Games derived from a generalized Thue-Morse word

Aviezri S. Fraenkel,
Dept. of Computer Science and Applied Mathematics,
Weizmann Institute of Science,
Rehovot 76100, Israel; fraenkel@wisdom.weizmann.ac.il
Michel Rigo,
University of Liège Department of Mathematics,
Allée de la découverte 12, B-4000 Liège, Belgium;
m.rigo@ulg.ac.be

December 31, 2015

Abstract

For fusing combinatorial game theory with combinatorics on words, we begin with some relevant background on words and automata theory, followed by devising and analyzing a triple of games derived from a generalization of the Thue-Morse word.

1 Introduction

It is always nice to make connections between different areas of mathematics. There are already some connections between combinatorial game theory and combinatorics on words. But they are few [5, 6, 7, 8]. We thought that it would be interesting to explore the possibility of deriving a game from the ubiquitous Thue-Morse word or even from a generalization thereof.

In this spirit we formulate three game rule sets for three new games derived from a generalization of the Thue-Morse word, and present their winning strategies.

In section 2 we give some basics on infinite words. In section 3 we present connections between words and games, and in the central section 4 we present and analyse a triple of games derived from the generalized Thue-Morse word. We wrap up in the final section 5.

2 Basics on infinite words

We assume that the reader is more familiar with combinatorial game theory than with words and automata theory. Let us start with a few classical definitions from combinatorics on words. For a general reference, see, for instance, [10, 12].

Definition 1. An *alphabet* is a finite set. The elements of the alphabet are usually called *letters*. Let A be an alphabet. A *finite word* of length $\ell \geq 1$ is a finite sequence of letters or, formally, a map $w : \{0, \ldots, \ell - 1\} \to A$. We usually write w_i instead of $w(i)$, where $w(i)$ denotes the i-th letter of w. The *length* of w is denoted by $|w|$. The *empty word* is denoted by ε. It has length zero and corresponds to the empty sequence. The set of finite words over A is denoted by A^*. Endowed with the product of concatenation of words, A^* is a monoid with neutral element ε. An *infinite word* over A is an element in $A^{\mathbb{N}}$, i.e., a map from $\mathbb{N}$ to A.

Definition 2. A *morphism* (or precisely, an endomorphism) of A^* is a map $f : A^* \to A^*$ such that $f(uv) = f(u)f(v)$ for all $u, v \in A^*$. If $f(a) = au$ with $a \in A$ and $u \in A^*$ and if $\lim_{n \to +\infty} |f^n(a)| = +\infty$, then f is said to be *prolongable* on a.

In this paper, we only deal with *non-erasing morphisms*, i.e., $|f(a)| \geq 1$ for all letters $a \in A$.

Example 1. The morphism $f_2 : \{0, 1\}^* \to \{0, 1\}^*$ defined by $f(0) = 01$ and $f(1) = 10$ is prolongable on both letters 0 and 1. Indeed, $|f^n(0)| = |f^n(1)| = 2^n$ for all $n > 0$. This morphism is called the *Thue–Morse morphism*.

Let $\mathbf{x}, \mathbf{y} \in A^{\mathbb{N}}$. One can define in a natural way a distance d over $A^{\mathbb{N}}$ by

$$d(\mathbf{x}, \mathbf{y}) = 2^{-|p(\mathbf{x}, \mathbf{y})|}$$

where $p(\mathbf{x}, \mathbf{y})$ is the longest common prefix of the two words. If $\mathbf{x} = \mathbf{y}$, we set $d(\mathbf{x}, \mathbf{y}) = 0$. We can therefore define converging sequences of elements in $A^{\mathbb{N}}$. We now extend this definition to a sequence of finite words converging to a limit infinite word.

