Détection d'erreurs par codage CRC

Durée : deux heures

Introduction

On se propose d'étudier deux techniques de détection d'erreurs dans la transmission de données sur des moyens de communication non fiables. L'adjonction d'un *bit de parité* est utilisée un peu partout : c'est la «clé» dans un numéro de sécurité sociale ou de compte bancaire, ou le treizième chiffre d'un code barre de supermarché. Quant au *codage CRC*, c'est la principale méthode de détection d'erreurs utilisée dans les télécommunications.
Les données transitant sur le réseau sont des séquences de *bits* que l'on notera 0 ou 1. Ces séquences de bits sont découpés en mots $b_0 b_1 \cdots b_{n-1}$ de longueur $n \geqslant 1$, que l'on représentera par un tableau d'entiers $[|b_0; b_1; \cdots; b_{n-1}|]$, avec $b_i \in \{0,1\}$. Le nombre d'erreurs de transmission du mot b est le nombre de bits ayant changé de valeur après transmission.

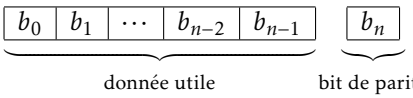
Partie I. Bit de parité

Le «ou exclusif» de deux bits x et y est noté $x \oplus y$ et est défini par la table de valeurs suivante :

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

Question 1. Écrire la fonction `ou_exclusif`, de type `int -> int -> int`, calculant $x \oplus y$ lorsque x et y sont pris comme arguments.

Question 2. La technique du bit de parité consiste à rajouter un bit b_n (le bit de parité) aux données utiles $b_0 b_1 \cdots b_{n-1}$ de façon à ce que le mot $b_0 b_1 \cdots b_{n-1} b_n$ ait un nombre pair de bits égaux à 1. Ainsi, pour $n = 7$, le bit de parité du tableau $[|1;1;0;1;1;1;0|]$ est 1, et le mot transmis sur le réseau est : 11011101.



a. Écrire la fonction `bit_de_parite`, de type `int vect -> int`, qui calcule le bit de parité d'un tableau b.

b. Combien d'erreurs de transmission la technique du bit de parité permet-elle de détecter dans un mot de longueur n ?

Partie II. Le codage CRC

Les lois $\oplus$ et $\times$ (la multiplication usuelle) confèrent à $\mathbb{B} = \{0,1\}$ une structure de corps ; ceci permet de définir l'anneau $\mathbb{B}[X]$ des polynômes à coefficients dans $\mathbb{B}$: on peut considérer un tableau b de n bits comme les coefficients du polynôme :

$$P(X) = b_0 X^{n-1} + b_1 X_{n-2} + \cdots + b_{n-2} X + b_{n-1}.$$

En particulier, la somme de deux polynômes $P(X)$ et $Q(X)$ de $\mathbb{B}[X]$ représentés par les tableaux $[|b_0; b_1; \cdots; b_{n-1}|]$ et $[|c_0; c_1; \cdots; c_{n-1}|]$ est définie par :

$$P(X) \oplus Q(X) = (b_0 \oplus c_0) X^{n-1} + (b_1 \oplus c_1) X^{n-2} + \cdots + (b_{n-1} \oplus c_{n-1}).$$

On notera que $P(X) \oplus P(X) = 0$; tout polynôme est son propre opposé.

Le *degré* du polynôme $P(X)$ est la valeur maximale de k tel que b_{n-1-k} ne soit pas nul.

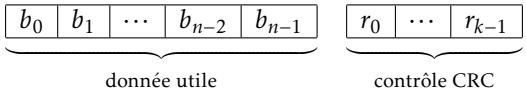
Question 3. Écrire la fonction `degre`, de type `int vect -> int`, prenant en argument un tableau b représentant le polynôme $P(X)$ de $\mathbb{B}[X]$, et retournant le degré de $P(X)$. Par convention, le degré du polynôme nul vaudra -1.

Question 4. Écrire la fonction `plus`, de type `int vect -> int vect -> int -> int -> int -> unit`, prenant comme arguments deux tableaux b et c, deux indices i et j et une longueur ℓ, et qui modifie le tableau b pour remplacer ses coefficients $b_i, b_{i+1}, \ldots, b_{i+\ell-1}$ par : $b_i \oplus c_j, b_{i+1} \oplus c_{j+1}, \ldots, b_{i+\ell-1} \oplus c_{j+\ell-1}$.

La technique appelée CRC (*Cyclic Redundancy Check*) ajoute plusieurs bits de contrôle à chaque mot transmis en procédant comme suit : on choisit un polynôme G(X) (appelé *polynôme générateur*), de degré $k \geqslant 1$, donné une fois pour toutes, et on calcule, pour le mot b à transmettre correspondant au polynôme P(X), le reste de la division euclidienne de $X^k P(X)$ par G(X) :

$$R(X) = X^k P(X) \bmod G(X).$$

Si $[|r_0; r_1; \cdots; r_{k-1}|]$ est le tableau représentant le polynôme R(X), le mot transmis sera :



Question 5. Pour vérifier l'intégrité du message reçu, le destinataire calcule le reste de la division euclidienne du polynôme associé au mot de longueur $n+k$ reçu, par G(X).

a. Montrer que si le mot a été transmis sans erreur, le résultat de ce calcul est nul.

b. Réciproquement, si ce résultat est nul, cela signifie-t-il forcément que le mot a été transmis sans erreur ?

