TOPOLOGICAL CONJUGACY OF CONSTANT LENGTH
SUBSTITUTION DYNAMICAL SYSTEMS

ETHAN M. COVEN, F. MICHEL DEKKING, AND MICHAEL S. KEANE

Abstract. Primitive constant length substitutions generate minimal symbolic dynamical systems. In this article we present a general procedure which decides whether two given systems arising from the same constant length substitution are topologically conjugate. We show that each conjugacy class contains only finitely many injective substitutions, as well as infinitely many which are not injective. An effective method is given for listing all injective substitutions in any given class, as well as all injective substitutions (of the same constant length) whose systems are factors of systems of the original class. As examples, the Toeplitz conjugacy class contains three injective substitutions (two on two symbols and one on three symbols), and the length two Thue-Morse conjugacy class contains twelve substitutions, among which are two on six symbols. Together, they constitute a list of all primitive substitutions of length two with infinite minimal systems which are factors of the Thue-Morse system.

Key words: Substitution dynamical system; conjugacy; sliding block code;
Thue-Morse substitution; Toeplitz substitution

MSC: 37B10, 54H20

1. Introduction

In the article [3] published in 1971, the minimal dynamical systems arising from primitive substitutions on a binary alphabet having the same constant length were classified, yielding for a given such substitution a list of all substitutions of the same length generating topologically conjugate systems. Here we extend this classification to arbitrary finite alphabets. More recently, the articles [4] and [5] exhibit characterizations of such systems; these only implicitly yield corresponding topological conjugacies. Also, in [12] our goal has been partially accomplished—with measure-theoretic conjugacy—for a restricted class of constant length substitutions.

If two constant length substitution systems are topologically conjugate, then the lengths of the substitutions are powers of the same integer ([10], [3]). Therefore, by taking suitable powers we can, and do, restrict our attention to substitutions of the same length $L$. 

Date: January 3, 2014.
The second author is partially supported by the department of Mathematics and Computer Science of Wesleyan University.
In this contribution, we present a complete classification, giving explicit conjugacies in an algorithmic manner. More specifically, we address and solve the following two problems, in which $L$ denotes a fixed integer larger than one.

**Problem 1.1.** Let $\alpha$ and $\beta$ be two substitutions of the same length $L$, both primitive and nonperiodic. Decide whether the dynamical systems $(X_\alpha, \sigma)$ and $(X_\beta, \sigma)$ are topologically conjugate.

**Problem 1.2.** Let $\alpha$ be a substitution of length $L$, primitive and nonperiodic. Give a list of all the injective substitutions $\beta$ of length $L$ such that the dynamical systems $(X_\alpha, \sigma)$ and $(X_\beta, \sigma)$ are topologically conjugate.

Finite systems are elementary, and we restrict attention everywhere to the nonperiodic case of primitive substitutions with corresponding infinite minimal sets.

We show that to any primitive substitution of constant length whose minimal set is infinite, there are always infinitely many primitive substitutions of the same constant length having topologically conjugate minimal systems, but only finitely many of these are injective. Thus, the list produced by our algorithmic solution of Problem 1.2 will, starting from any given primitive substitution of constant length, consist of all injective substitutions of that length with dynamical systems topologically conjugate to the initial system. Clearly, since the list in Problem 1.2 is finite, Problem 1.1 has then also been solved.

2. Substitutions and standard forms

We begin by recalling the basic definitions and known results without proof for primitive substitutions and their corresponding minimal systems, referring the reader to the standard reference [14].

Let $A$ be a finite set (an alphabet) with $c \geq 2$ elements which are symbols, or letters. Elements of $A^* = \cup_{n=0}^{\infty} A^n$ are called words. A substitution is a mapping $\alpha : A \longrightarrow A^*$.

The substitution $\alpha$ is of constant length $L$ if $\alpha(a) \in A^L$ for each $a \in A$. It is natural to view $A^*$ as a semigroup under juxtaposition, thus extending $\alpha$ to mappings from $A^*$ to $A^*$, $A^N$ to $A^N$, and $A^Z$ to $A^Z$ - no confusion results if we also denote them by $\alpha$, and they can be iterated, defining $\alpha^n$ for each $n \in \mathbb{N}$.

**Definition** The substitution $\alpha$ is primitive if for some $n > 0$ and for every $a \in A$ the word $\alpha^n(a)$ contains each of the letters of $A$. The language of $\alpha$ is the subset of $A^*$ consisting of those words appearing as consecutive letters, subwords, or factors, of images under powers of $\alpha$. Words from the language of $\alpha$ are called $\alpha$-admissible, or simply admissible. We denote by $A^N_\alpha$ the set of $\alpha$-admissible words of length $N$. 
We write $X_\alpha$ for the compact subset of $A^\mathbb{Z}$ of bilaterally infinite sequences each of whose finite factors belongs to the language of $\alpha$. Under the left shift $\sigma$ on $A^\mathbb{Z}$, it is a minimal symbolic system whenever $\alpha$ is primitive. If in addition, $X_\alpha$ is infinite, then $\alpha$ is recognizable [13]. For constant length $L$ substitutions, this is equivalent to the existence of a conjugacy from the minimal system $(X_\alpha, \sigma)$ to the rotation by 1 on the compact group of $L$-adic integers, which describes a unique hierarchical structure for each of the sequences belonging to $X_\alpha$.

