Descriptive complexity for pictures languages

Etienne Grandjean[1] and Frédéric Olive[2]

1 **Université de Caen Basse-Normandie / ENSICAEN / CNRS**
GREYC – Caen, France
etienne.grandjean@info.unicaen.fr

2 **Aix-Marseille Université / CNRS**
LIF – Marseille, France
frederic.olive@lif.univ-mrs.fr

Abstract

This paper deals with logical characterizations of picture languages of any dimension by syntactical fragments of existential second-order logic. Two classical classes of picture languages are studied:

- the class of *recognizable* picture languages, i.e. projections of languages defined by local constraints (or tilings): it is known as the most robust class extending the class of regular languages to any dimension;
- the class of picture languages recognized on *nondeterministic cellular automata in linear time* : cellular automata are the simplest and most natural model of parallel computation and linear time is the minimal time-bounded class allowing synchronization of nondeterministic cellular automata.

We uniformly generalize to any dimension the characterization by Giammarresi et al. [7] of the class of *recognizable* picture languages in existential monadic second-order logic.

We state several logical characterizations of the class of picture languages recognized in linear time on nondeterministic cellular automata. They are the first machine-independent characterizations of complexity classes of cellular automata.

Our characterizations are essentially deduced from normalization results we prove for first-order and existential second-order logics over pictures. They are obtained in a general and uniform framework that allows to extend them to other 'regular' structures.

1998 ACM Subject Classification F.1.1 Models of Computation, F.1.3 Complexity Measures and Classes, F.4.3 Formal Languages

Keywords and phrases Picture languages; locality and tiling; recognizability; linear time; cellular automata; logical characterizations; second-order logic.

Digital Object Identifier 10.4230/LIPIcs.CSL.2012.274

1 Introduction

One goal of descriptive complexity is to establish logical characterizations of natural classes of problems in finite model theory. Many results in this area involve second-order logic (SO) and its restrictions, monadic second-order logic (MSO) and existential second-order logic (ESO): see *e.g.* [5, 11] for descriptive complexity of formal languages and [5, 9, 8, 11] for the one of complexity classes.

It is important to recall that the complexity class defined by a logic often depends heavily on the underlying class of structures: words, trees, graphs, ordered or unordered structures, etc. E.g., for words, a classical result by Büchi, Elgot and Trahtenbrot [2, 5, 11] states that the class of languages definable in MSO equals the class of regular languages, in short,

Computer Science Logic 2012 (CSL'12).
Editors: Patrick Cégielski, Arnaud Durand; pp. 274–288
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

MSO = REG, whereas the same logic or even its existential second-order fragment EMSO can define some NP-complete problems on graphs, e.g., the 3-colorability problem.

We are interested in descriptive complexity of *picture languages.* A d-picture language is a set of d-pictures, i.e., d-dimensional words (or colored grids). First, notice the following points:

- In a series of papers culminating in [7], Giammarresi et al. proved that a 2-picture language is *recognizable*, i.e. is the projection of a local (that means: *tilable*) 2-picture language, iff it is definable in EMSO. In short: $\mathrm{REC}^2 = \mathrm{EMSO}$.
- In fact, the class REC^2 contains some NP-complete problems. More generally, one observes that, for each dimension $d \geq 1$, the class REC^d of recognizable languages *can be defined as* the class of d-picture languages recognized by nondeterministic d-dimensional *cellular automata* in *constant time*[1].

The present work originates from two questions about word/picture languages:

1. How can we generalize the proof of the above-mentioned theorem of Giammarresi et al. to any dimension? That is, can we establish the equality $\mathrm{REC}^d = \mathrm{EMSO}$ for any $d \geq 1$?
2. Can we obtain logical characterizations of time complexity classes of cellular automata[2]?

In this paper, a d-picture language is a set of d-pictures $p : [1, n]^d \to \Sigma$, for a finite alphabet Σ, i.e., d-dimensional Σ-words[3], and we use two natural representations of a d-picture p as a first-order structure:

- as a *pixel structure:* on the *pixel* domain $[1, n]^d$ where the sets $p^{-1}(s)$, $s \in \Sigma$, are encoded by *unary* relations $(Q_s)_{s\in\Sigma}$ and the underlying d-dimensional grid is encoded by d successor functions (see Definition 2);
- as a *coordinate structure:* on the *coordinate* domain $[1, n]$ where the sets $p^{-1}(s)$ are encoded by *d-ary* relations $(Q_s)_{s\in\Sigma}$; moreover, one uses the natural linear order of the coordinate domain $[1, n]$ and its associate successor function (see Definition 3).

We establish logical characterizations of two classes of d-picture languages, for all dimensions $d \geq 1$:

1. *On pixel structures:* $\mathrm{REC}^d = \mathrm{ESO}(\mathrm{arity}\, 1) = \mathrm{ESO}(\mathrm{var}\, 1) = \mathrm{ESO}(\forall^1, \mathrm{arity}\, 1)$. That means a d-picture language is *recognizable iff* it is definable in monadic ESO (resp. in ESO with 1 first-order variable, or in monadic ESO with 1 universally quantified first-order variable) [4].
2. *On coordinate structures:* $\mathrm{NLIN}^d_{\mathrm{ca}} = \mathrm{ESO}(\mathrm{var}\, d{+}1) = \mathrm{ESO}(\forall^{d+1}, \mathrm{arity}\, d{+}1)$; that means a d-picture language is recognized by a nondeterministic d-dimensional cellular automaton in *linear time* (see, e.g., [3, 14]) *iff* it is definable in ESO with $d + 1$ distinct first-order variables (resp. ESO with second-order variables of arity at most $d + 1$ and a prenex first-order part of prefix $\forall^{d+1}$).

---

[1] That means: for such a picture language L, there is some constant integer c such that each computation stops at instant c, and $p \in L$ iff it has at least one computation that stops *with each cell in an accepting state*: see Sommerhalder et al. [13], which, to our knowledge, was the first paper involving this notion.

[2] This originates from a question that J. Mazoyer asked us in 2000 (personal communication): give a logical characterization of the linear time complexity class of nondeterministic cellular automata.

[3] More generally, the domain of a d-picture is of the "rectangular" form $[1, n_1] \times \ldots \times [1, n_d]$. For simplicity and uniformity of presentation, we have chosen to present our results in the particular case of "square" pictures of domain $[n]^d$. Fortunately, they also hold with the same proofs for general domains $[1, n_1] \times \ldots \times [1, n_d]$.

[4] It is interesting to compare this result with some results by Borchert [1].

Items 1 and 2 proceed from normalization results of, respectively, first-order and ESO logics that we prove over picture languages.

Significance of our results:

1. The *normalization equality* $\mathrm{ESO}(\text{arity}\,1) = \mathrm{ESO}(\forall^1, \text{arity}\,1)$ of Item 1 is a consequence of the fact that, on pixel structures (and, more generally, on structures that consist of bijective functions and unary relations), any first-order formula is equivalent to a boolean combination of *cardinality formulas* of the form: 'there exist k distinct elements x such that $\psi(x)$', where ψ is a quantifier-free formula with only *one* variable. The normalization equality explicitly expresses the local feature of EMSO on pictures – using only *one* first-order variable. Using almost exclusively logical tools, the results of Item 1 can be regarded as an explicitation/simplification (using only *one* first-order variable) and uniformisation of the (more combinatorial) proof and ideas of the main result of Giammarresi et al. [7, 6]; this allows us to generalize it to *any* dimension and, potentially, to other *regular* structures.
2. Intuitively, our characterization $\mathrm{NLIN}^d_{\mathrm{ca}} = \mathrm{ESO}(\forall^{d+1}, \text{arity}\,d+1)$ of Item 2 naturally reflects a *symmetry* property of the time-space diagram of any computation of a *non-deterministic* d-dimensional cellular automaton: informally, the single first-order variable representing time *cannot be distinguished* from any of the d variables that represent the d-dimensional space; in other words, the $d+1$ variables can be permuted without increasing the expressive (or computational) power of the formula. This is the sense of the inclusion $\mathrm{ESO}(\forall^{d+1}, \text{arity}\,d+1) \subseteq \mathrm{NLIN}^d_{\mathrm{ca}}$ whose proof is far from trivial.

