理论与编码学习辅导及精选题解/傅祖芸编著. —北京:电子工业出版社,2004.7

信息与电子学科硕士研究生学习辅导丛书

ISBN 7-120-00106-X

I. 信... II. 傅... III. ①信息论—研究生—教学参考资料②信源编码—编码理论—研究生—教学参考资料③信道编码—编码理论—研究生—教学参考资料 IV. TN911.2

中国版本图书馆CIP数据核字(2004)第052215号

责任编辑:陈晓莉 特约编辑:李双庆

印 刷:北京牛山世兴印刷厂

出版发行:电子工业出版社

北京市海淀区万寿路173信箱 邮编100036

经 销:各地新华书店

开 本:787×980 1/16 印张:21 字数:504千字

印 次:2004年7月第1次印刷

印 数:5000册 定价:28.00元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系。联系电话:(010)68279077。质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

前 言

信息理论与编码是国内外高等院校电子信息、通信与电子科学等专业硕士研究生必修的一门专业基础课。有些院校也将其列为博士学位研究生入学考试的专业基础课。尽管各院校开设的课程名称、内容组织有所不同,但都是以香农信息理论为其核心内容。

信息理论与编码是一门理论性和系统性很强的课程。它涉及到众多学科,需要广泛的数学知识。为了能透彻理解和掌握信息与编码理论的基本概念、基本理论和分析方法,做习题是不可缺少的。通过习题的训练,可以加深对理论和概念的理解,并增强分析和解决实际问题的能力。为此,作者在编著《信息论——基础理论与应用》一书的同时编写了本书,它是该书的扩展。

全书共分8章,各章均由基本内容和学习要点、习题详解、精选习题解答三部分组成(其中,第1章只编排了基本内容和学习要点)。每章的基本内容和学习要点部分是对各章节的基本概念、基本定理和计算公式的深入概括,以便读者在解题时做到概念清晰,思路正确。习题详解部分,给出了各题的解题思路、推演过程或较规范的证明步骤。有的题还提供两种以上解题方法,并以提示方式指出易混淆的错误概念或易出现错误求解的原因。精选习题解答部分挑选了国内外同类的经典教材的习题,直接用英文原题列出,并给出详细的解题方法。

本书收集的习题是作者近二十年教学的实践期间,从国内外有关书籍中精心挑选、反复推敲和设计而逐渐积累起来的。多数题意在加深学生对概念和理论的理解,培养分析和解决问题的综合能力。少数题用以开拓和深化信息与编码理论实际应用方面的内容,以提高学生的学习兴趣。本书各章的习题都是些中等难题和难题,适合研究生或有一定基础的读者使用。有关大量的基本题是在本书的姊妹篇——供本科生所用的习题解答一书中给出。

任何一本习题解答无疑是学习该门课程的重要辅导材料。但作者认为,只有在读者独立思考、亲自独立运算习题的基础上,习题解答书籍才能起到它应有的作用。而且,本书中某些习题的巧妙解题思路和方法,也是作者在几十年教学实践中受学生的启发,总结提高得来的。因此在独立思考和独立解题的基础上,将本书作为参考和辅助资料,并希望通过你的演绎给出一些更简捷的解题方法。

本书部分解题曾得到研究生沈青、刘泉、彭一凡、李晓东、牛华宁、施海航、王君庭、周治、许晓东等的协助,同时也得到袁东风、董克俭、路而红等老师对本书编写的建议和帮助,张晋民老师对书稿进行认真的审读,在此一并表示诚挚的谢意。