For constant length $L$ substitutions, it is clear that the names we give to the individual symbols of their alphabets are not essential - different namings will produce conjugate systems. This leads us to restricting an alphabet of $c$ symbols to the alphabet $A = \{1, \ldots, c\}$. Even then, there is a permutational ambiguity, since permuting $A$ will yield up to $c!$ different substitutions, which we view as essentially the same. We find it useful in the following to single out one of these permutations as the one yielding the standard form of a substitution, as follows. If $\alpha$ is a constant length $L$ substitution on the alphabet of size $c$, then we define its characteristic word to be the word $\alpha(1) \cdots \alpha(c)$ of length $Lc$. For constant length substitutions, permutations yielding different substitutions then possess different characteristic words, and we call the substitution with the lexicographically smallest characteristic word the standard form of the substitution $\alpha$.

3. Letter-to-letter maps

Let $A$ and $B$ be finite alphabets. A map 

$$\pi : B \to A$$

is called a letter-to-letter map; by juxtaposition it clearly extends to maps from (finite or infinite) sequences on $B$ to sequences of the same lengths on $A$. We also denote this extension by the same symbol $\pi$. It will appear that the following easily proved lemma is the key to understanding the properties of conjugacies.

Lemma 3.1. If $\alpha : A \to A^*$ and $\beta : B \to B^*$ are substitutions, and if $\pi$ satisfies the intertwining equation $\alpha \pi = \pi \beta$, then for each positive integer $n$

$$\alpha^n \pi = \pi \beta^n.$$ 

Under the hypotheses of the lemma, the word $\beta^n(b)$ is mapped by $\pi$ to the word $\alpha^n(a)$, with $a = \pi(b)$, for any positive $n$. In particular, the language of $\beta$ is mapped to the language of $\alpha$, and we have:

Corollary 3.1. $\pi(X_\beta) \subseteq X_\alpha$, with equality whenever $\pi$ is surjective. In particular, if $\pi$ is surjective, then primitivity of $\beta$ implies primitivity of $\alpha$ and minimality of $X_\beta$ implies minimality of $X_\alpha$. 
4. \(N\)-Block presentations

In this section, we produce for a given primitive constant length substitution \(\alpha\) with infinite associated minimal system \((X_\alpha, \sigma)\), an infinite number of different primitive substitutions of the same constant length with topologically conjugate systems.

Let \(A\) be a finite alphabet, and let \(\alpha\) be a constant length primitive substitution on \(A\). We denote its length by \(L\), an integer greater than one. Further, let \(N\) denote any positive integer, fixed for the moment. A well-known and elementary but important result from ([11]) is that the symbolic systems \((A^Z, \sigma)\) and \(((A^N)^Z, \sigma)\) are topologically conjugate in a very simple manner, via the letter-to-letter map \(\pi\) from \(A^N\) to \(A\) which associates to each word of \(A^N\) its first letter. The inverse \(\psi\) of this conjugacy is generated by the \(N\)–block code which replaces each letter of \(A\) in an infinite sequence by the word given by that letter followed by the next \(N-1\) letters of the infinite sequence.

If we now concentrate our attention on \(X_\alpha\), and define the alphabet \(B = A_\alpha^N\), then the map \(\pi\) from \(B\) to \(A\) satisfies the intertwining condition \(\alpha\pi = \pi\beta\) of the previous section, if we define the substitution \(\beta\) on \(A\) properly. Namely, if \(b = a_1 \ldots a_N\) is an element of \(B\), then it belongs to the language of \(\alpha\). We can therefore apply \(\alpha\) to \(b\), obtaining a word \(\alpha(a_1 \ldots a_N)\) of length \(LN\). Now choose and fix any integer \(i\) with \(1 \leq i \leq (L-1)(N-1)\), so that the factor \(w\) of length \(L + N - 1\) of \(\alpha(a_1 \ldots a_N)\) starting with the \(i^{th}\) symbol is well–defined, and apply the “local version” of \(\psi\) to \(w\) to obtain an element of \(A^N\), which we define now to be \(\beta(b)\):

\[
\begin{array}{ccc}
X_\alpha & \xrightarrow{\alpha} & X_\alpha \\
\uparrow & & \downarrow \psi \\
X_\beta & \xrightarrow{\beta} & X_\beta 
\end{array}
\]

The reader is encouraged to examine the simple example below.

Now, Lemma 3.1 implies that \(\beta\) is primitive and that, since \(\psi\) is a conjugacy with inverse \(\pi\), \((X_\alpha, \sigma)\) and \((X_\beta, \sigma)\) are topologically conjugate. Each choice of \(i\) in the designated range yields a primitive substitution, so that we produce in principle \((L-1)(N-1)\) substitutions \(\beta\) of length \(L\), all with systems topologically conjugate to \((X_\alpha, \sigma)\).

Example Let \(A = \{1, 2, 3\}\), and let \(\alpha\) be given by

\[
\alpha(1) = 23, \; \alpha(2) = 13, \; \alpha(3) = 12.
\]

Then the words of length three in the language of \(\alpha\) are 122, 131, 223, 231, 312 and 313. Coding these in lexicographical order to a standard alphabet gives \(B = \{1, 2, \ldots, 6\}\). Since \(\alpha(122) = 231313\), we have \(\beta(1) = 46\). Performing this for all
symbols we find that
\[ \beta(1) = 46, \beta(2) = 45, \beta(3) = 26, \beta(4) = 25, \beta(5) = 13, \beta(6) = 13. \]

If \( X_\beta \) is infinite, then clearly the alphabets \( B \) grow larger and larger with \( N \).
This proves:

**Theorem 4.1.** For any primitive constant length substitution with infinite associated symbolic system there exist infinitely many primitive substitutions of the same length with symbolic systems topologically conjugate to the given system.