2 Preliminaries

All along the paper, we denote by Σ, Γ some finite alphabets and by d a positive integer. For any positive integer n, we set $[n] := \{1, \dots, n\}$. We are interested in sets of pictures of any fixed dimension d.

▶ **Definition 1.** A *d-dimensional picture* or *d-picture* on Σ is a function $p : [n]^d \to \Sigma$ where n is a positive integer. The set $dom(p) = [n]^d$ is called the *domain* of picture p and its elements are called *points*, *pixels* or *cells* of the picture. A set of d-pictures on Σ is called a *d-dimensional language*, or *d-language*, on Σ.

Notice that 1-pictures on Σ are nothing but nonempty words on Σ.

2.1 Pictures as model theoretic structures

Along the paper, we will often describe d-languages as sets of models of logical formulas. To allow this point of view, we must settle on an encoding of d-pictures as model theoretic structures.

For logical aspects of this paper, we refer to the usual definitions and notations in logic and finite model theory (see [5] or [11], for instance). A *signature* (or *vocabulary*) σ is a finite set of relation and function symbols each of which has a fixed arity. A (finite) *structure* S of vocabulary σ, or σ-structure, consists of a finite domain D of cardinality $n \geq 1$, and, for any symbol $s \in \sigma$, an interpretation of s over D, often denoted by s for simplicity. The tuple of the interpretations of the σ-symbols over D is called the *interpretation* of σ over D and, when no confusion results, it is also denoted σ. We will often deal with *tuples* of objects. We denote them by bold letters.

Let us define the two natural representations of a picture as a logical structure.

▶ **Definition 2.** Given $p : [n]^d \to \Sigma$, we denote by *pixel*$^d(p)$ the structure

$$\mathsf{pixel}^d(p) = ([n]^d, (Q_s)_{s\in\Sigma}, (\mathrm{succ}_i)_{i\in[d]}, (\min_i)_{i\in[d]}, (\max_i)_{i\in[d]}).$$

Here:

- succ_j is the *(cyclic) successor function* according to the j^{th} dimension of $[n]^d$, mapping each $(a_1, \ldots, a_d) \in [n]^d$ on $(a_1^{(j)}, \ldots, a_d^{(j)}) \in [n]^d$, where we set : $a_j^{(j)} = a_j + 1$ if $a_j < n$, and $a_j^{(j)} = 1$ otherwise; $a_i^{(j)} = a_i$ for each $i \neq j$; in other words, for $a \in [n]^d$, $\mathrm{succ}_j(a)$ is the d-tuple $a^{(j)}$ obtained from a by 'increasing' its j^{th} component according to the cyclic successor on $[n]$;
- the $\min_i$'s, $\max_i$'s and Q_s's are the following unary (monadic) relations: $\min_i = \{a \in [n]^d : a_i = 1\}$; $\max_i = \{a \in [n]^d : a_i = n\}$; $Q_s = \{a \in [n]^d : p(a) = s\}$.

▶ **Definition 3.** Given $p : [n]^d \to \Sigma$, we denote by *coord*$^d(p)$ the structure

$$\mathsf{coord}^d(p) = \langle [n], (Q_s)_{s\in\Sigma}, <, \mathrm{succ}, \min, \max\rangle. \tag{1}$$

Here:

- each Q_s is the d-ary relation that is the set of cells of p labelled by an s, in other words: $Q_s = \{a \in [n]^d : p(a) = s\}$;
- $<$, min, max are relations of respective arities 2, 1, 1, that are respectively the sets $\{(i, j) : 1 \leq i < j \leq n\}$, $\{1\}$ and $\{n\}$;
- succ is the cyclic successor, that is: $\mathrm{succ}(i) = i + 1$ for $i < n$ and $\mathrm{succ}(n) = 1$.

For a d-language L, we set *pixel*$^d(L) = \{\mathsf{pixel}^d(p) : p \in L\}$ and *coord*$^d(L) = \{\mathsf{coord}^d(p) : p \in L\}$.

▶ Remark. Several details are irrelevant in Definitions 2 and 3, i.e. our results still hold for several variants, in particular:

- in Definition 3, the fact that the linear order $<$ and the equality $=$ are allowed or not and the fact that min, max are represented by individual constants or unary relations;
- in both definitions, the fact that the successor function(s) is/are cyclic or not and is/are completed or not by predecessor(s) function(s).

At the opposite, it is essential that, in both definitions,

- the successor(s) is/are represented by *function(s)* and not by (binary) relation(s),
- the min, max are explicitly represented.

2.2 Logics under consideration

All formulas considered hereafter belong to *relational Existential Second-Order logic*. Given a signature σ, indifferently made of relation and function symbols, a relational existential second-order formula of signature σ has the shape $\Phi \equiv \exists \mathbf{R} \varphi(\sigma, \mathbf{R})$, where $\mathbf{R} = (R_1, \ldots, R_k)$ is a tuple of relation symbols and φ is a first-order formula of signature $\sigma \cup \{\mathbf{R}\}$. We denote by ESO^σ the class thus defined. We will often omit to mention σ for considerations on these logics that do not depend on the signature. Hence, ESO stands for the class of all formulas belonging to ESO^σ for some σ.

We will pay great attention to several variants of ESO. In particular, we will distinguish formulas of type $\Phi \equiv \exists \mathbf{R} \varphi(\sigma, \mathbf{R})$ according to: the number of distinct first-order variables involved in φ, the arity of the second-order symbols $R \in \mathbf{R}$, and the quantifier prefix of some prenex form of φ.

With the logic $\mathrm{ESO}^\sigma(\forall^d, \mathrm{arity}\,\ell)$, we control these three parameters: it is made of formulas of which first-order part is prenex with a universal quantifier prefix of length d, and where existentially quantified relation symbols are of arity at most ℓ. In other words, $\mathrm{ESO}^\sigma(\forall^d, \mathrm{arity}\,\ell)$ collects formulas of shape $\exists\mathbf{R}\forall\mathbf{x}\theta(\sigma, \mathbf{R}, \mathbf{x})$ where θ is quantifier free, $\mathbf{x}$ is a d-tuple of first-order variables, and $\mathbf{R}$ is a tuple of relation symbols of arity at most ℓ. Relaxing some constraints of the above definition, we set:

$$\mathrm{ESO}^\sigma(\forall^d) = \bigcup_{\ell>0} \mathrm{ESO}^\sigma(\forall^d, \mathrm{arity}\,\ell) \text{ and } \mathrm{ESO}^\sigma(\mathrm{arity}\,\ell) = \bigcup_{d>0} \mathrm{ESO}^\sigma(\forall^d, \mathrm{arity}\,\ell).$$

Finally, we write $\mathrm{ESO}^\sigma(\mathrm{var}\,d)$ for the class of formulas that involve at most d first-order variables, thus focusing on the sole number of distinct first-order variables (possibly quantified several times).

3 A logical characterization of recognizable picture languages

In this section, we define the class of *local* (resp. *recognizable*) picture languages and establish the logical characterizations of the class of recognizable picture languages. In order to define a notion of locality based on sub-pictures we need to mark the border of each picture.

▶ **Definition 4.** By $\Gamma^\sharp$ we denote the alphabet $\Gamma \cup \{\sharp\}$ where $\sharp$ is a special symbol not in Γ. Let p be any d-picture of domain $[n]^d$ on Γ. The *bordered d-picture* of p, denoted by $p^\sharp$, is the function $p^\sharp : [0, n+1]^d \to \Gamma^\sharp$ defined by $p^\sharp(a) = p(a)$ if $a \in dom(p)$; $p^\sharp(a) = \sharp$ otherwise. Here, 'otherwise' means that a is on the border of $p^\sharp$, i.e. some component a_i of a is 0 or $n+1$.

Let us now define our notion of *local picture language* or *tilings language*. It is based on some sets of allowed patterns (called tiles) of the bordered pictures and is a simple generalization to any dimension of the notion of *hv-local* 2-dimensional picture language of [10] (see also [6]).

