

王新梅 编著

纠错码浅说

人民邮电出版社

内 容 简 介

这本小册子通俗地介绍各种纠错码的构成原理和实现编、译码的基本方法。全书共十章，分三部分：第一部分介绍纠错码在数字通信中的地位和编码公式；第二部分介绍各种线性纠错码和卷积码；第三部分介绍纠错码的某些实用问题和在通信、计算机中的应用概况。

纠 错 码 浅 说

王新梅 编

人民邮电出版社出版

北京东长安街27号

天津市第一印刷厂印刷

新华书店北京发行所发行

各地新华书店经售

开本：787×1092 1/32 1976年7月第一版

印张：6 8/32 页数：100 1976年7月天津第1次印刷

字数：141千字 印数：1—23,000册

统一书号：15045·总2089—无618

定 价：0.51 元

毛主席语录

什么“三项指示为纲”，安定团结不是不要阶级斗争，阶级斗争是纲，其余都是目。

中国人民有志气，有能力，一定要在不远的将来，赶上和超过世界先进水平。

在生产斗争和科学实验范围内，人类总是不断发展的，自然界也总是不断发展的，永远不会停止在一个水平上。因此，人类总得不断地总结经验，有所发现，有所发明，有所创造，有所前进。

前 言

纠错编码是五十年代提出、六十年代发展起来的、提高数据传输可靠性的一种技术,也是通信理论中一个最基本的问题。近十几年来,由于数字通信、特别是卫星通信的飞速发展,以及作为数据处理、交换中心的电子数字计算机的应用,对数据传输的可靠性,提出了越来越高的要求。特别是对高速、大容量的数字通信系统,如何提高传输数据的可靠性,更成为一个需要迫切解决的问题。

在一般情况下,用通常的方法不能根本或有效地解决数据传输的可靠性,因此纠错编码作为解决这一问题的新途径和新方向,越来越受到通信工作者的重视。

一切真知都是从直接经验发源的。如同所有学科一样,纠错编码理论也来源于长期通信的实践,并在实践中不断得到发展。六十年代后半期以来,数字通信迅速发展,大规模集成电路成批生产,使纠错编码不仅在理论上得到了进一步的发展和完善,而且更在实践中得到了越来越多的应用。

在伟大的无产阶级文化大革命推动下,我国工人阶级和科技人员,坚持独立自主,自力更生,艰苦奋斗,勤俭建国,先后制成了各种类型的纠错码编译码器,并在实际运用中取得了良好的效果。

由于纠错编码主要用于数字通信中,因此为了更好地了解它在数字通信中的地位和作用,以及它的一般原理和方法,本书首先介绍一下通信、特别是数字通信的某些有关的基本概念,然后再介绍某些纠错码如分组码、大数逻辑可译码、纠突发错误码、卷积码以及计算机上应用的某些纠错码的基本原

理，最后介绍一下纠错编码的实用问题和在通信及计算机中当前使用的概况。

要深入了解纠错编码的原理，需要较多的数学、特别是现代代数的知识。但是，要理解它的基本概念并不困难。编写这本小册子的目的，就是想用通俗易懂的文字介绍一点有关纠错编码的最基本原理和实现方法，使它能够为广大读者所了解。但由于作者水平有限，也有可能达不到这个目的。衷心希望同志们提出宝贵意见。

作者

1975.10.6

目 录

第一章 通信系统模型	1
1.1 数字通信概述	1
1.2 二进制编码与十进制编码	4
1.3 数字通信系统模型	7
1.4 信道模型与错误图样	12
第二章 差错控制的基本思想和分类	16
2.1 基本思想	16
2.2 差错控制的分类	17
2.3 抗干扰码的分类	20
第三章 抗干扰码的一般概念	23
3.1 分组码的基本概念	24
3.2 抗干扰码的基本原理和码的距离	25
3.3 编码定理	30
3.4 常用检错码	39
第四章 线性分组码	43
4.1 线性分组码的基本概念	43
4.2 伴随式	54
4.3 纠单个错误的码	61
附录 矩阵的基本知识.....	64
第五章 循环码	70
5.1 循环码的基本概念	73
5.2 BCH码	83
5.3 多项式除法电路	86
5.4 纠一个错误的编译码器	90

