Codage de Hamming

Introduction

Le codage de Hamming est un type de codage qui peut être utilisé pour corriger automatiquement des erreurs (comme la partie verticale horizontale). Le codage utilise quelques opérations et concepts qui vont être introduits dans ce support.

La somme modulo-2

La somme modulo-2, est une opération mathématique effectuée sur des nombres binaires (ou chaînes de bits). Elle ressemble à l'opération d'addition ordinaire, mais sans retenue. Lorsque vous ajoutez deux nombres binaires, si la somme dépasse 1, la retenue est ignorée.

Si les deux chaînes ont un seul bit, voici le résultat de la somme modulo-2 :

1 +
1
—
0 (sans retenue)

1 +
0
—
1

0 +
1
—
1

0 +
0
—
0

Example 1

Calculons la somme modulo-2 des deux chaînes de bits suivantes : $C_1 = 0\,1\,0\,1\,0\,1\,1$ et $C_2 = 1\,1\,1\,1\,1\,1\,1$

0 1 0 1 0 1 1 +
1 1 1 1 1 1 1
1 0 1 0 1 0 0

Le poids de Hamming

Le poids de Hamming est le nombre de bits égaux à 1 dans une chaîne binaire donnée. Par exemple, si vous avez la chaîne binaire 110101, le poids de Hamming est de 4, car il y a 4 bits égaux à 1.

Example 2

Quel est le poids de Hamming de la chaîne 0 1 1 1 0 1 1 1 0 0 ?

Il suffit de compter les uns dans la chaîne. Le poids de Hamming est égal é 6 dans ce cas-ci.

Distance de Hamming

Il est facile de comprendre le concept de distance entre deux objets. Par exemple, la distance entre deux poteaux d'éclairage peut être mesurée en nombre de pas pour arriver d'un poteau à l'autre.
Nous pouvons étendre le concept de distance à de objets abstraits. Par exemple, nous pouvons définir la distance entre deux chaînes de bits de la même longueur comme le nombre de changements de bit nécessaires pour arriver de la première chaîne à l'autre chaîne. Cette "distance" est appelé la distance de Hamming. En d'autres mots, la distance de Hamming est une mesure de la différence entre deux chaînes de bits de longueur égale.

La manière la plus facile de déterminer la distance de Hamming est de compter le nombre de positions où les bits correspondants sont différents.

Example 3

Supposons que nous avons les deux chaînes de bits suivants :

$$C_1 = 0\,1\,1$$

et

$$C_2 = 0\,1\,0$$

La distance entre ces deux chaînes de bits est égale au nombre de positions où les bits correspondants sont différents. Nous remarquons que sur les trois positions (premier bit, deuxième bit et troisième bit), c'est seulement à la troisième que le bit est différent. La distance de Hamming entre ces deux blocs est égale à 1.

Example 4

Supposons que les chaînes sont maintenant les suivants :

$$C_1 = 0\,1\,0\,1\,0\,1\,1$$

et

$$C_2 = 1\,1\,1\,1\,1\,1\,1$$

Dans cet exemple, c'est le premier, le troisième et le cinquième bit qui sont différents. La distance de Hamming est donc égale à 3.

Il est aussi possible de calculer la distance de Hamming utilisant les concepts de somme modulo-2 et poids de Hamming introduits précédemment.
En effet, la distance de Hamming entre deux chaînes de bits de la même longueur est égale au poids de Hamming de la somme modulo-2 des deux chaînes.
Pour voir ceci, utilisons cette méthode pour les deux chaînes de exemple 4 : $C_1 = 0\,1\,0\,1\,0\,1\,1$ et $C_2 = 1\,1\,1\,1\,1\,1\,1$

Effectuons d'abord la somme modulo-e des deux chaînes. Nous l'avions déjà fait dans l'exemple 1 :

0 1 0 1 0 1 1 +
<u>1 1 1 1 1 1 1</u>
1 0 1 0 1 0 0

Maintenant, il suffit de trouver le poids de Hamming du résultat en comptant le nombre de 1. Le résultat est 3. La distance de Hamming est donc égale à 3. Ce résultat est le même que nous avions trouvé à l'exemple 4.

