#A112 INTEGERS 24 (2024)

DUMONT-THOMAS COMPLEMENT NUMERATION SYSTEMS FOR $\mathbb{Z}$

Sébastien Labbé
Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, Talence, France
sebastien.labbe@labri.fr

Jana Lepšová
FNSPE, CTU in Prague, Praha, Czech Republic
jana.lepsova@labri.fr

Received: 5/22/23, Revised: 5/21/24, Accepted: 11/27/24, Published: 12/9/24

Abstract

We extend the well-known Dumont-Thomas numeration systems to $\mathbb{Z}$ using an approach inspired by the two's complement numeration system. Integers in $\mathbb{Z}$ are canonically represented by a finite word (starting with 0 when nonnegative and with 1 when negative). The systems are based on two-sided periodic points of substitutions as opposed to the right-sided fixed points. For every periodic point of a substitution, we construct an automaton which returns the letter at position $n \in \mathbb{Z}$ of the periodic point when fed with the representation of n in the corresponding numeration system. The numeration system naturally extends to $\mathbb{Z}^d$. We give an equivalent characterization of the numeration system in terms of a total order on a regular language. Lastly, using particular periodic points, we recover the well-known two's complement numeration system and the Fibonacci analogue of the two's complement numeration system.

1. Introduction

On a finite size memory representing unsigned integers with base-10 digits, incrementing by 1 the largest representable number gives

$$\begin{array}{r} \texttt{9999999999999} \\ +1 \\ \hline \texttt{0000000000000} \end{array}$$

if we ignore the overflow error caused by the propagation of the carry beyond the memory limit. Therefore, it makes sense to identify the number $\texttt{999}\cdots\texttt{9}$ with the

DOI: 10.5281/zenodo.14340125

value -1 since adding one to it gives zero. Likewise, $\texttt{999}\cdots\texttt{98}$ can be identified with the value -2, $\texttt{999}\cdots\texttt{97}$ with the value -3, and so on, just like negative p-adic integers [17]. This numeration system is called ten's-complement. As instructively explained by Knuth [22, Section 4.1], the same can be done in an arbitrary integer base $b \geq 2$. When $b = 2$, it is called the two's complement numeration system. This system is still used nowadays to represent signed integers in the architecture of modern processors [21, Section 4.2.1] due to its efficiency at performing arithmetic operations.

In this article, we show that the concept of complement numeration systems goes beyond numeration systems in an integer base. The theory of numeration systems studies and describes the various ways of representing numbers (integers, real numbers, Gaussian integers, etc.) by sequences of digits [14, 13, 15, 5, 20, 4, 31]. One of these ways gives rise to the numeration systems based on substitutions which were proposed by Dumont and Thomas [10]. The Dumont-Thomas numeration system associated with a substitution provides a canonical representation for every nonnegative integer. It may also be used to represent real numbers in a certain interval. It turns out there exists a natural complement version of the Dumont-Thomas numeration systems allowing to represent all integers in $\mathbb{Z}$ and not only those that are nonnegative.

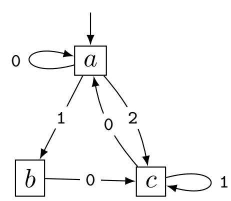
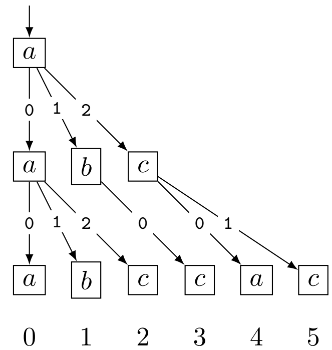


Figure 1: The graph associated with the substitution $a \mapsto abc, b \mapsto c, c \mapsto ac$.

In practical terms, the Dumont-Thomas numeration system can be defined by the set of finite paths in a directed graph starting from some fixed vertex. For example, consider the directed graph shown in Figure 1 with vertices a, b and c where the outgoing edges of every vertex are labeled with consecutive nonnegative integers starting with zero. The set of paths of fixed length starting with some chosen vertex can be unfolded into a tree; see Figure 2 (left). A path in the tree is uniquely identified with the sequence of labels of its edges starting from the root. Among the set of paths of a given length ordered lexicographically, the n-th one can be regarded as a representation of the nonnegative integer n. Considering arbitrarily long finite paths starting from the initial vertex in the directed graph, we obtain a canonical representation of all nonnegative integers after removing leading zeros in their representation (assuming the initial vertex has a loop labeled with $\texttt{0}$); see

Figure 2 (right). We refer to such a numeration system as the *Dumont-Thomas*



n	path	representation of n
0	00	ε
1	01	1
2	02	2
3	10	10
4	20	20
5	21	21

Figure 2: The set of paths starting in state a in the directed graph provide a canonical representation of the nonnegative integers after removing leading zeroes.

numeration system for $\mathbb{N}$ associated with the substitution $\mu : a \mapsto abc, b \mapsto c, c \mapsto ac$. The directed graph shown in Figure 2 (as well as the automaton shown in Figure 1) is derived from the substitution μ following a well-known construction for automatic sequences [1]: $\alpha \xrightarrow{i} \beta$ is an edge of the graph if and only if β is the i-th letter of the image of the letter α, for every integer i such that $0 \leq i < \ell$ where ℓ is the length of the image of the letter α. Among other properties, this numeration system gives a direct description of the right-infinite fixed point $\mathbf{t} = \mu(\mathbf{t}) = abccacacabc\ldots$ of the substitution μ as an automatic sequence.

The Dumont-Thomas numeration systems were later explained using the so-called prefix-suffix automata associated with primitive substitutions [6] by considering the cylinders of finite length words [6, Corollary 6.2]. The main motivation of Canterini and Siegel was to prove that every dynamical system generated by a substitution of Pisot type on d letters admits a minimal translation on the torus $\mathbb{T}^{d-1}$ as a topological factor [7]. As a consequence, they obtained a numeration system representing the elements of $\mathbb{T}^{d-1}$ by infinite paths in a prefix-suffix automaton; see [12].

In more generality, every regular language over a totally ordered alphabet leads to what is called an *abstract numeration system*, which may be used to represent nonnegative integers [27] or real numbers in an interval [26]; see also [28, Section 7], [16, Section 4], [3], [4, Section 3], [8].

Recently, a very general framework was proposed to extend the Dumont-Thomas numeration systems to all integers based on the notion of coding prescription, which allows the image of letters to be scattered words of nonconsecutive letters [33]. Another recent article extending these numeration systems concerns also the β-numeration of real numbers in an interval [34].

The extension of the Dumont-Thomas numeration systems to all integers in $\mathbb{Z}$

that we propose is inspired by integer base complement numeration systems; see Definition 4.3. It is derived from the two-sided periodic points of substitutions as opposed to the right-infinite fixed points. In a Dumont-Thomas complement numeration system, the representations of nonnegative integers start with the digit 0 whereas the representations of negative integers start with the digit 1. The proofs provided here follow as much as possible the approach originally proposed by Dumont and Thomas [10].

The main results of this contribution are Theorem 6.1, where we prove that two-sided periodic points of substitutions are automatic sequences with respect to the Dumont-Thomas complement numeration systems for $\mathbb{Z}$, and Theorem 8.4, where we characterize these numeration systems by means of a total order on the language recognized by an automaton. Finally, we show that the well-known two's complement numeration system can be constructed as a Dumont-Thomas complement numeration system for $\mathbb{Z}$ (Proposition 9.1) and similarly for the Fibonacci analogue of the two's complement numeration system (Proposition 9.2).

Also, we extend the Dumont-Thomas complement numeration systems to $\mathbb{Z}^d$; see Definition 7.2. The need for extending the theory of numeration systems based on substitutions from $\mathbb{N}$ to $\mathbb{Z}$ and to $\mathbb{Z}^d$ for $d \geq 2$ was motivated by the study of aperiodic Wang tilings of the plane. In [25], configurations in a particular aperiodic Wang shift based on 16 Wang tiles were described by an automaton derived from a two-dimensional substitution. The automaton takes as input the representation of a position in $\mathbb{Z}^2$ using a Fibonacci analogue of the two's complement numeration system and outputs the index of the Wang tile to place at this position. This example belongs to a family of the Dumont-Thomas complement numeration systems for $\mathbb{Z}^2$.

The authors believe further extensions beyond Dumont-Thomas based on a single substitution can be expected including S-adic sequences [11]. For instance, in a Bratteli–Vershik diagram [19, 18, 29], one may think of the maximal path in the diagram as a representation of -1 and the minimal path as a representation of 0. The representation of the other negative and nonnegative integers can be deduced from the order of a Bratteli–Vershik diagram and its natural successor map.

The article is structured as follows. Preliminaries and notation are presented in Section 2. Section 3 recalls numeration systems for $\mathbb{N}$ defined by Dumont and Thomas and presents some extensions of their results. In Section 4, we extend a theorem of Dumont and Thomas to the right-infinite and left-infinite periodic points of substitutions. We use it to define numeration systems for $\mathbb{Z}$ based on the two-sided periodic points of substitutions. In Section 5, we show some examples. In Section 6, we describe periodic points of substitutions as automatic sequences. In Section 7, we show how to extend the Dumont-Thomas numeration systems to $\mathbb{Z}^d$. In Section 8, we present a total order on $\{\mathtt{0},\mathtt{1}\} \odot \mathcal{D}^*$, where $\mathcal{D}$ is some alphabet of integers and $\odot$ is the concatenation of words within the monoid $\mathcal{D}^*$. We characterize the Dumont-Thomas complement numeration systems for $\mathbb{Z}$ with respect to this total order. In Section 9, we show that the well-known two's complement numeration system is an instance of a Dumont-Thomas numeration system for $\mathbb{Z}$ and similarly for the Fibonacci analogue of the two's complement numeration system.

2. Preliminaries

An *alphabet* A is a finite set and its elements $a \in A$ are called *letters*. A *finite word* $u = u_0 u_1 \cdots u_{n-1}$ is a concatenation of letters $u_i \in A$ for every $i \in \{0, 1, \ldots, n-1\}$ and $|u|$ denotes its *length*. When it is more convenient, we denote the i-th letter of u by $u[i]$ instead of u_i. The *empty word* is denoted by ε. The set of all finite words over the alphabet A is denoted by A^* and the set of all nonempty words over the alphabet A is denoted by $A^+ = A^* \setminus \{\varepsilon\}$. We define the concatenation $\odot$ as the following binary operation:

$$\odot : A^* \times A^* \to A^*, u \odot v \mapsto uv.$$

The set A^* with the concatenation as operation forms a monoid with ε as the neutral element.

A *morphism* over A is a map $\eta : A^* \to A^*$ such that $\eta(u \odot v) = \eta(u) \odot \eta(v)$ for all words $u, v \in A^*$. A *substitution* $\eta : A^* \to A^*$ is a morphism such that $\eta(a) \in A^+$ is nonempty for every $a \in A$ and there exists $a \in A$ such that a is *growing*, that is, $\lim_{k \to +\infty} |\eta^k(a)| = +\infty$. A morphism η is said *primitive* if there exists $k \in \mathbb{N}$ such that for every $a, b \in A$ the letter a appears in $\eta^k(b)$. A morphism η is said d-*uniform* for some nonnegative integer d if $|\eta(a)| = d$ for every letter $a \in A$.

We call $u_0 u_1 u_2 \cdots \in A^{\mathbb{Z}_{\geq 0}}$ a *right-infinite word* and $\cdots u_{-3} u_{-2} u_{-1} \in A^{\mathbb{Z}_{<0}}$ a *left-infinite word*. We call $u \in A^{\mathbb{Z}}$ a *two-sided word* and we separate by a vertical bar its elements u_{-1} and u_0 to indicate the origin, i.e., $u = \cdots u_{-3} u_{-2} u_{-1} | u_0 u_1 u_2 \cdots$.

