A Succinct Study of Positionality for Dumont–Thomas Numeration Systems

Savinien Kreczman[1(✉)][0000−0002−1928−0028], Sébastien Labbé[2][0000−0002−7330−2718], and Manon Stipulanti[1][0000−0002−2805−2465]

[1] Department of Mathematics, ULiège, Liège, Belgium
{savinien.kreczman,m.stipulanti}@uliege.be
[2] LaBRI, Université de Bordeaux, Talence, France
sebastien.labbe@labri.fr

Abstract. Numeration systems are maps between a set of numbers and a set of words that act as representations of these numbers. One desirable property is positionality: the ability to relate positions in the words to values of the numbers. In general, positionality is hard to decide. In this article, we obtain a criterion to decide the positionality of so-called Dumont–Thomas numeration systems, arising from substitutions. Then, we particularize this criterion to some well-behaved classes of substitutions, allowing us to link the related systems to existing literature.

Keywords: Morphism · Substitution · Periodic points · Numeration system · Positionality · Bertrand property · Fabre substitution

1 Introduction

The need of representing natural numbers (or expressing them in writing) has occupied humans for centuries. In its most general form, this leads to so-called *abstract numeration systems*, introduced by Lecomte and Rigo in 2001 [10] (see also [1, Chapter 3]). Such a numeration system is defined by a triple $S = (L, A, \prec)$ where A is an alphabet ordered by the total order $\prec$ and L is an infinite *regular* language over A (i.e., accepted by a deterministic finite automaton). We say that L is the *numeration language* of S. When we order the words of L with the *genealogical* order (i.e., first by length, then using the dictionary order) induced by $\prec$, we obtain a one-to-one correspondence rep_S between $\mathbb{N}$ and L. The *(S-) representation* of the non-negative integer n is then the $(n+1)$st word of L, and the inverse map, called the *(S-)evaluation map*, is denoted by val_S. A simple example is given by the abstract numeration system S built on the language $L = \mathtt{1}^*\mathtt{2}^*$ over the ordered alphabet $\{\mathtt{1}, \mathtt{2}\}$. The first few words in the language are $\varepsilon, \mathtt{1}, \mathtt{2}, \mathtt{11}, \mathtt{12}, \mathtt{22}, \mathtt{111}$. For example, $\mathrm{rep}_S(5) = \mathtt{22}$ and $\mathrm{val}_S(\mathtt{111}) = 6$.

In general, a numeration system $S = (L, A, S)$ is *positional* if the underlying alphabet A is a set of consecutive integers $\{\mathtt{0}, \mathtt{1}, \ldots, \mathtt{c}\}$ for some $\mathtt{c} \in \mathbb{N}$ and if there exists a sequence $(U_i)_{i\geq 0}$ of non-negative integers such that the evaluation map is of the form $\mathrm{val}_S \colon A^* \to \mathbb{N}, w_{k-1} \cdots w_0 \mapsto \sum_{i=0}^{k-1} w_i U_i$. We use the term *positional* because the *positions* of letters in representations are used to generate numbers. Observe that the numeration system built on $L = \mathtt{1}^*\mathtt{2}^*$ cannot be positional: indeed, we have $\mathrm{rep}_S(3) = \mathtt{11}$ and $\mathrm{rep}_S(5) = \mathtt{22}$, so there is no hope to find an integer sequence $(U_i)_{i\geq 0}$ such that $3 = 1 \cdot U_1 + 1 \cdot U_0$ and $5 = 2 \cdot U_1 + 2 \cdot U_0$ (see also [1, Example 3.1.12]). We thus raise the following question (see also [1, Exercise 3.13]):

Question 1. What are the conditions for an abstract numeration system to be positional?

As this question seems difficult to answer in its full generality, we consider a particular case with Question 2 below. Roughly, we consider numeration systems that are derived from *substitutions*, i.e., maps sending sequences to sequences and satisfying some mild properties. These numeration systems are due to Dumont and Thomas [4] in 1989 and are quite classical. See [6,15] for some examples of applications. In this paper, we exhibit conditions on the underlying substitution so that the corresponding Dumont–Thomas numeration for $\mathbb{Z}$ is positional.

The outline of the paper is as follows. Section 2 gives the necessary background and preliminary results, including the extension of Dumont–Thomas numerations to all integers and all words lengths. In Section 3, we study which Dumont–Thomas numeration systems are positional to answer Question 2. We start with a sketch of the argument in Section 3.1, then we state our main result in Section 3.2. We turn to particular cases in Section 3.3 and we finish by discussing the properties of our Dumont–Thomas numeration systems related to existing literature, e.g., the property of a numeration system to be Bertrand [2,3]. All proofs omitted due to space constraints are available online in [7].

2 Preliminaries

General combinatorics on words. We assume that the reader is familiar with basic notions of combinatorics on words –alphabet, word, length, factor, prefix, left- and right-infinite words, lexicographic order– and we refer the unfamiliar reader to [1, Section 1.2] for an introduction to these terms. We set some notation: we let A denote an alphabet, $|w|$ the length of a finite word w, w_i the i-th letter of w (indexed from 0), $w_{[i,j]}$ the factor of w going from positions i to j included, and $\prec_{\mathrm{lex}}$ is the lexicographic order.

We let $A^{\mathbb{D}}$ be the set of words indexed by $\mathbb{D}$ for $\mathbb{D} \in \{\mathbb{N}, \mathbb{Z}_{<0}, \mathbb{Z}\}$. We speak respectively of right-infinite, left-infinite and two-sided words. For convenience, we separate by a vertical bar the -1-th and 0-th elements of a two-sided word to indicate the origin, i.e., $u = \cdots u_{-3}u_{-2}u_{-1}|u_0u_1u_2\cdots$ when $u \in A^{\mathbb{Z}}$.

