



普通高等教育“十二五”规划教材

# 信息论基础与应用

XINXILUN JICHU YU YINGYONG

王燕妮 编著



北京邮电大学出版社  
[www.buptpress.com](http://www.buptpress.com)



普通高等教育“十二五”规划教材

# 信息论基础与应用

王燕妮 编 著



北京邮电大学出版社  
[www. buptpress. com](http://www.buptpress.com)

## 内 容 简 介

《信息论基础与应用》系统地讲述了信息论与编码的基本理论,共8章,内容包括:香农信息论的基本概念,各种离散信源、连续信源及其信息测度,离散信道、连续信道及其信道容量的计算,无失真信源编码定理、限失真信源编码定理及其常用的信源编码方法,有噪信道编码定理及其编码方法,信息论相关学科的应用与发展。在离散信源的信息熵、离散信道、离散信源编码以及信道编码章节后面都提供了相应内容的 MATLAB 仿真源程序,供教学使用。每章《信息论基础与应用》都力求内容精练、易读,强调掌握信息论与编码的基本理论与方法以及在通信系统中的应用。

《信息论基础与应用》可作为高等院校通信与信息类及相关专业本科生或低年级研究生学习的教材或教学参考书,也可供从事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

信息论基础与应用 / 王燕妮编著. --北京:北京邮电大学出版社, 2015. 1

ISBN 978-7-5635-4269-7

I. ①信… II. ①王… III. ①信息论 IV. ①G201

中国版本图书馆 CIP 数据核字(2015)第 002380 号

---

书 名: 信息论基础与应用

著作责任者: 王燕妮 编著

责任编辑: 孔 玥

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: 19

字 数: 492 千字

印 数: 1—2 000 册

版 次: 2015 年 1 月第 1 版 2015 年 1 月第 1 次印刷

---

ISBN 978-7-5635-4269-7

定 价: 38.00 元

· 如有印装质量问题,请与北京邮电大学出版社发行部联系 ·



# 前言

信息论基础即香农信息论,是主要研究信息传输和信息处理的一般规律的科学。自从美国数学家 C. E. Shannon 于 1948 年创立信息论以来,经过众多通信科技工作者的不断努力研究和创新,信息理论取得了巨大的成果。在信息的度量、信息的传输、信源的压缩编码与纠错编码等方面都有重大突破,极大地促进了通信技术的不断发展。

在参考了大量国内外信息论优秀教材和参考资料的基础上,作者结合多年的教学、科研实践经验编写了《信息论基础与应用》。本书以信息熵为基本概念,以香农三个编码定理——无失真信源编码定理、有噪信道编码定理和限失真信源编码定理为核心内容,研究通信系统中信息的度量、信源的编码以及信息通过信道有效和可靠传输等理论问题。

本课程是通信工程、电子信息工程等相关专业高年级本科生及研究生必修的专业基础课。本书以使学生更好地理解信息论基本原理、掌握基本方法、提高分析能力为目的,力求简单易懂,由浅入深、以学生最容易理解的方式介绍信息论的基本内容及应用。

全书共有 8 章。第 1、2、3、4 章是信息理论基础章节,主要阐述信息的概念、信息的度量,信息熵、平均互信息以及数据处理定理;信息传输中的损失熵、噪声熵、平均互信息与各类熵的关系,各种离散信道容量的计算;连续信源的信息测度,连续信道的信息传输率与信道容量。第 5、6、7 章是信息理论编码章节,是香农信息论的核心内容,主要介绍无失真信源编码、有噪信道编码和限失真信源编码的基本概念、定理及编码方法。第 8 章讨论信息理论的应用与其他学科的交叉发展。每章内容都配有适当的例题以及 MATLAB 上机实验环节,为学生提供更多的信息论基本问题解题示范,开阔学生的解题思路,进一步提高信息论课程理论教学的质量。书中一些公式、定义、定理的证明,力图推算简洁,重点突出公式、定理所表达的物理含义。在每章课后,有适量的配套习题,这对于学生深入理解基本概念,掌握和运用基本的解题方法,提高解决与信息论有关的基础性或综合性问题的能力是很有必要的。附录为上机作业。

本书在编写过程中参阅的国内外文献以及网站资料已尽量列出,若有遗漏深表歉意,在此对本书所引用文献的作者深表感谢。西安建筑科技大学李昌华教授、于军琪教授、王慧琴教授、王民教授、张立材副教授对本书的编写提出了许多宝贵的意见,在此一并表示衷心的感谢。