5.5	纠一个错误的译码器	92
第六章	大数逻辑可译码	98
6.1	大数逻辑可译码的基本概念	98
6.2	大数逻辑译码器	103
第七章	纠正突发错误的码	109
7.1	纠正单个突发错误的循环码	110
7.2	纠单个突发错误的译码	113
7.3	纠正突发错误和随机错误的码	119
第八章	卷积码	126
8.1	卷积码的一般概念	129
8.2	卷积码的大数逻辑译码	138
8.3	纠突发错误的卷积码	144
8.4	卷积码的概率译码	149
第九章	纠错码的实用问题	161
9.1	用户的要求	161
9.2	信道的错误统计特性	162
9.3	差错控制方式的选择	165
9.4	码的选择及其主要参数的确定	166
第十章	纠错编码实用概况	170
10.1	纠错编码在通信中的应用	171
10.2	纠错编码在计算机中的应用	182

第一章 通信系统模型

1.1 数字通信概述

人们在阶级斗争、生产斗争和科学实验这三项社会实践中，需要互通消息，于是产生了各种互通消息的方法，这就是通信的起源。

为了尽快与可靠地传送消息，在古代，让传送消息的人携带着写在竹帛上的文件、命令，或者其他物证如调动兵力的虎符等，乘坐最快的骏马，通过专门修建的道路——驿道，把消息送到一个站，然后换乘另一匹骏马再往下一站传送，这样一站接一站地把消息传送到收信者手里。这是最原始的通信方式。以后，随着生产和科学技术的发展，人们不仅可以把需要传送的消息写成书信形式，用轮船、汽车、火车、飞机等交通工具，送往收信者，而且可以用电话、电报以更快的速度把消息告诉对方。书信、电话、电报，是人们在现代生活中最常用的几种通信方式。今天，在我们伟大的社会主义祖国，为三大革命服务的，还远不止这三种通信方式。例如：清晨，当我们打开收音机的时候，就可收听到中央人民广播电台传来的《东方红》的庄严歌声，可听到播音员同志向我们传达毛主席党中央的战斗号令，和报告我国各条战线上学理论、抓革命、促生产的大好形势；晚上，当我们打开电视机的时候，在舞台上演的革命样板戏就立刻进入我们眼中；在边远省份，人们可以看见传真版的当天《人民日报》等等。这些都是很有效的通信方式。因为它们是以各自的特色，用语言、文字、图像等向人们

传播消息。

从上述这些通信方式中我们看到，通信的目的就是发信者要把接收者不知道的消息，尽快地和可靠地告诉对方。并且，所有这些方式中，都有以下几个特点：首先，都有一个发信者和一个收信者（广播、电视、报纸杂志的收信者可以有很多人）；其次，有一条把消息告诉对方的通路——信道。这就是上面提到的驿道、运载信件、报纸杂志的火车、飞机、汽车、轮船所需的铁路、航线、公路、航路等；传送电话、电报所需的电话、电报线路等；传送无线电广播、电视所需的无线电频段等；第三，有把传送的消息变成适合在不同信道传送的信号的设备，及把信号变成消息的还原设备。例如把声音变成适合在电话线路上传输的电信号的话筒，把电信号变为声音的耳机。把数字、文字、声音、音乐、图像变成适合在无线或有线信道传播的电信号（这种过程称为调制）的电传机、发射机（包括调制器和功率放大器），以及把这些电信号还原为消息（称这过程为解调）的电传机、接收机（它包括高放和解调器）和电视接收机、收音机等；第四，电信号在信道传送过程中都会受到不同程度的干扰，以致收到的消息与原发送的消息不完全相同，甚至产生错误。例如打电话时，耳机里有时会听到“卡嚓”声音；听广播，特别是听较远的无线电广播时，有很多噪声，甚至突然会听不见声音；在电视机的屏幕上有时能看到闪闪的亮点，妨碍人们观看；收到的电报有时会出现错误等。

“就人类认识运动的秩序说来，总是由认识个别的和特殊的事物，逐步地扩大到认识一般的事物。”我们根据上面列举的这些通信方式的共同目的和共性，可用图(1-1)的一个通信系统模型来概括它们。



图(1-1) 通信系统模型

上述的一些通信方式虽然有共同点，但它们彼此之间也有明显的不同之处。例如书信、电报、报纸杂志等是以文字或数字的形式表示的，也就是传送消息的单位是一个个的文字或数字，代表这些消息的电信号的某一参数（如正弦波或脉冲的振幅、频率、相位等），只取有限个数值，并且不直接或不完全与消息相对应。例如普通的电报是以四个数字代表某一汉字，且代表每个数字的电传机送出的脉冲信号，其高度只能取A和B两个值，表示空号和传号（通常用0和1表示）。一般来说，这些消息或信号不仅在取值上是离散的，而且在时间上也是离散的，称这种消息为离散消息或数字消息，其信号称数字信号。而广播、电话、电视是以声音或图像来表示消息的，而代表这些消息的电信号的某一参数，可取无限多个数值且直接与消息相对应，例如电话机话筒输出的电压，其幅值随说话人的语言连续变化，通常这些消息及代表消息的信号在时间上也是连续的，所以称它们为模拟信号。当然，这二者之间的差别仅只是相对的，在一定条件下可以互相转化。