Definition 3. Let α be a new letter not in A. If x is a finite word, we denote by $x\alpha^\omega$ the infinite word obtained by concatenating infinitely many copies of α to the right of x. Formally, it is an infinite word defined by $(x\alpha^\omega)_i = x_i$ if $i < |x|$ and $(x\alpha^\omega)_i = \alpha$ for $i \geq |x|$. Doing this, A^* is embedded into $(A \cup \{\alpha\})^{\mathbb{N}}$.

A sequence $(x^{(n)})_{n\ge 0}$ of finite words over A converges to an infinite word $\mathbf{y} \in A^{\mathbb{N}}$, if the sequence $(x^{(n)}\alpha^\omega)_{n\ge 0}$ of infinite words converges to $\mathbf{y}$. Otherwise stated, for every ℓ, there exists N such that for all $n \ge N$, the words $x^{(n)}$ and $\mathbf{y}$ share a common prefix of length ℓ.

Example 2. Continuing Example 1, the first few iterations of f_2 gives the words

$$0,\ f_2(0) = 01,\ f_2^2(0) = 0110,\ f_2^3(0) = 01101001, \dots,$$

and it is easy to see that $f_2^n(0)$ is the prefix of length 2^n of $f_2^{n+1}(0)$. This will ensure the convergence to a limit infinite word.

In particular, a morphism f of A^* can be extended to a map from $A^{\mathbb{N}}$ to $A^{\mathbb{N}}$. If $\mathbf{w} = w_0w_1w_2\cdots$ is an infinite word, consider the converging sequence $(f(w_0\cdots w_i))_{i\ge 0}$ and we define $f(\mathbf{w})$ as its limit. For references for the next proposition, see, for instance, [1, p. 10] or [12, Section 2.1].

Proposition 1. *Let $f : A^* \to A^*$ be a morphism prolongable on $a \in A$. The sequence $(f^n(a))_{n\ge 0}$ converges to an infinite word $\mathbf{w}$ denoted by $f^\omega(a)$. This infinite word is a fixed point of f, i.e., $f(\mathbf{w}) = \mathbf{w}$.*

Definition 4. An infinite word $\mathbf{w}$ is *pure morphic*, if there exists a morphism $f : A^* \to A^*$ prolongable on a letter $a \in A$ such that $f^\omega(a) = \mathbf{w}$.

The k-automatic words are obtained by iterating a morphism of constant length k (and an extra coding is allowed). See, for instance, [1, 13].

Definition 5. Let $k \ge 2$ be an integer. An infinite word $\mathbf{w} \in A^{\mathbb{N}}$ is *k-automatic* if there exists an infinite word $\mathbf{v} \in B^{\mathbb{N}}$, a morphism $f : B^* \to B^*$ prolongable on a letter $a \in B$ and a morphism $g : B^* \to A^*$ such that $f^\omega(a) = \mathbf{v}$, $g(\mathbf{v}) = \mathbf{w}$ and $|f(b)| = k$ and $|g(b)| = 1$ for all $b \in B$.

Example 3. The Thue–Morse word $\mathbf{t}$ is the fixed point of the morphism f_2 given in Example 1. A prefix of $\mathbf{t} = f_2^\omega(0)$ is given by

$$01101001100101101001011001101001$$

The Thue–Morse word is an example of a pure morphic word and it is also a 2-automatic word. With the notation of the previous definition, g is the identity map.

A theorem of Cobham links k-automatic words with finite automata and base-k expansions.

Definition 6. A *deterministic finite automaton with output* (DFAO) is a 6-tuple $\mathcal{M} = (Q, q_0, A, \delta, B, \tau)$ where Q is a finite set of states, $q_0 \in Q$ is the initial state, $\delta : Q \times A \to Q$ is the transition function, $\tau : Q \to B$ is the output function and A and B are respectively the input and the output alphabets. As usual δ can be extended to $Q \times A^*$ by $\delta(q, \varepsilon) = q$ and $\delta(q, aw) = \delta(\delta(q, a), w)$ for all $q \in Q$, $a \in A$, $w \in A^*$.