**Remark** Some caution is necessary here, as we have examples of topologically conjugate substitution systems which cannot be obtained in this manner.

5. FOR SUBSTITUTION MINIMAL SETS 3-BLOCK CODES SUFFICE

Here we give a new proof of a known result ([4]).

**Theorem 5.1.** Let \( \alpha \) and \( \beta \) each be primitive injective substitutions of constant length \( L > 1 \), whose minimal systems \( (X_\alpha, \sigma) \) and \( (X_\beta, \sigma) \) are infinite. If there exists a semi-conjugacy from \( (X_\alpha, \sigma) \) to \( (X_\beta, \sigma) \), then there is such a semi-conjugacy which is given by a three-block block code.

**Proof.** Denote by \( \Phi \) the hypothesized semi-conjugacy. As noted in Section [4] this semi-conjugacy is given by a finite block code \( \phi \), and we may assume without loss of generality that this map is an \( L^n \)-block code with memory 0 for some integer \( n \).

Recall that \( A_3^n \) denotes the set of admissible words of length three, and let \( B \) be the alphabet of \( \beta \). The proof now consists of two steps:

Step 1. Construction of a three-block code \( \psi \) from \( A_3^n \) to \( B \).

Choose any three-block \( ijk \in A_3^n \). The block \( \alpha^n(ijk) \) is an admissible \( 3L^n \)-block of \( \alpha \), to which we can apply \( \phi \), obtaining an admissible \( (2L^n+1) \)-block of \( \beta \).

By recognizability, there is a unique \( \beta^n \)-block, say \( \beta^n(p) \), occurring at a fixed position (independent of the choice of \( ijk \)) in this block. We then define \( \psi(ijk) := p \).

Step 2. The block code \( \psi \) defines a map \( \Psi \) from \( X_\alpha \) to a closed, shift-invariant set \( Y \) of sequences from the alphabet \( B \), so that \( \Psi \) is a semi-conjugacy from \( (X_\alpha, \sigma) \) to \( (Y, \sigma) \). We show in this step that \( Y = X_\beta \).

To verify this, choose any \( x \in X_\alpha \), apply \( \alpha^n \) to \( x \), then apply \( \Phi \), and finally “decode” using recognizability of \( \beta^n \). The resulting sequence must then be an element of \( X_\beta \), and by minimality all elements of this set occur.

**Corollary.** If the semi-conjugacy of the 3-block Theorem is a conjugacy, then the three-block code which results from the proof is also a conjugacy.

**Proof.** If \( x \) and \( x' \) are different points in \( X_\alpha \), it is obvious that their images under \( \Psi \) are also different, so that a conjugacy results.
Remark In [12] it is shown for a rather special class of substitutions that the measure-theoretic semi-conjugacies are given by 2-block codes. The example of the Thue-Morse substitution (see Section 11) shows that 3-block codes are sometimes necessary.

6. INJECTIVE SUBSTITUTIONS

A key ingredient in our classification result is that we may suppose that the substitutions are injective. This is based on the following result.

Theorem 6.1. (1) Any primitive, nonperiodic substitution which is not injective is conjugate to a primitive, nonperiodic substitution that is injective.

The proof given in [1] is constructive, and yields what we call the canonical injective version of a substitution. The construction amounts to identifying (iteratively) those letters which have equal images. For example, the substitution $\beta$ given by

$\beta(1) = 46, \beta(2) = 45, \beta(3) = 26, \beta(4) = 25, \beta(5) = 13, \beta(6) = 13$ from the example in Section 4 transforms in a first step to

$\beta'(1) = 45, \beta'(2) = 45, \beta'(3) = 25, \beta'(4) = 25, \beta'(5) = 13,$

and then in a second step to the injective substitution

$\beta''(1) = 35, \beta''(3) = 15, \beta''(5) = 13.$

Remark It is interesting to us that recognizability is equivalent to invertibility of the substitution under consideration as a map from its minimal system to its image. We see this as the key point of the proof.

7. SUBSTITUTIONS AND GRAPH HOMOMORPHISMS

Let $x$ be an infinite sequence over an alphabet $A$. Here we study the general question whether $x$ can be generated by a substitution of length $L$.

We consider graphs $G = (V, E)$, $G' = (V', E')$, and graph homomorphisms $\varphi : G \to G'$, i.e., maps $\varphi : V \to V'$ having the property that $(u, v) \in E$ implies that $(\varphi(u), \varphi(v)) \in E'$.

Let $W_2 = \{ab : ab = x_kx_{k+1} \text{ for some } k \in \mathbb{Z}\}$, be the set of 2-blocks occurring in $x$, and for $0 \leq M \leq L - 1$ let $W_{L,M} = \{a_1\ldots a_L : a_1\ldots a_L = x_{kL+M}\ldots x_{kL+M+L-1} \text{ for some } k \in \mathbb{Z}\}$ be the set of of $L$-blocks occurring in $x$ at positions $M \mod L$.

