Subshifts, Languages and Logic

Emmanuel Jeandel*[1] and Guillaume Theyssier**[2]

[1] Laboratoire d'informatique fondamentale de Marseille (LIF)
Aix-Marseille Université, CNRS
39 rue Joliot-Curie, 13 453 Marseille Cedex 13, France
[2] LAMA (Université de Savoie, CNRS)
Campus Scientifique, 73376 Le Bourget-du-lac cedex FRANCE

Abstract We study the Monadic Second Order (MSO) Hierarchy over infinite pictures, that is tilings. We give a characterization of existential MSO in terms of tilings and projections of tilings. Conversely, we characterise logic fragments corresponding to various classes of infinite pictures (subshifts of finite type, sofic subshifts).

1 Introduction

There is a close connection between words and monadic second-order (MSO) logic. Büchi and Elgot proved for finite words that MSO-formulas correspond exactly to regular languages. This relationship was developed for other classes of labeled graphs; trees or infinite words enjoy a similar connection. See [20,13] for a survey of existing results. Colorings of the entire plane, i.e tilings, represent a natural generalization of biinfinite words to higher dimensions, and as such enjoy similar properties. We plan to study in this paper tilings for the point of view of monadic second-order logic.

Tilings and logic have a shared history. The introduction of tilings can be traced back to Hao Wang [21], who introduced his celebrated tiles to study the (un)decidability of the $\forall\exists\forall$ fragment of first order logic. The undecidability of the domino problem by his PhD Student Berger [3] lead then to the undecidability of this fragment [5]. Seese [10,18] used the domino problem to prove that graphs with a decidable MSO theory have a bounded tree width. Makowsky[12,15] used the construction by Robinson [16] to give the first example of a finitely axiomatizable super-stable theory that is super-stable. More recently, Oger [14] gave generalizations of classical results on tilings to locally finite relational structures. See the survey [2] for more details.

Previously, a finite variant of tilings, called tiling pictures, was studied [6,7]. Tiling pictures correspond to colorings of a *finite* region of the plane, this region being bordered by special '#' symbols. It is proven for this particular model that language recognized by EMSO-formulas correspond exactly to so-called finite tiling systems, i.e. projections of finite tilings.

* Emmanuel.Jeandel@lif.univ-mrs.fr
** guillaume.theyssier@univ-savoie.fr

The equivalent of finite tiling systems for infinite pictures are so-called *sofic subshifts* [22]. A *sofic subshift* represents intuitively local properties and ensures that every point of the plane behaves in the same way. As a consequence, there is no general way to enforce that some specific color, say $\square$ appears at least once. Hence, some simple first-order existential formulas have no equivalent as sofic subshift (and even subshift). This is where the border of # for finite pictures play an important role: Without such a border, results on finite pictures would also stumble on this issue.

We deal primarily in this article with subshifts. See [1] for other acceptance conditions (what we called subshifts of finite type correspond to A-acceptance in this paper).

Finally, note that all decision problems in our context are non-trivial : To decide if a universal first-order formula is satisfiable (the domino problem, presented earlier) is not recursive. Worse, it is Σ_1^1-hard to decide if a tiling of the plane exists where some given color appears infinitely often [9,1]. As a consequence, the satisfiability of MSO-formulas is at least Σ_1^1-hard.

2 Symbolic Spaces and Logic

2.1 Configurations

Let $d \geq 1$ be a fixed integer and consider the discrete lattice $\mathbb{Z}^d$. For any finite set Q, a Q-configuration is a function from $\mathbb{Z}^d$ to Q. Q may be seen as a set of *colors* or *states*. An element of $\mathbb{Z}^d$ will be called a *cell*. A configuration will usually be denoted C, M or N.

Fig. 1 shows an example of two different configurations of $\mathbb{Z}^2$ over a set Q of 5 colors. As a configuration is infinite, only a finite fragment of the configurations is represented in the figure. The reader has to use his imagination to decide what colors do appear in the rest of the configuration. We choose not to represent which cell of the picture is the origin $(0,0)$ (we use only translation invariant properties).

A *pattern* is a partial configuration. A pattern $P : X \to Q$ where $X \subseteq \mathbb{Z}^2$ occurs in $C \in Q^{\mathbb{Z}^d}$ at position z_0 if

$$\forall z \in X,\ C(z_0 + z) = P(z).$$

We say that P occurs in C if it occurs at some position in C. As an example the pattern P of Fig 2 occurs in the configuration M but not in N (or more accurately not on the finite fragment of N depicted in the figure). A finite pattern is a partial configuration of finite domain. All patterns in the following will be finite. The *language* $\mathcal{L}(C)$ of a configuration C is the set of finite patterns that occur in C. We naturally extend this notion to sets of configurations.

A *subshift* is a natural concept that captures both the notion of *uniformity* and *locality*: the only description "available" from a configuration C is the finite patterns it contains, that is $\mathcal{L}(C)$. Given a set $\mathcal{F}$ of patterns, let $X_{\mathcal{F}}$ be the set

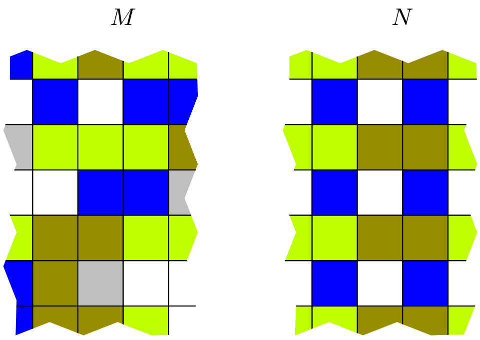

Figure 1. Two configurations

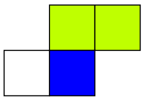

Figure 2. A pattern P. P appears in M but presumably not in N

of all configurations where no patterns of $\mathcal{F}$ occurs.

$$X_{\mathcal{F}} = \{C|\mathcal{L}(C) \cap \mathcal{F} = \emptyset\}$$

$\mathcal{F}$ is usually called the set of forbidden patterns or the *forbidden language*. A set of the form $X_{\mathcal{F}}$ is called a *subshift*.