Question 6. On désigne par T(X) le polynôme correspondant au mot transmis, et par $\widetilde{T}(X)$ celui correspondant au mot reçu. Le polynôme $E(X) = T(X) \oplus \widetilde{T}(X)$ indique les bits qui ont étés transmis de manière erronée.

a. Montrer que les erreurs sont détectées dès lors que G(X) ne divise pas E(X).

b. Montrer que si G(X) n'est pas un monôme, une erreur sur un seul bit est toujours détectée.

c. Montrer que si G(X) est un multiple de $X+1$, les erreurs en nombre impair sont détectées.

Question 7. Un paquet d'erreurs de longueur ℓ est une suite de ℓ bits dans le message dont le premier et le dernier bit sont faux, les bits intermédiaires pouvant être vrais ou faux.
On suppose dans cette question que le coefficient constant de G(X) n'est pas nul.

a. Montrer que tout paquet d'erreurs de longueur $\ell \leqslant k$ est détecté.

b. Montrer qu'il existe un seul paquet d'erreurs de longueur $k+1$ qui ne soit pas détecté par le CRC. Montrer que la probabilité d'une telle éventualité est égale à $\dfrac{1}{2^{k-1}}$.

c. Montrer que tout paquet d'erreurs de longueur supérieure ou égale à $k+2$ est détecté avec une probabilité égale à $1 - \dfrac{1}{2^k}$.

d. Un polynôme générateur souvent utilisé est le CRC 16 : $G(X) = X^{16} + X^{15} + X^2 + 1$.
Montrer qu'avec ce générateur, on détecte toutes les erreurs en nombre impair, et tous les paquets d'erreurs de longueur inférieure ou égale à 16. Quelle est la probabilité de détecter un paquet d'erreur de longueur 17 ? et de longueur supérieure ou égale à 18 ?

Pour calculer le reste R(X) de la division de $X^k P(X)$ par G(X), on utilise l'algorithme classique de la division : on aligne les bits de plus haut degré du dividende et du diviseur, puis on retranche le diviseur au dividende (grâce à l'opération $\oplus$). On recommence alors la division en prenant le résultat comme nouveau dividende, et ce jusqu'à ce que son degré soit strictement inférieur à k. Ainsi, pour les tableaux $b = [|0;1;1;1;0;1|]$ et $g = [|0;1;0;1|]$ (et donc $k = 2$) les étapes successives de la division donnent :

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

D'où la valeur 11 pour le CRC.

Question 8.

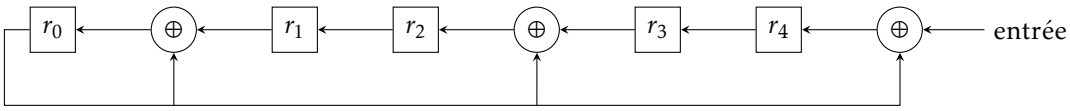
a. Écrire une fonction `crc` de type `int vect -> int vect -> int vect` qui prend en argument deux tableaux b et g et qui retourne le tableaux de bits correspondant aux coefficients du CRC. La valeur de l'argument b ne doit pas être modifiée.

b. Donner un ordre de grandeur du coût en temps et en mémoire pris par la fonction `crc` en fonction de la longueur n de b et du degré k de g.

Question 9. Pour tout $i \in [\![0, n+k-1]\!]$, on pose : $\mathrm{R}_i(\mathrm{X}) = (b_0\mathrm{X}^i + b_1\mathrm{X}^{i-1} + \cdots + b_{i-1}\mathrm{X} + b_i) \bmod \mathrm{G}(\mathrm{X})$, avec pour convention : $b_i = 0$ si $i \geqslant n$. Montrer que si $\mathrm{R}_i(\mathrm{X}) = \alpha_0\mathrm{X}^{k-1} + \alpha_1\mathrm{X}^{k-2} + \cdots + \alpha_{k-1}$, alors :

$$\mathrm{R}_{i+1}(\mathrm{X}) = \Big(\mathrm{X}\mathrm{R}_i(\mathrm{X}) + b_{i+1}\Big) \oplus \Big(\alpha_0\mathrm{G}(\mathrm{X})\Big).$$

Pour calculer rapidement les valeurs du CRC, on réalise des circuits électroniques dédiés à cet usage. Par exemple, le circuit utilisé pour calculer le CRC lorsque $\mathrm{G}(\mathrm{X}) = \mathrm{X}^5 + \mathrm{X}^4 + \mathrm{X}^2 + 1$ est représenté par le schéma suivant :



Initialement, tous les r_i sont nuls ; les bits du mot b d'entrée arrivent sur la droite, suivis de 4 bits valant 0 ; le circuit est synchronisé par une horloge globale qui à chaque tranche de temps décale tous les r_i d'un cran vers la gauche ; le résultat est le tableau de bits $[|r_0; r_1; r_2; r_3; r_4|]$ une fois que tous les bits d'entrée ont étés lus.

Question 10.

a. Expliquer comment fonctionne le circuit ci-dessus, et en déduire le circuit de calcul du CRC lorsque : $\mathrm{G}(\mathrm{X}) = \mathrm{X}^7 + \mathrm{X}^5 + \mathrm{X}^4 + \mathrm{X} + 1$.

b. Montrer que la technique du bit de parité est un cas particulier de la méthode de détection d'erreurs par CRC. Quel est le polynôme générateur correspondant à la méthode du bit de parité ? Et le circuit associé ?

♦ ♦
♦