

微型计算机硬件技术基础

Hardware Fundamentals of Modern Computer

Topic: Error Detection & Correction

陈启军，张 伟

Email: zhang_wi@mail.tongji.edu.cn

Dept. Of Automation, TongJi University



内容

● 从系统谈起——分而制之的还原论思想

● 基本概念：

- 接口与实现、接口与协议，总线与交换

● 原理：基于总线的互连技术

- 基本总线互连技术
- 基本总线技术的改进：多总线结构，局部总线，...
- 如何基于总线技术构建高性能系统
- 常见总线标准

● 原理：基于交换的互连技术

- 比较：总线技术和交换技术

内 容

- 基本问题描述
- 概念：检错和纠错，码距，数据校验码，编码和解码，漏检问题
- 奇偶校验码
- 海明码
- 循环冗余校验码(CRC)
- 三种编码的比较

对每种编码都应重点理解其编码原理，检错原理和技术实现算法

检错和纠错

- 问题：数据在传输过程中不可避免的会发生各种错误，那么数据的接收方如何判断收到的数据是否有错(检错)？若出错能否自行修正(纠错)？
- 概念：检错和纠错(Detection & Correction)
 - 相关概念：编码和解码(encoding & decoding)
- 应用：
 - 串行口数据传输：奇偶校验
 - 硬盘数据传输：CRC校验
 - 内存数据检错和传输：奇偶校验/海明校验
 - 无线数据传输：CRC校验/Turbo码.....
 - 磁盘阵列RAID技术

数据校验码

- 概念：数据校验码，编码和解码，漏检问题
- 原理：检错和纠错是建立在要对原始数据进行编码和解码的基础上的
 - 通过函数 f 对数据进行计算，以产生一种代码，代码和数据都被存储和传输
 - 因此，若原始数据字长为 M 位，校验码长为 K 位，则实际存储的字长应该是 $M+K$ 位。
 - 当原先存储的字读出时，这个代码用于检错和纠错，在 M 位数据中产生一组新的 K 位代码，与取出的代码进行比较：
 - 结果一致，无差错，取出的数据位传送出去；
 - 检测到差错，并可以纠正，数据位和纠错位一起送入纠正器，然后产生一组正确的 M 位数据位；
 - 检测到差错，但无法纠正，报告出错

奇偶校验码

● 编码原理

- 不管数据位长度多少，校验位只有一位。
- 若数据位和校验位一起包含奇数个“1”，则称为**奇校验**。
- 若数据位和校验位一起包含偶数个“1”，则称为**偶校验**。

例：数据	奇校验的编码	偶校验的编码
00000000	1 00000000	0 00000000
01010100	0 01010100	1 01010100
01111111	0 01111111	1 01111111

● 检错(校验)原理

- 对采用奇校验的数据，如接收端收到是偶码，则表示传送传输有误，采用偶校验的情形类似。

• => 因此，奇偶校验可发现1位错，但无法定位错误发生的位置

奇偶校验码

实现

- 无论编码还是纠错，都可方便的采用异或运算
奇性= $a_0 \oplus a_1 \oplus a_2 \oplus \dots \oplus a_n$
- 既可以方便的软件实现，也可以方便的硬件实现

1位纠错海明码(原理)

海明码

- 不仅能发现1位错，而且还能指出错误位置，因此可支持纠错。

原理/思路

- 设有R位校验位，则共有0到 2^R-1 个共 2^R 个组合。若用0表示无差错，则剩余 2^R-1 个值表示有差错，并指出错在第几位。由于差错可能发生在K个数据位中或R个校验位中，因此必然有：

$$2^R - 1 \geq R + K$$

海明校验码(校验位数)

— 海明码需要几位校验码?

- k —数据位的位数 r —校验位的位数
数据位 k 与校验位 r 的关系: $2^r \geq k+r+1$

数据位 k	校验位 r	总位数 n
1	2	3
2~4	3	5~7
5~11	4	9~15
12~26	5	17~31
27~57	6	33~63
58~120	7	65~127

海明校验码(校验位置)

— 校验位和数据位是如何排列的

校验位排列在 2^{i-1} ($i=0,1,2,\dots$) 的位置上

例：有一个BCD码 为 $D_4D_3D_2D_1$, 由此生成一个海明码

7	6	5	4	3	2	1
D_4	D_3	D_2	P_3	D_1	P_2	P_1
			2^2		2^1	2^0

有一字节的信息需生成海明码

D_8	D_7	D_6	D_5	P_4	D_4	D_3	D_2	P_3	D_1	P_2	P_1
				8				4		2	1

校验位取值公式(校验计算)

