A Hierarchy of Automatic Words having a Decidable MSO Theory

Vince Bárány
Mathematische Grundlagen der Informatik, RWTH Aachen

vbarany@informatik.rwth-aachen.de

July 2006

Abstract

We investigate automatic presentations of infinite words. Starting points of our study are the works of Rigo and Maes, and Carton and Thomas concerning the lexicographic presentation, respectively the decidability of the MSO theory of morphic words. Refining their techniques we observe that the lexicographic presentation of a (morphic) word is canonical in a certain sense. We then go on to generalize our techniques to a hierarchy of classes of infinite words enjoying the above mentioned properties. We introduce k-lexicographic presentations, and morphisms of level k stacks and show that these are inter-translatable, thus giving rise to the same classes of k-lexicographic or level k morphic words. We prove that these presentations are also canonical, which implies decidability of the MSO theory of every k-lexicographic word as well as closure of these classes under restricted MSO interpretations, e.g. closure under deterministic sequential mappings. The classes of k-lexicographic words are shown to form an infinite hierarchy.

1 Introduction

This paper is concerned with infinite words, of type ω, which are finitely presentable using automata and have a decidable monadic second-order theory. As such it is connected to two not so distant lines of research around the theme of using automata to decide logical theories.

Büchi originated the use of automata, omega-automata invented for this purpose, in deciding the MSO theory of the naturals with successor $(\mathbb{N}, succ)$. The fundamental result underlying this method is the convertibility of MSO formuli into Büchi automata and vice versa, which are equivalent in a natural way. The same correspondence holds for extensions of $(\mathbb{N}, <)$ by unary predicates P_a $(a \in \Sigma)$, which can be assumed to partition $\mathbb{N}$. Deciding the MSO theory of such extensions is, by the above, equivalent to the problem of deciding acceptance of the corresponding ω-word by any given Büchi automaton. Elgot and Rabin [10] have invented the *method of contractions* to reduce this problem, for suitable ω-words, to the case of ultimately periodic ones, for which it is trivially solvable. This was also the approach taken by Siefkes. However, little is known as to the applicability of the contraction method. Elgot and Rabin have illustrated their technique by proving the MSO decidability of the characteristic sequences of the factorial predicate, of k-th powers, and of powers of a fixed k. Fairly recently, Carton and Thomas [9] have used a very similar technique to prove the MSO decidability of *morphic words*. These results are elegantly rounded up in [19] and in the forthcoming [20].

The study of morphic words, having applications in combinatorics of words, goes back to Thue. An ω-word is said to be *morphic* if it is the homomorphic image of a fixed point of a morphism of finite words. It is thus associated to an HD0L system in a natural way, namely, as the limit of the sequence of words generated by it, i.e. by iterated application of the latter morphism (cf. [13]). Morphic words have been intensively studied in both of these contexts.

In this paper we define *morphisms of k-stacks* and classes of ω-words arising in a similar way by iterating such a morphism. We extend the result of Carton and Thomas to these classes of words. Additionally, we prove that for each k the class of k-morphic words is closed under MSO-definable re-colorings, e.g. under d.g.s.m. mappings. This seems to be a new result even for morphic words.

We use the formalism of *automatic presentations* of word structures. An (injective) automatic presentation of $(\mathbb{N}, <)$ consists of a regular set D of names, a synchronized rational relation $\prec$, and a bijective valuation function $\nu : D \to \mathbb{N}$ such that $n < m$ iff $\nu^{-1}(n) \prec \nu^{-1}(m)$ for every $n, m \in \mathbb{N}$. In the rich field of *generalized numeration systems* [6] the length-lexicographic ordering is the only natural choice for $\prec$. We consider automatic presentations of $(\mathbb{N}, <)$ where the ordering is a generalization of the length-lexicographical one. Given an ordered alphabet, we define the *k-lexicographic ordering* as a kind of k times nested length-lexicographical ordering.

A word structure is an extension $(\mathbb{N}, <, \{P_a\}_{a\in\Sigma})$ of the above by a unary predicate for each color according to the positions of the word where each symbol occurs. We generalize the result of Rigo and Maes [21] by showing that words representable using the k-lexicographic ordering are precisely those generated by iterated applications of a morphism of k-stacks followed by a homomorphic mapping. Thus, we obtain a hierarchy of k-lexicographic or k-morphic words. Indeed, we show that these classes form a strictly increasing infinite hierarchy.