A subshift can be equivalentely defined by topology considerations. Endow the set of configurations $Q^{\mathbb{Z}^d}$ with the product topology: A sequence $(C_n)_{n\in\mathbb{N}}$ of configurations converges to a configuration C if the sequence ultimately agree with C on every $z \in \mathbb{Z}^2$. Then a subshift is a closed subset of $Q^{\mathbb{Z}^d}$ also closed by shift maps.

A *subshift of finite type* (or *tiling*) correspond to a finite set $\mathcal{F}$: it is the set of configurations C such that no pattern in $\mathcal{F}$ occurs in C. If all patterns of $\mathcal{F}$ are of diameter n, this means that we only have to see a configuration through a window of size n to know if it is a tiling, hence the locality.

Given two state sets Q_1 and Q_2, a projection is a map $\pi : Q_1 \to Q_2$. We naturally extend it to $\pi : Q_1^{\mathbb{Z}^d} \to Q_2^{\mathbb{Z}^d}$ by $\pi(C)(z) = \pi(C(z))$. A *sofic subshift* of state set Q_2 is the image by some projection π of some subshift of finite type of state set Q_1. It is also a subshift (clearly closed by shift maps, and topologically closed because projections are continuous maps on a compact space). A sofic subshift is a natural object in tiling theory, although quite never mentioned explicitly. It represents the concept of *decoration*: some of the tiles we assemble to obtain the tilings may be decorated, but we forgot the decoration when we observe the tiling.

2.2 Structures

From now on, we restrict to dimension 2. A configuration will be seen in this article as an infinite structure. The signature τ contains four unary maps North, South, East, West and a predicate P_c for each color $c \in Q$.

A configuration M will be seen as a structure $\mathfrak{M}$ in the following way:

- The elements of $\mathfrak{M}$ are the points of $\mathbb{Z}^2$.
- North is interpreted by $\mathsf{North}^{\mathfrak{M}}((x,y)) = (x, y+1)$, East is interpreted by $\mathsf{East}^{\mathfrak{M}}((x,y)) = (x+1, y)$. $\mathsf{South}^{\mathfrak{M}}$ and $\mathsf{West}^{\mathfrak{M}}$ are interpreted similarly
- $P_c^{\mathfrak{M}}((x,y))$ is true if and only if the point at coordinate (x,y) is of color c, that is if $M(x,y) = c$.

As an example, the configuration M of Fig. 1 has three consecutive cells with the color $\color{yellowgreen}\blacksquare$. That is, the following formula is true:

$$\mathfrak{M} \models \exists z, P_{\color{yellowgreen}\blacksquare}(z) \wedge P_{\color{yellowgreen}\blacksquare}(\mathsf{East}(z)) \wedge P_{\color{yellowgreen}\blacksquare}(\mathsf{East}(\mathsf{East}(z)))$$

As another example, the following formula states that the configuration has a vertical period of 2 (the color in the cell (x,y) is the same as the color in the cell $(x, y+2)$). The formula is false in the structure $\mathfrak{M}$ and true in the structure $\mathfrak{N}$ (if the reader chose to color the cells of N not shown in the picture correctly):

$$\forall z, \begin{cases} P_{\color{olive}\blacksquare}(z) \implies P_{\color{olive}\blacksquare}(\mathsf{North}(\mathsf{North}(z))) \\ P_{\square}(z) \implies P_{\square}(\mathsf{North}(\mathsf{North}(z))) \\ P_{\color{blue}\blacksquare}(z) \implies P_{\color{blue}\blacksquare}(\mathsf{North}(\mathsf{North}(z))) \\ P_{\color{yellowgreen}\blacksquare}(z) \implies P_{\color{yellowgreen}\blacksquare}(\mathsf{North}(\mathsf{North}(z))) \\ P_{\color{gray}\blacksquare}(z) \implies P_{\color{gray}\blacksquare}(\mathsf{North}(\mathsf{North}(z))) \end{cases}$$

2.3 Monadic Second-Order Logic

This paper studies connection between subshifts (seen as structures as explained above) and monadic second order sentences. First order variables (x, y, z, ...) are interpreted as points of $\mathbb{Z}^2$ and (monadic) second order variables (X, Y, Z, ...) as subsets of $\mathbb{Z}^2$.

Monadic second order formulas are defined as follows:

- a term is either a first-order variable or a function (South, North, East, West) applied to a term ;
- atomic formulas are of the form $t_1 = t_2$ or $X(t_1)$ where t_1 and t_2 are terms and X is either a second order variable or a color predicate ;
- formulas are build up from atomic formulas by means of boolean connectives and quantifiers $\exists$ and $\forall$ (which can be applied either to first-order variables or second order variables).

A formula is *closed* if no variable occurs free in it. A formula is FO if no second-order quantifier occurs in it. A formula is EMSO if it is of the form

$$\exists X_1, \ldots, \exists X_n, \phi(X)$$

where ϕ is FO. Given a formula $\phi(X_1, \ldots, X_n)$ with no free first-order variable and having only $X_1, \ldots, X_n$ as free second-order variables, a configuration M together with subsets $E_1, \ldots, E_n$ is a model of $\phi(X_1, \ldots, X_n)$, denoted

$$(M, E_1, \ldots, E_n) \models \phi(X_1, \ldots, X_n),$$

if ϕ is satisfied (in the usual sense) when M is interpreted as $\mathfrak{M}$ (see previous section) and E_i interprets X_i.