▶ **Definition 5.**
1. Given a d-picture p and an integer $j \in [d]$, two cells $a = (a_i)_{i\in[d]}$ and $b = (b_i)_{i\in[d]}$ of p are *j-adjacent* if they have the same coordinates, except the j^{th} one for which $|a_j - b_j| = 1$.
2. A *tile* for a d-language L on Γ is a pair in $(\Gamma^\sharp)^2$.
3. A picture p *is j-tiled* by a set of tiles $\Delta \subseteq (\Gamma^\sharp)^2$ if for any two j-adjacent points $a, b \in dom(p^\sharp)$: $(p^\sharp(a), p^\sharp(b)) \in \Delta$.
4. Given d sets of tiles $\Delta_1, \ldots, \Delta_d \subseteq (\Gamma^\sharp)^2$, a d-picture p *is tiled by* $(\Delta_1,\ldots,\Delta_d)$ if p is j-tiled by Δ_j for each $j \in [d]$.
5. We denote by $L(\Delta_1,\ldots,\Delta_d)$ the set of d-pictures on Γ that are tiled by $(\Delta_1,\ldots,\Delta_d)$.
6. A d-language L on Γ is *local* if there exist $\Delta_1, \ldots, \Delta_d \subseteq (\Gamma^\sharp)^2$ such that $L = L(\Delta_1,\ldots,\Delta_d)$. We then say that L is $(\Delta_1, \ldots, \Delta_d)$*-local*, or $(\Delta_1, \ldots, \Delta_d)$*-tiled*.

▶ Remark. Our notion of *locality* (that generalizes the one of [10] to any dimension) is more restrictive than the one given by Giammarresi and al. [7]. At the opposite, the *locality* notion defined by Borchert [1] is the most general one: its is defined by the presence/absence or some patterns/sub-pictures of *any size* in the picture, and, as he proved, his *locality* is equivalent to definability by some universally quantified *one-variable* first-order sentence using non cyclic *successor functions* and *minimal* and *maximal* predicates. Fortunately, the notion of *recognizability* as defined below, is a *robust* notion that remains equivalent to the one defined by either one of the locality notions of [1] and [7].

▶ **Definition 6.** A d-language L on Σ is *recognizable* if it is the projection (i.e. homomorphic image) of a *local* d-language over an alphabet Γ. It means there exist a surjective function $\pi : \Gamma \to \Sigma$ and a local d-language L_{loc} on Γ such that $L = \{\pi \circ p : \ p \in L_{\text{loc}}\}$. Because of the last item of Definition 5, one can also write: L is recognizable if there exist a surjective function $\pi : \Gamma \to \Sigma$ and d sets $\Delta_1, \ldots, \Delta_d \subseteq (\Gamma^\sharp)^2$ such that $L = \{\pi \circ p : \ p \in L(\Delta_1, \ldots, \Delta_d)\}$. We write REC^d for the class of recognizable d-languages.

A characterization of recognizable languages of dimension 2 by a fragment of existential monadic second-order logic was proved by Giammarresi et al. [7]. They established:

▶ **Theorem 7** ([7])**.** *For any 2-language L: $L \in \text{REC}^2 \Leftrightarrow \mathsf{pixel}^2(L) \in \text{ESO}(\text{arity}\,1)$.*

In this section, we come back to this result. We simplify its proof, refine the logic it involves, and generalize its scope to any dimension.

▶ **Theorem 8.** *For any $d > 0$ and any d-language L, the following assertions are equivalent:*

1. $L \in \text{REC}^d$;
2. $\mathsf{pixel}^d(L) \in \text{ESO}(\forall^1, \text{arity}\,1)$;
3. $\mathsf{pixel}^d(L) \in \text{ESO}(\text{arity}\,1)$.

Theorem 8 is a straightforward consequence of Propositions 9 and 11 below.

▶ **Proposition 9.** *For any $d > 0$ and any d-language L on Σ: $L \in \text{REC}^d \Leftrightarrow \mathsf{pixel}^d(L) \in \text{ESO}(\forall^1, \text{arity}\,1)$.*

Sketch of proof. $\boxed{\Rightarrow}$ A picture belongs to L if there exists a tiling of its domain whose projection coincides with its content. In the logic involved in the proposition, the 'arity 1' corresponds to formulating the existence of the tiling, while the $\forall^1$ is the syntactic resource needed to express that the tiling behaves as expected. Let us detail these considerations.

By Definition 5, there exist an alphabet Γ (which can be assumed disjoint from Σ), a surjective function $\pi : \Gamma \to \Sigma$ and d subsets $\Delta_1, \ldots, \Delta_d \subseteq (\Gamma^\sharp)^2$ such that L is the set $\{\pi \circ p' : p' \in L(\Delta_1, \ldots, \Delta_d)\}$.

The belonging of a picture $p' : [n]^d \to \Gamma$ to $L(\Delta_1, \ldots, \Delta_d)$ is easily expressed on $\mathsf{pixel}^d(p') = \langle [n]^d, (Q_s)_{s\in\Gamma}, \ldots \rangle$ with a first-order formula which asserts, for each dimension $i \in [d]$, that for any pixel x of p', the couple $(x, \text{succ}_i(x))$ can be tiled with some element of Δ_i. Because it deals with each cell x separately, this formula has the form $\forall x \Psi(x, (Q_s)_{s\in\Gamma})$, where Ψ is quantifier-free.

Now, a picture $p : [n]^d \to \Sigma$ belongs to L iff it results from a π-renaming of a picture $p' \in L(\Delta_1, \ldots, \Delta_d)$. It means there exists a Γ-labeling of p (that is, a tuple $(Q_s)_{s\in\Gamma}$ of subsets of $[n]^d$) corresponding to a picture of $L(\Delta_1, \ldots, \Delta_d)$ (i.e. fulfilling $\forall x \Psi(x, (Q_s)_{s\in\Gamma})$) and from which the actual Σ-labeling of p (that is, the subsets $(Q_s)_{s\in\Sigma}$) is obtained *via* π (easily expressed by a formula of the form $\forall x \Psi'(x, (Q_s)_{s\in\Sigma}, (Q_s)_{s\in\Gamma})$).

Finally, the formula $(\exists Q_s)_{s\in\Gamma} \forall x : \Psi \wedge \Psi'$ conveys the desired property and fits the required form.

$\boxed{\Leftarrow}$ In order to prove the converse implication, it is convenient to first normalize the sentences of $\text{ESO}(\forall^1, \text{arity}\,1)$. This is the role of the technical result below, which asserts that on pixel encodings, each such sentence can be rewritten in a very local form where the first-order part alludes only pairs of adjacent pixels of the bordered picture. We state it without proof (see also [1]):

▶ **Fact 10.** *On pixel structures, any $\varphi \in \mathrm{ESO}(\forall^1, \mathrm{arity}\, 1)$ is equivalent to a sentence of the form:*

$$\exists \mathbf{U} \forall x \bigwedge_{i\in[d]} \left\{ \begin{array}{ccc} min_i(x) & \to & \mathsf{m}_i(x) \quad \wedge \\ max_i(x) & \to & \mathsf{M}_i(x) \quad \wedge \\ \neg max_i(x) & \to & \Psi_i(x) \end{array} \right\}. \tag{2}$$

Here, $\mathbf{U}$ is a list of monadic relation variables and m_i, M_i, Ψ_i are quantifier-free formulas such that

- *atoms of m_i and M_i have all the form $Q(x)$;*
- *atoms of Ψ_i have all the form $Q(x)$ or $Q(succ_i(x))$,*

where, in both cases, $Q \in \{(Q_s)_{s\in\Sigma}, \mathbf{U}\}$.