Automatic structures are known to have a decidable first-order theory and to be closed under first-order interpretations [5]. We use "higher-order" automata constructions to extend these properties to monadic second-order logic over word structures having a k-lexicographic presentation. The construction follows the factorization of the ω-word provided by the k-lexicographic presentation, and the key step consists of showing that the *contraction* of an ω-word wrt. a given $(k + 1)$-lexicographic presentation and a given automaton is k-lexicographic. In other words, the automatic presentation guides us in applying the contraction method. This allows us to argue inductively, or, conversely and more intuitively, to reduce the MSO theory of a k-lexicographic word in k contraction steps to questions about ultimately periodic words.

Acknowledgements I thank Didier Caucal for correspondence, Tobias Ganzow for careful corrections, Łukasz Kaiser for numerous stimulating discussions and suggestions considerably simplifying the notation, Alexander Rabinovich for his [19], Sasha Rubin for pointing out [21], and Wolfgang Thomas for a preprint of [20].

2 Preliminaries

Words Let Σ be a finite alphabet. Σ^* denotes the set of finite words over Σ. The length of a word $w \in \Sigma^*$ is written $|w|$, the empty word is ε, for every $0 \leq i < |w|$ the ith symbol of w is written as $w[i]$, and when I denotes some interval of positions then wI (e.g. $w[n, m)$) is the factor of w on these positions. Note that we start indexing with 0. Accordingly, for every $n \in \mathbb{N}$, we let $[n] = \{0, \dots, n-1\}$.

Morphisms We denote by $\mathsf{Hom}(M, N)$ the set of homomorphisms from the monoid M to N. Each $\varphi \in \mathsf{Hom}(\Sigma^*, \Sigma^*)$ can be specified by the images $\varphi(a)$ of individual symbols $a \in \Sigma$. The length of φ, denoted $|\varphi|$, is the maximum of all the $|\varphi(a)|$, and φ is *uniform*, when $|\varphi(a)| = |\varphi|$ for every $a \in \Sigma$.

Automata A finite labelled transition system (TS) is a tuple $\mathcal{T} = (Q, \Sigma, \Delta)$, where Q is a finite, nonempty set of states, Σ is a finite set of labels, and $\Delta \subseteq Q \times \Sigma \times Q$ is the transition relation. $\mathcal{T}$ is deterministic (DTS), when Δ is a function of type $Q \times \Sigma \to Q$, in this case we write δ instead of Δ, and δ^* for the unique homomorphic extension of δ to all words over Σ. Alternatively, each deterministic transition system can be represented as a pair (φ, M) where $M = (Q \to Q, \circ, \mathsf{id})$ is the monoid of (partial) functions from Q to Q with composition as product and $\varphi \in \mathsf{Hom}(\Sigma^*, M)$ is such that $\varphi(a)(q) = \delta(q, a)$ for every $a \in \Sigma$

and $q \in Q$. From (φ, M) one can again obtain the presentation (Q, Σ, δ). A *finite automaton* (FA) is a finite transition system together with sets of initial and final states $\mathcal{A} = (\mathcal{T}, I, F) = (Q, \Sigma, \Delta, I, F)$. $\mathcal{A}$ is deterministic (DFA) when $\mathcal{T}$ is and when I contains a single initial state q_0. The unfolding of a DFA $\mathcal{A}$ from its initial state is a Σ-branching Q-labelled regular tree, i.e. one having only finitely many subtrees up to isomorphism. Conversely, each such regular tree determines a DFA having the subtree-types as its states. The *completion* of a DFA $\mathcal{A}$ is the DFA $\overline{\mathcal{A}}$ obtained by introducing a new state $\perp$ and setting it the target of all yet undefined transitions. Thus, the transition function $\overline{\delta}$ of $\overline{\mathcal{A}}$ is defined for all pairs (q, a) with $q \in Q \cup \{\perp\}$.

Multi-tape automata Let Σ be a finite alphabet. We consider relations on words, i.e. subsets R of $(\Sigma^*)^n$ for some $n > 0$. *Asynchronous* n-tape automata accept precisely the *rational relations*, i.e., rational subsets of the product monoid $(\Sigma^*)^n$. Finite *transducers*, recognizing *rational transductions* [3], are asynchronous 2-tape automata. A relation $R \subseteq (\Sigma^*)^n$ is *synchronized rational* [11] or *regular* [15] if it is accepted by a *synchronous* n-tape automaton. Finally, $R \subseteq (\Sigma^*)^n$ is *semi-synchronous rational* [2] if it is accepted by an n-tape automaton reading each of its tapes at a fixed speed. A *deterministic generalized sequential machine* (d.g.s.m.) $\mathcal{S} = (\mathcal{T}, q_0, \mathcal{O})$ consists of a DTS, an initial state, and an output function $\mathcal{O} : Q \times \Sigma \to \Gamma^*$ and computes, in a natural way, a function $S : \Sigma^* \to \Gamma^*$.