2.4 Definability

This paper studies the following problems: Given a formula ϕ of some logic, what can be said of the configurations that satisfy ϕ? Conversely, given a subshift, what kind of formula can characterise it?

Definition 2.1 *A set S of Q-configurations is defined by ϕ if*

$$S = \left\{ M \in Q^{\mathbb{Z}^2} \middle| \mathfrak{M} \models \phi \right\}$$

Two formulas ϕ and ϕ' are equivalent iff they define the same set of configurations.

A set S is $\mathcal{C}$-definable if it is defined by a formula $\phi \in \mathcal{C}$.

Note that a definable set is always closed by shift (a shift between 2 configurations induces an isomorphism between corresponding structures). It is not always closed: The set of $\{\blacksquare, \square\}$-configurations defined by the formula $\phi : \exists z, P_{\blacksquare}(z)$ contains all configurations except the all-white one, hence is not closed.

When we are dealing with MSO formulas, the following remark is useful: second-order quantifiers may be represented as projection operations on sets of configurations. We formalize now this notion.

If $\pi : Q_1 \mapsto Q_2$ is a projection and S is a set of Q_1-configurations, we define the two following operators:

$$E(\pi)(S) = \left\{ M \in (Q_2)^{\mathbb{Z}^2} \middle| \exists N \in (Q_1)^{\mathbb{Z}^2}, \pi(N) = M \wedge N \in S \right\}$$
$$A(\pi)(S) = \left\{ M \in (Q_2)^{\mathbb{Z}^2} \middle| \forall N \in (Q_1)^{\mathbb{Z}^2}, \pi(N) = M \implies N \in S \right\}$$

Note that A is a dual of E, that is $A(\pi)(S) = {}^cE(\pi)({}^cS)$ where ^{c} represents complementation.

Proposition 2.2

- *A set S of Q-configurations is EMSO-definable if and only if there exists a set S' of Q' configurations and a map $\pi : Q' \mapsto Q$ such that $S = E(\pi)(S')$ and S' is FO-definable.*
- *The class of MSO-definable sets is the closure of the class of FO-definable sets by the operators E and A.*

Proof (Sketch). We prove here only the first item.

- Let $\phi = \exists X, \psi$ be a EMSO formula that defines a set S of Q-configurations. Let $Q' = Q \times \{0,1\}$ and π be the canonical projection from Q' to Q.
  Consider the formula ψ' obtained from ψ by replacing $X(t)$ by $\vee_{c \in Q} P_{(c,1)}(t)$ and $P_c(t)$ by $P_{(c,0)}(t) \vee P_{(c,1)}(t)$.
  Let S' be a set of Q' configurations defined by ψ'. Then is it clear that $S = E(\pi)(S')$. The generalization to more than one existential quantifier is straightforward.
- Let $S = E(\pi)(S')$ be a set of Q configurations, and S' FO-definable by the formula ϕ. Denote by $c_1 \dots c_n$ the elements of Q'. Consider the formula ϕ' obtained from ϕ where each P_{c_i} is replaced by X_i. Let

$$\psi = \exists X_1, \dots, \exists X_n, \begin{cases} \forall z, \vee_i X_i(z) \\ \forall z, \wedge_{i \neq j} (\neg X_i(z) \vee \neg X_j(z)) \\ \forall z, \wedge_i \left(X_i z \implies P_{\pi(c_i)}(z) \right) \\ \phi' \end{cases}$$

  Then ψ defines S. Note that the formula ψ constructed above is of the form $\exists X_1, \dots, \exists X_n (\forall z, \psi'(z)) \wedge \phi'$. This will be important later. □

Second-order quantifications will then be regarded in this paper either as projections operators or sets quantifiers.

3 Hanf Locality Lemma and EMSO

The first-order logic has a property that makes it suitable to deal with tilings and configurations: it is local. This is illustrated by Hanf's lemma [8,4,11].

Definition 3.1 *Two Q-configurations M and N are (n,k)-equivalent if for each Q-pattern P of size n:*

- *If P appears in M less than k times, then P appears the exact same number of times in M and in N*
- *If P appears in M more than k times, then P appears in N more than k times*

This notion is indeed an equivalence relation. Given n and k, it is clear that there is only finitely many equivalence classes for this relation.

The Hanf's local lemma can be formulated in our context as follows:

Theorem 3.2 *For every FO formula ϕ, there exists (n,k) such that*

$$\textit{if } M \textit{ and } N \textit{ are } (n,k) \textit{ equivalent, then } \mathfrak{M} \models \phi \iff \mathfrak{N} \models \phi$$

Corollary 3.3 *Every FO-definable set is a (finite) union of some (n,k)-equivalence classes.*

This is theorem 3.3 in [7], stated for finite configurations. Lemma 3.5 in the same paper gives a proof of Hanf's Local Lemma in our context.

Given (P, k) we consider the set $S_{=k}(P)$ of all configurations such that the pattern P occurs exactly k times (k may be taken equal to 0). The set $S_{\geq k}(P)$ is the set of all configurations such that the pattern P occurs more than k times.

We may rephrase the preceding corollary as:

Corollary 3.4 *Every FO-definable set is a positive combination (i.e. unions and intersections) of some $S_{=k}(P)$ and some $S_{\geq k}(P)$*

Theorem 3.5 *Every EMSO-definable set can be defined by a formula ϕ of the form:*

$$\exists X_1, \ldots, \exists X_n,\ \big(\forall z_1,\ \phi_1(z_1, X_1, \ldots, X_n)\big) \\ \wedge \big(\exists z_1, \ldots, \exists z_p,\ \phi_2(z_1 \ldots z_p, X_1, \ldots, X_n)\big),$$

where ϕ_1 and ϕ_2 are quantifier-free formulas.