Morphisms and substitutions. Given alphabets A, B, a *morphism* is a map $\mu\colon A^* \to B^*$ such that $\mu(uv) = \mu(u)\mu(v)$ for all words $u, v \in A^*$. A morphism is entirely determined by the images of the letters of A. A *substitution* is a morphism $\mu\colon A^* \to A^*$ such that the image $\mu(a)$ is non-empty for every letter $a \in A$ and there exists a *growing* letter $a \in A$, i.e., $\lim_{n\to+\infty} |\mu^n(a)| = +\infty$. A morphism $\mu\colon A^* \to A^*$ is *primitive* if there exists an integer $k \in \mathbb{N}$ such that for all $a, b \in A$, the letter a appears in $\mu^k(b)$. In this case, we also have that a appears in $\mu^\ell(b)$ for all $a, b \in A$ and $\ell \geq k$.

With a morphism $\mu\colon A^* \to A^*$ and a letter $a \in A$ we can associate a directed ordered tree $\mathcal{T}_{\mu,a}$ as follows. The root is labeled by a, and if a node of the tree is labeled by x and $\mu(x) = y_0 \cdots y_\ell$ then that node has $\ell + 1$ children labeled $y_0, \ldots, y_\ell$, with the edge from x to y_i labeled by i. Note that the k-th level of $\mathcal{T}_{\mu,a}$ stores the k-th iteration of μ on the letter a. We say that a node is *in column* n if it is the n-th node of its level in the order of the tree (indexed from 0).

Substitutions can naturally be applied to two-sided words by setting

$$\mu(\cdots u_{-3}u_{-2}u_{-1}|u_0u_1u_2\cdots) = \cdots \mu(u_{-3})\mu(u_{-2})\mu(u_{-1})|\mu(u_0)\mu(u_1)\mu(u_2)\cdots$$

Let $\mathbb{D} \in \{\mathbb{N}, \mathbb{Z}, \mathbb{Z}_{<0}\}$ and consider a substitution μ over A. A word $u \in A^{\mathbb{D}}$ is a *periodic point of* μ if there exists an integer $p \geq 1$ such that $\mu^p(u) = u$. In this case, p is called a *period* of the periodic point u. The smallest such integer is called *the period* of u. A periodic point of μ with period $p = 1$ is called a *fixed point* of μ. We let $\mathrm{Per}_{\mathbb{D}}(\mu) = \{u \in A^{\mathbb{D}} \mid \mu^p(u) = u \text{ for some } p \geq 1\}$ denote the set of periodic points of μ. If $u \in \mathrm{Per}_{\mathbb{Z}}(\mu)$, then the *seed* of u is the pair of letters $u_{-1}|u_0$. If both letters of the seed of a two-sided periodic point are growing, then the periodic point is defined entirely by its seed. More precisely, we have $u = \lim_{n\to+\infty} \mu^{np}(u_{-1})|\lim_{n\to+\infty} \mu^{np}(u_0)$, where p is a period of u.

By extension of the tree associated with a substitution and a letter as in the previous paragraph, we consider two-sided trees as follow. For a substitution μ over A and two letters $a, b \in A$, the tree $\mathcal{T}_{\mu,b|a}$ is obtained by setting a start root having two children: the left (resp., right) one is reached with an edge of label $\mathtt{1}$ (resp., $\mathtt{0}$) and is the root of the tree $\mathcal{T}_{\mu,b}$ (resp., $\mathcal{T}_{\mu,a}$). We count depth in the tree such that the root is at depth -1. There is exactly one increasing bijection between depth k in this tree and an interval of $\mathbb{Z}$ that maps the rightmost node to $\left|\mu^k(a)\right| - 1$. We say that a node is *in column* n if this bijection maps it to n. Examples of such trees are given later in the paper.

Numeration systems. A *numeration system* over the *domain* $\mathbb{D} \in \{\mathbb{N}, \mathbb{Z}\}$ is a pair of maps between $\mathbb{D}$ and a set of words, i.e., the *representation map* $\mathrm{rep}\colon \mathbb{D} \to A^*$ for some alphabet A, and the *evaluation map* $\mathrm{val}\colon L \to \mathbb{D}$, where $\mathrm{rep}(\mathbb{D}) \subset L \subset A^*$, such that $\mathrm{val} \circ \mathrm{rep} = id_{\mathbb{D}}$. The set $\mathrm{rep}(\mathbb{D})$ is the *language* of the numeration system, while L contains some additional, non-canonical representations.

In general, a numeration system over $\mathbb{D} = \mathbb{N}$ is *positional* if the underlying alphabet A is a set of consecutive integers $\{\mathtt{0}, \mathtt{1}, \ldots, \mathtt{c}\}$ for some $\mathtt{c} \in \mathbb{N}$ and the evaluation map is of the form $\mathrm{val}\colon A^* \to \mathbb{N}, w_{k-1} \cdots w_0 \mapsto \sum_{i=0}^{k-1} w_i U_i$ for some sequence $U = (U_i)_{i\geq 0} \in \mathbb{N}^{\mathbb{N}}$. Over $\mathbb{D} = \mathbb{Z}$, a numeration system is *positional* if the underlying alphabet A is a set of consecutive integers $\{\mathtt{0}, \mathtt{1}, \ldots, \mathtt{c}\}$ for some $\mathtt{c} \in \mathbb{N}$ and the evaluation map is of the form $\mathrm{val}\colon A^* \to \mathbb{Z}, w_{k-1} \cdots w_0 \mapsto \sum_{i=0}^{k-2} w_i U_i - w_{k-1} V_{k-1}$ for some sequences $U = (U_i)_{i\geq 0}, V = (V_i)_{i\geq 0} \in \mathbb{N}^{\mathbb{N}}$. The sequences U and V are the sequences of *weights* of the numeration system. Every position has a given weight, while the presence of an additional sequence V helps deal with the representation of negative numbers, in a fashion similar to the usual two's complement numeration system. Such a numeration system (see [8, Section 1]) can be defined by the evaluation map $w_{k-1}\cdots w_0 \mapsto -w_{k-1}2^{k-1}+\sum_{i=0}^{k-2} w_i 2^i$ together with the language $L = A^* \setminus (\mathtt{00}A^* \cup \mathtt{11}A^*)$ as the evaluation map is a bijection between L and $\mathbb{Z}$. The system is positional with weights $(2^i)_{i\geq 0}$.