由于水平有限,书中难免有不足和不妥之处,恳请读者批评指正。

作 者  
2014 年 6 月



# 目录

<b>第 1 章 绪论</b> .....	1
1.1 信息的基本概念 .....	1
1.1.1 信息的定义 .....	2
1.1.2 信息的特征 .....	4
1.1.3 信息、消息和信号的关系 .....	4
1.2 信息论研究的内容 .....	5
1.2.1 通信系统模型 .....	5
1.2.2 信息论研究的内容 .....	6
1.3 信息论的建立、发展与应用 .....	8
习题一 .....	11
<b>第 2 章 离散信源</b> .....	12
2.1 信源的数学模型及分类 .....	12
2.1.1 随机变量描述信源输出的消息 .....	13
2.1.2 随机序列描述信源输出的消息 .....	14
2.1.3 随机过程描述信源输出的消息 .....	15
2.2 离散信源的信息熵 .....	16
2.2.1 自信息 .....	16
2.2.2 信息熵 .....	20
2.2.3 信息熵的基本性质 .....	25
2.3 信息熵的唯一性定理 .....	32
2.4 离散无记忆的扩展信源 .....	33
2.5 离散平稳信源 .....	34
2.5.1 离散平稳信源的数学定义 .....	34
2.5.2 二维平稳信源及其信息熵 .....	35
2.5.3 离散平稳信源的极限熵 .....	38
2.6 马尔可夫信源 .....	40
2.6.1 马尔可夫信源的定义 .....	40
2.6.2 $m$ 阶马尔可夫信源 .....	42

2.6.3 马尔可夫信源的信息熵	42
2.7 信源冗余度与自然语言的熵	48
2.7.1 信源冗余度的定义	48
2.7.2 冗余的利用	49
2.8 离散信源的 MATLAB 实现	51
2.8.1 离散信源的相关原理	51
2.8.2 MATLAB 实验内容	51
本章小结	54
习题二	55
<b>第 3 章 离散信道</b>	<b>59</b>
3.1 信道的基本概念	59
3.1.1 信道的分类	60
3.1.2 离散信道的数学模型	60
3.1.3 单符号离散信道	62
3.2 平均互信息	64
3.2.1 信道疑义度	64
3.2.2 平均互信息	65
3.2.3 平均条件互信息	72
3.2.4 平均互信息的特性	74
3.3 信道容量的计算方法	77
3.3.1 离散无噪信道的信道容量	77
3.3.2 对称离散信道的信道容量	79
3.3.3 一般离散信道的信道容量	82
3.4 离散无记忆扩展信道及其信道容量	85
3.5 串联信道的互信息和数据处理定理	89
3.6 信源与信道匹配	91
3.7 信道容量的 MATLAB 实现	92
3.7.1 信道容量的相关原理	92
3.7.2 MATLAB 实验内容	92
本章小结	94
习题三	94
<b>第 4 章 连续信源和连续信道</b>	<b>97</b>
4.1 连续信源的熵及其特性	97
4.1.1 连续信源熵的定义	97
4.1.2 连续信源熵的特性	100
4.2 几种连续信源的熵	101

4.3	连续信源的最大熵	105
4.4	熵功率	107
4.5	连续熵的变换	108
4.6	连续信道及其信道容量	109
4.6.1	连续信道的分类	109
4.6.2	连续信道的平均互信息	110
4.6.3	连续信道平均互信息的特性	110
4.6.4	连续信道的信道容量	112
4.6.5	连续信道编码定理	114
	本章小结	115
	习题四	115
<b>第5章</b>	<b>无失真信源编码</b>	<b>116</b>
5.1	信源编码的基本概念	117
5.1.1	信源编码器	117
5.1.2	信源编码的分类	117
5.1.3	唯一可译变长码与即时码	118
5.2	信源编码定理	122
5.2.1	渐近等分割性和 $\epsilon$ 典型序列	122
5.2.2	等长信源编码定理	124
5.2.3	变长信源编码定理	127
5.3	信源编码方法	130
5.3.1	香农编码	131
5.3.2	霍夫曼码	131
5.3.3	费诺码	135
5.3.4	香农-费诺-埃利斯码	136
5.4	几种实用的无失真信源编码方法	137
5.4.1	游程编码和 MH 编码	137
5.4.2	算术编码	140
5.4.3	LZ 编码	142
5.5	无失真信源编码的 MATLAB 实现	143
5.5.1	无失真信源编码相关原理	143
5.5.2	MATLAB 实验内容	143
	本章小结	148
	习题五	149
<b>第6章</b>	<b>有噪信道编码</b>	<b>151</b>
6.1	信道编码的概念	152