本书所参阅的国内外一些经典著作,均列于本书参考书目中,在此谨向有关作者表示深切的谢意。

电子工业出版社陈晓莉同志对本书的出版做了大量工作并提出了宝贵的意见,使本书顺利出版,在此也深表感谢。

书中的不妥和错误之处,恳请广大读者给予批评指正。

作者
2004 年 4 月

目 录

第 1 章 绪论	1
1.1 基本内容和学习要点	1
1.1.1 信息的概念及常用术语	1
1.1.2 通信系统的统一模型	2
1.1.3 信息传输系统的模型	3
第 2 章 离散信源及其信息测度	4
2.1 基本内容和学习要点	4
2.1.1 信源的分类	4
2.1.2 基本信源的数学模型	4
2.1.3 离散信源的信息熵	5
2.1.4 信息熵的基本性质	5
2.1.5 离散无记忆扩展信源的信息熵	7
2.1.6 离散平稳信源的信息熵	7
2.1.7 马尔可夫信源及其信息熵	9
2.1.8 信源剩余度	11
2.1.9 加权熵	11
2.2 习题详解	12
2.3 精选习题解答	24
第 3 章 离散信道及其信道容量	52
3.1 基本内容和学习要点	52
3.1.1 信道的分类	52
3.1.2 离散信道的数字模型	52
3.1.3 离散信道的平均互信息及其特性	54
3.1.4 多个随机变量之间的平均互信息	57
3.1.5 离散信道的信道容量及其计算方法	58
3.1.6 常见信道的平均互信息和信道容量	59
3.1.7 无记忆 N 次扩展信道的 $I(\mathbf{X}; \mathbf{Y})$ 和容量 C^N	60
3.1.8 独立并联信道的互信息和信道容量	60
3.1.9 串接信道互信息和信道容量	60
3.1.10 信道剩余度	60

3.2 习题详解	60
3.3 精选习题解答	85
第4章 波形信源和波形信道	108
4.1 基本内容和学习要点	108
4.1.1 连续信源的差熵	108
4.1.2 多维连续平稳信源的信息熵	108
4.1.3 波形信源的差熵	110
4.1.4 差熵的性质	110
4.1.5 最大差熵定理	111
4.1.6 连续信源熵的变换	111
4.1.7 熵功率和熵功率不等式	112
4.1.8 波形信道和连续信道的分类和数学模型	112
4.1.9 连续信道和波形信道的平均互信息及其特性	113
4.1.10 连续信道和波形信道的信道容量	115
4.1.11 连续信道与信息传输的匹配	116
4.2 习题详解	117
4.3 精选习题解答	141
第5章 无失真信源编码	164
5.1 基本内容和学习要点	164
5.1.1 信源编码和码的类型	164
5.1.2 即时码及其树图构造法	164
5.1.3 惟一可译码存在的充要条件	165
5.1.4 惟一可译码的判断法	165
5.1.5 渐近等分割性和 ϵ 典型序列	165
5.1.6 无失真等长信源编码定理	166
5.1.7 无失真变长信源编码定理(香农第一定理)	167
5.1.8 无失真信源编码定理和数据压缩	167
5.1.9 最佳二元码	168
5.1.10 霍夫曼码和其他编码方法	168
5.1.11 几种实用的无失真信源编码方法	170
5.2 习题详解	171
5.3 精选习题解答	191
第6章 有噪信道编码	204
6.1 基本内容和学习要点	204
6.1.1 译码准则	204
6.1.2 平均错误概率	205
6.1.3 费诺不等式	205
6.1.4 信道编码的编、译基本原则	206

6.1.5 联合 ϵ 典型序列	206
6.1.6 有噪信道编码定理及其逆定理	207
6.1.7 信源—信道编码定理	207
6.1.8 纠错码的分类	208
6.1.9 线性分组码的基本概念	208
6.1.10 线性分组码	209
6.1.11 汉明码及其扩展码	210
6.1.12 循环码	211
6.2 习题详解	212
6.3 精选习题解答	220
第 7 章 保真度准则下的信源编码	237
7.1 基本内容和学习要点	237
7.1.1 失真度(失真函数)与平均失真度	237
7.1.2 信息率失真函数	240
7.1.3 信息率失真函数的特性	240
7.1.4 $R(D)$ 函数的参量表述及其计算	241
7.1.5 常见信源的 $R(D)$ 函数	244
7.1.6 保真度准则下信源编码定理(香农第三定理)	245
7.1.7 信息—传输定理(联合有失真信源信道编码定理)	245
7.2 习题详解	246
7.3 精选习题解答	271
第 8 章 网络信息论	284
8.1 基本内容和学习要点	284
8.1.1 通信网信道的分类	284
8.1.2 多个随机变量的联合典型序列	284
8.1.3 多址接入信道	285
* 8.1.4 广播信道	287
* 8.1.5 中继信道	288
8.1.6 反馈信道	289
8.1.7 相关信源无失真编码	289
8.1.8 相关信源的多址接入信道	290
* 8.1.9 具有边信息的无失真信源编码	290
* 8.1.10 具有边信息的率失真信源编码	291
8.2 习题详解	291
参考书目	328