With $x$ we associate a family of graphs. The simplest is $G_1^x = (V_1, E_1)$, given by

$V_1 = A, \quad E_1 = \{(a, b) : ab \in W_2\}.$

The graphs $G_{L,M}^x = (V_{L,M}, E_{L,M})$ for $M = 0, \ldots, L - 1$ are defined by

$V_{L,M} = W_{L,M}, \quad E_{L,M} = \{(a_1\ldots a_L, b_1\ldots b_L) : a_1\ldots a_L b_1\ldots b_L \in W_{2L,M}\}.$

We follow the convention of calling a surjective homomorphism an epimorphism.
Theorem 7.1. Let $x$ be sequence over $A$, and let $\varphi$ be a primitive substitution of length $L$ over $A$. Then $x$ is in $X_\varphi$ if and only if $\varphi$ is a graph epimorphism, $\varphi : G_1^x \to G_{L,M}^x$ for some $0 \leq M \leq L - 1$.

Proof. $\Rightarrow$ Suppose $x$ is in $X_\varphi$. Then $x$ can be written as a concatenation of $\varphi$-blocks (recognizability is not needed for this). Define $M$ as the first cutting position at or after 0. Let $y$ be such that $x = \sigma^M \varphi(y)$. By minimality of $X_\varphi$, all $\varphi$-blocks have to occur in $x$. So for each $\varphi$-block in $x$ there is a letter associated to it in $y$. This gives surjectivity of $\varphi$ considered as a graph homomorphism.

$\Leftarrow$ If all words occurring in $x$ are $\varphi$-admissible, then (by minimality of $X_\varphi$) $x$ is an element of $X_\varphi$. If not, let $w$ be the smallest word occurring in $x$ which is not $\varphi$-admissible. We can cut $x$ up in words of length $L$, which are vertices of $G_{L,M}^x$. Let $u_1 \ldots u_r$ be the smallest block consisting of such words that contains $w$. Because $\varphi$ is a graph epimorphism, these words are of the form $u_k = \varphi(v_k)$ for $k = 1, \ldots, r$, where the $v_k$ are from $A$. But then, if $|w| \geq 3$, $v_1 \ldots v_r$ is a shorter word than $w$, which is not $\varphi$-admissible. This would contradict the minimality of the length of $w$. But words $w$ of length 1 are certainly $\varphi$-admissible, and all words of length 2 will occur in the words from $W_{2L,M}$, which are also $\varphi$-admissible by the definition of $\varphi$. 

The theorem implies that a sequence $x$ has to be very rigid to be generated by a substitution of length $L$: this can only be the case if the number of $L$-blocks occurring in $x$ at positions $M$ modulo $L$ for some $M$ is no more than the number of symbols in the alphabet.

Example: Thue-Morse sequence

We consider the Thue-Morse sequence $x = 0110100110010110 \ldots$. It is easy to write down the graphs of the letters and the 2-blocks:

Note that $G_{2,1}^x$ has too many vertices, and with $G_{2,0}^x$ we find two surjective graph homomorphisms: $\varphi(0) = 01, \varphi(1) = 10$, corresponding to the usual substitution, but also $\varphi'(0) = 10, \varphi'(1) = 01$. Note that both are in standard form.
8. A solution to the list problem

**Theorem 8.1.** Let $\alpha$ be a primitive substitution generating an infinite dynamical system $(X_\alpha, \sigma)$. There is an algorithm to generate all primitive injective substitutions $\beta$ of the same length such that $(X_\beta, \sigma)$ is a factor of $(X_\alpha, \sigma)$.

**Proof:** By Theorem 6.1 we may suppose that $\alpha$ is injective. By Theorem 5.1 we may furthermore suppose that the factor map is a 3-block map. Theorem 7.1 is then used to make the list of all factors. Start with the 3-block presentation $X_\alpha^{[3]}$ of $\alpha$ from Section 3. Take any sequence $u$ from $X_\alpha^{[3]}$. For all (including the identity) letter-to-letter maps from $X_\alpha^{[3]}$ to another shift space consider the image $x$ of $u$ under the factor map. For any $M = 0, ..., L - 1$ determine the graphs $G^x_1$ and $G^x_{L,M}$. Then determine all epimorphisms $\varphi$ from $G^x_1$ to $G^x_{L,M}$. If the corresponding substitution is primitive, then $x$ is an element of $X_\varphi$ by Theorem 7.1 and by minimality the factor is $X_\varphi$. Suppose on the other hand, that the factor is generated by some primitive substitution $\eta$. Let $x$ be the same sequence as above. Then Theorem 7.1 tells us that $\eta$ is some epimorphism from $G^x_1$ to $G^x_{L,M}$, i.e., by finding all epimorphisms we find all $\eta$'s. \qed

For our solution of the list problem we still need another ingredient. A dynamical system is called coalescent if every endomorphism is an automorphism, i.e., every topological semi-conjugacy from the system onto itself is a topological conjugacy. It was shown for a two symbol alphabet in [6] and for a general alphabet in [8] that primitive, not necessarily constant length, substitutions generate coalescent dynamical systems. We omit our simple proof for constant length substitutions.

**Theorem 8.2.** Let $\alpha$ be a primitive substitution generating an infinite dynamical system $(X_\alpha, \sigma)$. There is an algorithm to generate all primitive injective substitutions $\beta$ of the same length such that $(X_\beta, \sigma)$ is conjugate $(X_\alpha, \sigma)$.

