



On languages of one-dimensional overlapping tiles

David Janin

► To cite this version:

David Janin. On languages of one-dimensional overlapping tiles. SOFSEM, Jan 2013, Špindlerův Mlýn, Czech Republic. pp.244-256, 10.1007/978-3-642-35843-2_22 . hal-00659202v4

HAL Id: hal-00659202

<https://hal.science/hal-00659202v4>

Submitted on 11 Jul 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



LaBRI, CNRS UMR 5800
Laboratoire Bordelais de Recherche en Informatique

Rapport de recherche RR-1457-12

On languages of one-dimensional overlapping tiles

July 11, 2012

David Janin,
LaBRI, IPB, Université de Bordeaux

Abstract

A one-dimensional tile with overlaps is a standard finite word that carries some more information that is used to say when the concatenation of two tiles is legal. Known since the mid 70's in the rich mathematical field of inverse monoid theory, this model of tiles with the associated partial product have yet not been much studied in theoretical computer science despite some implicit appearances in studies of two-way automata in the 80's.

In this paper, after giving an explicit description of McAlister monoid, we define and study several classical classes of languages of tiles: from recognizable languages (REC) definable by morphism into finite monoids up to languages definable in monadic second order logic (MSO).

We show that the class of MSO definable languages of tiles is both *simple*: these languages are finite sums of Cartesian products of rational languages, and *robust*: the class is closed under product, iterated product (star), inverse and projection on context tiles. A equivalent notion of regular expression is deduced from these results.

The much smaller class of REC recognizable languages of tiles is then studied. We describe few examples and we prove that these languages are tightly linked with covers of periodic bi-infinite words.

1 Introduction

1.1 Background

In this paper, we study languages of one-dimensional discrete overlapping tiles. These tiles already appear in the 70's as elements of McAlister's monoid in the rich mathematical field of inverse monoid theory (see 9.4. in [7]). In particular, though sometimes implicitly, they are used for studying the structure of (zigzag) covers of finite, infinite or bi-infinite words [1]. McAlister's monoid also appears in studies of the structure of tiling (in the usual sense with no overlaps) of the d -dimensional Euclidian space \mathbb{R}^d [6, 1].

In a field more closely related with computer science, overlapping tiles also appear decades ago in study of two-way automata on words (they are called word bisections in [11]). But there the underlying monoid structure is left at most implicit. Only recently, it has been shown that one can relevantly defined two-way word automata semantics by mapping partial runs to elements of McAlister's Monoid [3].

Oddly enough to be mentioned, our interest in studying languages of positive tiles came from application perspectives in computational music theory [5]. In particular,

tiles and their continuous variants, can be used to describe advanced synchronization mechanisms between musical patterns [2].

Our purpose here is to provide the computer science flavored basis for the systematic study of languages of tiles. In this paper, we especially focus our attention on characterizing two classical classes of languages that are defined on tiles. The class REC and the class MSO . It must be mentioned that the study of other intermediate classes of languages such as the class of languages definable by means of Kleene's regular expressions, in connection with two-way automata, can be found in a companion paper [3].

1.2 Outline

We first define the monoid of overlapping tiles and prove it isomorphic to the inverse monoid of McAlister. Three classes of languages of tiles are then considered:

- the class REC of languages definable as pre-images of monoid morphisms into finite monoids,
- the class REG_C of languages definable by means of finite Kleene's regular expressions extended with some inverse and context projection operators,
- the class MSO of languages of tiles definable by means of formula of monadic second order logic.

The largest class MSO is shown to be both *simple*: these languages are finite sums of Cartesian products of rational languages of words (Theorem 4), and *robust*: this class is closed under sum, product, star, inverse and context projection operators (Theorem 5). As a consequence of robustness, we show that $REG_C \subseteq MSO$ (corollary 6). As a consequence of simplicity, we show that $MSO \subseteq REG_C$ (corollary 7).

The class REC is studied in the last part. Compared to the class MSO , it really collapses. A simple example shows that it is strictly included into MSO (Lemma 8). Still a non trivial example of an onto morphism from the monoid of tiles to a finite monoid is provided (Theorem 11). It illustrates the fact, generalized in Theorem 12, that recognizable languages of tiles are necessarily sort of languages of covers of finitely many bi-infinite words.

1.3 Preliminaries

The free monoid. Given a finite alphabet A , A^* denotes the free monoid generated by A , 1 denotes the neutral element. The concatenation of two words u and v is denoted by uv .

Prefix and suffix lattices. \leq_p stands for the prefix order over A^* , \leq_s for the suffix order. \vee_p (resp. \vee_s) denotes the joint operator for the prefix (resp. suffix) order: thus for all words u and v , $u \vee_p v$ (resp. $u \vee_s v$) is the least word whose both u and v are prefixes (resp. suffixes).