Substitutions can be applied naturally to two-sided words $u \in A^{\mathbb{Z}}$ by setting

$$\eta(\ldots u_{-3} u_{-2} u_{-1} | u_0 u_1 u_2 \cdots) = \cdots \eta(u_{-3}) \eta(u_{-2}) \eta(u_{-1}) | \eta(u_0) \eta(u_1) \eta(u_2) \cdots .$$

Let $\mathbb{D} \in \{\mathbb{Z}, \mathbb{Z}_{\geq 0}, \mathbb{Z}_{<0}\}$. A word $u \in A^{\mathbb{D}}$ is called a *periodic point* of the substitution η if there exists an integer $p \geq 1$ such that $\eta^p(u) = u$, and in this case, p is called a *period* of the periodic point. The minimum integer $p \geq 1$ such that $\eta^p(u) = u$ is called *the* period of u. A periodic point with period $p = 1$ is called a *fixed point* of η. The set of periodic points of η is denoted by $\mathrm{Per}_{\mathbb{D}}(\eta) = \{u \in A^{\mathbb{D}} \mid \eta^p(u) = u \text{ for some } p \geq 1\}$. Since we are mostly interested in two-sided words in this contribution, we omit the domain when $\mathbb{D} = \mathbb{Z}$ and we write $\mathrm{Per}(\eta) = \mathrm{Per}_{\mathbb{Z}}(\eta)$.

If $u \in \mathrm{Per}(\eta)$ is a two-sided periodic point of a substitution η, then we say that the pair of letters $u_{-1} | u_0$ is the *seed* of u; see [2, Section 4.1]. If the seed letters of a two-sided periodic point are growing, then the periodic point is defined entirely by its seed. More precisely, $u = \lim_{k \to +\infty} \eta^{pk}(u_{-1}) | \eta^{pk}(u_0)$, where p is a period of u.

Let $u = u_0u_1 \cdots \in \mathrm{Per}_{\mathbb{N}}(\eta)$ with a seed $u_0 = a$. The previous terminology is inspired by [6], where a prefix-suffix automaton is associated with η. However, for our goal an automaton associated with η as in [4] is sufficient. Let $\mathcal{D}$ denote the alphabet $\mathcal{D} = \{\mathtt{0}, ..., \max_{c\in A} |\eta(c)| - 1\}$ whose elements are integers. The set $\mathcal{D}^*$ is a monoid for the operation $\odot$ of concatenation. The deterministic finite automaton with output (DFAO) associated with the substitution η and letter a is the 5-tuple $\mathcal{A}_{\eta,a} = (A, \mathcal{D}, \delta, a, A)$, where the transition function $\delta : A \times \mathcal{D} \to A$ is a partial function such that $\delta(b, i) = c$ if and only if $c = w_i$ and $\eta(b) = w_0 \dots w_{|\eta(b)|-1}$. The transition function δ is naturally extended to $A \times \mathcal{D}^*$ by $\delta(b, \varepsilon) = b$ for every $b \in A$, and, for every $b \in A$, $i \in \mathcal{D}$ and $w \in \mathcal{D}^*$, $\delta(b, i \odot w) = \delta(\delta(b, i), w)$. For some state $b \in A$ and word $w \in \mathcal{D}^*$, we let $\mathcal{A}_{\eta,a}(b, w)$ denote $\delta(b, w)$. In particular, we let $\mathcal{A}_{\eta,a}(w)$ denote $\delta(a, w)$. We let $\mathcal{L}(\mathcal{A}_{\eta,a})$ denote the set of words accepted by the automaton $\mathcal{A}_{\eta,a}$. Finally, for every $q \in \mathbb{N}$, we let $\mathcal{L}_q(\mathcal{A}_{\eta,a})$ denote the set of words in $\mathcal{L}(\mathcal{A}_{\eta,a})$ of length q.

3. Dumont-Thomas Numeration System for $\mathbb{N}$

In this section, we recall Dumont-Thomas numeration system for $\mathbb{N}$, which was based on substitutions having a right-infinite fixed point [10]. It uses the definition of admissible sequences.

Definition 3.1 (admissible sequence [10])**.** Let $\eta : A^* \to A^*$ be a substitution. Let $a \in A$ be a letter, k an integer and, for each integer i, $0 \le i \le k$, (m_i, a_i) be an element of $A^* \times A$. We say that the finite sequence $(m_i, a_i)_{i=0,\dots,k}$ is *admissible with respect to* η if and only if, for all i, $1 \le i \le k$, $m_{i-1}a_{i-1}$ is a prefix of $\eta(a_i)$. We say that this sequence is *a-admissible with respect to* η if it is admissible with respect to η and, moreover, $m_k a_k$ is a prefix of $\eta(a)$.

As done in [10], when the substitution is clear from the context, we write that a sequence is *admissible* or *a-admissible* without specifying the substitution.

Dumont and Thomas proved the following result, which we rewrite in our notation.

Theorem 3.2 ([10, Theorem 1.5])**.** *Let $a \in A$ and let $\eta : A^* \to A^*$ be a substitution. Let $u = \eta(u)$ be a right-infinite fixed point of η with growing seed $u_0 = a$. For every integer $n \ge 1$, there exists a unique integer $k = k(n)$ and a unique sequence $(m_i, a_i)_{i=0,\dots,k}$ such that*

- *this sequence is a-admissible and $m_k \ne \varepsilon$,*
- $u_0u_1 \cdots u_{n-1} = \eta^k(m_k)\eta^{k-1}(m_{k-1}) \cdots \eta^0(m_0)$.

The proof of the above theorem was based on the following lemmas, which we cite here as we need them in what follows.

Lemma 3.3 ([10, Lemma 1.1])**.** *Let $\eta : A^* \to A^*$ be a substitution and $k \geq 0$ be an integer. If $(m_i, a_i)_{i=0,\ldots,k}$ is an admissible sequence, then*

$$\sum_{j=0}^{k} |\eta^j(m_j)| < |\eta^k(m_k a_k)|.$$

Lemma 3.4 ([10, Lemma 1.3])**.** *Let $\eta : A^* \to A^*$ be a substitution and $k \geq 0$ be an integer. Let $b \in A$, $(m_i, a_i)_{i=0,\ldots,k}$ and $(m'_i, a'_i)_{i=0,\ldots,k}$ be two b-admissible sequences and n be an integer such that*

$$n = \sum_{j=0}^{k} |\eta^j(m_j)| = \sum_{j=0}^{k} |\eta^j(m'_j)|.$$

Then for every i, $0 \leq i \leq k$, we have $(m_i, a_i) = (m'_i, a'_i)$.

Lemma 3.5 ([10, Lemma 1.4])**.** *Let $\eta \colon A^* \to A^*$ be a substitution. Let $\ell \geq 1$ be an integer, $a \in A$ a letter and $m \in A^*$ a proper prefix of the word $\eta^\ell(a)$. Then there exist $(m', a') \in A^* \times A$ and $m'' \in A^*$ such that $m'a'$ is a prefix of $\eta(a)$, m'' is a proper prefix of $\eta^{\ell-1}(a')$ and $m = \eta^{\ell-1}(m')m''$.*

3.1. Some Extensions of Dumont-Thomas Results

In this subsection, we propose some extensions of Dumont-Thomas lemmas. Firstly, we observe that admissible sequences are related to automata as follows.

Lemma 3.6. *Let $\eta : A^* \to A^*$ be a substitution, $k \geq 1$ be an integer and $x \in A$. If $(m_i, a_i)_{i=0,\ldots,k-1}$ is an x-admissible sequence, then*

$$a_i = \mathcal{A}_{\eta,x}(|m_{k-1}| \odot |m_{k-2}| \odot \ldots \odot |m_i|) \qquad \text{for every } i = 0, \ldots, k-1.$$

Remark 3.7. The notation $(\odot)$ in the above equation and the proof that follows stands for the concatenation of words within the monoid $\mathcal{D}^*$. Since the elements of $\mathcal{D}$ are integers, we write this notation explicitly to avoid misinterpreting it with the multiplication of integers.

Proof. The proof is carried out by induction on i. If $i = k-1$, then $a_i = a_{k-1} = \eta(x)[|m_{k-1}|] = \mathcal{A}_{\eta,x}(|m_{k-1}|)$. If $i < k-1$, then

$$\begin{aligned} a_i &= \eta(a_{i+1})[|m_i|] = \eta\left(\mathcal{A}_{\eta,x}(|m_{k-1}| \odot \ldots \odot |m_{i+1}|)\right)[|m_i|] \\ &= \mathcal{A}_{\eta,x}(|m_{k-1}| \odot \ldots \odot |m_{i+1}| \odot |m_i|). \end{aligned}$$

□

Lemma 3.8. *Let $\eta : A^* \to A^*$ be a substitution, $k \geq 1$ be an integer and $x \in A$. If $v_{k-1}v_{k-2}\cdots v_0 \in \mathcal{L}(\mathcal{A}_{\eta,x})$, then there exists an x-admissible sequence $(m_i, a_i)_{i=0,\ldots,k-1}$ such that $|m_i| = v_i$ for every $i = 0, \ldots, k-1$.*

Proof. We carry out the proof by induction on k. If $k = 1$, then $v_0 \in \mathcal{L}(\mathcal{A}_{\eta,x})$ implies that $0 \leq v_0 < |\eta(x)|$. Denote m_0 the proper prefix of $\eta(x)$ of length v_0 and let $a_0 \in A$ so that $m_0 a_0$ is a prefix of $\eta(x)$. The length-1 sequence $(m_i, a_i)_{i=0}$ is x-admissible and satisfies the condition that $|m_0| = v_0$.

Assume that $k \geq 1$ is an integer such that for every word $w_{k-1}w_{k-2}\cdots w_0 \in \mathcal{L}(\mathcal{A}_{\eta,x})$ of length k, there exists an x-admissible sequence $(m_i, a_i)_{i=0,\ldots,k-1}$ such that $|m_i| = w_i$ for every $i = 0, \ldots, k-1$. Let $v_k v_{k-1} \cdots v_0 \in \mathcal{L}(\mathcal{A}_{\eta,x})$. Then from the induction hypothesis applied on $v_k v_{k-1} \cdots v_1 \in \mathcal{L}(\mathcal{A}_{\eta,x})$, which is of length k, we have an x-admissible sequence $(m_i, a_i)_{i=1,\ldots,k}$ such that $|m_i| = v_i$ for every $i = 1, \ldots, k$. We have from Lemma 3.6 that $a_1 = \mathcal{A}_{\eta,x}(v_k v_{k-1} \cdots v_1)$ and we have from the definition of the automaton $\mathcal{A}_{\eta,x}$ that $v_0 < |\eta(a_1)|$. Denote m_0 the proper prefix of $\eta(a_1)$ of length v_0 and let $a_0 \in A$ so that $m_0 a_0$ is a prefix of $\eta(a_1)$. Then $|m_0| = v_0$ and $(m_i, a_i)_{i=0,\ldots,k}$ is an x-admissible sequence. □

Lemma 3.5 can be used to construct an admissible sequence from a prefix of the image of a letter under the p-th power of a substitution.

Lemma 3.9. *Let $\eta : A^* \to A^*$ be a substitution and $p \geq 1$ be an integer. If $m \in A^*$ and $x \in A$ are such that m is a proper prefix of $\eta^p(x)$, then there exists a unique x-admissible sequence $(m_i, a_i)_{i=0,\ldots,p-1}$ such that*

$$|m| = \textstyle\sum_{j=0}^{p-1} |\eta^j(m_j)|. \tag{1}$$

Moreover, $m = \eta^{p-1}(m_{p-1})\eta^{p-2}(m_{p-2})\cdots\eta^0(m_0)$.

Proof. (Uniqueness) Let $(m_i, a_i)_{i=0,\ldots,p-1}$ and $(m'_i, a'_i)_{i=0,\ldots,p-1}$ be two x-admissible sequences satisfying the hypothesis. Then

$$\textstyle\sum_{j=0}^{p-1} |\eta^j(m_j)| = |m| = \sum_{j=0}^{p-1} |\eta^j(m'_j)|.$$

By Lemma 3.4, $(m_i, a_i)_{i=0,\ldots,p-1} = (m'_i, a'_i)_{i=0,\ldots,p-1}$.

(Existence) We carry out the proof by induction on p. If $p = 1$, then m is a proper prefix of $\eta(x)$. Let $m_0 = m$ and $a_0 \in A$ be such that ma_0 is a prefix of $\eta(x)$. The length-1 sequence $(m_i, a_i)_{i=0}$ is x-admissible and satisfies the condition that $m = \eta^0(m_0)$.

