第二章 数据的机器层次表示

目录

目录

2.6 数据校验码

- 差错控制
  - 数据传输中，对数据采用某种编码法，通过少量的附加电路，使之能发现某些错误，甚至能确定出错位置，进而实现自动改错的能力。
- 数据校验码
  - 是一种常用的带有发现某些错误或自动改错能力的数据编码方法。
- 实现原理
  - 加进一些冗余码，使合法数据编码出现某些错误时，就成为非法编码。这样，就可以通过检测编码的合法性来达到发现错误的目的。
  - 差错控制是以增加冗余信息为代价的。

2.6 数据校验码

- 码距是两个码字C1与C2之间不同的比特数。
  - 如：1100与1010的码距为2;
    1111与0000的码距为4。
- 一个编码系统的码距就是整个编码系统中任意(所有)两个码字的最小距离。
  - 例如：一个编码系统有四种编码分别为：0000，0011，1100， 1111，此编码系统中0000与1111的码距4；0000与0011的码距为2，是此编码系统的最小码距。因此该编码系统的码距为2。

2.6 数据校验码

- 8421BCD码的码距1，不具有检错能力，从一个编码改变一位（出错）会变成另一个合法的编码。
- 增加编码的位数可以增加码距：因为位数增加使可以表示的状态增加，将合法的码点分散到不同的状态。
- 如表示两个数用2位二进制，增加一个冗余位：
  - {11，10，01，00}；
  - 合法码：{11,00}；
  - 非法码：{01,10}；

2.6 数据校验码

- 假如我们现在要对A，B两个字母进行编码。我们可以选用不同长度的编码，以产生不同码距的编码，分析它们的检错纠错能力。
- 方案1：用1位长度的二进制编码。A=1，B=0。这样A，B之间的最小码距为1。
  - 合法码：{0,1};
  - 非法码：{ };
  - 此编码无检错纠错能力。因为这种编码无论由1错为0，或由0错为1，接收端都无法判断是否有错，因为1，0都是合法的编码。

2.6 数据校验码

- 方案2：用2位长度的二进制编码，若以A=11，B=00，A、B之间的最小码距为2。
  - 合法码：{11,00}；
  - 非法码：{01,10}；
- 特点：
  - 可以检测出1位错。
  - 无法纠错。
  - 不能发现两位错误。
- 此种编码的检错能力为1位，纠错能力为0位。

2.6 数据校验码

- 具有检、纠错能力的数据校验码的实现原理是：
  - 在编码时增加冗余的信息位数，（增加编码的位数可以增加码距）从而使在编码中，除去合法的码字外，增加一些非法的码字，当某个合法码字出现错误时，就变为非法码字。从而可以检测出错误。