The extended monoid $A^* + \{0\}$ (with $0u = u0 = 0$ for every word u), ordered by \leq_p (extended with $u \leq_p 0$ for every word u), is a lattice; in particular, $u \vee_p v = 0$ whenever neither u is a prefix of v , nor v is a prefix of u . Symmetric properties hold in the suffix lattice.

Syntactic inverses. Given \bar{A} a disjoint copy of A , $u \mapsto \bar{u}$ denotes the mapping from $(A + \bar{A})^*$ to itself inductively defined by $\bar{1} = 1$, for every letter $a \in A$, \bar{a} is the copy of a in \bar{A} and $\overline{\bar{a}} = a$, for every word $u \in (A + \bar{A})^*$, $\overline{\bar{u}} = u$. The mapping $u \mapsto \bar{u}$ is involutive ($\overline{\bar{u}} = u$ for every word u); it is an antimorphism of the free monoid $(A + \bar{A})^*$, i.e. for all words u and $v \in (A + \bar{A})^*$, $\overline{uv} = \bar{v}.\bar{u}$.

Free group. The free group $FG(A)$ generated by A is the quotient of $(A + \bar{A})^*$ by the least congruence \simeq such that, for every letter $a \in A$, $a\bar{a} \simeq 1$ and $\bar{a}a \simeq 1$.

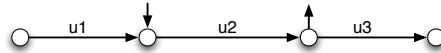
2 The monoid of overlapping tiles

We give in this section a description of the monoid of overlapping tiles. It is shown to be isomorphic to McAlister's monoid [8].

2.1 Positive, negative and context tiles

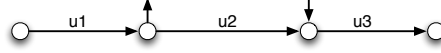
A *tile* over the alphabet A is a triple of words $u = (u_1, u_2, u_3) \in A^* \times (A^* + \bar{A}^*) \times A^*$ such that, if $u_2 \in \bar{A}^*$, its inverse \bar{u}_2 is a suffix of u_1 and a prefix of u_3 . When $u_2 \in A^*$ we say that u is a *positive tile*. When $u_2 \in \bar{A}^*$ we say that u is a *negative tile*. When $u_2 = 1$, i.e. when u is both positive and negative, we say that u is a *context tile*.

A positive tile $u = (u_1, u_2, u_3)$ is conveniently drawn as a (linear, unidirectional and left to right) Munn's birooted word tree [10]:

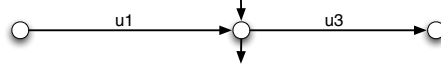


where the dangling input arrow (marking the beginning of the root) appears on the left of the dangling output arrow (marking the end of the root). A negative tile of

the form $u = (u_1 u_2, \overline{u_2}, u_2 u_3) \in A^* \times \overline{A}^* \times A^*$ is also drawn as a birooted word tree



where the dangling input arrow appears on the right of the dangling output arrow. A context tile of the form $u = (u_1, 1, u_3) \in A^* \times 1 \times A^*$ is then drawn as follows:

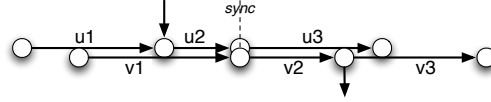


The *domain* of a tile $u = (u_1, u_2, u_3)$ is the reduced form of $u_1 u_2 u_3$ (always a word of A^*). Its *root* is the word u_2 .

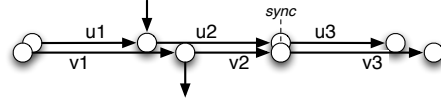
Sets T_A , T_A^+ , T_A^- and C_A will respectively denote the set of tiles, the set of positive tiles, the set of negative tiles and the set of context tiles over A .

2.2 A product of tiles

Intuitively, the sequential product of two tiles is their superposition in such a way that the end of the root of the first tile coincides with the beginning of the root of the second tile; the superposition requires pattern-matching conditions to the left and to the right of the synchronization point. When both tiles are positive, this can be drawn as follows:



The product can be extended to arbitrary tiles, as illustrated by the following figure (positive u and negative v):



Formally, we extend the set T_A with a zero tile to obtain $T_A^0 = T_A + \{0\}$. The sequential product of two non-zero tiles $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ is defined as

$$u.v = ((u_1 u_2 \vee_s v_1) \overline{u_2}, u_2 v_2, \overline{v_2} (u_3 \vee_p v_2 v_3))$$

when both $u_1 u_2 \vee_s v_1 \neq 0$ and $u_3 \vee_p v_2 v_3 \neq 0$, and $u.v = 0$ otherwise, where, in this formula, words in $(A + \overline{A})^*$ are implicitly considered as reduced elements of $FG(A)$. We also let $u.0 = 0.u = 0$ for every $u \in T_A^0$.

Remark. Let a, b, c and $d \in A$ be distinct letters. Then $(a, b, c).(b, c, d) = (a, bc, d)$ whereas $(a, b, c).(a, c, d) = 0$. In the latter case, the left matching constraint is

violated because $a \neq b$.

