Investigation of Periodic Sequences with Maximum Nonlinear Complexity

Chunlei Li

Joint work with Zhimin Sun, Xiangyong Zeng and Tor Helleseth

University of Bergen

July 3 - 8, BFA-2017

Outline

Why not 'truly' random sequences

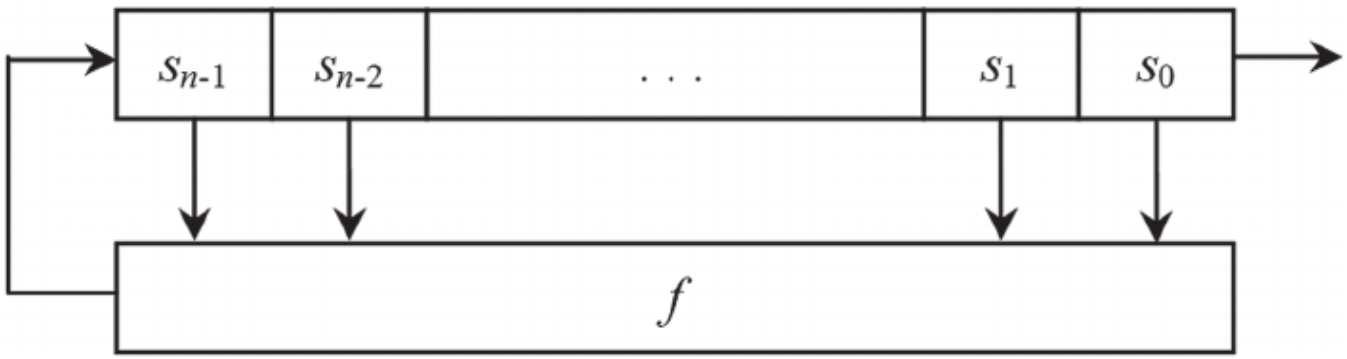
Pseudorandom Sequences

Sequences that are generated by a **deterministic** algorithm and **look random** are called **pseudorandom**

Desirable *randomness properties* depend on the application.

- cryptography: unpredictability
- simulation: uniform distribution
- radar: distinction from reflected signal
- ...

Feedback Shift Registers (FSRs)



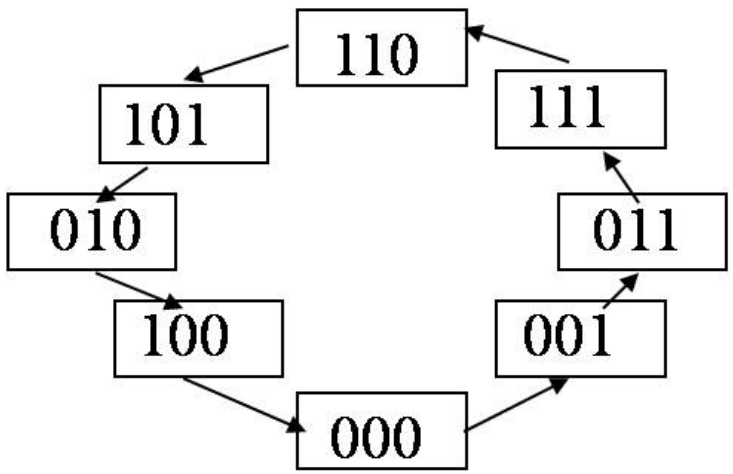
- an initial state $(s_0, s_1, \cdots, s_{n-1})$
- a feedback function: $f(x_0, x_1, \cdots, x_{n-1})$
- FSR sequecnes: for initial states $(s_0, s_1, \cdots, s_{n-1})$, an FSR generates sequences $\mathbf{s} = \{s_i\}$ via the recursion

$$s_{i+n} = f(s_i, s_{i+1}, \cdots, s_{i+n-1}), i \geq 0$$

A toy example

Let $f_1(x_0, x_1, x_2) = 1 + x_0 + x_1 + x_1x_2$.

i	(x_i, x_{i+1}, x_{i+2})	x_{i+3}
0	000	1
1	001	1
2	011	1
3	111	0
4	110	1
5	101	0
6	010	0
7	100	0



The output sequence: 00011101...

