第5讲 纠错编码的基本概念

第1章 纠错编码的基本概念

第1章 纠错编码的基本概念

- ◈数字通信和纠错码
- ◈检错编码和纠错编码
- ◈检错能力和纠错能力

数字通信系统

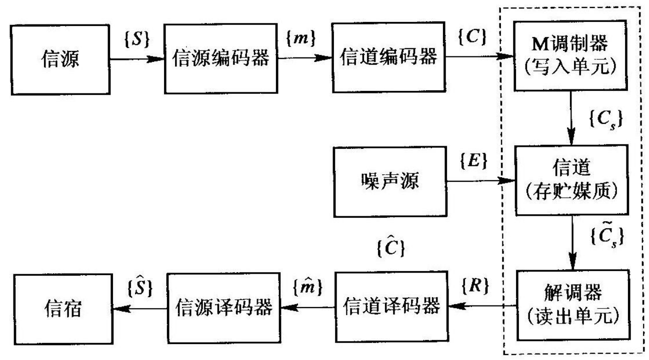

图1：数字通信系统框图

数字通信和纠错码

编码有信源编码和信道编码。

信源编码的目的是压缩冗余度，提高信息的传输速率。

信道编码的目的是提高信息传输时的抗干扰能力以增加信息传输的可靠性。

纠错编码即信道编码。包括具有抗干扰能力的码类的构造及编码译码方法。

信源编码

信源信息--->数字信息（信源编码）

例：{26个字母 空格 ， 。？！ -}可以用$V_5(F_2)$中的向量表示如下：

元素组	英文字母或标点符号
00000	空格
00001	a
00010	b
00011	c
00100	d
00101	e
00110	f
⋮	⋮
11100	.
11101	?
11110	!
11111	-

信道编码（抗干扰编码）

抗干扰编码：
σ $(c_0,c_1,c_2,c_3,c_4)=(c_0,c_1,c_2,c_3,c_4,\sum^4_{i=0}c_i)$

元素组	校验位	英文字母或标点符号
00000	0	空格
00001	1	a
00010	1	b
00011	0	c
00100	1	d
00101	0	e
00110	0	f
⋮	⋮	⋮
11100	1	.
11101	0	?
11110	0	!
11111	1	-

7

信道编码

抗干扰编码：
σ $(c_0,c_1,c_2,c_3,c_4)=(c_0,c_1,c_2,c_3,c_4,\sum^4_{i=0}c_i)$

原始数字信息	码字	英文字母或标点符号
00000	000000	空格
00001	000011	a
00010	000101	b
00011	000110	c
00100	001001	d
00101	001010	e
00110	001100	f
⋮	⋮	⋮
11100	111001	.
11101	111010	?
11110	111100	!
11111	111111	-

8

信道编码

消息集合：$V_5(F_2)$

编码：σ：$V_5(F_2)$-->$V_6(F_2)$

例如：$\sigma(c_0,c_1,c_2,c_3,c_4)=(c_0,c_1,c_2,c_3,c_4,\sum^4_{i=0}c_i)$

$\sigma(V_5(F_2))\subset V_6(F_2)$

码：$C=\sigma(V_5(F_2))$
码字：$C=\sigma(V_5(F_2))$中的向量

字：$V_6(F_2)$中的向量

*q*元码：码字中的每个分量在F_q中取值的码；
二元码：码字中的每个分量在F_2中取值的码；

9

检错编码

消息集合：$V_5(F_2)$

编码：σ：$V_5(F_2)$-->$V_6(F_2)$

例如：$\sigma(c_0,c_1,c_2,c_3,c_4)=(c_0,c_1,c_2,c_3,c_4,\sum^4_{i=0}c_i)$

$\sigma(V_5(F_2))\subset V_6(F_2)$

上述码字有个特点：

$\sum^5_{i=0}c_i=0$

若收到的字：$r=(r_0,r_1,r_2,r_3,r_4,r_5)$
如果$\sum^5_{i=0}r_i\neq 0$，则在传输中一定出现了差错，(1个，3个，5个)
如果$\sum^5_{i=0}r_i=0$，则在传输中没有出现差错，或者出现了2个或4个差错。