第 1 章 绪 论

1.1 基本内容和学习要点

1.1.1 信息的概念及常用术语

1. 情报

情报是人们对于某个特定对象所见、所闻、所理解而产生的知识。

2. 知识

知识是一种具有普遍和概括性质的高层次的信息。

3. 消息

用文字、符号、数据、语言、音符、图片、图像等能够被人们的感觉器官所能感知的形式,把客观物质运动和主观思维活动的状态表达出来就成为消息。消息中包含信息,它是信息的载体,一则消息可以载荷不同的信息,可以通过得到消息从而获得信息。

4. 信号

把消息变换成适合信道传输的物理量,这种物理量就称为信号。信号携带着消息,它是消息的运载工具。同一消息可用不同的信号来表示,同一信号也可表示不同的消息。

5. 香农信息

香农信息定义为事物运动状态或存在方式的不确定性的描述。

6. 可靠性

信息传输的可靠性就是要使信源发出的消息经过信道传输后,尽可能准确地、不失真地在接收端复现。

7. 有效性

信息传输的有效性就是用尽可能短的时间和尽可能少的设备或费用来传送一定数量的信息。

8. 保密性

信息传输的保密性是指在信息传输过程中要隐蔽和保护所传送的消息,使它只能被授权接收者获取,而不能被未授权者接收和理解。

9. 认证性

信息传输的认证性是指接收者能够判断所接收的消息的正确性,验证消息的完整性,

而没有被伪造和篡改。

可靠性、有效性、保密性和认证性体现了现代通信系统对信息传输的全面要求。

1.1.2 通信系统的统一模型

各种通信系统的基本任务是相同的,即“在通信的一端准确或近似地复现从另一端选择出来的消息”。香农将各种通信系统中具有共同特性的部分抽取出来,概括成一个统一的理论模型如图 1.1 所示。

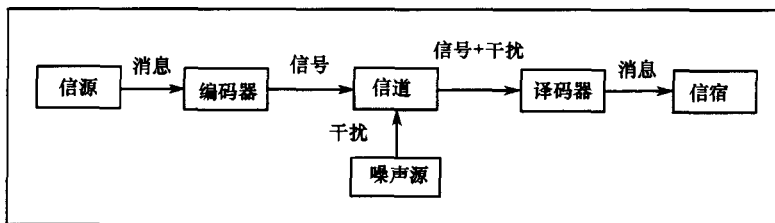


图 1.1 通信系统模型

该模型主要有五个部分组成：

1. 信源

信源是产生消息和消息序列的源。信源输出的消息是随机的、不确定的,但又有一定的规律性。

2. 编码器

编码器的作用是把消息变换成信号,并将信源和信道进行匹配。编码器可分为信源编码器和信道编码器。信源编码器对信源输出的消息进行适当的变换和处理,以提高信息传输的效率;信道编码器对消息进行适当的变换和处理,以提高信息传输的可靠性。编码器还应包括换能、调制、发射等各种变换处理。