Dumont–Thomas numeration systems. Dumont and Thomas [4] originally defined numeration systems based on substitutions and their right-infinite fixed points for the natural domain $\mathbb{D} = \mathbb{N}$. The intuition behind these systems is that the number n can be represented by the label of the shortest path from the root to a node in column n. See [7, Section 2.1] for a gentle introduction. Recently, the authors of [9] proposed an extension of these numeration systems to all integers (so $\mathbb{D} = \mathbb{Z}$) and based instead on some periodic points of substitutions. By relaxing the condition of [9] on the lengths of representations, we go even beyond and use all periodic points of substitutions to define Dumont–Thomas numeration systems.

To formally introduce our object of study, we need some additional definitions and notation. Let $\mu\colon A^* \to A^*$ be a substitution. Fix a letter $a \in A$ and an integer k. For every $i \in \{0,\dots,k\}$, we consider a pair $(m_i, a_i) \in A^* \times A$. The sequence $((m_i, a_i))_{i=0,\dots,k}$ is *admissible (with respect to* μ*)* if for every $i \in \{1,\dots,k\}$, $m_{i-1}a_{i-1}$ is a prefix of $\mu(a_i)$. This sequence is *a-admissible (with respect to* μ*)* if it is admissible with respect to μ and $m_k a_k$ is a prefix of $\mu(a)$.

Theorem 1 (Extension of [9, Theorem 4.1]). *Let $\mu : A^* \to A^*$ be a substitution with growing letter $a \in A$. Consider a right-infinite periodic point $u \in \mathrm{Per}_{\mathbb{N}}(\mu)$ with $u_0 = a$ and period $p \geq 1$. Fix a residue $r \in \{0, 1, \dots, p-1\}$ modulo p and define $v_r = \mu^r(u)$. For every integer $n \geq 0$, there exist a unique integer $k = k(n)$ with $k \equiv r \bmod p$ and a unique sequence $((m_i, a_i))_{i=0,\dots,k-1}$ such that the sequence is a-admissible,*

$$m_{k-1}m_{k-2}\cdots m_{k-p} \neq \varepsilon \text{ if } k \geq p,$$

and $(v_r)_{[0,n-1]} = \mu^{k-1}(m_{k-1})\mu^{k-2}(m_{k-2})\cdots\mu^0(m_0)$.

Theorem 2 (Extension of [9, Theorem 4.2]). *Let $\mu : A^* \to A^*$ be a substitution with growing letter $b \in A$. Consider a left-infinite periodic point $u \in \mathrm{Per}_{\mathbb{Z}_{<0}}(\mu)$ with $u_{-1} = b$ and period $p \geq 1$. Fix a residue $r \in \{0, 1, \dots, p-1\}$ modulo p and define $v_r = \mu^r(u)$. For every integer $n \leq -1$, there exist a unique integer $k = k(n)$ with $k \equiv r \bmod p$ and a unique sequence $((m_i, a_i))_{i=0,\dots,k-1}$ such that the sequence is b-admissible,*

$$\mu^{p-1}(m_{k-1})\mu^{p-2}(m_{k-2})\cdots\mu^0(m_{k-p})a_{k-p} \neq \mu^p(b) \text{ if } k \geq p, \tag{1}$$

and $(v_r)_{[-|\mu^k(b)|,n-1]} = \mu^{k-1}(m_{k-1})\mu^{k-2}(m_{k-2})\cdots\mu^0(m_0)$.

From these results, we generalize the definition of Dumont–Thomas numeration systems [4,9].

Definition 1. *Let $\mu\colon A^* \to A^*$ be a substitution and let $u \in \mathrm{Per}_{\mathbb{Z}}(\mu)$ be a two-sided periodic point with growing seed $u_{-1}|u_0$ and period $p \geq 1$. Let $r \in$*

$\{0, 1, \ldots, p-1\}$. *Define* $\mathtt{c} = \max_{a \in A} |\mu(a)| - 1$ *and the set* $D = \{\mathtt{0}, \mathtt{1}, \ldots, \mathtt{c}\}$. *The* Dumont–Thomas complement numeration system associated with μ, u and r *is defined by the map* $\mathrm{rep}_{u,r} \colon \mathbb{Z} \to \{\mathtt{0}, \mathtt{1}\}D^*, n \mapsto \mathrm{rep}_{u,r}(n)$ *where*

$$\mathrm{rep}_{u,r}(n) = \begin{cases} \mathtt{0} \cdot |m_{k-1}| \cdot |m_{k-2}| \cdots |m_0|, & \text{if } n \geq 0; \\ \mathtt{1} \cdot |m_{k-1}| \cdot |m_{k-2}| \cdots |m_0|, & \text{if } n \leq -1; \end{cases}$$

such that $k = k(n)$ *is the unique integer with* $k \equiv r \bmod p$ *and* $((m_i, a_i))_{i=0,\ldots,k-1}$ *is the unique sequence obtained from Theorem 1 (resp., Theorem 2) applied on the right-infinite periodic point* $u|_{\mathbb{N}} = u_0 u_1 \cdots$ *(resp., the left-infinite periodic point* $u|_{\mathbb{Z}_{<0}} = \cdots u_{-2} u_{-1}$*) with period* p*. Note that when the context is clear, we drop the dependence on* μ*,* u *and* r*.*

The intuition behind this system is to use the tree $\mathcal{T}_{\mu,b|a}$ and represent n with the label of a shortest path of length congruent to $r+1$ modulo p between the root and a node in column n. Note that representations in this system have length congruent to $r+1$ modulo p, and their first digits depend only on the sign of the represented number. Note that we use a special font for the digits representing numbers to distinguish them from integers.