Now, consider L such that $\mathsf{pixel}^d(L) \in \mathrm{ESO}(\forall^1, \mathrm{arity}\, 1)$. Fact 10 ensures that $\mathsf{pixel}^d(L)$ is characterized by a sentence of the form (2) above. We have to prove that L is the projection of some local d-language L_{loc} on some alphabet Γ, that is a $(\Delta_1,\ldots,\Delta_d)$-tiled language for some $\Delta_1,\ldots,\Delta_d \subseteq \Gamma^2$. Let $U_1,\ldots,U_k$ denote the list of (distinct) elements of the set $\{(Q_s)_{s\in\Sigma}, \mathbf{U}\}$ of unary relation symbols of φ so that the first ones $U_1,\ldots,U_m$ are the Q_s's (here, $\min_i$ and $\max_i$ symbols are excluded). The trick is to put each subformula $\mathsf{m}_i(x)$, $\mathsf{M}_i(x)$ and $\Psi_i(x)$ of φ into its *complete disjunctive normal form* with respect to $U_1,\ldots,U_k$. Typically, each subformula $\Psi_i(x)$ whose atoms are of the form $U_j(x)$ or $U_j(\mathrm{succ}_i(x))$, for some $j \in [k]$, is transformed into the following 'complete disjunctive normal form':

$$\bigvee_{(\epsilon,\epsilon')\in\Delta_i} \left(\bigwedge_{j\in[k]} \epsilon_j U_j(x) \wedge \bigwedge_{j\in[k]} \epsilon'_j U_j(\mathrm{succ}_i(x)) \right).$$

Here, the following conventions are adopted:

- $\epsilon = (\epsilon_1,\ldots,\epsilon_k) \in \{0,1\}^k$ and similarly for ϵ';
- for any atom α and any bit $\epsilon_j \in \{0,1\}$, $\epsilon_j\alpha$ denotes the literal α if $\epsilon_j = 1$, the literal $\neg\alpha$ otherwise.

For $\epsilon \in \{0,1\}^k$, we denote by $\Theta_\epsilon(x)$ the 'complete conjunction' $\bigwedge_{j\in[k]} \epsilon_j U_j(x)$. Intuitively, $\Theta_\epsilon(x)$ is a complete description of x and the set $\Gamma = \bigcup_{i\in[m]} \{0^{i-1}10^{m-i}\} \times \{0,1\}^{k-m}$ is the set of possible colours (remember that the Q_s's that are the U_j's for $j \in [m]$ form a partition of the domain). The complete disjunctive normal form of $\Psi_i(x)$ can be written into the suggestive form: $\bigvee_{(\epsilon,\epsilon')\in\Delta_i} (\Theta_\epsilon(x) \wedge \Theta_{\epsilon'}(\mathrm{succ}_i(x)))$.

If each subformula $\mathsf{m}_i(x)$ and $\mathsf{M}_i(x)$ of φ is similarly put into complete disjunctive normal form, that is $\bigvee_{(\sharp,\epsilon)\in\Delta_i} \Theta_\epsilon(x)$ and $\bigvee_{(\epsilon,\sharp)\in\Delta_i} \Theta_\epsilon(x)$, respectively (there is no ambiguity in our implicit definition of the Δ_i's, since $\sharp \notin \Gamma$), then the whole formula φ becomes the following equivalent formula:

$$\varphi' = \exists \mathbf{U} \forall x \bigwedge\nolimits_{i\in[d]} \left\{ \begin{array}{cccr} \min_i(x) & \to & \bigvee_{(\sharp,\epsilon)\in\Delta_i} \Theta_\epsilon(x) & \wedge \\ \max_i(x) & \to & \bigvee_{(\epsilon,\sharp)\in\Delta_i} \Theta_\epsilon(x) & \wedge \\ \neg\max_i(x) & \to & \bigvee_{(\epsilon,\epsilon')\in\Delta_i} (\Theta_\epsilon(x) \wedge \Theta_{\epsilon'}(\mathrm{succ}_i(x))) & \end{array} \right\}$$

Finally, let L_{loc} denote the d-language over Γ defined by the first-order sentence φ_{loc} obtained by replacing each Θ_ϵ by the new unary relation symbol Q_ϵ in the first-order part of

φ'. In other words, $\mathsf{pixel}^d(L_{\text{loc}})$ is defined by the following first-order sentence:

$$\varphi_{\text{loc}} = \forall x \bigwedge_{i\in[d]} \left\{ \begin{array}{lcll} \min_i(x) & \to & \bigvee_{(\sharp,\epsilon)\in\Delta_i} Q_\epsilon(x) & \wedge \\ \max_i(x) & \to & \bigvee_{(\epsilon,\sharp)\in\Delta_i} Q_\epsilon(x) & \wedge \\ \neg\max_i(x) & \to & \bigvee_{(\epsilon,\epsilon')\in\Delta_i} (\, Q_\epsilon(x) \wedge Q_{\epsilon'}(\text{succ}_i(x))\,) & \end{array} \right\}$$

Hence, $L_{\text{loc}} = L(\Delta_1, \ldots, \Delta_d)$. That is, L_{loc} is indeed local and the corresponding sets of tiles are the Δ_i's of the previous formula. It is now easy to see that our initial d-language L is the projection of the local L_{loc} by the projection $\pi : \Gamma \to \Sigma$ defined as follows: $\pi(\epsilon) = s$ iff $\epsilon_i = 1$ for $i \in [m]$ and U_i is Q_s. This completes the proof. ◀

▶ **Proposition 11.** ESO(arity 1) $\subseteq$ ESO($\forall^1$, arity 1) *on pixel structures, for any* $d > 0$*.*

Proof. In a pixel structure, each function symbol is interpreted as a *bijective* function (namely, a cyclic successor). It has been proved in [4, 12] that any first-order formula on such a structure can be rewritten as a so-called *cardinality formula*, that is as a boolean combination of sentences of the form $\psi^{\geq k} = \exists^{\geq k} x\psi(x)$ (for $k \geq 1$) where $\psi(x)$ is a quantifier-free formula (using the 'bijective' function symbols f and their inverses f^{-1}) with the single variable x and where the quantifier $\exists^{\geq k}x$ means 'there exist at least k elements $x\ldots$'. Therefore, it is easily seen that proving the proposition amounts to showing that each sentence of the form $\psi^{\geq k}$ or $\neg\psi^{\geq k}$ can be translated in ESO($\forall^1$, arity 1) on pixel structures.

This is done as follows: for a given sentence $\exists^{\geq k} x\psi(x)$, we introduce new unary relations $U^{=1}, U^{=2}, \ldots, U^{=k-1}$ and $U^{\geq k}$, with the intended meaning:

A pixel $a \in [n]^d$ *belongs to* $U^{=j}$ *(resp.* $U^{\geq k}$*) iff there are exactly* j *(resp. at least* k*) pixels* $b \in [n]^d$ *lexicographically smaller than or equal to* a *such that* $\mathsf{pixel}^d(p) \models \psi(b)$*.*

Then we have to compel these relation symbols to fit their expected interpretations, by means of a universal first-order formula with a single variable. First, we demand the relations to form a partition of the domain:

(1) $\displaystyle\bigwedge_{i<j<k} \big(\neg U^{=i}(x) \vee \neg U^{=j}(x)\big) \wedge \bigwedge_{i<k} \big(\neg U^{=i}(x) \vee \neg U^{\geq k}(x)\big)\,.$

Let us temporarily denote by $\leq_{\text{lex}}$ the lexicographic order on $[n]^d$ inherited from the natural order on $[n]$, and by succ_{lex}, $\min_{\text{lex}}$, $\max_{\text{lex}}$ its associated successor function and unary relations corresponding to extremal elements. Then the sets described above can be defined inductively by the conjunction of the following six formulas:

(2) $(\min_{\text{lex}}(x) \wedge \neg\psi(x)) \to U^{=0}(x)$

(3) $(\min_{\text{lex}}(x) \wedge \psi(x)) \to U^{=1}(x)$

(4) $\displaystyle\bigwedge_{i<k} \big((\neg\max_{\text{lex}}(x) \wedge U^{=i}(x) \wedge \neg\psi(\text{succ}_{\text{lex}}(x))\big) \to U^{=i}(\text{succ}_{\text{lex}}(x))$

(5) $\displaystyle\bigwedge_{i<k-1} \big((\neg\max_{\text{lex}}(x) \wedge U^{=i}(x) \wedge \psi(\text{succ}_{\text{lex}}(x))\big) \to U^{=i+1}(\text{succ}_{\text{lex}}(x))$

(6) $\big((\neg\max_{\text{lex}}(x) \wedge U^{=k-1}(x) \wedge \psi(\text{succ}_{\text{lex}}(x))\big) \to U^{\geq k}(\text{succ}_{\text{lex}}(x))$

(7) $\big((\neg\max_{\text{lex}}(x) \wedge U^{\geq k}(x)\big) \to U^{\geq k}(\text{succ}_{\text{lex}}(x))$
Hence, under the hypothesis (1) $\wedge \ldots \wedge$ (7), the sentences $\psi^{\geq k}$ and $\neg\psi^{\geq k}$ are equivalent, respectively, to $\forall x(\max_{\text{lex}}(x) \to U^{\geq k}(x))$ and $\forall x(\max_{\text{lex}}(x) \to \neg U^{\geq k}(x))$.