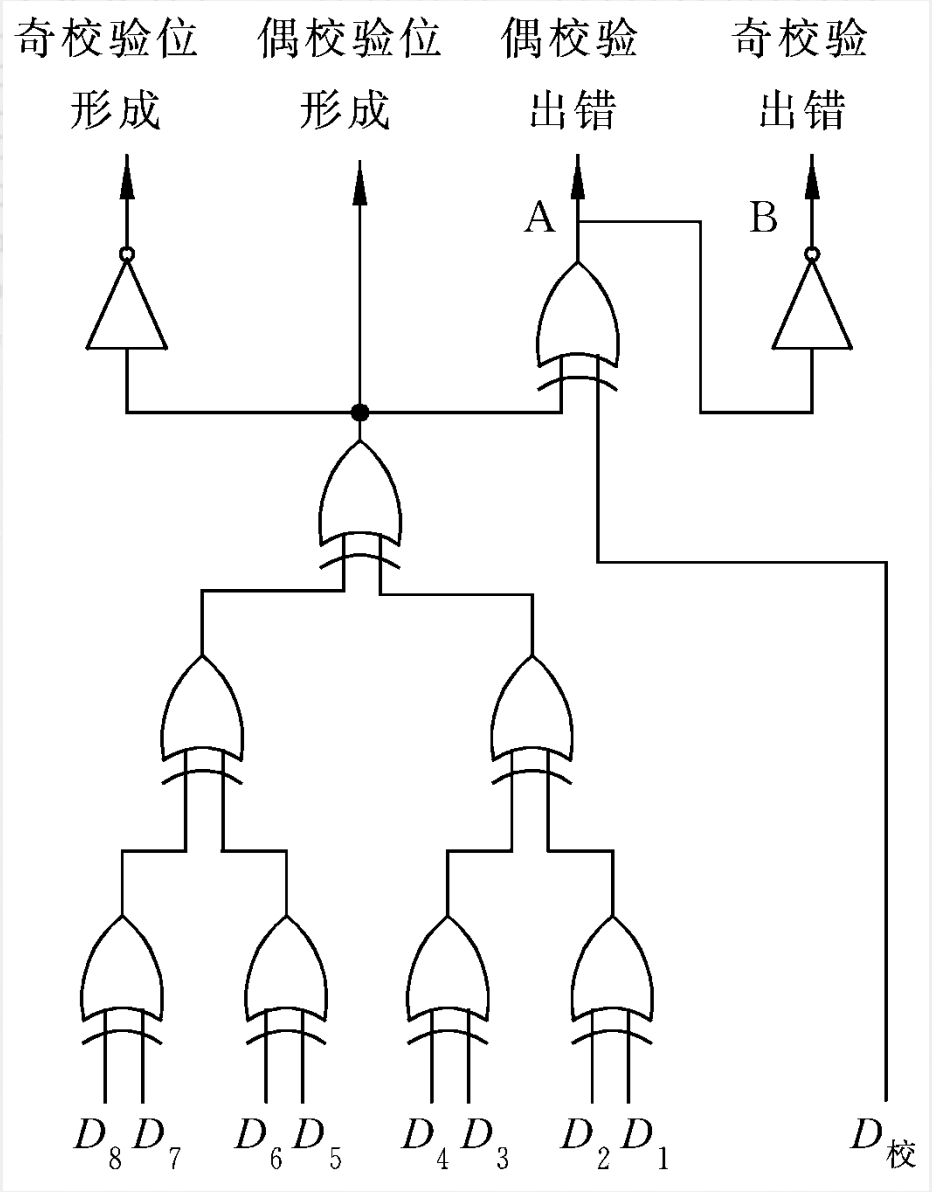
2.6.1 奇偶校验码

- 它可以检测出一位（或奇数位）错误。
- 奇偶校验实现方法
  - 在有效信息（如一个字节）中再加上一个二进制位（校验位）组成校验码，然后根据校验码的奇偶性质进行校验。
  - **奇校验——整个校验码（有效信息位和校验位）中“1”的个数为奇数。**
  - **偶校验——整个校验码中“1”的个数为偶数。**

2.6.1 奇偶校验码

- 校验位的产生
  - 偶校验：各数据位异或运算
  - 奇校验：各数据位异或运算后取反
- 校验过程
  - 偶校验：所有位异或，结果为0，无错误，否则出错
    - ¤ $P' = C_7 \oplus C_6 \oplus C_5 \oplus C_4 \oplus C_3 \oplus C_2 \oplus C_1 \oplus C_0 \oplus P$
  - 奇校验：所有位异或后取反，结果为0，无错误，否则出错

2.6.1 奇偶校验码



练习

- 1. 假定下列字符码中有奇偶校验位，但没有数据错误，采用奇校验的字符码是（D）
- A. 11001010　　B. 11010111
- C. 11001100　　D. 11001011

2.6.2 海明校验码

- 其实现原理是：在k位信息位中为，增加r位冗余位，构成一个n=k+r位的码字，附加的r位冗余位被编在传输码字的特定位置上，将数据代码的码距均匀地拉大，每一个这种冗余位对码字中的若干位进行奇偶校验，并且数据的每一二进制位分配在几个奇偶校验组中，这样当某一位出错后，会引起几个校验位的值的变化。这样，不但能够检测出错误，而且能够为进一步纠错提供依据。