Linear Feedback Shift Registers (LFSRs)

The feedback function f is linear, namely, having the form

$$f(x_0, x_1, \cdots, x_{n-1}) = c_0 x_0 + c_1 x_1 + \cdots + c_{n-1} x_{n-1}, \quad c_i \in \mathbb{F}_q$$

The theory of LFSR is well developed (by Ward, Golomb, Selmer, Zierler, etc)

- linear recurrence $s_{t+n} = c_{n-1} s_{t+n-1} + \cdots + c_1 s_1 + c_0 s_0$
- the output sequences $(s_0 s_1 s_2 \cdots)$ can be studied via the characteristic polynomial

$$f(x) = x^n + c_0 x^{n-1} + \cdots + c_1 x + c_0$$

- $per(f) :=$ the smallest integer e such that $f(x) | (x^e - 1)$

Two Fundamental Identities

1. $$g(x) = \sum_{i=0}^{\infty} s_i x^i = \frac{\varphi^*(x)}{f^*(x)}$$

   where $\varphi(x) = \sum_{i=0}^{n-1} \left(\sum_{j=0}^{n-1-i} c_{i+j+1} s_j \right) x^i$

2. for a periodic sequence $(s_0 s_1 s_2 \cdots)$ with period ε,

$$g(x) = \sum_{i=0}^{\infty} s_i x^i = \frac{s_0 + s_1 x + \cdots + s_{\varepsilon-1} x^{\varepsilon-1}}{1 - x^{\varepsilon}} = \frac{\sigma^*(x)}{1 - x^{\varepsilon}}$$

Periods of LFSR sequences

$$g(x) = \frac{\varphi^*(x)}{f^*(x)} = \frac{\sigma^*(x)}{1 - x^\varepsilon}$$

- Let $per(f) = e$ and $F(x) = (x^e - 1)/f(x)$,

$$g(x) = \sum_{i=0}^{\infty} s_i x^i = \frac{\varphi^*(x)}{f^*(x)} = \frac{\varphi^*(x)F^*(x)}{1 - x^e}$$

  $\Rightarrow$ all nontrivial output sequences $\mathbf{s}$ generated from f have $per(f)$ as a general period, i.e., $per(\mathbf{s})|per(f)$
- when $f(x)$ is irreducible, $f(x)|x^\varepsilon - 1 \Rightarrow per(f)|per(\mathbf{s})$
  $\Rightarrow per(f) = per(\mathbf{s})$ for all nontrivial output sequences

Periods of LFSR sequences

$$g(x) = \frac{\varphi^*(x)}{f^*(x)} = \frac{\sigma^*(x)}{1 - x^\varepsilon}$$

$$\Downarrow$$

$$f(x)\sigma(x) = (x^\varepsilon - 1)\varphi(x)$$

- Let $per(f) = e$ and $F(x) = (x^e - 1)/f(x)$,

$$g(x) = \sum_{i=0}^{\infty} s_i x^i = \frac{\varphi^*(x)}{f^*(x)} = \frac{\varphi^*(x)F^*(x)}{1 - x^e}$$

  $\Rightarrow$ all nontrivial output sequences $\mathbf{s}$ generated from f have $per(f)$ as a general period, i.e., $per(\mathbf{s})|per(f)$
- when $f(x)$ is irreducible, $f(x)|x^\varepsilon - 1 \Rightarrow per(f)|per(\mathbf{s})$
  $\Rightarrow per(f) = per(\mathbf{s})$ for all nontrivial output sequences

Periods of LFSR sequences

$$g(x) = \frac{\varphi^*(x)}{f^*(x)} = \frac{\sigma^*(x)}{1 - x^\varepsilon}$$

$$\Downarrow$$

$$f(x)\sigma(x) = (x^\varepsilon - 1)\varphi(x)$$

- Let $per(f) = e$ and $F(x) = (x^e - 1)/f(x)$,

$$g(x) = \sum_{i=0}^{\infty} s_i x^i = \frac{\varphi^*(x)}{f^*(x)} = \frac{\varphi^*(x)F^*(x)}{1 - x^e}$$

  $\Rightarrow$ all nontrivial output sequences $\mathbf{s}$ generated from f have $per(f)$ as a general period, i.e., $per(\mathbf{s})|per(f)$
- when $f(x)$ is irreducible, $f(x)|x^\varepsilon - 1 \Rightarrow per(f)|per(\mathbf{s})$
  $\Rightarrow per(f) = per(\mathbf{s})$ for all nontrivial output sequences

Periods of LFSR sequences

When $f(x)$ is primitive, i.e., $per(f) = 2^n - 1 \Rightarrow$ the well-known maximum-length sequences (m-sequence)

The m-sequences have very good statistical property (satisfying the Golomb's random postulates):

- *balancedness*
- *run-property*
- *2-level ideal autocorrelation*

The m-sequences numerous applications in cryptography, sequence design, coding theory, radar system, GPS, $\cdots$

They lead us to many interesting problems in these fields

when it comes to nonlinear feedback functions, the world has dramatically changed ...