See [20, Corollary 4.1] or [19, Corollary 4.2] for a similar result. This result is an easy consequence of [17, Theorem 3.2] (see also the corrigendum). We include here a full proof.

Proof. Let $\mathcal{C}$ be the set of such formulas. We proceed in three steps:

- Every EMSO-definable set is the projection of a positive combination of some $S_{=k}(P)$ and $S_{\geq k}(P)$ (using prop. 2.2 and the preceding corollary)
- Every $S_{=}(P, k)$ (resp. $S_{\geq}(P, k)$) is $\mathcal{C}$-definable
- $\mathcal{C}$-definable sets are closed by (finite) union, intersection and projections.

$\mathcal{C}$-definable sets are closed by projection using the equivalence of prop. 2.2 in the two directions, the note at the end of the proof and some easy formula equivalences. The same goes for intersection.

Now we prove that $\mathcal{C}$-definable sets are closed by union. The difficulty is to ensure that we use only one universal quantifier. Let ϕ and ϕ' be two $\mathcal{C}$-formulas defining sets S_1 and S_2. We can suppose that ϕ and ϕ' use the same numbers of second-order quantifiers and of first-order existential quantifiers.

Then the formula

$$\exists X, \exists X_1, \ldots, \exists X_n, \forall z_1, \begin{cases} X(z_1) \iff X(\mathsf{North}(z_1)) \\ X(z_1) \iff X(\mathsf{East}(z_1)) \\ X(z_1) \implies \phi_1(z_1, X_1 \ldots X_n) \\ \neg X(z_1) \implies \phi'_1(z_1, X_1 \ldots X_n) \end{cases} \\ \wedge \exists z_1, \ldots, \exists z_p \bigvee \begin{matrix} X(z_1) \wedge \phi_2(z_1 \ldots z_p, X_1 \ldots X_n) \\ \neg X(z_1) \wedge \phi'_2(z_1 \ldots z_p, X_1 \ldots X_n) \end{matrix}$$

defines $S_1 \cup S_2$ (the disjunction is obtained through variable X which is forced to represent either the empty set or the whole plane $\mathbb{Z}^2$).

It is now sufficient to prove that a $S_{=k}(P)$ set (resp. a $S_{\geq k}(P)$ set) is definable by a $\mathcal{C}$-formula. Let $\phi_P(z)$ be the quantifier-free formula such that $\phi_P(z)$ is true if and only if P appears at position z.

Then $S_{=k}(P)$ is definable by

$$\exists X_1 \dots \exists X_k \exists A_1, \dots, \exists A_k, \forall x \begin{cases} \wedge_i A_i(x) \iff [A_i(\mathsf{North}(x)) \wedge A_i(\mathsf{East}(x))] \\ \wedge_i X_i(x) \iff [A_i(x) \wedge \neg A_i(\mathsf{South}(x)) \wedge \neg A_i(\mathsf{West}(x))] \\ \wedge_{i \neq j} X_i(x) \implies \neg X_j(x) \\ (\vee_i X_i(x)) \iff \phi_P(x) \end{cases}$$
$$\wedge \exists z_1, \dots, \exists z_k, X_1(z_1) \wedge \dots \wedge X_k(z_k)$$

The formula ensures indeed that A_i represents a quarter of the plane, X_i being a singleton representing the corner of that plane. If $k = 0$ this becomes $\forall x, \neg\phi_P(x)$. To obtain a formula for $S_{\geq k}(P)$, change the last $\iff$ to a $\implies$ in the formula. □

4 Logic Characterization of SFT and Sofic Subshifts

We start by a characterization of subshifts of finite type (SFTs, i.e tilings). The problem with SFTs is that they are closed neither by projection nor by union. As a consequence, the corresponding class of formulas is not very interesting:

Theorem 4.1 *A set of configurations is a SFT if and only if it is defined by a formula of the form*

$$\forall z, \psi(z)$$

where ψ is quantifier-free.

Note that there is only one quantifier in this formula. Formulas with more than one universal quantifier do not always correspond to SFT: This is due to SFTs not being closed by union.

Proof. Let $P_1 \dots P_n$ be patterns. To each P_i we associate the quantifier-free formula $\phi_{P_i}(z)$ which is true if and only if P_i appears at the position z. Then the subshifts that forbids patterns $P_1 \dots P_n$ is defined by the formula:

$$\forall z, \neg\phi_{P_1(z)} \wedge \dots \wedge \neg\phi_{P_n(z)}$$

Conversely, let ψ be a quantifier-free formula. Each term t_i in ψ is of the form $f_i(z)$ where f_i is some combination of the functions North, South, East and West, each f_i thus representing somehow some vector z_i ($f_i(z) = z + z_i$). Let Z be the collection of all vectors z_i that appear in the formula ψ. Now the fact that ψ is true at the position z only depends on the colors of the configurations in points $(z + z_1), \dots, (z + z_n)$, i.e. on the *pattern* of domain Z that occurs at position z. Let $\mathcal{P}$ be the set of patterns of domain Z that makes ψ false. Then the set S defined by ψ is the set of configurations where no patterns in $\mathcal{P}$ occurs, hence a SFT. □

Theorem 4.2 *A set S is a sofic subshift if and only if it is definable by a formula of the form*

$$\exists X_1, \ldots, \exists X_n, \forall z_1, \ldots, \forall z_p,\ \psi(X_1, \ldots, X_n, z_1 \ldots z_p)$$

where ψ is quantifier-free. Moreover, any such formula is equivalent to a formula of the same form but with a single universal quantifier ($p = 1$).