Theorem 1. *[1, 3] Let $k \geq 2$ be an integer. An infinite word*

$$\mathbf{w} = w_0 w_1 w_2 \cdots \in B^\omega$$

is k-automatic if and only if there exists a DFAO

$$\mathcal{M} = (Q, q_0, \{0, \ldots, k-1\}, \delta, B, \tau)$$

such that, for all $n \geq 0$,

$$w_n = \tau(\delta(q_0, \rho_k(n))),$$

where $\rho_k(n)$ denotes the usual base-k representation of n.

The proof of this theorem is constructive. A DFAO can be derived from the morphism and conversely.

Example 4. The nth letter occurring in the Thue–Morse word $\mathbf{t}$ can be obtained by feeding the DFAO depicted in Figure 1 with the base-2 expansion of n. The initial state is represented with an in-going arrow. The output function is written inside the states. The labels of the edges give the transition function. Otherwise stated, if $s_2(n)$ denotes the sum-of-digits of the

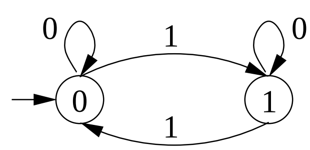

Figure 1: A DFAO over $\{0, 1\}$ generating the Thue–Morse word

base-2 expansion of n, then

$$t_n = s_2(n) \mod 2. \tag{1}$$

The Thue–Morse word is ubiquitous. See, for instance, [2] for a nice survey. It has applications in combinatorics on words: it is an overlap-free word (i.e., it does not contain any *factor* (subword) of the form $auaua$

where a is a letter and u is a word or empty), in number theory: it provides a solution to the Prouhet–Tarry–Escott problem, in symbolic dynamics, in chess, in differential geometry, etc.

One can think about several generalizations of this word. In this paper, motivated by (1), we will consider the following family of infinite words. Let $m \geq 2$ be an integer. The infinite word $\mathbf{t}_m$ over $\{0, \ldots, m-1\}$ is defined by

$$\mathbf{t}_m(n) = s_m(n) \mod m$$

where $s_m(n)$ is the sum-of-digits of the base-m expansion of n. In particular, for $m = 2$ we get the usual Thue–Morse word. The word $\mathbf{t}_3$ starts with

$$012120201120201012201012120 \cdots . \tag{2}$$

For instance, this latter word is called "a generalized Thue–Morse word" in [11] and arithmetic progressions of maximal length occurring in this word are characterized.

Let us briefly describe how the word $\mathbf{t}_m$ can be obtained. It is an easy exercise about automatic words. Let $\Sigma_m := \{0, \ldots, m-1\}$.

Proposition 2. *Let $m \geq 2$. The word $\mathbf{t}_m := (s_m(n) \mod m)_{n \geq 0}$ is the fixed point over $\{0, \ldots, m-1\}$ starting with 0 of the morphism f_m defined by*

$$f_m(i) = i\,(i+1 \mod m) \cdots (i+m-1 \mod m), \quad \forall i \in \Sigma_m.$$

Otherwise stated, $\mathbf{t}_m = f_m^\omega(0)$.

It is an m-automatic word generated by the DFAO $\mathcal{M}_m$ having Σ_m as set of states, 0 as initial state, $\delta(i,j) = i + j \mod m$ as transition function for all $i, j \in \Sigma_m$ and the identity as output function.

Corollary 1. *Let $m \geq 2$. The word $\mathbf{t}_m$ is the concatenation of words of length m containing exactly one occurrence of every letter in $\{0, \ldots, m-1\}$. Each of these factors is a cyclic permutation of $01 \cdots (m-1)$.*

Example 5. Consider the word $\mathbf{t}_3$. It is the fixed point starting with 0 of the morphism f_3 defined by $f_3(0) = 012$, $f_3(1) = 120$ and $f_3(2) = 201$. The nth letter in $\mathbf{t}_3$ (starting with 0 as first index) is obtained by feeding the DFAO depicted in Figure 2 with the base-3 expansion of n.