General knowledge about NFSRs is rather limited

- the output sequences are periodic iff. f is nonsingular, i.e.,

$$f = x_0 + g(x_1, \cdots, x_{n-1})$$

- the maximum period of an NFSR sequence is q^n, which is a q-ary DeBruijn sequence of order n
- the total number of such sequence is $\frac{(q!)^{q^{n-1}}}{q^n}$
- when $q = 2$, the number is $2^{2^{n-1}-n}$

Problems with NFSRs are challenging

Periods of NFSR seqences

- hard problem in general
- rather few general results on the period
- some nontrivial result in the case that the feedback function is symmetric NFSRs
  (by Kjeldsen, Søreng during 1970-80s)
- Proofs are in general very technical and hard to read
- Mykkeltveit (1979) used arithmetic codes to study periods of NFSR

Problems with NFSRs are challenging

Generation of NFSR sequences (with prescribed periods)?

Only some results for the extremity case: deBruijn sequences

- comprehensive survey by Fredricksen in 1982
- algorithmic methods (expensive for large n)
- mathematic approaches (some progress with cycle joining method)
  - starting with LFSRs
  - investigate adjacent cycles
  - characterize conjugate paris and join small cycles
  - progress in recent years
    - $(1+x)p(x)$, $(1+x^2)p(x)$ (Mykkeltveit, Hemmati)
    - $(1+x)^3p(x)$, $(1+x^3)p(x)$ (Hellseth, Hu, Li, L. Zeng)
    - $(1+x)\prod_i p_i(x)$ for primitive/irreducible polynomials (Hellseth, Li, Li, Lin, L. Zeng, etc)
    - general polynomial $\prod_i p_i^{e_i}(x)$ (Lin et al.)

FSRs $\Longrightarrow$ pseudo-random sequences

pseudo-random sequences $\Longrightarrow$ FSRs

Linear Complexity

Let $\mathbf{s}_n = (s_0, s_1, \cdots, s_{n-1})^\infty$ be a periodic sequence over $\mathcal{F}$.

The **linear complexity** $lc(\mathbf{s}_n)$ is the length L of the shortest LFSR that generate the sequence $\mathbf{s}_n$

- Berlekamp-Massey algorithm (initially from coding theory)
- find the (unique) shortest LFSR that generate the sequence if $n \geq 2lc(\mathbf{s}_n)$
- theoretic approach: $lc(\mathbf{s}_n) = \deg(f(x))$ and

$$\frac{\phi(x)}{f(x)} = \frac{\sigma(x)}{x^n - 1}$$

$$\Rightarrow lc(\mathbf{s}_n) = \deg\left(\frac{x^n-1}{\gcd(x^n-1, \sigma(x))}\right) = n - \deg(\gcd(x^n - 1, \sigma(x)))$$

Desirable properties for Applications

Sequences for cryptographic use should not have low linear complexity.
However, high linear complexity **does not** guarantee cryptographic strength

- E.g., $0 \cdots 01$ has maximum linear complexity, but poor cryptographic quality
- A sequence with high linear complexity can probably be generated by a (much) shorter FSR with nonlinear feedback function

Other Linear Complexity Measures

The k-**th linear complexity** $lc(\mathbf{s}_n, k)$, $1 \leq N \leq n-1$, is the length L of the shortest LFSR that generate $(s_0, s_1, \cdots, s_{N-1})$

The k-**th error linear complexity** $lc_k(\mathbf{s}_n)$ is the smallest linear complexity that can be obtained by altering at most k positions in $\mathbf{s}_n$

Higher order complexity

The k-**th order nonlinear complexity** of $\mathbf{s} = (s_0, s_1, \cdots, s_{l-1})$ over an alphabet $\mathcal{A}$ is the length of

the shortest FSR with feedback function of degree $\leq k$

that can generate the sequence $\mathbf{s}$.

- $k = 1$: linear complexity
- $k = 2$: quadratic complexity
- $k = 3$: cubic complexity
- ...

Nonlinear Complexity

Maximum Order Complexity (by C. Jansen)

The **nonlinear complexity** of a sequence $\mathbf{s} = (s_0, s_1, \cdots, s_{l-1})$ over a field $\mathcal{F}$ is *the length of the shortest (arbitrary) feedback shift register* that can generate the sequence $\mathbf{s}$.