Set T_A^0 equipped with the above sequential product is a monoid. But the proof of that fact is postponed to the proof that it is even isomorphic to the monoid of McAlister.

2.3 The inverse monoid of McAlister

McAlister's monoid is defined as the quotient of the free inverse monoid $FIM(A)$ by the ideal \perp of non unidirectional and non linear tiles.

More precisely, following Munn's result [10], elements of $FIM(A)$ are seen as birooted word trees, i.e. pairs (P, u) where P is a non empty finite and prefix-closed subset of (reduced elements of) the free group $FG(A)$ generated by A , with $u \in P$. The product of two birooted trees (P, u) and (Q, v) is defined by $(P, u).(Q, v) = (P \cup uQ, uv)$.

A birooted word tree (P, u) is said *unidirectional* when $P \subseteq A^* + \overline{A}^*$, and *linear* when both $P \cap A^*$ and $P \cap \overline{A}^*$ are totally ordered by the prefix order. It is straightforward that the set \perp of non-unidirectional or non-linear birooted trees is an ideal.

The monoid of McAlister is then defined as the Rees quotient $FIM(A)/\perp$, i.e. the monoid obtained from $FIM(A)$ by merging into a single zero all elements of \perp . In that monoid, given two linear and unidirectional birooted word trees (P, u) and (Q, v) , the product of these two tiles is defined to be $(P \cup uQ, uv)$ as in $FIM(A)$ when the resulting birooted tree is linear and unidirectional, and 0 otherwise.

Theorem 1 *The set T_A^0 equipped with the sequential product is a monoid isomorphic to the monoid of McAlister.*

Proof. (sketch of) For all non zero tile $u = (u_1, u_2, u_3) \in T_A$ let $t_u = (P_u, u_2)$ be the resulting birooted tree defined by $P_u = \{x \in \overline{A}^* : x \leq_p \overline{u}_1\} \cup \{x \in A^* : x \leq_p u_2u_3\}$.

We observe that t_u is a well-defined unidirectional and linear birooted tree. Indeed, when u is a positive tile, we have $u_2 \leq u_2u_3$ hence $u_2 \in P_u$. When u is a negative tile, i.e. with $u_2 \in \overline{A}^*$ we have both u_1u_2 and $u_2u_3 \in A^*$ and thus $u_2 \leq_p \overline{u}_1$ hence $u_2 \in P_u$.

We conclude then by showing that the mapping $\varphi : T_A^0 \rightarrow FIM(A)/\perp$ defined by $\varphi(0) = 0$ and for any non-zero tile $u = (u_1, u_2, u_3) \in T_A$, $\varphi(u) = (P_u, u_2)$ is an isomorphism.

First, it is easy to check that it is a bijection. In fact, given a linear and unidirectional tile (P, u) one check that $\varphi^{-1}((P, u))$ is the tile (u_1, u_2, u_3) defined by $u_2 = u$, $u_1 = \bigvee_s \overline{P \cap \overline{A}^*}$ and $u_3 = \overline{u}_2 \bigvee_p P \cap A^*$.

It remains to show φ preserves products. For this, it is enough to check that for any two non zero tiles $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3) \in T_A$, one indeed has $t_{u.v} = (P_u \cup u_2 P_v, u_2 v_2)$ which essentially follows from definitions. \square

Corollary 2 *Monoid T_A^0 is an inverse monoid with (pseudo) inverses given by $0^{-1} = 0$ and for every non zero tile $u = (u_1, u_2, u_3) \in T_A$, $u^{-1} = (u_1 u_2, \bar{u}_2, u_2 u_3)$.*

Proof. (sketch of) For every $u = (u_1, u_2, u_3) \in T_A$, given $t_u = (P_u, u_2)$, this amounts to check that $t_{u^{-1}} = (t_u)^{-1} = (u_2 P_u, \bar{u}_2)$ and thus this just amounts to check that $u_2 P_u = P_{u^{-1}}$. \square

An immediate property worth being mentioned:

Lemma 3 *The mapping $u \mapsto (1, u, 1)$ from A^* to T_A is a one-to-one morphism and monoid T_A^0 is finitely generated from image of letters of A and their inverses.*

In other words, the free monoid A^* can be seen as a submonoid of T_A^0 . In the remainder of the text we may use the same notation for words of A^* and their images in T_A^0 .

One can also observe that positive (resp. negative) tiles extended with 0 form a submonoid T_A^{+0} (resp. T_A^{-0}) of T_A^0 . However, this submonoid *is not* finitely generated. One need extra operators called canonical left and right context. For every $u = (u_1, u_2, u_3) \in T^+$, let $u_L = u^{-1}u = (u_1 u_2, 1, u_3)$, the canonical *left context tile* associated to tile u and let $u_R = uu^{-1} = (u_1, 1, u_2 u_3)$, the canonical *right context tile* associated to tile u . By construction $u_R u = u u_L = u$.