Automatic structures The idea to use automata to represent structures goes back to Büchi. The general notion was first introduced and studied by Hodgson [12] and was then rediscovered by Khoussainov and Nerode [14]. Since then it has been subject of some theses and numerous publications, see e.g. [4, 5, 15, 22] for an overview. Note that a great deal of attention has been given to natural automatic presentations of specific structure classes including groups [8], semigroups, and automatic sequences [7, 6, 1]. We shall take all structures to be relational with functions represented by their graphs.

Definition 2.1 (Automatic structures [14]). An automatic presentation of a structure $\mathfrak{A} = (A, \{R_i\}_i)$ consists of a collection of synchronous automata $\mathfrak{d} = (\mathcal{A}_D, \mathcal{A}_\varepsilon, \{\mathcal{A}_{R_i}\}_i)$ and a naming-, or coordinate function $\nu : L(\mathcal{A}_D) \to A$, such that $\varepsilon = L(\mathcal{A}_\varepsilon)$ is the equivalence relation $\{(x, y) \mid \nu(x) = \nu(y)\}$ and ν is a homomorphism from $(L(\mathcal{A}_D), \{L(\mathcal{A}_{R_i})\}_i)$ onto $\mathfrak{A}$, hence $\mathfrak{A} \cong (L(\mathcal{A}_D), \{L(\mathcal{A}_{R_i})\}_i)/_\varepsilon$. AutStr designates the class of automatic structures.

One effectively obtains an injective presentation from any given automatic presentation by restricting $L(\mathcal{A}_D)$ to a set of unique (e.g. length-lexicographically least) representants of each ε-class. In this paper we will only consider injective presentations, omitting ε, and quite often tacitly consider a tuple of regular relations $(D, \{R_i\}_i)$ as an automatic presentation.

Logics We use the abbreviation FO and MSO for first-order and for monadic second-order logic, respectively, and $\mathsf{FO}^{\infty,\mathrm{mod}}$ for the extension of FO by infinity $(\exists^\infty)$ and modulo-counting quantifiers $(\exists^{(r,m)})$. The meaning of the formuli $\exists^\infty x\,\theta$ and $\exists^{(r,m)} x\,\theta$ is that there are infinitely many elements x, respectively r many elements x modulo m, such that θ holds. We shall make extensive use, often without direct reference, of the well-known relationships of automata and logics (cf. [23]) as well as of the following facts.

Theorem 2.2. *(Consult [4, 5] and [16, 22].)*

i) Let $(\mathfrak{d}, \nu)$ be an aut. pres. of $\mathfrak{A} \in$ AutStr. Then for each $\mathsf{FO}^{\infty,\mathrm{mod}}$-formula $\varphi(\vec{a}, \vec{x})$ with parameters $\vec{a}$ from $\mathfrak{A}$, defining a k-ary relation R over $\mathfrak{A}$, one can construct a k-tape synchronous automaton recognizing $\nu^{-1}(R)$.

ii) The $\mathsf{FO}^{\infty,\mathrm{mod}}$-theory of every automatic structure is decidable.

iii) AutStr is effectively closed under $\mathsf{FO}^{\infty,\mathrm{mod}}$-interpretations.

In this paper we extend these results to MSO over word structures having automatic presentations of a certain kind to be introduced later. It will be convenient to consider automatic presentations up to equivalence.

Definition 2.3 (Equivalence of automatic presentations).
Two presentations $(\mathfrak{d}_1, \nu_1)$ and $(\mathfrak{d}_2, \nu_2)$ of some $\mathfrak{A} \in$ AutStr are *equivalent* when for every relation R over $\mathfrak{A}$, $\nu_1^{-1}(R)$ is regular iff $\nu_2^{-1}(R)$ is regular.

In other words, two automatic presentations are equivalent if there is no difference between them in terms of representability of relations via automata, i.e. if they are expressively equivalent. In [2] we have shown that two presentations are equivalent iff the transduction $T = \{(x, y) \in D \times D' \mid \nu_1(x) = \nu_2(y)\}$, translating names of elements from one presentation to the other, is semi-synchronous rational. Thus, equivalent presentations are truly identical modulo such a simple coding, i.e. expressive equivalence coincides with computational equivalence.

3 Word structures

An ω-word over Σ is a function $w : \mathbb{N} \to \Sigma$. The set of ω-words over Σ is denoted Σ^ω. To every $w \in \Sigma^\omega$ we associate its *word structure* $W_w = (\mathbb{N}, <, \{P_a\}_{a\in\Sigma})$, where $P_a = w^{-1}(a)$ for each $a \in \Sigma$. Word structures of finite words are defined similarly. Note that we consider the ordering, as opposed to the successor relation, as given in our word structures. When one is working with monadic second-order logic, there is of course no difference in terms of expressiveness. However, as we are engaging in an investigation of automatically presentable word structures, the presence of the ordering is not without significance.