To complete the proof, it remains to get rid of symbols succ_{lex}, $\min_{\text{lex}}$ and $\max_{\text{lex}}$ that are not allowed in our language. It is done by referring to these symbols implicitly rather than explicitly. For instance, since $\text{succ}_{\text{lex}}(x) = \text{succ}_i\text{succ}_{i+1}\ldots\text{succ}_d(x)$ for the smallest $i \in [d]$ such that $\bigwedge_{j>i} \max_j(x)$, each formula φ involving $\text{succ}_{\text{lex}}(x)$ actually corresponds to the conjunction:

$$\bigwedge_{i\in[d]} \left((\neg\max_i(x) \wedge \bigwedge_{i<j\leq d} \max_j(x)) \to \varphi_i \right),$$

where φ_i is obtained from φ by the substitution $\text{succ}_{\text{lex}}(x) \rightsquigarrow \text{succ}_i \ldots \text{succ}_d(x)$. Similar arguments allow to get rid of $\min_{\text{lex}}$ and $\max_{\text{lex}}$. ◀

▶ Remark. In this proof, two crucial features of a structure of type $\mathsf{pixel}^d(p)$ are involved: its 'bijective' nature, that allows to rewrite first-order formulas as cardinality formulas; the regularity of its predefined arithmetics (the functions succ_i defined on each dimension), that endows $\mathsf{pixel}^d(p)$ with a grid structure: it enables us to implicitly define an order on the whole domain $dom(p)$ by means of first-order formulas with a single variable, which in turn allows to express cardinality formulas by 'cumulative' arguments, *via* the sets $U^{=i}$. Proposition 11 straightforwardly generalizes to all structures – and there are a lot – that fulfill these two properties.

4 A logical characterization of $\mathrm{NLIN}_{\mathsf{ca}}$

The second main concept studied in this paper is the classical notion of linear time complexity on nondeterministic cellular automata of any dimension (e.g., see [3, 14]). For simplicity of notation, we only present here the notion of *one-way d-dimensional cellular automaton*, instead of the more usual notion of *two-way d-dimensional cellular automaton*, but it is a folklore result that in the nondeterministic case, the two linear-time complexity classes so defined coincide (see [14]).

There are some technicalities in our definition of the transition function of a cellular automaton here below. This is due to the need to distinguish the different possible positions of the pixels of a picture w.r.t. its border: the one-way neighborhood of a cell $\mathbf{x} = (x_1, \ldots, x_d)$, that is the set of cells $\mathbf{y} = (y_1, \ldots, y_d)$ such that $0 \leq y_i - x_i \leq 1$ for each $i \in [d]$, may be *incomplete* according to the *position* of the cell $\mathbf{x}$ w.r.t. the border of the picture.

▶ **Definition 12.** A pixel $\mathbf{x} = (x_1, \ldots, x_d) \in [n]^d$ is *in position* $a = (a_1, \ldots, a_d) \in \{0,1\}^d$ in a picture $p : [n]^d \to \Gamma$ or in the domain $[n]^d$ if, for all $i \in [d]$, we have $a_i = 0$ if $x_i = n$ and $a_i = 1$ if $x_i < n$.

We are going to define the transition function on a pixel x of a picture p according to some 'neighborhood' (sub-picture) denoted $p_{a,x}$ whose domain, denoted by Dom_a, depends on the position a of the pixel in the picture.

▶ **Definition 13.** For each $a = (a_1, \ldots, a_d) \in \{0,1\}^d$, let us define the *a-domain* as $\mathsf{Dom}_a = [0, a_1] \times \cdots \times [0, a_d]$.

The *a-neighborhood* of some pixel $x \in [n]^d$ in position a in a picture $p : [n]^d \to \Gamma$ is the function $p_{a,x} : \mathsf{Dom}_a \to \Gamma$ defined as $p_{a,x}(b) = p(x + b)$, where $x + b$ denotes the sum of the vectors x and b.

We denote by $\textsf{\textit{neighb}}_a(\Gamma)$ the set of all possible a-neighborhoods on an alphabet Γ, that is the set of functions $\nu : \mathsf{Dom}_a \to \Gamma$.

We are now ready to define the 'transition function' of a cellular automaton:

▶ **Definition 14.** A *one-way nondeterministic d-dimensional cellular automaton* (d-*automaton*, for short) over an alphabet Σ is a tuple $\mathcal{A} = (\Sigma, \Gamma, \delta, F)$, where

- the finite alphabet Γ called the *set of states* of $\mathcal{A}$ includes the *input alphabet* Σ and the set F of *accepting states*: $\Sigma, F \subseteq \Gamma$;
- δ is the (nondeterministic) *transition function* of $\mathcal{A}$: it is a family of a-*transition functions* $\delta = (\delta_a)_{a\in\{0,1\}^d}$ of the form $\delta_a : \mathsf{neighb}_a(\Gamma) \to \mathcal{P}(\Gamma)$.

Let us now define a computation.

▶ **Definition 15.** Let $\mathcal{A} = (\Sigma, \Gamma, \delta, F)$ be a d-automaton and $p, p' : [n]^d \to \Gamma$ be two d-pictures on Γ. We say that p' is a *successor* of p for $\mathcal{A}$, denoted by $p' \in \mathcal{A}(p)$, if, for each position $a \in \{0,1\}^d$ and each point x of position a in $[n]^d$, $p'(x) \in \delta_a(p_{a,x})$. The set of j^{th}*-successors* of p for $\mathcal{A}$, denoted by $\mathcal{A}^j(p)$, is defined inductively:

$$\mathcal{A}^0(p) = \{p\} \text{ and, for } j \geq 0,\ \mathcal{A}^{j+1}(p) = \bigcup_{p'\in\mathcal{A}^j(p)} \mathcal{A}(p').$$

▶ **Definition 16.** A *computation* of a d-automaton $\mathcal{A}$ on an input d-picture p is a sequence $p_1, p_2, p_3, \ldots$ of d-pictures such that $p_1 = p$ and $p_{i+1} \in \mathcal{A}(p_i)$ for each i. A computation is *accepting* if it is finite – it has the form $p_1, p_2, \ldots, p_k$ for some k – and the cell of minimal coordinates, $1^d = (1, \ldots, 1)$, of its last configuration is in an accepting state: $p_k(1^d) \in F$.

▶ **Definition 17.** Let $\mathcal{A} = (\Sigma, \Gamma, \delta, F)$ be a d-automaton and let $T : \mathbb{N} \to \mathbb{N}$ be such that $T(n) > n$. A d-picture p on Σ is *accepted by $\mathcal{A}$ in time $T(n)$* if $\mathcal{A}$ admits an accepting computation of length $T(n)$ on p. That means, there exists a computation $p = p_1, p_2, \ldots, p_{T(n)} = \mathcal{A}^{T(n)-1}(p)$ of $\mathcal{A}$ on p such that $p_{T(n)}(1^d) \in F$. A d-language L on Σ is *accepted*, or *recognized*, by $\mathcal{A}$ *in time* $T(n)$ if it is the set of d-pictures accepted by $\mathcal{A}$ *in time* $T(n)$. That is $L = \{p : \exists p' \in \mathcal{A}^{T(n)-1}(p) \text{ such that } p'(1^d) \in F\}$.

If $T(n) = cn + c'$, for some integers c, c', then L is said to be *recognized in linear time* and we write $L \in \mathrm{NLIN}^d_{ca}$.

▶ Remark. The nondeterministic linear time class $\mathrm{NLIN}^d_{\mathrm{ca}}$ is very robust, i.e. is not modified by many changes in the definition of the automaton or in its time bound. In particular, it is a folklore result that the constants c, c' defining the bound $T(n) = cn + c'$ can be fixed arbitrarily, provided $T(n) > n$. For example, the class $\mathrm{NLIN}^d_{\mathrm{ca}}$ does not change if we take the *minimal* time $T(n) = n + 1$, called *real time*, i.e. the minimal time for that the information of any pixel of p can be communicated to the reference pixel, 1^d (see [14]).