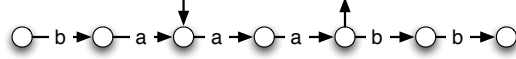
3 MSO-definable and regular languages of tiles

We consider in this section the class MSO of languages of tiles definable by means of monadic second order formulae.

3.1 MSO definability

We need FO-models for positive tiles. For this, we use a typical encoding of words into FO-structures that amounts to encode each letter $a \in A$ as a relation between elements of the domain. This way, there is no need of end markers and the empty word is simply modeled by the structure with singleton domain and empty relations. Then we raise models of words to models of tiles just by marking (as pictured in birooted trees) entry and exit points.

For instance, the triple $u = (ba, aa, bb)$ is modeled as indicated by the following picture



where, as before, a dangling input arrow marks the entry point and a dangling output arrow marks the exit point.

The model of a tile u is denoted by t_u . The associated domain of its underlying FO-structure is written $dom(t_u)$, the entry point is written $in(t_u)$ and the exit point is written $out(t_u)$.

A language $L \subseteq T_A$ is MSO definable when there is a MSO formula of the form $\varphi_L(U, x, y)$ where U is a set variable and x and y are two FO-variables such that, for all $t \in T_A$, $t \in L$ if and only if $t \models \varphi_L(dom(t), in(t), out(t))$.

3.2 A word congruence for languages of tiles

We aim at achieving a simple characterization of MSO definable language of tiles. For this purpose, we first define a notion of congruence relation over A^* which is defined for all language of tiles. It occurs that this congruence is of finite index if and only if the language of tile is definable in MSO.

Given a language $L \subseteq T_A$ on non zero tiles, we define the *word* congruence \simeq_L associated to L that tells when two words can be replaced one by the other in any tile without altering membership in L .

Formally, \simeq_L is the least relation over words such that, for all u_0 and $v_0 \in A^*$, $u_0 \simeq_L v_0$ when for all w_0, w_2, w_3 and $w_4 \in A^*$, if $u = (w_1 u_0 w_2, w_3, w_4)$ and $v = (w_1 v_0 w_2, w_3, w_4)$, or if $u = (w_1, w_2 u_0 w_3, w_4)$ and $v = (w_1, w_2 v_0 w_3, w_4)$, or if $u = (w_1, w_2, w_3 u_0 w_4)$ and $v = (w_1, w_2, w_3 v_0 w_4)$ then $u \in L \Leftrightarrow v \in L$ and $u^{-1} \in L \Leftrightarrow v^{-1} \in L$.

Relation \simeq_L is a congruence on words. For every $u \in A^*$, let $[u]_L$ be the equivalence class of u defined by $[u]_L = \{v \in A^* : u \simeq_L v\}$.

Theorem 4 *For every language $L \subseteq T_A$ of non zero tiles:*

$$L = \sum_{(u_1, u_2, u_3) \in L \cap T_A^+} [u_1]_L \times [u_2]_L \times [u_3]_L + \sum_{(u_1, u_2, u_3)^{-1} \in L \cap T_A^-} ([u_1]_L \times [u_2]_L \times [u_3]_L)^{-1}$$

Moreover, L is definable in MSO if and only if relation \simeq_L is of finite index.

Proof. The first statement is an immediate consequence of the definition of \simeq_L . Moreover, if \simeq_L is of finite index, then this sum is finite, and any language of the

form $[w]_L \subseteq A^*$ with $w \in A^*$ is rational henceforth MSO definable. It follows that L is thus definable in MSO.

Conversely, assume L is definable in MSO. Given $L^+ = L \cap T_A^+$ and $L^- = L \cap T_A^-$, we observe that both L^+ and $(L^-)^{-1} \subseteq A^* \times A^* \times A^*$ can be encoded into languages of words M^+ and $M^- \subseteq A_L^* A_C^* A_R^*$ where A_L , A_C and A_R are three disjoint copies of the alphabet A for encoding the left, center and right elements of each tiles.

Now, since L is definable in MSO, so are L^+ and L^- and thus, their encoding M^+ and M^- are also definable in MSO. By Büchi theorem, this means that they are regular and thus their syntactic congruences \simeq_{M^+} and \simeq_{M^-} are of finite index. This implies that \simeq_L is also of finite index. Indeed, for all word u and $v \in A^*$, we have $u \simeq_L v$ if and only if $u_X \simeq_{M^+} v_X$ and $u_X \simeq_{M^-} v_X$ for X being L , C or R and with w_X denoting the renaming of any word $w \in A^*$ in the copy alphabet A_X . \square

3.3 Extended regular languages

For every language L and $M \subseteq T_A$, let $L + M = L \cup M$, $L.M = \{u.v \in T_A : u \in L, v \in M, u.v \neq 0\}$, $L^* = \bigcup_{k \geq 0} L^k$, $L^{-1} = \{u^{-1} \in T_A : u \in L\}$ and $L^C = \{u \in T_A : u \in L \cap C_A\}$ (called context projection).

