Texte de préparation à l'épreuve de modélisation de l'agrégation, option C
Autour des codes BCH

25 janvier 2013

Mots clés : corps finis, polynômes cyclotomiques, polynôme minimal, factorisation de polynômes sur un corps fini, algorithme d'Euclide, codes correcteurs d'erreurs.

1 Généralités

Les codes ont été inventés pour détecter et corriger des erreurs sur des canaux de transmission d'information. On les retrouve dans les satellites, les disques compacts, les numéros de billets de banque etc ...

Un *code cyclique binaire* C de longueur impaire n peut être vu comme l'ensemble de tous les polynômes $c(x)$ de $\mathbb{F}_2[x]$ de degrés $< n$ tels que $c(\xi) = 0$ pour tous les ξ dans un ensemble donné $\mathcal{S}$ de racines $n$ièmes de l'unité vivant dans une extension de $\mathbb{F}_2$. L'ensemble $\mathcal{S}$ sera appelé *système de racines* de C par la suite.

La *distance minimale* du code est alors le nombre minimum de termes qui apparaissent dans les polynômes non nuls $c(x)$. Dire qu'un code est *t-correcteur d'erreurs* signifie que la distance minimale du code est $\geq 2t + 1$.

Le ppcm des polynômes minimaux sur $\mathbb{F}_2$ des éléments de S est un diviseur de $x^n - 1$ appelé *polynôme générateur* de C.

Un problème fondamental en théorie des codes est de déterminer ou de borner la distance minimale d'un code cyclique connaissant son système de racines $\mathcal{S}$ ou son polynôme générateur.

On a un premier résultat :

Proposition 1 *Soit $s \in \mathbb{N}^*$ et soit $n = 2^s - 1$. Soit $\xi \in \mathbb{F}_{2^s}$ une racine primitive nième de l'unité. Si l'ensemble $\mathcal{S}$ est égal à $\{\xi\}$, alors le code cyclique défini par $\mathcal{S}$ est un code $[n, n - s, 3]$.*

Le code défini précédemment n'est pas très intéressant car il ne permet de corriger qu'une seule erreur. Les codes BCH ont été inventés par Bose, Chaudhuri et Hocquenghem dans les années 70. Ce sont des codes cycliques pour lesquels on peut garantir à l'avance une borne sur la distance minimale. De plus, ils ont de bonnes propriétés de décodage.

Définition 1 *Un code cyclique binaire de longueur $n = 2^s - 1$ (s entier non nul) sur $\mathbb{F}_2$ est un* code BCH (primitif) de distance prescrite $\delta \geq 2$ *si son système de racines $\mathcal{S}$ est*

$$\mathcal{S} = \{\xi, \xi^2, \ldots, \xi^{\delta-1}\}$$

où ξ est une racine primitive n-ième de l'unité.

Proposition 2 *La distance minimale d'un code BCH de distance prescrite δ est $\geq \delta$.*

2 Construction du polynôme générateur

Pour construire le polynôme générateur d'un code BCH de longueur n, on a besoin d'une racine primitive n-ième de l'unité qui est caractérisée par son polynôme minimal. Ce polynôme est un polynôme irréductible divisant le $n$ième polynôme cyclotomique $\Phi_n(x)$.

2.1 Construction des polynômes cyclotomiques

Soit n un entier non nul et soit un corps fini contenant des racines n-ièmes de l'unité. Le $n$ième polynôme cyclotomique, $\Phi_n(x)$, est le produit des $x - w$ où w est une racine primitive n-ième de l'unité.

Ils vérifient deux propriétés fondamentales

$$x^n - 1 = \prod_{d|n} \Phi_d$$

et par la forme multiplicative du théorème d'inversion de Mobius :

$$\Phi_n(x) = \prod_{d|n} (x^d - 1)^{\mu(n/d)}$$

où μ désigne la fonction de Mobius définie par

$$\mu(n) = \begin{cases} 1 & \text{si } n = 1 \\ (-1)^k & \text{si } n \text{ est le produit de } k \text{ premiers distincts} \\ 0 & \text{si } n \text{ n'est pas sans facteur carré.} \end{cases}$$

De cette écriture de $\Phi_n(x)$, on déduit les lemmes suivants.

Lemme 1 *1. Soient p un nombre premier et r un entier non nul non multiple de p.*

(a) $\Phi_p(x) = x^{p-1} + \cdots + x + 1$.

(b) $\Phi_{rp}(x) = \dfrac{\Phi_r(x^p)}{\Phi_r(x)}$.

