Master 2 Mathematics and Computer Science
Symbolic Dynamics. Lecture 5

Marie-Pierre Béal

University Gustave Eiffel
Laboratoire d'informatique Gaspard-Monge UMR 8049

Overview

- Substitutions, substitution shifts.
- Primitive substitutions, linear recurrence, minimality, and uniform recurrence, return words, block complexity.

Substitutions, substitution shifts

Substitutions

Given two finite alphabets A, B, a *substitution* $\sigma\colon A^* \to B^*$ is a monoid morphism from A^* to B^*.
Thus $\sigma(\varepsilon) = \varepsilon$ and $\sigma(uv) = \sigma(u)\sigma(v)$ for every $u, v \in A^*$.
The substitution is determined by the images $\sigma(a)$ of the letters $a \in A$.
Indeed, we have $\sigma(a_0 a_1 \cdots a_{n-1}) = \sigma(a_0)\sigma(a_1)\cdots\sigma(a_{n-1})$ for $a_i \in A$ and $n \geq 0$.

Extension to infinite sequences

Such a substitution σ extends to a partial map from $A^{\mathbb{N}}$ to $B^{\mathbb{N}}$.
It is defined by $\sigma(x_0x_1\cdots) = \sigma(x_0)\sigma(x_1)\cdots$ if the righthand side is infinite, that is, if $\sigma(x_0)\sigma(x_1)\cdots \in B^{\mathbb{N}}$.
The right-hand side is possibly a finite word (if $\sigma(x_n) = \varepsilon$ for all n sufficiently large). In this case, $\sigma(x_0x_1\cdots)$ is undefined.
A substitution $\sigma\colon A^* \to B^*$ also extends to a partial map from $A^{\mathbb{Z}}$ to $B^{\mathbb{Z}}$ by

$$\sigma(\cdots x_{-1} \cdot x_0x_1\cdots) = \cdots\sigma(x_{-1})\cdot\sigma(x_0)\sigma(x_1)\cdots \tag{1}$$

The result is a two-sided infinite sequence provided $\sigma(x_n) \neq \varepsilon$ for an infinite number of negative and positive indices n.

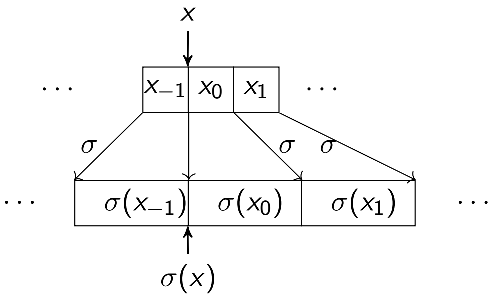


Figure: The two-sided sequence $\sigma(x)$.

A substitution $\sigma\colon A^* \to B^*$ is *non-erasing* if $\sigma(a)$ is nonempty for every $a \in A$.
A substitution $\sigma\colon A^* \to B^*$ has *constant length* k (or is *uniform*) if $|\sigma(a)| = k$ for every $a \in A$. A substitution of constant length $k \geq 1$ is non-erasing.

Lettre coding

A substitution $\sigma\colon A^* \to B^*$ is a *letter coding* if it is of constant length 1. Letter codings, also called *letter-to-letter* substitutions, play an important role in the definition of morphic sequences (see later).
They are the substitutions preserving length, meaning that $|\sigma(w)| = |w|$ for every $w \in A^*$. They also correspond to 1-block sliding block codes.

For a substitution $\sigma\colon A^* \to B^*$, we define

$$|\sigma| = \max_{a\in A} |\sigma(a)|, \quad \text{and} \quad \langle\sigma\rangle = \min_{a\in A} |\sigma(a)| \tag{2}$$

Composition matrix

Let $\sigma\colon A^* \to B^*$ be a substitution. The *composition matrix* of σ is the $(B \times A)$-matrix $M = M(\sigma)$ defined by

$$M_{b,a} = |\sigma(a)|_b,$$

where $|\sigma(a)|_b$ is the number of occurrences of the letter b in the word $\sigma(a)$. Thus, the composition vector of each $\sigma(a)$ is the column of index a of the matrix $M(\sigma)$.
If $\sigma\colon B^* \to C^*$ and $\tau\colon A^* \to B^*$ are substitutions, we have

$$M(\sigma \circ \tau) = M(\sigma)M(\tau).$$

Indeed, for every $a \in A$ and $c \in C$, we have

$$M(\sigma \circ \tau)_{c,a} = |\sigma \circ \tau(a)|_c = \sum_{b \in B} |\sigma(b)|_c |\tau(a)|_b = (M(\sigma)M(\tau))_{c,a}.$$

The transpose of $M(\sigma)$ is called the *adjacency matrix*.