2.6.2 海明校验码

海明码编码规律（假设编成海明码为$H_mH_{m-1}...H_2H_1$）：

- (a)冗余位数：假设校验位的位数为r位，设有效信息码组为k位，

应有关系：

$$2^r-1\geq k+r$$

例如：一个字节的数据需要4个校验位。

2.6.2 海明校验码

海明码编码规律（假设编成海明码为$H_mH_{m-1}...H_2H_1$）：

- (b) 校验位分布位置：

在m位的海明码中，各校验位分布在位号为2^{i-1}的位置，即校验位的位置分别为1,2,4,8,...，其余为数据位。数据位按原来的顺序关系排列。如有效信息码为$...D_5D_4D_3D_2D_1$，则编成的海明码为$...D_5P_4D_4D_3D_2P_3D_1P_2P_1$，其中$P_i$为第 i个校验位。

2.6.2 海明校验码

海明码编码规律（假设编成海明码为$H_mH_{m-1}...H_2H_1$）：

- (c) 校验关系：
  - 每个校验位对海明码中的部分数据位（包括校验位本身）进行奇偶校验。

    校验位所在位置决定了它要校验和跳过的比特位顺序。具体关系如下：
    - 位置1：校验1位，跳过1位，校验1位，跳过1位 (需要校验1,3,5,7,9,11,13,15,...位)
    - 位置2：校验2位，跳过2位，校验2位，跳过2位 (2,3,6,7,10,11,14,15,...)
    - 位置4：校验4位，跳过4位，校验4位，跳过4位 (4,5,6,7,12,13,14,15,20,21,22,23,...)
    - 位置8：校验8位，跳过8位，校验8位，跳过8位 (8-15,24-31,40-47,...)
    - ...

2.6.2 海明校验码

海明码编码规律（假设编成海明码为$H_mH_{m-1}...H_2H_1$）：

- (d) 校验位的取值：（与奇偶校验的产生方法相同）
  - 如进行偶校验则如果全部校验的位置中有奇数个1，把该奇偶校验位置为1；如果全部校验的位置中有偶数个1，把该奇偶校验位置为0。

海明校验码

- 海明码编码举例：
- 设信息位为D7,D6,D5,D4,D3,D2,D1,D0
  - 编成的海明码为H12,H11,H10,H9,H8,H7,H6,H5,H4,H3,H2,H1

    D7, D6, D5, D4,P4,D3,D2,D1,P3, D0,P2,P1
    - P1=H1⊕H3⊕H5⊕H7⊕H9⊕H11=P1⊕D0⊕D1⊕D3⊕D4⊕D6(P1初值取0)
    - P2=H2⊕H3⊕H6⊕H7⊕H10⊕H11=P2⊕D0⊕D2⊕D3⊕D5⊕D6(P2初值取0)
    - P3=H4⊕H5⊕H6⊕H7⊕H12=P3⊕D1⊕D2⊕D3⊕D7(P3初值取0)
    - P4=H8⊕H9⊕H10⊕H11⊕H12=P4⊕D4⊕D5⊕D6⊕D7(P4初值取0)

262 海明校验码

- 海明码的接收端的公式
  - P1=H1⊕H3⊕H5⊕H7⊕H9⊕H11=P1⊕D0⊕D1⊕D3⊕D4⊕D6
  - P2=H2⊕H3⊕H6⊕H7⊕H10⊕H11=P2⊕D0⊕D2⊕D3⊕D5⊕D6
  - P3=H4⊕H5⊕H6⊕H7⊕H12=P3⊕D1⊕D2⊕D3⊕D7
  - P4=H8⊕H9⊕H10⊕H11⊕H12=P4⊕D4⊕D5⊕D6⊕D7
    - P4,P3,P2,P1等于0,0,0,0表示传送的数据流正确；
    - 如P4,P3,P2,P1等于0,0,1,0则表示传送的数据流中第2位出错；
    - 如P4,P3,P2,P1等于0,0,1,1则表示传送的数据流中第3位出错；
    - 依次类推。