Now let $m \in A^*$ and $x \in A$ be such that m is a proper prefix of $\eta^{p+1}(x)$. From Lemma 3.5, there exist $(m_p, a_p) \in A^* \times A$ and $m'' \in A^*$ such that $m_p a_p$ is a prefix of $\eta(x)$, m'' is a proper prefix of $\eta^p(a_p)$ and $m = \eta^p(m_p)m''$. By the induction hypothesis, there exists an a_p-admissible sequence $(m_i, a_i)_{i=0,\ldots,p-1}$ such that

$$m'' = \eta^{p-1}(m_{p-1})\eta^{p-2}(m_{p-2})\cdots\eta^0(m_0).$$

Therefore,

$$m = \eta^p(m_p)m'' = \eta^p(m_p)\eta^{p-1}(m_{p-1})\eta^{p-2}(m_{p-2})\cdots\eta^0(m_0).$$

The extended sequence $(m_i, a_i)_{i=0,\ldots,p}$ is x-admissible since $m_{p-1}a_{p-1}$ is a prefix of $\eta(a_p)$ and $m_p a_p$ is a prefix of $\eta(x)$. □

Let $\eta : A^* \to A^*$ be a substitution and $\mathcal{D} = \{0, ..., \max_{c \in A} |\eta(c)| - 1\}$. Lemma 3.9 allows us to define a map for every integer $p \geq 1$ and $x \in A$ as follows

$$\begin{array}{llll} \mathrm{tail}_{\eta,p,x} : & \{0, 1, \ldots, |\eta^p(x)| - 1\} & \to & \mathcal{D}^p \\ & n & \mapsto & |m_{p-1}| \odot |m_{p-2}| \odot \ldots \odot |m_0|, \end{array}$$

where $(m_i, a_i)_{i=0,\ldots,p-1}$ is the unique x-admissible sequence satisfying Equation (1) with m being the prefix of length n of $\eta^p(x)$. The map $\mathrm{tail}_{\eta,p,x}$ will be used in Section 4.

Example 3.10. Consider the Tribonacci substitution $\psi_T : a \mapsto ab, b \mapsto ac, c \mapsto a$ [30]. The successive images of a under the substitution ψ_T are illustrated as a tree in Figure 3. The path from the root of the tree to a node of depth p at x-position

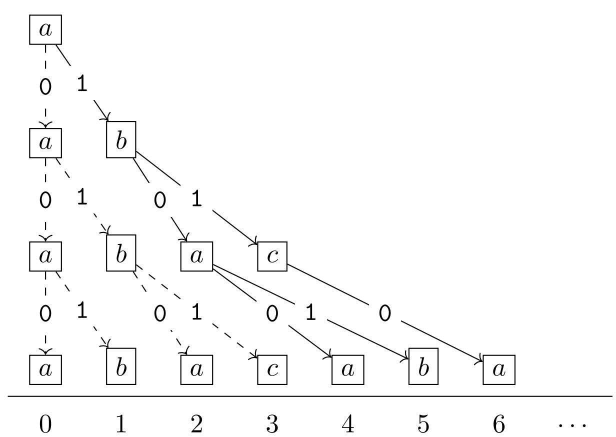


Figure 3: The successive images of the letter a under the Tribonacci substitution.

$n \in \mathbb{N}$ is labeled by $\mathrm{tail}_{\psi_T,p,a}(n)$. Their values are illustrated in Table 1.

Lemma 3.11. *Let $\eta : A^* \to A^*$ be a substitution and $p \geq 1$ be an integer. Let $x \in A$. Then for every $\ell \in \{0, 1, \ldots, |\eta^p(x)| - 1\}$ we have*

$$\eta^p(x)[\ell] = \mathcal{A}_{\eta,x}(\mathrm{tail}_{\eta,p,x}(\ell)).$$

Proof. Let m be the prefix of $\eta^p(x)$ of length ℓ. From Lemma 3.9, there exists a unique x-admissible sequence $(m_i, a_i)_{i=0,\ldots,p-1}$ such that

$$m = \eta^{p-1}(m_{p-1})\eta^{p-2}(m_{p-2}) \cdots \eta^0(m_0).$$

n	$\mathrm{tail}_{\psi_T,1,a}(n)$	$\mathrm{tail}_{\psi_T,2,a}(n)$	$\mathrm{tail}_{\psi_T,3,a}(n)$
0	0	00	000
1	1	01	001
2		10	010
3		11	011
4			100
5			101
6			110

Table 1: The tail map for the Tribonacci substitution ψ_T of depth $p \in \{1, 2, 3\}$ associated with letter $x = a$ for integers n from 0 to 6.

The word ma_0 is a prefix of $\eta^p(x)$, thus $\eta^p(x)[\ell] = a_0$. From Lemma 3.6,

$$\eta^p(x)[\ell] = a_0 = \mathcal{A}_{\eta,x}(|m_{p-1}| \odot |m_{p-2}| \odot \ldots \odot |m_0|) = \mathcal{A}_{\eta,x}(\mathrm{tail}_{\eta,p,x}(\ell)).$$

□

In the next lemma, we consider the total order $(\mathcal{D}^*, <_{lex})$, where $u <_{lex} v$ means that u is lexicographically less than v. Recall that given a totally ordered set $(\mathcal{D}, <)$, and two words $u, v \in \mathcal{D}^*$ such that v is nonempty, then one has that u is lexicographically less than v, if u is a proper prefix of v, or there exist words $r, s, t \in \mathcal{D}^*$ and letters $a, b \in \mathcal{D}$ such that $u = ras$ and $v = rbt$ with $a < b$.

Lemma 3.12. *Let $\eta : A^* \to A^*$ be a substitution and $p \geq 1$ be an integer. Let $n, n' \in \{0, 1, \ldots, |\eta^p(x)| - 1\}$. Then*

(i) $n = n'$ if and only if $\mathrm{tail}_{\eta,p,x}(n) = \mathrm{tail}_{\eta,p,x}(n')$,

(ii) $n < n'$ if and only if $\mathrm{tail}_{\eta,p,x}(n) <_{lex} \mathrm{tail}_{\eta,p,x}(n')$.

Proof. Let $(m_i, a_i)_{i=0,\ldots,p-1}$ and $(m'_i, a'_i)_{i=0,\ldots,p-1}$ be two x-admissible sequences such that $n = \sum_{j=0}^{p-1} |\eta^j(m_j)|$ and $n' = \sum_{j=0}^{p-1} |\eta^j(m'_j)|$. Thus $\mathrm{tail}_{\eta,p,x}(n) = |m_{p-1}| \odot |m_{p-2}| \odot \ldots \odot |m_0|$ and $\mathrm{tail}_{\eta,p,x}(n') = |m'_{p-1}| \odot |m'_{p-2}| \odot \ldots \odot |m'_0|$.

(i) If $n = n'$, then $\mathrm{tail}_{\eta,p,x}(n) = \mathrm{tail}_{\eta,p,x}(n')$. Conversely, if $\mathrm{tail}_{\eta,p,x}(n) = \mathrm{tail}_{\eta,p,x}(n')$, then $m_{p-1}a_{p-1} = m'_{p-1}a'_{p-1}$ since both are prefixes of the same length of $\eta(x)$. Thus $m_{p-1} = m'_{p-1}$ and $a_{p-1} = a'_{p-1}$. Similarly, $m_{p-2}a_{p-2} = m'_{p-2}a'_{p-2}$ since both are prefixes of the same length of $\eta(a_{p-1})$. Thus $m_{p-2} = m'_{p-2}$ and $a_{p-2} = a'_{p-2}$. By induction, we obtain $(m_i, a_i)_{i=0,\ldots,p-1} = (m'_i, a'_i)_{i=0,\ldots,p-1}$. Thus $n = \sum_{j=0}^{p-1} |\eta^j(m_j)| = \sum_{j=0}^{p-1} |\eta^j(m'_j)| = n'$.

(ii) Suppose that $|m_{p-1}| \odot |m_{p-2}| \odot \ldots \odot |m_0| <_{lex} |m'_{p-1}| \odot |m'_{p-2}| \odot \ldots \odot |m'_0|$. Then there exists an integer ℓ such that $0 \leq \ell \leq p-1$, $|m_j| = |m'_j|$ for every integer j such that $\ell < j \leq p-1$ and $|m_\ell| < |m'_\ell|$. Since $|m_{p-1}| = |m'_{p-1}|$ and $m_{p-1}a_{p-1}$ and $m'_{p-1}a'_{p-1}$ are prefixes of $\eta(x)$ we have that $m_{p-1} = m'_{p-1}$ and $a_{p-1} = a'_{p-1}$.

Similarly, we have $m_j = m'_j$ and $a_j = a'_j$ for every j such that $\ell < j \le p-1$. Thus $m_\ell a_\ell$ and $m'_\ell a'_\ell$ must both be prefixes of the image under η of the same letter. This letter is x if $\ell = p-1$ or otherwise is $a_{\ell+1} = a'_{\ell+1}$. Since $|m_\ell| < |m'_\ell|$, we have that $m_\ell a_\ell$ is a prefix of m'_ℓ. Using Lemma 3.3, we have

$$\begin{aligned} n - n' &= \textstyle\sum_{j=0}^{p-1} |\eta^j(m_j)| - \sum_{j=0}^{p-1} |\eta^j(m'_j)| \\ &= \textstyle\sum_{j=0}^{\ell} |\eta^j(m_j)| - \sum_{j=0}^{\ell} |\eta^j(m'_j)| \le |\eta^\ell(m_\ell\, a_\ell)| - |\eta^\ell(m'_\ell)| \le 0. \end{aligned}$$

Then $n \le n'$. If $n = n'$, we obtain a contradiction from part (i). Thus, we conclude that $n < n'$.

Now suppose that $n < n'$ and suppose by contradiction that $\mathrm{tail}_{\eta,p,x}(n) \not<_{lex} \mathrm{tail}_{\eta,p,x}(n')$. If $\mathrm{tail}_{\eta,p,x}(n) = \mathrm{tail}_{\eta,p,x}(n')$, then we obtain from part (i) that $n = n'$, a contradiction. If $\mathrm{tail}_{\eta,p,x}(n) >_{lex} \mathrm{tail}_{\eta,p,x}(n')$, then we obtain from above that $n > n'$, a contradiction. Therefore, we conclude that $\mathrm{tail}_{\eta,p,x}(n) <_{lex} \mathrm{tail}_{\eta,p,x}(n')$. □

4. Dumont-Thomas Complement Numeration Systems for $\mathbb{Z}$ Based on Periodic Points

In this section, we prove extensions of Theorem 3.2 to right-infinite and left-infinite periodic points of substitutions from which we deduce a numeration system for $\mathbb{Z}$ associated with any two-sided periodic point with growing seed of a substitution.

Theorem 4.1. *Let $\eta : A^* \to A^*$ be a substitution with growing letter $a \in A$. Let $u \in \mathrm{Per}_{\mathbb{Z}_{\ge 0}}(\eta)$ such that $u_0 = a$. Let $p \ge 1$ be a period of u. For every integer $n \ge 1$, there exists a unique integer $k = k(n)$ such that p divides k and a unique sequence $(m_i, a_i)_{i=0,\dots,k-1}$ such that*

(i) this sequence is a-admissible and $m_{k-1}m_{k-2}\cdots m_{k-p} \ne \varepsilon$,

(ii) $u_0u_1\cdots u_{n-1} = \eta^{k-1}(m_{k-1})\eta^{k-2}(m_{k-2})\cdots\eta^0(m_0)$.

Proof. Since u is a periodic point of period p, we have that $u_0 = a$ is a prefix of $\eta^p(a)$. Also, since a is growing, we have that $\eta^p(a) \in aA^+$. Thus $(|\eta^{p\ell}(a)|)_{\ell\in\mathbb{N}}$ is a strictly increasing sequence starting with value 1 when $\ell = 0$. Let $n \ge 1$ be an integer. There exists a unique integer $\ell \ge 1$ such that $|\eta^{p(\ell-1)}(a)| \le n < |\eta^{p\ell}(a)|$. Let $k = p\ell$ so that we have

$$|\eta^{k-p}(a)| \le n < |\eta^k(a)|. \tag{2}$$

The word $m = u_0u_1\cdots u_{n-1}$ is thus a proper prefix of $\eta^k(a)$. From Lemma 3.9, there exists a unique a-admissible sequence $(m_i, a_i)_{i=0,\dots,k-1}$ such that

$$m = \eta^{k-1}(m_{k-1})\eta^{k-2}(m_{k-2})\cdots\eta^0(m_0).$$

Assume by contradiction that $m_{k-1}m_{k-2}\cdots m_{k-p} = \varepsilon$. Then $a_{k-p} = a$ and from Lemma 3.3, we have

$$\begin{aligned} n = |m| &= \textstyle\sum_{j=0}^{k-1} |\eta^j(m_j)| = \sum_{j=0}^{k-p-1} |\eta^j(m_j)| \\ &< |\eta^{k-p-1}(m_{k-p-1}a_{k-p-1})| \le |\eta^{k-p-1}(\eta(a_{k-p}))| = |\eta^{k-p}(a)|, \end{aligned}$$

a contradiction with (2). Thus $m_{k-1}m_{k-2}\cdots m_{k-p} \neq \varepsilon$. □

We now adapt Dumont-Thomas's theorem to the left-infinite periodic points.

