Complexité d'ordre maximal des suites Thue–Morse et de Rudin–Shapiro

Pierre Popoli

Université de Lorraine

EJCIM 2020, 16 Juin 2020

Contexte: p premier, $\mathcal{S} = (s_n)_n$ suite sur $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ et $N \geq 1$.

Complexité linéaire au rang N

$L(\mathcal{S}, N)$ est le plus petit entier L tel que

$$s_{i+L} = c_0 s_i + \cdots + c_{L-1} s_{i+L-1},$$

avec $c_j \in \mathbb{F}_p$ et $0 \leq i \leq N - L - 1$. On peut engendrer les N premiers termes à partir des L premiers par une récurrence linéaire.

Contexte: p premier, $\mathcal{S} = (s_n)_n$ suite sur $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ et $N \geq 1$.

Complexité linéaire au rang N

$L(\mathcal{S}, N)$ est le plus petit entier L tel que

$$s_{i+L} = c_0 s_i + \cdots + c_{L-1} s_{i+L-1},$$

avec $c_j \in \mathbb{F}_p$ et $0 \leq i \leq N - L - 1$. On peut engendrer les N premiers termes à partir des L premiers par une récurrence linéaire.

Complexité d'ordre maximal au rang N

$M(\mathcal{S}, N)$ est le plus petit entier M tel que

$$s_{i+M} = f(s_i, \ldots, s_{i+M-1}),$$

avec $f(X_1, \ldots, X_M) \in \mathbb{F}_p[X_1, \ldots, X_M]$ et $0 \leq i \leq N - M - 1$.

On a trivialement

$$M(\mathcal{S}, N) \leq L(\mathcal{S}, N).$$

$G(x)$ la série génératrice de $\mathcal{S}$: $G(x) = \sum_{n \geq 0} s_n x^n$.

Complexité d'expansion au rang N

$E(\mathcal{S}, N)$ est le plus petit degré total de $h(x, y) \in \mathbb{F}_p[X, Y]$ tel que

$$h(x, G(x)) \equiv 0 \mod x^N.$$

$G(x)$ la série génératrice de $\mathcal{S}$: $G(x) = \sum_{n \geq 0} s_n x^n$.

Complexité d'expansion au rang N

$E(\mathcal{S}, N)$ est le plus petit degré total de $h(x, y) \in \mathbb{F}_p[X, Y]$ tel que

$$h(x, G(x)) \equiv 0 \mod x^N.$$

Résultat (Mérai, Niederreiter et Winterhof)

$$E(\mathcal{S}, N) \leq L(\mathcal{S}, N) + 1.$$

Question naturelle

La complexité d'ordre maximal est-elle une mesure plus fine que la complexité d'expansion ?

$G(x)$ la série génératrice de $\mathcal{S}$: $G(x) = \sum_{n \geq 0} s_n x^n$.

Complexité d'expansion au rang N

$E(\mathcal{S}, N)$ est le plus petit degré total de $h(x, y) \in \mathbb{F}_p[X, Y]$ tel que

$$h(x, G(x)) \equiv 0 \mod x^N.$$

Résultat (Mérai, Niederreiter et Winterhof)

$$E(\mathcal{S}, N) \leq L(\mathcal{S}, N) + 1.$$

Question naturelle

La complexité d'ordre maximal est-elle une mesure plus fine que la complexité d'expansion ? Non, il existe des contres exemples.

Complexité en sous-mots

Pour $k \geq 0$, on définit $p_{\mathcal{S}}$ par:

$$p_{\mathcal{S}}(k) = \text{Card}\{(b_0, \ldots, b_{k-1}) \in \mathbb{F}_p^k : \exists i, s_i = b_0, \ldots, s_{i+k-1} = b_{k-1}\}.$$

Complexité en sous-mots

Pour $k \geq 0$, on définit $p_{\mathcal{S}}$ par:

$$p_{\mathcal{S}}(k) = \text{Card}\{(b_0, \ldots, b_{k-1}) \in \mathbb{F}_p^k : \exists i, s_i = b_0, \ldots, s_{i+k-1} = b_{k-1}\}.$$

Une suite est dite *normale* si pour tout bloc $(b_0, \ldots, b_{k-1}) \in \mathbb{F}_p^k$:

$$\lim_{N \to +\infty} \text{Card}\{i < N, s_i = b_0, \ldots, s_{i+k-1} = b_{k-1}\} = \frac{1}{p^k}$$

La complexité en sous-mots est maximale, i.e. chaque bloc apparaît, et chaque bloc apparaît avec la bonne fréquence.