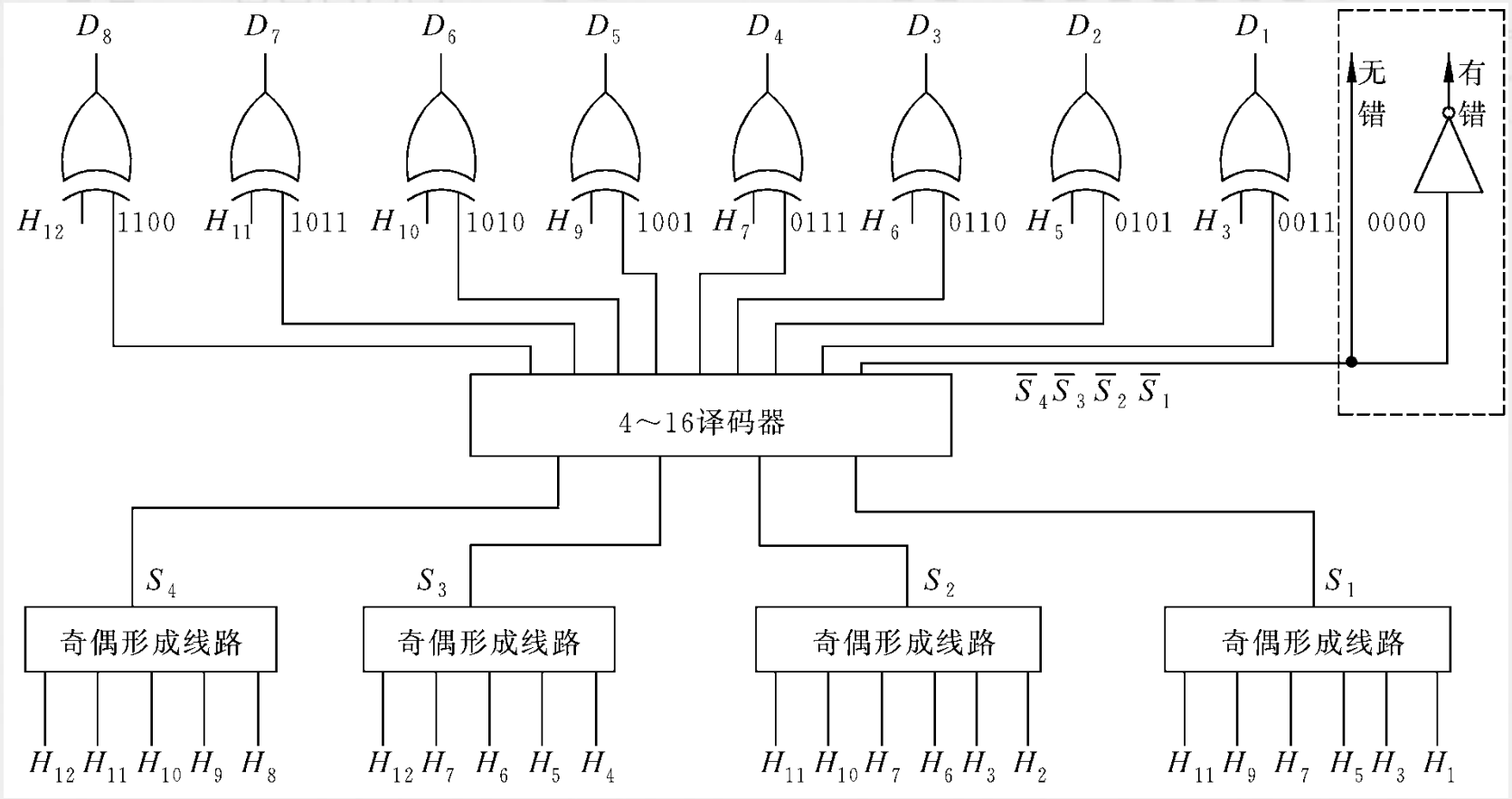
表 3.8　数据位 k 与校验位 r 的对应关系

k 值	最小的 r 值
1～4	4
5～11	5
12～26	6
27～57	7
58～120	8

表 3.9　出错的海明码位号和校验位位号的关系

海明码位号	数据位/校验位	参与校验的校验位位号	被校验位的海明码位号 = 校验位位号之和
H_1	P_1	1	1=1
H_2	P_2	2	2=2
H_3	D_1	1,2	3=1+2
H_4	P_3	4	4=4
H_5	D_2	1,4	5=1+4
H_6	D_3	2,4	6=2+4
H_7	D_4	1,2,4	7=1+2+4
H_8	P_4	8	8=8
H_9	D_5	1,8	9=1+8

海明校验码



(12，8)分组码海明校验线路

习题

- 1.(2013)用海明码对长度为8位的数据进行检/纠错时，若能纠正一位错，则校验位数至少是（C）

  –A. 2　　B. 3　　C. 4　　D. 5

- 2. 求数据10110100110的海明码，可检/纠错一位错。

习题

- 信息位11位，校验位4位，海明码位数11+4 = 15位

校验位分布在 2 的整数次方位置上。

H15 H14 H13 H12 H11 H10 H9 H8 H7 H6 H5 H4 H3 H2 H1

1 0 1 1 0 1 0 P4 0 1 1 P3 0 P2 P1

P1 = P1⊕H3⊕H5⊕H7⊕H9⊕H11⊕H13⊕H15 = 0⊕0⊕1⊕0⊕0⊕0⊕1⊕1=1

P2 = P2⊕H3⊕H6⊕H7⊕H10⊕H11⊕H14⊕H15=0⊕0⊕1⊕0⊕1⊕0⊕0⊕1=1

P3 = P3⊕H5⊕H6⊕H7⊕H12⊕H13⊕H14⊕H15=0⊕1⊕1⊕0⊕1⊕1⊕0⊕1=1

P4 = P4⊕H9⊕H10⊕H11⊕H12⊕H13⊕H14⊕H15=0⊕0⊕1⊕0⊕1⊕1⊕0⊕1=0