Example 1. Consider the substitution $\mu \colon a \mapsto ccd, b \mapsto cd, c \mapsto ab, d \mapsto a$ and its periodic point $u \in \mathrm{Per}_{\mathbb{N}}(\mu)$ with growing seed $a|a$ and period $p = 2$. The tree $\mathcal{T}_{\mu,a|a}$ is depicted in Fig. 1. Depending on the choice of even or odd length for representations, we obtain different numeration systems as illustrated in the table of Fig. 1.

The numeration systems with $r = 0$ were studied in Lepšová's PhD thesis [11]. Notably, [11, Example 6.5.6] presents two substitutions associated with the silver mean $1 + \sqrt{2}$ such that one gives rise to a positional numeration system and the other does not. Lepšová raised the following natural question [11].

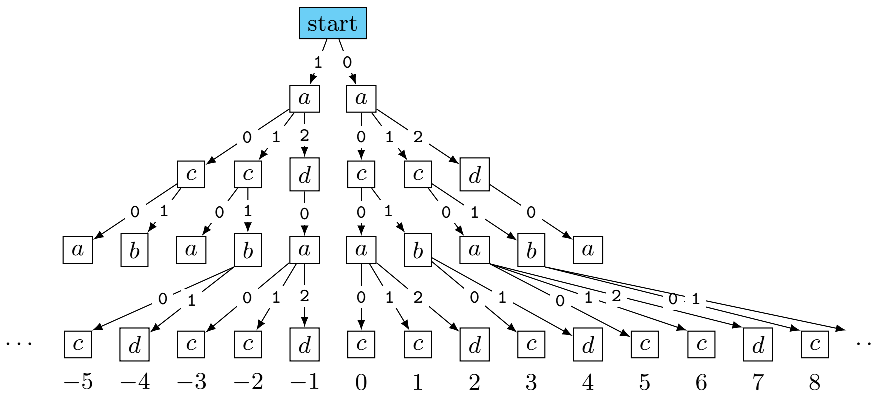
Question 2. [11, Question 6.5.7] What are the conditions for a complement Dumont–Thomas numeration system to be positional?

3 Positional Dumont–Thomas Numeration Systems

In this section, we study when a substitution generates a Dumont–Thomas numeration system that is positional to answer Question 2. We first sketch our argument, and give some intuition for the statement of the main result (Theorem 3). After stating the result in the most general form, we present some corollaries in Section 3.3. We also link our Dumont–Thomas numeration systems to previous literature.

3.1 Sketch of the Argument

We informally sketch the argument that we used to solve Question 2. We then present examples where this argument fails, which allows us to motivate the



n	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9
$\mathrm{rep}_{u,0}(n)$	101	110	111	1	0	001	010	011	020	00100	00101	00110	01000	01001
$\mathrm{rep}_{u,1}(n)$	1111	10	11	12	00	01	02	0010	0011	0100	0101	0102	0110	0111

Fig. 1: The tree $\mathcal{T}_{\mu,a|a}$ for the substitution $\mu\colon a \mapsto ccd, b \mapsto cd, c \mapsto ab, d \mapsto a$ and the periodic point u of period $p = 2$ and seed $a|a$. Below, depending on the residue $r \in \{0, 1\}$, we obtain a Dumont–Thomas numeration system and we give $(\mathrm{rep}_{u,r}(n))_{-4 \le n \le 9}$ whose lengths are congruent to $r + 1 \bmod p$.

technicalities that are introduced in Section 3.2 and to explain the reasoning without them getting in the way of the explanation.

Sketch 1. We let $\mu\colon A^* \to A^*$ be a substitution, $u \in \mathrm{Per}_{\mathbb{Z}}(\mu)$ be a two-sided periodic point of μ with growing seed $b|a$ and period $p \ge 1$, and $r \in \{0, 1, \ldots, p-1\}$ be a residue. We also consider the corresponding Dumont–Thomas complement numeration system associated with μ, u and r.

Now consider two words $wt0^\ell$ and $w(t+1)0^\ell$ that label paths in the tree $\mathcal{T}_{\mu,b|a}$ (see Fig. 2, where we have arbitrarily chosen to represent the situation with $\mathcal{T}_{\mu,a}$ instead) and assume these words are the representations of the integers n_1 and n_2 respectively, i.e., the paths end in columns n_1 and n_2, in addition to which we have some conditions for admissibility. We also let c and d be the letters attained after reading wt and $w(t + 1)$ respectively.

