

# 计算机组成原理与系统结构

## 第三章 信息编码与数据表示

<http://jpkc.hdu.edu.cn/computer/zcyl/dzkjdx/>





## 第三章 信息编码与数据表示

3.

数值数据的表示

3.

数据格式

3.3

定点机器数的表示

方法

3.4

浮点机器数的表示

方法

3.

非数值数据的表示

3.

校验码

3.7

现代计算机系统的数据表

示

本章小结

BACK



## 3.6 校验码

一

校验码概述

二

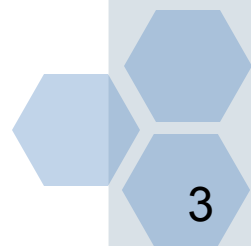
奇偶校验码

三

海明校验码

四

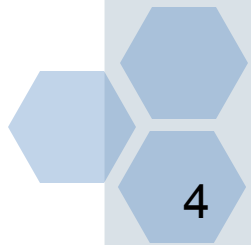
CRC校验码





# 一、校验码概述

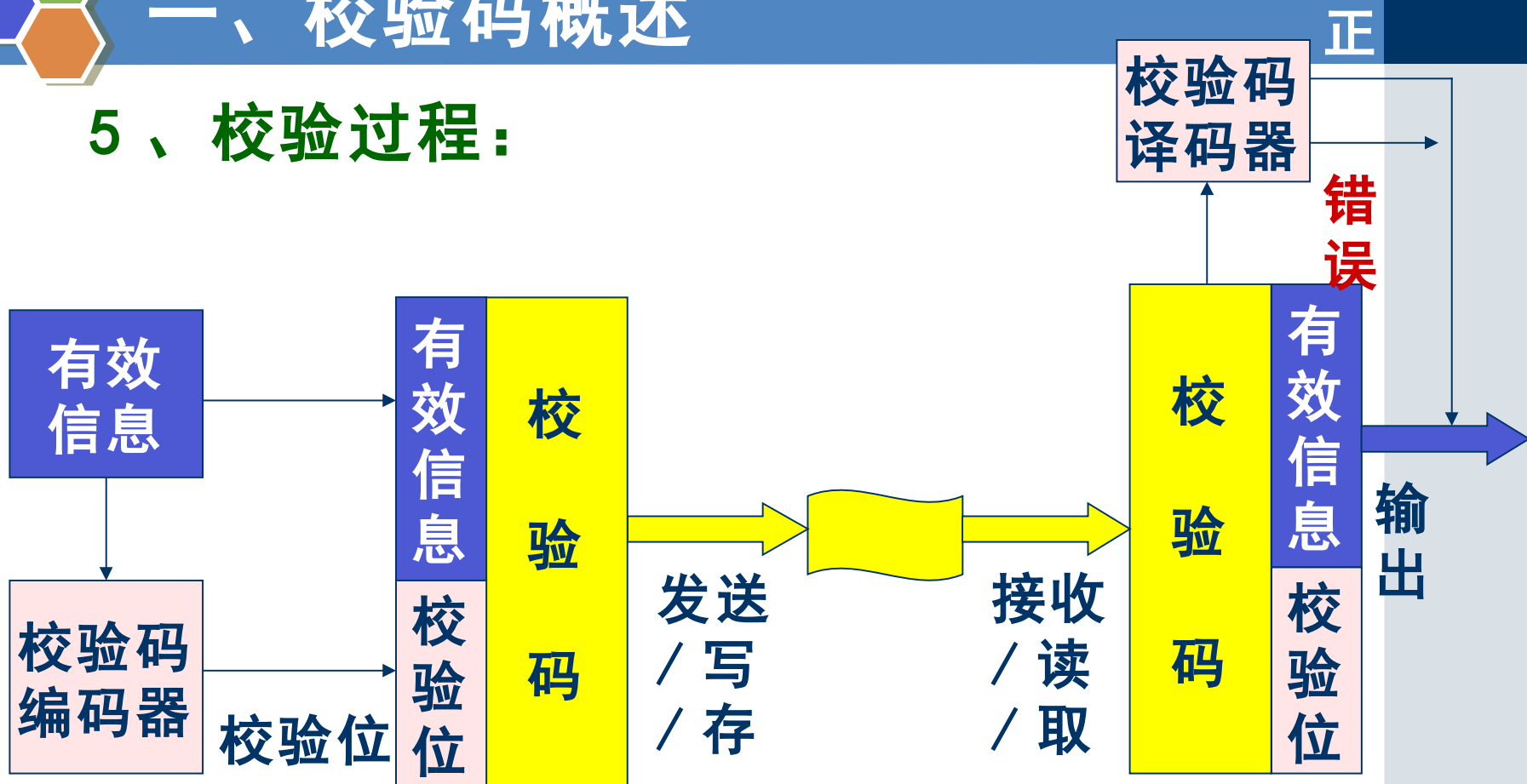
- 1、**校验码定义**：是一种具有发现某些错误或自动改正错误能力的一种数据编码方法。
- 2、**校验码目的**：用于**检查或纠正**在存取、读写和传送数据的过程中可能出现的**错误**。
- 3、**校验码的基本思想**：“**冗余校验**”，即通过在有效信息代码的基础上，添加一些冗余位来构成整个校验码。
- 4、**校验码的构成**：**有效信息+校验位**（由有效信息产生的冗余位）