Basic Facts - (non-periodic) sequences

Given a sequence $\mathbf{s}_n = (s_0, s_1, \cdots, s_{n-1})$ over $\mathcal{F}$,

- $nlc(\mathbf{s})$ = the shortest length l, such that all the subsequences of $\mathbf{s}$ of length l, have unique successor
- $nlc(\mathbf{s})$ = the length-plus-one of the longest subsequences of $\mathbf{s}$ that occurs (at least) twice with different successor
- range of $nlc(\mathbf{s}) : 0 \leq nlc(\mathbf{s}) \leq l - 1$
  - $nlc(\mathbf{s}) = 0$ iff. the sequence $\mathbf{s} = (\alpha, \cdots, \alpha)$
  - $nlc(\mathbf{s}) = l - 1$ iff. the sequence $\mathbf{s} = (\alpha, \cdots, \alpha, \beta)$ for $\alpha \neq \beta$
- a nonlinear complexity $nlc(\mathbf{s}) = c$
  $\Rightarrow$ all l-long subsequence of $\mathbf{s}$ are distinct for any for $l > c$
- calculation of $nlc(\mathbf{s})$: Blumer's algorithm for *directed acyclic word graph* (DAWG)

Basic Facts - periodic sequences

For a periodic sequence $\mathbf{s}_n = (s_0, s_1, \cdots, s_{n-1})^\infty$ over $\mathcal{F}$,

- the nonlinear complexity satisfies the inequality

$$\lceil \log_{|\mathcal{F}|}(n) \rceil \leq nlc(\mathbf{s}_n) \leq n - 1$$

- the nonlinear complexity of a period sequence $\mathbf{s}_n$ is the same as that of its
  - shift equivalent sequences

$$(s_i, s_{i+1}, \cdots, s_{i+n-1})^\infty, \quad i = 0, 1, \cdots, n-1$$

  - transposed sequence

$$T\mathbf{s}_n = (Ts_0, Ts_1, \cdots, Ts_{n-1})^\infty$$

    with an injection $T : \mathcal{A} \to \mathcal{B}$ for alphabet $\mathcal{B}$ with $|\mathcal{B}| \geq |\mathcal{A}|$
  - reciprocal sequence $(s_{n-1}, \cdots, s_1, s_0)^\infty$
- a complexity $nlc(\mathbf{s}_n) = c$
  $\Rightarrow$ there are n distinct l-long subsequence for any for $l \geq c$

Basic Facts - periodic sequences

Feedback Functions Equivalents

For a periodic sequence $\mathbf{s}_n = (s_0, s_1, \cdots, s_{n-1})^\infty$ over a finite field $GF(q)$, suppose the nonlinear complexity of $\mathbf{s}_n$ is c, there exist in total

$$q^{q^c - n}$$

feedback functions that can the sequence $\mathbf{s}_n$

- for $n = q^c$, the sequence $\mathbf{s}_n$ is a DeBruijn sequence of order c
- for $q = 2$, this number is equal to the number of binary DeBruijn sequences of order n

Basic Facts - periodic sequences (cont.)

Given a periodic sequence $\mathbf{s}_n = (s_0, s_1, \cdots, s_{n-1})^\infty$, let $\mathbf{s} = (s_0, s_1, \cdots, s_{n-1})$ (a single period of $\mathbf{s}_n$).

- $nlc(\mathbf{s}_n) \geq nlc(\mathbf{s})$
- $nlc(\mathbf{s}_n) = nlc(\mathbf{s}^t)$ for $t \geq 2$, $\mathbf{s}^t$ denotes the t copies of $\mathbf{s}$

In order to calculate $nlc(\mathbf{s}_n)$,

- it suffices to compute $nlc(\mathbf{s}^2)$
- more precisely, if $nlc(\mathbf{s}_n) = c$, one only needs to investigate

$$nlc(s_0, s_1, \cdots, s_{n-1}, s_0, \cdots, s_{c-2})$$

- by $nlc(\mathbf{s}_n) \leq n - 1$, it suffices to calculate

$$nlc(s_0, s_1, \cdots, s_{n-1}, s_0, \cdots, s_{n-3}),$$

Motivation

- high linear/nonlinear complexity is desirable for (cryptographic) applications
- non-periodic sequences: only one sequence $(\alpha \cdots \alpha\beta)$ has maximum nonlinear complexity
- for periodic sequences,
  - $(\alpha \cdots \alpha\beta)^\infty$ has maximum nonlinear complexity
  - there are other periodic sequences with maximum nonlinear complexity

Question

Can we characterize the periodic sequences with maximum nonlinear complexity?