Here is the second main result of this paper.

▶ **Theorem 18.** *For any $d > 0$ and any d-language L, the following assertions are equivalent:*

1. $L \in \mathrm{NLIN}^d_{ca}$*;*
2. $\textsf{\textit{coord}}^d(L) \in \mathrm{ESO}(\forall^{d+1}, \mathrm{arity}\, d + 1)$*;*
3. $\textsf{\textit{coord}}^d(L) \in \mathrm{ESO}(\mathrm{var}\, d + 1)$.

This theorem is a straightforward consequence of Propositions 19 and 20 below.

▶ **Proposition 19.** *For any $d > 0$ and any d-language L,*

$$L \in \mathrm{NLIN}^d_{ca} \Leftrightarrow \textsf{\textit{coord}}^d(L) \in \mathrm{ESO}(\forall^{d+1}, \mathrm{arity}\,(d + 1)).$$

Sketch of proof. $\boxed{\Rightarrow}$

Let $L \in \mathrm{NLIN}^d_{\mathrm{ca}}$. By the Remark preceding Theorem 18, L is recognized by a d-automaton $\mathcal{A} = (\Sigma, \Gamma, \delta, F)$ in *real time*, i.e. in time $n+1$. The sentence in $\mathrm{ESO}(\mathrm{var}\, d+1)$ that we are going to construct is of the form $\exists (R_s)_{s\in\Gamma}\ \forall \mathbf{x} \forall t\ \psi(\mathbf{x}, t)$, where:

- ψ is a quantifier-free formula that uses a list of exactly $d+1$ first-order variables $\mathbf{x} = (x_1, \ldots, x_d)$ and t. Intuitively, the d first ones represent the coordinates of any point in $dom(p) = [n]^d$ and the last one represents any of the first n instants $t \in [n]$ of the computation (the last instant $n+1$ is not explicitly represented);

- ψ uses, for each state $s \in \Gamma$, a relation symbol R_s of arity $d+1$. Intuitively, $R_s(a_1, \ldots, a_d, t)$ holds, for any $a = (a_1, \ldots, a_d) \in [n]^d$ and any $t \in [n]$, iff the state of cell a at instant t is s.

- ψ is the conjunction $\psi(\mathbf{x}, t) = \textsc{init}(\mathbf{x}, t) \wedge \textsc{step}(\mathbf{x}, t) \wedge \textsc{end}(\mathbf{x}, t)$ of three formulas whose intuitive meaning is the following.

- $\forall \mathbf{x} \forall t\ \textsc{init}(\mathbf{x}, t)$ describes the first configuration of $\mathcal{A}$, i.e. at initial instant 1, that is the input picture $p_1 = p$;
- $\forall \mathbf{x} \forall t\ \textsc{step}(\mathbf{x}, t)$ describes the computation between the instants t and $t+1$, for $t \in [n-1]$, i.e. describes the $(t+1)^{th}$ configuration p_{t+1} from the t^{th} one p_t, i.e. says $p_{t+1} \in \mathcal{A}(p_t)$;
- $\forall \mathbf{x} \forall t\ \textsc{end}(\mathbf{x}, t)$ expresses that the n^{th} configuration p_n leads to a (last) $(n+1)^{th}$ configuration $p_{n+1} \in \mathcal{A}(p_n)$ which is accepting, i.e. with an accepting state in cell 1^d: $p_{n+1}(1^d) \in F$.

Let us only give explicitly the second formula, $\textsc{step}$, which is the most central one (the last formula, $\textsc{end}$, is similar; the first one, $\textsc{init}$, is easy to construct):

$$\textsc{step}(\mathbf{x}, t) \equiv$$

$$\bigwedge_{a\in\{0,1\}^d} \ \bigwedge_{\nu \in \mathsf{neighb}_a(\Gamma)} \left\{ (\ \neg\max(t) \wedge P_a(\mathbf{x}) \wedge \bigwedge_{b \in \mathsf{Dom}_a} R_{\nu(b)}(\mathbf{x}+b, t)\ ) \to \bigoplus_{s\in\delta_a(\nu)} R_s(\mathbf{x}, \mathrm{succ}(t)) \right\}$$

Here, $\bigoplus$ denotes the exclusive disjunction. Furthermore:

- For $\mathbf{x} \in [n]^d$ and $a = (a_1, \ldots, a_d) \in \{0,1\}^d$, the formula $P_a(\mathbf{x})$ claims that the pixel $\mathbf{x}$ is in position a. Namely: $P_a(\mathbf{x}) \equiv \bigwedge_{i\in[d]} (\neg_i)\max(x_i)$, where $(\neg_i)$ is $\neg$ if $a_i = 1$, and *nothing* otherwise.
- For $b = (b_1, \ldots, b_d) \in \{0,1\}^d$, $\mathbf{x} + b$ abbreviates the tuple of terms $(\theta_1, \ldots, \theta_d)$ where, for each i, the term θ_i is x_i if $b_i = 0$, and $\mathrm{succ}(x_i)$ otherwise.

It is easy to verify that the formula $\forall \mathbf{x}\ \textsc{step}(\mathbf{x}, t)$ means $p_{t+1} \in \mathcal{A}(p_t)$, as claimed.

$\boxed{\Leftarrow}$ Assume $\mathsf{coord}^d(L) \in \mathrm{ESO}(\forall^{d+1}, \mathrm{arity}\, d+1)$. That is, there is some sentence Φ in $\mathrm{ESO}(\forall^{d+1}, \mathrm{arity}\, d+1)$ such that $p \in L \Leftrightarrow \mathsf{coord}^d(p) \models \Phi$. We want to prove $L \in \mathrm{NLIN}^d_{\mathrm{ca}}$, i.e. L is recognized by some d-automaton in linear time. Let us give the main idea of the proof for the simplest case $d = 1$ and a formula $\Phi \in \mathrm{ESO}(\forall^2, \mathrm{arity}\, 2)$ of the form

$$\Phi = \exists R \forall x \forall y \psi(x, y)$$

where R is a binary relation symbol and ψ is a quantifier-free formula where the only atoms in which R occurs, called R atoms, are of the following forms (1-4):

(1) $R(x, y)$; (2) $R(succ(x), y)$; (3) $R(x, succ(y))$; (4) $R(y, x)$.

First, notice that if the only atoms where R occurs are of the forms (1-3), i.e. the variables x, y only appear in this *unique* order in the arguments of R, then formula Φ has a *local* *behaviour*: points (x, y), $(succ(x), y)$ and $(x, succ(y)$ are *neighbours*, i.e. adjacent each other. This allows to construct a 1-automaton (nondeterministic cellular automaton of dimension 1) $\mathcal{A}$ that mimics Φ. Roughly, $\mathcal{A}$ successively guesses 'rows' $R(i, \ldots)$, for $i = 1, 2, \ldots n$, of R, and in the same time, it checks locally the coherence of each instantiation $\psi(i, j)$: more precisely, at instant i, the state of each cell j, $1 \leq j \leq n$, of $\mathcal{A}$ contains both values $R(i, j)$ and $R(i + 1, j)$. So, in case atoms of R are of the forms (1-3), the language L is recognized by such a 1-automaton $\mathcal{A}$ in linear time as claimed.

Now, let us consider the general case where the formula includes all the forms (1-4). Of course, the pixel of the form (4), $R(y, x)$, is not adjacent to pixels of the form (1-3) but is their symmetric (more precisely, is symmetric of $R(x, y)$) with respect to the diagonal $x = y$. The intuitive idea is to cut or to fold the 'picture' R along this diagonal: R is replaced by its two 'half pictures' denoted R_1 and R_2, that are superposed in the half square $x \leq y$ above the diagonal. More precisely, R_1 and R_2 are binary relations whose intuitive meaning is the following: for points (x, y) such that $x \leq y$, one has the equivalence $R_1(x, y) \Leftrightarrow R(x, y)$ and the equivalence $R_2(x, y) \Leftrightarrow R(y, x)$. By this transformation, each pixel $R_2(x, y)$ that represents the original pixel $R(y, x)$ lies at the same point (x, y) as pixel $R_1(x, y)$ that represents pixel $R(x, y)$, for $x \leq y$. The case $y \leq x$ is similar. This solves the problem of vicinity.