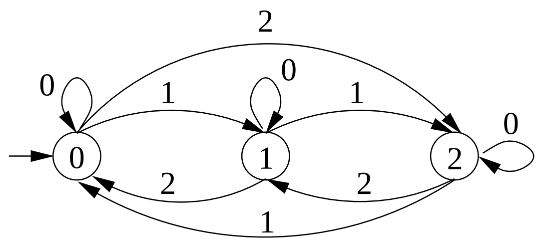

Figure 2: A DFAO over $\{0,1,2\}$ generating $\mathbf{t}_3$.

3 Origin of the problem

The idea about this paper came from the reading of [13] by the first author, and follows his participation in the workshop *Automatic Sequences* held in Liège in May 2015. Only a few recent papers deal with the connection between combinatorial game theory and combinatorics on words [5, 6, 7, 8]. It is desirable to have a game whose set of P-positions is coded by an infinite word with particular properties. One may ask whether a generalized Thue–Morse word corresponds to some 'interesting' game. As an example, the celebrated Wythoff game is *coded* by the Fibonacci word $\mathbf{f}$ as presented below. See, for instance, [4].

Example 6. The Fibonacci word is the fixed point of the morphism $a \mapsto ab$, $b \mapsto a$. It starts with

$$abaababaabaababaababaabaababaabaab\cdots$$

If positions inside $\mathbf{f}$ are counted from 1, then the position of the nth letter a (resp. b) is denoted A_n (resp. B_n), $n \geq 1$. For instance, $A_1 = 1$, $A_2 = 3$, $A_3 = 4$, $B_1 = 2$, $B_2 = 5$, $B_3 = 7$.

In [5] the following characterization of Wythoff's sequence using the Fibonacci word is given.

Proposition 3. *The sequence $(A_n, B_n)_{n\geq 1}$ defined in Example 6 is exactly the Wythoff sequence, i.e., the set of P-position of the Wythoff game.*

Remark 1. When generating k-automatic words, it is more convenient to have indices starting with 0, but when dealing with codings of P-positions, it is sometimes useful to start with index 1 as in the previous example. This does not make any difference about the considered infinite words but we have to take this shift into account.

Given the generalized Thue–Morse word $\mathbf{t}_m$, similarly to Proposition 3, our aim is to define a subtraction game on m piles of tokens such that the set of P-positions of this game is coded by $\mathbf{t}_m$.

4 A triple of subtraction games

- Throughout m denotes the number of piles from which we remove tokens. Throughout we assume $m \geq 2$.
- Throughout x_i denotes the size of pile i. Throughout the piles are unordered. For convenience we list them by nondecreasing size. Specifically, for games 1 and 2, game positions are denoted by $(x_1, \ldots, x_m)$ with $x_1 \leq \cdots \leq x_m$. For game 3, game positions are denoted by $(x_0, \ldots, x_{m-1})$ with $x_0 < \cdots < x_{m-1}$. For the three games, after removing tokens from some piles, the piles are reordered to satisfy these inequalities.
- Throughout we consider two-player games; and *normal* play, that is, the player first unable to move loses.

4.1 Game 1

The following is the rule-set for the game $\mathcal{G}^1_{\mathbf{t}_m}$, played on $m \geq 2$ piles of tokens.

1. The end position is $(1, \ldots, m)$. Thus for every position $(x_1, \ldots, x_m)$ we have $x_i \geq i$, $1 \leq i \leq m$.
2. If all the piles are of distinct size, a player may remove any positive number of tokens from up to $m-1$ piles, subject to rule 1.
3. If two (or more) piles have the same size, a player may remove any positive number of tokens from up to all m piles, subject to rule 1.