6.1.1	译码规则 .....	152
6.1.2	编码方法 .....	156
6.2	香农第二定理 .....	163
6.2.1	有噪信道编码定理 .....	163
6.2.2	有噪信道编码逆定理 .....	164
6.2.3	信源信道编码定理 .....	165
6.3	差错控制概述 .....	165
6.3.1	二元对称信道和错误格式 .....	165
6.3.2	差错控制分类 .....	165
6.3.3	常用差错控制码 .....	166
6.4	信道编码方法 .....	167
6.4.1	线性分组码 .....	167
6.4.2	汉明码 .....	177
6.4.3	循环码 .....	178
6.4.4	卷积码 .....	184
6.5	MATLAB 实现循环码编码 .....	186
	本章小结 .....	187
	习题六 .....	189
<b>第 7 章</b>	<b>限失真信源编码</b> .....	<b>191</b>
7.1	失真测度与失真矩阵 .....	192
7.1.1	失真函数 .....	192
7.1.2	失真矩阵 .....	192
7.1.3	平均失真度 .....	194
7.2	信息率失真函数 .....	194
7.2.1	$D$ 失真许可信道 .....	194
7.2.2	信息率失真函数的定义 .....	195
7.2.3	信息率失真函数的性质 .....	195
7.3	信息率失真函数的计算 .....	199
7.4	限失真信源编码定理 .....	202
7.4.1	信源编码定理 .....	203
7.4.2	信源编码逆定理 .....	206
7.4.3	联合信源信道编码定理 .....	207
7.4.4	限失真信源编码定理的应用 .....	207
7.5	限失真信源编码 .....	210
7.5.1	预测编码 .....	210
7.5.2	变换编码 .....	212
7.6	信源编码的应用 .....	213



7.6.1	语音信息编码 .....	213
7.6.2	图像信息编码 .....	213
7.6.3	视频信息编码 .....	214
7.7	MATLAB 实现率失真函数曲线 .....	217
7.7.1	实验原理 .....	217
7.7.2	实验内容 .....	217
	本章小结 .....	219
	习题七 .....	219
<b>第 8 章</b>	<b>信息理论的应用与发展 .....</b>	<b>222</b>
8.1	信息理论的应用 .....	222
8.2	信息论与其他学科交叉发展 .....	224
8.2.1	密码学的应用 .....	225
8.2.2	金融工程中的信息论问题 .....	228
8.2.3	信息论在决策树中的应用 .....	230
8.2.4	人工智能 .....	236
8.2.5	生命科学 .....	237
8.2.6	统计学 .....	239
8.2.7	网络信息论 .....	239
<b>附录</b>	<b>.....</b>	<b>241</b>
实验一	离散信源的信息熵 .....	241
一、	实验设备 .....	241
二、	实验目的 .....	241
三、	实验原理 .....	241
四、	实验内容及要求 .....	241
实验二	离散信道 .....	243
一、	实验目的 .....	243
二、	实验原理 .....	243
三、	实验内容及要求 .....	243
实验三	离散信源编码 .....	244
一、	实验目的 .....	244
二、	实验原理 .....	244
三、	实验内容及要求 .....	244
实验四	信道编码 .....	244
一、	实验目的 .....	244
二、	实验原理 .....	244
三、	实验内容及要求 .....	245

习题一参考答案	246
习题二参考答案	247
习题三参考答案	254
习题四参考答案	261
习题五参考答案	262
习题六参考答案	268
习题七参考答案	273
<b>上机作业参考答案</b>	<b>281</b>
实验一 离散信源的信息熵	281
实验二 离散信道	283
实验三 离散信源编码	284
实验四 信道编码	287
<b>参考文献</b>	<b>291</b>