If the numeration system is positional, then $n_2 - n_1$ must equal U_ℓ, the weight in position ℓ, as the representations of n_1 and n_2 differ only by one unit in position ℓ. However, $n_2 - n_1$ is the number of columns between those two nodes in the tree, which is $\left|\mu^\ell(c)\right|$, i.e., the number of level-ℓ descendants of c in the tree (see again Fig. 2). Since this reasoning could be applied to any letter c that has a sibling to its right in $\mathcal{T}_{\mu,b|a}$, this points towards the following implication: *If the Dumont–Thomas numeration system associated with μ, u, and r is positional, then $\left|\mu^\ell(c)\right| = U_\ell$ for every letter c that has a younger sibling in $\mathcal{T}_{\mu,b|a}$.*

The converse implication is also justified, almost by definition of a Dumont–Thomas numeration system. If the positive integer $n \ge 1$ is represented by the word $\mathtt{0} \cdot w_{k-1} \cdots w_0$, this means that there is an a-admissible sequence

$((m_i, a_i))_{i=0,\ldots,k-1}$ such that

$$\mu^r(u)_{[0,n-1]} = \mu^{k-1}(m_{k-1}) \ldots \mu^0(m_0) \tag{2}$$

and $|m_i| = w_i$ for every $i \in \{0, \ldots, k-1\}$. Because all letters in m_i have a_i as a younger sibling, their image by μ^ℓ has length U_ℓ by assumption. Taking the length in Eq. (2) yields $n = \sum_{i=0}^{k-1} U_i w_i$, which corresponds to the numeration system being positional with weights $(U_i)_{i\geq 0}$. The case of negative numbers is similar, with one correcting term corresponding to the value of V_{k-1}.

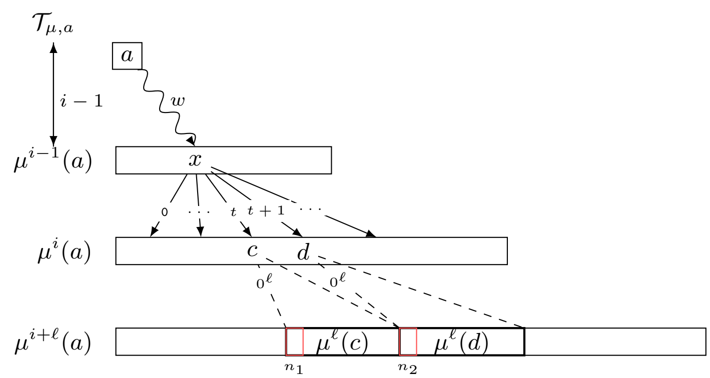


Fig. 2: Comparing the values of $wt0^\ell$ and $w(t+1)0^\ell$ in the right part of $\mathcal{T}_{\mu,b|a}$.

The above sketch leads us to formulate the next conjecture: *The Dumont–Thomas numeration system associated with μ, u, and r is positional if and only if $c \mapsto |\mu^\ell(c)|$ is constant on all letters c that have a younger sibling in $\mathcal{T}_{\mu,b|a}$, in which case this constant is the weight U_ℓ.* However, trying to prove this conjecture reveals two issues with Sketch 1, which the following examples highlight.

Example 2. Recall the substitution μ from Example 1. The letters having a younger sibling in $\mathcal{T}_{\mu,a}$ are a and c, but the sequences of the lengths of their consecutive images under μ are $(|\mu^j(a)|)_{j\geq 0} = 1, 3, 5, 13, 21, 55, 89, 233, 377, \ldots$ and $(|\mu^j(c)|)_{j\geq 0} = 1, 2, 5, 8, 21, 34, 89, 144, 377, \ldots$ respectively. Despite this, the corresponding Dumont–Thomas numeration system is positional for both values of r, with weights $1, 2, 5, 8, 21, 34, \ldots$ for $r = 0$ and $1, 3, 5, 13, 21, 55, \ldots$ for $r = 1$.

This example illustrates that, in the case where μ is not primitive, we may only control some image lengths among all different letters. Since a and c are never both present on some level of the tree, they may have different behaviors. Fixing this is the purpose of the sets E_j in Definition 2.

Example 3. Consider the substitution $\mu\colon a \mapsto bca$, $b \mapsto bb$, $c \mapsto b$ and the associated numeration system with the seed $a|b$. The letters b and c both appear in the tree with a younger sibling, and they have images of different lengths. However, the numeration system is still positional, with weights $U_i = 2^i$ and $V_0 = 1$, $V_i = 3 \cdot 2^{i-1}$ for every $i \geq 1$.

Our sketched argument fails, because if wt is a path to a node labeled by c in the tree, this node is in column -2. Then, $w(t+1)$ leads to a node in column -1 and $w(t+1)0^\ell$ is never the representation of any number, due to Eq. (1) from Theorem 2. Thus we cannot constrain the lengths of the images of c. This case and some similar ones may occur when a letter appears only in column -2, which motivates the separate treatment of column -2 found in Definition 3.

3.2 Main Result

To state the main result, we need to define some particular sets of letters. We again consider a substitution $\mu\colon A^* \to A^*$ and a two-sided periodic point $u \in \mathrm{Per}_{\mathbb{Z}}(\mu)$ of μ with growing seed $b|a$ and period $p \geq 1$. We draw the tree $\mathcal{T}_{\mu,b|a}$. For a fixed residue $r \in \{0, \dots, p-1\}$, we also consider the corresponding Dumont–Thomas numeration system $\mathrm{rep}_{u,r}\colon \mathbb{Z} \to \{\mathtt{0},\mathtt{1}\}D^*$ from Definition 1.