Let $\mathcal{P}^1_{\mathbf{t}_m}$ be the set consisting of all m-tuples of adjacent integers whose smallest member is $\equiv 1 \pmod{m}$. Thus

$$\mathcal{P}^1_{\mathbf{t}_m} = \{(1, \ldots, m), (m+1, \ldots, 2m), (2m+1, \ldots, 3m), \ldots\}.$$

More succinctly, $\mathcal{P}^1_{\mathbf{t}_m} = \{km+1, km+2, \ldots, (k+1)m \mid k \geq 0\}$.

Let us explain the connection between the latter set and the infinite word $\mathbf{t}_m$. By Corollary 1, for all $k \geq 0$, the position of the $(k+1)$th occurrence of any symbol i in $\mathbf{t}_m$ belongs to $\{km+1,\dots,(k+1)m\}$ (if positions in $\mathbf{t}_m$ are counted from 1) and the factor $(km+1)\cdots(k+1)m$ of $\mathbf{t}_m$ contains exactly one occurrence of i. As an example, for the word $\mathbf{t}_3$ (see 2), the first few occurrences of 0 are $1,6,8,12,14$; the first occurrences of 1 are $2,4,9,10,15$ and those for 2 are $3,5,7,11,13$. Proceed as in Proposition 3 and consider the sequence of m-tuples made of the $(k+1)$th occurrences of the symbols $0,\dots,m-1$. For $m=3$, this sequence starts with $(1,2,3)$, $(6,4,5)$, $(8,9,7)$, $(12,10,11)$, $(14,15,13)$. Recall that for Game 1, game positions are denoted by $(x_1,\dots,x_m)$ with $x_1\le\cdots\le x_m$, thus we reorder the m-tuples accordingly. For instance, we replace the triple $(6,4,5)$ by $(4,5,6)$. In that respect, we can say that $\mathcal{P}^1_{\mathbf{t}_m}$ is coded (up to reordering) by the generalized Thue-Morse word $\mathbf{t}_m$.

Theorem 2. *The game $\mathcal{G}^1_{\mathbf{t}_m}$ defined above has $\mathcal{P}^1_{\mathbf{t}_m}$ as its set of P-positions.*

Proof. The set $\mathcal{P}^1_{\mathbf{t}_m}$ is absorbing: We prove that for any position $(x_1,\dots,x_m)$ not in $\mathcal{P}^1_{\mathbf{t}_m}$, there is an *option* (direct follower) in $\mathcal{P}^1_{\mathbf{t}_m}$.

We have $x_1=km+j$, $k\ge 0$, $0\le j<m$; k, j not both 0, since $x_1\ge 1$.

Assume first that all the pile sizes x_i are distinct: $1\le x_1<\cdots<x_m$. If $j=0$, reduce the sizes of all x_i, $i>1$, to produce the P-position $((k-1)m+1,(k-1)m+2,\dots,km)$. If $j>0$, we also leave x_1 put, and reduce the sizes of at most $m-1$ piles to fill in the positions $(km+1,km+2,\dots,(k+1)m)$.

Secondly, assume that there are two piles of the same size. If $j=0$, then the above argument holds, and we can again move to $((k-1)m+1,\dots,km)$. So we may assume $j>0$.

If $x_i\ge km+i$ for all $1\le i\le m$, then also the above argument is valid and by reducing at most $m-1$ piles we reach $(km+1,\dots,(k+1)m)$. However, we may have, say, $x_i=km+i=x_{i+1}<km+i+1$, or, say, $x_i=km+i+1=x_{i+1}=x_{i+2}<km+i+2$. In these cases, $(km+1,\dots,(k+1)m)$ cannot be reached. But then we can reach, say, $((k-1)m+1,\dots,km)$, by reducing *all* m piles.

Let $(x_1,\dots,x_m)=(km+1,\dots,(k+1)m)$ be an arbitrary P-position with $k>0$. To move it to another P-position $(\ell m+1,\dots,(\ell+1)m)$ with $\ell<k$, we clearly need to reduce *all* m pile sizes $x_1,\dots,x_m$. But rule 3 is not applicable to P-positions. Thus $\mathcal{P}^1_{\mathbf{t}_m}$ is independent. □

Example 7. For $m=3$, the following are some examples of moves from N- to P-positions, following the order of cases dealt with in the proof.