海明码：101101000111011

2.6.3 循环冗余校验码

- （1）循环冗余校验的基本原理：
  - 发送方：在N位信息码后再拼接K位的校验码。接受方将收到的信息码用同一个生成多项式去除，合法的循环冗余校验码应当能被生成多项式整除。如果环冗余校验码不能被生成多项式整除，就说明出现了信息差错。并且，有信息差错时，循环冗余校验码被生成多项式整除所得到的余数与出错位有对应关系，因而能确定出错位置。

2.6.3 循环冗余校验码

- （2）生成CRC校验码的过程

步骤1：将待编码的n位信息码组$C_{n-1}C_{n-2}\cdots C_i\cdots C_2C_1C_0$表达为一个n-1阶的信息多项式M(x)：

$$M(x)=C_{n-1}x^{n-1}+C_{n-2}x^{n-2}+\ldots+C_ix^i+\ldots++C_1x^1+C_0x^0$$

步骤2：将信息码组左移k位，成$M(x)\cdot x^k$，即成n+k位的信息码组：

$$C_{n-1+\ k}C_{n-2+\ k}\cdots C_{i+\ k}\cdots C_{2+\ k}C_{1+\ k}C_k00\ldots00$$

步骤3：用一个预先选定的k+1位的生成多项式G(x)对$M(x)\cdot x^k$作模2除，得到一个商Q(x)和一个余数R(x)；

步骤4：将左移k位的待编码有效信息与余数R(x)（余数就是校验位）作模2加，即形成循环冗余校验码。

2.6.3 循环冗余校验码

- 例：对四位有效信息1100求循环冗余校验码，选择生成多项式G(x)为1011(因此k=3)。

  步骤1：$M(x)=x^3+x^2=1100$

  步骤2：$M(x)\cdot x^3=x^6+x^5=1100000$

  (k=3，即左移3位，加了3个0)