Definition 2. *Let $j \in \{0, \dots, p-1\}$. We let E_j be the set of letters $c \in A$ such that there exist some integer $k \geq 1$ and some sequence $((m_i, a_i))_{i=0,\dots,k-1} \in (A^* \times A)^k$ that verifies the following:*

- *the sequence $((m_i, a_i))_{i=0,\dots,k-1}$ is a- or b-admissible;*
- *we have $k \equiv j \bmod p$;*
- *the letter c appears at the end of the word m_0;*
- *if $((m_i, a_i))_{i=0,\dots,k-1}$ is b-admissible, then $\mu^{k-1}(m_{k-1}) \cdots \mu^0(m_0)a_0 \neq \mu^k(b)$.*

Thinking in terms of the tree $\mathcal{T}_{\mu,b|a}$, the first three conditions simply describe letters that appear in the tree at some level congruent to $j \bmod p$ and have a younger sibling on that level. The last condition excludes letters c that only appear in column -2 in the tree $\mathcal{T}_{\mu,b|a}$. To deal with these letters, a dedicated condition is required, as follows.

Definition 3. *Let $j \in \{0, \dots, p-1\}$. On level j in the tree $\mathcal{T}_{\mu,b|a}$, we let c be the letter in column -2 and d be the letter in column -1 (i.e., d is immediately to the right of c). If c and d share the same parent, we consider two cases to modify E_j. If $\left|\mu^{p-j}(d)\right| > 1$, then we add c to E_j if it was not already present. If $\left|\mu^{p-j}(d)\right| = 1$ and $j \leq r < p$, then we add the following condition:*

$$\left|\mu^{r-j}(c)\right| \text{ must be equal to } \left|\mu^{r-j}(e)\right| \text{ for every letter } e \in E_j. \tag{3}$$

Example 4. For the substitution μ from Example 1, consider its two-sided periodic point with seed $a|a$ and period $p = 2$. In this case, we obtain $E_0 = \{a\}$ and $E_1 = \{c\}$. Condition (3) must only be checked for $r = 1$ and the letter c, and it is trivially satisfied.

Consider the substitution $\mu\colon a \mapsto bcd$, $d \mapsto ba$, $b \mapsto b^2$, $c \mapsto b$ and its two-sided periodic point with seed $a|b$. If we go only by Definition 2, we will find

$E_0 = E_1 = \{b\}$, but if we add Definition 3, c is added to E_1, and we now correctly find that the system is not positional (which we can also see from the representations of -5, -3 and -2).

Theorem 3. *Let $\mu\colon A^* \to A^*$ be a substitution and let $u \in \mathrm{Per}_{\mathbb{Z}}(\mu)$ be a two-sided periodic point with growing seed $b|a$ and period $p \geq 1$. The Dumont–Thomas complement numeration system associated with μ, u, and r is positional if and only if for every $j \in \{0, \ldots, p-1\}$, the map $c \mapsto \left|\mu^\ell(c)\right|$ is constant over E_j for every ℓ such that $\ell + j \equiv r \bmod p$, and condition (3) is satisfied for the letters where it was added.*

In this case, the sequences U, V of weights of the numeration system are given as follows: for every $\ell \geq 0$, we define $U_\ell = \left|\mu^\ell(c)\right|$ for a letter $c \in E_j$ where $j \in \{0, \ldots, p-1\}$ and $\ell + j \equiv r \bmod p$; and $V_\ell = \left|\mu^\ell(b)\right|$ for every $\ell \in \mathbb{N}$.

Remark 1. It may be, although rarely, that E_j is empty. (For example, consider the case of $\mu\colon a \mapsto b$, $b \mapsto aa$ and the set E_1.) In this case, if Condition (3) applies to some letter c, we may give the weight $\left|\mu^{r-j}(c)\right|$ to position $r - j$. Otherwise, this means that only the digit 0 appears at positions congruent to $j \bmod p$ in this numeration system, and as such the weight given to these positions is arbitrary.

Remark 2. Replacing the substitution by one of its powers may lead to the loss of positionality. This is the case of the substitution μ from Example 1 and its square, for which the numeration systems associated with seed $a|a$ are once positional, once not positional (consider e.g. the representations of 5 and 8).

3.3 Particular Cases

We now highlight some general cases where the technicalities of Section 3.2 do not occur, leading to results that are more concise and legible. We also discuss possible simplifications of the substitution at play, as well as a parallel to Bertrand numeration systems.

From now on, we assume that the alphabet A of the substitution μ is *minimal*, i.e, all the letters in A are present in $\mu^n(b|a)$ for some n. Given a substitution μ, we say that a letter c is *non-final* if there exist $d \in A$, $x \in A^*$, and $y \in A^+$ such that $\mu(d) = xcy$. We let E_μ denote the set of non-final letters of μ. For example, with the Fibonacci substitution $\mu\colon a \mapsto ab, b \mapsto a$, we have $E_\mu = \{a\}$.

When $\mathbb{D} = \mathbb{N}$ and the substitution μ has a fixed point or when μ is primitive (even over $\mathbb{D} = \mathbb{Z}$), the reasoning in Sketch 1 holds, as stated in the next results.

Corollary 1. *Let $\mu\colon A^* \to A^*$ be a substitution and $u = u_0 u_1 \cdots$ be a right-infinite fixed point of μ with growing seed a. Then the Dumont–Thomas numeration system (for $\mathbb{N}$) associated with μ, u and $r = 0$ is positional if and only if the map $c \mapsto \left|\mu^\ell(c)\right|$ is constant over E_μ for every ℓ, in which case the sequence of weights is equal to $U_\ell = \left|\mu^\ell(a)\right|$ for every $\ell \geq 0$.*

Corollary 2. *Let $\mu\colon A^* \to A^*$ be a primitive substitution and let $u \in \mathrm{Per}_{\mathbb{Z}}(\mu)$ be a two-sided periodic point with growing seed $b|a$ and period $p \geq 1$. The Dumont–Thomas complement numeration system associated with μ, u, and r is positional* *if and only if the map* $c \mapsto \left|\mu^\ell(c)\right|$ *is constant over* E_μ *for every* $\ell \geq 0$*. In this case, for every* $\ell \geq 0$*,* U_ℓ *is the constant value of* $\left|\mu^\ell(\cdot)\right|$ *over* E_μ *and* $V_\ell = \left|\mu^\ell(b)\right|$*.*

In both of these cases, we may use the following lemma to simplify the substitution at play by reducing the number of non-final letters to one.