Theorem 4.2. *Let $\eta : A^* \to A^*$ be a substitution with growing letter $b \in A$. Let $u \in \mathrm{Per}_{\mathbb{Z}_{<0}}(\eta)$ such that $u_{-1} = b$. Let $p \geq 1$ be a period of u. For every integer $n \leq -2$, there exists a unique integer $k = k(n)$ such that p divides k and a unique sequence $(m_i, a_i)_{i=0,\dots,k-1}$ such that*

(i) this sequence is b-admissible and

$$\eta^{p-1}(m_{k-1})\eta^{p-2}(m_{k-2})\cdots\eta^0(m_{k-p})a_{k-p} \neq \eta^p(b), \tag{3}$$

(ii) $u_{-|\eta^k(b)|}\cdots u_{n-2}u_{n-1} = \eta^{k-1}(m_{k-1})\eta^{k-2}(m_{k-2})\cdots\eta^0(m_0)$.

Proof. Since u is a periodic point of period p, we have that $u_{-1} = b$ is a suffix of $\eta^p(b)$. Also, since b is growing, we have that $\eta^p(b) \in A^+b$. Thus $(-|\eta^{p\ell}(b)|)_{\ell\in\mathbb{N}}$ is a strictly decreasing sequence starting with value -1 when $\ell = 0$. Let $n \leq -2$ be an integer. There exists a unique integer $\ell \geq 1$ such that $-|\eta^{p\ell}(b)| \leq n < -|\eta^{p(\ell-1)}(b)|$. Let $k = p\ell$ so that we have

$$-|\eta^k(b)| \leq n < -|\eta^{k-p}(b)|. \tag{4}$$

Therefore the word $m = u_{-|\eta^k(b)|}\cdots u_{n-2}u_{n-1}$ of length

$$|m| = |\eta^k(b)| + n < |\eta^k(b)| - |\eta^{k-p}(b)| \leq |\eta^k(b)| \tag{5}$$

is a proper prefix of the word $\eta^k(b)$. From Lemma 3.9, there exists a unique b-admissible sequence $(m_i, a_i)_{i=0,\dots,k-1}$ such that

$$m = \eta^{k-1}(m_{k-1})\eta^{k-2}(m_{k-2})\cdots\eta^0(m_0).$$

By contradiction, assume that (3) is an equality. Then $a_{k-p} = b$ and

$$\begin{aligned} |m| &= |\eta^{k-p}(\eta^p(b))| - |\eta^{k-p}(a_{k-p})| + \textstyle\sum_{j=0}^{k-p-1} |\eta^j(m_j)| \\ &\geq |\eta^k(b)| - |\eta^{k-p}(b)|, \end{aligned}$$

a contradiction with (5). □

We may now define a numeration system for $\mathbb{Z}$ using the previous results.

Definition 4.3 (Dumont-Thomas complement numeration systems for $\mathbb{Z}$)**.** Let $\eta : A^* \to A^*$ be a substitution and $u \in \mathrm{Per}(\eta)$ be a two-sided periodic point with growing seed $u_{-1}|u_0$. Let $p \geq 1$ be the period of u. Let $\mathcal{D} = \{\mathtt{0}, ..., \max_{c\in A} |\eta(c)| - 1\}$. We define

$$\begin{aligned} \mathrm{rep}_u : \quad \mathbb{Z} &\to \{\mathtt{0}, \mathtt{1}\} \odot \mathcal{D}^* \\ n &\mapsto \begin{cases} \mathtt{0} \odot |m_{k-1}| \odot |m_{k-2}| \odot \ldots \odot |m_0|, & \text{if } n \geq 1, \\ \mathtt{0}, & \text{if } n = 0, \\ \mathtt{1}, & \text{if } n = -1, \\ \mathtt{1} \odot |m_{k-1}| \odot |m_{k-2}| \odot \ldots \odot |m_0|, & \text{if } n \leq -2, \end{cases} \end{aligned}$$

where $k = k(n) \geq 0$ is the unique integer and $(m_i, a_i)_{i=0,\ldots,k-1}$ is the unique sequence obtained from Theorem 4.1 (Theorem 4.2) applied on the right-infinite periodic point $u|_{\mathbb{Z}_{\geq 0}}$ (on the left-infinite periodic point $u|_{\mathbb{Z}_{<0}}$) if $n \geq 1$ (if $n \leq -2$, respectively) both with period p.

Note that the period $p \in \mathbb{N}$ of u divides $|\mathrm{rep}_u(n)|-1$ for every $n \in \mathbb{Z}$. Also, one may observe that

$$\mathrm{rep}_u(n) = \begin{cases} \mathtt{0} \odot \mathrm{tail}_{\eta,k,u_0}(n), & \text{if } n \geq 0, \\ \mathtt{1} \odot \mathrm{tail}_{\eta,k,u_{-1}}(n), & \text{if } n < 0. \end{cases} \tag{6}$$

Remark 4.4. In Definition 4.3, the numeration system rep_u could be defined with any period of the two-sided periodic point u and the main result, Theorem 6.1, would still hold. A choice is made here to keep it simple and always take the period of the periodic point u.

Remark 4.5. If $u \in \mathrm{Per}(\eta)$ is a two-sided periodic point of period p with growing seed, then its restriction $u|_{\mathbb{Z}_{\geq 0}}$ to the nonnegative integers is also a periodic point, but its period might be smaller than p (in general, a divisor of p). For example, this is what happens for the Fibonacci substitution $\varphi : a \mapsto ab, b \mapsto a$ or the Thue-Morse substitution $\psi_{TM} : a \mapsto ab, b \mapsto ba$. Both have two-sided periodic points of period 2 and right-infinite fixed points. In Definition 4.3, the numeration system is defined with the period of the two-sided periodic point u when applying Theorem 4.1 on $u|_{\mathbb{Z}_{\geq 0}}$ and Theorem 4.2 on $u|_{\mathbb{Z}_{<0}}$.

When $u = \eta^p(u)$ is a periodic point of a substitution η, then it is also a fixed point of the substitution η^p. Thus, Theorem 3.2 may be used to define a numeration system for $\mathbb{N}$, but it leads to a much larger alphabet size $\#\mathcal{D}$. One advantage of Definition 4.3 is that the size of the alphabet $\mathcal{D}$ is independent of the period p.

Example 4.6. Consider the Tribonacci substitution $\psi_T : a \mapsto ab, b \mapsto ac, c \mapsto a$ [30]. The successive images of the seed $c|a$ under the substitution ψ_T are illustrated as a tree in Figure 4. Let $\omega = \cdots abac|abacaba \cdots$ be the two-sided periodic point

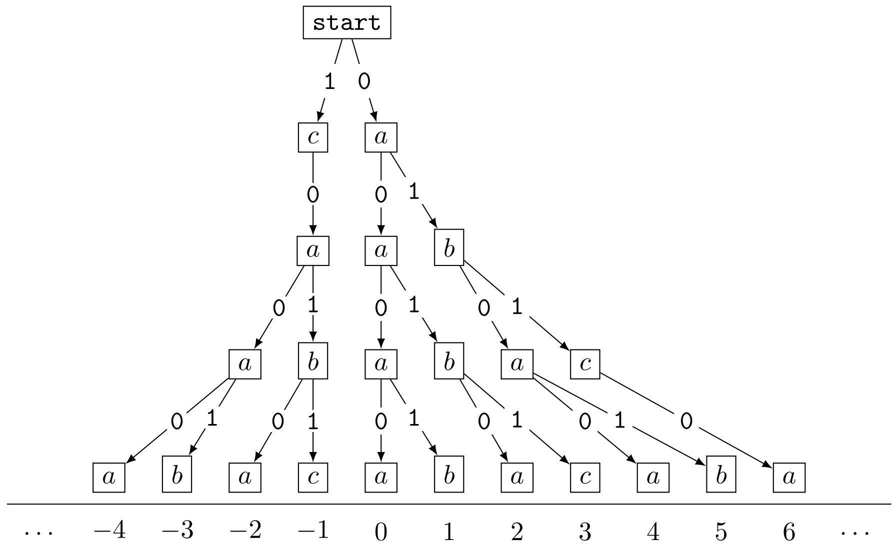


Figure 4: The successive images of the seed $c|a$ under the Tribonacci substitution.

of ψ_T of period 3 with seed $c|a$. In Figure 4, the representation $\mathrm{rep}_\omega(n)$ of n labels the shortest path from the root of the tree to a node at x-position $n \in \mathbb{N}$. The representation of small integers based on the periodic point ω is illustrated in Table 2.

n	$\mathrm{rep}_\omega(n)$	n	$\mathrm{rep}_\omega(n)$
-7	1010100	0	0
-6	1010101	1	0001
-5	1010110	2	0010
-4	1000	3	0011
-3	1001	4	0100
-2	1010	5	0101
-1	1	6	0110

Table 2: Representation of small integers in the Dumont-Thomas complement numeration system based on the periodic point ω with seed $c|a$ of the Tribonacci substitution ψ_T.

Definition 4.7 (quotient, remainder)**.** Let $\eta : A^* \to A^*$ be a substitution and $u \in \mathrm{Per}(\eta)$ be a two-sided periodic point with growing seed $s = u_{-1}|u_0$. Let $p \geq 1$ be the period of u. Let $n \in \mathbb{Z} \setminus \{-1, 0\}$ be an integer and $k = k(n)$ be the unique integer and $(m_i, a_i)_{i=0,\ldots,k-1}$ be the unique sequence obtained from Theorem 4.1

(Theorem 4.2) applied on $u|_{\mathbb{Z}_{\geq 0}}$ ($u|_{\mathbb{Z}_{<0}}$) if $n \geq 1$ (if $n \leq -2$, respectively) both with period p. We define the *u-quotient* of n as

$$q = \begin{cases} |\eta^{k-p-1}(m_{k-1})\eta^{k-p-2}(m_{k-2})\cdots\eta^0(m_p)|, & \text{if } n \geq 1, \\ |\eta^{k-p-1}(m_{k-1})\eta^{k-p-2}(m_{k-2})\cdots\eta^0(m_p)| - |\eta^{k-p}(u_{-1})|, & \text{if } n \leq -2, \end{cases}$$

and the *u-remainder* of n as $r = |\eta^{p-1}(m_{p-1})\eta^{p-2}(m_{p-2})\cdots\eta^0(m_0)|$.

Notice that the u-quotient q and u-remainder r of an integer $n \in \mathbb{Z} \setminus \{-1, 0\}$ fulfill the condition that if $n \geq 1$ then $0 \leq q < n$ and if $n \leq -2$ then $n < q \leq -1$. Consequently, $|q| < |n|$. Also, if η is d-uniform, then the u-quotient and u-remainder of n correspond to the quotient and remainder of the division of n by d^p.

Remark 4.8. Note that if we know the u-quotient q and the u-remainder r, we can recover the sequence $|m_{p-1}| \odot |m_{p-2}| \odot \ldots \odot |m_0|$. Indeed, it is equal to $\mathrm{tail}_{\eta,p,u_q}(r)$.

Lemma 4.9. *Let $\eta : A^* \to A^*$ be a substitution and $u \in \mathrm{Per}(\eta)$ be a two-sided periodic point with growing seed. Let $p \geq 1$ be the period of u. Let $n \in \mathbb{Z} \setminus \{-1, 0\}$ be an integer. If $q \in \mathbb{Z}$ is the u-quotient and $r \in \mathbb{N}$ is the u-remainder of n, then*

$$u_n = \eta^p(u_q)[r] \quad \textit{and} \quad \mathrm{rep}_u(n) = \mathrm{rep}_u(q) \odot \mathrm{tail}_{\eta,p,u_q}(r).$$

Proof. Let $a, b \in A$ denote the letters $b = u_{-1}$ and $a = u_0$. Let $n \in \mathbb{Z} \setminus \{-1, 0\}$ and let q be the u-quotient and r the u-remainder of n.