Composition matrix

For a word $w \in A^*$, we denote by $\ell(w)$ the column vector $(|w|_a)_{a \in A}$, called the *composition vector* of w.
The composition matrix satisfies, for every $w \in A^*$, the equation

$$\ell(\sigma(w)) = M(\sigma)\ell(w). \qquad (3)$$

Example

The composition matrix of $\sigma : a \mapsto ab, b \mapsto aa$ is

$$M(\sigma) = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}.$$

Iteration of a substitution

A substitution $\sigma \colon A^* \to A^*$ from A^* into itself is an endomorphism of the monoid A^*. It can be iterated, that is, its powers σ^n for $n \geq 1$ are also substitutions.

Let $\sigma \colon A^* \to A^*$ be an iterable substitution. The *language* of σ, denoted by $\mathcal{L}(\sigma)$ is the set of words occurring as blocks in the words $\sigma^n(a)$ for some $n \geq 0$ and some $a \in A$.
It follows from the definition that

$$\sigma(\mathcal{L}(\sigma)) \subseteq \mathcal{L}(\sigma). \tag{4}$$

The language $\mathcal{L}(\sigma)$ is decidable (exercise).

Let $\sigma \colon A^* \to A^*$ be an iterable substitution.
The *substitution shift* defined by σ is the shift space $\mathsf{X}(\sigma)$ consisting of all $x \in A^{\mathbb{Z}}$ whose finite blocks belong to $\mathcal{L}(\sigma)$.

Show that it is a shift space.

Since $\sigma(\mathcal{L}(\sigma)) \subseteq \mathcal{L}(\sigma)$ by (4), we have also

$$\sigma(\mathsf{X}(\sigma)) \subseteq \mathsf{X}(\sigma). \tag{5}$$

Example

The substitution $\sigma : a \mapsto ab, b \mapsto a$ is the *Fibonacci substitution*. The shift $X = \mathsf{X}(\sigma)$ is the *Fibonacci shift*.

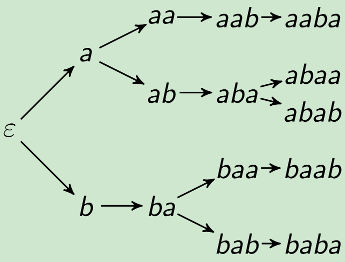


Figure: The words of $\mathcal{B}(X)$ for the Fibonacci shift.

Example: Thue-Morse

Example

The substitution $\sigma: a \mapsto ab, b \mapsto ba$ is the *Thue-Morse substitution*.
The shift $X = \mathrm{X}(\sigma)$ is the *Thue-Morse shift*.

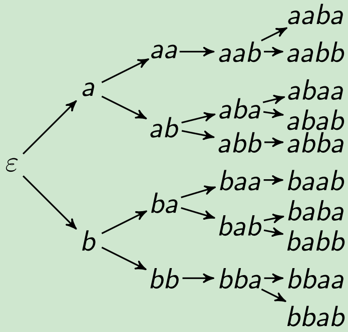


Figure: The words of $\mathcal{B}(X)$ for the Thue-Morse shift.

Note that $\mathcal{B}(\mathsf{X}(\sigma)) \subseteq \mathcal{L}(\sigma)$, but the converse inclusion may not hold, as shown in the example below.

Example

Consider the substitution $\sigma\colon a \mapsto ab, b \mapsto b$. We have $\mathcal{L}(\sigma) = ab^* \cup b^*$ but $\mathsf{X}(\sigma) = b^\infty$, and thus $\mathcal{B}(\mathsf{X}(\sigma)) = b^*$.

Erasable and growing letters

Let $\sigma \colon A^* \to A^*$ be an iterable substitution. A letter $a \in A$ is *erasable* if $\sigma^n(a) = \varepsilon$ for some $n \geq 1$.
A word is *erasable* if it is formed of erasable letters.

A word $w \in A^*$ is *growing* for σ if the sequence $(|\sigma^n(w)|)_n$ is unbounded.
A word is growing if and only if at least one of its letters is growing.
The substitution σ itself is said to be *growing* if all letters are growing.
We have the following property of growing letters.

Proposition

If $a \in A$ is growing for σ, then for every $r \geq 0$, $\sigma^{r\,\mathrm{Card}(A)}(a)$ contains at least $r+1$ non-erasable letters. In particular, $\lim_{n \to +\infty} |\sigma^n(a)| = +\infty$.

Lemma growing

Proof.

Set $k = \mathrm{Card}(A)$.

- Assume first that $\sigma^k(a)$ contains only one non-erasable letter. Then this letter has to be growing.
  Next, by the pigeonhole principle, there are i, p with $i + p \leq k$ and $p \geq 1$ such that $\sigma^i(a) = ubv$ and $\sigma^p(b) = rbs$ with u, v, r, s erasable and b a growing letter.
  Since r, s are erasable, $\sigma^k(r) = \varepsilon$ and $\sigma^k(s) = \varepsilon$.
  Set $w = \sigma^{kp}(b) = \sigma^{(k-1)p}(r) \cdots \sigma^p(r) rbs \sigma^p(s) \cdots \sigma^{(k-1)p}(s)$.
  Then $\sigma^p(w) = w$, a contradiction with the fact that b is growing. This proves the statement for $r = 1$.
- Assume that $\sigma^{rk}(a)$ contains $s \geq r + 1$ non-erasable letters $a_1, \ldots, a_s$. One of them, say a_i, must be growing. Then each of the $\sigma(a_1), \ldots, \sigma(a_s)$ contains a non-erasing letter and $\sigma^k(a_i)$ contains at least two due (case $r = 1$). Therefore, $\sigma^{(r+1)k}(a)$ contains at least $r + 2$ non-erasing letters.