Lemma 1. *Let* $\mu\colon A^* \to A^*$ *be a substitution such that the map* $c \mapsto \left|\mu^\ell(c)\right|$ *is constant over* E_μ *for every integer* $\ell \geq 0$*. Then there exist* $B \subseteq A$*, a substitution* $\nu\colon B^* \to B^*$ *with* $|E_\nu| = 1$*, and* $a', b' \in B$ *such that the trees* $\mathcal{T}_{\mu,b|a}$ *and* $\mathcal{T}_{\nu,b'|a'}$ *differ only by their labeling. These objects can be computed effectively.*

Example 5. Consider the primitive substitution $\mu\colon a \mapsto ab, b \mapsto ba$ (often referred to as the Thue–Morse substitution). We note that both a, b are non-final letters and $\left|\mu^\ell(c)\right| = 2^\ell$ for $c \in \{a, b\}$ and for every $\ell \geq 0$. The substitution given by Lemma 1 is $\nu\colon a \mapsto a^2$. The corresponding Dumont–Thomas numeration system (over $\mathbb{N}$) is the usual binary system.

We now link some of our numeration systems to existing literature. Our target numeration systems represent natural numbers and use words of every length, so they correspond to Corollary 1. Due to Lemma 1, it is enough to consider substitutions with only one non-final letter e, which must be the seed of the fixed point. The image of any letter is defined only by the number of repetitions of e and the choice of final letter. As a result, if we operate on the minimal alphabet, the substitution is equivalent to one of the form

$$\mu : a_1 \mapsto a_1^{d_1} a_2,\ a_2 \mapsto a_1^{d_2} a_3,\ \ldots, a_n \mapsto a_1^{d_n} a_k, \tag{4}$$

where $n \geq 1$ is an integer, $\{a_1, \ldots, a_n\}$ is the alphabet, $k \in \{1, \ldots, n\}$, $d_1 > 0$ and d_i is a non-negative integer for every $i \in \{1, \ldots, n\}$. An example is the Tribonacci substitution $\tau\colon a \mapsto ab, b \mapsto ac, c \mapsto a$, for which $n = 3$, $a_1 = a$, $a_2 = b$, $a_3 = c$, $d_1 = d_2 = 1$, $d_3 = 0$, and $k = 1$. The similarity with the substitutions studied by Fabre in [5] is striking, so we call *Fabre-like* the substitutions of the form (4). We now study this particular class of substitutions in additional detail.

We quickly recall some other milestones of numeration systems. For the reader not familiar with the classical theory, see, e.g., [1, Chapter 2] and [13, Chapter 2]. In 1957, Rényi [14] introduced β-numeration systems to represent positive real numbers with a real base $\beta > 1$. We let $d_\beta(1)$ (resp., $d^*_\beta(1)$) denote the β-representation (resp., quasi-greedy β-representation) of 1. Of particular interest are the so-called *Parry numbers*, which are real numbers β such that $d_\beta(1)$ is either finite (β is then a *simple Parry* number) or ultimately periodic [12]. An article of Bertrand-Mathis [2], later corrected by Charlier, Cisternino and Stipulanti [3], studies the greedy positional numeration systems such that the associated language L verifies $w \in L \Leftrightarrow w\mathtt{0} \in L$. As it turns out, these systems are exactly those that verify the relation $U_i = d_1 U_{i-1} + d_2 U_{i-2} + \ldots + d_i U_0 + 1$ for all $i \geq 1$, where $d_1 d_2 \cdots$ is either $d^*_\beta(1)$ for some $\beta > 1$ (called *canonical* Bertrand numeration systems), or $d_\beta(1)$ for some simple Parry number β (called *non-canonical* Bertrand numeration systems and introduced in [3]), or $d_1 d_2 \cdots = 10^\omega$

(called the *trivial* Bertrand numeration system, also from [3]). In 1995, Fabre [5] introduced another way to approach canonical Bertrand numeration systems. If β is a Parry number with $d_\beta^*(1) = d_1 \cdots d_n (d_{n+1} \cdots d_{n+m})^\omega$ for some m, n, we introduce the substitution μ_β mapping $1 \mapsto 1^{d_1} 2, \ldots, (n+m-1) \mapsto 1^{d_{n+m-1}}(n+m)$, and $(n+m) \mapsto 1^{d_{n+m}}(n+1)$. Note that simple Parry numbers correspond to $n = 0$. Fabre then shows that $|\mu_\beta^\ell(1)|$ is the integer U_ℓ defined for the canonical Bertrand numeration system, and his [5, Theorem 2] establishes the equality between the Dumont–Thomas numeration system associated with μ_β and the canonical Bertrand numeration system associated with β.

Let us now go back to the study of our Dumont–Thomas numeration systems based on Fabre-like substitutions. We first note that Bertrand numeration systems based on a Parry number are a particular case of Dumont–Thomas numeration systems (and not just the canonical ones as proven by Fabre).

Proposition 1. *Every Bertrand numeration system associated with a Parry number is equal to some Dumont–Thomas numeration system associated with a Fabre-like substitution.*

Although our Dumont–Thomas numeration systems verify the property $w \in L \Leftrightarrow w\mathtt{0} \in L$ that characterizes Bertrand numeration systems, the converse of Proposition 1 is not true, as we see in the following example.