— 海明码的校验位 P_i 和数值位 D_i 的关系

例：BCD码为1011

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

$$0 = 1 \oplus 0 \oplus 1$$

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

$$0 = 1 \oplus 0 \oplus 1$$

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

$$1 = 1 \oplus 1 \oplus 1$$

	7	6	5	4	3	2	1
2^2	D_4	D_3	D_2	P_3			
2^1	D_4	D_3			D_1	P_2	
2^0	D_4		D_2		D_1		P_1

最后 海明码为1010101

1位纠错海明码的实现

● 把 $M+K=m$ 个数据 $\Rightarrow H_m H_{m-1} \dots H_2 H_1$ (海明码), 每个校验位 P_i 在海明码中被分配在 2^{i-1} 位置上。

● H_i 由多个校验位校验 (多重校验思想)

- 每个海明码的位号要等于参与校验它的几个校验位的位号之和。
- 目的: 尽量增大合法码的码距。

● 举例: $M=8$, $K=4$, 则海明码为:

$H_{12} H_{11} H_{10} H_9 H_8 H_7 H_6 H_5 H_4 H_3 H_2 H_1$

P_4, P_3, P_2, P_1 根据规则, 分别对应于: H_8, H_4, H_2, H_1

这样, $P_1 = M_1 \oplus M_2 \oplus M_4 \oplus M_5 \oplus M_7$

$P_2 = M_1 \oplus M_3 \oplus M_4 \oplus M_6 \oplus M_7$

$P_3 = M_2 \oplus M_3 \oplus M_4 \oplus M_8$

$P_4 = M_5 \oplus M_6 \oplus M_7 \oplus M_8$

1位纠错海明码的实现(续)

	12	11	10	9	8	7	6	5	4	3	2	1
2^0		M ₇		M ₅		M ₄		M ₂		M ₁		P ₁
2^1		M ₇	M ₆			M ₄	M ₃			M ₁	P ₂	
2^2	M ₈					M ₄	M ₃	M ₂	P ₃			
2^3	M ₈	M ₇	M ₆	M ₅	P ₄							

$$P_1 = M_1 \oplus M_2 \oplus M_4 \oplus M_5 \oplus M_7$$

$$P_2 = M_1 \oplus M_3 \oplus M_4 \oplus M_6 \oplus M_7$$

$$P_3 = M_2 \oplus M_3 \oplus M_4 \oplus M_8$$

$$P_4 = M_5 \oplus M_6 \oplus M_7 \oplus M_8$$

基于海明码的纠错

海明码的接收端的公式:

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

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

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

- 假定 海明码1010101在传送中变成了1000101

$$S_3 = P_3 \oplus D_4 \oplus D_3 \oplus D_2 = 0 \oplus 1 \oplus 0 \oplus 0 = 1$$

$$S_2 = P_2 \oplus D_4 \oplus D_3 \oplus D_1 = 0 \oplus 1 \oplus 0 \oplus 1 = 0$$

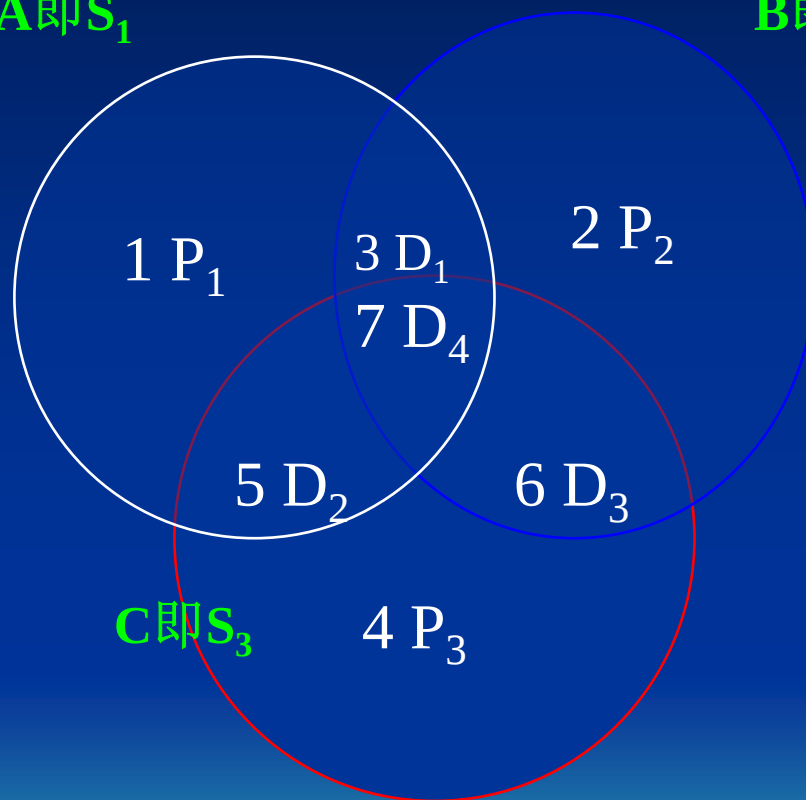
$$S_1 = P_1 \oplus D_4 \oplus D_2 \oplus D_1 = 1 \oplus 1 \oplus 0 \oplus 1 = 1$$