# 一、校验码概述

## 5、校验过程：

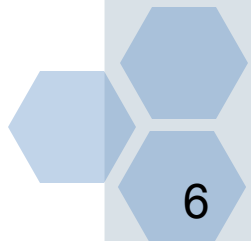




# 一、校验码概述

## 6、校验码原理：通过判断代码的合法性来检错的。

- 只有当合法码之间的码距  $d \geq 2$  时，校验码才具有检错能力，当码距  $d \geq 3$  时，校验码才具有纠错能力。
- 码距：一种码制的码距是指该码制中所有代码之间的最小距离。
- 两个代码之间的距离：在一种编码中，在任何两个代码之间逐位比较，对应位值不同的个数。





# 一、校验码概述

6、**校验码原理**：通过判断代码的**合法性**来检错的。

■ 校验码的**检错纠错能力与码距的关系**如下：

① 若码距  $d$  为奇数，**如果**只用来检查错误，则可以发现  $d - 1$  位错误；**如果**用来纠正错误，则能够纠正  $(d - 1) / 2$  位错误。

② 若码距  $d$  为偶数，则可以发现  $d/2$  位错误，**并**能够纠正  $(d/2 - 1)$  位错误。

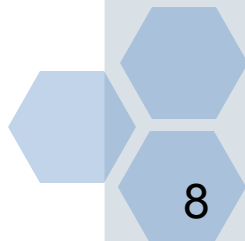




# 一、校验码概述

## 7、常见校验码：

- **奇偶校验码**：码距  $d=2$ ，检错码，**能检验奇数位错误**；通常用于磁带或者串行通信中。
- **海明校验码**：码距  $d \geq 3$ ，纠错码，**能纠正 1 位或多位错误**；通常用于磁盘冗余阵列中。
- **CRC 校验码**：码距  $d=3$ ，纠错码，**能纠正 1 位错误**；通常用于磁盘或数据块的校验。





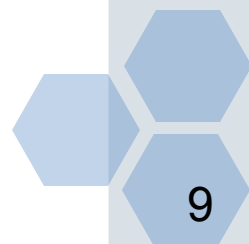


## 二、奇偶校验码

- ❖ 在有效信息位的前面或者后面添加一位奇（偶）校验位就组成了奇（偶）校验码。
- ❖ 奇（偶）校验码的编码和译码在硬件上通常采用异或非门（异或门）实现。

### 1、编码

- 奇校验位的取值应该使整个奇校验码中“1”的个数为奇数，偶校验位的取值应该使整个偶校验码中“1”的个数为偶数。





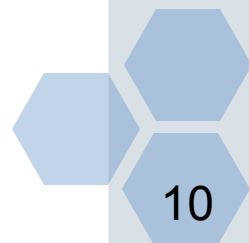
## 二、奇偶校验码

### 1、编码

- 假设在发送端，要发送七位 ASCII 码（ $B_6 B_5 B_4 B_3 B_2 B_1 B_0$ ），在 ASCII 码前面添加一位奇校验位  $P_{\text{奇}}$  或偶校验位  $P_{\text{偶}}$  变为一个字节的奇偶校验码，则它们的生成表达式为

$$P_{\text{奇}} = \overline{B_6 \oplus B_5 \oplus B_4 \oplus B_3 \oplus B_2 \oplus B_1 \oplus B_0}$$