Theorem 5 *For every languages L and $M \subseteq T_A$ on non zero tiles, if L and M are MSO definable then so are $L + M$, $L.M$, L^* , L^{-1} and L^C .*

Proof. Let $\varphi_L(U, x, y)$ and $\varphi_M(U, x, y)$ be two formulae defining respectively the language of tiles L and M . We assume that these formulae also check that both x and y belongs to U that, moreover, U is connected.

Case of $L + M$: take $\psi(U, x, y)$ defined by $\varphi_L(U, x, y) \vee \varphi_M(U, x, y)$.

Case of $L.M$: take $\psi(U, x, y)$ stating that there exist two sets X and Y such that $U = X \cup Y$ and there is z such that both $\varphi_L(X, x, z)$ and $\varphi_M(Y, z, y)$ hold.

Case of L^* : in order to define $\varphi(U, x, y)$, the main idea is to consider the reflexive transitive closure $R^+(x, y)$ of the binary relation $R(x_1, x_2)$ defined by $\exists X \varphi_L(X, x_1, x_2)$; one must take care, however, that set U is completely covered by (sub)tiles' domains; this is equivalent to the fact, as domains necessarily overlap, that each extremity (left most or right most element) of the domain U belongs to one of these sets X at least. This is easily encoded by a disjunction of the three possible cases: extremities are reached in a single intermediate tile, left extremity is reached first or right extremity is reached first.

Case of L^{-1} : take $\psi(U, x, y)$ defined by $\varphi_L(U, y, x)$.

Case of L^C : take $\psi(U, x, y)$ defined by $(x = y) \wedge \varphi_L(U, x, y)$. \square

Let then REG_C be the class of extended regular languages of tiles that is the class of languages definable by means of finite sets of non zero tiles, sum, product, iterated product (or star), inverse and context projection.

Since finite tile languages are indeed definable in MSO we can state the following corollary of Theorem 5:

Corollary 6 *Every extended regular languages of tiles is MSO definable.*

Moreover, since for every regular languages L, M and $R \subseteq A^*$, one has $L \times C \times R = (L^{-1}.L)^C.M.(R.R^{-1})^C$ where, on the right side, words are embedded into tiles as in Lemma 3, we can also state immediate corollary of Theorem 4:

Corollary 7 *Every MSO definable languages of tiles is extended regular.*

In other words, $REG_C = MSO$.

Observe that, in the above proof, we use the fact that for every languages $L \subseteq A^*$ of (embeddings of) words into tiles, $(L.L^{-1})^C = \{uu^{-1} : u \in L\}$ and $(L^{-1}.L)^C = \{u^{-1}u : u \in L\}$. This property is not true for arbitrary languages of T_A .

4 Recognizable languages of tiles

In this section, we consider languages of tiles that are recognizable in the algebraic sense. Although the theory of tiles can be seen as part of the theory of inverse monoid, the results we obtain rather differ from the former studies of languages of words recognized by finite inverse monoids [9] or free inverse monoid languages [12]. Morphisms from T_A^0 (or even T_A^{+0}) to arbitrary monoids turns out to be even more constraints than morphisms from A^* to finite inverse monoids.

4.1 A non recognizable language of tiles

The next result, negative, tells us that rather simple (MSO definable) languages of tiles are not recognizable.

Lemma 8 *Language $L = \{(ba^m, a^n, 1) \in T_A^0 : m, n \in \mathbb{N}\}$ with a and b two distinct letters, is not recognizable.*

Proof. We prove that the syntactical congruence \simeq_L (and not the word congruence defined above) associated to L in T_A^0 is of infinite index.

For all $m \in \mathbb{N}$, let u_m be the tile $u_m = (ba^m, 1, 1)$. It is an easy observation that for all m and $n \in \mathbb{N}$, $u_m \simeq_L u_n$ if and only if $m = n$ hence the claim. Indeed, for all $k \in \mathbb{N}$ let $v_k = (a^k, 1, 1)$. We have for any given $m \in \mathbb{N}$, $u_m v_k \in L$ if and only if $k \leq m$. Now if for some m and $n \in \mathbb{N}$ one has $u_m \simeq_L u_n$ then for all $k \in \mathbb{N}$, $u_m v_k \in L$ if and only if $u_n v_k \in L$. It follows that, for all k , $k \leq m$ if and only if $k \leq n$, hence $m = n$. \square

Since language L defined above is obviously MSO, we have:

Corollary 9 $REC \neq MSO$

4.2 A (non-trivial) recognizable language

Before studying in the next section recognizable languages in full generality, we provide in this section a non trivial example of such a language. It illustrates the main characteristic of all recognizable languages of tiles : a strong link with tiles's cover of periodic bi-infinite words.