  步骤3：模2除

  $M(x)\cdot x^k/G(x)=1100000/1011=1110+010/1011$，

  即余数为 $R(x)=010$

  步骤4：模2加，得到循环冗余校验码

$M(x)\cdot x^3=Q(x)\cdot G(x)+R(x)=110000+010=1100010$

2.6.3 循环冗余校验码

- 模2除：按模2减求部分余数的除法。其余和算术除法类似，具体规则如下：
  - (a)上商原则：当部分余数的首位为1时商为1，当部分余数的首位为0时商为0。
  - (b)每求一位商，应使部分余数减少一位。（左移，移出最高位）
  - (c)部分余数的位数小于除数的位数时，该余数即为最后的余数。
- 模2减：不考虑进位的减法，即异或操作

【例】1111000除以1101：

1 0 1 1————商

————————

1 1 1 1 0 0 0-----被除数

1 1 0 1———————— 除数

————————

0 0 1 0 0 0 0------部分余数，减少一位。商为0

  0 0 0 0

————————

  0 1 0 0 0 0

    1 1 0 1

————————

    0 1 0 1 0————————余数

⊕   1 1 0 1

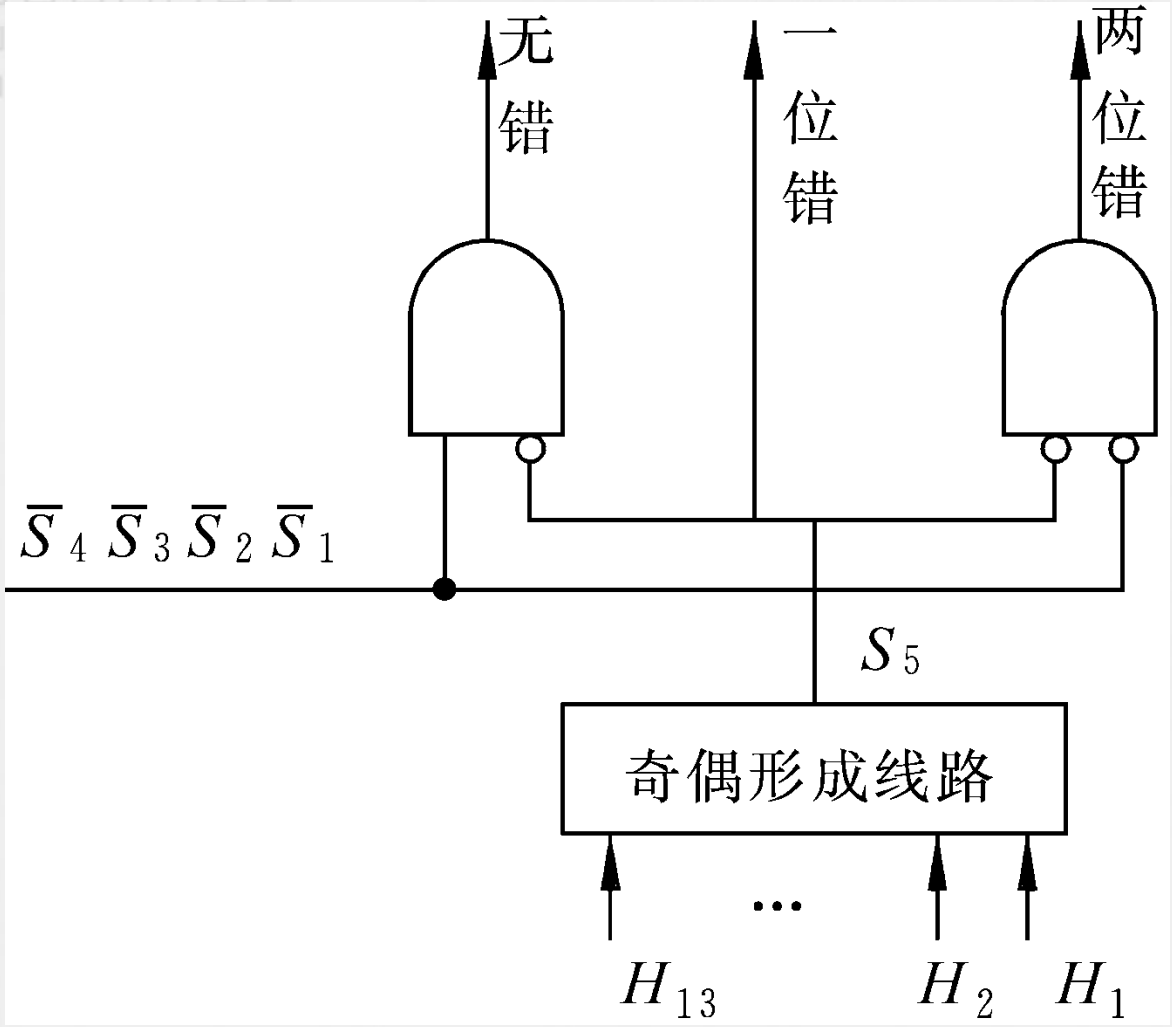
⊕ ————————

⊕   0 1 1 1 ————————余数

2.6.3 循环冗余校验码

- （5）关于生成多项式
  - 并不是任何一个多项式都可以作为生成多项式。从检错和纠错的要求出发，生成多项式应能满足下列要求：
    - 任何一位发生错误都应使余数不为0。
    - 不同位发生错误应使余数不同。
    - 对余数继续作模2运算，应使余数循环。
  - 生成多项式的选择主要靠经验。有三种多项式已经成为标准，具有极高的检错率。它们是：
    - CRC-12=$x^{12}+x^{11}+x^3+x^2+x+1$
    - CRC-16=$x^{16}+x^{15}+x^2+1$
    - CRC-CCITT=$x^{16}+x^{12}+x^5+1$