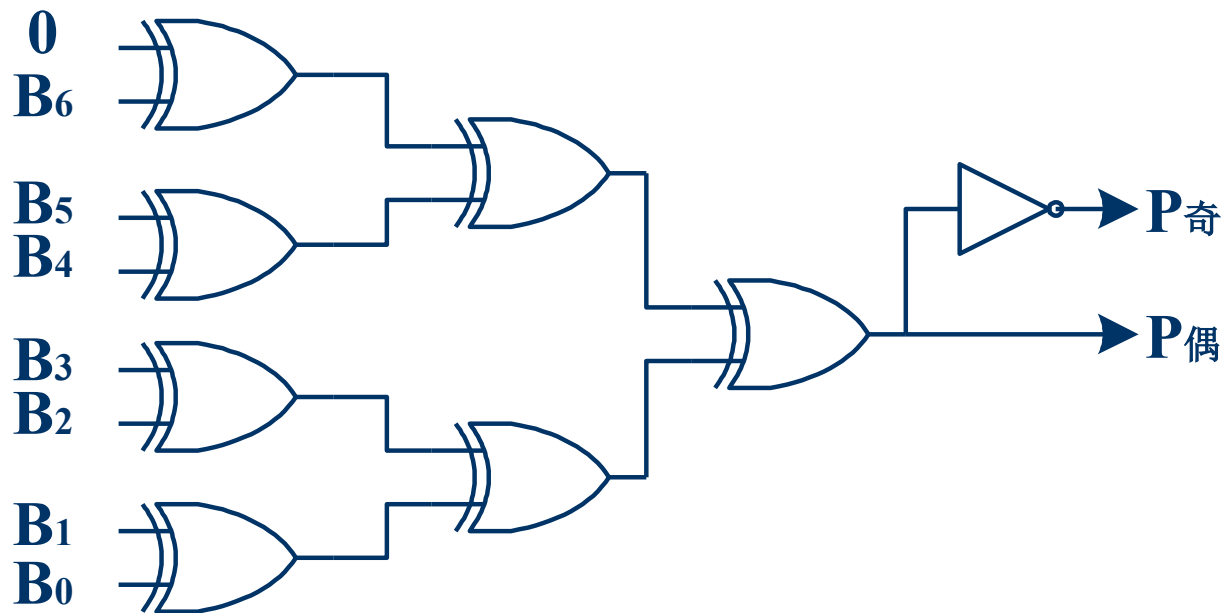
$$P_{\text{偶}} = B_6 \oplus B_5 \oplus B_4 \oplus B_3 \oplus B_2 \oplus B_1 \oplus B_0$$





## 二、奇偶校验码

### 奇偶校验的编码电路

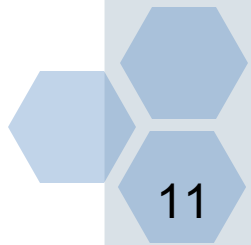


❖ 例如：字符“A”的ASCII码为  $41H = 1$

$00\ 0001B$

❖ 奇校验码为  $C1H = 1100\ 0001$

❖ 偶校验码为  $41H = 0100\ 0001$





## 二、奇偶校验码

### 2、译码

- 接收端**检验**接收到的**校验码信息的奇偶性**，奇校验码中“1”的个数应该为奇数；偶校验码中“1”的个数应该为偶数；否则出错。
- 设  $E_{\text{奇}}$  为奇校验码出错信号， $E_{\text{偶}}$  为偶校验码出错信号，为 1 出错，为 0 正确，则它们的表达式为