Building such an example essentially amounts to provide a (onto) monoid morphism from T_A^0 onto some non-trivial finite (inverse) monoid M . Here, the main idea is to type tiles, by means of a monoid morphism, according to their capacity to cover the bi-infinite word ${}^\omega(ab)(ab)^\omega$ with a and b two distinct letters.

In order to do so, let $M = \{0, 1, (a, 1, b), (b, 1, a), (b, a, b), (a, b, a)\}$ with product \odot defined as expected for 0 and 1 and defined according to the following product table:

\odot	$(a, 1, b)$	$(b, 1, a)$	(b, a, b)	(a, b, a)
$(a, 1, b)$	$(a, 1, b)$	0	0	(a, b, a)
$(b, 1, a)$	0	$(b, 1, a)$	(b, a, b)	0
(b, a, b)	(b, a, b)	0	0	$(b, 1, a)$
(a, b, a)	0	(a, b, a)	$(a, 1, b)$	0

Lemma 10 *Monoid (M, \odot) is an inverse monoid.*

Proof. We easily check that product \odot is associative hence M is a monoid. Given $E(M) = \{0, 1, (a, 1, b), (b, 1, a)\}$ the set of idempotents of S , the commutation of idempotents immediately follows from unique non trivial case $(a, 1, b) \odot (b, 1, a) = (b, 1, a) \odot (a, 1, b) = 0$. Last, we check that $(a, b, a) \odot (b, a, b) \odot (a, b, a) = (a, b, a)$ and $(b, a, b) \odot (a, b, a) \odot (b, a, b) = (b, a, b)$. It follows that $(a, b, a)^{-1} = (b, a, b)$ and $(b, a, b)^{-1} = (a, b, a)$. All other element is idempotent and thus self-inverse. \square

The expected monoid morphism $\varphi : T_A^0 \rightarrow M$ is then defined by $\varphi(0) = 0$, $\varphi(1) = 1$ and for all $(u, v, w) \in T_A^0$ such that $uvw \neq 1$, $\varphi(u, v, w) = 0$ when uvw is not a factor of $(ab)^\omega$ and, otherwise, when u is a positive tile:

1. $\varphi(u, v, w) = (a, 1, b)$ when $|v|$ is even with $a \leq_s u, b \leq_p v, a \leq_s v$ or $b \leq_p w$,
2. $\varphi(u, v, w) = (b, 1, a)$ when $|v|$ is even with $b \leq_s u, a \leq_p v, b \leq_s v$ or $a \leq_p w$,
3. $\varphi(u, v, w) = (b, a, b)$ when $|v|$ is odd with $a \leq_p v$,
4. $\varphi(u, v, w) = (a, b, a)$ when $|v|$ is odd with $b \leq_p v$,

and $\varphi(u, v, w) = (\varphi(uv, \bar{v}, vw))^{-1}$ when (u, v, w) is a negative tile.

Theorem 11 *The mapping $\varphi : T_A^0 \rightarrow M$ is a onto morphism.*

Proof. This follows from the fact that, for all u and $v \in T_A^0$, $\varphi(u) \odot \varphi(v) = \varphi(uv) = \varphi(\varphi(u)\varphi(v))$. \square

Given $L_S = (ab)^* + b(ab)^*$, given $L_C = (ab)^*$, given $L_P = (ab)^* + (ab)^*a$, this theorem says, in particular, that the non trivial tile language $L_S \times L_C \times L_P - 1$ is recognizable since it equals $\varphi^{-1}((b, 1, a))$.

4.3 More on recognizable languages of tiles

We conclude our study by showing that recognizable languages of tiles are essentially generalization of the example described above: languages of REC are essentially definable out of finitely many periodic bi-infinite words.

Let $\varphi : T_A^0 \rightarrow M$ be a monoid morphism with finite M . Since we can always restrict M to $\varphi(T_A^0)$ and $\varphi(0)$ is a zero in the submonoid $\varphi(T_A^0)$, we assume, without loss of generality, that $M = \varphi(T_A^0)$ with $\varphi(0) = 0$. Now, by complement, understanding the structure of languages of tiles recognizable by φ amounts to understand the structure of languages of the form $\varphi^{-1}(s)$ for all non-zero element $s \in S$.

Moreover, we can restrict our attention to recognizable languages of positive tiles. It can indeed be shown that $\varphi(T_A^0)$ is necessarily an inverse monoid and, for every $s \in M - 0$, $\varphi^{-1}(s) = \varphi^{-1}(s) \cap T_A^+ + (\varphi^{-1}(s^{-1}) \cap T_A^+)^{-1}$.

In the theorem below, we even consider the more difficult case of a onto monoid morphism $\varphi : T_A^{+0} \rightarrow M$ from the submonoid of positive tiles to a finite monoid M . However, in order to avoid undesirable effect due to the fact that, contrary to T_A^0 , T_A^{+0} is not finitely generated, we assume that φ is *safe* in the sense, for every $u \in T_A^+$, if $\varphi(u) \neq 0$ then both $\varphi_L(u) = \varphi(u^{-1}u) \neq 0$ and $\varphi_R(u) = \varphi(uu^{-1}) \neq 0$. One can check that an onto morphism from all tiles T_A^0 to a monoid M is always safe.