# 第 1 章

## 绪 论

### 一、学习内容

1. 信息的概念。
2. 信息论研究的对象、目的和内容。
3. 信息论的发展简史和现状。

### 二、学习目标

1. 理解信息的概念及通信系统模型。
2. 理解信息论研究的主要目的和内容。
3. 了解信息论的发展简史和现状。

### 三、重点与难点

**重点:**信息的概念及其研究的主要目的和内容。

**难点:**信息的概念。

信息论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。信息论将信息的传递作为一种统计现象来考虑,给出了估算通信信道容量的方法。信息传输和信息压缩是信息论研究中的两大领域,这两个方面又由信息传输定理、信源-信道编码定理相互联系。美国数学家、信息论的创始人香农被称为是“信息论之父”。人们通常将香农于1948年10月发表于《贝尔系统技术学报》上的论文 *A Mathematical Theory of Communication* (通信的数学理论) 作为现代信息论研究的开端。随着信息理论的迅猛发展和信息概念的不断深化,信息论涉及的内容早已超越了狭义的通信工程,进入了信息科学这一更广阔、更新兴的领域。

任何一门科学都有它自身的基本概念,理解和掌握这些基本概念是学习这门科学的基础。传统科学的基本概念是物质和能量,而信息论最基本和最重要的概念就是信息。

本章的主要内容是讨论信息的概念,信息论研究的对象和内容以及信息论的建立与发展。

## 1.1 信息的基本概念

信息(Information)是物质运动规律总和。它是客观事物状态和运动特征的一种普遍形

式,客观世界中大量地存在、产生和传递着以这些方式表示出来的各种各样的信息。

南唐李中《暮春怀故人》诗:“梦断美人沉信息,目穿长路倚楼台。”宋陈亮《梅花》诗:“欲传春信息,不怕雪埋藏。”《水浒传》第四四回:“宋江大喜,说道:“只有贤弟去得快,旬日便知信息。”巴金《家》三一:“二表哥的事情怎样了?为什么连信息也不给我一个?”

现代科学指事物发出的消息、指令、数据、符号等所包含的内容。人类通过获得、识别自然界和社会的不同信息来区别不同事物,得以认识和改造世界。在一切通信和控制系统中,信息是一种普遍联系的形式。陈原《社会语言学》4.1:“按物理学的观念,信息只不过是按一定方式排列起来的信号序列。”在社会交际活动中,这个定义还不够:信息还必须有一定的意义,或者说信息必须是“意义的载体”。《科学画报》1983年第1期:“整个地铁的运行靠一台中心电子计算机控制,它赖以控制的信息,来自装在车站以及车站之间线路上的高灵敏度仪器。”

当今社会,“信息”一词在各种场合都被广泛采用,但如同数学中的“集合”一词一样,要给它下一个严格的定义却异常之难。即使是信息论的奠基人香农在其著名论文《通信的数学理论》中,也没有给信息下一个明确的定义。香农论文发表之后,由于其方法新颖,引来许多专家学者对信息进行深入研究,研究中碰到的首要问题就是要给“信息”一词下一个明确的定义。很多学者都给“信息”下过定义,流行的说法不下百种,而且对此还展开了一些重要的哲学争论。各种说法一则出发点不同,二则所站角度不一样,有些甚至带有较明显的学科倾向,但都在一定层面上对信息概念做了描述。

信息既是信息论的出发点,也是它的归宿。具体地说,信息论的出发点是认识信息的本质和它的运动规律;它的归宿则是利用信息来达到某种具体的目的。

什么是信息?

### 1.1.1 信息的定义

信息论中常常把消息中有意义的内容称为信息。1948年,香农在题为《通信的数学理论》的论文中指出:“信息是用来消除随机不定性的东西”。同年,美国著名数学家、控制论的创始人维纳在《控制论》一书中,指出:“信息就是信息,既非物质,也非能量。”

信息是有价值的,就像不能没有空气和水一样,人类也离不开信息。因此人们常说,物质、能量和信息是构成世界的三大要素。所以说,信息的传播是极其重要与有效的。

信息是事物的运动状态和过程以及关于这种状态和过程的知识。它的作用在于消除观察者在相应认识上的不确定性,它的数值则以消除不确定性的多少,或等效地以新增知识的多少来度量。虽然有着各式各样的传播活动,但所有社会传播活动的内容从本质上说都是信息。

信息相关资料:图片信息(又称作讯息或资讯)是一种消息,通常以文字或声音、图像的形式来表现,是数据按有意义的关联排列的结果。信息由意义和符号组成。文献是信息的一种,即通常讲到的文献信息。信息就是指以声音、语言、文字、图像、动画、气味等方式所表示的实际内容。