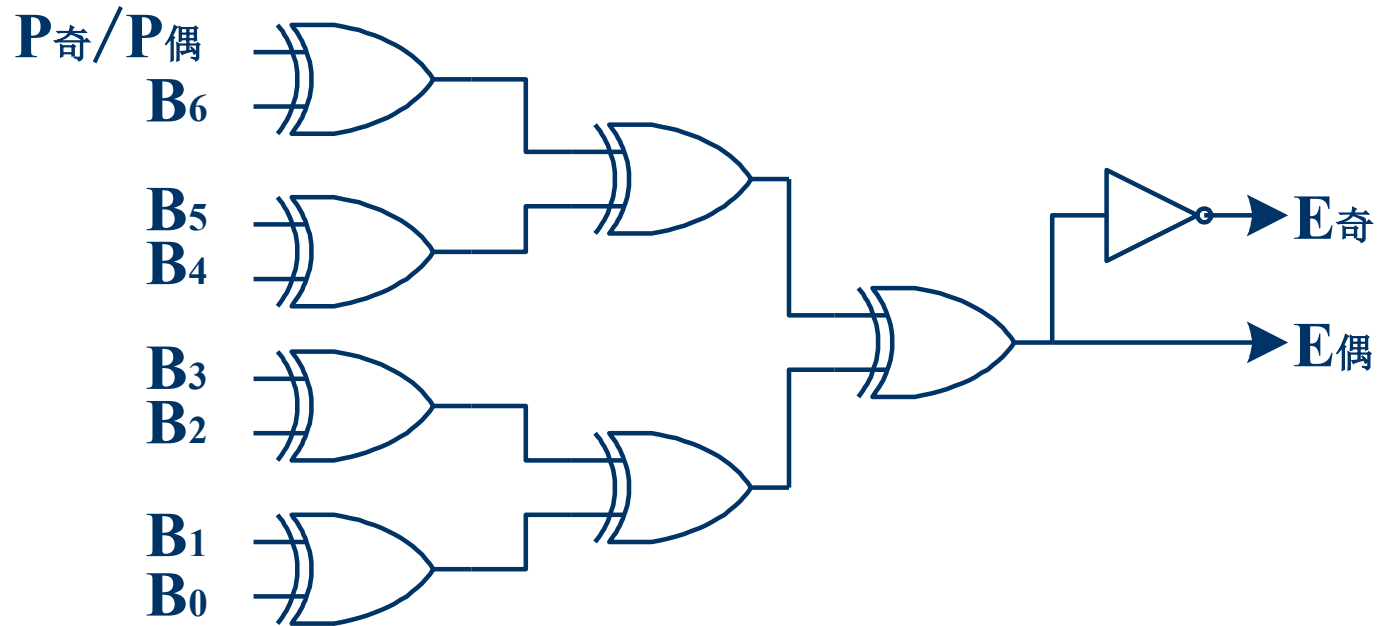
$$E_{\text{奇}} = B_6 \oplus B_5 \oplus B_4 \oplus B_3 \oplus B_2 \oplus B_1 \oplus B_0 \oplus P_{\text{奇}}$$

$$E_{\text{偶}} = B_6 \oplus B_5 \oplus B_4 \oplus B_3 \oplus B_2 \oplus B_1 \oplus B_0 \oplus P_{\text{偶}}$$



## 二、奇偶校验码

### 奇偶校验的译码电路





## 三、海明码

### ❖ 1、编码：步骤如下：

- (1) 计算校验位的位数
- (2) 确定有效信息和校验位的位置
- (3) 分组
- (4) 进行奇偶校验，合成海明码

### ❖ 2、译码



## (1) 计算校验位的位数

- ❖ 假设信息位为  $k$  位，增加  $r$  位校验位，构成  $n = k + r$  位海明码字。若要求海明码能纠正一位错误，用  $r$  位校验位产生的  $r$  位指误字来区分无错状态及码字中  $n$  个不同位置的一位错误状态，则要求  $r$  满足：
- ❖  $2^r \geq k + r + 1$
- ❖ 计算出  $k$  位有效信息时，必须添加的能纠错一位的海明校验码的校验位的位数

k 值	r 最小值
1 ~ 4	3
5 ~ 11	4
12 ~ 26	5
27 ~ 57	6
58 ~ 120	7





## (2) 确定有效信息和校验位的

- ❖ 假设  $k$  位有效信息从高到低为  $D_k D_{k-1} \cdots D_2 D_1$
- ❖ 添加的  $r$  位校验位为  $P_r P_{r-1} \cdots P_2 P_1$
- ❖ 它们构成  $n=k+r$  位的海明码为  $H_n H_{n-1} \cdots H_2 H_1$
- ❖  $H$  的下标被称为海明位号，则第  $i$  位的校验位  $P_i$  必须位于位号为  $2^{i-1}$  的位置，即：
- ❖  $P_i = H_j$ ， $j = 2^{i-1}$ ，其中， $i = 1, 2, \cdots, r$ ；
- ❖ 有效信息则在其余的海明码位置上顺序排列。