Theorem 12 Let $\varphi : T_A^{+0} \rightarrow M$ be a safe monoid morphism. Let $s \in M$ be a non-zero element. Let $L_s = \varphi^{-1}(s)$. There is $x \in A^*$, $y \in A^+$ and $k \geq 0$ such that, either L_s is finite with $L_s \subseteq \{u \in T_A^+ : v \leq u\}$ for some $v \in T_A^+$ or L_s is a co-finite subset of one of the following set of non-zero tiles:

1. $S({}^\omega(xy)) \times x \times P(w)$ for some finite $w <_p (yx)^\omega$,
2. $S(w) \times x \times P((yx)^\omega)$ for some finite $w <_s {}^\omega(xy)$,
3. $S({}^\omega(xy)) \times x \times P((xy)^\omega)$,
4. or $S({}^\omega(xy)) \times x(yx)^k(yx)^* \times P((yx)^\omega)$,

with, for all $w \in A^* + {}^\omega A$, $S(w) = \{z \in A^* : z <_s w\}$, i.e. the set of strict suffix of w , and for all $w \in A^* + A^\omega$, $P(w) = \{z \in A^* : z <_p w\}$, i.e. the set of strict prefix of w .

In all cases, tiles of L_s are compatible with the bi-infinite periodic word ${}^\omega(xy)x(yx)^\omega$.

Proof. The rest of this section is dedicated to the proof of this theorem. In order to do so, we first prove several closure properties of L_s .

Lemma 13 For all $u = (u_1, u_2, u_3) \in L_s$ and $v = (v_1, v_2, v_3) \in L_s$, $(u_1 \vee_s v_1, u_2, u_3) \in L_s$ and $(u_1, u_2, u_3 \vee_p v_3) \in L_s$.

Proof. We know that $(v_1)_L v = v$ hence $(v_1)u \in L_s$ since $\varphi((v_1)_L v) = \varphi_L(v_1)\varphi(v) = \varphi_L(v_1)\varphi(u) = \varphi((v_1)_L u)$. Moreover, $0 \notin L_s$ hence $(v_1)_L u = (u_1 \vee_s v_1, u_2, u_3)$ with $u_1 \vee_s v_1 \neq 0$. Symmetrical arguments prove the prefix case. \square

Lemma 14 For all $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3) \in L_s$ such that $|u_2| \leq |v_2|$, either $u_2 = v_2$ or there exists $x \in A^*$, $y \in A^+$ and $k \geq 0$ such that $u_2 = x(yx)^k$, $v_2 = x(yx)^{k+1}$. In that latter case, there is $u' \in L_s$ such that $u'((xy)_C)^* \subseteq L_s$.

Proof. Let u and v as above. We have $(v_2)_R v (v_2)_L = v$ hence, by a similar argument as in Lemma 13, $(v_2)_R u (v_2)_L \in L_s$. Since $|u_2| \leq |v_2|$ this means that $u_2 \leq_p v_2$ and $u_2 \leq_s v_2$, i.e. roots of elements of L_s are totally by both prefix and suffix.

In the case $u_2 \neq v_2$ this means $v_2 = wu_2 = u_2w'$ for w and $w' \in A^+$. Let then $k \geq 0$ be the greatest integer such that $|w^k| < |u_2|$.

When $k = 0$, this means $w = u_2y$ for some $y \in A^*$ and we take $x = u_2$. Otherwise, by a simple inductive argument over k , this means $u_2 = w^kx$ for some $x \in A^*$ with $|x| < |w|$. In that latter case, we have $v_2 = w^{k+1}x = w^kxw'$. Since

$|w'| = |w|$ it follows that $w' = yx$ for some $y \in A^+$ and thus $w = xy$. In all cases, $u_2 = x(yx)^k$ and $v_2 = x(yx)^{k+1}$.

By applying Lemma 13, with $v'_1 = u_1 \vee_s v_1$, $v'_3 = u_3 \vee_p v_3$ and $v' = (v'_1, v_2, v'_3)$, we have $v' \in L_s$. But we know that $(v_2 v'_3)_L v' = v'$ hence we also have $(v_2 v'_3)_L u \in L_s$. By applying product definition, this means that $u' = (v'_1, u_2, yxv'_3) \in L_s$ hence $v' = u'(yx)_C \in L_s$ hence, by an immediate pumping argument, $u'((yx)_C)^* \subseteq L_s$. \square

Lemma 14 already proves the finite case of Theorem 12. We assume now L_s is infinite.

In the case a single root appears in elements of L_s , for all $u = (u_1, x, u_3) \in L_s$ we have $(u_1)_L u(u_3)_L = u$ hence, because $s \neq 0$, $\varphi_L(u_1) \neq 0$ and $\varphi_L(u_3) \neq 0$.