10

检错编码

消息集合：$V_5(F_2)$

编码：σ：$V_5(F_2)$-->$V_6(F_2)$

例如：$\sigma(c_0,c_1,c_2,c_3,c_4)=(c_0,c_1,c_2,c_3,c_4,\sum^4_{i=0}c_i)$

$\sigma(V_5(F_2))\subset V_6(F_2)$

若假设码字在信道中传输时最多出现1个差错，则

检错措施	结论
$\sum_{i=0}^{5}r_i\neq 0$	有错
$\sum_{i=0}^{5}r_i=0$	无错

则称码$C=\sigma(V_5(F_2))$为能检查出小于等于1个差错的检错码。σ称为检错编码。

相关概念：信息位，校验位，检错编码，检错措施。

11

检错编码

消息集合：$V_k(F_q)$

编码：σ：$V_k(F_q)$-->$V_n(F_q)$

$\sigma(a_0,a_1,\ldots,a_{k-1})=(c_0,c_1,\ldots,c_{n-1})$

$\sigma(V_k(F_q))\subset V_n(F_q)$

码：$C=\sigma(V_k(F_q))$ （$\subset V_n(F_q)$）
码字：$C=\sigma(V_k(F_q))$中的向量
码元：码字的分量
字：$V_n(F_q)$中的向量

如果传输过程中有小于等于t个位置的码元发生差错时，收方从收到的字就能判断出有没有错误发生，那么C就叫码长为n的可以检查出t个差错的检错码。而σ称为检错编码。

12

纠错编码

消息集合：$V_k(F_q)$

编码：σ：$V_k(F_q) \dashrightarrow V_n(F_q)$

$\sigma(a_0,a_1,\ldots,a_{k-1})=(c_0,c_1,\ldots,c_{n-1})$

$\sigma(V_k(F_q)) \subset V_n(F_q)$

码：$C=\sigma(V_k(F_q))$ （$\subset V_n(F_q)$）

码字：$C=\sigma(V_k(F_q))$中的向量

码元：码字的分量

字：$V_n(F_q)$中的向量

如果传输过程中有小于等于t个位置的码元发生差错时，收方从收到的字仍能正确地译出发方发送的码字就能判断出有没有错误发生，那么C就叫码长为n的可以纠正t个差错的纠错码。而σ称为纠错编码。

13

纠错编码

如果传输过程中有小于等于t个位置的码元发生差错时，收方从收到的字仍能正确地译出发方发送的码字就能判断出有没有错误发生，那么C就叫码长为n的可以纠正t个差错的纠错码。而σ称为纠错编码。

例：设原始数字消息集合：$V_2(F_2)=\{00, 01, 10, 11\}$；

进行纠错编码：

$\sigma(00)=(10010)$

$\sigma(01)=(01001)$

$\sigma(10)=(10101)$

$\sigma(11)=(01110)$

则 $C=\sigma(V_2(F_2))=\{10010, 01001, 10101, 01110\}$

信息率：2/5；

$(k/n)\log_2 q$ 为$C=\sigma(V_k(F_q))$的信息率。

14

纠错编码

发方发送一个码字，收方收到一个字$r=(r_0,r_1,r_2,r_3,r_4)\in V_5(F_2)$

问：将r译为哪个码字?

答：“像”-----极大似然译码

Hamming距离：$a, b \in V_n(F_q)$,

$$H(a,b) = \sum_{a_i \neq b_i} 1$$

性质：

（1）自反性：$H(a,a)=0$

（2）对称性：$H(a,b)=H(b,a)$

（3）三角不等式：$H(a,b)+H(b,c) \geq H(a,c)$

15

纠错编码

发方发送一个码字，收方收到一个字$r=(r_0,r_1,r_2,r_3,r_4)\in V_5(F_2)$

问：将r译为哪个码字?

答：“像”-----极大似然译码

纠错编码σ:

$\sigma(00)=(10010)$

$\sigma(01)=(01001)$

$\sigma(10)=(10101)$

$\sigma(11)=(01110)$

则 $C=\{10010, 01001, 10101, 01110\}$

$r=(10110)$ $c=(10010)$

$r=(01101)$ $c=(01001)$

$r=(11011)$ $c=?$

16

纠错编码

译码表：

码字	10010	01001	10101	01110
其余的字	00010 11010 10110 10000 10011	11001 00001 01101 01011 01000	00101 11101 10001 10111 10100	11110 00110 01010 01100 01111
	11011 00011	00000 11000	00111 11111	11100 00100

17

检错、纠错能力

定理1：设C为码长为n的一个码，

(1) 若任意两个码字的距离$\geq t+1$，则C是可检出t个差错的检错码；

证明：$\forall\ a, b \in C$，$H(a,b) \geq t+1$；

发a，收r，

如果 $H(a,r) \leq t$；则不存在$b \in C$，使得$r=b$；

$r=a$，说明没有传错，将r译为a；

$r \neq a$，说明传错了。

所以，C是可检出t个差错的检错码。

18

检错、纠错能力

定理1：设C为码长为n的一个码，
(1) 若任意两个码字的距离≥t+1，则C是可检出t个差错的检错码；
(2) 若确有两个码字的距离=t+1，则C不能检出t+1个差错；