Necessary Conditions

If a periodic sequence $\mathbf{s}_n = (s_0, s_1, \cdots, s_{n-1})^\infty$ with $nlc = c$,

- there exist two identical $(c-1)$-long subsequence with different successors in $(s_0, s_1, \cdots, s_{n-1}, s_0, \cdots, s_{c-2})$;
- each c-long subsequence of $\mathbf{s}_n$ are distinct

If $nlc(\mathbf{s}_n) = n - 1$, then there is a integer $1 \leq p < n$ such that

- $(s_0, s_1, \cdots, s_{n-3}) = (s_p, s_{p+1}, \cdots, s_{p+n-3})$;
- $s_{n-2} \neq s_{p+n-2}$; moreover, $s_{n-1} \neq s_{p+n-1} = s_{p-1}$

Necessary Conditions (cont.)

$nlc(\mathbf{s}_n) = n-1 \Leftrightarrow \exists\, 1 \le p < n$ such that $s_i = s_{i+p}$ iff. $i \in \mathbb{Z}_{n-2}$

Question: what are such integers p?

Example: Let $n = 10$ and $T_i = \{k \in \mathbb{Z}_n : s_k = s_i\}$ for $i \in \mathbb{Z}_n$

$p = 1$	$T_0 = \{0,1,2,3,4,5,6,7,8\}$	$s_8 \neq s_9,\ s_9 \neq s_0$
$p = 2$	$T_0 = \{0,2,4,6,8\}$ $T_1 = \{1,3,5,7,9\}$	$s_8 \neq s_0,\ s_9 \neq s_1$
$p = 3$	$T_0 = T_1 = \{0,3,6,9,1,4,7,0\}$ $T_2 = \{2,5,8\}$	$s_8 \neq s_1,\ s_9 \neq s_2$
$p = 4$	$T_0 = \{0,4,8\},\ T_1 = \{1,5,9\}$ $T_2 = \{2,6,0\},\ T_3 = \{3,7,1\}$	$s_8 \neq s_2,\ s_9 \neq s_3$
$p = 5$	$T_0 = \{0,5\}, T_1 = \{1,6\}, T_2 = \{2,7\}$ $T_3 = \{3,8\},\ T_4 = \{4,9\}$	$s_8 \neq s_3,\ s_9 \neq s_4$

- Continuing the above process, p cannot be $6, 8$
- It appears that p needs to be coprime to n

Necessary Conditions (cont.)

Take $f = n/\gcd(p, n)$ and define

$$K = \{(p-1) + tp(\mathrm{mod} n) \, : \, t = 0, 1, \cdots, f-2\}.$$

It follows that

- $n - 1 = (p-1) + (f-1)p(\mathrm{mod}\ n) \notin K$;
- $n - 2$ belongs to K; otherwise,
  $s_{p-1} = \cdots = s_{(p-1)+(f-2)p} = s_{(p-1)+(f-1)p} = s_{n-1}$
- thus $n - 2 = (p-1) + t_0 p(\mathrm{mod}\ n)$ for some t_0

A periodic sequence $\mathbf{s}_n$ has $nlc(\mathbf{s}_n) = n - 1$

$\Leftrightarrow$ $\exists\, 1 \le p < n$ such that $s_i = s_{i+p}$ holds iff. $i \in \mathbb{Z}_{n-2}$

$\Rightarrow$ $\gcd(p, n) = 1$

Necessary Conditions (cont.)

Furthermore,

- $\gcd(p, n) = 1 \Rightarrow$ there exists a unique pair $(u, v) \in \mathbb{Z}_p^* \times \mathbb{Z}_n^*$ such that $un - vp = 1$.
- $\mathbb{Z}_{n-1}$ can be partitioned as $\mathbb{Z}_{n-1} = H_1 \cup H_2$ with

$$\begin{aligned} H_1 &= \{(p-1) + tp (\mathrm{mod}\, n) \,:\, t = 0, 1, \cdots, v-1\} \\ H_2 &= \{(p-1) + tp (\mathrm{mod}\, n) \,:\, t = v, \cdots, n-1\} \\ &= \{(p-2) + tp (\mathrm{mod}\ n) \,:\, t = 0, 1, \cdots, n-v-1\} \end{aligned}$$

- $n - 2 \equiv vp - 1 = (p-1) + (v-1)p (\mathrm{mod}\, n)$ in H_1
- $n - 1 \equiv (p-2) + (n-v-1)p (\mathrm{mod}\, n)$ is H_2

$s_i = s_{i+p}$ holds iff. $i \in \mathbb{Z}_{n-2}$ implies

- $s_i = s_{n-2}$ for $i \in H_1$ and $s_i = s_{n-1}$ for $i \in H_2$
- $s_{n-1} \neq s_{n-2}$