$(6,8,11)\to(4,5,6)$; $(5,8,11)\to(4,5,6)$. We reduced $m-1=2$ piles in these two examples. $(5,6,6)\to(4,5,6)$. Also here we had to reduce only two piles. Notice that $(5,6,6)$ is not reachable from any P-position. But it can be an initial game-position. $\{(4,4,6),(4,5,5),(5,5,5),(7,8,8)\} \to (1,2,3)$. Here all $m=3$ piles have to be reduced. ($(7,8,8)$ may also be reduced to $(4,5,6)$.)

4.2 Game 2

For the game $\mathcal{G}^2_{\mathbf{t}_m}$, played on $m \geq 2$ piles of tokens, we define the following rule-set:

1. The end position is $(1,\ldots,m)$. Thus for every position $(x_1,\ldots,x_m)$ we have $x_i \geq i$, $1 \leq i \leq m$.

2. If all the piles are of distinct size, a player may remove any positive number of tokens from up to $m-1$ piles, subject to rule 1.

3. If $2 \leq i < m$ piles have the same size, a player may remove any positive number of tokens from up to all m piles, subject to rule 1.

4. From a position $(x,\ldots,x)$, $x > m$, only the following moves are permitted:

   $$(x,\ldots,x) \to (x-(m+i), x-(m+i-1), x-(m+i-2),\ldots,x-(i+1)),$$

   for all $i = 0,\ldots,m-2$.

Notice that the first two rules are identical to those for $\mathcal{G}^1_{\mathbf{t}_m}$; rule 3 and the new rule 4 are *restrictions* of rule 3 of $\mathcal{G}^1_{\mathbf{t}_m}$: if all piles have the same size we cannot anymore remove from all piles *any* number of tokens; we are constrained by the restriction of rule 4.

We let $\mathcal{P}^2_{\mathbf{t}_m} = \mathcal{P}^1_{\mathbf{t}_m}$.

Theorem 3. *The game $\mathcal{G}^2_{\mathbf{t}_m}$ defined above has $\mathcal{P}^2_{\mathbf{t}_m}$ as its set of P-positions.*

Proof. The set $\mathcal{P}^2_{\mathbf{t}_m}$ is absorbing: the first part is proved exactly as in the proof of Theorem 2. It remains only to deal with rule 4. Consider the position $(x,\ldots,x)$, $x > m$. Let $x = km + j$, $k \geq 1$, $0 \leq j \leq m$.

If $j = 0$, one copy of x stays put as the largest pile of the P-position we will move to. The remaining $m-1$ copies are reduced so as to reach the position $((k-1)m+1,(k-1)m+2,\ldots,km)$.

We may thus assume $j \geq 1$. Choose $i = j-1$. Then $0 \leq i \leq m-2$, consistent with rule 4. Moreover, by rule 4,

$$
\begin{aligned}
&(x, \ldots, x) \to \\
&\qquad (km + j - (m + j - 1), km + j - (m + j - 2), \ldots, km + j - j) \\
&\qquad = ((k-1)m + 1, (k-1)m + 2, \ldots, km) \in \mathcal{P}^2_{\mathbf{t}_m}.
\end{aligned}
$$

The fact that the set $\mathcal{P}^2_{\mathbf{t}_m}$ is independent is proved as in the proof of Theorem 2, since rule 4, same as rule 3, does not apply to P-positions. $\square$

Example 8. As for Example 7, we illustrate moves from N- to P-positions.

For $m = 2$, rule 3 does not apply. For rule 4 we have $i = 0$. For $k \geq 1$, $(4k, 4k) \to (4k - 1, 4k)$ $(j = 0)$, but $(3k, 3k) \to (3k - 2, 3k - 1)$ (rule 4).