Suppose $n \geq 1$. From Theorem 4.1, there exists a unique a-admissible sequence $(m_i, a_i)_{i=0,\ldots,k-1}$ such that $u_0 \ldots u_{n-1} = \eta^{k-1}(m_{k-1}) \ldots \eta^0(m_0)$. Also, $\eta^{k-1}(m_{k-1}) \ldots \eta^0(m_0)a_0$ is a prefix of $\eta^k(a)$, which is a prefix of $u_0 u_1 \cdots u_{|\eta^k(a)|-1}$, thus $u_n = a_0$. Since u has period p, the word

$$\eta^{k-p-1}(m_{k-1})\eta^{k-p-2}(m_{k-2})\cdots\eta^0(m_p)a_p$$

is a prefix of $\eta^{k-p}(a)$, which is a prefix of $u_0 u_1 \cdots u_{|\eta^{k-p}(a)|-1}$. Thus $a_p = u_q$. Since $\eta^{p-1}(m_{p-1})\cdots\eta^0(m_0)a_0$ is a prefix of $\eta^p(a_p)$, we deduce that $u_n = a_0 = \eta^p(a_p)[r] = \eta^p(u_q)[r]$.

Suppose $n \leq -2$. From Theorem 4.2, there exists a unique b-admissible sequence $(m_i, a_i)_{i=0,\ldots,k-1}$ such that $u_{-|\eta^k(b)|} \ldots u_{n-1} = \eta^{k-1}(m_{k-1}) \ldots \eta^0(m_0)$. Also, $\eta^{k-1}(m_{k-1}) \ldots \eta^0(m_0)a_0$ is a prefix of $\eta^k(b)$, which is a prefix of $u_{-|\eta^k(b)|} \ldots u_{-1}$, thus $u_n = a_0$. Since u has period p, the word

$$\eta^{k-p-1}(m_{k-1})\eta^{k-p-2}(m_{k-2})\cdots\eta^0(m_p)a_p$$

is a prefix of $\eta^{k-p}(a)$, which is a prefix of $u_{-|\eta^{k-p}(b)|} \ldots u_{-1}$, thus $a_p = u_q$. Since $\eta^{p-1}(m_{p-1})\cdots\eta^0(m_0)a_0$ is a prefix of $\eta^p(a_p)$, we deduce that $u_n = a_0 = \eta^p(a_p)[r] = \eta^p(u_q)[r]$.

To finish the proof for both cases simultaneously, if $n \geq 1$ ($n \leq -2$), applying Theorem 4.1 (Theorem 4.2) on the u-quotient q gives for $\mathtt{d} = \mathtt{0}$ ($\mathtt{d} = \mathtt{1}$)

$$\operatorname{rep}_u(q) = \mathtt{d} \odot |m_{k-1}| \odot |m_{k-2}| \odot \ldots \odot |m_p|.$$

As $n \geq 1$ if and only if $q \geq 0$, we have

$$\begin{aligned}\operatorname{rep}_u(n) &= \mathtt{d} \odot |m_{k-1}| \odot |m_{k-2}| \odot \ldots \odot |m_p| \odot |m_{p-1}| \odot \ldots \odot |m_0| \\ &= \operatorname{rep}_u(q) \odot |m_{p-1}| \odot \ldots \odot |m_0| = \operatorname{rep}_u(q) \odot \operatorname{tail}_{\eta,p,u_q}(r).\end{aligned}$$

□

5. More Examples

We consider the following substitutions:

$$\psi_{TM} : \begin{cases} a \mapsto ab, \\ b \mapsto ba, \end{cases} \qquad \psi_2 : \begin{cases} a \mapsto ab, \\ b \mapsto cb, \\ c \mapsto ac, \end{cases} \qquad \varphi : \begin{cases} a \mapsto ab, \\ b \mapsto a, \end{cases}$$

(Thue-Morse) (some 2-uniform) (Fibonacci)

$$\psi_T : \begin{cases} a \mapsto ab, \\ b \mapsto ac, \\ c \mapsto a, \end{cases} \qquad \rho : \begin{cases} a \mapsto ac, \\ b \mapsto cb, \\ c \mapsto c. \end{cases}$$

(Tribonacci) (non-primitive)

We let

- $\alpha \in \operatorname{Per}(\psi_{TM})$ denote the periodic point with the seed $a|a$ and period 2,
- $\beta \in \operatorname{Per}(\psi_2)$ denote the periodic point with the seed $b|a$ and period 1,
- $\gamma, \delta \in \operatorname{Per}(\varphi)$ denote the periodic point of period 2 with, respectively, the seeds $b|a$ and $a|a$,
- $\tau \in \operatorname{Per}(\psi_T)$ denote the periodic point with the seed $c|a$ and period 3,
- $\chi \in \operatorname{Per}(\mu)$ denote the periodic point with the seed $c|a$ and period 1 of the substitution $\mu : a \mapsto abc, b \mapsto c, c \mapsto ac$ defined in the introduction,
- $\xi \in \operatorname{Per}(\rho)$ denote the periodic point with the seed $b|a$ and period 1.

The numeration systems derived from these two-sided periodic points are shown in Table 3.

substitution images	T.-Morse (ab, ba)	2-uniform (ab, cb, ac)	Fibo. (ab, a)	Fibo. (ab, a)	Tribo. (ab, ac, a)	Intro. (abc, c, ac)	non-prim. (ac, cb, c)
per. point	α	β	γ	δ	τ	χ	ξ
seed	$a\|a$	$b\|a$	$b\|a$	$a\|a$	$c\|a$	$c\|a$	$b\|a$
period	2	1	2	2	3	1	1
n	$\mathrm{rep}_\alpha(n)$	$\mathrm{rep}_\beta(n)$	$\mathrm{rep}_\gamma(n)$	$\mathrm{rep}_\delta(n)$	$\mathrm{rep}_\tau(n)$	$\mathrm{rep}_\chi(n)$	$\mathrm{rep}_\xi(n)$
10	01010	01010	0010010	0010010	0001011	0202	01000000000
9	01001	01001	0010001	0010001	0001010	0201	0100000000
8	01000	01000	0010000	0010000	0001001	0200	010000000
7	00111	0111	01010	01010	0001000	0101	01000000
6	00110	0110	01001	01001	0110	0100	0100000
5	00101	0101	01000	01000	0101	021	010000
4	00100	0100	00101	00101	0100	020	01000
3	011	011	00100	00100	0011	010	0100
2	010	010	010	010	0010	02	010
1	001	01	001	001	0001	01	01
0	0	0	0	0	0	0	0
-1	1	1	1	1	1	1	1
-2	110	10	100	101	1010	10	10
-3	101	101	10010	100	1001	102	100
-4	100	100	10001	10101	1000	101	1000
-5	11011	1011	10000	10100	1010110	100	10000
-6	11010	1010	1001010	10010	1010101	1021	100000
-7	11001	1001	1001001	10001	1010100	1020	1000000
-8	11000	1000	1001000	10000	1010011	1010	10000000
-9	10111	10111	1000101	1010101	1010010	1002	100000000
-10	10110	10110	1000100	1010100	1010001	1001	1000000000

Table 3: Numeration systems for periodic points α, β, γ, δ, τ, χ, ξ with given seed.

6. Periodic Points as Automatic Sequences

Let $\eta : A^* \to A^*$ be a substitution and $u \in \mathrm{Per}(\eta)$ be a two-sided periodic point with growing seed $s = u_{-1}|u_0$. Let $\mathcal{D} = \{\mathtt{0}, ..., \max_{c\in A} |\eta(c)| - 1\}$. We associate an automaton $\mathcal{A}_{\eta,s}$ with (η, s) by adding a new state $\mathtt{start}$ and two additional edges to the automaton $\mathcal{A}_{\eta,a}$ defined in [4]. The automaton $\mathcal{A}_{\eta,s} = (A \cup \{\mathtt{start}\}, \mathcal{D}, \delta, \mathtt{start}, A)$ has the transition function $\delta : A \cup \{\mathtt{start}\} \to A$ such that

- $\delta(\mathtt{start}, \mathtt{0}) = s_0 = u_0, \quad \delta(\mathtt{start}, \mathtt{1}) = s_{-1} = u_{-1}$,
- for every $c, d \in A$, every $w = w_0 w_1 \dots w_{\ell-1} \in A^\ell$ and every $i \in \mathcal{D}$, it holds that $\delta(c, i) = d$ if and only if $\eta(c) = w$ and $w_i = d$.

Examples of automata associated with the Fibonacci substitution are shown in Figure 5.

If the seed is $s = b|a$, the automaton $\mathcal{A}_{\eta,s}$ is related to the usual automata $\mathcal{A}_{\eta,a}$ and $\mathcal{A}_{\eta,b}$ according to the following equalities for every $w \in \mathcal{D}^*$:

$$\mathcal{A}_{\eta,s}(\mathtt{0} \odot w) = \mathcal{A}_{\eta,a}(w) \qquad \text{and} \qquad \mathcal{A}_{\eta,s}(\mathtt{1} \odot w) = \mathcal{A}_{\eta,b}(w). \tag{7}$$

Also if $\mathcal{A}_{\eta,s}(w) = a$ for some $w \in \mathcal{D}^+$, then for every $u \in \mathcal{D}^*$

$$\mathcal{A}_{\eta,a}(u) = \mathcal{A}_{\eta,s}(w \odot u). \tag{8}$$

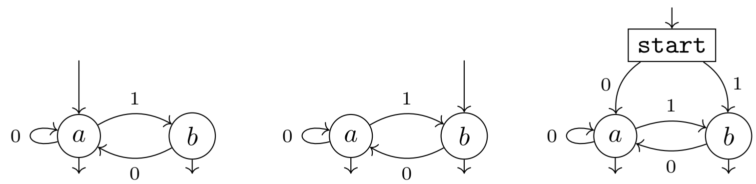


Figure 5: Automata $\mathcal{A}_{\varphi,a}$, $\mathcal{A}_{\varphi,b}$ and $\mathcal{A}_{\varphi,s}$ for $\varphi : a \mapsto ab$, $b \mapsto a$ and $s = b|a$.

A theorem of Cobham [9] says that a sequence $u = (u_n)_{n\geq 0}$ is k-automatic with $k \geq 2$ if and only if it is the image, under a coding, of a fixed point of a k-uniform morphism [1, Section 6]. It was extended to abstract numeration systems based on regular languages which includes numeration systems based on non-uniform morphisms [32]; see also [4, Section 3]. The following result extends Cobham's theorem to the case of two-sided periodic points of non-uniform substitutions.

Theorem 6.1. *Let $\eta : A^* \to A^*$ be a substitution and $u \in \mathrm{Per}(\eta)$ be a two-sided periodic point with growing seed $s = u_{-1}|u_0$. Then for every $n \in \mathbb{Z}$*

$$u_n = \mathcal{A}_{\eta,s}(\mathrm{rep}_u(n)).$$

Proof. If $n \in \{0, -1\}$ then by definition we have $u_n = s_n = \mathcal{A}_{\eta,s}(\mathrm{rep}_u(n))$.

The proof is done by induction. Let $n \in \mathbb{Z} \setminus \{0, -1\}$. Assume that for every $m \in \mathbb{Z}$ such that $|m| < |n|$ it holds that $x_m = \mathcal{A}_{\eta,s}(\mathrm{rep}_u(m))$. Let q be the u-quotient and r the u-remainder of n. As $|q| < |n|$, q fulfills the induction hypothesis, i.e., $u_q = \mathcal{A}_{\eta,s}(\mathrm{rep}_u(q))$. Let $p \geq 1$ be the period of u. From Lemma 4.9 we have $u_n = \eta^p(u_q)[r]$ and $\mathrm{rep}_u(n) = \mathrm{rep}_u(q) \odot \mathrm{tail}_{\eta,p,u_q}(r)$. Using Lemma 3.11 and Equation (8), we have

$$\begin{aligned} u_n &= \eta^p(u_q)[r] = \mathcal{A}_{\eta,u_q}(\mathrm{tail}_{\eta,p,u_q}(r)) \\ &= \mathcal{A}_{\eta,s}(\mathrm{rep}_u(q) \odot \mathrm{tail}_{\eta,p,u_q}(r)) = \mathcal{A}_{\eta,s}(\mathrm{rep}_u(n)). \end{aligned}$$

□

7. Numeration Systems for $\mathbb{Z}^d$ Based on Periodic Points

A numeration system for $\mathbb{Z}^d$ can be deduced from the numeration system for $\mathbb{Z}$ based on a periodic point. Since not all integers are represented by words of the same length, we propose here a way to pad them to a common length.