Necessary Conditions

A periodic sequence $\mathbf{s}_n$ has $nlc(\mathbf{s}_n) = n - 1$

$\Leftrightarrow$ $\exists\, 1 \leq p < n$ such that $s_i = s_{i+p}$ iff. $i \in \mathbb{Z}_{n-2}$

$\Rightarrow$ $\gcd(p, n) = 1$

- $\exists\, (u, v) \in \mathbb{Z}_p^* \times \mathbb{Z}_n^*$ such that $un - vp = 1$
- partition $\mathbb{Z}_n = H_1 \cup H_2$ with
  $H_1 = \{(p-1) + tp (\mathrm{mod}\, n) \,:\, t \in \mathbb{Z}_v\}$ and
  $H_2 = \{(p-2) + tp (\mathrm{mod}\, n) \,:\, t \in \mathbb{Z}_{n-v}\}$

$\Rightarrow$ $s_i = \begin{cases} s_{n-2} & \text{for } i \in H_1 \\ s_{n-1} & \text{for } i \in H_2 \end{cases}$ and $s_{n-2} \neq s_{n-1}$

- $p = 1$: $\mathbf{s}_n = (\alpha \cdots \alpha\beta)$
- $p = 2$ and n is odd: $\mathbf{s}_n = (\alpha\beta \cdots \alpha\beta\beta) = (\alpha\beta)^{\frac{n-1}{2}}\beta$

Question: What does $\mathbf{s}_n$ look like for other p coprime to n?

A recursive construction

For two integers r_0, r_1 with $r_0 > r_1$ and $\gcd(r_0, r_1) = 1$, applying the Euclidean algorithm on them gives

$$r_{i+1} = r_{i-1} - m_i r_i, \quad i = 1, 2, \cdots, k,$$

with $r_1 > r_2 > \cdots > r_{k+1} = 1$.

Define a class of periodic sequences as follows:

$$\begin{cases} \mathbf{s}_{r_0} &= (\mathbf{s}_{r_1})^{m_1} \mathbf{s}_{r_2}, \\ \mathbf{s}_{r_1} &= (\mathbf{s}_{r_2})^{m_2} \mathbf{s}_{r_3}, \\ &\vdots \\ \mathbf{s}_{r_{k-1}} &= (\mathbf{s}_{r_k})^{m_k} \mathbf{s}_{r_{k+1}} \end{cases}$$

where $\mathbf{a}^t$ denotes the concatenation of t copies of $\mathbf{a}$

The periodic sequences are given by

$$\mathbf{s}_{r_{i-1}} = (\mathbf{s}_{r_i})^{m_i}\mathbf{s}_{r_{i+1}}, \quad i = 0, 1, \cdots, k$$

with $r_{i+1} = r_{i-1} - m_i r_i$.

- $\mathbf{s}_{r_{k-1}}$ is uniquely determined by $\mathbf{s}_{r_k}$ and $\mathbf{s}_{r_{k+1}}$ with m_k
- $\mathbf{s}_{r_{k-2}}$ is uniquely determined by $\mathbf{s}_{r_{k-1}}$ and $\mathbf{s}_{r_k}$ with m_{k-1}
- $\cdots$
- recursively, $\mathbf{s}_{r_0}$ is determined by $\mathbf{s}_{r_k}$, $\mathbf{s}_{r_{k+1}}$ and the integers $m_k, \cdots, m_1$ and $r_k, \cdots, r_1$

The periodic sequences are given by

$$\mathbf{s}_{r_{i-1}} = (\mathbf{s}_{r_i})^{m_i}\mathbf{s}_{r_{i+1}}, \quad i = 0, 1, \cdots, k$$

with $r_{i+1} = r_{i-1} - m_i r_i$.

Theorem 1

If $\mathbf{s}_{r_k} = ((\alpha)^{r_k-1}\beta)$, $\mathbf{s}_{r_{k+1}} = (\alpha)$, then $nlc(\mathbf{s}_{r_{k-1}}) = r_{k-1} - 1$. Moreover,

$$\begin{aligned} nlc(\mathbf{s}_{r_{k-2}}) &= r_{k-2} - 1 \\ &\vdots \\ nlc(\mathbf{s}_{r_1}) &= r_1 - 1 \\ nlc(\mathbf{s}_{r_0}) &= r_0 - 1 \end{aligned}$$

$nlc(\mathbf{s}_n) = n-1 \Leftrightarrow \exists\, 1 \le p < n$ such that $s_i = s_{i+p}$ iff. $i \in \mathbb{Z}_{n-2}$

Proof of Theorem 1.