因此,由 $S_3S_2S_1 = 101$,指出第5位错,应由0变1

校验位用文氏图计算举例

A即 S_1

B即 S_2

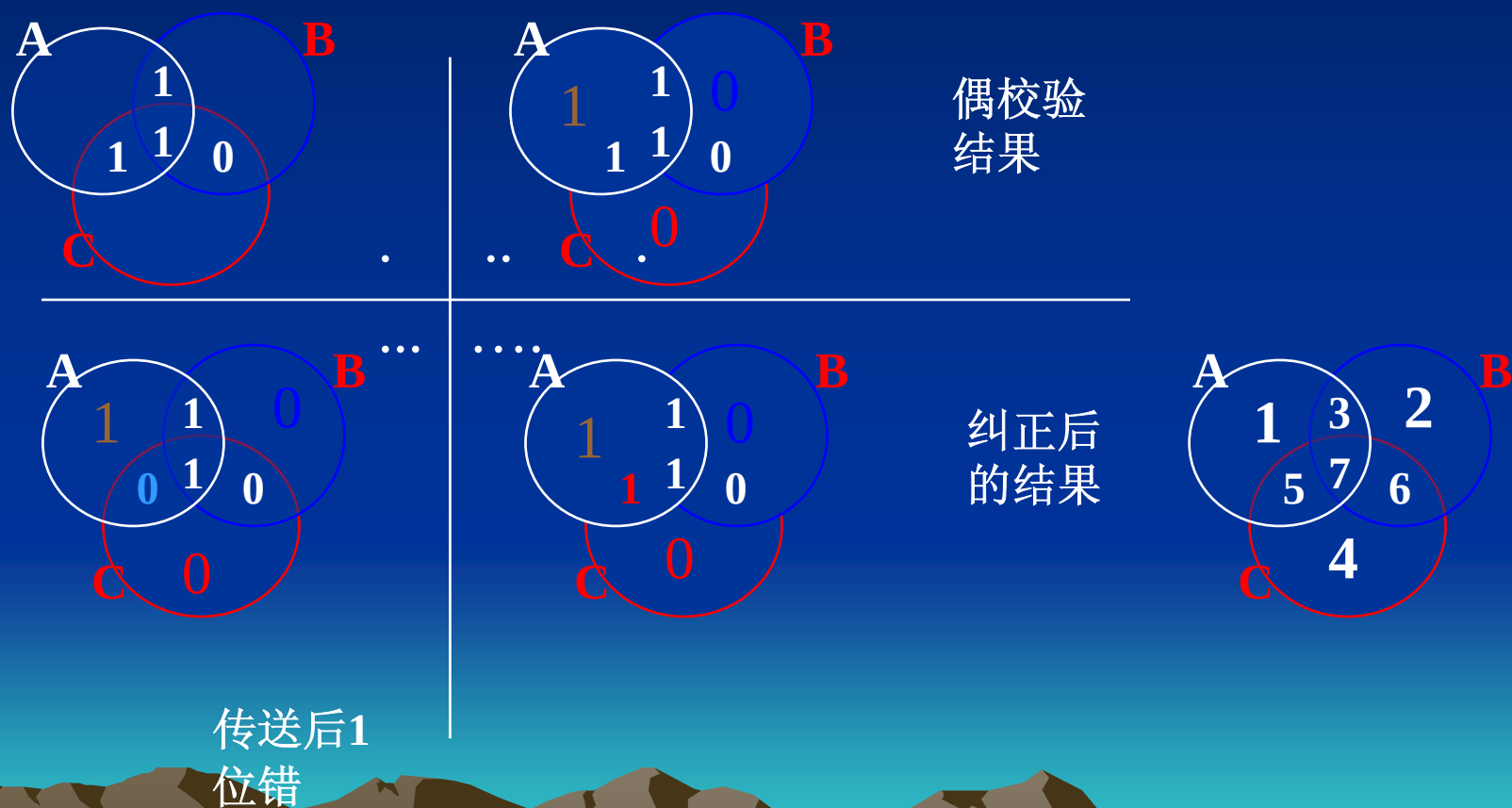


用文氏图计算举例

7	6	5	4	3	2	1
D_4	D_3	D_2	P_3	D_1	P_2	P_1

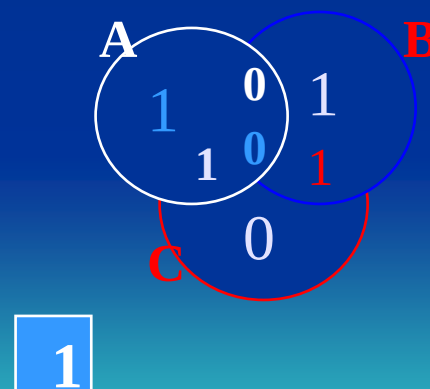
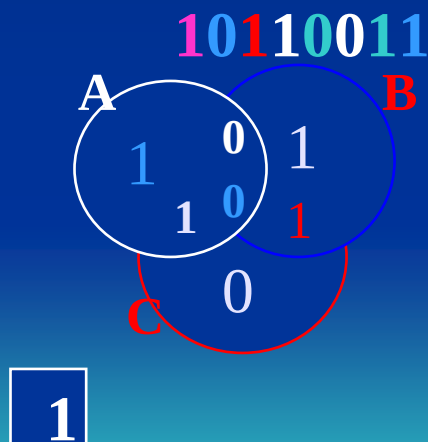
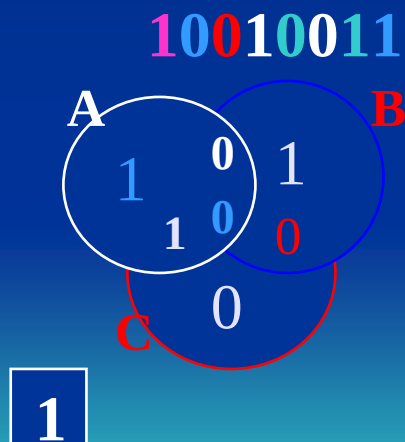
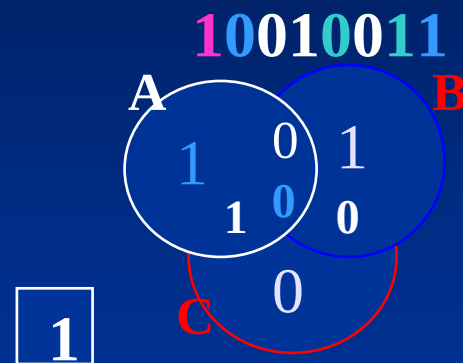