By *safety assumption*, this means that $\varphi_C(u_1) \neq 0$ and $\varphi_C(u_3) \neq 0$. Since M is *finite*, while L_s is finite, this means that two distinct constraints have same images by φ_C . Applying Lemma 14 this implies all domains of tiles of L_s are factors of the same periodic bi-infinite word of the form $^\omega(xy)x(yx)^\omega$ for some x and $y \in A^*$ with $xy \neq 0$.

Depending on the case elements of L_s have infinitely many left constraints, right constraints or both left and right constraints, we conclude by showing that for all $v_1 \leq_s {}^\omega(xy)$ and $v_3 \leq_p (yx)^\omega$ there is $v \in L_s$ such that either $(v_1)_L v$, $v(v_3)_R$ or $(v_1)_L v(v_3)_R \in L_s$ hence, given a given fixed $u = (u_1, x, u_3) \in L_s$, either $(v_1)_L u$, $u(v_3)_R$ or $(v_1)_L u(v_3)_R \in L_s$. This proves the co-finiteness inclusion in case 1, 2 and 3 in Theorem 12.

In the case two distinct roots (hence infinitely many) appears in tiles of L_s . Applying Lemma 14 there is $x \in A^*$, $y \in A^+$ and $k \geq 0$ with $u = (u_1, u_2, xyu_3)$ such that $u((xy)_C)^* \subseteq L_s$. By choosing properly the initial u and v in Lemma 14, we may assume that $u_2 = x(yx)^k$ is the least root of elements of L_s and that $y(yx)^{k+1}$ is the second least. Then we claim that all roots of elements of L_s belongs to $x(yx)^k(yx)^*$.

Indeed, let then $(v_1, v_2, v_3) \in L_s$ and let $m \geq 0$ be the unique integer such that $v_2 = x(yx)^{m+k}x'$ with $x' <_p x$ hence $x = x'y'$ for some $y' \in A^+$. By an argument similar with the argument in the proof of Lemma 14, we can show that there is $u' = (v'_1, x(yx)^{m+k}, x'v'_3) \in L_s$ with $v' = (v'_1, v_2, v'_3) \in L_s$ henceforth with $v' = u'x'_C \in L_s$. But this also means that $ux'_C = (u_1, u_2x', y'yu_3) \in L_s$ which, by minimality of u_2xy as second least roots, forces x' to be equal to 1. This shows the inclusion stated in case 4 in Theorem 12.

Co-finiteness of the inclusion follows from the finiteness of M with arguments similar to arguments for case 3 above. \square

In order to conclude, let us mention that, at least for languages of positive tiles, a weakening of the notion of recognizability by means of premorphisms instead of morphisms have been studied [4]. How such a more expressive notion can be

relevantly applied to languages of arbitrary tiles is still an open problem.

References

- [1] Filipa Soares de Almeida. *Algebraic Aspects of Tiling Semigroups*. PhD thesis, Universidade de Lisboa, Faculdade de Ciências Departamento de Matemática, 2010.
- [2] F. Berthaut, D. Janin, and B. Martin. Advanced synchronization of audio or symbolic musical patterns. In *Sixth IEEE International Conference on Semantic Computing*, 2012 (to appear).
- [3] A. Dicky and D. Janin. Two-way automata and regular languages of overlapping tiles. Technical Report RR-14XX-12, LaBRI, Université de Bordeaux, 2012.
- [4] D. Janin. Quasi-recognizable vs MSO definable languages of one-dimensionnal overlapping tiles. In *37th International Symposium on Mathematical Foundations of Computer Science (MFCS)*, 2012 (to appear).
- [5] D. Janin. Vers une modélisation combinatoire des structures rythmiques simples de la musique. *Revue Francophone d’Informatique Musicale*, 2, 2012 (to appear).
- [6] J. Kellendonk. The local structure of tilings and their integer group of coinvariants. *Comm. Math. Phys.*, 187:115–157, 1997.
- [7] M. V. Lawson. *Inverse Semigroups : The theory of partial symmetries*. World Scientific, 1998.
- [8] Mark V. Lawson. McAlister semigroups. *Journal of Algebra*, 202(1):276 – 294, 1998.
- [9] Stuart W. Margolis and Jean-Eric Pin. Languages and inverse semigroups. In *ICALP*, volume 172 of *Lecture Notes in Computer Science*, pages 337–346. Springer, 1984.
- [10] W. D. Munn. Free inverse semigroups. *Proceedings of the London Mathematical Society*, 29(3):385–404, 1974.
- [11] Jean-Pierre Pécuchet. Automates boustrophedon, semi-groupe de birget et monoïde inversif libre. *ITA*, 19(1):71–100, 1985.

- [12] Pedro V. Silva. On free inverse monoid languages. *ITA*, 30(4):349–378, 1996.