- $nlc(\mathbf{s}_{r_k}) = r_k - 1$ and $nlc(\mathbf{s}_{r_{k+1}}) = r_{k+1} - 1 = 0$
- for the periodic sequence

$$\mathbf{s}_{r_{k-1}} = (\mathbf{s}_{r_k})^{m_k}\mathbf{s}_{r_{k+1}} = (\alpha^{r_k-1}\beta)^{m_k}\beta$$

  - denote $\mathbf{s} = (\mathbf{s}_{r_{k-1}})^2$
  - $\mathbf{s}[0 : r_{k-1} - 1] = \mathbf{s}_{r_{k-1}} = (\mathbf{s}_{r_k})^{m_k-1}\mathbf{s}_{r_k}\mathbf{s}_{r_{k+1}}$
  - $\mathbf{s}[r_k : r_k + r_{k-1} - 1] = (\mathbf{s}_{r_k})^{m_k-1}\mathbf{s}_{r_{k+1}}\mathbf{s}_{r_k}$
  - $\mathbf{s}_{r_k}\mathbf{s}_{r_{k+1}} = (\alpha^{r_k-1}\beta)\alpha = (\alpha^{r_k-1})\beta\alpha$
  - $\mathbf{s}_{r_{k+1}}\mathbf{s}_{r_k} = \alpha(\alpha^{r_k-1}\beta) = (\alpha^{r_k-1})\alpha\beta$
  - $\mathbf{s}[0 : r_{k-1} - 3] = \mathbf{s}[r_k : r_k + r_{k-1} - 3]$
  - $\mathbf{s}[r_{k-1} - 2 : r_{k-1} - 1] \neq \mathbf{s}[r_k + r_{k-1} - 2 : r_k + r_{k-1} - 1]$
- $nlc(\mathbf{s}_{r_i}) = r_i - 1$ by induction

Main Theorem

A periodic sequence $\mathbf{s}_n$ over $\mathcal{F}$ has $nlc(\mathbf{s}_n) = n - 1$ if and only if it can, up to shift equivalence, be represented as one of the following two forms:

1)
$$\mathbf{s}_n = ((\alpha)^{n-1}\beta) \text{ for } p = 1,$$

2) $\mathbf{s}_n = \mathbf{s}_{r_0} = (\mathbf{s}_{r_1})^{m_1}\mathbf{s}_{r_2}$ for certain integer $r_1 \in \mathbb{Z}_n^*$ with

$$\mathbf{s}_{r_{i-1}} = (\mathbf{s}_{r_i})^{m_i}\mathbf{s}_{r_{i+1}}, \ i = 1, 2, \cdots, k,$$

and

$$\mathbf{s}_{r_k} = ((\alpha)^{r_k-1}\beta), \mathbf{s}_{r_{k+1}} = (\alpha),$$

where the integers m_i, r_{i+1} are derived from $r_{i+1} = r_{i-1} - m_i r_i$ with $r_1 > r_2 > \cdots > r_{k+1} = 1$,

where α, β are any two different elements of $\mathcal{F}$.

Proof of Necessity.

A periodic sequence $\mathbf{s}_n$ has $nlc(\mathbf{s}_n) = n - 1$

$\Leftrightarrow$ $\exists\, 1 \leq p < n$ such that $s_i = s_{i+p}$ iff. $i \in \mathbb{Z}_{n-2}$

$\Rightarrow$ $\gcd(p, n) = 1$

- $\exists\, (u, v) \in \mathbb{Z}_p^* \times \mathbb{Z}_n^*$ such that $un - vp = 1$
- partition $\mathbb{Z}_n = H_1 \cup H_2$ with
  $H_1 = \{(p-1) + tp (\mathrm{mod}\, n) \,:\, t \in \mathbb{Z}_v\}$ and
  $H_2 = \{(p-2) + tp (\mathrm{mod}\, n) \,:\, t \in \mathbb{Z}_{n-v}\}$

$\Rightarrow$ $s_i = \begin{cases} s_{n-2} & \text{for } i \in H_1 \\ s_{n-1} & \text{for } i \in H_2 \end{cases}$ and $s_{n-2} \neq s_{n-1}$

Proof of Necessity.