由于生产和科学技术的发展，特别是电子数字计算机的大量生产和普遍应用，现代通信范围远远超出了人与人之间的联系，而发展到人与机器、机器与机器之间的通信。例如指挥中心对卫星、宇宙飞船、导弹等指令的传送，以及从这些飞行体送回到地面的观察数据；工矿企业、科学研究单位的生产数据、科学实验数据等，往往要送到电子计算机进行加工、计算，并把结果送到相应的设备；又例如自动化防空系统中的防

空雷达，它送出一个探测信号去探测空中的目标，并把探测结果接收下来送到计算机。计算机算出目标的性质、飞行高度、速度、方位距离等数据，并且算出我方的飞机、导弹迎击敌机的时间、方位、速度等参数，并把相应的指令送到机场或导弹发射场，指挥飞机和导弹迎击来犯的敌机或导弹。在这里，传送这些数据、指令的通信，都是在机器与机器，或机器与人之间进行的。因此，从更广泛的通信意义上来说，图(1-1)中的发信者应改为信息源（简称信源），它可以发送语言、音乐、图像和各种数据，受信者可为人、机器等。

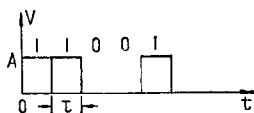
1.2 二进制编码与十进制编码

由上可知，与消息相应的信号可分为数字信号或模拟信号。我们把传输数字信号的通信称为数字通信；传输模拟信号的通信称为模拟通信。对于通信系统，也相应地分为数字通信系统与模拟通信系统。所有这些区分也只是相对的；并且模拟信号经过数字化后，也可在数字通信系统中传输，如数字电话就是一例。

前面说过，数字通信中代表消息的电信号的某一参数只取有限个数值；而在信道中传送的信号，在每一时间单位，其参数值只是其中之一。如果这有限个数为 M ，则称为 M 进制信号。例如由普通电传机送出的脉冲信号，其幅度只取二个值，因此这种电传信号就是二进制信号。

通常在数字通信系统中，从信源发出的或信道中传输的信号，大都采用二进制信号形式，也就是一个消息或数据，仅以二个符号0和1的组合来表示。如图(1-2)中表示用五个0和1的组合组成了一个数字信号，它代表普通电传信号中的数字

“2”。我们把每个持续时间为 τ 秒的0或1的符号,称为一个二进制码元;而把这五个0、1组成的序列称为码长为五的一个二进制码组,它代表某一个数字符号或指令。长为五的二进制序列共有 $2^5 = 32$ 种,故有32个长为五的二进制码组,能代表32种不同的数字符号或指令。像这种表示方法称为二进制编码。通常把这32个码组称为五单位(元)码。表1中列出了国际上通用的五单位码所代表的数字和符号,称为国际第二号五单位码,也就是通常所说的电传码。一般来说,码长为 n 的二进制序列,共有 2^n 种可能的不同组合;能代表 2^n 个不同的符号或指令。



图(1-2) 数字信号

由表1看出,所有数字、英文字母以及电报业务上常用的勤务语言,都可用五单位码来表示。但我们常用的汉字共有几千个,如何用这种五单位电传码来进行通报呢?以前曾谈到过,每个汉字可用四个数字来表示。例如“北”字用“0554”表示,“京”字用“0079”表示。像这种用四个十进制数字的组合来表示汉字的方法,就是一种十进制编码。通过十进制编码把汉字化为数字来表示,然后把每个数字再用其相应的五单位码来传送,即再加一次二进制编码,从而就能以二进制数字通信的方式进行汉字之间的电传通报。例如“北京”两字的十进制码为“0554 0079”,其相应的五单位码为:

01101, 00001, 00001, 01010, 01101, 01101, 11100,
00011。

而“北京”的拼音字“BEIJING”的电传码是:

10011 (B), 10000 (E), 01100 (I), 11010 (J),
01100 (I), 00110 (N), 01011 (G)。

表1

五单位电传码与3:4等
重码3:2数字保护码对照表

号 码	字 母	5 单 位 码	3:4 码	数字保护码
-	A	1 1 0 0 0	0 0 1 1 0 1 0	
?	B	1 0 0 1 1	0 0 1 1 0 0 1	
:	C	0 1 1 1 0	1 0 0 1 1 0 0	
你是谁	D	1 0 0 1 0	0 0 1 1 1 0 0	
3	E	1 0 0 0 0	0 1 1 1 0 0 0	1 0 1 1 0
%	F	1 0 1 1 0	0 0 1 0 0 1 1	
%	G	0 1 0 1 1	1 1 0 0 0 0 1	
	H	0 0 1 0 1	1 0 1 0 0 1 0	
8	I	0 1 1 0 0	1 1 1 0 0 0 0	0 1 1 1 0
元	J	1 1 0 1 0	0 1 0 0 0 1 1	
(K	1 1 1 1 0	0 0 0 1 0 1 1	
)	L	0 1 0 0 1	1 1 0 0 0 1 0	
.	M	0 0 1 1 1	1 0 1 0 0 0 1	
,	N	0 0 1 1 0	1 0 1 0 1 0 0	
9	O	0 0 0 1 1	1 0 0 0 1 1 0	1 0 0 1 1
0	P	0 1 1 0 1	1 0 0 1 0 1 0	0 1 1 0 1
1	Q	1 1 1 0 1	0 0 0 1 1 0 1	0 1 0 1 1
4	R	0 1 0 1 0	1 1 0 0 1 0 0	1 1 0 1 0
'	S	1 0 1 0 0	0 1 0 1 0 1 0	
5	T	0 0 0 0 1	1 0 0 0 1 0 1	0 0 1 1 1
7	U	1 1 1 0 0	0 1 1 0 0 1 0	1 1 1 0 0
=	V	0 1 1 1 1	1 0 0 1 0 0 1	
2	W	1 1 0 0 1	0 1 0 0 1 0 1	1 1 0 0 1
/	X	1 0 1 1 1	0 0 1 0 1 1 0	
6	Y	1 0 1 0 1	0 0 1 0 1 0 1	1 0 1 0 1
"	Z	1 0 0 0 1	0 1 1 0 0 0 1	
回行 >		0 0 0 1 0	1 0 0 0 0 1 1	
换行 ≡		0 1 0 0 0	1 0 1 1 0 0 0	
字母键		1 1 1 1 1	0 1 0 0 1 1 0	
数字键		1 1 0 1 1	0 0 0 1 1 1 0	
间隔		0 0 1 0 0	1 1 0 1 0 0 0	
		0 0 0 0 0	0 0 0 0 1 1 1	
R Q			0 1 1 0 1 0 0	
α			0 1 0 1 0 0 1	
β			0 1 0 1 1 0 0	

“北京”的英语“PEKING”的电传码为：

01101 (P)，10000 (E)，11110 (K)，01100 (I)，
00110 (N)，01011 (G)。

像这种用一定长度的二进制数字的不同组合，来表示数字、拼音字母、各种符号，以及用一定长度的十进制数字的组合表示汉字的编码方法，在通信理论中，看成是一种信源编码。一般来说，对声音、图像的模拟信号也可用信源编码的方法，把它们变为用二进制序列来表示的数字信号。

为什么要把各种数字、文字、符号、声音、图像等最后都变为用二进制序列来表示呢？这是由于二进制序列中的0、1二个数字，特别易于用电流的大小，电压的高低或高频电磁波的频率、振幅或相位的两个不同值来表示，并且能很容易地在机器中进行处理、传输或加工。即使是对于多进制的信号序列，也往往先转换成二进制信号序列，例如四进制的四个数字0，1，2，3，可用00，01，10，11来表示。因此，目前数字通信中或数字电子计算机内部运算、传输的信号序列，大都采用二进制数字序列。此外，数字通信还具有高速、保密、抗干扰、易于自动化等一系列优点。因此，在现代通信中，数字通信越来越占有重要地位，并得到了极其迅速的发展。

1.3 数字通信系统模型

随着我国社会主义革命和建设的飞速发展，要求传送的消息越来越多。因此，一个数字通信系统在单位时间内能够传送的消息越多越好，不仅如此，而且要求传送的消息用户都能够可靠地接收到。也就是说，我们要求一个通信系统能“多快好省”地传送消息或信号。但是，事物总是一分为二的，在一切