Example 6. Consider the Fabre-like substitution $\mu\colon a \mapsto aab,\ b \mapsto aaaa$ with fixed point $u = aabaabaaaa\cdots$. The corresponding Dumont–Thomas numeration system is positional, with the sequence of weights starting by $1, 3, 10, 32, \ldots$. We note that $\mathrm{rep}_{r,0}(9) = \mathtt{23}$, but this cannot happen in a Bertrand numeration system as the representation of 9 would be $\mathtt{30}$ with the given weights.

This is because our systems are not greedy in the usual sense. This phenomenon can be understood with an adaptation of the *Parry condition*. To recall, this condition (first seen in [12, Corollary 1]) states that a word $d_1 d_2 \cdots$ is equal to $d_\beta(1)$ for some $\beta > 1$ if and only if $d_1 d_2 \cdots \succ 10^\omega$ and $d_1 d_2 \cdots \succ d_i d_{i+1} \cdots$ for every $i \geq 1$. In the case of Example 6, if the substitution μ were the Fabre substitution associated with some Parry number β, we would have $d_\beta^*(1) = (23)^\omega$ and $d_\beta(1) = 240^\omega$, which contradicts the Parry condition. Thus the system cannot be equal to a Bertrand numeration system. In fact, the Parry condition (adapted for use with $d_\beta^*(1)$ instead of $d_\beta(1)$) is all that is necessary to guarantee that the system is greedy and thus equal to a Bertrand numeration system.

Proposition 2. *Let μ be a Fabre-like substitution as in* (4)*. Construct the right-infinite word $d_1 d_2 \cdots = d_1 \cdots d_{k-1}(d_k \cdots d_n)^\omega$. The Dumont–Thomas numeration system associated with μ and the seed a_1 is equal to a Bertrand numeration system if and only if we have $d_i d_{i+1} \cdots \preccurlyeq_{lex} d_1 d_2 \cdots$ for each $i \geq 1$.*

Acknowledgments. Savinien Kreczman is supported by the FNRS Research Fellow grant 1.A.789.23F. Sébastien Labbé is supported by France's Agence Nationale de la Recherche (ANR) project IZES (ANR-22-CE40-0011). Manon Stipulanti is supported by the FNRS Research grant 1.C.104.24F. The authors want to thank Émilie Charlier for useful discussions.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Berthé, V., Rigo, M. (eds.): Combinatorics, Automata, and Number Theory, Encyclopedia of Mathematics and its Applications, vol. 135. Cambridge University Press, Cambridge (2010). https://doi.org/10.1017/CBO9780511777653
2. Bertrand-Mathis, A.: Comment écrire les nombres entiers dans une base qui n'est pas entière. Acta Math. Hungar. **54**(3-4), 237–241 (1989). https://doi.org/10.1007/BF01952053
3. Charlier, E., Cisternino, C., Stipulanti, M.: A full characterization of Bertrand numeration systems. In: Developments in language theory, Lecture Notes in Comput. Sci., vol. 13257, pp. 102–114. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-05578-2_8
4. Dumont, J.M., Thomas, A.: Systèmes de numération et fonctions fractales relatifs aux substitutions. Theoret. Comput. Sci. **65**(2), 153–169 (1989). https://doi.org/10.1016/0304-3975(89)90041-8
5. Fabre, S.: Substitutions et β-systèmes de numération. Theoret. Comput. Sci. **137**(2), 219–236 (1995). https://doi.org/10.1016/0304-3975(95)91132-A
6. Gheeraert, F., Romana, G., Stipulanti, M.: String attractors of fixed points of k-Bonacci-like morphisms. In: Combinatorics on words, Lecture Notes in Comput. Sci., vol. 13899, pp. 192–205. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-33180-0_15
7. Kreczman, S., Labbé, S., Stipulanti, M.: Positionality of Dumont–Thomas numeration systems for integers (2025), preprint available at http://arxiv.org/abs/2503.04487
8. Labbé, S., Lepšová, J.: A Fibonacci analogue of the two's complement numeration system. RAIRO Theor. Inform. Appl. (RAIRO:ITA) **57**, Paper No. 12, 23 (2023). https://doi.org/10.1051/ita/2023007
9. Labbé, S., Lepšová, J.: Dumont-Thomas complement numeration systems for $\mathbb{Z}$. Integers **24**, Paper No. A112, 27 (2024). https://doi.org/10.5281/zenodo.14340125
10. Lecomte, P., Rigo, M.: Numeration systems on a regular language. Theory Comput. Syst. **34**(1), 27–44 (2001). https://doi.org/10.1007/s002240010014
11. Lepšová, J.: Substitutive structures in combinatorics, number theory, and discrete geometry. PhD Thesis, Université de Bordeaux and Czech Technical University in Prague (2024), available online at https://theses.fr/2024BORD0083 or https://theses.hal.science/tel-04679032
12. Parry, W.: On the β-expansions of real numbers. Acta Math. Acad. Sci. Hung. **11**, 401–416 (1960). https://doi.org/10.1007/BF02020954
13. Rigo, M.: Formal languages, automata and numeration systems. Vol. 2. Networks and Telecommunications Series, ISTE, London; John Wiley & Sons, Inc., Hoboken, NJ (2014). https://doi.org/10.1002/9781119042853
14. Rényi, A.: Representations for real numbers and their ergodic properties. Acta Math. Acad. Sci. Hung. **8**(3-4), 477–493 (Sep 1957). https://doi.org/10.1007/BF02020331
15. Surer, P.: Substitutive number systems. Int. J. Number Theory **16**(8), 1709–1751 (2020). https://doi.org/10.1142/S1793042120500906