The graph $G(\sigma)$

We associate with an iterable substitution $\sigma\colon A^* \to A^*$ the graph $G(\sigma)$ having A as the set of vertices and $|\sigma(a)|_b$ edges from a to b. The adjacency matrix of $G(\sigma)$ is the adjacency matrix of σ.

Example

Let $\sigma\colon a \mapsto ab, b \mapsto a$ be the Fibonacci substitution.

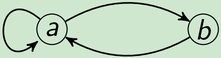


Figure: The graph $G(\sigma)$.

The adjacency matrix of σ is

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Primitive substitutions, linear recurrence, minimality, and uniform recurrence, return words, block complexity.

Primitive substitutions

An iterable substitution $\sigma \colon A^* \to A^*$ is *primitive* if there is an integer $n \geq 1$ such that for every $a, b \in A$ one has $|\sigma^n(a)|_b \geq 1$.

For a primitive substitution σ, except the trivial case $A = \{a\}$ and $\sigma(a) = a$, every letter is growing and $\mathcal{L}(\sigma) = \mathcal{B}(\mathrm{X}(\sigma))$ (exercise).

A substitution shift $X = \mathrm{X}(\sigma)$ is *primitive* if σ is primitive, and not the identity on a one-letter alphabet.

Exercise

Show that $\mathcal{L}(\sigma) = \mathcal{B}(\mathsf{X}(\sigma))$ if and only if $\mathcal{L}(\sigma)$ is extendable, *i.e.* if for each $u \in \mathcal{L}(\sigma)$, there are letters a, b such that $aub \in \mathcal{L}(\sigma)$.

Minimal shift spaces

A shift space X is *minimal* if it is nonempty and if, for every subshift $Y \subseteq X$, one has $Y = \emptyset$ or $Y = X$.

Equivalently, X is minimal if and only if the closure of the orbit $\mathcal{O}(x) = \{S^n(x) \mid n \in \mathbb{Z}\}$ of x is equal to X, for every $x \in X$.

A shift space is minimal if and only if the closure $\mathcal{O}^+(x) = \{S^n(x) \mid n \in \mathbb{N}\}$ of x is equal to X, for every $x \in X$.

Indeed, if X is minimal and Y equal to the closure of $\mathcal{O}^+(x)$, then $Z = \cap_{n \geq 0} S^n(Y)$ is nonempty shift contained in X, thus equal to X. (It is nonempty by compacity as a decreasing sequence of nonempty compact sets).

Let X be a shift space. Given a word $u \in \mathcal{B}(X)$, a *return word* to u in X is a nonempty word w such that $wu \in \mathcal{B}(X)$ and wu has exactly two occurrences of u: one as a prefix and one as a suffix.

By convention, a return word to the empty word is a letter. The set of return words to u in X is denoted by $\mathcal{R}_X(u)$.

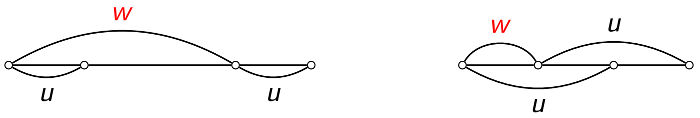
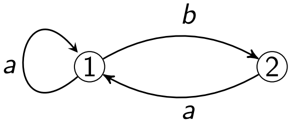


Figure: Return word to u.

The set of return words to u is a *suffix code*, that is, a set S of nonempty words such that no element of S is a proper suffix of another one.

Example

The set of return words to b in the golden mean shift X is $\mathcal{R}_X(b) = ba^+$.



Uniform recurrence

A nonempty shift space X is *recurrent* if it is irreducible, that is, for every $u, v \in \mathcal{B}(X)$ there is a block $w \in \mathcal{B}(X)$ such that $uwv \in \mathcal{B}(X)$.

A nonempty shift space X is *uniformly recurrent* if for every $w \in \mathcal{B}(X)$ there is an integer $n \geq 1$ such that w occurs in every word of $\mathcal{B}_n(X)$.

As an equivalent definition, a shift space X is uniformly recurrent if for every $n \geq 1$ there is an integer $N = R_X(n)$ such that every word of $\mathcal{B}_n(X)$ occurs in every word of $\mathcal{B}_N(X)$. The function R_X is called the *recurrence function* of X.

Remark: Uniform recurrence implies recurrence

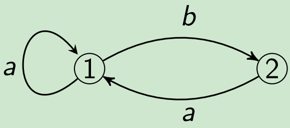
Uniform recurrence implies recurrence.