Complexité en sous-mots

Pour $k \geq 0$, on définit p_S par:

$$p_S(k) = \text{Card}\{(b_0, \ldots, b_{k-1}) \in \mathbb{F}_p^k : \exists i, s_i = b_0, \ldots, s_{i+k-1} = b_{k-1}\}.$$

Une suite est dite *normale* si pour tout bloc $(b_0, \ldots, b_{k-1}) \in \mathbb{F}_p^k$:

$$\lim_{N \to +\infty} \text{Card}\{i < N, s_i = b_0, \ldots, s_{i+k-1} = b_{k-1}\} = \frac{1}{p^k}$$

La complexité en sous-mots est maximale, i.e. chaque bloc apparaît, et chaque bloc apparaît avec la bonne fréquence.

Suite de Champernowne

$\omega = 0\ 1\ 10\ 11\ 100\ \cdots$, $s(n) = \omega[n]$ est une suite normale de $\mathbb{F}_2$.

Suite de Thue–Morse $\mathcal{T} = (t(n))_n$

$t(0) = 0$ puis $\begin{cases} t(2n) = t(n) \\ t(2n+1) = 1 - t(n). \end{cases}$ $\mathbf{t} = 01101001\cdots$

$$n = \varepsilon_r \dots \varepsilon_0 \implies \begin{cases} 2n & = \varepsilon_r \dots \varepsilon_0 0 \\ 2n+1 & = \varepsilon_r \dots \varepsilon_0 1. \end{cases}$$

Alors $t(n) \equiv s_1(n) \pmod 2$ la somme des chiffres de n en base 2.

Suite de Thue–Morse $\mathcal{T} = (t(n))_n$

$t(0) = 0$ puis $\begin{cases} t(2n) = t(n) \\ t(2n+1) = 1 - t(n). \end{cases}$ $\mathbf{t} = 01101001\cdots$

$$n = \varepsilon_r \dots \varepsilon_0 \implies \begin{cases} 2n & = \varepsilon_r \dots \varepsilon_0 0 \\ 2n+1 & = \varepsilon_r \dots \varepsilon_0 1. \end{cases}$$

Alors $t(n) \equiv s_1(n) \pmod 2$ la somme des chiffres de n en base 2.

Suite de motifs $\mathcal{P}_k = (p_k(n))_n$

$k \geq 1$, $p_k(n) \equiv s_k(n) \pmod 2$ où $s_k(n)$ compte le nombre de fois que le motif $1^{(k)}$ apparaît.

Pour $k = 1$ on trouve la suite de Thue–Morse, $k = 2$ celle de Rudin–Shapiro. Ce sont toutes des suites automatiques.

Notation: $\gg$ est la notation de Vinogradov.

Suite de Thue–Morse

Pour $N \geq 4$, on a

$$M(\mathcal{T}, N) \gg N.^{\text{i}}$$

$$p_{\mathcal{T}}(k) \leq \frac{10}{3}k.^{\text{ii}}$$

$$E(\mathcal{T}, N) \leq 5.$$

[i] Sun et Winterhof [4]

[ii] Allouche et Shallit [1]

Notation: $\gg$ est la notation de Vinogradov.

Suite de Thue–Morse

Pour $N \geq 4$, on a

$$M(\mathcal{T}, N) \gg N.^{\text{i}}$$

$$p_{\mathcal{T}}(k) \leq \frac{10}{3}k.^{\text{ii}}$$

$$E(\mathcal{T}, N) \leq 5.$$

La mesure de la complexité d'ordre maximal est grande mais les deux autres mesures de complexité ne sont pas assez grande comparé à ce que l'on pourrait attendre d'une suite pseudo-aléatoire. Les résultats sont similaires pour les suites de motifs.