Let $\eta : A^* \to A^*$ be a substitution and $u \in \mathrm{Per}(\eta)$ with period $p \geq 1$ and growing seed. Let $\mathtt{W}_{\min}$ and $\mathtt{W}_{\max}$ be the following minimum and the maximum element under the tail map with particular parameters:

$$\begin{aligned}\mathtt{W}_{\min} &= \operatorname{tail}_{\eta,p,u_0}(0) = \mathtt{0}^p,\\ \mathtt{W}_{\max} &= \operatorname{tail}_{\eta,p,u_{-1}}(|\eta^p(u_{-1})|-1).\end{aligned}$$

The words $\mathtt{W}_{\min}$ and $\mathtt{W}_{\max}$ play the role of neutral words in the numeration system as illustrated in the next lemma. Below the words $\mathtt{W}_{\min}$ and $\mathtt{W}_{\max}$ are concatenated with others words from $\mathcal{D}^*$ using the binary operation $\odot$, which is not explicitly written to avoid heavy notation.

Lemma 7.1. *Let $\eta : A^* \to A^*$ be a substitution and $u \in \operatorname{Per}(\eta)$ be a two-sided periodic point with growing seed $s = u_{-1}|u_0$. Let $w \in \mathcal{L}(\mathcal{A}_{\eta,s})$. Then*

$$\mathcal{A}_{\eta,s}(w) = \begin{cases}\mathcal{A}_{\eta,s}(\mathtt{0}(\mathtt{W}_{\min})^i v), & \text{if } w = \mathtt{0}v,\\ \mathcal{A}_{\eta,s}(\mathtt{1}(\mathtt{W}_{\max})^i v), & \text{if } w = \mathtt{1}v,\end{cases}$$

for every integer $i \geq 0$.

Proof. Let $i \geq 0$ be an integer and $p \geq 1$ be the period of u. Let $w \in \mathcal{L}(\mathcal{A}_{\eta,s})$.

Suppose that w starts with letter $\mathtt{0}$. Let $v \in \mathcal{D}^*$ such that $w = \mathtt{0}v$. We have $\mathcal{A}_{\eta,u_0}(\mathtt{0}^p) = u_0$. Thus $\mathcal{A}_{\eta,s}(\mathtt{0}(\mathtt{W}_{\min})^i) = u_0$. From Equation (7) and Equation (8) we obtain

$$\mathcal{A}_{\eta,s}(w) = \mathcal{A}_{\eta,s}(\mathtt{0}v) \overset{(7)}{=} \mathcal{A}_{\eta,u_0}(v) \overset{(8)}{=} \mathcal{A}_{\eta,s}(\mathtt{0}(\mathtt{W}_{\min})^i v).$$

Suppose that w starts with letter $\mathtt{1}$. Let $v \in \mathcal{D}^*$ such that $w = \mathtt{1}v$. We have $\mathcal{A}_{\eta,u_{-1}}(\mathtt{W}_{\max}) = u_{-1}$. Thus $\mathcal{A}_{\eta,s}(\mathtt{1}(\mathtt{W}_{\max})^i) = u_{-1}$. From Equation (7) and Equation (8) we obtain

$$\mathcal{A}_{\eta,s}(w) = \mathcal{A}_{\eta,s}(\mathtt{1}v) \overset{(7)}{=} \mathcal{A}_{\eta,u_{-1}}(v) \overset{(8)}{=} \mathcal{A}_{\eta,s}(\mathtt{1}(\mathtt{W}_{\max})^i v).$$

□

It is useful to pad words to a certain length using neutral words as follows using a pad function. Let $s = u_{-1}|u_0$. Let $w \in \mathcal{L}_{\ell p+1}(\mathcal{A}_{\eta,s})$ for some $\ell \in \mathbb{N}$. Let $t \in \mathbb{N}$ such that $t \geq |w|$ and $t \bmod p = 1$. We define

$$\operatorname{pad}_t(w) = \begin{cases}\mathtt{0}(\mathtt{W}_{\min})^m v, & \text{if } w = \mathtt{0}v,\\ \mathtt{1}(\mathtt{W}_{\max})^m v, & \text{if } w = \mathtt{1}v,\end{cases}$$

where $m = (t - |w|)/p$. The padding map can be used to pad words so that they all have the same length. This allows us to represent coordinates in $\mathbb{Z}^d$ in dimension $d \geq 1$.

Definition 7.2 (Numeration system for $\mathbb{Z}^d$)**.** Let $\eta : A^* \to A^*$ be a substitution and $u_1, u_2, \ldots, u_d \in \mathrm{Per}(\eta)$ be periodic points with growing seeds and of the same period. For every $\boldsymbol{n} = (n_1, n_2, \ldots, n_d) \in \mathbb{Z}^d$, we define

$$\mathrm{rep}_u(\boldsymbol{n}) = \begin{pmatrix} \mathrm{pad}_t(\mathrm{rep}_{u_1}(n_1)) \\ \mathrm{pad}_t(\mathrm{rep}_{u_2}(n_2)) \\ \cdots \\ \mathrm{pad}_t(\mathrm{rep}_{u_d}(n_d)) \end{pmatrix} \in \{\mathtt{0},\mathtt{1}\}^d(\mathcal{D}^d)^*,$$

where $t = \max\{|\mathrm{rep}_{u_i}(n_i)| : 1 \le i \le d\}$.

Remark 7.3. In Definition 7.2, considering different periodic points with the same period of the same 1-dimensional substitution in each dimension can be necessary for instance to describe the different 2-dimensional periodic points of 2-dimensional substitutions. This is what happens when one wants to describe the 8 configurations of Wang tiles presented in [23] which are the periodic points of a 2-dimensional substitution.

Of course, it is possible and simpler to use the same periodic point to represent the entries of an integer vector. This is what is done in the example that follows.

Example 7.4. Consider the Tribonacci substitution $\psi_T : a \mapsto ab, b \mapsto ac, c \mapsto a$ as in Example 3.10. Let $\tau \in \mathrm{Per}(\psi_T)$ be the periodic point with period $p = 3$ and seed $c|a$. We have

$$\mathtt{W}_{\min} = \mathrm{tail}_{\psi_T,p,u_0}(0) = \mathrm{tail}_{\psi_T,3,a}(0) = \mathtt{0}^3 = \mathtt{000},$$
$$\mathtt{W}_{\max} = \mathrm{tail}_{\psi_T,p,u_{-1}}(|\psi_T^p(u_{-1})| - 1) = \mathrm{tail}_{\psi_T,3,c}(|\psi_T^3(c)| - 1) = \mathrm{tail}_{\psi_T,3,c}(3) = \mathtt{011}.$$

The words $\mathtt{W}_{\min}$ and $\mathtt{W}_{\max}$ can be used to pad words to a given length which is a multiple of 3 plus 1. For instance, we illustrate in Table 4 the padding of the Dumont-Thomas representation based on the periodic point τ. The representation of integers from -10 to 10 is padded to words of length 7. Thus, the coordinate $(-1, 8) \in \mathbb{Z}^2$ can thus be written as a word

$$\begin{aligned}\mathrm{rep}_\tau(-1,8) &= \begin{pmatrix} \mathrm{pad}_7(\mathrm{rep}_\tau(-1)) \\ \mathrm{pad}_7(\mathrm{rep}_\tau(8)) \end{pmatrix} = \begin{pmatrix} \mathtt{1011011} \\ \mathtt{0001001} \end{pmatrix} \\ &= \begin{pmatrix} \mathtt{1} \\ \mathtt{0} \end{pmatrix}\begin{pmatrix} \mathtt{0} \\ \mathtt{0} \end{pmatrix}\begin{pmatrix} \mathtt{1} \\ \mathtt{0} \end{pmatrix}\begin{pmatrix} \mathtt{1} \\ \mathtt{1} \end{pmatrix}\begin{pmatrix} \mathtt{0} \\ \mathtt{0} \end{pmatrix}\begin{pmatrix} \mathtt{1} \\ \mathtt{0} \end{pmatrix}\begin{pmatrix} \mathtt{1} \\ \mathtt{1} \end{pmatrix} \\ &\in \left\{ \begin{pmatrix} \mathtt{0} \\ \mathtt{0} \end{pmatrix}, \begin{pmatrix} \mathtt{0} \\ \mathtt{1} \end{pmatrix}, \begin{pmatrix} \mathtt{1} \\ \mathtt{0} \end{pmatrix}, \begin{pmatrix} \mathtt{1} \\ \mathtt{1} \end{pmatrix} \right\}^*\end{aligned}$$

whose alphabet of size 4 is the Cartesian product of the alphabet $\{\mathtt{0},\mathtt{1}\}$ with itself.

n	$\mathrm{rep}_\tau(n)$	$\mathrm{pad}_7(\mathrm{rep}_\tau(n))$
10	0001011	0001011
9	0001010	0001010
8	0001001	0001001
7	0001000	0001000
6	0110	0000110
5	0101	0000101
4	0100	0000100
3	0011	0000011
2	0010	0000010
1	0001	0000001
0	0	0000000
-1	1	1011011
-2	1010	1011010
-3	1001	1011001
-4	1000	1011000
-5	1010110	1010110
-6	1010101	1010101
-7	1010100	1010100
-8	1010011	1010011
-9	1010010	1010010
-10	1010001	1010001

Table 4: The Dumont-Thomas complement representation of integers from -10 to 10 using the periodic point of seed $c|a$ of the Tribonacci substitution ψ_T can be padded to obtain words of length 7.

8. A Total Order

In this section, we define a total order on $\{0,1\}\mathcal{D}^* := \{0,1\} \odot \mathcal{D}^*$ and we show that rep_u is increasing with respect to this order.

The radix order on a language $L \subset \mathcal{D}^*$ is a total order $(L, <_{rad})$ such that $u <_{rad} v$ if and only if $|u| < |v|$ or $|u| = |v|$ and $u <_{lex} v$, where $<_{lex}$ denotes the lexicographic order. For example, over the alphabet $\{0,1\}$, the minimum elements for the radix order are:

$$\varepsilon <_{rad} 0 <_{rad} 1 <_{rad} 00 <_{rad} 01 <_{rad} 10 <_{rad} 11 <_{rad} 000 <_{rad} 001 <_{rad} \cdots .$$

We define the reversed-radix order as a total order such that $u <_{rev} v$ if and only if $|u| > |v|$ or $|u| = |v|$ and $u <_{lex} v$. For example, over the alphabet $\{0,1\}$, the maximum elements for the reverse-radix order are:

$$\cdots <_{rev} 110 <_{rev} 111 <_{rev} 00 <_{rev} 01 <_{rev} 10 <_{rev} 11 <_{rev} 0 <_{rev} 1 <_{rad} \varepsilon.$$

Let us stress that the reversed-radix order behaves in the same manner as the radix order on the words of the same length. Also if L has infinite cardinality, then L has no maximal element for the radix order and has no minimum element for the reverse-radix order.

The radix order and the reverse-radix order can be used jointly to define a total order on a language with no minimum nor maximum element. Below, we use the first letter of a word in $\{0,1\}\mathcal{D}^*$ to split the two cases.

Definition 8.1 (total order $\prec$)**.** For every $u, v \in \{\mathtt{0}, \mathtt{1}\}\mathcal{D}^*$, we define $u \prec v$ if and only if

- $u \in \mathtt{1}\mathcal{D}^*$ and $v \in \mathtt{0}\mathcal{D}^*$, or
- $u, v \in \mathtt{0}\mathcal{D}^*$ and $u <_{rad} v$, or
- $u, v \in \mathtt{1}\mathcal{D}^*$ and $u <_{rev} v$.

Thus, if $\mathcal{D} = \{\mathtt{0}, \mathtt{1}\}$, we get

$$\cdots \prec \mathtt{100} \prec \mathtt{101} \prec \mathtt{110} \prec \mathtt{111} \prec \mathtt{10} \prec \mathtt{11} \prec \mathtt{1} \prec \mathtt{0} \prec \mathtt{00} \prec \mathtt{01} \prec \mathtt{000} \prec \mathtt{001} \prec \cdots .$$

The total order $\prec$ makes sense with respect to Dumont-Thomas complement numeration systems for $\mathbb{Z}$ because of the following result.