Morphic words A particularly well understood class of ω-words is that of the so called *morphic words*. The basic idea, successfully applied by Thue, is to obtain an infinite word via iteration of a suitable morphism $\tau : \Sigma^* \to \Sigma^*$. Suitability is expressed by the condition that $\tau(a)[0] = a$ for some $a \in \Sigma$. In this case τ is said to be *prolongable on a*. This ensures that the sequence $(\tau^n(a))_{n\in\mathbb{N}}$ converges to either a finite or infinite word, which is a fixed point of τ, denoted $\tau^\omega(a)$. An ω-word $w \in \Gamma^*$ is morphic, if $w = \sigma(\tau^\omega(a))$ for some τ prolongable on a and some $\sigma \in \mathsf{Hom}(\Sigma^*, \Gamma^*)$ extended in the obvious way to ω-words.

Example 3.1. Consider $\tau : a \mapsto ab, b \mapsto ccb, c \mapsto c$ and $\sigma : a, b \mapsto 1, c \mapsto 0$ both homomorphically extended to $\{a, b, c\}^*$. The fixed point of τ starting with a is the word $abccbccccbc^6b\ldots$, and its image under σ, $11001000010^61\ldots$, is the characteristic sequence of the set of squares.
In general, as was shown in [9], the characteristic sequence of every set of the form $\{\sum_{k=0}^n s_k \mid n \in \mathbb{N}\}$, where $0 < (s_k)$ is an $\mathbb{N}$-rational sequence is morphic. This result follows trivially from the characterization of [21], cf. Proposition 4.2.

Example 3.2. Let $\phi : a \mapsto ab, b \mapsto a$. Its fixed point $\phi^\omega(a)$ is the *Fibonacci word* $f = abaababaabaababaababa\ldots$, so called for the recursive dependence $\phi^{n+2}(a) = \phi^{n+1}(a)\cdot\phi^n(a)$ implying that $|\phi^n(a)|$ is the n^{th} Fibonacci number.

Automatic presentations

In accordance with Definition 2.1 an automatic presentation $(D, R, \{P_a\}_{a\in\Sigma})$ of W_w as above comprises a regular set D partitioned by the regular sets P_a for each $a \in \Sigma$ over some alphabet Γ, together with a regular relation R, which is a linear ordering of type ω over D such that the i-th word in this ordering belongs to P_a iff the i-th symbol of w is a.
The most frequently, if not exclusively, used regular ordering of type ω is the *length-lexicographic* ordering, also called military-, radix-, or genealogical ordering by some or shortlex by others. Starting point of

our investigation is the observation that those words possessing an automatic presentation using the length-lexicographic ordering are precisely those *morphic* (cf. [21]). Nevertheless there are other choices of ordering worth investigating. Indeed, as we shall see, increasing the complexity of the ordering widens the class of words thus presentable. First we define the key concept of canonicity and derive extensions of Theorem 2.2 to MSO over word structures having a canonical presentation.

Definition 3.3 (Canonical presentations).
An automatic presentation $\mathfrak{d} = (D, <, \{P_a\}_{a\in\Sigma})$ of some infinite word $w \in \Sigma^\omega$ is *canonical* when there is an algorithm, which constructs for every homomorphism $\psi \in \mathsf{Hom}(\Sigma^*, M)$ into a finite monoid M and for every monoid element $m \in M$ a synchronous two-tape automaton recognizing the relation

$$B_m = \{(x, y) \in D^2 \mid x < y \wedge \psi(w[\nu(x), \nu(y)]) = m\} \ .$$

Thus, canonicity means that membership of finite factors of w in a regular language can be decided by an effectively constructible automaton reading the representations of the two endpoints of the factor.