Note that the real difficulty in the proof of this theorem is to treat the only binary predicate, the equality ($=$). The reader might try to find a sofic subshift corresponding to the following formula before reading the proof:

$$\forall x, y, \left(P_{\color{lime}\blacksquare}(x) \wedge P_{\color{blue}\blacksquare}(\mathsf{East}(y))\right) \implies x = y$$

Proof. A sofic subshift being a projection of a SFT, one direction of the first assertion follows from the previous theorem and proposition 2.2.

Let $\mathcal{C}$ be the class of formulas of the form:

$$\exists X_1, \ldots, \exists X_n, \forall z_1, \ldots, \forall z_p,\ \psi(X_1, \ldots, X_n, z_1 \ldots z_{p_i})$$

Now we prove by induction on the number p of universal quantifiers that each formula of $\mathcal{C}$ is equivalent to a formula with only one universal quantifier. There is nothing to prove for $p = 1$.

First, we rewrite the formula in conjunctive normal form:

$$\exists X_1, \ldots, \exists X_n, \forall z_1, \ldots, \forall z_p,\ \wedge_i \psi_i(X_1, \ldots, X_n, z_1 \ldots z_p)$$

where ψ_i is disjunctive. This is equivalent to

$$\exists X_1, \ldots, \exists X_n, \wedge_i \forall z_1, \ldots, \forall z_p,\ \psi_i(X_1, \ldots, X_n, z_1 \ldots z_p) \equiv \exists X_1, \ldots, \exists X_n, \wedge_i \eta_i$$

Now we treat each η_i separately. ψ_i is a disjunction of four types of formulas:

$$\bullet P_c(f(x)) \qquad \bullet \neg P_c(f(x)) \qquad \bullet f(x) = y \qquad \bullet f(x) \neq y$$

because terms are made only of bijective functions (compositions of North, South, East, West). We may suppose the last case never happens: $\forall x, y, z f(x) \neq y \vee \psi(x, y, z)$ is equivalent to $\forall x, z, \psi(x, f(x), z)$. We may rewrite

$$\psi_i(z_1 \ldots z_p) \equiv \epsilon(z_p) \vee z_p = f(z_{k_1}) \vee \cdots \vee z_p = f(z_{k_m}) \vee \theta(z_1 \ldots z_{p-1})$$

(we forgot the second-order variables to simplify notations)

We may suppose that no formula is of the form $z_p = z_p$. Now is the key argument: Suppose that there are strictly more that m values of z such that $\epsilon(z)$ is false. Then given $z_1 \ldots z_{p-1}$ we may find a z_p such that the formula $\epsilon(z_p) \vee (z_p = f(z_{k_1})) \vee \cdots \vee (z_p = f(z_{k_m}))$ is false. That is, if there are more than m values of z so that $\epsilon(z)$ is false, then

$$\forall z_1, \ldots, \forall z_{p-1},\ \theta(z_1 \ldots z_{p-1})$$

must be true.

As a consequence, our formula η_i is equivalent to the disjunction of the formula

$$\forall z_1, \dots, \forall z_{p-1},\ \theta(z_1 \dots z_{p-1})$$

and the formula

$$\exists S_1, \dots, \exists S_m, \begin{cases} \Psi_i \\ \forall z, \vee_i S_i(z) \iff \neg\epsilon(z) \\ \forall z_1, \dots, \forall z_{p-1},\ S_1(f(z_{k_1})) \vee \dots \vee S_m(f(z_{k_m})) \vee \theta(z_1 \dots z_{p-1}) \end{cases}$$

where Ψ_i express that S_i has at most one element and is defined as follows:

$$\Psi_i \stackrel{def}{=} \exists A, \forall x \begin{cases} A(x) \iff A(\mathsf{North}(x)) \wedge A(\mathsf{East}(x)) \\ S_i(x) \iff A(x) \wedge \neg A(\mathsf{South}(x)) \wedge \neg A(\mathsf{West}(x)) \end{cases}$$

Simplifying notations, our formula η_i is equivalent to

$$\forall z_1, \dots, \forall z_{p-1},\ \theta(z_1 \dots z_{p-1}) \vee \exists P_1, \dots, \exists P_q \forall z_1, \dots, \forall z_{p-1},\ \zeta(z_1 \dots z_{p-1})$$

which is equivalent to

$$\exists X, \exists P_1, \dots, \exists P_q \forall z_1, \dots, \forall z_{p-1}, \begin{cases} X(z_1) \iff X(\mathsf{North}(z_1)) \\ X(z_1) \iff X(\mathsf{East}(z_1)) \\ X(z_1) \implies \theta(z_1, \dots z_{p-1}) \\ \neg X(z_1) \implies \zeta(z_1, \dots z_{p-1}) \end{cases}$$

Now report this new formula instead of η_i to obtain a formula

$$\exists X_1, \dots, \exists X_n, \wedge_i \exists R_1, \dots, \exists R_{q_i}, \forall z_1, \dots, \forall z_{p-1},\ \theta_i(z_1 \dots z_{p_i}, R_1 \dots R_{q_i})$$

equivalent to

$$\exists X_1, \dots, \exists X_n, \exists R_{11}, \dots, \exists R_{kq_k}, \forall z_1, \dots, \forall z_{p-1},\ \wedge_i \theta_i(z_1 \dots z_{p_i}, R_{i1} \dots R_{iq_i})$$

We finally obtain a formula of $\mathcal{C}$ with $p-1$ universal quantifiers, and we may conclude by induction.