**Proof:** Use Theorem 8.1 to determine all primitive injective substitutions $\beta$ with the same length that generate factors of $(X_\alpha, \sigma)$. Make the list for $\beta$, and check whether $\alpha$ is on it. If it is, then $(X_\alpha, \sigma)$ is conjugate to $(X_\beta, \sigma)$, by coalescence; if not, then $(X_\alpha, \sigma)$ is not conjugate to $(X_\beta, \sigma)$. \qed

9. A list free solution to Problem 1.1

We remark that Problem 1.1 can be solved without using our solution to Problem 1.2. In particular we formulate

**Lemma 9.1.** Let $\alpha$ on $A$ and $\beta$ on $B$ be primitive nonperiodic injective substitutions of constant length $L > 1$. Then $(X_\alpha, \sigma)$ is topologically semi-conjugate to $(X_\beta, \sigma)$ via a letter-to-letter map $f$ if and only if there exists an integer $M$ with $0 \leq M < L$ such that for every $\beta$-admissible $bb'$ there exists an $\alpha$-admissible word $aa'a''$
satisfying

\[ (\ast) \quad \beta(bb') = f(\alpha(aa'a''))^M_{M+2L-1}. \]

Here \( w_k^{\alpha} \) denotes the factor \( w_k \ldots w_m \) of a word \( w_1 \ldots w_p \).

**Proof.** \( \Rightarrow \) Take any \( x \in X_\beta \). Then \( x \) can be written as a concatenation of \( \beta \)-blocks. Define \( M \) as the first cutting position at or after 0. Let \( y \) be such that \( x = \sigma_M f(y) \). Then in particular \( (\ast) \) has to hold for all admissible \( bb' \).

\( \Leftarrow \) By minimality of \( (X_\beta, \sigma) \), it is enough to show that \( f(X_\alpha) \subset X_\beta \). Suppose not. Let \( w \) be the shortest word occurring in some sequence \( x \in f(X_\alpha) \) which is not \( \beta \)-admissible. The length of \( w \) is at least 3, since all words \( f(\alpha(aa'a'')) \) are \( \beta \)-admissible according to \( (\ast) \). Let \( v_1 \ldots v_r \) be the smallest block such that \( u_1 \ldots u_r := \beta(v_1 \ldots v_r) \) contains \( w \). Then \( u_1 \ldots u_r \) can not be \( \beta \)-admissible, and neither can \( v_1 \ldots v_r \). Since the length of \( w \) is at least 3, \( v_1 \ldots v_r \) is a shorter word than \( w \), and we have a contradiction. \( \square \)

Combining this lemma with Theorem 6.1, Theorem 5.1, and coalescence the following result can be proved. In the result we write \( \alpha_{[3]} \) for the canonical substitution on the 3-blocks with \( i = 0 \) in Section 4.

**Theorem 9.1.** Let \( \alpha \) and \( \beta \) be primitive nonperiodic substitutions of the same length \( L \). Then \( (X_\alpha, \sigma) \) and \( (X_\beta, \sigma) \) are conjugate if and only if there exists a map \( f : A_\alpha^3 \to B \) and an integer \( M \) with \( 0 \leq M < L \) such that for every \( \beta \)-admissible \( bb' \) there exists an \( \alpha_{[3]} \)-admissible word \( aa'a'' \) satisfying

\[ \beta(bb') = f(\alpha_{[3]}(aa'a''))^M_{M+2L-1}, \]

and the same formula holds with \( \alpha \) and \( \beta \) interchanged (with different \( f \) and \( M \)).

**10. The conjugacy class of the Toeplitz substitution**

Here we use Theorem 5.1 to determine the injective substitutions of length two that generate factors of the Toeplitz system \( (X_\alpha, \sigma) \) generated by the substitution

\[ \alpha(0) = 01, \quad \alpha(1) = 00. \]

Actually, the property of \( \alpha \) that the first letters of the two \( \alpha \)-blocks are equal implies that that for any \( n \alpha^n(0) \) and \( \alpha^n(1) \) only differ in their final letter. It then suffices to restrict ourselves to 2-block codes.

The set of \( \alpha \)-admissible words of length two is equal to \( A_\alpha^2 = \{00, 01, 10\} \), so we code the 2-blocks by \( B = \{1, 2, 3\} \). The graphs \( G_1 = G_1^T, G_{2,0} = G_{2,0}^T \) and \( G_{2,1} = G_{2,1}^T \) of a sequence \( x \) in the 2-block presentation \( X_\alpha^{[2]} \) are given by
There are two surjective graph homomorphisms $\varphi : G_1 \to G_{2,0}$ which give a primitive substitution:

$\varphi(1) = 23$, $\varphi(2) = 23$, $\varphi(3) = 11$, and $\varphi(1) = 23$, $\varphi(2) = 11$, $\varphi(3) = 23$.

After injectivization the first one gives the substitution $\beta$ given by $\beta(1) = 13$, $\beta(3) = 11$, whose standard form is the Toeplitz substitution. The second one injectivizes to the substitution $\beta$ given by $\beta(1) = 21$, $\beta(2) = 11$, which is the rotated Toeplitz substitution.

There is exactly one surjective graph homomorphism $\varphi : G_1 \to G_{2,1}$ which gives the primitive substitution:

$\varphi(1) = 32$, $\varphi(2) = 31$, $\varphi(3) = 12$,

which has standard form given by $\beta(1) = 23$, $\beta(2) = 13$, $\beta(3) = 12$. We call this substitution 3-symbol Toeplitz.