Nombre d'erreurs subis par une chaîne

Quand nous transmettant une chaîne de bits para un canal de communication, celle-ci peut subir des erreurs dû au perturbation électromagnétiques. Si nous connaissons la chaîne transmise et la chaîne reçue, nous pouvons déterminer le nombre de bit erronés en calculons la distance de Hamming entre les deux.
Par exempt, supposez qu'on vous a transmis la chaîne $c = 0\,1\,1\,1\,0\,1\,1\,0\,1\,0\,1\,1$ et que la chaîne reçue est $c' = 0\,1\,1\,1\,1\,1\,1\,0\,1\,0\,1\,1$.

La distance de Hamming est égale au pods de Hamming de la somme des deux chaînes :

0 1 1 1 0 1 1 0 1 0 1 1 +
<u>0 1 1 1 1 1 1 0 1 0 1 1</u>
0 0 0 0 1 0 0 0 0 0 0 0

Le poids de Hamming du résultat de ls somme est égal à 1. Le nombre de bits erronés est donc 1.

On peut voir que ce résultat est correct en comparant directement les deux chaînes.

Le résultat de la somme modulo-2 de la chaîne transmise et la chaîne reçu s'appelle "la chaîne-erreur" ou le "mot-erreur". Le mot-erreur est une chaine où les positions des erreurs sont marquées par de bits à 1. Le mot-erreur 0 0 0 0 1 0 0 0 0 0 0 0 indique donc qu'il y a eu une seule erreur et que celle-ci est survenue au cinquième bit.

Notez que nous avions utilisé le terme "chaîne de bits" et que nous avons commencer à utiliser le term "mot" pour les désigner. Dans le cadre de systèmes de communication et dans l'informatique, il est très courant d'utiliser le terme "mot" pour se référer à des chaînes de bits.

Mot-code transmis, mot-code reçu, mot-information, bits de contrôle

Ces termes sont essentiels pour comprendre la manière dont les systèmes de communication détectent et corrigent les erreurs dans la transmission de données, en particulier dans le contexte des codes de détection et de correction d'erreurs.

Mots-code transmis :
Les mots-code transmis sont des séquences de bits qui sont envoyées d'un émetteur à un récepteur dans un système de communication. Ils sont généralement le résultat de la codification d'informations pour la transmission, y compris les données à transmettre ainsi que des bits de contrôle pour la détection et la correction d'erreurs.
Mot-code reçu :
Les mots-code reçus sont les séquences de bits reçues par le récepteur d'un système de communication. Ils peuvent être sujets à des erreurs de transmission ou de bruit, et le récepteur doit les traiter pour récupérer les informations d'origine.

Mot-information :
Un mot-information est la partie du mot-code qui contient les données réelles ou les informations à transmettre. Il exclut les bits de contrôle ajoutés pour la détection et la correction d'erreurs.

Bits de contrôle, aussi connus comme bots de redondance ou bits de parité :
Les bits de contrôle sont des bits ajoutés aux mots d'information avant la transmission. Leur rôle est de permettre au récepteur de détecter ou, dans certains cas, de corriger les erreurs de transmission qui peuvent se produire pendant la communication. Ils sont calculés en fonction des bits d'information pour vérifier l'intégrité des données lors de la réception.

Placement des bits de contrôle :
Si les bits de contrôles sont groupés et ajoutés tous ensemble à la fin du mot-information, le code s'appelle systématique. Un exemple est le bit de parité simple qui est ajouté à la fin du mot information de façon ce que le nombre de 1 soit paire.
Si les bits sont au contraire intercalés permis les bits d'information, le code s'appelle non-systématique. Un exemple de code non-systématique est les parties verticales-horizontales.