$H_{12}$	$H_{11}$	$H_{10}$	$H_9$	$H_8$	$H_7$	$H_6$	$H_5$	$H_4$	$H_3$	$H_2$	$H_1$
$D_8$	$D_7$	$D_6$	$D_5$	$P_4$	$D_4$	$D_3$	$D_2$	$P_3$	$D_1$	$P_2$	$P_1$







### ( 3 ) 分组

- ❖ 海明码是分组进行奇偶校验的，每一组通过一个监督表达式来监督有效信息的变化
- ❖ 分组必须使得监督表达式得出的指误字能够反映出错位的位号。
- ❖ **分组的原则：**
  - 校验位只参加一组奇偶校验，有效信息则参加至少两组的奇偶校验。
  - 若  $D_i = H_j$  ，则  $D_i$  参加那些位号之和等于  $j$  的校验位的分组校验。



# $k=8$ , $r=4$ 的海明码分组

序号 分组	$H_{12}$	$H_{11}$	$H_{10}$	$H_9$	$H_8$	$H_7$	$H_6$	$H_5$	$H_4$	$H_3$	$H_2$	$H_1$
	$D_8$	$D_7$	$D_6$	$D_5$	$P_4$	$D_4$	$D_3$	$D_2$	$P_3$	$D_1$	$P_2$	$P_1$
$P_4$	✓	✓	✓	✓	✓							
$P_3$	✓					✓	✓	✓	✓			
$P_2$		✓	✓			✓	✓			✓	✓	
$P_1$		✓		✓		✓		✓		✓		✓





## (4) 进行奇偶校验，合成海明

- ① 按照分组和奇偶校验的规律将每个校验位的生成表达式写出
- ② 带入有效信息的值，依次得出校验位的取值
- ③ 将校验位按各自的位置插入，与有效信息一起合成海明码。

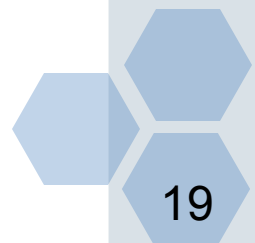
**例如：**有效信息为 11101001，则可以纠错一位的海明码为 1 1 1 0 1 1 0 0 0 1 0 1

$$P_4 = D_8 \oplus D_7 \oplus D_6 \oplus D_5$$

$$P_3 = D_8 \oplus D_4 \oplus D_3 \oplus D_2$$

$$P_2 = D_7 \oplus D_6 \oplus D_4 \oplus D_3 \oplus D_1$$

$$P_1 = D_7 \oplus D_5 \oplus D_4 \oplus D_2 \oplus D_1$$





## 2、译码

- ❖ 在接收端收到每个海明码后，也必须按上述分组检验每组的奇偶性有无发生变化；方法：
  - 按照监督关系式算出指误字  $s_r \ s_{r-1} \ \cdots \ s_2 \ s_1$
  - 若为全零，则说明各组奇偶性全部无变化，信息正确，将相应的有效信息位析取出来使用；
  - 若不为零，则指误字的十进制值，就是出错位的海明位号。



## 2、译码

❖ **例如：**收到的海明码为 110011000101，得到  $S_4 S_3 S_2 S_1 = (1010)_2 = (10)_{10}$ ，则表明是  $H_{10}$  ( $D_6$ ) 出错，将  $H_{10}$  取反，得正确海明码为 11**1**011000101。

$$S_4 = P_4 \oplus D_8 \oplus D_7 \oplus D_6 \oplus D_5$$

$$S_3 = P_3 \oplus D_8 \oplus D_4 \oplus D_3 \oplus D_2$$

$$S_2 = P_2 \oplus D_7 \oplus D_6 \oplus D_4 \oplus D_3 \oplus D_1$$

$$S_1 = P_1 \oplus D_7 \oplus D_5 \oplus D_4 \oplus D_2 \oplus D_1$$





## 四、循环冗余码 CRC

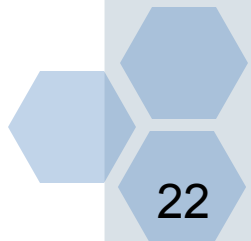
- ❖ 循环冗余码 CRC (Cyclic Redundancy Code)，又称为多项式码。