Note that the systems generated by rotated and 3 symbol Toeplitz are not only factors, but conjugate to Toeplitz since the 2-block presentation is conjugate to the original system.

To finish, we still have to examine the possibilities of letter-to-letter maps $\pi : \{1, 2, 3\} \to \{\hat{1}, \hat{2}\}$, where $\{\hat{1}, \hat{2}\}$ is a two letter alphabet. There are three of these maps $\pi_k$ given by

$\pi_1 : 1 \to \hat{1}$, $2 \to \hat{2}$, $3 \to \hat{1}$, $\pi_2 : 1 \to \hat{1}$, $2 \to \hat{2}$, $3 \to \hat{1}$, $\pi_3 : 1 \to \hat{2}$, $2 \to \hat{1}$, $3 \to \hat{1}$.

Let $t_k, k = 1, 2, 3$ be a sequence from $\pi_k(X_\varphi)$, where $\varphi$ is 3-symbol Toeplitz on $\{1, 2, 3\}$. The graphs $G_1^1 = G_{1,1}^1$, $G_{2,0}^1 = G_{1,0}^1$, and $G_{2,1}^1 = G_{1,1}^2$ are given by

There are obvious graph homomorphisms from $G_1^1$ to $G_{2,0}^1$ and to $G_{2,1}^1$. The first one again yields the Toeplitz substitution, the second one yields the substitution

$\varphi(\hat{1}) = \hat{2} \bar{1}$, $\varphi(\hat{2}) = \bar{1} \hat{1}$.
whose standard form is rotated Toeplitz.

One can check that the letter-to-letter map \( \pi \) gives similar results, and that the graph \( G_\pi \) has two loops, which prevents graph homomorphisms in this case.

Conclusion: the conjugacy class of Toeplitz consists of three substitutions: Toeplitz, rotated Toeplitz, and 3-symbol Toeplitz.

11. The length 2 substitution factors of the Thue-Morse system

Let \( \alpha \) be the Thue-Morse substitution

\[
\alpha(0) = 01, \quad \alpha(1) = 10.
\]

The set of admissible words of length 3 is then

\[
A_\alpha^3 = \{001, 010, 011, 100, 101, 110\}.
\]

The usual coding—which happens to be the binary coding—gives the 3-block alphabet \( B = \{1, 2, 3, 4, 5, 6\} \). The graph \( G_1 = G_\alpha \) of a sequence \( x \) in the 3-block presentation is given by

![Graph G1]

The graphs \( G_{2,0} = G_\alpha^{x_0} \) and \( G_{2,1} = G_\alpha^{x_1} \) of a sequence \( x \) in the 3-block presentation \( X_\alpha \) are given by

![Graphs G2,0 and G2,1]

To find all graph epimorphisms from \( G_1 \) to \( G_{2,i} \) for \( i = 0, 1 \), we exploit the following simple lemma.

**Lemma 11.1.** Let \( \varphi : G \to G' \) be a graph homomorphism. Suppose \( G' \) has no loops. Then 2-cycles and 3-cycles in \( G \) are mapped to 2-cycles, respectively 3-cycles in \( G' \).

We start with finding all \( \varphi : G_1 \to G_{2,1} \). By the lemma, \( \{\varphi(2), \varphi(5)\} \) equals \( \{13, 64\} \). If \( \varphi(2) = 13 \), then \( \varphi(4) = 65 \) and \( \varphi(1) = 24 \), and also \( \varphi(5) = 64 \), \( \varphi(3) = 12 \) and \( \varphi(6) = 53 \), since \( (2, 4, 1) \) and \( (5, 3, 6) \) form 3-cycles. If \( \varphi(2) = 64 \), then in the same way we obtain a second epimorphism

\[
1 \to 53, \quad 2 \to 64, \quad 3 \to 65, \quad 4 \to 12, \quad 5 \to 13, \quad 6 \to 24.
\]
Next we consider all \( \varphi : G_1 \to G_{2,0} \). Now \( \{ \varphi(2), \varphi(5) \} \) equals \( \{36, 41\} \). If \( \varphi(2) = 36 \), then \( \varphi(4) = 52 \) and \( \varphi(1) = 41 \), or \( \varphi(4) = 41 \) and \( \varphi(1) = 25 \).

In the first case necessarily \( (5, 3, 6) \to (41, 25, 36) \), and in the second case \( (5, 3, 6) \to (41, 36, 52) \).

If \( \varphi(2) = 41 \), then in the same way we obtain a third and fourth epimorphism

\[
\begin{align*}
1 \to 36, & \quad 2 \to 41, 3 \to 52, 4 \to 25, 5 \to 36, 6 \to 41, \\
1 \to 52, & \quad 2 \to 41, 3 \to 41, 4 \to 36, 5 \to 36, 6 \to 25.
\end{align*}
\]

After injectivization this yields 4 substitutions on a four symbol alphabet.

We now do the letter-to-letter maps. This is much more involved than in the case of the Toeplitz substitution.

Note that the letter-to-letter maps from \( A_6^{[3]} = \{1, 2, 3, 4, 5, 6\} \) to another alphabet are in one to one correspondence with the set of all partitions of \( \{1, 2, 3, 4, 5, 6\} \). Hence there are \( B_6 = 203 \) of such maps. Since \( M \) can take the values 0 and 1, this means that there are 406 cases of candidate epimorphisms to consider.