证明：$\exists\, a, b \in$ C，H(a, b) = t+1；

发a, 收r (=b),
H(a, r) = t+1 ,
H(b, r) = 0 ,
则认为没有传错，错译为b，则C检查不出t+1个差错。

19

检错、纠错能力

定理1：设C为码长为n的一个码，
(3) 若任意两个码字的距离≥ 2t+1，则C是可纠正t个差错的纠错码；

证明：$\forall\, a, b \in$ C，H(a, b) ≥ 2t+1；

发a, 收r ,
如果 H(a, r) ≤ t ,
则$\forall\, b \in$ C，H(a, r) < H(b, r)；
（因为2t+1≤ H(a, b) ≤ H(a, r)+H(b, r)，所以H(b, r) ≥ t+1；）
根据极大似然译码法则，将r译为a；
所以，C是可纠正t个差错的纠错码；

20

检错、纠错能力

定理1：设C为码长为n的一个码，
(3) 若任意两个码字的距离≥ 2t+1，则C是可纠正t个差错的纠错码；
(4) 若确有两个码字的距离= 2t+1，则C不能纠正t+1个差错。

证明：$\exists a, b \in$ C，H(a, b) = 2t+1；

发a, 收r ,
如果 H(a, r) = t +1, 而 H(b, r) = t ,
则$\forall\, b \in$ C，H(a, r) < H(b, r)；
（例a=111000, b=111111，r=111110 ）
根据极大似然译码法则，将r译为b，发生错误；
所以，C不能纠正t+1个差错。

21

检错、纠错能力

定理1：设C为码长为n的一个码，
(1) 若任意两个码字的距离≥t+1，则C是可检出t个差错的检错码；
(2) 若确有两个码字的距离=t+1，则C不能检出t+1个差错；
(3) 若任意两个码字的距离≥ 2t+1，则C是可纠正t个差错的纠错码；
(4) 若确有两个码字的距离= 2t+1，则C不能纠正t+1个差错。

由定理1，码C的两两码字的距离的极小值，

$$\min C=\min\{H(a,b)|a,b\in C, a\neq b\}$$

是衡量C的检错和纠错能力的一个数，称为C的极小距离。

22

检错、纠错能力

定理1：设C为码长为n的一个码，
(1) 若任意两个码字的距离≥t+1，则C是可检出t个差错的检错码；
(2) 若确有两个码字的距离=t+1，则C不能检出t+1个差错；
(3) 若任意两个码字的距离≥ 2t+1，则C是可纠正t个差错的纠错码；
(4) 若确有两个码字的距离= 2t+1，则C不能纠正t+1个差错。

由定理1，码C的两两码字的距离的极小值，

$$\min C=\min\{H(a,b)|a,b\in C, a\neq b\}$$

是衡量C的检错和纠错能力的一个数，称为C的极小距离。

例：设C为码长为5的一个二元码，
设 C={10010，01001，10101，01110}，求该码的最小距离。
并给出该码的检错能力和纠错能力。
min(C)= 3；
检2 错；
纠1 错。

23

纠错编码

译码表：

码字	10010	01001	10101	01110
其余的字	00010	11001	00101	11110
	11010	00001	11101	00110
	10110	01101	10001	01010
	10000	01011	10111	01100
	10011	01000	10100	01111
	……	……	……	……
	11011	00000	00111	11100
	00011	11000	11111	00100

说明：从译码表译码工作量太大，故希望码具有某种代数结构，从而利用码的这种代数特性来译码。

24

检错、纠错能力

定理1： 设C为码长为n的一个码，
(1) 若任意两个码字的距离$\geq t+1$，则C是可检出t个差错的检错码；
(2) 若确有两个码字的距离$= t+1$，则C不能检出$t+1$个差错；
(3) 若任意两个码字的距离$\geq 2t+1$，则C是可纠正t个差错的纠错码；
(4) 若确有两个码字的距离$= 2t+1$，则C不能纠正$t+1$个差错。

由定理1，码C的两两码字的距离的极小值，

$$\min C=\min\{H(a,b)|a,b\in C,\ a\neq b\}$$

是衡量C的检错和纠错能力的一个数，称为C的**极小距离**。

课堂练习： 设C为码长为6的一个二元码，
设 C={101010，101101，011011，001110}，求该码的最小距离。并给出该码的检错能力和纠错能力。

min(C)= 2；
检1 错；
纠0 错。

25