3.2 Synthesis of LFSRs

In this section we treat the problem of how to find an LFSR of shortest length that generates a given finite bit sequence. In section 2.6 we described a method of finding linear relations for sequence elements from a quite general generator. This might result in an LFSR, but anyway the linear relations might change from step to step and there appears no easy way of getting an optimal LFSR.

Here we follow another approach that solves our problem in a surprisingly easy way: the BM-algorithm, named after Berlekamp (1968 in a different context) and Massey (1969).

We don't use any special properties of the field $\mathbb{F}_2$, so we work over an arbitrary field K. Our goal is to construct a homogeneous linear generator of the smallest possible recursion depth l that generates a given finite sequence $u \in K^N$.

We consider a homogeneous linear generator whose recursion formula is

$$u_k = a_1 u_{k-1} + \cdots + a_l u_{k-l} \quad \text{for } k = l, \ldots, N-1. \tag{1}$$

Its coefficient vector is $(a_1, \ldots, a_l) \in K^l$. The polynomial

$$\varphi = 1 - a_1 T - \cdots - a_l T^l \in K[T]$$

is called **feedback polynomial**.

Note Don't confuse this polynomial with the feedback function

$$s(u_0, \ldots, u_{l-1}) = a_1 u_{l-1} + \cdots + a_l u_0.$$

The feedback polynomial is the reciprocal polynomial of the characteristic polynomial

$$\chi = \mathrm{Det}(T \cdot 1 - A) = T^l - a_1 T^{l-1} - \cdots - a_l$$

of the companion matrix

$$A = \begin{pmatrix} 0 & 1 & \ldots & 0 \\ & \ddots & \ddots & \\ & & & 1 \\ a_l & a_{l-1} & \ldots & a_1 \end{pmatrix}.$$

These two polynomials are related by the formula

$$\varphi = T^l \cdot \chi(\frac{1}{T}).$$

Lemma 12 *Let the sequence $u = (u_0, \ldots, u_{n-1}) \in K^n$ be a segment of the output of the linear generator (1), but not the sequence $\hat{u} = (u_0, \ldots, u_n) \in K^{n+1}$. Then every homogeneous linear generator of length $m \geq 1$ that generates $\hat{u}$ has $m \geq n + 1 - l$.*

Proof. **Case 1:** $l \geq n$. Then obviously $l + m \geq n + 1$.

Case 2: $l \leq n - 1$. Assume $m \leq n - l$. We have

$$u_j = a_1 u_{j-1} + \cdots + a_l u_{j-l} \quad \text{for } l \leq j \leq n-1.$$

Let $(b_1, \ldots, b_m)$ be the coefficient vector of a homogeneous linear generator that produces $\hat{u}$. Then

$$u_j = b_1 u_{j-1} + \cdots + b_m u_{j-m} \quad \text{for } m \leq j \leq n.$$

We deduce

$$\begin{aligned} u_n &\neq a_1 u_{n-1} + \cdots + a_l u_{n-l} \\ &= \sum_{i=1}^{l} a_i \cdot \underbrace{\sum_{k=1}^{m} b_k u_{n-i-k}}_{u_{n-i}} \quad [\text{since } n - l \geq m] \\ &= \sum_{k=1}^{m} b_k \cdot \underbrace{\sum_{i=1}^{l} a_i u_{n-k-i}}_{u_{n-k}} = u_n, \end{aligned}$$

contradiction. ◇

Consider a sequence $u \in K^N$. For $0 \leq n \leq N$ let $\lambda_n(u) = \lambda_n$ be the smallest recursion depth for which a homogeneous linear generator exists that produces $(u_0, \ldots, u_{n-1})$.

Lemma 13 *For every sequence $u \in K^N$ we have:*

(i) $\lambda_{n+1} \geq \lambda_n$ *for all* n.

(ii) *There is a homogeneous linear generator of recursion depth λ_n that produces $(u_0, \ldots, u_n)$ if and only if $\lambda_{n+1} = \lambda_n$.*

(iii) *If there is no such generator, then*

$$\lambda_{n+1} \geq n + 1 - \lambda_n.$$

Proof. (i) Every generator that produces $(u_0, \ldots, u_n)$ a forteriori produces $(u_0, \ldots, u_{n-1})$.

(ii) follows from (i).

(iii) The precondition of Lemma 12 is true for every generator of $(u_0, \ldots, u_{n-1})$. ◇

Proposition 10 [Massey] *Let* $u \in K^N$ *and* $0 \le n \le N-1$. *Let* $\lambda_{n+1}(u) \neq \lambda_n(u)$. *Then*

$$\lambda_n(u) \le \frac{n}{2} \quad \textit{and} \quad \lambda_{n+1}(u) = n + 1 - \lambda_n(u).$$

Thus the linear complexity may jump only if λ_n (we often omit u in the notation) is "below the diagonal," and then it jumps to the symmetric position "above the diagonal." An illustration is in Figure 3.2.