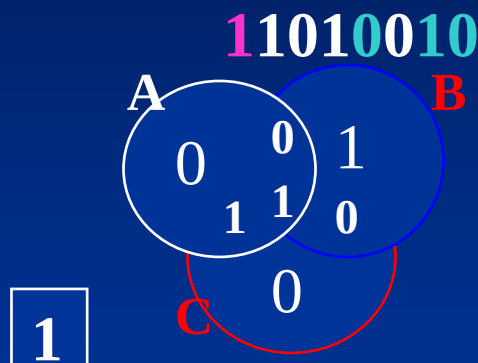
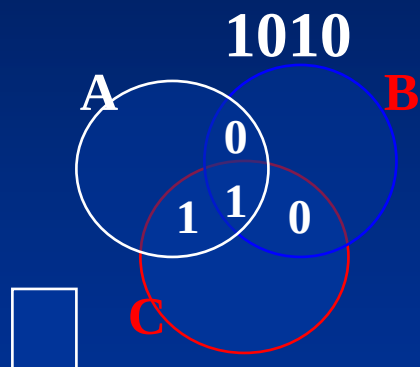
海明码的纠错原理(文氏图)

4位海明纠错码的纠错原理（文氏图）



1位纠错2位海明码的实现

如出现两位错,必须再增加一位奇偶校验位



1位纠错2位海明码的实现

- 上述这种码称为单纠错码（SEC）。通常半导体存储器采用SEC-DED（单纠错-双检错码）。
- 双检错码与SEC相比需要增加1个附加位。
- 实例
 - 在IBM3000系列中，主存64位数据采用8位SEC-DED码进行校验，占了12%的系统开销；
 - VAX计算机中32位字长机器，采用7位SEC-DED码，占了22%的系统开销