For $m = 3$, $i \in \{0, 1\}$. $(6, 6, 6) \to (4, 5, 6)$ $(j = 0)$; $(7, 7, 7) \to (4, 5, 6)$ $(j = 1, i = 0)$; $(8, 8, 8) \to (4, 5, 6)$ $(j = 2, i = 1)$.

Remark 2. It is of independent interest to note that $\mathcal{G}^1_{\mathbf{t}_m}$ and $\mathcal{G}^2_{\mathbf{t}_m}$ are two games with distinct game rules, that, nevertheless, share an identical set of P-positions.

4.3 Game 3

Informally, a subtraction game is *invariant* if its rule-set is independent of the game positions; in other words, every move can be made from any position, provided only that the result is nonnegative. Otherwise the game is *variant*. Invariance gives a game some measure of robustness. The notion was introduced in [9] and expanded on in [6].

The two games defined above have some deficiencies:

- Both are variant: moves depend on whether or not there are piles of the same size.
- The end position $(1, \ldots, m)$ is connived: further subtractions could be made. Usually in subtraction games the end position is such that any further subtraction would lead to a negative result or a result that would upset a rule *other* than the end position rule.
- The first element of the generalized Thue-Morse word is tagged by the number 1. In combinatorics on words it is normally tagged by 0 (see, for instance, [1, Lemma 6.3.1]).

In the following game, all three deficiencies disappear.

The following is the rule-set for the game $\mathcal{G}^3_{\mathbf{t}_m}$, played on $m \geq 2$ piles of tokens.

1. The end position is $(0, \ldots, m-1)$. Thus for every position $(x_0, \ldots, x_{m-1})$ we have $x_i \geq i$, $0 \leq i \leq m-1$.

2. Throughout play, the piles have distinct sizes.

3. A player may remove any positive number of tokens from up to $m-1$ piles, subject to rules 1 and 2.

Let $\mathcal{P}^3_{\mathbf{t}_m}$ be the set consisting of all m-tuples of adjacent integers whose smallest member is $\equiv 0 \pmod{m}$. Thus

$$\mathcal{P}^3_{\mathbf{t}_m} = \{(0, \ldots, m-1), (m, \ldots, 2m-1), (2m, \ldots, 3m-1), \ldots\}.$$

More succinctly, $\mathcal{P}^3_{\mathbf{t}_m} = \{km, km+1, \ldots, (k+1)m-1 \mid k \geq 0\}$. The set $\mathcal{P}^3_{\mathbf{t}_m}$ is but a *left shift* of $\mathcal{P}^1_{\mathbf{t}_m} = \mathcal{P}^2_{\mathbf{t}_m}$. Again, by Corollary 1, for all $k \geq 0$, the position of the $(k+1)$th occurrence of any symbol i in $\mathbf{t}_m$ belongs to $\{km, \ldots, (k+1)m-1\}$ (if positions in $\mathbf{t}_m$ are counted from 0) and the factor $km \cdots (k+1)m-1$ of $\mathbf{t}_m$ contains exactly one occurrence of i. Recall that for Game 3, game positions are denoted by $(x_0, \ldots, x_{m-1})$ with $x_0 < \cdots < x_{m-1}$, thus we reorder the m-tuples accordingly. In that respect, we can say that $\mathcal{P}^3_{\mathbf{t}_m}$ is coded (up to reordering) by the generalized Thue-Morse word $\mathbf{t}_m$.

Remark 3. (i) The game is invariant. (ii) The end position is natural, since removing any token from it would result in a position violating move rule 2. (iii) The P-positions now correspond to a word whose initial element is tagged by 0. (iv) The rule-set for game 3 is simpler than for games 1 and 2. One may argue whether this is an advantage or disadvantage.