Proof. First we consider the easy case $\lambda_n = 0$: Here $u_0 = \ldots = u_{n-1} = 0$. If $u_n = 0$, then $\lambda_{n+1} = \lambda_n = 0$, leaving nothing to prove. Otherwise $u_n \neq 0$, and then $\lambda_{n+1} = n + 1 = n + 1 - \lambda_n$ by remark 5 in 3.1.

In general the first statement follows from the second one: We have $\lambda_n < \lambda_{n+1}$, hence $2\lambda_n < \lambda_n + \lambda_{n+1} = n + 1$.

Now we prove the second statement by induction on n. In the case $n = 0$ we have $\lambda_0 = 0$—this case is already settled.

Now let $n \ge 1$. We may assume $l := \lambda_n \ge 1$. Let

$$u_j = a_1 u_{j-1} + \cdots + a_l u_{j-l} \quad \text{for } j = l, \ldots, n-1;$$

hence the feedback polynomial is

$$\varphi := 1 - a_1 T - \cdots - a_l T^l \in K[T].$$

Let the "n-th discrepancy" be defined as

$$d_n := u_n - a_1 u_{n-1} - \cdots - a_l u_{n-l}.$$

If $d_n = 0$, then the generator outputs u_n as the next element, and there is nothing to prove. Otherwise let $d_n \neq 0$. Let r be the length of the segment before the last increase of linear complexity, thus

$$t := \lambda_r < l, \quad \lambda_{r+1} = l.$$

By induction $l = r + 1 - t$. We have a relation

$$u_j = b_1 u_{j-1} + \cdots + b_t u_{j-t} \quad \text{for } j = t, \ldots, r-1,$$

the corresponding feedback polynomial is

$$\psi := 1 - b_1 T - \cdots - b_t T^t \in K[T],$$

and the corresponding r-th discrepancy,

$$d_r := u_r - b_1 u_{r-1} - \cdots - b_t u_{r-t} \neq 0.$$

In the case $t = 0$ we have $\psi = 1$ and $d_r = u_r$. Now we form the polynomial

$$\eta := \varphi - \frac{d_n}{d_r} \cdot T^{n-r} \cdot \psi = 1 - c_1 T - \cdots - c_m T^m \in K[T]$$

with $m = \deg \eta$. What is the output of the corresponding homogeneous linear generator? We have

$$\begin{aligned} u_j - \sum_{i=1}^{m} c_i u_{j-i} &= u_j - \sum_{i=1}^{l} a_i u_{j-i} - \frac{d_n}{d_r} \cdot \left[u_{j-n+r} - \sum_{i=1}^{t} b_i u_{j-n+r-i} \right] \\ &= 0 \quad \text{for } j = m, \ldots, n; \end{aligned}$$

for $j = m, \ldots, n-1$ this follows directly, for $j = n$ via the intermediate result $d_n - [d_n/d_r] \cdot d_r$. Hence the output is $(u_0, \ldots, u_n)$. Now we have

$$\lambda_{n+1} \leq m \leq \max\{l, n - r + t\} = \max\{l, n + 1 - l\}.$$

Since linear complexity grows monotonically we conclude $m > l$, and by Lemma 12 we get $m \geq n + 1 - l$. Hence $m = n + 1 - l$ and $\lambda_{n+1} = m$. This proves the proposition. ◇

Corollary 1 *If $d_n \neq 0$ and $\lambda_n \leq \frac{n}{2}$, then*

$$\lambda_{n+1} = n + 1 - \lambda_n > \lambda_n.$$

Proof. By Lemma 12 we have $\lambda_{n+1} \geq n + 1 - \lambda_n$, thus $\lambda_{n+1} \geq \frac{n}{2} + 1 > \lambda_n$. By Proposition 10 we conclude $\lambda_{n+1} = n + 1 - \lambda_n$. ◇

During the successive construction of a linear generator in the proof of the proposition, in each iteration step one of two cases occurs:

- $d_n = 0$: then $\lambda_{n+1} = \lambda_n$.
- $d_n \neq 0$: then
  - $\lambda_{n+1} = \lambda_n$ if $\lambda_n > \frac{n}{2}$,
  - $\lambda_{n+1} = n + 1 - \lambda_n$ if $\lambda_n \leq \frac{n}{2}$.

In particular we always have:

- If $\lambda_n > \frac{n}{2}$, then $\lambda_{n+1} = \lambda_n$.
- If $\lambda_n \leq \frac{n}{2}$, then $\lambda_{n+1} = \lambda_n$ or $\lambda_{n+1} = n + 1 - \lambda_n$.

By the way we found an alternative method of predicting LFSRs:

Corollary 2 *If $u \in \mathbb{F}_2^N$ is generated by an LFSR of length $\leq l$, then one such LFSR may be determined from $u_0, \ldots, u_{2l-1}$.*

Proof. Assume n is the first index $\geq 2l$ such that $d_n \neq 0$. Then $\lambda_n \leq l \leq \frac{n}{2}$, thus $\lambda_{n+1} = n + 1 - \lambda_n \geq l + 1$, contradiction. ◇