Indeed, let $u, v \in \mathcal{B}(X)$ and $n \geq 1$ such that u and v occur in every word of $\mathcal{B}_n(X)$.
Then every word w in $\mathcal{B}_{2n}(X)$ contains a block uzv for some block z, since u appears in the first half of w and v in the second half.

Example

Example

The golden mean shift X is recurrent but not uniformly recurrent since b is in $\mathcal{B}(X)$ although b does not occur in any $a^n \in \mathcal{B}(X)$.



Minimality and uniform recurrence

Proposition

A shift space is minimal if and only if it is uniformly recurrent.

Proof.

Assume first that X is a minimal shift space and consider $u \in \mathcal{B}(X)$. Since X is minimal, the forward orbit $\mathcal{O}^+(x) = \{S^n(x) \mid n \geq 0\}$ of every $x \in X$ is dense, and thus the integer $n(x) = \min\{n > 0 \mid S^n x \in [u]_X\}$ exists.

The map $x \mapsto n(x)$ is continuous since the set of x such that $n(x) = n$ is the open set $S^{-n}([u]_X) \setminus \cup_{i=1}^{n-1} S^{-i}([u]_X)$. Since the map $x \mapsto n(x)$ is continuous on a compact space, the integers $n(x)$ are bounded. Then u occurs in every word $w \in \mathcal{B}(X)$ of length $|u| + \max n(x)$. Thus, X is uniformly recurrent.

Conversely, if X is uniformly recurrent, the orbit of every $x \in X$ is dense, and thus X is minimal. □

Example

The golden mean shift is not minimal since it contains the one-point set $\{a^\infty\}$ which is closed and shift-invariant.

Example

The periodic shift generated by $(abc)^\infty$ is minimal.

We define $u^\infty = \cdots uu \cdot uuu \cdots$

Primitive substitution shifts are minimal

Proposition

Let $\sigma\colon A^* \to A^*$ be a substitution distinct from the identity on a one-letter alphabet. If σ is primitive, then it is growing, and $\mathsf{X}(\sigma)$ is minimal. The converse is true if, additionally, every letter is in $\mathcal{B}(X)$.

Proof.

Let $\sigma\colon A^* \to A^*$ be primitive. Since the trivial case $A = \{a\}$ and $\sigma(a) = a$ is excluded, we have $\mathcal{B}(\mathsf{X}(\sigma)) = \mathcal{L}(\sigma)$.
Let $n \geq 1$ be such that every $b \in A$ occurs in every $\sigma^n(a)$ for $a \in A$. □

Primitive substitution shifts are minimal

Proof.

For $u \in \mathcal{L}(\sigma)$, let $m \geq 1$ and $b \in A$ be such that u occurs in $\sigma^m(b)$.
Then u occurs in every $\sigma^{n+m}(a)$.
Let v be a block of $X(\sigma)$ of length $2|\sigma|^{n+m}$.
Then v is a block of eveny $\sigma^{n+m+p}(c)$ for $c \in A$ and $p \geq 0$ large enough.
Thus, v is a block of some $\sigma^{n+m}(z)$, with $z \in A^*$.
Since the size of v is larger than or equal to the size of $\sigma^{n+m}(ab)$, for any letters a, b, it contains $\sigma^{n+m}(a)$, for some letter a, and thus it contains u as a block.
This shows that $\mathrm{X}(\sigma)$ is uniformly recurrent, and thus minimal. □

.

Primitive substitution shifts are minimal

Proof.

Conversely, if $X = \mathrm{X}(\sigma)$ is minimal, and every letter is in $\mathcal{B}(X)$ and there is an $n \geq 1$ such that every letter appears in every word of $\mathcal{B}_n(X)$.
Since σ is growing, there is m such that $\langle \sigma^m \rangle \geq n$. Then every letter $b \in A$ occurs in every $\sigma^m(a)$ with $a \in A$. □

.

Examples

Example

The Fibonacci substitution $\sigma : a \mapsto ab, b \mapsto a$ is primitive.
According to the proposition, the Fibonacci shift $X(\sigma)$ is minimal.

Example

The Thue-Morse substitution $\sigma : a \mapsto ab, b \mapsto ba$, is primitive.
Accordingly to the proposition, the Thue-Morse shift $X(\sigma)$ is minimal.

Prolongable

A substitution $\sigma\colon A^* \to A^*$ is *prolongable* (or *right prolongable*) on $u \in A^+$ if $\sigma(u)$ begins with u and u is growing.

In this case, there is a unique right-infinite sequence, denoted $\sigma^\omega(u)$ such that each $\sigma^n(u)$ is a prefix of $\sigma^\omega(u)$.

One has, of course $\sigma^\omega(u) = \lim_{n\to\infty} \sigma^n(u)$.

Note also that $\sigma^\omega(u)$ is a right-infinite fixed point of σ.

Example

The substitution $\sigma \colon a \mapsto ab, b \mapsto a$ is the Fibonacci substitution and $x = \sigma^{\omega}(a)$ is the *Fibonacci sequence*.

$$x = \sigma^{\omega}(a) = abaababaabaababaababaabaab\ldots$$

Example

The substitution $\sigma\colon a \mapsto aaba, b \mapsto b$ is the *Chacon binary substitution* and $x = \sigma^\omega(a)$ is the *Chacon binary sequence*.