2. Soit n un entier non nul et soit m sa partie sans facteur carré. On a

$$\Phi_n(x) = \Phi_m(x^{\frac{n}{m}}).$$

L'algorithme 1 qui découle de ce lemme permet de calculer efficacement le n-ième polynôme cyclotomique.

2.2 Factorisation de $\Phi_n(x)$

Proposition 3 *Soit $n = 2^s - 1$ avec $s > 0$. Le polynôme $\Phi_n(x)$ se factorise en un produit de $\varphi(n)/s$ polynômes irréductibles sur $\mathbb{F}_2$ de degré s.*

Pour chercher un facteur irréductible de $f(x) = \Phi_n(x)$, on va utiliser l'algorithme 2, de Cantor-Zassenhauss, qui permet de trouver un facteur d'un polynôme dont tous les facteurs irréductibles ont même degré.

Notons

$$f = f_1 \cdots f_r, \deg(f_i) = s, f_i \text{ irreductible}, r = \varphi(n)/s$$

D'après le théorème chinois, il existe un isomorphisme d'anneau Ξ

$$\begin{array}{rcccl} \Xi & : & R = \mathbb{F}_2[x]/(f) & \rightarrow & \mathbb{F}_2[x]/(f_1) \times \cdots \times \mathbb{F}_2[x]/(f_r) \\ & & a \bmod f & \mapsto & (a \bmod f_1, \ldots, a \bmod f_r) = (\Xi_1(a), \ldots, \Xi_r(a)) \end{array}$$

Si $b \in \mathbb{F}_2[x]$ est tel qu'il existe $i \neq j$ vérifiant $\Xi_i(b) = 0$ et $\Xi_j(b) \neq 0$ alors $pgcd(b, f)$ est un facteur non trivial de f. On va maintenant construire un b pour lequel $pgcd(f, b)$ a de "grandes chances" d'être un facteur non trivial de f.

On définit le d-ième polynôme trace sur $\mathbb{F}_2$ par

$$T_d(x) = x^{2^{d-1}} + \cdots + x^4 + x^2 + x$$

On a

$$T_d(x)(T_d(x) + 1) = x^{2^d} + x$$

De plus $T_d(x)$ et $T_d(x) + 1$ ont même degré 2^{d-1} et pour tout α de $\mathbb{F}_{2^d}$, on a $T_d(\alpha)(T_d(\alpha) + 1) = 0$ donc pour $\alpha \in \mathbb{F}_{2^d}$, on a $T_d(\alpha) = 0$ ou $T_d(\alpha) = 1$ avec la même probabilité.

Soit $a \in \mathbb{F}_2[x]$ premier avec f. Soit $i \in \{1, \ldots, r\}$, $\Xi_i(a) \in \mathbb{F}_{2^d}$ donc $T_d(\Xi_i(a))$ prend la valeur 0 ou 1 avec la même probabilité.

Connaissant un facteur irréductible de $\Phi_n(x)$, on peut en déduire un polynôme primitif qui sera polynôme minimal d'une racine primitive n-ième de l'unité ξ.

Le calcul du polynôme générateur en découle immédiatement : c'est le ppcm des polynômes minimaux de $\xi, \ldots \xi^{\delta-1}$.

3 Décodage

Pour le décodage, nous allons utiliser la notion d'approximant de Padé.

3.1 Approximants de Padé

Soit F un corps et soit $S = \sum_{i \geq 0} S_i z^i \in F[[z]]$. Un approximant de Padé pour S est une fonction rationnelle

$$\frac{R}{T} \in F(z), R, T \in F[z], z \not| T$$

qui approxime S à une puissance de z suffisamment grande. Plus précisément, pour n, k entiers tels que $1 \leq k \leq n$, $\frac{R}{T}$ est un approximant de Padé $(k, n-k)$ de S si

$$\begin{cases} \frac{r}{t} \equiv S \ [z^n] \\ \\ z \not| T \\ \deg(R) < k \\ \deg(T) \leq n-k. \end{cases}$$

Le théorème suivant fournit un moyen simple pour calculer un approximant de Padé basé sur l'algorithme d'Euclide étendu appliqué à z^n et S mod z^n.

Théorème 1 *Soit $f \in F[z]$, $\deg(f) \leq n$, $k \in \{0, \ldots, n\}$.*

Soit r_i la suite des restes successifs dans l'algorithme d'Euclide appliqué à $r_0 = z^n$ et $r_1 = f$.