[i] Sun et Winterhof [4]

[ii] Allouche et Shallit [1]

En revanche pour $\mathcal{T}_2 = (t(n^2))_n$, la suite le long des carrés on a

Le long des carrés

$$M(\mathcal{T}_2, N) \gg N^{1/2}.$$[i]

$\mathcal{T}_2$ est une suite normale.[ii]

$E(\mathcal{T}_2, N) \to +\infty$, théorème de Christol

[i] Sun et Winterhof [5]

[ii] Drmota, Mauduit et Rivat [2]

En revanche pour $\mathcal{T}_2 = (t(n^2))_n$, la suite le long des carrés on a

Le long des carrés

$$M(\mathcal{T}_2, N) \gg N^{1/2}.$$[i]

$\mathcal{T}_2$ est une suite normale.[ii]

$E(\mathcal{T}_2, N) \to +\infty$, théorème de Christol

Donc la suite de Thue–Morse le long des carrés semble être un meilleur candidat pour être une suite pseudo-aléatoire.

[i] Sun et Winterhof [5]

[ii] Drmota, Mauduit et Rivat [2]

Soit $P \in \mathbb{Z}[X]$, $P(\mathbb{N}) \subset \mathbb{N}$ de degré $d \geq 2$. Soit $\mathcal{T}_P = (t(P(n)))_n$, on sait que

La complexité en sous-mots de $\mathcal{T}_P$ est exponentielle. [i]

$E(\mathcal{T}_P, N) \to +\infty$.

[i] Moshe [3]

Soit $P \in \mathbb{Z}[X]$, $P(\mathbb{N}) \subset \mathbb{N}$ de degré $d \geq 2$. Soit $\mathcal{T}_P = (t(P(n)))_n$, on sait que

La complexité en sous-mots de $\mathcal{T}_P$ est exponentielle. [i]

$E(\mathcal{T}_P, N) \to +\infty$.

Extension du résultat sur la complexité d'ordre maximal

Soit $P \in \mathbb{Z}[X]$, $P(\mathbb{N}) \subset \mathbb{N}$ de degré d unitaire. Soit $\mathcal{T}_P = (t(P(n)))_n$ et $\mathcal{P}_{k,P} = (p_k(P(n)))_n$, alors on a pour $N \geq N_0(k, P)$

$$M(\mathcal{T}_p, N) \gg N^{1/d} \tag{1}$$

$$M(\mathcal{P}_{k,P}, N) \gg N^{1/d} \tag{2}$$

[i] Moshe [3]

Preuve pour Thue–Morse Étape 1/3

Pour prouver un tel résultat il faut trouver deux blocs de taille identiques et $\gg N^{1/d}$ qui ont deux successeurs différents dans les N premiers termes. Pour cela on utilise le fait bien connu que pour $a, b \geq 0$ et $b < 2^r$ on a $s_1(a2^r + b) = s_1(a) + s_1(b)$.

$$\begin{array}{rccl} & (a)_2 & 0\cdots 0 & = a2^r \\ + & & (b)_2 & = b \\ \hline & (a)_2 & 0\cdots(b)_2 & = a2^r + b \end{array}$$

On dit que la somme de a à b est non interférente dans ce cas.

Preuve pour Thue–Morse Étape 1/3

Pour prouver un tel résultat il faut trouver deux blocs de taille identiques et $\gg N^{1/d}$ qui ont deux successeurs différents dans les N premiers termes. Pour cela on utilise le fait bien connu que pour $a, b \geq 0$ et $b < 2^r$ on a $s_1(a2^r + b) = s_1(a) + s_1(b)$.

$$\begin{array}{rccl} & (a)_2 & 0\cdots 0 & = a2^r \\ + & & (b)_2 & = b \\ \hline & (a)_2 & 0\cdots(b)_2 & = a2^r + b \end{array}$$

On dit que la somme de a à b est non interférente dans ce cas.

On montre alors que pour pour tout $r > 0$ et $0 \leq n < c2^l$, on a

$$t(P(n + 2^{dl})) = t(P(n + 2^{dl+r}))$$

Preuve pour Thue–Morse Étape 2/3

On va chercher y, r tels que

$$t(P(1 + y2^l + 2^{dl})) \equiv t(P(1 + y2^l + 2^{dl+r})) + 1 \pmod 2$$

Après étude des interférences possibles cela se ramène à trouver (y, r) tels que

$$t(y^d + z) \equiv t(y^d + 2^r z) + 1 \pmod 2$$

avec $z = P'(1)$. On peut forcer $z \geq 0$ à l'aide d'une translation.