3. 信道

信道是传输信号的通道。在通信系统中,实际信道可以有线的,也可以是无线的,并有多重传输媒介。信道同时对信号产生各种干扰和噪声。

4. 译码器

译码器把信道输出的编码信号(已叠加了干扰和噪声)进行反变换。它分为信源译码器和信道译码器。

5. 信宿

信宿是消息传送的对象,即为接收消息的人或机器。

1.1.3 信息传输系统的模型

信息传输系统的基本模型全面概括了现代通信系统对信息传输的要求,如图 1.2 所示。

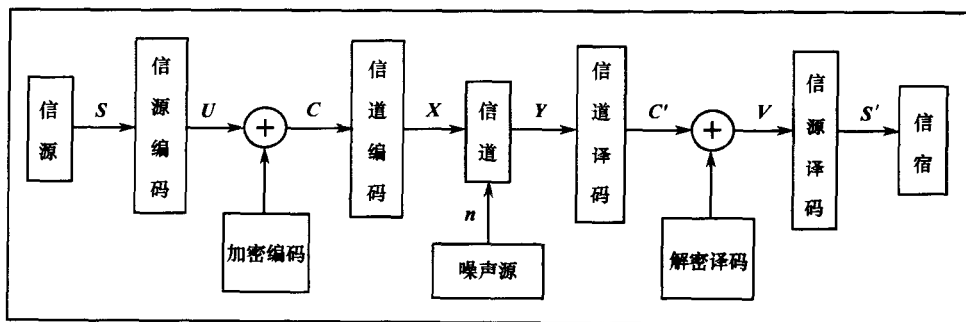


图 1.2 信息传输系统模型

研究这个信息传输系统的目的,就是要找到信息传输过程的共同规律,以提高信息传输的可靠性、有效性、保密性和认证性,使信息传输系统达到最优化。

第 2 章 离散信源及其信息测度

2.1 基本内容和学习要点

2.1.1 信源的分类

信源是信息的来源,是产生消息或消息序列的源泉。

不同的信源输出的消息其随机性质不同。根据消息所具有的随机性质的不同,对信源进行如下分类:

按照消息取值集合以及取值时刻集合的离散性和连续性,信源可分为离散信源(数字信源)和波形信源(模拟信源);

按照某取值时刻消息的取值集合的离散性和连续性,信源可分为离散信源和连续信源;

按照信源输出消息所对应的随机序列的平稳性,信源可分为平稳信源和非平稳信源;

按照信源输出的消息所对应的随机序列中随机变量前后之间有无统计依赖关系,信源可分为无记忆信源和有记忆信源。

(请参见参考书目[1]中图 2.1。)

2.1.2 基本信源的数学模型

根据信源输出消息所对应的不同的随机特性就有不同的信源数学模型。而基本的信源数学模型有以下几种。

1. 离散信源

信源输出的是单个符号或代码的消息,信源符号集的取值是有限的,或可数的,可以用一维离散型随机变量来描述。信源的数学模型就是离散型随机变量 X 的概率空间,表示为

$$\left[\begin{array}{c} X \\ P(x) \end{array} \right] = \left[\begin{array}{cccc} a_1, & a_2, & \cdots, & a_q \\ P(a_1), & P(a_2), & \cdots, & P(a_q) \end{array} \right] \quad \text{并满足} \quad \sum_{i=1}^q P(a_i) = 1$$

其中,样本空间为 $\{a_1, a_2, \cdots, a_q\}$, $q \in I$, I 为正整数集;符号 a_i ($i = 1, 2, \cdots, q$) 出现的概率为 $P(a_i) = P(x = a_i)$ 。信源的概率空间是一个完备集。

这是最基本的离散信源。

2. 连续信源

信源输出的是单个符号或代码的消息,但信源符号集的取值是连续的,可以用一维连续型随机变量来描述。相应的信源的数学模型就是连续型随机变量的概率空间,表示为

$$\left[\begin{array}{c} X \\ p(x) \end{array} \right] = \left[\begin{array}{c} (a, b) \\ p(x) \end{array} \right] \quad \text{或} \quad \left[\begin{array}{c} \mathbf{R} \\ p(x) \end{array} \right]$$

并满足 $\int_a^b p(x)dx = 1$ 或 $\int_{\mathbf{R}} p(x)dx = 1$

其中, (a, b) 是连续随机变量 X 的取值区间, $\mathbf{R}$ 表示全实数集,而 $p(x)$ 是连续随机变量 X 的概率密度函数。

2.1.3 离散信源的信息熵

1. 自信息

自信息即为某事件 a_i 发生所含有的信息量。事件的自信息定义为

$$I(a_i) \triangleq -\log P(a_i)$$

式中 $P(a_i)$ 是事件 a_i 发生的概率。自信息的单位有几种:以 2 为底的对数时单位是比特(bit);以 e 为底的对数时单位是奈特(nat);以 10 为底的对数时单位是哈特(hart)。

2. 信息熵

离散随机变量 X 的信息熵就是其概率空间中每个事件所含有的自信息量的数学期望,即

$$H(X) = E[I(a_i)] = - \sum_{i=1}^q P(a_i) \log P(a_i)$$

其单位是:在以 2 为底的对数时为比特 / 符号(bit/symbol);以 e 为底的对数时为奈特 / 符号(nat/symbol);以 10 为底的对数时为哈特 / 符号(hart/symbol)。