Proposition 8.2. *Let $\eta : A^* \to A^*$ be a substitution and $u \in \mathrm{Per}(\eta)$ be a two-sided periodic point with growing seed. The map $\mathrm{rep}_u : \mathbb{Z} \to \{\mathtt{0}, \mathtt{1}\}\mathcal{D}^*$ is increasing with respect to the order $\prec$ on $\{\mathtt{0}, \mathtt{1}\}\mathcal{D}^*$.*

Proof. Let $n, n' \in \mathbb{Z}$ be two integers such that $n < n'$.

Assume that $n < 0 \leq n'$. Then $\mathrm{rep}_u(n) \in \mathtt{1}\mathcal{D}^*$ and $\mathrm{rep}_u(n') \in \mathtt{0}\mathcal{D}^*$ so that $\mathrm{rep}_u(n) \prec \mathrm{rep}_u(n')$.

Assume that $0 \leq n < n'$. Then $\mathrm{rep}_u(n) \in \mathtt{0}\mathcal{D}^*$ and $\mathrm{rep}_u(n') \in \mathtt{0}\mathcal{D}^*$. The case $|\mathrm{rep}_u(n)| > |\mathrm{rep}_u(n')|$ is impossible. Indeed, suppose that $|\mathrm{rep}_u(n)| = k + 1$ and $|\mathrm{rep}_u(n')| = k' + 1$ for some integers k and k'. If $|\mathrm{rep}_u(n)| > |\mathrm{rep}_u(n')|$, then $k - p \geq k'$, where p is the period of u. From Equation (2), we have

$$n' < |\eta^{k'}(a)| \leq |\eta^{k-p}(a)| \leq n,$$

a contradiction. If $|\mathrm{rep}_u(n)| < |\mathrm{rep}_u(n')|$, then $\mathrm{rep}_u(n) \prec \mathrm{rep}_u(n')$. Suppose now that $|\mathrm{rep}_u(n)| = |\mathrm{rep}_u(n')| = k + 1$ for some integer k. From Lemma 3.12, we have

$$\mathrm{rep}_u(n) = \mathtt{0} \odot \mathrm{tail}_{\eta,k,u_0}(n) <_{lex} \mathtt{0} \odot \mathrm{tail}_{\eta,k,u_0}(n') = \mathrm{rep}_u(n').$$

Thus $\mathrm{rep}_u(n) \prec \mathrm{rep}_u(n')$.

Assume that $n < n' < 0$. Then $\mathrm{rep}_u(n) \in \mathtt{1}\mathcal{D}^*$ and $\mathrm{rep}_u(n') \in \mathtt{1}\mathcal{D}^*$. The case $|\mathrm{rep}_u(n)| < |\mathrm{rep}_u(n')|$ is impossible. Indeed, suppose that $|\mathrm{rep}_u(n)| = k + 1$ and $|\mathrm{rep}_u(n')| = k' + 1$ for some integers k and k'. If $|\mathrm{rep}_u(n)| < |\mathrm{rep}_u(n')|$, then $k' - p \geq k$, where p is the period of u. From Equation (4), we have

$$n' < -|\eta^{k'-p}(b)| \leq -|\eta^{k}(b)| \leq n,$$

a contradiction. If $|\mathrm{rep}_u(n)| > |\mathrm{rep}_u(n')|$, then $\mathrm{rep}_u(n) \prec \mathrm{rep}_u(n')$. Suppose that $|\mathrm{rep}_u(n)| = |\mathrm{rep}_u(n')| = k + 1$ for some integer k. From Lemma 3.12, we have

$$\mathrm{rep}_u(n) = \mathtt{1} \odot \mathrm{tail}_{\eta,k,u_{-1}}(n) <_{lex} \mathtt{1} \odot \mathrm{tail}_{\eta,k,u_{-1}}(n') = \mathrm{rep}_u(n').$$

Thus $\mathrm{rep}_u(n) \prec \mathrm{rep}_u(n')$. ∎

It follows from Proposition 8.2 that $\operatorname{rep}_u : \mathbb{Z} \to \{\mathtt{0},\mathtt{1}\}\mathcal{D}^*$ is injective. Therefore it is a bijection onto its image. The next result describes the image of the map rep_u.

Lemma 8.3. *Let $\eta : A^* \to A^*$ be a substitution and $u \in \operatorname{Per}(\eta)$ be a two-sided periodic point with growing seed $s = u_{-1}|u_0$. Let $p \geq 1$ be the period of u. Then*

$$\operatorname{rep}_u(\mathbb{Z}) = \bigcup_{\ell\in\mathbb{N}} \mathcal{L}_{\ell p+1}(\mathcal{A}_{\eta,s}) \setminus \{\mathtt{0W}_{\min}, \mathtt{1W}_{\max}\}\mathcal{D}^*.$$

Proof. $(\subseteq)$. It follows from Theorem 6.1 that $\operatorname{rep}_u(\mathbb{Z}) \subset \mathcal{L}(\mathcal{A}_{\eta,s})$. Also for every $n \in \mathbb{Z}$, $\operatorname{rep}_u(n)$ is a word of length $\ell p + 1$ for some $\ell \in \mathbb{N}$. Thus $\operatorname{rep}_u(\mathbb{Z}) \subset \bigcup_{\ell\in\mathbb{N}} \mathcal{L}_{\ell p+1}(\mathcal{A}_{\eta,s})$. It remains to show that $\operatorname{rep}_u(\mathbb{Z}) \cap \{\mathtt{0W}_{\min}, \mathtt{1W}_{\max}\}\mathcal{D}^* = \varnothing$. Suppose by contradiction that there exists $n \in \mathbb{Z}$ such that $\operatorname{rep}_u(n) \in \mathtt{0W}_{\min}\mathcal{D}^*$. We have $\operatorname{rep}_u(n) = \mathtt{0} \odot |m_{k-1}| \odot |m_{k-2}| \odot \ldots \odot |m_0|$, where $k = \ell p$. Then $|m_{k-1}| \odot \ldots \odot |m_{k-p}| = \mathtt{0}^p$, which implies $m_{k-1}m_{k-2}\ldots m_{k-p} = \varepsilon$, thus contradicting Theorem 4.1. On the other hand, suppose by contradiction that there exists $n \in \mathbb{Z}$ such that $\operatorname{rep}_u(n) \in \mathtt{1W}_{\max}\mathcal{D}^*$. We have $\operatorname{rep}_u(n) = \mathtt{1} \odot |m_{k-1}| \odot |m_{k-2}| \odot \ldots \odot |m_0|$. Then $|m_{k-1}| \odot \ldots \odot |m_{k-p}| = \operatorname{tail}_{\eta,p,u_{-1}}(|\eta^p(u_{-1})| - 1)$, which implies $\eta^{p-1}(m_{k-1})\eta^{p-2}(m_{k-2}) \cdots \eta^0(m_{k-p})$ is the prefix of $\eta^p(u_{-1})$ of length $|\eta^p(u_{-1})| - 1$, a contradiction with Theorem 4.2.

$(\supseteq)$. Let $\ell \in \mathbb{N}$ and $k = \ell p$. Let $v = v_{k-1} \cdots v_0$ such that $\mathtt{d} \odot v \in \mathcal{L}_{\ell p+1}(\mathcal{A}_{\eta,s}) \setminus \{\mathtt{0W}_{\min}, \mathtt{1W}_{\max}\}\mathcal{D}^*$.

Suppose that $\mathtt{d} = \mathtt{0}$. We have $v \in \mathcal{L}(\mathcal{A}_{\eta,u_0})$. From Lemma 3.8, there exists a u_0-admissible sequence $(m_i, a_i)_{i=0,\ldots,k-1}$ such that $|m_i| = v_i$ for every $i = 0, \ldots, k-1$. Let $n = \sum_{i=0}^{k-1} |\eta^i(m_i)|$. From Theorem 4.1 and using $v \notin \mathtt{W}_{\min}\mathcal{D}^*$, we have $\operatorname{rep}_u(n) = \mathtt{0} \odot v$. Thus $\mathtt{d} \odot v \in \operatorname{rep}_u(\mathbb{Z})$.

Suppose that $\mathtt{d} = \mathtt{1}$. We have $v \in \mathcal{L}(\mathcal{A}_{\eta,u_{-1}})$. From Lemma 3.8, there exists a u_{-1}-admissible sequence $(m_i, a_i)_{i=0,\ldots,k-1}$ such that $|m_i| = v_i$ for every $i = 0, \ldots, k-1$. Let $n = -|\eta^k(u_{-1})| + \sum_{i=0}^{k-1} |\eta^i(m_i)|$. From Theorem 4.2, and using $v \notin \mathtt{W}_{\max}\mathcal{D}^*$, we have $\operatorname{rep}_u(n) = \mathtt{1} \odot v$. Thus $\mathtt{d} \odot v \in \operatorname{rep}_u(\mathbb{Z})$. □

Results similar to Proposition 8.2 exist for other numeration systems; see [6, Section 5] and [16, Section 4]. In some other works on numeration systems, such an increasing bijection is not a consequence but rather a hypothesis. For example, a bijection $\mathbb{N} \to \mathcal{L}$ serves as the definition of abstract numeration systems in [27]. Similarly, we have the following characterization of Dumont-Thomas complement numeration systems for $\mathbb{Z}$ in terms of the total order $\prec$ on the language recognized by an automaton.

Theorem 8.4. *Let $\eta : A^* \to A^*$ be a substitution and $u \in \operatorname{Per}(\eta)$ be a two-sided periodic point with growing seed $s = u_{-1}|u_0$. Let $p \geq 1$ be the period of u. Let $f : \mathbb{Z} \to \{\mathtt{0},\mathtt{1}\}\mathcal{D}^*$ be some map. The following items are equivalent:*

- $f = \operatorname{rep}_u$,

- *f is increasing with respect to $\prec$, its image is $f(\mathbb{Z}) = \bigcup_{\ell\in\mathbb{N}} \mathcal{L}_{\ell p+1}(\mathcal{A}_{\eta,s}) \setminus \{\mathtt{0W}_{\min}, \mathtt{1W}_{\max}\}\mathcal{D}^*$ and $f(0) = \mathtt{0}$.*

Proof. Suppose that $f = \mathrm{rep}_u$. Then f is increasing from Proposition 8.2. Its image was computed in Lemma 8.3, Also, $f(0) = \mathtt{0}$ from Definition 4.3.

Let $f : \mathbb{Z} \to \{\mathtt{0}, \mathtt{1}\}\mathcal{D}^*$. Suppose f is increasing, its image is

$$f(\mathbb{Z}) = \bigcup_{\ell\in\mathbb{N}} \mathcal{L}_{\ell p+1}(\mathcal{A}_{\eta,s}) \setminus \{\mathtt{0W}_{\min}, \mathtt{1W}_{\max}\}\mathcal{D}^*$$

and $f(0) = \mathtt{0}$. The map rep_u satisfies the same properties. Since there is a unique increasing bijection $\mathbb{Z} \to f(\mathbb{Z})$ such that $f(0) = \mathtt{0}$, we conclude that $f = \mathrm{rep}_u$. □

9. Relation with Existing Complement Numeration Systems

In this section, we show that two existing complement numeration systems can be recovered as a Dumont-Thomas complement numeration system using the some well-chosen substitutions. The involved substitutions are part of the examples presented in Section 5.

9.1. Two's Complement Numeration System

Let $\mathcal{D} = \{\mathtt{0}, \mathtt{1}\}$. In the two's complement representation of integers the value of a binary word $w = w_{k-1}w_{k-2}\cdots w_0 \in \mathcal{D}^k$ is $\mathrm{val}_{2c}(w) = \sum_{i=0}^{k-1} w_i 2^i - w_{k-1}2^k$; see [22, Section 4.1]. For every $n \in \mathbb{Z}$ there exists a unique word $w \in \mathcal{D}^+ \setminus (\mathtt{00}\mathcal{D}^* \cup \mathtt{11}\mathcal{D}^*)$ such that $n = \mathrm{val}_{2c}(w)$. The word w is called the *two's complement representation* of the integer n, and we denote it by $\mathrm{rep}_{2c}(n)$. Observe that the map $\mathrm{rep}_{2c} : \mathbb{Z} \to \mathcal{D}^+ \setminus (\mathtt{00}\mathcal{D}^* \cup \mathtt{11}\mathcal{D}^*)$ is an increasing bijection with respect to the order $\prec$:

$$\begin{array}{ccccccccccccccccccccccc}
\cdots & < & -5 & < & -4 & < & -3 & < & -2 & < & -1 & < & 0 & < & 1 & < & 2 & < & 3 & < & 4 & < & \cdots \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
\cdots & \prec & \mathtt{1011} & \prec & \mathtt{100} & \prec & \mathtt{101} & \prec & \mathtt{10} & \prec & \mathtt{1} & \prec & \mathtt{0} & \prec & \mathtt{01} & \prec & \mathtt{010} & \prec & \mathtt{011} & \prec & \mathtt{0100} & \prec & \cdots .
\end{array}$$

We now show that the two's complement numeration system coincides with the Dumont-Thomas complement numeration system associated with a two-sided fixed point of 2-uniform substitution.