The Chacon binary substitution is not primitive, but the shift $\mathsf{X}(\sigma)$, called the *Chacon binary shift*, is minimal.

This can be proved either directly (Exercise) or by exhibiting a primitive substitution τ such that $\mathsf{X}(\sigma)$ is conjugate to $\mathsf{X}(\tau)$ (Exercise, next slide).

Example

The primitive substitution $\tau : 0 \to 0012, 1 \to 12, 2 \to 012$ is the *Chacon ternary substitution*.
Show that $w_n = \tau^n(0)$ satisfies the recurrence relation $w_{n+1} = w_n w_n 1 w_n'$, where w_n' is obtained from w_n by changing the initial letter 0 into a 2.
Deduce from this that the 1-block map $\theta \colon 0, 2 \to 0, 1 \to 1$ defines a conjugacy from the substitution shift $\mathsf{X}(\tau)$ called the *Chacon ternary shift*, to the Chacon binary shift $\mathsf{X}(\sigma)$.
As a consequence, the Chacon binary shift is minimal.

Proof.

Set $w_n = 0t_n$ and thus $w'_n = 2t_n$ for $n \geq 0$. Then we have

$$w_{n+1} = \tau(w_n) = 0012\tau(t_n) = 0\tau(2t_n) = 0\tau(w'_n)$$

showing that $t_{n+1} = \tau(w'_n)$ for $n \geq 0$. Thus

$$\begin{aligned} w_{n+1} &= \tau(w_{n-1}w_{n-1}1w'_{n-1}) = w_n w_n 12\tau(w'_{n-1}) \\ &= w_n w_n 12 t_n = w_n w_n 1 w'_n. \end{aligned}$$

The map θ sends the infinite word $\tau^\omega(0)$ to $\sigma^\omega(a)$ and thus maps $\mathrm{X}(\tau)$ to $\mathrm{X}(\sigma)$. Its inverse is the map that replaces 0 by 2 whenever it is immediately preceded by 1. Thus θ is a conjugacy. □

Proposition

A shift space X is uniformly recurrent if and only if it is irreducible, and for every $u \in \mathcal{B}(X)$ the set of return words to u is finite.

Proof.

Assume first that X is uniformly recurrent. Let $u \in \mathcal{B}_n(X)$ and let $v \in \mathcal{B}(X)$ be of length $R_X(n) - n + 1$ with $vu \in \mathcal{B}(X)$. Then vu has length $R_X(n) + 1$ and thus u has a second occurrence in vu. This shows that v has a suffix in $\mathcal{R}_X(u)$. Thus $\max\{|w| + n - 1 \mid w \in \mathcal{R}_X(u), u \in \mathcal{B}_n(X)\} \leq R_X(n)$ and $\mathcal{R}_X(u)$ is finite.

□

Proof.

Conversely, let $N = \max\{|w| + n - 1 \mid w \in \mathcal{R}_X(u), u \in \mathcal{B}_n(X)\}$. Let $u \in \mathcal{B}_n(X)$ and $r \in \mathcal{B}_N(X)$. Since X is irreducible, there are words s, t such that $usrtu \in \mathcal{B}(X)$. If r has no block equal to u, this implies that r occurs in some uvu, where r is not a prefix or a suffix of uvu and $uv \in \mathcal{R}_X(u)$. This implies
$|uv| + n - 1 \leq N = |r| \leq |uvu| - 2$, whence
$n - 1 \leq |u| - 2 = n - 2$, a contradiction. Thus, $R_X(n) \leq N$. □

Example

Let $\sigma : a \mapsto ab, b \mapsto ba$ be the Thue-Morse substitution and $X = \mathrm{X}(\sigma)$ be the Thue-Morse shift.

We have

$$\begin{aligned}\mathcal{R}_X(ab) &= \{ab, aba, abb, abba\},\\ \mathcal{R}_X(aa) &= \{aababb, aababbab, aabb, aabbab\}\end{aligned}$$

with the elements of $\mathcal{R}_X(ab)ab$ colored in green in Figure 6 and those of $\mathcal{R}_X(aa)aa$ colored in red. Thus, the maximal length R of return words of length 2 is 8.

Thue-Morse

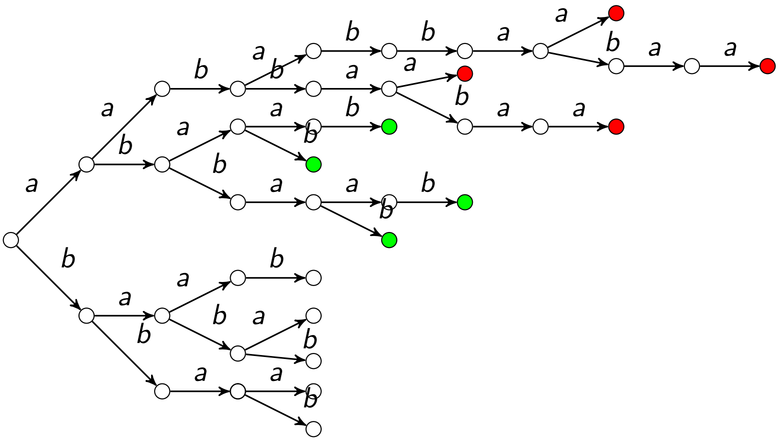


Figure: The words of $\mathcal{B}(X)$ for the Thue-Morse shift.