3. 信息熵的物理含义

- (1) 信息熵 $H(X)$ 表示了信源输出前,信源的平均不确定性;
- (2) 信息熵 $H(X)$ 表示了信源输出后,每个消息或符号所提供的平均信息量;
- (3) 信息熵 $H(X)$ 反映了随机变量 X 的随机性。

2.1.4 信息熵的基本性质

用一个概率矢量 $\mathbf{P}$ 表示离散信源 X 的概率分布,即令

$$\mathbf{P} = (P(a_1), P(a_2), \dots, P(a_q)) = (p_1, p_2, \dots, p_q)$$

则信息熵可表示为 $\mathbf{P}$ 的函数

$$H(X) = - \sum_{i=1}^q P(a_i) \log P(a_i) = - \sum_{i=1}^q p_i \log p_i = H(p_1, p_1, \dots, p_1) = H(\mathbf{P})$$

又称 $H(\mathbf{P})$ 为熵函数。熵函数具有以下性质:

1. 对称性

$$H(p_1, p_2, \dots, p_q) = H(p_2, p_1, p_3, \dots, p_q) = \dots = H(p_q, p_1, p_2, \dots, p_{q-1})$$

2. 确定性

$$H(1, 0) = H(1, 0, 0) = \dots = H(1, 0, 0, \dots, 0) = 0$$

3. 非负性

$$H(X) \geq 0$$

4. 扩展性

$$\lim_{\epsilon \rightarrow 0} H_{q+1}(p_1, p_2, \dots, p_q - \epsilon, \epsilon) = H_q(p_1, p_2, \dots, p_q)$$

5. 可加性

统计独立的两个信源 X 和 Y , 有 $H(XY) = H(X) + H(Y)$

6. 强可加性

任意两个相互关联的信源 X 和 Y 有

$$H(XY) = H(X) + H(Y|X)$$

或

$$H(XY) = H(Y) + H(X|Y)$$

7. 递增性

$$\begin{aligned} H_{n+m-1}(p_1, p_2, \dots, p_{n-1}, q_1, q_2, \dots, q_m) \\ = H_n(p_1, p_2, \dots, p_{n-1}, p_n) + p_n H_m\left(\frac{q_1}{p_n}, \frac{q_2}{p_n}, \dots, \frac{q_m}{p_n}\right) \end{aligned}$$

其中

$$\sum_{i=1}^n p_i = 1, \sum_{j=1}^m q_j = p_n$$

8. 极值性, 即最大离散熵定理

$$H(p_1, p_2, \dots, p_q) \leq H\left(\frac{1}{q}, \frac{1}{q}, \dots, \frac{1}{q}\right) = \log q$$

9. 上凸性

对任意概率矢量 $\mathbf{P}_1 = (p_1, p_2, \dots, p_q)$ 和 $\mathbf{P}_2 = (p'_1, p'_2, \dots, p'_q)$, 及任意 $0 < \theta < 1$ 有

$$H[\theta \mathbf{P}_1 + (1 - \theta) \mathbf{P}_2] > \theta H(\mathbf{P}_1) + (1 - \theta) H(\mathbf{P}_2)$$

由于 $H(\mathbf{P})$ 是概率矢量 $\mathbf{P}$ 上的严格上凸函数, 所以熵函数存在极大值。

2.1.5 离散无记忆扩展信源的信息熵

1. 离散无记忆扩展信源的数学模型

若信源输出的消息是取值离散的平稳随机序列,并且序列中各随机变量之间彼此统计独立,则此信源称为平稳离散无记忆信源。离散无记忆信源的数学模型与基本离散信源的数学模型相同,也用 $[X, P(x)]$ 概率空间来描述。

离散无记忆信源 X 的 N 次扩展信源记为 X^N ,它的输出消息由 N 个符号序列组成,并且前后符号的出现是彼此无依赖的、统计独立的。它的数学模型是 $[X, P(x)]$ 的 N 重概率空间 $[X^N, P(\alpha_i)]$ 。

设离散无记忆信源的概率空间为

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} a_1, & a_2, & \cdots, & a_q \\ p_1, & p_2, & \cdots, & p_q \end{bmatrix} \quad \sum_{i=1}^q p_i = 1$$

则信源 X 的 N 次扩展信源 X^N 的 N 重概率空间为

$$\begin{bmatrix} X^N \\ P(\alpha_i) \end{bmatrix} = \begin{bmatrix} \alpha_1, & \alpha_2, & \cdots, & \alpha_q^N \\ P(\alpha_1), & P(\alpha_2), & \cdots, & P(\alpha_q^N) \end{bmatrix}$$

