Master 2 Mathematics and Computer Science
Symbolic Dynamics. Lecture 5

Marie-Pierre Béal

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

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

$$\ell(\sigma(w)) = M(\sigma)\ell(w). \tag{3}$$

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

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

Blocks of a substitution 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 $\sigma : a \mapsto ab, b \mapsto b$. 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, $\lim_{n\to+\infty} |\sigma^n(a)| = +\infty$.

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 $\mathcal{R}_X(u)$.

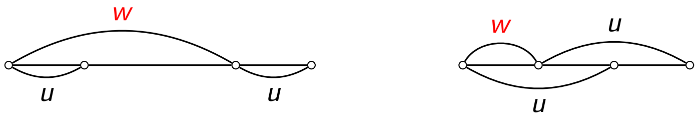
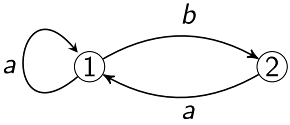


Figure: Return word to u.

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.

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. $\square$

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 $a \in A$. $\square$

Examples

Example

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

Example

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

Return words

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.

$\square$

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)

$\triangleright$ u is a prefix of $x = \sigma^\omega(a)$ and $w \in \mathcal{R}_X(u)$
$\triangleright$ Returns in R the set $\mathcal{R}_X(u)$
$R \leftarrow \emptyset$
$S \leftarrow \{w\}$
$\triangleright$ S is the set of return words to be processed
while $S \neq \emptyset$ **do**
    $r \leftarrow$ an element of S
    $S \leftarrow S \setminus \{r\}$
    $R \leftarrow R \cup \{r\}$
    $r(1), \cdots, r(k) \leftarrow \sigma(r)$
    $\triangleright$ 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 $\sigma: a \mapsto ab, b \mapsto ba$ 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 $\max p_x(n)$.

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* $\mathcal{B}_2(\mathsf{X}(\sigma))$*.*

Proposition

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

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 $n \geq n_0$.