Lemma 3.4. Let $\mathfrak{d} = (D, <, \{P_a\}_{a\in\Sigma})$ and ν constitute a canonical presentation of $w \in \Sigma^\omega$. Then for every deterministic Muller automaton $\mathcal{A}$ an automaton recognizing the following set can be effectively constructed.

$$E_{\mathcal{A}} = \{x \in D \mid w[\nu(x), \infty) \in L(\mathcal{A})\}$$

Corollary 3.5. *Let w be an ω-word having a canonical automatic presentation. Then the* MSO*-theory of W_w is decidable.*

Let φ be an MSO sentence in a language of word structures and let x, y be first-order variables not occurring in any subformula of φ. We define three kinds of *relativizations of* φ: $\varphi^{[0,x]}$, $\varphi^{[x,y]}$, and $\varphi^{[x,\infty)}$ obtained by relativizing all first- and second-order quantifications to the noted intervals. For instance $(\exists z\vartheta)^{[x,y]} = \exists z(x \leq z \wedge z \leq y \wedge \vartheta^{[x,y]})$, and $(\forall Z\vartheta)^{[x,\infty)} = \forall Z(\forall z(z \in Z \to x \leq z) \to \vartheta^{[x,\infty)})$. The relevance of relativization is expressed by the equivalence $W_w \models \varphi^I \iff W_{wI} \models \varphi$, where I is an interval of any of the three kinds.

Lemma 3.6 (Normal Form of MSO formuli over word structures).
Every MSO formula $\varphi(\vec{x})$ having free first-order variables $x_0, \ldots, x_{n-1}$ and no free second-order variables is equivalent to a boolean combination of formuli $x_i < x_j$ and relativized MSO sentences $\vartheta^{[0,x_i]}$, $\vartheta^{[x_i,x_j]}$, and $\vartheta^{[x_i,\infty)}$ with $i, j \in [n]$.

Theorem 3.7 (MSO definability). *Let w be an ω-word having a canonical presentation $\mathfrak{d}$ having domain D and bijective coordinate function $\nu : D \to \mathbb{N}$. Then there is an algorithm transforming every* MSO *formula $\varphi(\vec{x})$ having n free first-order variables (and no free set variables) into an n-tape synchronous automaton $\mathcal{A}$ such that for every $u_1, \ldots, u_n \in D$*

$$W_w \models \varphi[\nu(\vec{u})] \iff \vec{u} \in L(\mathcal{A})$$

Proof
Using the above Lemma, we transform φ into a boolean combination of relativized sentences and comparison formuli $x_i < x_j$. Canonicity and Lemma 3.4 deliver automata recognizing the relations defined by relativized sentences $\vartheta^{[0,x_i]}$, $\vartheta^{[x_i,x_j]}$, respectively $\vartheta^{[x_i,\infty)}$. Thus, by the appropriate combination of the automaton recognizing $<$ and of the automata recognizing the relativized subformulae of the normal form we obtain $\mathcal{A}$ as required. □

Note that a set $X \subseteq \mathbb{N}$ is definable by an MSO formula $\psi(X)$ in W_w iff it is pointwise definable by one of the form $\varphi(x)$. Thus, (W_w, X) is automatic for every canonically presentable W_w and for every in W_w MSO-definable X.

4 k-lexicographic presentations

Let Σ be a finite non-empty alphabet. To each word $u = a_0 a_1 \dots a_{n-1} \in \Sigma^*$ of length n and to each $0 < k$ we associate its k-*split* $(u^{(1)}, u^{(2)}, \dots, u^{(k)})$ defined as follows. Let t be such that $tk \le n < (t+1)k$. Then the ith word of the k-split is $u^{(i+1)} = a_i a_{k+i} a_{2k+i} \dots a_{tk+i}$ for each $i < k$. Conversely, the k-merge of the components produces the original word $u = \otimes_k(u^{(1)}, \dots, u^{(k)})$. Additionally, we define $u^{(0)} = |u| \in \mathbb{N}$ or in unary presentation as $1^{|u|}$, whichever is more convenient. For $0 \le i < k$ we define the equivalence

$$u =_i v \overset{\text{def}}{\iff} \forall j \le i \quad u^{(j)} = v^{(j)} \ .$$

This implies, in particular, $|u| = |v|$. Let now $<$ be a linear ordering of Σ, and let $<_{\text{lex}}$ denote the induced lexicographic ordering. For each $0 \le k$ we define the k-length-lexicographic ordering ($<_{k\text{-llex}}$) of Σ^* as

$$u <_{k\text{-llex}} v \overset{\text{def}}{\iff} |u| < |v| \lor \exists i < k : \ u =_i v \land u^{(i+1)} <_{\text{lex}} v^{(i+1)} \ .$$

Definition 4.1 (k-lexicographic words). An ω-word $w \in \Sigma^\omega$ is k-*lexicographic* (short: k-*lex*) if there is an automatic presentation $(D, <_{k\text{-llex}}, \{P_a\}_{a\in\Sigma})$ of the associated word structure W_w. For each k, the class of k-lexicographic words is denoted $\mathcal{W}_k$, and we also let $\mathcal{W} = \bigcup_k \mathcal{W}_k$.