Computation of $\mathcal{R}_X(u)$ when $X = \mathsf{X}(\sigma)$ is minimal, u is a **prefix** of a fixed point x of σ and $w \in \mathcal{R}_X(u)$.
The word w can be an arbitrary element of $\mathcal{R}_X(u)$, for instance the prefix of x in $\mathcal{R}_X(u)$.

Computation of the return words of prefixes of a fixed point

ReturnWords(u, w)

▷ u is a prefix of $x = \sigma^\omega(a)$ and $w \in \mathcal{R}_X(u)$
▷ Returns in R the set $\mathcal{R}_X(u)$
$R \leftarrow \emptyset$
$S \leftarrow \{w\}$
▷ S is the set of return words to be processed
while $S \neq \emptyset$ **do**
    $r \leftarrow$ an element of S
    $S \leftarrow S \setminus \{r\}$
    $R \leftarrow R \cup \{r\}$
    $r(1), \cdots, r(k) \leftarrow \sigma(r)$
    ▷ The words $r(i)$ are the decomposition of $\sigma(r)$ in return words to u
    for $i \leftarrow 1$ **to** k **do**
        if $r(i) \notin R \cup S$ **then**
            $S \leftarrow S \cup r(i)$
return R

Example

Let $\sigma\colon a \mapsto ab, b \mapsto ba$ be the Thue-Morse substitution.
$\sigma^\omega(a) = abbabaabbaababba\ldots$.
$u = ab$.
$w = abb$. $S = \{abb\}$.

1. $r = abb$. $S = \emptyset$. $R = \{abb\}$. $\sigma(abb) = abb\,aba$. $S = \{aba\}$
2. $r = aba$. $S = \emptyset$. $R = \{abb, aba\}$. $\sigma(aba) = abba\,ab$.
   $S = \{abba, ab\}$
3. $r = ab$. $S = \{abba\}$. $R = \{abb, aba, abba, ab\}$.
   $\sigma(ab) = abba$. $S = \{abba\}$
4. $r = abba$. $S = \emptyset$. $R = \{abb, aba, abba, ab\}$.
   $\sigma(abba) = abb\,aba\,ab$. $S = \emptyset$

Thus, $\mathcal{R}_X(ab) = \{ab, aba, abb, abba\}$.

Block complexity

The *block complexity*, or just *complexity*, of a shift space X is the sequence $(p_X(n))_{n\geq 0}$ with $p_X(n) = \mathrm{Card}(\mathcal{B}_n(X))$.

We also write $p_x(n) = \mathrm{Card}(\mathcal{B}_n(x))$ for an individual sequence x.

Theorem (Morse, Hedlund)

Let x be a two-sided sequence. The following conditions are equivalent.

(i) *For some $n \geq 1$, one has $p_x(n) \leq n$.*

(ii) *For some $n \geq 1$, one has $p_x(n) = p_x(n+1)$.*

(iii) *x is periodic.*

Moreover, in this case, the least period of x is $\max p_x(n)$.

Proof.

(i) $\Rightarrow$ (ii). If $p_x(1) = 1$, then $p_x(n) = 1$ for all n. Assume $p_x(1) > 1$. Note that $p_x(n) \leq p_x(n+1)$ for all $n \geq 0$. If the inequality is strict for all $n \geq 1$, we have $p_x(n) > n$ for all $n \geq 1$. Thus $p_x(n) \leq n$ for some $n \geq 1$ implies $p_x(n) = p_x(n+1)$ for some $n \geq 1$.
(ii) $\Rightarrow$ (iii). For every $w \in \mathcal{B}_n(x)$, there is a unique letter $a \in A$ such that $wa \in \mathcal{B}_{n+1}(x)$. This implies that two consecutive occurrences of a word u of length n in x are separated by a fixed word depending only on u and thus that x is periodic.
(iii) $\Rightarrow$ (i) is obvious.
Let n be the least period of x. Since a primitive word of length n has n distinct conjugates, we have $p_x(n) = n$ and $p_x(m) = n$ for all $m \geq n$. This proves the final assertion. □

A shift space is *linearly recurrent* if it is minimal and if there is an integer $n \geq 1$ and a real number $K \geq 0$ such that, for every $u \in \mathcal{B}_{\geq n}(X)$, the length of every return word to u in X is bounded by $K|u|$.
We say that X is (K, n)-linearly recurrent.

We say that X is linearly recurrent with constant K. We say that X is linearly recurrent if it is K-linearly recurrent for some $K \geq 1$.