More precisely, the sentence $\Phi = \exists R \forall x \forall y \psi(x, y)$ is normalized as follows. Let $\mathsf{coherent}(x, y)$ denote the formula $x = y \rightarrow (R_1(x, y) \leftrightarrow R_2(x, y))$ whose universal closure ensures the coherence of R_1 and R_2 on the common part of R they both represent, that is the diagonal $x = y$. Using R_1 and R_2, it is not difficult to construct a formula

$$\psi'(x, y) = \mathsf{coherent}(x, y) \wedge \begin{pmatrix} x < y \rightarrow \psi_<(x, y) & \wedge \\ x = y \rightarrow \psi_=(x, y) & \wedge \\ x > y \rightarrow \psi_>(x, y) & \end{pmatrix}$$

such that the sentence $\Phi' = \exists R_1 \exists R_2 \forall x \forall y \psi'(x, y)$ in $\mathrm{ESO}(\forall^2, \mathrm{arity}\, 2)$ is equivalent to Φ. Let us describe and justify its precise form and meaning.

Table 1 Replacement of R-atoms by R_1- or R_2-atoms.

case	formula	$R(x, y)$	$R(\mathrm{succ}(x), y)$	$R(x, \mathrm{succ}(y))$	$R(y, x)$
$x < y$	$\psi_<(x, y)$	$R_1(x, y)$	$R_1(\mathrm{succ}(x), y)$	$R_1(x, \mathrm{succ}(y))$	$R_2(x, y)$
$x = y$	$\psi_=(x, y)$	$R_1(x, y)$	$R_2(x, \mathrm{succ}(y))$	$R_1(x, \mathrm{succ}(y))$	$R_1(x, y)$
$x > y$	$\psi_>(x, y)$	$R_2(y, x)$	$R_2(y, \mathrm{succ}(x))$	$R_2(\mathrm{succ}(y), x)$	$R_1(y, x)$

The formulas $\psi_<(x, y)$, $\psi_=(x, y)$ and $\psi_>(x, y)$ are obtained from formula $\psi(x, y)$ by substitution of R-atoms by R_1- or R_2-atoms according to the cases described in Table 1. It is easy to check that each replacement is correct according to its case. For instance, it is justified to replace each atom of the form $R(x, \mathrm{succ}(y))$ in ψ by $R_2(\mathrm{succ}(y), x)$ when $x > y$ (in order to obtain the formula $\psi_>(x, y)$), because when $x > y$ we get $\mathrm{succ}(y) \leq x$ and hence the equivalence $R(x, \mathrm{succ}(y)) \leftrightarrow R_2(\mathrm{succ}(y), x)$ holds, by definition of R_2.

Notice that the variables x, y always occur in this order in each R_1- or R_2-atom of the formulas $\psi_<$ and $\psi_=$ (see Table 1). At the opposite, they always occur in the reverse order y, x in the formula $\psi_>(x, y)$. This is not a problem because, by symmetry, the roles of x and y can be exchanged and the universal closure $\forall x \forall y (x > y \rightarrow \psi_>(x, y))$ is trivially equivalent to $\forall x \forall y (y > x \rightarrow \psi_>(y, x))$. So, the above sentence Φ' – and hence, the original sentence

Φ – is equivalent to the sentence denoted Φ" obtained by replacing in Φ' the subformula $x > y \to \psi_>(x, y)$ by $y > x \to \psi_>(y, x)$. By construction, relation symbols R_1, R_2 only occur in Φ in atoms of the three required 'sorted' forms: $R_i(x, y)$, $R_i(\mathrm{succ}(x), y)$ or $R_i(x, \mathrm{succ}(y))$.

Finally, to be precise, there remain two difficulties so that a 1-automaton can simulate the 'sorted' sentence Φ" in linear time, by the informal algorithm described above:

- the presence of equalities and inequalities in the sentence;
- the forms of the atoms involving input relation symbols.

It is easy to get rid of equalities and inequalities by introducing new binary relation symbols defined and used in a 'sorted' manner too (see (1-3)). Concerning the second point, we can assume, without loss of generality, that the only atoms involving the input relation symbols $(Q_s)_{s\in\Sigma}$ are of the two forms $Q_s(x)$ or $Q_s(y)$. As we do for equalities and inequalities, we can get rid of atoms of the form $Q_s(y)$ by introducing new binary ESO relation symbols : intuitively, they convey each bit $Q_s(a)$ at each point of coordinates $(a, \dots)$ or $(\dots, a)$; those new binary relations are also defined and used in a 'sorted' manner. The fact that all the atoms involving the input are of the form $Q_s(x)$ allows to consider this input in the initial configuration of the computation of the 1-automaton *but in no later configuration* as required. So, the sketch of proof is complete for the case $d = 1$.

For the general case, i.e. for any dimension d, the ideas and the steps of the proof are exactly the same as for $d = 1$ but the notations and details of the proof are much more technical. To give an idea, let us succinctly describe the ESO relations of arity $d+1$ introduced in the main normalization step. Here again, each ESO relation symbol R of the original sentence Φ in $\mathrm{ESO}(\forall^{d+1}, \mathrm{arity}\, d+1)$ is replaced by – or, intuitively, 'divided into' – $(d+1)!$ new ESO relation symbol R_α of the same arity $d+1$, where α is a permutation of the set of indices $[d+1]$. The intended meaning of each relation R_α is the following: for each tuple $(a_1, \dots, a_{d+1}) \in [n]^{d+1}$ such that $a_1 \le a_2 \dots \le a_{d+1}$, the equivalence

$$R_\alpha(a_1, \dots, a_{d+1}) \leftrightarrow R(a_{\alpha(1)}, \dots, a_{\alpha(d+1)})$$

holds. Then, we introduce a partition of the domain $[n]^{d+1}$ into subdomains, similar to the partition of the domain $[n]^2$ described above for $d = 1$ into the diagonal $x = y$ and the two half domains over and under the diagonal $x < y$ and $x > y$. According to the case (i.e. subdomain of the partition), this allows to replace each R atom in Φ by an atom of one of the two following sorted forms, $R_\alpha(\mathbf{x})$ and $R_\alpha(\mathbf{x}^{(i)})$, where $\mathbf{x} = (x_1, \dots, x_{d+1})$, $1 \le i \le d+1$ and $\mathbf{x}^{(i)}$ is the tuple $\mathbf{x}$ where x_i is replaced by $\mathrm{succ}(x_i)$. Finally, the equalities and inequalities are similarly eliminated in the sentence and we normalize it with respect to the input d-ary relations $(Q_s)_{s\in\Sigma}$ by using new ESO relation symbols of arity $d+1$ to convey the input information: in the final sorted sentence all the Q_s atoms are of the unique form $Q_s(x_1, \dots, x_d)$. For such a sorted $\mathrm{ESO}(\forall^{d+1}, \mathrm{arity}\, d+1)$-sentence Φ, it is now easy to construct a d-automaton that generalizes the automaton described above in case $d = 1$, and checks in linear time whether $\mathsf{coord}^d(p) \models \Phi$. ◀

▶ **Proposition 20.** *For any $d > 0$, $\mathrm{ESO}(\mathrm{var}\, d) \subseteq \mathrm{ESO}(\forall^d, \mathrm{arity}\, d)$ on coordinate structures.*

Sketch of proof. We first prove a kind of Skolemization of $\mathrm{ESO}(\mathrm{var}\, d)$-formulas, thus providing a first normalization of these formulas, in which the first-order part is *universal* and includes *the same number of first-order variables* as the initial formula. To illustrate the procedure that performs this preliminary normalisation, let us run it on a very simple first-order formula with *two* variables: $\varphi \equiv \exists x\, (\forall y U(x, y) \vee \exists y D(x, y)\,)$. We introduce three new relation symbols R_1, R_2, R_3 corresponding to the quantified subformulas of φ.