Theorem 4. *The game $\mathcal{G}^3_{\mathbf{t}_m}$ defined above has $\mathcal{P}^3_{\mathbf{t}_m}$ as its set of P-positions.*

Proof. The set $\mathcal{P}^3_{\mathbf{t}_m}$ is absorbing: Let $(x_0, \ldots, x_{m-1}) \notin \mathcal{P}^3_{\mathbf{t}_m}$, $x_0 = km + j$, $k \geq 1$, $0 \leq j < m$. Leaving x_0 put, we reduce at most the remaining $m-1$ piles so as to reach $(km, km+1, \ldots, (k+1)m-1)$.

Let $(km, \ldots, (k+1)m-1) \in \mathcal{P}^3_{\mathbf{t}_m}$ be an arbitrary P-position, $k \geq 1$. To reduce it to a P-position $(\ell m, \ldots, (\ell+1)m-1)$, we clearly have $\ell < k$. Since $x_i \geq km$ for all i, all m piles have to be reduced, violating rule 3. Hence $\mathcal{P}^3_{\mathbf{t}_m}$ is independent. □

5 Conclusion

We gave some background on combinatorics on words, and considered a generalized Thue-Morse word. We then devised and analysed three games related to this word.

Suggested further work:

- Compute the Sprague-Grundy function of the games.
- Formulate a winning strategy for the games in misère play (first player unable to move wins).
- Find rule-sets for more games derived from the generalized Thue-Morse word.
- Consider the case of *ordered* piles: Find a subtraction game with a 'reasonable' rule-set (i.e., with no direct relation to $\mathbf{t}_m$) in such a way that its set of P-positions is exactly given by the m-tuples corresponding to the kth occurrences of the m symbols. As an example, for $m = 3$, devise a game whose first few P-positions are exactly $(1, 2, 3)$, $(6, 4, 5)$, $(8, 9, 7)$, $(12, 10, 11)$, $(14, 15, 13)$. So (6,4,5) means 6 tokens in the first pile, 4 in the second and 5 in the third, and we would not be allowed to reorder the piles.

References

[1] J.-P. Allouche, J. O. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, Cambridge, 2003.

[2] J.-P. Allouche, J. Shallit, The ubiquitous Prouhet-Thue-Morse sequence, in *Sequences and their applications* (Singapore, 1998), 116, Springer Ser. Discrete Math. Theor. Comput. Sci., Springer, London, 1999.

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

[4] E. Duchêne, A. S. Fraenkel, R. J. Nowakowski, M. Rigo, Extensions and restrictions of Wythoff's game preserving its P positions, *J. Combin. Theory Ser. A* **117** (2010), 545–567.

[5] E. Duchêne, M. Rigo, A morphic approach to combinatorial games: the Tribonacci case, *Theor. Inform. Appl.* **42** (2008), 375–393.

[6] E. Duchêne, M. Rigo, Invariant games, *Theoret. Comput. Sci.* **411** (2010), 31693180.

[7] E. Duchêne, M. Rigo, Pisot unit combinatorial games, *Monatsh. Math.* **155** (2008), no. 3-4, 217–249.

[8] A. S. Fraenkel, Complementary iterated floor words and the Flora game, *SIAM J. Discrete Math.* **24** (2010), no. 2, 570–588.

[9] S. W. Golomb, A mathematical investigation of games of "take-away". *J. Comb. Theory* **1** (1966) 443-458. 90.70

[10] M. Lothaire, *Combinatorics on words*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, (1997).

[11] O. G. Parshina, On Arithmetic Progressions in Generalized Thue-Morse Word, in *WORDS 2015*, Kiel, *Lect. Notes in Comp. Sci.* **9304** (2015), 191–196.

[12] M. Rigo, *Formal languages, automata and numeration systems, vol. 1. Introduction to combinatorics on words*, ISTE, London; John Wiley & Sons, Inc., Hoboken, NJ, 2014.

[13] E. Rowland, What is . . . an automatic sequence? *Notices Amer. Math. Soc.* **62** (2015), 274-276.