- $\mathbf{s}_{r_0} = r_0 - 1 \Rightarrow \exists\ r_1 \in \mathbb{Z}_{r_0}^*$ yielding $r_1 > \cdots > r_k > r_{k+1} = 1$
- $\gcd(r_i, r_{i+1}) = 1 \Rightarrow \exists\ u_i, v_i$ s.t. $u_i r_i - v_i r_{i+1} = 1$ for $i = 0, 1, \cdots, k$
- define sets
  $H_{i,1} = \{(r_{i+1} - 1 + t \cdot r_{i+1}) \pmod{r_i} \,|\, t \in \mathbb{Z}_{v_i}\}$,
  $H_{i,2} = \{(r_{i+1} - 2 + t \cdot r_{i+1}) \pmod{r_i} \,|\, t \in \mathbb{Z}_{r_i - v_i}\}$
- $\mathbb{Z}_{r_i} = H_{i,1} \cup H_{i,2}$ and $\mathbf{s}_{r_i}[t] = \begin{cases} \mathbf{s}_{r_i}[r_i - 2] & \text{for } t \in H_{i,1} \\ \mathbf{s}_{r_i}[r_i - 1] & \text{for } t \in H_{i,2} \end{cases}$

Proof of Necessity (cont).

$$\mathbb{Z}_{r_i} = H_{i,1} \cup H_{i,2} \text{ and } \mathbf{s}_{r_i}[t] = \begin{cases} \mathbf{s}_{r_i}[r_i - 2] & \text{for } t \in H_{i,1} \\ \mathbf{s}_{r_i}[r_i - 1] & \text{for } t \in H_{i,2} \end{cases}$$

How to determine $H_{i,1}$ and $H_{i,2}$?

- $H_{i,1} = \bigcup_{t \in H_{i+1,2}} \{x \in \mathbb{Z}_{r_i} \mid x \equiv t (\mathrm{mod}\, r_{i+1})\}$
- $H_{i,2} = \bigcup_{t \in H_{i+1,1}} \{x \in \mathbb{Z}_{r_i} \mid x \equiv t (\mathrm{mod}\, r_{i+1})\}$
- This implies

$$\begin{aligned} H_{k,1}, H_{k,2} &\Rightarrow H_{k-1,1}, H_{k-1,2} \\ &\vdots \\ H_{2,1}, H_{2,2} &\Rightarrow H_{1,1}, H_{1,2} \\ H_{1,1}, H_{1,2} &\Rightarrow H_{0,1}, H_{0,2} \end{aligned}$$

- for $r_k > r_{k+1} = 1$, $\mathbf{s}_{r_k} = (\alpha)^{r_k - 1}\beta$
- $H_{k,1} = \{0, 1, \cdots, r_k - 2\}$ and $H_{k-2} = r_k - 1$

Enumeration

The total number N, up to shift equivalence, of periodic sequence with maximum nonlinear complexity over $\mathcal{F}$ is

$$N = \phi(n)|\mathcal{F}|/2,$$

where $\phi(n)$ is the Euler's totient function.

Examples

n	Sequences	# Seq.
3	001, 011	2
4	0001, 0111	2
5	00001, 00101, 01011, 01111	4
6	000001, 011111	2
7	0000001, 0001001, 0010101, 0101011, 0110111, 0111111	6
8	00000001, 00100101, 01011011, 01111111	4
9	000000001, 000010001, 001010101, 010101011, 011101111, 011111111	6
10	0000000001, 0001001001, 0110110111, 0111111111	4
11	00000000001, 00000100001, 00010001001, 00100100101, 00101010101 01010101011, 01011011011, 01101110111, 01111011111, 01111111111	10
12	000000000001, 001010010101, 010101101011, 011111111111	4
13	0000000000001, 0000001000001, 0000100010001, 0001001001001 0010010100101, 0010101010101, 0101010101011, 0101101011011 0110110110111, 0111011101111, 0111110111111, 0111111111111	12
14	00000000000001, 00001000010001, 00100100100101 01011011011011, 01110111101111, 01111111111111	6
15	000000000000001, 000000010000001, 000100010001001, 001010101010101 010101010101011, 011011101110111, 011111101111111, 011111111111111	8
16	0000000000000001, 0000010000100001, 0001001001001001, 0010101001010101 0101010110101011, 0110110110110111, 0111101111011111, 0111111111111111	8

Randomness Analysis

For all binary periodic sequences $\mathbf{s}_n$ with $nlc(\mathbf{s}_n) = n - 1$,

- **balancedness:** $\mathbf{s}_n$ is (nearly) balanced only if $\mathbf{s}_n = (01)^{\frac{n-1}{2}}0$ for odd; others are far from being balanced
- **scalability:** there exist many subsequence with small nonlinear complexity
- **k-th error complexity:** changing a few bits dramatically decrease the nonlinear complexity

The randomness properties of such sequences are not sounding.

Thanks for your attention!