Exercices

1. Dans chaque cas, additionnez les bits correspondants des deux chaînes en modulo-2.

a)
10101101101
11011010100

b)
100100001
101111011

c)
11110011010
10101010101

d)
110011000011
100110111111

e)
1010100011011
1101011100101

2. Pour chacun de mots suivant, calculez le poids de Hamming :

a) 1101101

b) 1001001

c) 1010101010

d) 1010101010

e) 1111111111

3. Calculez la distance de Hamming entre les mots suivants :

a)

Mot A : 1101101
Mot B : 1011001

b)

Mot A : 1001001
Mot B : 1101011

c)

Mot A : 1110000
Mot B : 0110101

d)

Mot A : 1010101010
Mot B : 1010101010

e)

Mot A : 11110000
Mot B : 00001111

4. Calculez le mot-erreur dans chacun des cas suivants

a)

Mot-code : c = 1101011
Mot reçu : c’ = 1101111

b)

Mot-code : c = 10011001
Mot reçu : c’ = 10111101

c)

Mot-code : c = 1110000
Mot reçu : c’ = 1110100

d)

Mot-code : c = 1010101010
Mot reçu : c’ = 1010001010

e)

Mot-code : c = 1101101
Mot reçu : c’ = 1001101

5. Dans chaque cas ci-dessous, on vous donne le mot-code reçu et le mot-erreur. Trouvez le mot-code qui avait été transmis (c’est à dire, lo mot sans erreurs).

a)

Mot-code reçu : c’ = 1011001

Mot-erreur : e = 0010100

b)

Mot-code reçu : c’ = 1101011

Mot-erreur : e = 0100010

c)

Mot-code reçu : c’ = 1110010

Mot-erreur : e = 1110100

d)

Mot-code reçu : c’ = 10011001

Mot-erreur : e = 11000001

e)
Mot-code reçu : c’ = 1010110

Mot-erreur : e = 1011110

6. Dans les exercices 5. que vous venez de résoudre, additionnez le mot-code reçu c’ et le mot erreur e et comparez le résultat au mot-code transmis c que vous aviez trouvé. Que pouvez vous dire de la formule suivante ?

$$c = c' + e \qquad (1)$$

Codage de Hamming correcteur d’une erreur

Le codage de Hamming directeur d’une erreur est un code non-systématique.

Nous allons expliquer ici comment utiliser le codage de Hamming du point de vue mathématique. Une explication intuitive du fonctionnement sera ajoutée à ce support ultérieurement.

Les bits d’information et de contrôle
Dans le codage de Hamming pour la correction d’une erreur que nous allons étudier, le mot-information a 4 bits auxquels trois bits de contrôle sont ajoutés. Le code es non-systématique, ce qui veut dire, comme expliqué précédemment, que les bits de contrôle sont intercalés parmi les bits d’information.

Puisque le nombre de bits d'information est 4 et le nombre de bits de contrôle est 3, le mot-code a 7 bits. Nous représentons ce bits avec un vecteur come suit :

$$c = [c_1\, c_2\, c_3\, c_4\, c_5\, c_6\, c_7\,] \tag{2}$$

Les bits de contrôle sont placés aux positions correspondant aux puissances successives de 2. Le positions sont donc
$2^0 = 1$
$2^1 = 2$
et
$2^2 = 4$

Comme nous avons expliqué au début, cette explication du codage de Hamming ne donne que des informations sur comment appliquer le codage et il n'y a pour l'instant pas d'explication du pourquoi. Je vous demande donc d'voir de la patience si vous aimeriez comprendre l'origine de ces pas.

Puisque les positions des bits de contrôle sont 1, 2 et 4, les positions des bits d'information sont le reste, c'est à dire les positions 3, 5, 6 e 7.

Le code de Hamming correcteur d'une erreur utilise une matrice qui s'appelle la matrice de contrôle de parité ou simplement matrice de contrôle.

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \tag{3}$$

Pour coder, on multiplie la matrice H par le mot-code transposé (pour lequel on ne connaît que les bits d'information) par et on fonce le résultat à être égale à zéro :

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{4}$$

On procède en multipliant les lignes par les colonnes et on obtient

$$\begin{bmatrix} c_4 & c_5 & c_6 & c_7 \\ c_2 & c_3 & c_6 & c_7 \\ c_1 & c_3 & c_5 & c_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{5}$$

Puisque chaque élément de la matrice doit être égale à gauche et à droite, nous pouvons écrire

$$c_4 + c_5 + c_6 + c_7 = 0 \tag{6a}$$

$$c_2 + c_3 + c_6 + c_7 = 0 \tag{6b}$$

$$c_1 + c_3 + c_5 + c_7 = 0 \tag{6c}$$

Nous pouvons maintenant mettre en évidence dans ce système d'équations les trous bits de contrôle :

$$c_4 = c_5 + c_6 + c_7 \tag{7a}$$

$$c_2 = c_3 + c_6 + c_7 \tag{7b}$$

$$c_1 = c_3 + c_5 + c_7 \tag{7c}$$

Notez que nous avons gardé des signes positifs pour tous les bits. Ceci est mathématiquement légal dans l'arithmétique modulo-2.