香农将信息定义为事物存在方式或运动状态的不确定性的描述。这里的“事物”泛指存在于人类社会、思维活动和自然界中一切可能的对象。“存在方式”指事物的内部结构和外部联系。“运动状态”则是指事物在时间和空间上变化所展示的特征、态势和规律。

信息的基本概念在于它的不确定性,已确定的事物都不含信息。其特征有:

- 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;
- 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;

- 信息可以产生,也可以消失,同时信息可以被携带、存储及处理;
- 信息是可以度量的,信息量有多少的差别。

例如,气象预报:

$$\text{甲:} \quad \begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} \text{晴} & \text{阴} & \text{大雨} & \text{小雨} \\ 1/2 & 1/4 & 1/8 & 1/8 \end{bmatrix}$$

$$\text{乙:} \quad \begin{bmatrix} Y \\ p(y) \end{bmatrix} = \begin{bmatrix} \text{晴} & \text{阴} & \text{大雨} & \text{小雨} \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix}$$

“甲地晴”比“乙地晴”的不确定性小。

**例 1.1** 甲袋红、白球各 50 个,乙袋红、白、蓝、黑球各 25 个。比较从甲袋中取出一个球是红球的事件和从乙袋中取出一个球是红球的事件发生的难易程度,也就是事件发生的不确定性。

“甲袋中取红球”比“乙袋中取红球”的不确定性小。

由此可知某一事物状态出现的概率越大,其不确定性越小,反之,某一事物状态出现的概率越小,其不确定性越大。若某一事物状态出现的概率接近于 1,即预料中肯定会出现的事件,那它的不确定性就接近于 0。

某一事物状态的不确定性的,与该事物可能出现的不同状态数目和各状态出现的概率大小有关。既然不确定性的,大小能够度量,可见信息是可以测度的。

$$\text{概率空间:} \quad \begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_i & \cdots & x_n \\ p(x_1) & \cdots & p(x_i) & \cdots & p(x_n) \end{bmatrix}$$

样本空间指某事物各种可能出现的不同状态,先验概率  $p(x_i)$  即是选择符号  $x_i$  作为消息的概率。对  $x_i$  的不确定性可表示为先验概率  $p(x_i)$  的倒数的某一函数。

自信息:

$$I(x_i) = \log \frac{1}{p(x_i)} \quad (1.1)$$

互信息:

$$I(x_i; y_j) = \log \frac{1}{p(x_i)} - \log \frac{1}{p(x_i | y_j)} \quad (1.2)$$

后验概率  $p(x_i | y_j)$  指接收端收到消息  $y_j$  后而发送端发的是  $x_i$  的概率。

香农定义的信息概念在现有的各种理解中,是比较深刻的,它有许多优点。

- (1) 它是一个科学的定义,有明确的数学模型和定量计算。
- (2) 它与日常用语中信息的含义是一致的。

例如,设某一事件  $x_i$  发生的概率等于 1,即  $x_i$  是预料中一定会发生的必然事件,如果事件  $x_i$  果然发生了,收信者将不会得到任何信息(日常含义),不存在任何不确定性。

因为  $p(x_i) = 1$ ,所以

$$I(x_i) = \log \frac{1}{p(x_i)} = 0$$

即自信息等于零。反之如果  $x_i$  发生的概率很小,一旦  $x_i$  果然发生了,收信者就会觉得很意外和惊讶,获得的信息量很大。

因为  $p(x_i) \ll 1$ ,故得:

$$I(x_i) = \log \frac{1}{p(x_i)} \gg 1$$

(3) 定义排除了对信息一词某些主观上的含义。

### 1.1.2 信息的特征

信息作为客观世界存在的第三要素,具有以下特征。

(1) 可度量:信息可采用某种度量单位进行度量以表征其量值的大小和多少,并进行信息编码。如现代计算机使用的二进制。

(2) 可识别:信息可采取直观识别、比较识别和间接识别等多种方式来获知。

(3) 可转换:信息可以从一种形态转换为另一种形态。如自然信息可转换为语言、文字和图像等形态,也可转换为电磁波信号或计算机代码等。

(4) 可存储:信息可以存储。大脑就是一个天然信息存储器。人类发明的文字、摄影、录音、录像以及计算机存储器等都可以进行信息的存储以保留信息。

(5) 可处理:人脑就是最佳的信息处理器。人脑的思维功能可以进行决策、设计、研究、写作、改进、发明、创造等多种信息处理活动。计算机也具有信息处理功能。

(6) 可传递:信息的传递是与物质和能量的传递同时进行的。语言、表情、动作、报刊、书籍、广播、电视、电话等是人类常用的信息传递方式。

(7) 可再生:信息经过处理后,可以其他形式再生。如自然信息经过人工处理后,可用语言或图形等方式再生成信息。输入计算机的各种数据文字等信息,可用显示、打印、绘图等方式再生成信息。