2.6.3 循环冗余校验码


无
错
一
位
错
两
位
错
$\overline{S}_4\ \overline{S}_3\ \overline{S}_2\ \overline{S}_1$
S_5
奇偶形成线路
…
H_{13}
H_2
H_1

2.6.3 循环冗余校验码

- 循环冗余校验(CRC)码
  - 1. CRC码的编码方法
  - 2. CRC的译码与纠错

表 3.10　(7,4)循环码的出错模式(生成多项式 $G(x)=1011$)

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	余　数	出错位
正　确	1	1	0	0	0	1	0	0 0 0	无
错　误	1	1	0	0	0	1	1	0 0 1	7
	1	1	0	0	0	0	0	0 1 0	6
	1	1	0	0	1	1	0	1 0 0	5
	1	1	0	1	0	1	0	0 1 1	4
	1	1	1	0	0	1	0	1 1 0	3
	1	0	0	0	0	1	0	1 1 1	2
	0	1	0	0	0	1	0	1 0 1	1

2.6.3 循环冗余校验码

- 循环冗余校验(CRC)码
  - 3. 关于生成多项式

表 3.11　生成多项式

n	k	码距 d	$G(x)$多项式	$G(x)$二进制码
7	4	3	$G_1(x)=(x^3+x+1)$	1011
			或(x^3+x^2+1)	1101
	3	4	$G_2(x)=G_1(x)(x+1)$	
			$=(x^3+x+1)(x+1)$	11101
			或$(x^3+x^2+1)(x+1)$	10111
15	11	3	(x^4+x+1)	10011
	7	5	$(x^4+x+1)(x^4+x^3+x^2+x+1)$	111010001
31	26	3	(x^5+x^2+1)	100101
	21	5	$(x^5+x^2+1)(x^5+x^4+x^3+x^2+1)$	11101101001
63	57	3	(x^6+x+1)	1000011
	51	5	$(x^6+x+1)(x^6+x^4+x+1)$	1010000110101

THANK YOU