Observe that the 0-lexicographic ordering is just the ordering of words according to their length. Therefore, the domain of a 0-lex presentation has to be *thin*, i.e. containing at most one word of each length. All such presentations are easily seen to be equivalent to one over a unary alphabet. Thus, $\mathcal{W}_0$ is precisely the class of ultimately periodic words. Further, it is not hard to see, that an ω-word is 1-lex iff it is morphic.

Proposition 4.2.

0. $\mathcal{W}_0$ is the class of ultimately periodic words.
1. $\mathcal{W}_1$ is the class of morphic words (cf. [21]).

Let us now give an example of a 2-lexicographic word, which is not morphic.

Example 4.3. Consider the Champernowne word $s = 12345678910111213\dots$ (also called Smarandache sequence) obtained by concatenating all decimal numerals (without leading zeros) in ascending, i.e. length-lexicographic order. To give a natural 2-lex presentation of W_s we use words $\otimes_2(x^{(1)}, x^{(2)})$ such that $x^{(1)}$ is a decimal numeral (not starting with a zero) and $x^{(2)} \in 1^*01^*$. We use the single 0 in $x^{(2)}$ to mark a position within $x^{(1)}$. For each digit $d \in [10]$ we can thus define the unary predicate P_d as $([10]1)^* d0([10]1)^* \setminus 0[10]^*$.

We close this section with two simple but useful observations.

Proposition 4.4 (Closure under homomorphic mappings). The class of automatically presentable ω-words is closed under homomorphic mappings. In particular, if w is k-lexicographic, then so is $h(w)$ for every homomorphism h.

Lemma 4.5 (Normal Form Lemma). Let $1 < k \in \mathbb{N}$. Each k-lexicographic presentation $\mathfrak{d} = (D, <_{k\text{-llex}})$ of $(\mathbb{N}, <)$ over an alphabet Σ is equivalent to one $\mathfrak{d}' = (D', <_{k\text{-llex}})$ over some Γ such that $D' \subseteq (\Gamma^k)^*$. In fact, one can choose $\Gamma = \{0, 1\}$ in the above.

5 Canonicity, Closure and Decidability

Let a $(k+1)$-lex presentation $\mathfrak{d} = (D, <_{(k+1)\text{-llex}}, \{\mathcal{A}_a\}_{a\in\Sigma})$ of $w \in \Sigma^\omega$ in normal form over the alphabet Γ together with the bijective coordinate function $\nu : D \to \mathbb{N}$ as well as a homomorphism $\psi \in \mathsf{Hom}(\Sigma^*, M)$

into a finite monoid M be given. We associate to $\mathfrak{d}$ the DFA $\mathcal{A}_{\mathfrak{d}} = \overline{\prod_{a\in\Sigma}\mathcal{A}_a}$ consisting of the DTS $\mathcal{T}_{\mathfrak{d}} = (Q_{\mathfrak{d}}, \Gamma, \delta_{\mathfrak{d}})$ and having initial state $\vec{q}_0$. Further let $\sigma_{\mathfrak{d}} \in \mathsf{Hom}(Q_{\mathfrak{d}}^*, \Sigma^*)$ be such that $\sigma_{\mathfrak{d}}(\vec{q}) = a$ whenever the a^{th} component of $\vec{q}$ is in a final state (in which case a is uniquely determined) and $\sigma_{\mathfrak{d}}(\vec{q}) = \varepsilon$ otherwise. Finally, we set $w_{\mathfrak{d}} = \prod_{x\in\Gamma^*}^{<_{(k+1)\text{-llex}}} \delta_{\mathfrak{d}}^*(\vec{q}_0, x) \in Q_{\mathfrak{d}}^\omega$. Clearly, $w = \sigma_{\mathfrak{d}}(w_{\mathfrak{d}})$
For every $x = \otimes_{k+1}(x^{(1)}, \ldots, x^{(k+1)})$ let $x' = \otimes_k(x^{(1)}, \ldots, x^{(k)})$ be the projection of x onto its first k splitting components, when $k > 0$ and let $x' = x^{(0)} = 1^{|x|}$ when $k = 0$. We define $D' = \{x' \mid x \in D\}$ as the point-wise projection of D. The equivalence $=_k$ partitions the set D of indices into consecutive intervals. Let $c(x)$ denote the interval containing x, i.e. $c(x) = \{y \mid y' = x'\}$, and consider the factorization of w according to such intervals.

$$w = \prod_{x'\in D'}^{<_{k\text{-llex}}} w[\nu(c(x)\cap D)]$$

The *contraction* (compare with that of [10]) of w wrt. $\mathfrak{d}$ and ψ is the ω-word

$$c_{\mathfrak{d}}^{\psi}(w) = \prod_{x'\in D'}^{<_{k\text{-llex}}} \psi(w[\nu(c(x)\cap D)]) \quad \in M^\omega$$

indexed by elements of D' ordered by $<_{k\text{-llex}}$. We can prove that $c_{\mathfrak{d}}^{\psi}(w)$ is in fact automatically presentable over $(D', <_{k\text{-llex}})$.