(8) 可压缩:信息可以进行压缩,可以用不同的信息量来描述同一事物。人们常常用尽可能少的信息量描述一事物的主要特征。

(9) 可利用:信息是物质和事物固有的,是普遍存在的。它具有普遍性,因此信息具有可利用性。

(10) 可共享:由于信息是普遍存在的,因此信息是可共享的。

### 1.1.3 信息、消息和信号的关系

消息是表达客观物质运动和主观思维活动的状态,不讲述详细的经过和细节,以简要的语言文字迅速传播新近事实的新闻体裁,也是最广泛、最经常采用的新闻基本体裁,如文字、语言、图像等。消息传递过程即是消除不确定性的过程:收信者存在不确定,收信前,不知消息的内容;干扰使收信者不能判定消息的可靠性,收信者得知消息内容后,消除原先的“不确定”。

消息的结构:(一)标题:1. 引题(眉题、肩题):交代背景;2. 主标题:概括主要新闻或消息;3. 副标题:补充说明主标题。(二)导语:一般是对事件或事件中心的概述。(三)主体:承接导语,扣住中心,对导语所概括事实作比较具体的叙述,是导语内容的具体化。(四)背景:说明原因、条件、环境等。(五)结语:或小结,或指出事情发展方向等。消息的三个特点:真实性,实效性,传播性。

信息与消息的关系可描述为:形式上传输消息,实质上传输信息;消息具体,信息抽象;消息是表达信息的工具,信息载荷在消息中,同一信息可用不同形式的消息来载荷;消息可能包含丰富的信息,也可能包含很少的信息。

信号(也称为讯号)是运载消息的工具,是消息的载体。从广义上讲,它包含光信号、声信号和电信号等。例如,古代人利用点燃烽火台而产生的滚滚狼烟,向远方军队传递敌人入侵的消息,这属于光信号;当我们说话时,声波传递到他人的耳朵,使他人了解我们的意图,这属于

声信号;遨游太空的各种无线电波、四通八达的电话网中的电流等,都可以用来向远方表达各种消息,这属于电信号。把消息转换成适合信道传输的物理量,如光信号、电信号、声信号和生物信号等,人们通过对光、声、电信号进行接收,才知道对方要表达的消息。

对信号的分类方法很多,按数学关系、取值特征、能量功率、处理分析、所具有的时间函数特性、取值是否为实数等,信号可以分为确定性信号和非确定性信号(又称随机信号)、连续信号和离散信号、能量信号和功率信号、时域信号和频域信号、时限信号和频限信号、实信号和复信号等。

信息、消息和信号三者之间的关系可总结为:消息携带着信息,消息是信息的运载工具;消息是信息的表现形式,信息是消息的具体内容。信号是消息的物理体现。在通信系统中,系统传输的是信号,但本质的内容是消息。消息包含在信号之中,信号是消息的载体。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 信息论研究的内容

### 1.2.1 通信系统模型

通信的基本问题是在存储或通信等情况下,精确或者是近似再现信源发出的消息。在通信领域中,所需要研究的主要内容是通信中的有效性和可靠性,有的时候还要考虑信息传输的安全性。通信系统的一般模型如图 1.1 所示。

#### 1. 信源

信源是产生消息的来源,可以是文字、语言、图像等;可以是连续的,也可以是离散的。信源本身十分复杂,在信息论中一般只是对信源的输出进行研究。信源输出是以符号形式表示具体信息,是信息的载体。尽管信源