The lower bound of the numbers K such that X is K-linearly recurrent is called the *minimal constant* of linear recurrence.

Primitive substitution shifts are linearly recurrent

Proposition

A primitive substitution shift $\mathsf{X}(\sigma)$ *is linearly recurrent.*

Proposition

A primitive substitution shift $\mathsf{X}(\sigma)$ *is linearly recurrent with minimal constant* $K(\sigma) \leq kR|\sigma|$*, where* k *is such that* $|\sigma^n| \leq k\langle\sigma^n\rangle$ *for all* $n \geq 1$ *and* R *is the maximal length of a return word to a word of* $\mathcal{B}_2(\mathsf{X}(\sigma))$.

Primitive substitution shifts are linearly recurrent

Proof.

Let $\sigma \colon A^* \to A^*$ be a primitive substitution, and let $X = \mathrm{X}(\sigma)$ be the corresponding shift space. Since σ is primitive, it follows that there is a constant k such that, for all $n \geq 1$,

$$|\sigma^n| \leq k\langle \sigma^n \rangle \qquad (6)$$

Indeed, let $\lambda = \lambda_{M(\sigma)}$.
By the Perron-Frobenius theorem, the sequence $(M(\sigma)^n/\lambda^n)$ converges to the matrix yx where x, y are positive left and right eigenvectors relative to λ with $\sum_{a \in A} y_a = 1$ and $\sum_{a \in A} x_a y_a = 1$.

This implies that $\lim_{n\to\infty} \frac{|\sigma^n(a)|}{\lambda^n} = x_a$.

Indeed,
$\lim_{n\to\infty} \frac{|\sigma^n(a)|}{\lambda^n} = \sum_{b \in A} \lim_{n\to\infty} \frac{M(\sigma)^n_{b,a}}{\lambda^n} = \sum_{b \in A} (x \cdot y)_{b,a} = x_a.$ □

Primitive substitution shifts are linearly recurrent

Proof.

Consider $w \in \mathcal{B}(X)$. The substitution σ being primitive, the sequence $(\langle \sigma^n \rangle)_{n \geq 0}$ is nondecreasing and unbounded. There is an integer n such that

$$\langle \sigma^{n-1} \rangle \leq |w| \leq \langle \sigma^n \rangle. \tag{7}$$

Let v be a return word to w. Let $u \in \mathcal{B}(X)$ be such that vw occurs in $\sigma^n(u)$.
Since, by Inequality (7), $|w|$ is at most equal to every $|\sigma^n(a)|$, and we may assume that w occurs in the image by σ^n of the prefix ab of length 2 of u. We may also assume that u contains a second occurrence of ab. Set $\sigma^n(ab) = pwq$. Then the prefix of length $|v|$ of $\sigma^n(u)$ occurs in a word of $\sigma^n(\mathcal{R}_X(ab))$ (see Figure). □

Primitive substitution shifts are linearly recurrent

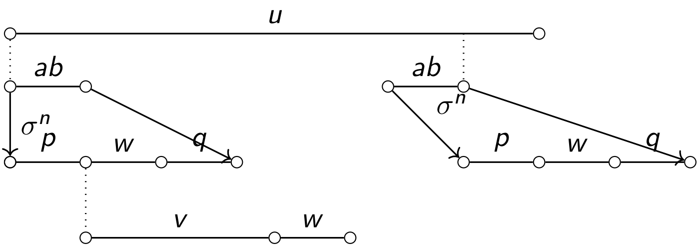


Figure: The words u, v, w.

Primitive substitution shifts are linearly recurrent

Proof.

Thus,

$$|v| \leq R|\sigma^n| \leq Rk\langle\sigma^n\rangle \leq Rk|\sigma|\langle\sigma^{n-1}\rangle \leq Rk|\sigma||w|,$$

where R is the maximal length of return words to a word of length 2. This shows that X is linearly recurrent with minimal constant

$$K(\sigma) \leq kR|\sigma|. \tag{8}$$

□

Left and right-special words

Let X be a shift space on A. For $w \in \mathcal{B}(X)$, let

$$\ell_X(w) = \mathrm{Card}\{a \in A \mid aw \in \mathcal{B}(X)\}, \quad r_X(w) = \mathrm{Card}\{a \in A \mid wa \in \mathcal{B}(X)\}$$

A word $w \in \mathcal{B}(X)$ is *left-special* with respect to X, or with respect to $\mathcal{B}(X)$, if $\ell_X(w) \geq 2$ (resp. $r_X(w) \geq 2$). It is *bispecial* if it is both left-special and right-special.

We define $s_X(n) = p_X(n+1) - p_X(n)$.

Left and right-special words

Then, one has

$$s_X(n) = \sum_{w \in \mathcal{B}_n(X)} (\ell_X(w) - 1) = \sum_{w \in \mathcal{B}_n(X)} (r_X(w) - 1). \qquad (9)$$

Indeed,

$$\begin{aligned} s_X(n) &= p_X(n+1) - p_X(n) = \mathrm{Card}(\mathcal{B}_{n+1}(X)) - \mathrm{Card}(\mathcal{B}_n(X)) \\ &= \sum_{w \in \mathcal{B}_n(X)} (\ell_X(w) - 1). \end{aligned}$$

If the sequence $s_X(n)$ is bounded, the complexity $p_X(n)$ is at most linear, that is $p_X(n) \leq kn$ for some $k \geq 1$. The converse is true, by an important result that we quote without proof.