Lemma 5.1 (Contraction Lemma).
Let $\mathfrak{d} = (D, <_{k+1\text{-llex}}, \{\mathcal{A}_a\}_{a\in\Sigma})$ be a $(k+1)$-lex presentation with coordinate function ν of the word structure of an ω-word $w \in \Sigma^\omega$. Then for every finite monoid M, every $\psi \in \mathsf{Hom}(\Sigma^*, M)$ and for each $m \in M$ the following relations are regular.

$$\begin{aligned} B'_m &= \{(x,y) \in D^2 \mid x \leq_{k+1\text{-llex}} y \wedge x =_k y \wedge \psi(w[\nu(x),\nu(y)]) = m\} \\ P'_m &= \{x' \in D' \mid \psi(w[\nu(c(x)\cap D)]) = m\} \end{aligned}$$

Whence, $(D', <_{k\text{-llex}}, \{P'_m\}_{m\in M})$ is a k-lexicographic presentation of $c_{\mathfrak{d}}^{\psi}(w)$.

In particular, the contraction of a morphic word wrt. any given lexicographic presentation and any given morphism into a finite monoid is an ultimately periodic sequence. This is already sufficient to yield MSO decidability of morphic words, and is essentially the proof given in [9]. Obviously, by iterating this contraction process starting from any given k-lex presentation of an ω-word we arrive after (at most) k contractions, at an ultimately periodic sequence. It is now easy to use this fact to prove MSO decidability of k-lexicographic words. However, we aim for the stronger canonicity result.

Main Theorem 5.2 (Canonicity of k-lex presentations).
All k-lexicographic presentations are canonical.

Proof
The proof is by induction on k, the base case being clear. For the induction step, we consider a $k+1$-lex presentation. Observe that if two $k+1$-lex presentations of the same ω-word are equivalent, then one is canonical iff the other one is. Therefore, by the Normal Form Lemma, it is sufficient to provide a proof for $k+1$-lex presentations in normal form. So let $\mathfrak{d} = (D, <_{k+1\text{-llex}}, \{P_a\}_{a\in\Sigma})$ be a $k+1$-lex presentation in normal form of an ω-word $w \in \Sigma^\omega$. Let a morphism $\psi \in \mathsf{Hom}(\Sigma^*, M)$ into a finite monoid M be given. We need to construct automata deciding, given words $x, y \in D$ with $x \leq_{k+1\text{-llex}} y$, whether $\psi(w[\nu(x),\nu(y)]) = m$. There are two cases. If $x' = y'$ then we simply verify $(x,y) \in B'_m$ as in the Contraction Lemma. When on the other hand $x' <_{k\text{-llex}} y'$ then we partition the interval $x \leq_{k+1\text{-llex}} z \leq_{k+1\text{-llex}} y$ into three segments according to whether $x' = z'$, $x' <_{k\text{-llex}} z' <_{k\text{-llex}} y'$ or $z' = y'$,

i.e. consider the factors $w[\nu(x),\nu(\hat{x})]$, $w[\nu(\{z\in D \mid x' <_{k\text{-llex}} z' <_{k\text{-llex}} y'\})]$ and $w[\nu(\hat{y}),\nu(y)]$, where $\hat{x}$ is the greatest element of $c(x)\cap D$ with respect to $<_{k+1\text{-llex}}$ and similarly $\hat{y}$ is the least element of $c(y)\cap D$. Note that both $\hat{x}$ and $\hat{y}$ are first-order definable, hence automaton computable from x, respectively from y. We can therefore compute $\psi(w[\nu(x),\nu(\hat{x})])$ as well as $\psi(w[\nu(\hat{y}),\nu(y)])$ by an automaton simultaneously verifying B'_m for both pairs $(x,\hat{x})$ and $(\hat{y},y)$ for all $m\in M$.
It remains to show that the value of the central segment is also automaton computable. By the Contraction Lemma we know that $\mathfrak{d}' = (D', <_{k\text{-llex}}, \{P'_m\}_{m\in M})$ is a k-lex presentation of $c^{\psi}_{\mathfrak{d}}(w)$. Thus, by the induction hypothesis, $\mathfrak{d}'$ is canonical. We use this fact to compute the value of the central segment. To this end, we employ the *multiplier morphism* $\mu_M \in \mathsf{Hom}(M^*,M)$ defined by stipulating that $\mu_M(m)=m$ for all $m\in M$. Let ν' denote the co-ordinate mapping associated to $\mathfrak{d}'$. By definition of a contraction $\psi(w[\nu(c(z)\cap D)]) = c^{\psi}_{\mathfrak{d}}(w)[\nu'(z')]$, therefore the value of the central segment $\psi(w[\nu(\{z\in D \mid x' <_{k\text{-llex}} z' <_{k\text{-llex}} y'\})])$ can be written as $\mu_M(c^{\psi}_{\mathfrak{d}}(w)(x',y'))$, which is by canonicity of $\mathfrak{d}'$ automaton computable. □