To finish the proof, a formula with only one universal quantifier

$$\exists X_1, \dots, \exists X_n, \forall z, \theta(z)$$

defines indeed a sofic subshift (use the proof of theorem 4.1 to conclude that this formula defines a projection of a SFT, hence a sofic subshift) □

5 Separation Result

Theorems 3.5 and 4.2 above suggest that EMSO-definable subshifts are not necessarily sofic. We will show in this section that the set of EMSO-definable subshifts is indeed strictly larger than the set of sofic subshifts. The proof is based on the analysis of the computational complexity of forbidden languages. It is well-known that sofic subshifts have a recursively enumerable forbidden language. The following theorem shows that the forbidden language of an MSO-definable subshift can be arbitrarily high in the arithmetical hierarchy.

This is not surprising since arbitrary Turing computation can be defined via first order formulas (using tilesets) and second order quantifiers can be used to simulate quantification of the arithmetical hierarchy. However, some care must be taken to ensure that the set of configurations obtained is a subshift.

Theorem 5.1 *Let E be an arithmetical set. Then there is an MSO-definable subshift with forbidden language $\mathcal{F}$ such that E reduces to $\mathcal{F}$ (for many-one reduction).*

Proof (sketch). Suppose that the complement of E is defined as the set of integers m such that:

$$\exists x_1, \forall x_2, \ldots, \exists/\forall x_n, R(m, x_1, \ldots, x_n)$$

where R is a recursive relation. We first build a formula ϕ defining the set of configurations representing a successful computation of R on some input $m, x_1, \ldots, x_n$. Consider 3 colors c_l, c and c_r and additional second order variables $X_1, \ldots, X_n$ and $S_1, \ldots, S_n$. The input $(m, x_1, \ldots, x_n)$ to the computation is encoded in unary on an horizontal segment using colors c_l and c_r and variables S_i as separators, precisely: first an occurrence of c_l then m occurrences of c, then an occurrence of c_r and, for each successive $1 \leq i \leq n$, x_i positions in X_i before a position of S_i. Let ϕ_1 be the FO formula expressing the following:

1. there is exactly 1 occurrence of c_l and the same for c_r and all S_i are singletons;
2. starting from an occurrence c_l and going east until reaching S_n, the only possible successions of states are those forming a valid input as explained above.

Now, the computation of R on any input encoded as above can be simulated via tiling constraints in the usual way. Consider sufficiently many new second order variables $Y_1, \ldots, Y_p$ to handle the computation and let ϕ_2 be the FO formula expressing that:

1. a valid computation starts at the north of an occurrence of c_l;
2. there is exactly one occurrence of the halting state (represented by some Y_i) in the whole configuration.

We define ϕ by:

$$\exists X_1, \forall X_2, \ldots, \exists/\forall X_n, \exists S_1, \ldots, \exists S_n, \exists Y_1, \ldots, \exists Y_p, \phi_1 \wedge \phi_2.$$

Finally let ψ be the following FO formula: $(\forall z, \neg P_{c_l}) \vee (\forall z, \neg P_{c_r})$. Let X be the set defined by $\phi \vee \psi$. By construction, a finite (unidimensional) pattern of the form $c_l c^m c_r$ appears in some configuration of X if and only if $m \notin E$. Therefore E is many-one reducible to the forbidden language of X.

To conclude the proof it is sufficient to check that X is closed. To see this, consider a sequence $(C_n)_n$ of configurations of X converging to some configuration C. C has at most one occurrence of c_l and one occurrence of c_r. If one of these two states does not occur in C then $C \in X$ since ψ is verified. If, conversely, both c_l and c_r occur (once each) then any pattern containing both occurrences also occurs in some configuration C_n verifying ϕ. But ϕ is such that any modification outside the segment between c_l and c_r in C_n does not change the fact that ϕ is satisfied provided no new c_l and c_r colors are added. Therefore ϕ is also satisfied by C and $C \in X$. □

The theorem gives the claimed separation result for subshifts of EMSO.

Corollary 5.2 *There are EMSO-definable subshifts which are not sofic.*

Proof. In the previous theorem, choose E, to be the complement of the set of integers m for which there is x such that machine m halts on empty input in less than x steps. E is not recursively enumerable and, using the construction of the proof above, it is reducible to the forbidden language of an EMSO-definable subshift. □

6 A Characterization of EMSO

EMSO-definable sets are projections of FO-definable sets (proposition 2.2). Besides, sofic subshifts are projections of subshifts of finite type (or tilings). Previous results show that the correspondence sofic↔EMSO fails. However, we will show in this section how EMSO can be characterized through projections of “locally checkable” configurations.

Corollary 3.4 expresses that FO-definable sets are essentially captured by counting occurrences of patterns up to some value. The key idea in the following is that this counting can be achieved by local checkings (equivalently, by tiling constraints), provided it is limited to a finite and explicitly delimited region. This idea was successfully used in [7] in the context of picture languages: pictures are rectangular finite patterns with a border made explicit using a special state (which occurs all along the border and nowhere else). We will proceed here quite differently. Instead of putting special states on borders of some rectangular zone, we will simply require that two special subsets of states Q_0 and Q_1 are present in the configuration: we call a (Q_0, Q_1)-*marked configuration* any configuration that contains both a color $q \in Q_0$ and some color $q' \in Q_1$ somewhere. By extension, given a subshift Σ over Q and two subsets $Q_0 \subseteq Q$ and $Q_1 \subseteq Q$, the *doubly-marked set* Σ_{Q_0,Q_1} is the set of (Q_0, Q_1)-marked configurations of Σ. Finally, a *doubly-marked set of finite type* is a set Σ_{Q_0,Q_1} for some SFT Σ and some Q_0, Q_1.