$$\begin{array}{lcll} \text{def}_1(R_1) \equiv \forall x: & R_1(x) & \leftrightarrow & \forall y\, U(x,y) \\ \text{def}_2(R_2) \equiv \forall x: & R_2(x) & \leftrightarrow & \exists y\, D(x,y) \\ \text{def}_3(R_3) \equiv \forall x: & R_3(x) & \leftrightarrow & R_1(x) \vee R_2(x) \end{array}$$

Hence our initial formula can be rewritten:

$$\exists R_1, R_2, R_3 : \left(\bigwedge_{1 \le i \le 3} \text{def}_i(R_i) \right) \wedge \exists x\ R_3(x). \tag{3}$$

It is easily seen that (3) can be written as a conjunction of prenex formulas, each of which involves no more than two variables and has a quantifier prefix of the shape $\forall x \forall y$ or $\forall x \exists y$ (we include in this later form the subformula $\exists x\ R_3(x)$). All in all, φ is equivalent to a formula of the form:

$$\exists R_1, R_2, R_3 : \forall x \forall y \psi(x, y, \mathbf{R}) \wedge \forall x \exists y \theta(x, y, \mathbf{R}), \tag{4}$$

where ψ and θ are quantifier-free. In order to put this conjunction under prenex form without adding a new first-order variable, we have to "replace" the existential quantifier by a universal one. (Afterward φ, as a conjunction of formulas of prefix $\forall x, y$, could be written under the requisite shape.) To proceed, we get use of the arithmetics embedded in coordinate structures. It allows to defining a binary relation W with intended meaning: $\mathrm{W}(x, y)$ iff there exists $z \le y$ such that $\theta(x, z)$ holds. This interpretation is achieved thanks to the formula:

$$\forall x, y : \{\min(y) \to \big(\mathrm{W}(x,y) \leftrightarrow \theta(x,y)\big)\} \wedge \{\mathrm{W}(x, \mathrm{succ}(y)) \leftrightarrow \big(\theta(x, \mathrm{succ}(y)) \vee \mathrm{W}(x,y)\big)\} \tag{5}$$

Under assumption (5), the assertion $\forall x \exists y \theta(x, y)$ is equivalent to $\forall x \forall y : \max(y) \to \mathrm{W}(x, y)$. This allows to rewriting (4), and hence φ, as the $\mathrm{ESO}(\forall^2)$-formula:

$$\exists R_1, R_2, R_3, \mathrm{W}\ \big(\ (5)\ \wedge \forall x \forall y \psi(x, y, \mathbf{R})\ \wedge \forall x \forall y \big(\max(y) \to \mathrm{W}(x, y)\big).\ \big) \tag{6}$$

Thus, the above considerations allow to show the normalization $\mathrm{ESO}(\mathrm{var}\, d) = \mathrm{ESO}(\forall^d)$ on coordinate structures. It remains to prove $\mathrm{ESO}(\forall^d) = \mathrm{ESO}(\forall^d, \mathrm{arity}\, d)$. It amounts to build, for each formula Φ of type $\exists \mathbf{R} \forall x_1, \ldots, x_d \varphi$, where φ is quantifier-free and $\mathbf{R}$ is a tuple of relation symbols of any arity, a formula Φ' with the same shape, but in which all relation symbols are of arity $\le d$, such that Φ and Φ' have the same models, as far as pixel structures are concerned. The possibility to replace a k-ary ($k \ge d$) relation symbol R of Φ by d-ary symbols rests in the limitation of the number of first-order variables in Φ: each atomic formula involving R has the form $R(t_1, \ldots, t_k)$ where the t_i's are terms built on $x_1, \ldots, x_d$. Therefore, although R is k-ary, *in each of its occurrences* it behaves as a d-ary symbol, dealing with the sole variables $x_1, \ldots, x_d$. Hence, the key is to create a d-ary symbol for each occurrence of R in Φ or, more precisely, for each k-tuple of terms $(t_1, \ldots, t_k)$ involved in a R-atomic formula. Let us again opt for a 'proof-by-example' choice and illustrate the procedure on a very simple case.

Let Φ be the $\mathrm{ESO}(\forall^2, \mathrm{arity}\, 3)$-formula $\exists R \forall x, y \varphi(x, y, R)$, where $\varphi \equiv R(x,y,x) \wedge \neg R(y,x,y)$. Introduce two new *binary* relation symbols $R_{(x,y,x)}$ and $R_{(y,x,y)}$ associated to the triple of terms (x, y, x) and (y, x, y) involved in Φ, and fix their interpretation as follows: for any $\forall a, b \in [n]$, $R_{(x,y,x)}(a, b) \Leftrightarrow R(a, b, a)$ and $R_{(y,x,y)}(a, b) \Leftrightarrow R(b, a, b)$. Then we get the equivalence: $\langle S, R \rangle \models \forall x, y : R(x,y,x) \wedge \neg R(y,x,y)$ iff $\langle S, \mathbf{R} \rangle \models \forall x, y : R_{(x,y,x)}(x,y) \wedge \neg R_{(y,x,y)}(x,y)$ which, in turn, yields the implication:

$$S \models \exists R \forall x, y : R(x,y,x) \wedge \neg R(y,x,y) \Rightarrow S \models \exists \mathbf{R} \forall x, y : R_{(x,y,x)}(x,y) \wedge \neg R_{(y,x,y)}(x,y) \tag{7}$$

The converse implication would immediatly complete the proof. Unfortunately, it does not hold, since the second formula has a model, while the first has not.

To get the right-to-left implication in (7), we have to strengthen the second formula with some assertion that compels the tuple $R_{(x,y,x)}, R_{(y,x,y)}$ to be, in some sense, the binary representation of some ternary relation. This last construction is more sophisticated than the preceding ones, and we can't detail it here. ◀

Acknowledgments

We warmly thank Gaétan Richard who has obtained in collaboration with us several results of this paper, in particular Proposition 19 and Proposition 20. We also thank the anonymous referees for their careful reading, corrections and suggestions that have helped to notably improve the readability of the paper.

References

1 Bernd Borchert. Formal language characterizations of P, NP, and PSPACE. *J. of Automata, Languages and Combinatorics*, 13(3/4):161–183, 2008.

2 J. R. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik Grundlagen Math.*, 6:66–92, 1960.

3 M. Delorme and J. Mazoyer. *Cellular automata: A parallel model.* Springer, 373 pages, Mathematics and Its Applications, 1998.

4 A. Durand and E. Grandjean. First-order queries on structures of bounded degree are computable with constant delay. *ACM Transactions on Computational Logic*, V:1–18, 2006.

5 H.-D. Ebbinghaus and J. Flum. *Finite Model Theory.* Springer-Verlag, 1995.

6 D. Giammarresi and A. Restivo. *Two-dimensional languages, in: Handbook of Theoretical Computer Science*, volume 3- Beyond words, chapter 4, pages 215–267. Springer-Verlag New York, 1997.

7 D. Giammarresi, A. Restivo, S. Seibert, and W. Thomas. Monadic second-order logic over rectangular pictures and recognizability by tiling systems. *Information and Computation*, 125(1):32 – 45, 1996.

8 E. Grädel, Ph. G. Kolaitis, L. Libkin, M. Marx, J. Spencer, M. Y. Vardi, Y. Venema, and S.Weinstein. *Finite Model Theory and Its Applications.* Texts in Theoretical Computer Science. Springer, 2007.

9 N. Immerman. *Descriptive Complexity.* Graduate Texts in Computer Science. Springer, 1999.

10 M. Latteux and D. Simplot. Recognizable picture languages and domino tiling. Internal Report IT-94-264, Laboratoire d'Informatique Fondamentale de Lille, Université de Lille, France, 1994.

11 L. Libkin. *Elements of Finite Model Theory.* Springer, 2004.

12 Steven Lindell. A normal form for first-order logic over doubly-linked data structures. *Int. J. Found. Comput. Sci.*, 19(1):205–217, 2008.

13 Rudolph Sommerhalder and S. Christian van Westrhenen. Parallel language recognition in constant time by cellular automata. *Acta Inf.*, 19:397–407, 1983.

14 V. Terrier. Language recognition by cellular automata. In *Handbook of Natural Computing, Section 1, Cellular Automata.* Springer-Verlag, 2011.