Shift Register Sequences

Mark Goresky
Linear recursion:

\[ a_n = q_1 a_{n-1} + q_2 a_{n-2} + \cdots + q_r a_{n-r} \]

or

\[ -a_n + q_1 a_{n-1} + q_2 a_{n-2} + \cdots + q_r a_{n-r} = 0 \]

Connection Polynomial:

\[ q(x) = -1 + q_1 x + q_2 x^2 + \cdots + q_r x^r \]
Theorem. (Golomb, 1955)
Also Dickson, 1903

The output sequence is

$$a_0 + a_1 x + a_2 x^2 + \cdots = \frac{p(x)}{q(x)} \in \mathbb{F}_2[[x]]$$

where

$$q(x) = -1 + q_1 x + q_2 x^2 + \cdots + q_r x^r$$

depends on the taps and

$$p(x) = \sum_{n=0}^{r-1} \left( \sum_{i=0}^{n} q_i a_{n-i} \right) x^n$$

depends on the initial loading.
Applications

- smallest shift register to generate a given sequence: Berlekamp-Massey algorithm = continued fraction expansion in $\mathbb{F}_2[[x]]$.

- sum of several LFSR sequences

- Same analysis works for LFSR over any commutative ring

Fibonacci numbers:

\[
1 + 1x + 2x^2 + 3x^3 + 5x^4 + \cdots = \frac{-1}{-1 + x + x^2}
\]
Choose a root $\alpha \in \mathbb{F}_{2^n}$, $q(\alpha^{-1}) = 0$

Choose a linear function $T : \mathbb{F}_{2^n} \rightarrow \mathbb{F}_2$.

The sequence of values

$$
\begin{array}{ccccc}
T(\alpha^{r-1}) & T(\alpha^{r-2}) & \cdots & T(\alpha) & T(1)
\end{array}
$$

satisfies the linear recursion.
Proof.

\[ 0 = q_0 \alpha^0 + q_1 \alpha^{-1} + \cdots + q_r \alpha^{-r} \]

\[ 0 = q_0 \alpha^r + q_1 \alpha^{r-1} + \cdots + q_r \alpha^0 \]

\[ T(\alpha^r) = q_1 T(\alpha^{r-1}) + \cdots + q_r T(1) \]

Similarly for any \( A \in \mathbb{F}_{2^n} \) the sequence \( T(A\alpha^i) \) satisfies the linear recursion.

Example: if \( q(x) \) is irreducible of degree \( r \),

\[ \mathbb{F}_{2^r} \cong (\mathbb{F}_2)^r \]

\[ A \mapsto (T(A\alpha^{r-1}), \ldots, T(A\alpha), T(A)) \]
Recall $\alpha \in \mathbb{F}_{2^r}$ is primitive if
\[
\mathbb{F}_{2^r} = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{r-2}\}
\]

**Corollary.** If $\alpha \in \mathbb{F}_{2^r}$ is a primitive element then the shift register will cycle through all possible nonzero states before it repeats. So $q(x)$ should be a primitive polynomial.

The result is an *m-sequence*
e.g. $q(x) = -1 + x^2 + x^3$

\[a = 00101110010111\cdots\]
Spread spectrum communication

Signature sequence: 0010111
Send 1: 0010111  Send 0: 1101000

00101110010111110100000101111101000
  0010111

The correlator counts

  #(agreements) - #(disagreements)

in each window
Given periodic sequences of period $T$
a = $(a_0, a_1, \ldots)$ and b = $(b_0, b_1, \ldots)$
Their cross correlation is

$$C_{a,b}(\tau) = \sum_{i=0}^{r-1} (-1)^{ai}(-1)^{bi+\tau}$$

Autocorrelation $A_a(\tau) = C_{a,a}(\tau)$

Idea: Each user has a different signature sequence.
These sequences have low cross-correlations. The receiver (correlator) knows your sequence. Received signal will look like this:
**Theorem.** If \( a = (a_0, a_1, \cdots) \) is an m-sequence of period \( P = 2^r - 1 \) then

\[
A_a(\tau) = \begin{cases} 
-1 & \tau \neq 0 \\
\frac{P}{\tau} & \tau = 0
\end{cases}
\]

**Proof.**

\[
\sum_{i=0}^{P-1} (-1)^T(\alpha^i) (-1)^T(\alpha^{i+\tau}) = \sum_{i=0}^{P-1} (-1)^T(\alpha^{i+\alpha^i}) = \sum_{i=0}^{P-1} (-1)^T(\alpha^i(1+\alpha^\tau)) = \sum_{x\in \mathbb{F}_{2^r}} (-1)^T(Ax) - 1 = 0 - 1
\]

where \( A = (1 + \alpha^\tau) \).
Problem. Compute the cross correlation of any two distinct m-sequences:

\[ C_{a,b} = \sum_{i=0}^{2^r-2} (-1)^T(\alpha^i)(-1)^T(\beta^i+\tau) \]

where \( \alpha, \beta \in \mathbb{F}_{2^r} \) are primitive. Say \( \beta = \alpha^k \).