其中 $X^N = (X_1 X_2 \cdots X_N)$, $\alpha_i = (a_{i_1} a_{i_2} \cdots a_{i_N})$, $a_{i_k} \in A = \{a_1, a_2, \cdots, a_q\}$
($k = 1, 2, \cdots, N$)

$$P(\alpha_i) = P(X^N = \alpha_i) = \prod_{k=1}^N p_{i_k} \quad \text{并满足} \quad \sum_{i=1}^q P(\alpha_i) = 1$$

2. 离散无记忆扩展信源的信息熵

信源 X^N 的信息熵与信源 X 的信息熵之间的关系为

$$H(X^N) = NH(X)$$

2.1.6 离散平稳信源的信息熵

1. 离散平稳信源的数学模型

若信源输出的消息是取值离散的随机序列,随机序列的任意有限维的概率分布不随时间平移而改变,则称为离散平稳信源。又根据随机序列中各随机变量有否依赖关系分有记忆信源和无记忆信源。

N 维离散平稳无记忆信源就是离散无记忆的扩展信源 X^N 。

而 N 维离散平稳有记忆信源 X 的数学模型为

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} \alpha_1, & \alpha_2, & \cdots, & \alpha_M \\ p(\alpha_1), & p(\alpha_2), & \cdots, & p(\alpha_M) \end{bmatrix} \quad \sum_{i=1}^M p(\alpha_i) = 1$$

其中 $\mathbf{X} = (X_1 X_2 \cdots X_N)$, $\mathbf{x} = (x_1 x_2 \cdots x_N)$, 又 $\alpha_i = (a_{i_1} a_{i_2} \cdots a_{i_N})$

$$P(\alpha_i) = P(a_{i_1} a_{i_2} \cdots a_{i_N}) = p(a_{i_1}) p(a_{i_2} | a_{i_1}) \cdots p(a_{i_N} | a_{i_1} a_{i_2} \cdots a_{i_{N-1}})$$

$$(i = 1, 2, \cdots, M)$$

其中 a_{i_k} ($k = 1, 2, \cdots, N$) 可以取自同一离散符号集, 也可取自不同符号集, 而一般取自同一符号集。

2. 离散平稳信源的信息测度

(1) 联合熵

联合熵是随机序列 $X_1 X_2$ 联合离散符号集上的每个符号对 $a_i a_j$ 的联合自信息量的数学期望, 记为 $H(X_1 X_2)$ 则有

$$H(X_1 X_2) = - \sum_{i=1}^q \sum_{j=1}^q P(a_i a_j) \log P(a_i a_j)$$

X_1, X_2 也可以是具有不同的概率空间的随机变量, 如

$$\begin{bmatrix} X_2 \\ P(x) \end{bmatrix} = \begin{bmatrix} b_1, & b_2, & \cdots, & b_r \\ P(b_1), & P(b_2), & \cdots, & P(b_r) \end{bmatrix} \quad \sum_{i=1}^r P(b_i) = 1$$

则

$$H(X_1 X_2) = - \sum_{i=1}^q \sum_{j=1}^r P(a_i b_j) \log P(a_i b_j)$$

相应地, N 维联合离散随机序列 $X_1 X_2 \cdots X_N$ 的联合熵记为 $H(\mathbf{X})$ 或 $H(X_1 X_2 \cdots X_N)$, 则有

$$H(\mathbf{X}) = - \sum_{i_1} \cdots \sum_{i_N} P(a_{i_1} a_{i_2} \cdots a_{i_N}) \log P(a_{i_1} a_{i_2} \cdots a_{i_N})$$