Nous pouvons maintenant appliquer ces équations pour trouver les bits de contrôle en fonction des bits 'information.

Example 5

Calculez le mot-code qui correspond au mot information $i = 0010$ Lelong le codage de Hamming.

Nous connaissons les positions des bits d'information dans le mot-code de Hamming :

$$c_3 = 0$$
$$c_5 = 0$$
$$c_6 = 1$$
$$c_7 = 0$$

Nous pouvons maintenant utiliser les équations 7 pour trouver les bits de contrôle :

$$c_4 = 0 + 1 + 0 = 1$$
$$c_2 = 0 + 1 + 0 = 1$$
$$c_1 = 0 + 0 + 0 = 0$$

Le mot-code es donc
0 1 0 1 0 1 0

Example 6

Calculez le mot-code qui correspond au mot information $i = 1110$ Lelong le codage de Hamming.

Nous connaissons les positions des bits d'information dans le mot-code de Hamming :

$$c_3 = 1$$
$$c_5 = 1$$
$$c_6 = 1$$
$$c_7 = 0$$

Nous pouvons maintenant utiliser les équations 7 pour trouver les bits de contrôle :

$$c_4 = 1 + 1 + 0 = 0$$

$c_2 = 1 + 1 + 0 = 0$
$c_1 = 1 + 1 + 0 = 0$

Le mot-code es donc
0 0 1 0 1 1 0

Exercice

Calculez le mot-code qui correspond au mot information $i = 0010$ selon le codage de Hamming.

Comment trouver l'erreur à partir du mot-code reçu

Pour trouver l'erreur (et la corriger) à partir d'un mot-code reçu, on procède de la façon suivante :

On calcule le syndrome en multipliant la matrice de contrôle H par le mod-code reçu transposé

$S = Hc^T$

Le syndrome donne, en binaire, la position du bit erroné. Si le syndrome à une valeur zéro, on peut conclure qu'il n'y a pas eu d'erreur lors de la transmission.

Example 7

Supposez que le mot-code reçu est 0101011. Calculez le syndrome, trouvez l'erreur s'il y a en une et corrigez-la.

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Le bit erroné se trouve à la position 7, qui est la valeur décimale qui correspond au syndrome 111.

Explication intuitive du codage

Le codage repose sur la condition que la multiplication de la matrice de contrôle H et le vecteur contenant le mot-code à transmettre donne zéro :

$$Hc^T = [0] \tag{8}$$

A la réception du mot-code reçu, on calcule le syndrome pour trouver la position à laquelle l'erreur a survenue :

$$Hc'^T = S \tag{9}$$

Nous savons que le mot-code reçu est la somme du mot-code transmis et le mot erreur :

$$c' = c + e \tag{10}$$

Remplaçons maintenant l'équation (10) dans l'équation (9) :

$$H(c + e)^T = S \tag{11}$$

que nous pouvons écrire comme

$$Hc^T + He^T = S \tag{12}$$

Nous savons par (8) que le premier term dans (12) est égal à zéro. Nous pouvons réécrire donc (12) comme suit :

$$He^T = S \tag{13}$$

Utilisons maintenant la matrice H dans l'équation (13)

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix} = S \tag{14}$$

S'il y a une seule erreur, tous les éléments du mot-erreur sauf 1 seront égaux à zéro. Ce fait a pour effect de sélectionner la colonne de la matrice H qui correspond à la position de l'erreur. Par exemple, si le seul bit erroné est à la deuxième position, le mot erreur sera $[0\ 0\ 1\ 0\ 0\ 0\ 0]$ et l'équation (14) devient

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \tag{15}$$

Puisque la matrice de contrôle est faite pour que la colonne i corresponde au numéro i en binaire, cette sélection (tous les éléments sont mutilais par zéro sauf ceux qui correspondent à la troisième colonne) est la raison pour laquelle le syndrome indique la position de l'erreur.