Preuve pour Thue–Morse Étape 3/3

On prend $y = 2^s$ tel que $t(y^d + z) = t(y^d) + t(z) = 1 + t(z)$.
Puis on translate z, à l'aide du r pour avoir la somme suivante avec $z = 1z'$:

$$\begin{array}{rcccl} & 1 & 0 \cdots 0 & & = 2^{sd} \\ + & & 1 & z' & = 2^r z \\ \hline & 1 & 0 & z' & = a2^r + b \end{array}$$

Preuve pour Thue–Morse Étape 3/3

On prend $y = 2^s$ tel que $t(y^d + z) = t(y^d) + t(z) = 1 + t(z)$.
Puis on translate z, à l'aide du r pour avoir la somme suivante avec $z = 1z'$:

$$\begin{array}{rcccl} & 1 & 0 \cdots 0 & & = 2^{sd} \\ + & & 1 & z' & = 2^r z \\ \hline & 1 & 0 & z' & = a2^r + b \end{array}$$

Alors $t(y^d + 2^r z) = t(z)$. En effet l'addition de 2^{sd} à $2^r z$ n'a pas changé le nombre de 1-bit à z.

Suites de motifs Étape 1/2

Pour les suites de motifs, il faut opérer différemment car il n'est pas sûr que z contienne le block $1 \cdots 1$.

Suites de motifs Étape 1/2

Pour les suites de motifs, il faut opérer différemment car il n'est pas sûr que z contienne le block $1 \cdots 1$.
Pour cela on pose $f_a(x) = ax^3 + ax^2 - x + a$, on a $f_a(x)^d = \sum_{0 \leq i \leq 3d} \mu_i x^i$ avec $\mu_i \geq 0$ pour $i \neq 1$ et $\mu_1 < 0$. On prend $x = 2^u$ pour avoir la décomposition binaire suivante

$$y^d = f_a(2^u)^d = \omega_1 0 \ldots 0 \omega_2 0 1^{(\alpha)} 0 \omega_3.$$

Il est important de noter que la taille du bloc de 1 est incrémentée de 1 par passage de u à $u+1$.

Suites de motifs Étape 2/2

On réalise le shift suivant

$$\begin{array}{lllll} & \omega_1 0\ldots 0 & \omega_2 & \overbrace{1\ldots 1}^{\alpha-i} & \overbrace{1\ldots 11}^{i} \quad 0\omega_3 = y^d \\ + & & & & 1\ldots 11 \quad 0\omega' = 2^s z \\ \hline & \omega_1 0\ldots 0 & (\omega_2+1) & 0\ldots 0 & 1\ldots 10 \quad 0(\omega_3+\omega') = y^d + 2^s z \end{array}$$

Alors on adapte u pour qu'un nombre impair de k-motifs soit détruit ou créé.

Problèmes ouverts

- Prendre n'importe quel motif à la place $1^{(k)}$.
- Normalité de la suite de Thue–Morse le long des suites polynomiales (problème difficile).

[1] J.-P. Allouche and J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, Cambridge, 2003.

[2] Michael Drmota, Christian Mauduit, and Joël Rivat, *Normality along squares*, J. Eur. Math. Soc. (JEMS) **21** (2019), no. 2, 507–548.

[3] Yossi Moshe, *On the subword complexity of Thue-Morse polynomial extractions*, Theoret. Comput. Sci. **389** (2007), no. 1-2, 318–329.

[4] Zhimin Sun and Arne Winterhof, *On the maximum order complexity of subsequences of the Thue-Morse and Rudin-Shapiro sequence*, Unif. Distrib. Theory **14** (2019), no. 2, 33–42.

[5] _____, *On the maximum order complexity of subsequences of the Thue-Morse and Rudin-Shapiro sequence along squares*, Int. J. Comput. Math. Comput. Syst. Theory **4** (2019), no. 1, 30–36.

Merci de votre attention !