### 1、编码

- ❖  $k$  位要发送的有效信息位可对应于一个  $k-1$  次多项式  $M(x)$ ， $r$  位冗余校验位对应于一个  $r-1$  次多项式  $R(x)$ 。由  $k$  位信息位后面加上  $r$  位冗余位组成的  $n=k+r$  位 CRC 码字则对应于一个  $n-1$  次多项式  $C(x)$ ，即：

$$C(x) = x^r \cdot M(x) + R(x)$$

该 CRC 码称为  $(n, k)$  循环码。



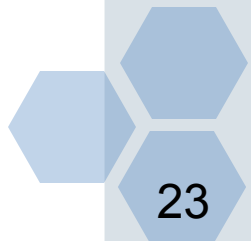


## 四、循环冗余码 CRC

- ❖ 由信息位产生冗余位的编码过程，就是已知  $M(x)$  求  $R(x)$  的过程。在 CRC 码中可以通过找到一个特定的多项式  $G(x)$  来实现。用  $G(x)$  去除  $x^r \cdot M(x)$  得到的余式就是  $R(x)$ ，假设商的多项式为  $Q(x)$ ，编码过程

$$\frac{x^r \cdot M(x)}{G(x)} = Q(x) + \frac{R(x)}{G(x)}$$

$$x^r \cdot M(x) = G(x) \cdot Q(x) + R(x)$$





## 四、循环冗余码 CRC

- ❖ 生成多项式  $G(x)$  应满足以下条件：
  - ① 必须是  $r$  次多项式，最高项  $x^r$  和  $x^0$  的系数为 1，即它对应的二进制编码是  $r+1$  位的。
  - ② CRC 校验码的任何一位发生错误，余数不为零；且不同位发生错误，余数不同。
  - ③ 对余数继续模 2 除，应使余数循环。
- ❖ 目前已经有多种生成多项式被列入国际标准中，如：CRC-4、CRC-12、CRC-16、CCITT-16、CRC-32 等。

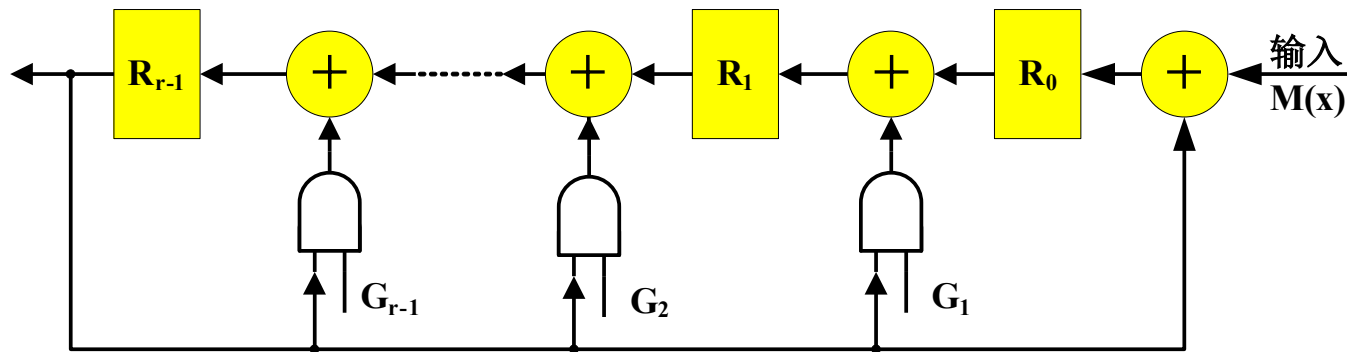




## 四、循环冗余码 CRC

- ❖ 目前常用的 CRC-16 多项式为  $x^{16}+x^{12}+x^5+1$  (记为 1021)，CCITT-16 多项式为  $x^{16}+x^{15}+x^2+1$  (记为 8005)。

CRC ( n, k ) 校验码串行生成电路原理图



图中  $G_i$  代表生成多项式  $G(x)$  各项的系数



## 四、循环冗余码 CRC

### 2、译码

- ❖ 接收端的校验过程就是用  $G(x)$  来除接收到的码字多项式的过程。
  - 若余式为零则认为传输无差错；
  - 若余式不为零则传输有差错。出错的位置与余数值是一一对应的关系，通过查找出错模式表，即可以确定是那一位出错。





**The End !**