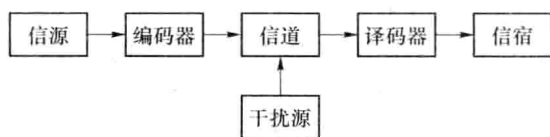


图 1.1 通信系统的一般模型

输出形式很多,但是可以对其进行分类,其表现形式要么是连续的,要么是离散的。如文字、符号、数字等符号或者符号序列,其符号的取值都是可数的,这样的消息就是离散的;对于语音、图像等在时间上连续变化的参量,符号的取值都是不可数的,这样的消息是连续的。无论信源输出的符号是连续的还是离散的,它们都必须是随机的,否则无论是信源的特征研究还是通信研究都没有意义。信源模块主要是研究信源发出消息的统计特征以及信源产生的信息速率。

#### 2. 编码器

编码器是将信源发出的符号转化为适合信道传输的信号的设备,一般包括信源编码、信道编码和调制器等。编码器的模型如图 1.2 所示。

① 信源编码器:主要解决有效性问题,在一定的准则下对信源输出进行变换和处理,目的是提高信息传输的效率,即通过去除信源输出符号的冗余,使信源输出的每个符号携带更多的信息量,从而降低信息传递所需要的符号数量,提高传输效率。

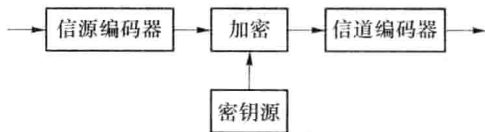


图 1.2 编码器的模型

② 信道编码器:由纠错编码器和调制器组成,目的在于充分利用信道的传输能力,并可靠

地传输信息。

**纠错编码器:**对信源输出进行变换处理,通过增加冗余提高对信道干扰的抵抗力,从而保证信息传输的可靠性。由于信道中存在干扰,数据传递的过程中会出现错误,信道编码可以提供检测或者是纠正数据传输错误的能力,从而提高数据传输的可靠性。

**调制器:**将信道编码的输出变换为适合信道传输要求的信号。信道编码器输出的数字信号并不适合信道的传输,需要对其进行相应的信号变换和调制,然后将变换后的信号送往信道进行传输。

**加密:**为了提高信息传输的安全性,有时需要进行加密处理,这就需要扩展码位。加密处理同时也会降低系统传输效率,即有效性。

### 3. 信道

信道是信息传输的媒质。信道将携带信息的信号从一个地方传送到另一个地方。常见的信道有明线、电缆、光纤、无线电波、纸张、图片甚至空气等。在水中,通信可以采用声波传输,声波传输的媒质是水,所以水也是信道。随着科学技术的发展,大量的信息需要存储,存储器也是信道。

### 4. 干扰源

通信系统中的各部分都会受到干扰,信号的类型不同,经过的信道不同,所遭受的噪声、干扰也有差异。将各种干扰等效成一个方框作用于信道。干扰源的统计特征是划分信道的重要因素,并是决定信道传输能力的决定因素。

干扰源的分类:

① **加性干扰:**由外界引入的随机干扰,如电磁干扰、设备内部噪声,它们与信道输入的信号统计特征无关,信道输出则是输入的干扰之和。

② **乘性干扰:**信号在传播过程中,由于物理条件的变化,如温度、电离层位置的随机变化引起的信号参量的随机变化,此时信道的输出是输入信号与某些随机变量相乘的结果。

信息论就是对干扰进行数学上的描述,确定它们对信号传输的影响,从而给出在无干扰的情况下,信道的传输能力。

### 5. 译码器

译码器是编码器的逆过程,其目的是为了准确或者近似再现信源发出的消息。与编码器相对应,译码器一般是由解调器、信道译码器和信源译码器组成。其作用就是从受干扰的信号里最大限度地提取出有关信源输出消息的信息,尽可能精确地恢复信源的输出并送给信宿。其中心问题就是研究各种可实现的解调和译码的方法。

### 6. 信宿

信宿是信息的载体,即接收消息的人或机器,与信源处于不同地点或存在于不同时刻。它要对传送过来的信息提出可接受的条件,即提出一定的准则,发端将以此来确定对信源处理时所要保留的最小信息量。信宿的数量可以是一个,也可以是多个,取决于具体应用的需要。

## 1.2.2 信息论研究的内容

信息论研究的主要内容是为设计有效的、可靠的通信系统提供理论依据。由于消息中包含着信息,所以消息的传输系统也是信息的传输系统,简称通信系统。人们通过消息的传输和处理过程来研究信息传输和处理过程中的共同规律。信息论是运用概率论与数理统计的方法研究信息传输和信息处理系统中一般规律的新兴学科。核心问题是信息传输的有效性和可靠