Soit (t_i) la suite définie par $t_0 = 0, t_1 = 1$ et $t_{i+1} = t_{i-1} - q_i t_i$ où q_i est tel que $r_{i+1} = r_{i-1} - q_i r_i$.

Soit j le plus petit entier tel que $\deg(r_j) < k$.

Si $\frac{R}{T}$ est un approximant de Padé $(k, n-k)$ de f avec $pgcd(R, T) = 1$, alors il existe τ dans F tel que $R = r_j/\tau$ et $T = t_j/\tau$.

3.2 Idée générale du décodage sur $\mathbb{F}_2$

Notons c un mot du code et $e \in \mathbb{F}_2^n$ de poids $r \leq t$. Soit $v = c + e$ soit le mot reçu. Il s'agit ici de retrouver c connaissant v.

On note $e_1, \ldots, e_r$ les positions des erreurs $0 \leq e_1 < \cdots < e_r \leq n-1$ et pour $i \in \{1, \ldots, r\}$, $X_i = \xi^{e_i}$. En particulier, on a $e(\xi^j) = \sum_{i=1}^r X_i^j$.

Théorème 2 *Soient $\sigma(z), w(z)$ et $S(z)$ les polynômes de $\mathbb{F}_{2^s}[z]$ définis par*

$$\sigma(z) = \prod_{j=1}^{r} (1 - X_j \, z)$$

$$w(z) = \sum_{i=1}^{r} X_i \prod_{j=1, j \neq i}^{r} (1 - X_j \, z)$$

$$S(z) = \sum_{j=1}^{\infty} e(\xi^j) z^{j-1}.$$

On a

- $deg(\sigma) = r$ *et* $deg(w) < r$,
- σ *et* w *premiers entre eux,*
- $w(z) = S(z)\,\sigma(z)$.

Dit autrement, $w(z)/\sigma(z)$ *est un approximant de Padé* (r, r) *de* $S(z)$.

Maintenant, il suffit de remarquer que l'on est capable de calculer $S(z)$ mod z^{2t} connaissant le mot reçu v, ainsi, on va pouvoir calculer un approximant de Padé (r, r) de $S(z)$ en utilisant le théorème de la section précédente. Une fois que σ est trouvé (son coefficient constant est égal à 1), une recherche exhaustive de ses racines fournit les positions d'erreurs $e_1, \ldots, e_r$.

Suggestions

1. Justifier les affirmations du texte laissées sans explication.
2. Montrer que l'algorithme du calcul des polynômes cyclotomiques est correct et évaluer son coût.
3. Programmer les algorithmes du calcul du polynôme cyclotomique et de la factorisation de Cantor-Zassenhaus. Comparer avec les commandes Maple.
4. Programmer un algorithme de construction de polynôme générateur des codes BCH.
5. Calculer tous les polynômes générateurs des codes BCH de longueur 15, 31, 61 et 123.
6. On pourra mettre en oeuvre l'algorithme de décodage présenté sur $\mathbb{F}_2$ et proposer une généralisation sur $\mathbb{F}_q$, $q \neq 2$.

Algorithme 1

Entrée : $n \in \mathbb{N}^*$

Sortie : n-ième polynôme cyclotomique

1: Déterminer les facteurs premiers distincts $p_1, \dots p_r$ de n.
2: $m \leftarrow p_1 \cdots p_r$
3: $f_0 \leftarrow x - 1$
4: **pour** $i = 1$ à r **faire**
5: $\quad f_i(x) \leftarrow f_{i-1}(x^{p_i})/f_{i-1}(x)$
6: **fin pour**
7: **rendre** $f_r(x^{n/m})$

Algorithme 2

Entrée : un polynôme unitaire f sans facteur carré de $\mathbb{F}_2[x]$ de degré $n > 0$ ayant tous ses facteurs irréductibles de même degré d

Sortie : un facteur de f ou faux

1: Choisir a dans $\mathbb{F}_2[x]$ au hasard de degré $< n$
2: $g \leftarrow pgcd(f, a)$
3: **si** $g \neq 1$ **alors**
4: $\quad$ **rendre** g
5: **fin si**
6: $b \leftarrow T_d(a) \bmod f$
7: $g \leftarrow pgcd(b - 1, f)$
8: **si** $g \neq 1$ et $g \neq f$ **alors**
9: $\quad$ **rendre** g
10: **sinon**
11: $\quad$ **rendre** faux
12: **fin si**