Lemma 6.1 *For any finite pattern P and any $k \geq 0$, $S_{=k}(P)$ is the projection of some doubly-marked set of finite type. The same result holds for $S_{\geq k}(P)$.*

Moreover, any positive combination (union and intersection) of projections of doubly-marked sets of finite type is also the projection of some doubly-marked sets of finite type.

Proof (sketch). We consider some base alphabet Q, some pattern P and some $k \geq 0$. We will build a doubly-marked set of finite type over alphabet $Q' = Q \times Q_+$ and then project back on Q. Q_+ is itself a product of different layers. The first layer can take values $\{0, 1, 2\}$ and is devoted to the definition of the marker subsets Q_0 and Q_1: a state is in Q_i for $i \in \{0, 1\}$ if and only if its value on the layer is i.

We first show how to convert the apparition in a configuration of two marked positions, by Q_0 and Q_1, into a locally identifiable rectangular zone. The zone is defined by two opposite corners corresponding to an occurrence of some state of Q_0 and Q_1 respectively. This can be done using only finite type constraints as follows. By adding a new layer of states, one can ensure that there is a unique occurrence of a state of Q_0 and maintain everywhere the following information:

1. $N_{Q_0}(z) \equiv$ the position z is at the north of the (unique) occurrence of a state from Q_0,
2. $E_{Q_0}(z) \equiv$ the position z is at the east of the occurrence of a state from Q_0.

The same can be done for Q_1. From that, the membership to the rectangular zone is defined at any position z by the following predicate (see figure 6):

$$Z(z) \equiv N_{Q_0}(z) \neq N_{Q_1}(z) \wedge E_{Q_0}(z) \neq E_{Q_1}(z).$$

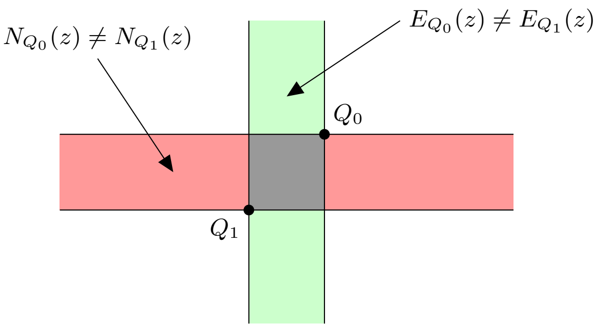

Figure 3. The rectangular zone in dark gray defined by predicate $Z(z)$.

We can also define locally the border of the zone: precisely, cells not in the zone but adjacent to it. Now define $P(z)$ to be true if and only if z is the lower-left position in an occurrence of the pattern P. We add k new layers, each one storing (among other things) a predicate $C_i(z)$ verifying

$$C_i(z) \Rightarrow Z(z) \wedge P(z) \wedge \bigwedge_{j \neq i} \neg C_j(z).$$

Moreover, on each layer i, we enforce that exactly 1 position z verifies $C_i(z)$: this can be done by maintaining north/south and east/west tags (as for Q_0 above) and requiring that the north (resp. south) border of the rectangular zone sees only the north (resp. south) tag and the same for east/west. Finally, we add the constraint:

$$P(z) \wedge Z(z) \Rightarrow \bigvee_i C_i$$

expressing that each occurrence of P in the zone mut be "marked" by some C_i. Hence, the only admissible (Q_0, Q_1)-marked configurations are those whose rectangular zone contains exactly k occurrences of pattern P. We thus obtain exactly $S_{\geq k}(P)$ after projection. To obtain $S_{=k}(P)$, it suffices to add the constraint:

$$P(z) \Rightarrow Z(z)$$

in order to forbid occurrences of P outside the rectangular zone.

To conclude the proof we show that finite unions or intersections of projections of doubly-marked sets of finite type are also projections of doubly-marked sets of finite type. Consider two SFT X over Q and Y over Q' and two pairs of marker subsets $Q_0, Q_1 \subseteq Q$ and $Q'_0, Q'_1 \subseteq Q'$. Let $\pi_1 : Q \to A$ and $\pi_2 : Q' \to A$ be two projections.

First, for the case of union, we can suppose (up to renaming of states) that Q and Q' are disjoint and define the SFT Σ over alphabet $Q \cup Q'$ as follows:

- 2 adjacent positions must be both in Q or both in Q';
- any pattern forbidden in X or Y is forbidden in Σ.

Clearly, $\pi(\Sigma_{Q_0 \cup Q'_0, Q_1 \cup Q'_1}) = \pi_1(X_{Q_0,Q_1}) \cup \pi_2(Y_{Q'_0,Q'_1})$ where $\pi(q)$ is $\pi_1(q)$ when $q \in Q$ and $\pi_2(q)$ else.

Now, for intersections, consider the SFT Σ over the fiber product

$$Q_\times = \{(q, q') \in Q \times Q' | \pi_1(q) = \pi_2(q')\}$$

and defined as follows: a pattern is forbidden if its projection on the component Q (resp. Q') is forbidden in X (resp. Y);

If we define π as π_1 applied to the Q-component of states, and if E is the set of configuration of Σ such that states from Q_0 and Q_1 appear on the first component and states from Q'_0 and Q'_1 appear on the second one, then we have:

$$\pi(E) = \pi_1(X_{Q_0,Q_1}) \cup \pi_2(Y_{Q'_0,Q'_1}).$$

To conclude the proof, it is sufficient to obtain E as the projection of some doubly-marked set of finite type. This can be done starting from Σ and adding a new component of states whose behaviour is to define a zone from two markers (as in the first part of this proof) and check that the zone contains occurrences of Q_0, Q_1, Q'_0 and Q'_1 in the appropriate components. □

Theorem 6.2 *A set is EMSO-definable if and only if it is the projection of a doubly-marked set of finite type.*

Proof. First, a doubly-marked set of finite type is an FO-definable set because SFT are FO-definable (theorem 4.1) and the restriction to doubly-marked configurations can be expressed through a simple existential FO formula. Thus the projection of a doubly-marked set of finite type is EMSO-definable.