性以及两者间的关系。

根据研究内容范围的大小,可对信息论进行分类:

狭义信息论也称经典信息论。主要研究:信息的测度、信道容量、信源编码、信道编码。这部分内容是信息论的基础理论,又称香农信息理论。如图 1.3 所示。

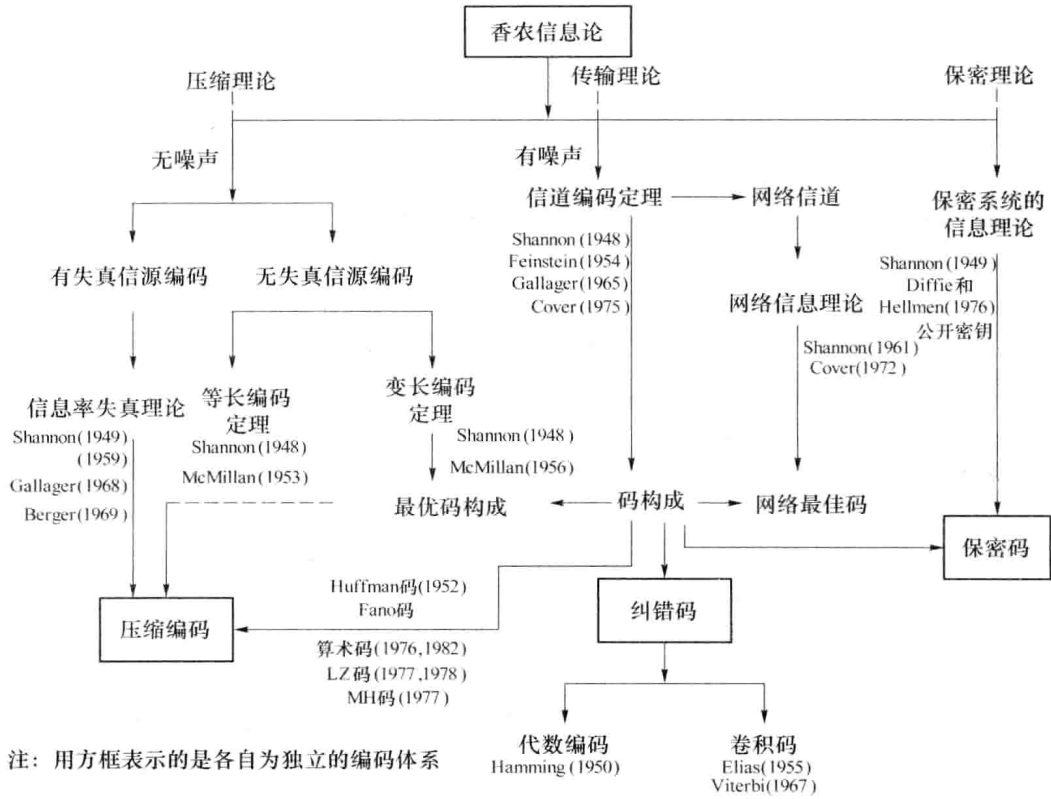


图 1.3 香农信息论的科学体系

一般信息论主要是研究信息传输和处理的问题。除了香农理论以外,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。后一部分内容是以美国科学家维纳为代表。

广义信息论不仅包括上述两方面的内容,而且包括所有与信息有关的自然和社会领域。如模式识别、计算机翻译、心理学、遗传学、神经生理学、语言学、语义学甚至包括社会学中有关信息的问题。

研究一个概括性很强的通信系统,其目的就是要找到信息传输过程的共同规律。一旦总结出这种共同的规律,就可以用来指导具体通信系统的设计,使设计出来的各种通信系统具有更高的可靠性和有效性。可靠性高,就是要使信源发出的信息经信道传输以后,尽可能的准确不失真地再现在接收端。有效性高,就是经济效果好,即用尽可能短的时间和尽可能少的设备来传送一定数量的信息。两者的结合就能使系统达到最优化。

提高可靠性和提高有效性常常会发生矛盾,这就要统筹兼顾。例如,为了兼顾有效性,有时就不一定要求绝对准确的在接收端再现原来的信息,而是允许一定的误差或一定的失真,或者说允许近似地再现原来的消息。

关于信息论研究的具体内容,有人认为信息论只是概率论的一个分支,因为香农信息论确