通信系统中毫无例外地都存在着妨碍多快好省传送消息的因素。例如通信机器（收、发信机等）各个部件的不理想性和内部噪声；信道中存在的各种干扰：如有线电话电报线路中的串音、瞬断、接插噪声；高频无线电信道和散射信道中的衰落、天电干扰以及人为干扰等；超高频接收机内部的热噪声等。所有这些干扰，都可使传输中的二进制数字序列的0变成1，1错成0。特别严重的是这种差错有时可能引起“数字互变”。如表1中，数字“7”的五单位码是“111100”。如果在传输中，第五位的0变成了1，那么接收端的电传打字机便打印出了数字“1”。像这种错误就称为“数字互变”，无论是人还是机器都很难发现。因而，这种错误造成的后果，对某些用户来说可能是很严重的。例如，在一个自动化防空体系中，搜索雷达把敌机的高度7000米这个数据，用“7”这个数字来代表，并用“7”的五单位码送往控制计算机。如果由于传输中的干扰，使“7”错成为“1”，而数字“1”又代表高度1000米，结果计算机就按照高度1000米的数据进行计算，并控制导弹或飞机起飞迎击敌机，从而可能造成使敌机漏网的严重后果。因此，克服这种数字传输中的错误，在数字通信系统中越来越重要。

此外，通信系统为了做到多、快、好、省，必须解决技术上的一些矛盾。例如，要求在一秒钟内传送的二进制码元越多，则每个码元所占的时间越短，波形越窄，能量越少，受到干扰后错误的可能性也越大，传送消息的可靠性就越低。反之，则使传送消息的速度减慢。因此，如何解决传输可靠性与速度这一对矛盾，是正确设计一个通信系统的关键之一。而通信理论本身，特别是数字通信理论，也在解决这一对矛盾中不断前进。

通常，我们以一秒钟内传送二进制码元的多少来表明传送信息的速度，用比特/秒（或b/秒）表示。如一秒钟内传送一万个二进制码元，则称传送信息的速率（简称传信率）为一万比特/秒，或用10kb/秒（K表示千）表示。而以误码率 p_e 表示传送信息的可靠性（准确度），由下式确定：

$$p_e = \frac{\text{错误接收的码元数}}{\text{接收的总码元数}} \quad (1-1)$$

例如，若接收到一万个码元，检查后错了一个，则误码率为万分之一，用 10^{-4} 表示。当然，也可用类似的式子定义误字率、误句率，来表示传输的质量。

目前由于数字通信中传送消息的速度越来越高，例如高达几百万比特/秒，甚至几千万或几兆、几十兆比特/秒。在这样高的速率下，只要信道中有很小一点干扰，就能影响很多码元的正确接收。目前，在一般调制制度（如残留边带调制、调频等）和中等传信率（600至4800比/秒）下，普通的电话电报线路中的误码率在 10^{-4} 至 10^{-6} 之间，而高频无线信道或散射信道只有 10^{-2} 至 10^{-4} 。这种传输质量，对某些用户来说可能已经够好，但对某些用户来说则远远满足不了要求。例如像上面提到的自动化防空体系，数据传输的误码率要求在 10^{-7} 或更低，至于计算机和计算机之间的数据传输误码率，则要求低于 10^{-9} 。

毛主席教导我们：“马克思主义的哲学认为十分重要的问题，不在于懂得了客观世界的规律性，因而能够解释世界，而在于拿了这种对于客观规律性的认识去能动地改造世界。”为了降低误码率，人们在长期的实践中，总结了很多方法。这些方法大致可归纳为二类：一是选择合理的调制制度和改造信道以提高接收机输出的信噪比，例如增加发射机的功率、增大天线面积或采用分集技术、低噪声器件等。若要进一步减低误码

率，或者在某些场合上述方法不适合或无能为力，这时可用另一类方法——抗干扰编码。

抗干扰编码就是把要传送的信息序列，通过一抗干扰码编码器，以一定的规则增加一些多余度码元，使这些多余度码元与传送的信息元之间建立一定关系（编码），然后把这些信息元及多余度码元一起送往发射机，经信道传输至接收机后，再送入抗干扰码译码器，检验它们之间的关系（译码），从而自动地发现错误，或者纠正错误。这种方法在数字通信中称为差错控制技术，所应用的码称抗干扰码或纠错码，也就是本小册子将要介绍的内容。

由上面叙述中，我们可把图(1-1)的通信系统模型，归结成图(1-3)所示的数字通信系统模型。为了讨论问题简单起见，图中把通信系统中各个部件的不理想性，以及存在于机器内和信道中的噪声和干扰，都等效成一个加于信道的噪声源。图中的信源编码器把信源发出的消息如语言、图像、文字等变换为二进制形式的信息序列，并且为了使传输有效，去掉了一些与传输信息无关的内容（有时为了保密，信源编码器后还可接上加密器）。信源译码器（其前可以有解密器）是把二进制形式的信息序列还原为消息。



图(1-3) 数字通信系统模型

上述的数字通信系统模型中，并没有显示出另一个很重要的组成部分——同步系统，但是它在数字通信中一般是不可缺少的。数字通信系统发送端送出的数字或脉冲序列，是按一定的