The opposite direction follows immediately from proposition 2.2 and corollary 3.4 and the lemma above. □

At this point, one could wonder whether considering simply-marked set of finite type is sufficient to capture EMSO via projections. In fact the presence of 2 markers is necessary in the above theorem: considering the set Σ_{Q_0,Q_1} where Σ is the full shift $Q^{\mathbb{Z}^2}$ and Q_0 and Q_1 are distinct singleton subsets of Q, a simple compactness argument allows to show that it is not the projection of any simply-marked set of finite type.

7 Open Problems

- Is the second order alternation hierarchy strict for MSO (considering our model-theoretic equivalence)?
- One can prove that theorem 4.1 also holds for formulas of the form:

$$\forall X_1 \ldots \forall X_n, \forall z,\ \psi(z, X_1 \ldots X_n)$$

  where ψ is quantifier-free. Hence, adding universal second-order quantifiers does not increase the expression power of formulas of theorem 4.1. More generally, let $\mathcal{C}$ be the class of formulas of the form

$$\forall X_1, \exists X_2, \ldots, \forall/\exists X_n, \forall z_1, \ldots, \forall z_p, \phi(X_1, \ldots, X_n, z_1, \ldots, z_p).$$

  One can check that any formula in $\mathcal{C}$ defines a subshift. Is the second-order quantifiers alternation hierarchy strict in $\mathcal{C}$? On the contrary, do all formulas in $\mathcal{C}$ represent sofic subshifts ?

References

1. Jan-Henrik Altenbernd, Wolfgang Thomas, and Stefan Wohrle. Tiling systems over infinite pictures and their acceptance conditions. In *Developments in Language Theory*, 2003.
2. A. Ballier and E. Jeandel. Tilings and model theory. *First Symposium on Cellular Automata Journées Automates Cellulaires.*, 2008.
3. R. Berger. The undecidability of the domino problem. *Memoirs American Mathematical Society*, 66:1966, 1966.
4. Heinz-Dieter Ebbinghaus and Jörg Flum. *Finite Model Theory.* Springer Monographs in Mathematics. Springer, Berlin, 1995.
5. Yuri Gurevich Ergon Börger, Erich Grädel. *The classical decision problem.* Springer-Verlag Telos, 1996.
6. Dora Giammarresi and Antonio Restivo. *Handbook of Formal Languages*, volume 3. Beyond Words, chapter Two-Dimensional Languages. Springer, 1997.

7. Dora Giammarresi, Antonio Restivo, Sebastian Seibert, and Wolfgang Thomas. Monadic Second-Order Logic over Rectangular Pictures and Recognizability by Tiling Systems. *Information and Computation*, 125:32–45, 1996.
8. William Hanf. Model-theoretic methods in the study of elementary logic. In *Symposium in the Theory of Models*, pages 132–145, 1965.
9. David Harel. Recurring Dominoes: Making the Highly Undecidable Highly Understandable. *Annals of Discrete Mathematics*, 24:51–72, 1985.
10. Dietrich Kuske and Markus Lohrey. Logical aspects of Cayley-graphs: the group case. *Annals of Pure and Applied Logic*, 131(1-3):263–286, January 2005.
11. Leonid Libkin. *Elements of Finite Model Theory*. Texts in Theoretical Computer Science, an EATCS Series. Springer Verlag, 2004.
12. J.A. Makowsky. On some conjectures connected with complete sentences. *Fund. Math.*, 81:193–202, 1974.
13. Oliver Matz and Nicole Schweikardt. Expressive power of monadic logics on words, trees, pictures, and graphs. In E. Gradel J. Flum and T. Wilke, editors, *Logic and Automata: History and Perspectives*, Texts in Logic and Games, pages 531–552. Amsterdam University Press, 2007.
14. Francis Oger. Algebraic and model-theoretic properties of tilings. *Theoretical computer science*, 319:103–126, 2004.
15. Bruno Poizat. Une théorie finiement axiomatisable et superstable. *Groupe d'étude des théories stables*, 3(1):1–9, 1980-1982.
16. Raphael M. Robinson. Undecidability and Nonperiodicity for Tilings of the Plane. *Inventiones Math.*, 12, 1971.
17. Thomas Schwentick and Klaus Barthelmann. Local Normal Forms for First-Order Logic with Applications to Games and Automata. *Discrete Mathematics and Theoretical Computer Science*, 3:109–124, 1999.
18. D. Seese. The structure of the models of decidable monadic theories of graphs. *Annals of pure and applied logic*, 53(2):169–195, 1991.
19. Wolfgang Thomas. On logics, tilings, and automata. In Springer Berlin, editor, *Proceedings of the 18th International Colloquium on Automata, Languages and Programming, ICALP*, volume 510 of *Lecture Notes in Computer Science*, pages 441–454, 1991.
20. Wolfgang Thomas. *Handbook of Formal Languages*, volume 3. Beyond Words, chapter Languages, Automata, and Logic. Springer, 1997.
21. Hao Wang. Proving theorems by pattern recognition ii. *Bell system technical journal*, 40:1–41, 1961.
22. B. Weiss. Subshifts of finite type and sofic systems. *Monatshefte für Mathematik*, 77:462–474, 1973.