To reduce this number, we note that there is the mirror symmetry \( 0 \to 1, 1 \to 0 \), which at the level of 3-blocks corresponds to the permutation

\[ P = (16)(25)(34). \]

Obviously a partition and its permuted version will generate (if any) a substitution with the same standard form.

To further speed up the process we can apply the following three simple tools.

(T1) If \( G_{L,M} \) has more nodes than \( G_1 \), then an epimorphism is not possible.

(T2) If the graph \( G_1 \) contains a loop then \( G_{L,M} \) contains a loop.

(T3) If \( G_1 \) and \( G_{L,M} \) have the same number of nodes, then they also must have the same number of edges.

When \( M = 0 \) tool (T2) is particularly useful, since loops in \( G_{2,0} \) (where the prime indicates that the symbols are merged according to the partition at hand) are relatively rare in this case.

When \( M = 1 \) tool (T1) is especially useful. For example, for all the 15 partitions of type 2/2/2, the number of nodes in \( G_{2,1} \) is at least 4, except for the partition \( \{1, 6\} \{2, 5\} \{3, 4\} \) which yields the substitution \( 1 \to 23, 2 \to 13, 3 \to 12 \). This gives a factor on three symbols, which is 3-symbol Toeplitz (see Section 10).

With aid of the tools one finds that there are 15 factors of the Thue-Morse system generated by injective substitutions of length 2. Their standard forms on the alphabets \( \{1, 2, \ldots, r\} \) are given by
Three substitutions (nrs. 2, 3 and 7) in the Thue-Morse factor list generate systems that are certainly not conjugate to the Thue-Morse system, as they are in the Toeplitz conjugacy class. To see whether the other 12 yield systems conjugate to the Thue-Morse system, according to Theorem 9.1 we would have to construct the factor list of each of these 12. This is quite involved, for example the 3-block presentations of the two factors on 5 symbols have 11 symbols.

However, there is a quicker way to determine whether these factors are conjugate to the Thue-Morse systems, by finding explicit semi-conjugacies from these factors to the Thue-Morse system. Then by coalescence the systems are conjugate.

The 6 substitutions associated to the partition \{1\}{2}\{3\}{4}\{5\}{6} generate conjugate systems, since the 3-block presentation of any system is conjugate to the original system (this also uses Theorem 6.1). So we only have to examine the substitutions on 3 and on 5 symbols. For the partition \{1, 4, 5\}{2, 3}\{6\} we consider the letter-to-letter map \(\pi_2\) given by

\[
\{1, 4, 5\} \rightarrow 0, \quad \{2, 3\} \rightarrow 1, \quad \{6\} \rightarrow 1.
\]

Here we use the notation \(\pi_2\), since \(\pi_2\) corresponds to projection on the second coordinate in the 3-block presentation. It is easily checked that the substitution on

| Nr. | Partition | \(M\) | Substitution |
|-----|-----------|------|-------------|
| 1.  | \{1, 3, 5\}{2, 4, 6} | 0    | 1 \(\rightarrow\) 12, 2 \(\rightarrow\) 21 |
| 2.  | \{1, 3, 5\}{2, 4, 6} | 0    | 1 \(\rightarrow\) 21, 2 \(\rightarrow\) 12 |
| 3.  | \{1, 6\}{2, 3, 4, 5} | 1    | 1 \(\rightarrow\) 21, 2 \(\rightarrow\) 11 |
| 4.  | \{3, 4\}{1, 2, 5, 6} | 1    | 1 \(\rightarrow\) 12, 2 \(\rightarrow\) 11 |
| 5.  | \{1, 4, 5\}{2, 3}\{6\} | 1    | 1 \(\rightarrow\) 12, 2 \(\rightarrow\) 31, 3 \(\rightarrow\) 21 |
| 6.  | \{1, 4, 5\}{2, 6}\{3\} | 1    | 1 \(\rightarrow\) 21, 2 \(\rightarrow\) 13, 3 \(\rightarrow\) 12 |
| 7.  | \{1, 6\}{2, 5}\{3, 4\} | 1    | 1 \(\rightarrow\) 23, 2 \(\rightarrow\) 13, 3 \(\rightarrow\) 12 |
| 8.  | \{1\}{2}\{3\}{4}\{5\}\{6\} | 0    | 1 \(\rightarrow\) 12, 2 \(\rightarrow\) 31, 3 \(\rightarrow\) 34, 4 \(\rightarrow\) 13 |
| 9.  | \{1\}{2}\{3\}{4}\{5\}\{6\} | 0    | 1 \(\rightarrow\) 21, 2 \(\rightarrow\) 13, 3 \(\rightarrow\) 43, 4 \(\rightarrow\) 31 |
| 10. | \{1\}{2}\{3\}{4}\{5\}\{6\} | 0    | 1 \(\rightarrow\) 23, 2 \(\rightarrow\) 14, 3 \(\rightarrow\) 21, 4 \(\rightarrow\) 12 |
| 11. | \{1\}{2}\{3\}{4}\{5\}\{6\} | 0    | 1 \(\rightarrow\) 23, 2 \(\rightarrow\) 13, 3 \(\rightarrow\) 41, 4 \(\rightarrow\) 31 |
| 12. | \{1, 5\}{2\{3\}{4}\{5\}\{6\} | 1    | 1 \(\rightarrow\) 12, 2 \(\rightarrow\) 31, 3 \(\rightarrow\) 45, 4 \(\rightarrow\) 35, 5 \(\rightarrow\) 14 |
| 13. | \{1\}{2, 3}\{4\}{5\} | 1    | 1 \(\rightarrow\) 21, 2 \(\rightarrow\) 13, 3 \(\rightarrow\) 45, 4 \(\rightarrow\) 51, 5 \(\rightarrow\) 43 |
| 14. | \{1\}{2\{3\}{4\}{5}\{6\} | 1    | 1 \(\rightarrow\) 23, 2 \(\rightarrow\) 14, 3 \(\rightarrow\) 21, 4 \(\rightarrow\) 56, 5 \(\rightarrow\) 63, 6 \(\rightarrow\) 54 |
| 15. | \{1\}{2\{3\}{4\}{5\} | 1    | 1 \(\rightarrow\) 23, 2 \(\rightarrow\) 13, 3 \(\rightarrow\) 41, 4 \(\rightarrow\) 56, 5 \(\rightarrow\) 46, 6 \(\rightarrow\) 25 |
\[
\{1, 4, 5\} \times \{2, 3\} \times \{6\},
\]
\(\pi_2\) and the Thue-Morse substitution satisfy the intertwining equation of Section 3, hence \(\pi_2\) is a semi-conjugacy from the system on three symbols \(\{1, 4, 5\}, \{2, 3\}\) and \(\{6\}\) to the Thue-Morse system. In the same way Nrs. 6, 12 and 13 can be treated.