(2) 平均符号熵

离散平稳信源输出 N 长的信源符号序列中平均每个信源符号所携带的信息量称为平均符号熵, 记为 $H_N(\mathbf{X})$, 则有

$$H_N(\mathbf{X}) = \frac{1}{N} H(X_1 X_2 \cdots X_N)$$

(3) 条件熵

随机序列 $X_1 X_2$ 的联合符号集上的条件自信息量的数学期望为条件熵, 记为 $H(X_1 | X_2)$, 它表示已知前面一个符号 (X_1 发出) 时, 信源将要输出下一个符号 (X_2 发出) 的平均不确定性, 则有

$$H(X_2 | X_1) = - \sum_{i=1}^q \sum_{j=1}^q P(a_i a_j) \log P(a_j | a_i)$$

同理, X_1, X_2 也可以是具有不同的概率空间的随机变量, 如

$$\begin{bmatrix} X_2 \\ P(x) \end{bmatrix} = \begin{bmatrix} b_1, & b_2, & \cdots, & b_r \\ P(b_1), & P(b_2), & \cdots, & P(b_r) \end{bmatrix} \quad \sum_{i=1}^r P(b_i) = 1$$

则

$$H(X_2 | X_1) = - \sum_{i=1}^q \sum_{j=1}^r P(a_i b_j) \log P(b_j | a_i)$$

对于 N 维离散随机序列 $X_1 X_2 \cdots X_N$, 当已知前 $n-1$ 个符号时, 后面将要出现第 n 个

符号的平均不确定性就是条件熵

$$H(X_n | X_1 X_2 \cdots X_{n-1}) = - \sum_{i_1, i_2, \dots, i_n} P(a_{i_1} a_{i_2} \cdots a_{i_n}) \log P(a_{i_n} | a_{i_1} a_{i_2} \cdots a_{i_{n-1}})$$

$$(n = 2, 3, \dots, N)$$

(4) 极限熵

若离散平稳信源当 N 趋于无穷时, 平均符号熵的极限存在, 则称此极限为离散平稳信源的极限熵(也称熵率), 记为 H_∞ , 即

$$H_\infty = \lim_{N \rightarrow \infty} H_N(\mathbf{X}) = \lim_{N \rightarrow \infty} H(X_N | X_1 X_2 \cdots X_{N-1})$$

(5) 离散平稳信源信息熵的性质

当 $H(\mathbf{X}) < \infty$ 时, 有下述性质成立:

① 条件熵 $H(X_N | X_1 X_2 \cdots X_{N-1})$ 随着 N 的增大是非递增的, 即

$$H(X_N | X_1 X_2 \cdots X_{N-1}) \leq H(X_{N-1} | X_1 X_2 \cdots X_{N-2}) \leq \cdots$$

$$\leq H(X_3 | X_1 X_2) \leq H(X_2 | X_1)$$

② N 给定时, 平均符号熵 $\geq$ 条件熵, 即

$$H_N(\mathbf{X}) \geq H(X_N | X_1 X_2 \cdots X_{N-1})$$

③ 平均符号熵 $H_N(\mathbf{X})$ 也是随着 N 的增大而非递增的, 即

$$H_N(\mathbf{X}) \leq H_{N-1}(\mathbf{X}) \leq \cdots \leq H_3(\mathbf{X}) \leq H_2(\mathbf{X}) \leq H(\mathbf{X})$$

④ H_∞ 存在, 并且有

$$H_\infty = \lim_{N \rightarrow \infty} H_N(\mathbf{X}) = \lim_{N \rightarrow \infty} H(X_N | X_1 X_2 \cdots X_{N-1})$$

(6) 各种信息熵之间的关系

① $H(X_2) \geq H(X_2 | X_1)$, 当且仅当 X_1 与 X_2 统计独立时等式成立。

$$\begin{aligned} \text{② } H(X_1 X_2) &= H(X_2 X_1) = H(X_1) + H(X_2 | X_1) \\ &= H(X_2) + H(X_1 | X_2) \end{aligned}$$

$$\begin{aligned} \text{③ } H(\mathbf{X}) &= H(X_1 X_2 X_3 \cdots X_N) \\ &= H(X_1) + H(X_2 | X_1) + \cdots + H(X_N | X_1 X_2 \cdots X_{N-1}) \\ &= \sum_{i=1}^N H(X_i | X_1 X_2 \cdots X_{i-1}) \end{aligned}$$

④ $H(\mathbf{X}) \leq \sum_{i=1}^N H(X_i)$, 当且仅当 $\mathbf{X}$ 中各变量彼此无依赖, 统计独立时, 等式成立。

2.1.7 马尔可夫信源及其信息熵

1. 马尔可夫信源的定义

马尔可夫信源是一类有限长度记忆的非平稳离散信源, 信源输出的消息是非平稳的