Proposition 9.1. *Let $\psi : A \to A^*$ be some 2-uniform substitution and let $\beta \in \mathrm{Per}(\psi)$ be some two-sided periodic point of period 1. Then rep_β is the two's complement numeration system, that is, $\mathrm{rep}_\beta = \mathrm{rep}_{2c}$.*

Proof. From Proposition 8.2, $\mathrm{rep}_\beta : \mathbb{Z} \to \{\mathtt{0}, \mathtt{1}\}\mathcal{D}^*$ is an increasing map with respect to the order $\prec$. Thus, it is an increasing bijection $\mathbb{Z} \to \mathrm{rep}_\beta(\mathbb{Z})$. From Lemma 8.3, we have

$$\begin{aligned}\operatorname{rep}_\beta(\mathbb{Z}) &= \bigcup_{\ell\in\mathbb{N}} \mathcal{L}_{\ell p+1}(\mathcal{A}_{\psi,s}) \setminus \{\texttt{0W}_{\min}, \texttt{1W}_{\max}\}\mathcal{D}^* = \mathcal{L}_{\geq 1}(\mathcal{A}_{\psi,s}) \setminus \{\texttt{00},\texttt{11}\}\mathcal{D}^* \\ &= \mathcal{D}^+ \setminus (\texttt{00}\mathcal{D}^* \cup \texttt{11}\mathcal{D}^*)\,,\end{aligned}$$

since $p = 1$, $\mathcal{L}(\mathcal{A}_{\psi,s}) = \mathcal{D}^*$, $\texttt{W}_{\min} = \texttt{0}$ and $\texttt{W}_{\max} = \texttt{1}$. Also, $\operatorname{rep}_\beta(0) = \texttt{0}$. On the other hand, the map $\operatorname{rep}_{2c} : \mathbb{Z} \to \mathcal{D}^+ \setminus (\texttt{00}\mathcal{D}^* \cup \texttt{11}\mathcal{D}^*)$ is an increasing bijection with respect to the order $\prec$ and $\operatorname{rep}_{2c}(0) = \texttt{0}$. From Theorem 8.4, we conclude $\operatorname{rep}_\beta = \operatorname{rep}_{2c}$. □

Note that the Thue-Morse substitution has no two-sided fixed point, so the above result does not hold for numeration systems based on fixed points of the Thue-Morse substitution; see Table 3.

9.2. Fibonacci Analogue of the Two's Complement Numeration System

In what follows, the Fibonacci sequence $(F_n)_{n\geq 0}$, $F_n = F_{n-1} + F_{n-2}$, for all $n \geq 2$, is defined with the initial conditions $F_0 = 1$, $F_1 = 2$. We let $\mathcal{D}$ denote the binary alphabet $\{\texttt{0},\texttt{1}\}$.

In [24], a Fibonacci analogue of the two's complement numeration system for nonnegative and negative integers was defined from the value map $\operatorname{val}_{\mathcal{F}c} : \mathcal{D}^* \to \mathbb{Z}$ by $\operatorname{val}_{\mathcal{F}c}(w) = \sum_{i=0}^{k-1} w_i F_i - w_{k-1} F_k$ for every binary word $w = w_{k-1} \cdots w_0 \in \mathcal{D}^k$. It is an analog of the two's complement value map val_{2c}, using Fibonacci numbers instead of powers of 2. It was proved in [24] that for every $n \in \mathbb{Z}$ there exists a unique odd-length word $w \in L = \mathcal{D}(\mathcal{D}\mathcal{D})^* \setminus (\mathcal{D}^* 11 \mathcal{D}^* \cup 000\mathcal{D}^* \cup 101\mathcal{D}^*)$ such that $n = \operatorname{val}_{\mathcal{F}c}(w)$. It defines the map $\operatorname{rep}_{\mathcal{F}c} : \mathbb{Z} \to L$ by the rule $n \mapsto w$.

We show that the Dumont-Thomas complement numeration system obtained from the two-sided Fibonacci word is the Fibonacci analogue of the two's complement numeration system introduced in [24].

Proposition 9.2. *Let $\varphi : a \mapsto ab, b \mapsto a$ be the Fibonacci substitution and let $\gamma \in \operatorname{Per}(\varphi)$ be the periodic point of period 2 with seed $s = b|a$. Then $\operatorname{rep}_\gamma$ is the Fibonacci analogue of the two's complement numeration system, that is, $\operatorname{rep}_\gamma = \operatorname{rep}_{\mathcal{F}c}$.*

Proof. From Proposition 8.2, $\operatorname{rep}_\gamma : \mathbb{Z} \to \{\texttt{0},\texttt{1}\}\mathcal{D}^*$ is an increasing map with respect to the order $\prec$. Thus, it is an increasing bijection $\mathbb{Z} \to \operatorname{rep}_\gamma(\mathbb{Z})$. Let $L = \mathcal{D}(\mathcal{D}\mathcal{D})^* \setminus (\mathcal{D}^*\texttt{11}\mathcal{D}^* \cup \texttt{000}\mathcal{D}^* \cup \texttt{101}\mathcal{D}^*)$. From Lemma 8.3, we have

$$\begin{aligned}\operatorname{rep}_\gamma(\mathbb{Z}) &= \bigcup_{\ell\in\mathbb{N}} \mathcal{L}_{\ell p+1}(\mathcal{A}_{\varphi,s}) \setminus \{\texttt{0W}_{\min}, \texttt{1W}_{\max}\}\mathcal{D}^* \\ &= (\mathcal{D}(\mathcal{D}\mathcal{D})^* \setminus \mathcal{D}^*\texttt{11}\mathcal{D}^*) \setminus \{\texttt{000},\texttt{101}\}\mathcal{D}^* = L,\end{aligned}$$

since $p = 2$, $\mathcal{L}(\mathcal{A}_{\varphi,s}) = \mathcal{D}^* \setminus \mathcal{D}^*\texttt{11}\mathcal{D}^*$, $\texttt{W}_{\min} = \texttt{00}$ and $\texttt{W}_{\max} = \texttt{01}$. From [24], the map $\operatorname{rep}_{\mathcal{F}c}$ is an increasing bijection $\mathbb{Z} \to L$ with respect to the order $\prec$. Moreover, $\operatorname{rep}_\gamma(0) = \texttt{0} = \operatorname{rep}_{\mathcal{F}c}(0)$. From Theorem 8.4, $\operatorname{rep}_{\mathcal{F}c} = \operatorname{rep}_\gamma$. □

We leave open the following question.

Question 9.3. If rep_γ is the Fibonacci analogue of the two's complement numeration system, then what is the meaning of rep_δ? Can we define it from some value map? Recall that $\delta \in \text{Per}(\varphi)$ is the periodic point of period 2 of the Fibonacci substitution φ with seed $a|a$; see Table 3.

Acknowledgements. The authors would like to thank the reviewer for their valuable comments leading in particular to an improved buildup in the introduction and clearer notation for the concatenation of integers (not to be misinterpreted as multiplication). This work was partially funded in France by ANR CODYS (ANR-18-CE40-0007) and ANR IZES (ANR-22-CE40-0011). The second author acknowledges financial support by the Barrande fellowship program and the Grant Agency of Czech Technical University in Prague (SGS20/183/OHK4/3T/14).

References

[1] J.-P. Allouche and J. Shallit, *Automatic Sequences*, Cambridge University Press, Cambridge, 2003.

[2] M. Baake and U. Grimm, *Aperiodic Order. Vol. 1*, Cambridge University Press, Cambridge, 2013.

[3] V. Berthé and M. Rigo, Odometers on regular languages, *Theory Comput. Syst.* **40** (1) (2007), 1-31.

[4] V. Berthé and M. Rigo, editors, *Combinatorics, Automata and Number Theory*, Cambridge University Press, Cambridge, 2010.

[5] V. Bruyère and G. Hansel, Bertrand numeration systems and recognizability, *Theoret. Comput. Sci.* **181** (1) (1997), 17-43.

[6] V. Canterini and A. Siegel, Automate des préfixes-suffixes associé à une substitution primitive, *J. Théor. Nombres Bordeaux* **13** (2) (2001), 353-369.

[7] V. Canterini and A. Siegel, Geometric representation of substitutions of Pisot type, *Trans. Amer. Math. Soc.* **353** (12) (2001), 5121-5144.

[8] E. Charlier, M. Le Gonidec, and M. Rigo, Representing real numbers in a generalized numeration system, *J. Comput. System Sci.* **77** (4) (2011), 743-759.

[9] A. Cobham, Uniform tag sequences, *Math. Systems Theory* **6** (1972), 164-192.

[10] J.-M. Dumont and A. Thomas, Systèmes de numération et fonctions fractales relatifs aux substitutions, *Theoret. Comput. Sci.* **65** (2) (1989), 153-169.

[11] S. Ferenczi, Systems of finite rank, *Colloq. Math.* **73** (1) (1997), 35-65.

[12] N. P. Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, Springer-Verlag, Berlin, 2002.

[13] A. S. Fraenkel, The use and usefulness of numeration systems, *Inform. and Comput.* **81** (1) (1989), 46-61.

[14] C. Frougny, Linear numeration systems of order two, *Inform. and Comput.* **77** (3) (1988), 233-259.

[15] C. Frougny and B. Solomyak, On representation of integers in linear numeration systems, in *Ergodic Theory of* $\mathbf{Z}^d$ *Actions*, Cambridge University Press, Cambridge, 1996, pp. 345-368.

[16] C. Fuchs and R. Tijdeman, Substitutions, abstract number systems and the space filling property, *Ann. Inst. Fourier (Grenoble)* **56** (7) (2006), 2345-2389.

[17] F. Q. Gouvêa, *p-adic Numbers. An Introduction*, Springer, Cham, 2020.

[18] P. J. Grabner, P. Liardet, and R. F. Tichy, Odometers and systems of numeration, *Acta Arith.* **70** (2) (1995), 103-123.

[19] R. H. Herman, I. F. Putnam, and C. F. Skau, Ordered Bratteli diagrams, dimension groups and topological dynamics, *Internat. J. Math.* **3** (6) (1992), 827-864.

[20] M. Hollander, Greedy numeration systems and regularity, *Theory Comput. Syst.* **31** (2) (1998), 111-133.

[21] Intel, *Intel 64 and IA-32 Architectures Software Developer's Manual: Basic Architecture*, 2023.

[22] D. E. Knuth, *The Art of Computer Programming. Vol. 2*, Addison-Wesley, Reading, 1998.

[23] S. Labbé, A self-similar aperiodic set of 19 Wang tiles, *Geom. Dedicata* **201** (2019), 81-109.

[24] S. Labbé and J. Lepšová, A Fibonacci analogue of the two's complement numeration system, *RAIRO, Theor. Inform. Appl.* **57** (2023), 23 pages.

[25] S. Labbé and J. Lepšová, A numeration system for Fibonacci-like Wang shifts, in *Combinatorics on words*, Springer, Cham, 2021, pp. 104-116.

[26] P. Lecomte and M. Rigo, On the representation of real numbers using regular languages, *Theory Comput. Syst.* **35** (1) (2002), 13-38.

[27] P. B. A. Lecomte and M. Rigo, Numeration systems on a regular language, *Theory Comput. Syst.* **34** (1) (2001), 27-44.

[28] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press, Cambridge, 2002.

[29] I. F. Putnam, *Cantor Minimal Systems*, American Mathematical Society, Providence, 2018.

[30] G. Rauzy, Nombres algébriques et substitutions, *Bull. Soc. Math. France* **110** (1982), 147-178.

[31] M. Rigo, *Formal Languages, Automata and Numeration Systems*, Wiley-ISTE, London, 2014.

[32] M. Rigo and A. Maes, More on generalized automatic sequences, *J. Autom. Lang. Comb.* **7** (3) (2002), 351-376.

[33] P. Surer, Substitutions, coding prescriptions and integer representation, *J. Number Theory* **190** (2018), 367-393.

[34] P. Surer, Substitutive number systems, *Int. J. Number Theory* **16** (8) (2020), 1709-1751.