Conclusion: there are 12 primitive injective substitutions of length 2 that generate a system conjugate to the Thue-Morse dynamical system.

13. Epilogue

An interesting extension of our result would be to consider also non-constant length substitutions. For example, let \(\theta\) be the ternary Thue-Morse substitution, defined by

\[
\theta(1) = 123, \quad \theta(2) = 13, \quad \theta(3) = 2.
\]

An application of Theorem 1 in Section V of [7] shows that \((X_\theta, \sigma)\) is conjugate to a substitution of constant length 2 on 6 symbols. Its canonical injective version is a substitution on 5 symbols, and taking the standard form of this substitution we find that it is on the Thue-Morse list.

In the recent preprint [15] a first step in this direction is taken: it shown there that modulo powers of the shift there are only finitely many conjugacies between the systems generated by two primitive substitutions whose matrices have the same Perron-Frobenius eigenvalue.

Primitive substitutions generate dynamical systems with unique shift invariant measure. One can consider Problem 1.2 for measure-theoretic conjugacy. We believe that the equivalence classes will not change. This has been proved for a subclass of constant length substitutions in [12].

14. Acknowledgement

We have profited from electronic discussions on coalescence with Fabien Durand and Reem Yassawi.

References

[1] François Blanchard et al (2004) Constant-length substitutions and countable scrambled sets Nonlinearity 17: 817-833.
[2] Ethan M. Coven and Michael S. Keane (1971). The structure of substitution minimal sets Trans. Amer. Math. Soc 162: 89-102.
[3] Ethan M. Coven, Michael S. Keane and Michelle LeMasurier (2008). A characterization of the Morse minimal set up to topological conjugacy Ergodic Theory and Dynamical Systems 28, 1443–1451.
[4] Ethan M. Coven, Andrew Dykstra, Michael S. Keane and Michelle LeMasurier (2013). Topological conjugacy to given constant length substitution minimal systems \(\text{arXiv:1209.2073v2}\).
[5] Ethan M. Coven, Andrew Dykstra, and Michelle LeMasurier (2014). A short proof of a theorem of Cobham on substitutions Rocky Mountain J. Math, to appear.
[6] Ethan M. Coven (1971). Endomorphisms of substitution minimal sets Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 20, 129–133.
[7] F.M. Dekking (1978). The spectrum of dynamical systems arising from substitution of constant length Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 41, 221–239.
[8] Fabien Durand (2000). Linearly recurrent subshifts have a finite number of nonperiodic factors Ergodic Theory and Dynamical Systems 20, 1061–1078.
[9] Fabien Durand (2003). Corrigendum and addendum to ‘Linearly recurrent subshifts have a finite number of nonperiodic factors’ Ergodic Theory and Dynamical Systems 23, 663–669.
[10] Fabien Durand (2008). Cobham-Semenov theorem and \( \mathbb{N}^d \)-subshifts Theoretical Computer Science 391, 20–38.
[11] G.A. Hedlund (1969). Endomorphisms and automorphisms of the shift dynamical system Mathematical Systems Theory 3, 320–375.
[12] Bernard Host and François Parreau (1989). Homomorphismes entre systèmes dynamiques définis par substitutions Ergodic Theory and Dynamical Systems 9, 469–477.
[13] Brigitte Mossé (1996). Reconnaissabilité des substitutions et complexité des suites automatiques Bull. Soc. Math. Fr. 124, 329–346.
[14] Martine Queffélec (2010). Substitution Dynamical Systems - Spectral Analysis Lecture Notes in Mathematics 1294, 2nd ed., Springer.
[15] Ville Salo and Ilkka Törmä (2013). Block Maps between Primitive Uniform and Pisot Substitutions arXiv:1306.3777.

Wesleyan University and Delft University of Technology.