\[ = \sum_{i=0}^{2^r-2} (-1)^T(\alpha^i+\alpha^{ki}\alpha^\tau) \]

\[ = \sum_{x \in \mathbb{F}_{2^r}^*} (-1)^T(x+Ax^k) \]

where \( A = \alpha^\tau \).

Trigonometric sum (Gauss, Hardy, Littlewood, + about 1000 other number theorists)

Lachaud and Wolfman: For \( T : \mathbb{F}_{2^r} \to \mathbb{F}_2 \) use the trace

\[ Tr(x) = x + x^2 + x^4 + \cdots + x^{2^r-1} \]
Lemma. $Tr(x) = 0$ if and only if

$$x = h(y) = y^2 - y$$

for some $y \in \mathbb{F}_{2^r}$.

Proof. If $h(y) = y(y - 1) = 0$ then $y \in \mathbb{F}_2$. And

$$Tr(h(y)) = -y + y^2 - y^2 + y^4 - \cdots + y^{2^r} = 0.$$ 

So this sequence is exact:

$$0 \rightarrow \mathbb{F}_2 \rightarrow \mathbb{F}_{2^r} \xrightarrow{h} \mathbb{F}_{2^r} \xrightarrow{Tr} \mathbb{F}_2 \rightarrow 0$$
Now consider the equation

\[ y^2 - y = Ax + x^k. \]

Let \( E \) denote the set of solutions \((x, y)\)
Whenever \( Tr(Ax + x^k) = 0 \) we get two points in \( E \). So

\[ \#E/2 = \# \left\{ x \mid Tr(Ax + x^k) = 0 \right\} \]

So

\[ A(\tau) = \#E/2 - (2^r - \#E/2) \]
\[ = \#E - 2^r (\pm 1) \]

Case \( k = -1 \) (the reverse sequence):

\[ y^2 - y = Ax + x^{-1} \]

This is an elliptic curve:

\[ xy^2 - xy = Ax^2 + 1 \]

Problem was completely solved by Honda-Tate.
Feedback with carry
joint with Andrew Klapper

\[ M_{n-1} \text{ div } 2 \text{ mod } 2 \]

\[ a_{n-1} \quad a_{n-2} \quad \cdots \quad a_{n-r} \]

\[ q_1 \quad q_2 \quad \cdots \quad q_r \]
Linear recursion with carry:

\[ 2M_n + a_n = M_{n-1} + \sum_{i=1}^{r} q_i a_{n-i} \]

Connection integer

\[ q = -1 + q_1 2 + q_2 2^2 + \cdots + q_r 2^r \]

**Theorem.** The output sequence is

\[ a_0 2^0 + a_1 2^1 + a_2 2^2 + \cdots = \frac{p}{q} \in \mathbb{Z}_2 \]

where

\[ p = \sum_{n=0}^{r-1} \sum_{i=0}^{n} q_i a_{n-i} 2^n - M_n 2^n \]
Applications

• smallest FCSR to generate a given sequence continued fraction expansion in \( \mathbb{Z}_2 \) may fail to converge! Instead, use approximation theory in 2-adic numbers (Mahler, de Weger).

• sum-with-carry of several FCSR sequences (we broke the summation combiner cipher)

• arithmetic cross-correlation analysis.

• Similar analysis for FCSR sequences over other \( p \)-adic fields.

• Other related architectures.
Exponential representation:

\[ a_i = A2^{-i} \pmod{q} \pmod{2} \]

\[ \mathbb{Z}/(q) \xrightarrow{\pmod{2}} \mathbb{Z}/(2) \]

compare LFSR case:

\[ a_i = T(A\beta^{-i}) \]

or, better: \[ a_i = Ax^{-i} \pmod{q} \pmod{x} \].

\[ \mathbb{F}_2[x]/(q) \xrightarrow{\pmod{x}} \mathbb{F}_2 \]

Maximal length sequences:

**Theorem.** If 2 is primitive modulo \( q \) the FCSR passes through all (nonzero) periodic states \( (q - 1) \) of them) before it repeats.
Algebraic shift registers

A setting that includes both LFSR and FCSR and others, e.g. the 2-FCSR involves ramified extensions of the 2-adic numbers:
Galois 2-FCSR.