

海明码的检错与纠错

核心思想与目的

- 目的：检测并纠正数据传输过程中可能发生的单比特错误。
- 核心思想：在数据位（Data Bits）中插入多个校验位（Parity Bits），并将这些校验位放置在海明码字的特定位置上。通过计算和比对校验位，不仅可以发现错误，还能精确定位错误发生的位置。

关键概念与前置知识

- 校验位（Parity Bit）：用于辨别是否出错，比如常见的奇偶检验
- 海明距离（Hamming Distance）：两个等长码字之间，对应位不同的数量。
例如：000 和 011 的海明距离是 2（第 2、3 位不同）
 - ▶ 要检测 d 个错误，需要海明距离为 $d+1$ 的编码方案
 - ▶ 要纠正 d 个错误，需要海明距离为 $2d+1$ 的编码方案
- 海明码的海明距离是 3，因此它可以检测 2 位错误或纠正 1 位错误

海明码的构造与计算步骤

假设我们要对 4 位数据位 $D_0 D_1 D_2 D_3$ 进行海明码编码。

第 1 步：确定校验位的数量 (k)

- 公式： $2^k \geq m + k + 1$
 - ▶ m = 数据位的长度 (这里是 4)
 - ▶ k = 所需校验位的数量 (通常取其最小值)
- 计算： $2^k \geq 4 + k + 1$ 解得 $k_{\min} = 3$
- 所以需要 3 个校验位 (记为 P_0, P_1, P_2)。
最终海明码的总长度 $n = m + k = 4 + 3 = 7$ 位

第 2 步：确定校验位和数据位的位置

- 海明码的位置从 1 开始编号，校验位 P_i 必须放在第 2^i 的位置上

- ▶ $P_0(i=0)$ 放在 $2^0 = 1$ 号位
- ▶ $P_1(i=1)$ 放在 $2^1 = 2$ 号位
- ▶ $P_2(i=2)$ 放在 $2^2 = 4$ 号位

□ 剩余的位置 (3, 5, 6, 7) 用来存放数据位 $D_0 D_1 D_2 D_3$

最终位置分配:

位置编号	1	2	3	4	5	6	7
内容	P_0	P_1	D_0	P_2	D_1	D_2	D_3
二进制	001	010	011	100	101	110	111

第 3 步: 计算每个校验位负责校验哪些位

□ 每个校验位负责校验位置编号的二进制表示中, 第 i 位为 1 的所有位

- ▶ P_0 (位置 1, 二进制 001): 负责所有最低位为 1 的位置: 1, 3, 5, 7
- ▶ P_1 (位置 2, 二进制 010): 负责所有中间位为 1 的位置: 2, 3, 6, 7
- ▶ P_2 (位置 4, 二进制 100): 负责所有最高位为 1 的位置: 4, 5, 6, 7

这个规则确保了每一个数据位都被多个校验位覆盖。

例如数据位 D_0 (位置 3), 编码为 011, 其中的两个 1 说明被后两位的校验位 P_0 和 P_1 覆盖, 数据为 D_3 (位置 7), 编码为 111, 被全部三个校验位覆盖

第 4 步: 计算每个校验位的值 (偶校验)

□ 根据上一步的分组, 让该组内所有位 (包括校验位自己) 进行异或 (XOR, 不同为 1, 相同为 0) 操作, 结果应为 0

Tip

这里的异或运算在本质上是采用了偶校验, 计数 1 的个数。这里也可以使用其他类型的校验方法, 只不过在纠错时也需要做对应的修改。

□ 但实际上, 我们是根据数据位的值来反推校验位的值, 使得 XOR 结果为 0

- ▶ $P_0 = D_0 \oplus D_1 \oplus D_3$
- ▶ $P_1 = D_0 \oplus D_2 \oplus D_3$
- ▶ $P_2 = D_1 \oplus D_2 \oplus D_3$

第 5 步: 纠错过程 (接收方)

□ 接收方收到一个 7 位的海明码, 对每一个校验组重新计算校验值 (syndrome 位):

- ▶ $S_0 = P_0 \oplus D_0 \oplus D_1 \oplus D_3$

$$\blacktriangleright S_1 = P_1 \oplus D_0 \oplus D_2 \oplus D_3$$

$$\blacktriangleright S_2 = P_2 \oplus D_1 \oplus D_2 \oplus D_3$$

☐ 将计算出的 syndrome 位组成一个二进制数 $S_2 S_1 S_0$

☐ 解读 syndrome:

$\blacktriangleright$ 如果 $S_2 S_1 S_0 = 000$, 则表示没有错误

$\blacktriangleright$ 如果 $S_2 S_1 S_0 \neq 000$, 则这个二进制数对应的十进制值, 就是出错位的位置编号

实例演示

任务: 对数据 $D_0 D_1 D_2 D_3 = 1011$ 进行海明码编码 (采用偶校验)

1. 确定位置: 7 位海明码, 校验位在 1, 2, 4 位

位置	1	2	3	4	5	6	7
内容	P_0	P_1	D_0	P_2	D_1	D_2	D_3
值	?	?	1	?	0	1	1

2. 计算校验位:

$$\square P_0 = D_0 \oplus D_1 \oplus D_3 = 1 \oplus 0 \oplus 1 = 0$$

$$\square P_1 = D_0 \oplus D_2 \oplus D_3 = 1 \oplus 1 \oplus 1 = 1$$

$$\square P_2 = D_1 \oplus D_2 \oplus D_3 = 0 \oplus 1 \oplus 1 = 0$$

3. 填入值, 得到最终海明码:

位置	1	2	3	4	5	6	7
内容	P_0	P_1	D_0	P_2	D_1	D_2	D_3
值	0	1	1	0	0	1	1

所以, 编码后的海明码为: 0110011

4. 假设传输后, 接收方收到的是 0110111 (第 5 位由 0 变成了 1)

位置	1	2	3	4	5	6	7
值	0	1	1	0	1	1	1

5. 计算 Syndrome:

$$\square S_0 = P_0 \oplus D_0 \oplus D_1 \oplus D_3 = 0 \oplus 1 \oplus 1 \oplus 1 = 1$$

$$\square S_1 = P_1 \oplus D_0 \oplus D_2 \oplus D_3 = 1 \oplus 1 \oplus 1 \oplus 1 = 0$$

$$\square S_2 = P_2 \oplus D_1 \oplus D_2 \oplus D_3 = 0 \oplus 1 \oplus 1 \oplus 1 = 1$$

6. 组成 Syndrome: $s_0 s_1 s_2 = 101$, 对应十进制的第 5 位

7. 结论: 第 5 位出错, 接收方只需翻转第 5 位的值 (将 1 变回 0), 即可纠正错误, 得到原始的正确码字 0110011

另一个案例

原码为 101101, 对应的海明码是什么?

$2^k \geq 6 + k + 1$ 解得 $k_{\min} = 4$, 即需要设立 4 个二进制位, 共 10 位, 其中 1、2、4、8 位为编码位。计算校验位:

位置	1	2	3	4	5	6	7	8	9	10
二进制	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010
编码位置	P_0	P_1	1	P_2	0	1	1	P_3	0	1

☐ $P_0 = 1 \oplus 0 \oplus 1 \oplus 0 = 0$ (1、3、5、7、9 位)

☐ $P_1 = 1 \oplus 1 \oplus 1 \oplus 1 = 0$ (2、3、6、7、10 位)

☐ $P_2 = 0 \oplus 1 \oplus 1 = 0$ (4、5、6、7 位)

☐ $P_3 = 0 \oplus 1 = 1$ (8、9、10 位)

因而, 编码后的数据为 0010 0111 01