Proposition (Cassaigne)

If the complexity of a shift X is at most linear, then $s_X(n)$ is bounded.

Proposition

If $\sigma : A^* \to A^*$ is a primitive substitution that is not the identity on a one-letter alphabet and such that $X = \mathsf{X}(\sigma)$ is not periodic, then $p_X(n) = \Theta(n)$.

Proof.

Since X is not periodic, we have $p_X(n) \geq n + 1$ for every $n \geq 1$ by the Morse-Hedlund theorem. Thus $p_X(n) = \Omega(n)$. □

Proof.

To prove the upper bound, let λ be the maximal eigenvalue of $M(\sigma)$.
Let $c, d > 0$ be such that $\ell_n = c\lambda^n \le |\sigma^n(a)| \le d\lambda^n = L_n$ for every $a \in A$. Changing σ for some of its powers, we may assume that $L_n \le \ell_{n+1}$.
In order to bound $p_X(k)$, consider n such that $\ell_n \le k < \ell_{n+1}$. Let $w \in \mathcal{L}_k(\sigma)$. No word $\sigma^{n+1}(a)$ with $a \in A$ can occur in w since otherwise $\ell_{n+1} \le k$, a contradiction. Thus there exist $a, b \in A$ such that $\sigma^{n+1}(ab) = pws$ with $|p| < |\sigma^{n+1}(a)|$. Since w is determined by a, b and $|p|$, this implies that

$$
\begin{aligned}
p_X(k) &\le \mathrm{Card}(A)^2 L_{n+1} \le \mathrm{Card}(A)^2 d\lambda^{n+1} \le \mathrm{Card}(A)^2 \lambda \frac{d}{c} c\lambda^n \\
&\le \mathrm{Card}(A)^2 \lambda \frac{d}{c} \ell_n \le \mathrm{Card}(A)^2 \lambda \frac{d}{c} k,
\end{aligned}
$$

showing that $p_X(k) = O(k)$. □

Example

The Fibonacci substitution $\sigma\colon a \mapsto ab, b \mapsto a$ is primitive. The complexity of the Fibonacci shift $X = \mathrm{X}(\sigma)$ is $p_X(n) = n + 1$.

Proof.

The words $F_n = \sigma^n(a)$ are left-special. Indeed, this is true for $n = 0$ since $aa, ba \in \mathcal{L}(\sigma)$. Next, $aF_{n+1} = \sigma(bF_n), \quad abF_{n+1} = \sigma(aF_n)$ show the claim by induction on n. It is easy to see (again by induction) that conversely, every left-special word is a prefix of some F_n. This implies that there is exactly one left-special word of each length and thus that $p_X(n) = n + 1$. □

Block complexity of linearly recurrent shift

Proposition

Every linearly recurrent shift has at most linear complexity. More precisely, a shift X is (K, n_0)-linearly recurrent if and only if, for $n \geq n_0$, every word of $\mathcal{B}_n(X)$ occurs in every word of $\mathcal{B}_m(X)$ when $m > (K+1)n - 2$. In this case, $p_X(n) \leq Kn$ for every $n \geq n_0$.

Block complexity of linearly recurrent shift

Proof.

Assume first that the shift X is (K, n_0)-linearly recurrent. Since, for $n \geq n_0$, the length of return words to $u \in \mathcal{B}_n(X)$ is at most Kn, the length of a word in $\mathcal{B}(X)$ without any occurrence of u is at most $n - 1 + Kn - 1 = (K+1)n - 2$. Thus, every word of length $m > (K+1)n - 2$ contains an occurrence u.

Conversely, if the condition is satisfied, let $n \geq n_0$ and let $u \in \mathcal{B}_n(X)$. Then X is minimal. Moreover, two consecutive occurrences of u in $\mathcal{B}(X)$ cannot be separated by more than Kn letters, and thus a return word to u is of length at most Kn.

We have $R_X(n) \geq p_X(n) + n - 1$ for $n \geq n_0$.
Indeed, the number of distinct words of length n occurring in a word of length N is at most $N - n + 1$. Therefore, $p_x(n) \leq R_X(n) - n + 1$. Hence $p_X(n) \leq Kn$ for $n \geq n_0$. □

Generalizations (admitted)

Theorem (Maloney and Rust 2018 for non-erasing substitutions)

Let σ be an iterable substitution. If $\mathsf{X}(\sigma)$ is minimal, then it is conjugate to $\mathsf{X}(\zeta)$, where ζ is a primitive substitution that is computable. Furthermore, for some $n \geq 1$, the dominant eigenvalues of σ^n and ζ coincide.

Corollary

A minimal substitution shift is linearly recurrent.