循环冗余校验(CRC)码

循环冗余校验码基本原理

- CRC 码是将一串二进制看成是系数为0或1的多项式，一个有 k 位组成的帧可以看成从 x^{k-1} 到 x^0 的 k 次多项式的系数序列，这个多项式的阶数为 $k-1$ 。最高位是 x^{k-1} 项的系数，次高位是 x^{k-2} 的系数，且多项式以2为模运算。
- 校验和的计算思路
 - (1) 设 $G(x)$ 为 r 阶，在帧的末尾附加 r 个零，使帧具有 $m+r$ 位数据，相应的多项式可以表示为 $x^r M(x)$ 。
 - (2) 按照模2除法，用 $x^r M(x)$ 除以对应的生成多项式 $G(x)$ 。
 - (3) 由上面的步骤得到一个商 $Q(x)$ 和一个余数 $R(x)$ ，将其中的余数 $R(x)$ 作为校验和。

循环冗余校验(CRC)码

CRC码的编码方法

- 模2加减: 异或逻辑 — $0 \oplus 0 = 0$; $0 \oplus 1 = 1$;
 $1 \oplus 0 = 1$; $1 \oplus 1 = 0$.

- 模2乘:

$$\begin{array}{r} 1010 \\ \times 101 \\ \hline 1010 \\ 0000 \\ 1010 \\ \hline 100010 \end{array}$$

- 模2除:

$$\begin{array}{r} 101 \\ 101 \overline{) 10000} \\ \underline{101} \\ 010 \\ 000 \\ \underline{100} \\ 101 \\ \underline{01} \end{array}$$

循环冗余校验(CRC)码

● 例: $M(x) = x^3 + x^2 = 1100$

$$M(x) \cdot x^3 = x^6 + x^5 = 1100000$$

$$G(x) = x^3 + x + 1 = 1011$$

$$\frac{M(x) \cdot x^3}{G(x)} = \frac{1100000}{1011} = 1110 + \frac{010}{1011}$$

$$\begin{aligned} M(x) \cdot x^3 + R(x) &= 1100000 + 010 \\ &= 1100010 \end{aligned}$$

编好的循环校验码称为(7,4)码,即 $n=7, k=4$

循环冗余校验(CRC)码

CRC的译码与纠错

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	余数
正确	1	1	0	0	0	1	0	000
A_7 错	1	1	0	0	0	1	1	001
A_6 错	1	1	0	0	0	0	0	010
A_5 错	1	1	0	0	1	1	0	100
A_4 错	1	1	0	1	0	1	0	011
A_3 错	1	1	1	0	0	1	0	110
A_2 错	1	0	0	0	0	1	0	111
A_1 错	0	1	0	0	0	1	0	101

将收到的CRC码用约定的生成多项式 $G(x)$ 去除,如果码字无误则余数应位0,如果有某一位出错,则余数不为0,不同位数出错余数不同.

如果对余数补个0继续除下去,我们将发现一个现象:各次余数将按上表顺序循环.

CRC码的纠错方法

● 根据不同的余数来纠正不同的出错位

● 循环除法:

- 余数添0继续除法,同时信息部分进行循环移位.
- 当最高位变成101时,则最高位取反,纠错.
- 继续余数除法,直至循环一遍,余数变成第一次的余数.
- 例A₅出错

1100110	100	0101100	101	继续循环
1001101	011	1011000	001	
0011011	110	0110001	010	
0110110	111	1100010	100	循环结束
1101100	101			
0				

余数为101
最高位取反,纠错

生成多项式G(x)

生成多项式G(x)应能满足下列要求:

- 任何一位发生错误都应使余数不为0.
- 不同位发生错误应当使余数不同.
- 对余数继续作模2除,应使余数循环.
- 例: $x^7+1=(x+1)(x^3+x+1)(x^3+x^2+1)$
 - $G(x)=x+1=11$ (7,6)码,判一位错
 - $G(x)=x^3+x+1=1011$ (7,4)码,判二位错或纠一位错
 - $G(x)=x^3+x^2+1=1101$ (7,4)码,判二位错或纠一位错
 - $G(x)=(x+1)(x^3+x+1)=11101$ (7,3)码,判二位错并纠一位错

三种编码的比较

● 奇偶校验码

- 简单，效率高，检1位错，漏检率高

● 海明码

- 复杂，效率低，检2或纠1，适用于字节/字数据

● 循环冗余校验码

- 复杂，检2或纠1，适用于块数据

参考

🌐 苏州大学计算机系讲义