Corollary 5.3 (MSO decidability).
The MSO theory of the word structure W_w associated to a k-lex word $w\in\mathcal{W}$ is decidable.

MSO interpretations are usually understood to be one-dimensional. We use the notation $\leq^{\mathcal{I}}_{\mathrm{mdMSO}}$ to stress that $\mathcal{I}$ might be multi-dimensional. Further, we say that a tuple $(\varphi(x),\{\varphi_b(x)\}_{b\in\Gamma})$ of MSO formuli, together with the formula $\varphi_<(x,y) = x<y$, form a *restricted* MSO *interpretation* $\mathcal{I}$ (the restriction being that $\mathcal{I}$ can only redefine the coloring, but not $<$) of a finite or infinite word structure $W_{w'} \leq^{\mathcal{I}}_{\mathrm{rMSO}} W_w$. From Theorem 5.2 and Theorem 2.2 we conclude the next corollaries.

Corollary 5.4 (Closure under MSO interpretations).
Let w be a k-lexicographic word. For every structure $\mathfrak{A}$ and word w' we have

1. $\mathfrak{A} \leq_{\mathrm{mdMSO}} W_w \Longrightarrow \mathfrak{A}$ *is automatic,*
2. $W_{w'} \leq_{\mathrm{rMSO}} W_w \Longrightarrow W_{w'}$ *is k-lexicographic.*

Corollary 5.5 (Closure under d.g.s.m. mappings).
For each $k\in\mathbb{N}$ the class $\mathcal{W}_k$ is closed under deterministic generalized sequential mappings.

As an example of what can be interpreted in a word consider the following.

Theorem 5.6 (Automatic equivalence structures). *Consider $\mathfrak{A}=(A,E)$ with E an equivalence relation on a countably infinite set A having only finite equivalence classes. Assume further that for each n there are $f(n)\in\mathbb{N}$ many equivalence classes of size n.*
Then $\mathfrak{A}\in$ AutStr if and only if there is a 2-lex word $w = 0^{m_0}10^{m_1}10^{m_2}1\ldots$ such that $f(n) = |\{i \mid m_i = n\}|$, in which case $\mathfrak{A} \leq^{\mathcal{I}}_{\mathrm{FO}} W_w$ for a fixed one-dimensional FO-interpretation $\mathcal{I}$, also implying that $\mathrm{Th}_{\mathrm{MSO}}(\mathfrak{A})$ is decidable.

6 Hierarchy Theorem

It is readily seen, that $\mathcal{W}_k$ is included in $\mathcal{W}_{k+1}$ for each k. Next we show that each $\mathcal{W}_k$ is properly included in the next one by exhibiting ω-words $s_{k+1}\in\mathcal{W}_{k+1}\setminus\mathcal{W}_k$. We call the s_k *stuttering words.* Each s_k is a word over the $(k+1)$-letter alphabet $\{a_0,\ldots,a_k\}$ and is defined as the infinite concatenation product $s_k = \prod_{n=0}^{\infty} s_{k,n}$, where $s_{0,n}=a_0$ and $s_{k+1,n} = (s_{k,n})^{2^n} a_{k+1}$ for every k and n. To give an illustration, we write for convenience $a,b,c,d\ldots$ instead of $a_0,a_1,a_2,a_3\ldots$ for small k. The first few stuttering words are

$$\begin{array}{rcl} s_0 & = & a^\omega \\ s_1 & = & abaabaaaaba^8ba^{16}b\dots \\ s_2 & = & abcaabaabc(aaaab)^4c(a^8b)^8c\dots \\ s_3 & = & abcd(aabaabc)^2d((aaaab)^4c)^4d((a^8b)^8c)^8d\dots \\ & \vdots & \end{array}$$

Theorem 6.1 (Hierarchy Theorem).
For each $k \in \mathbb{N}$ we have $s_{k+1} \in \mathcal{W}_{k+1} \setminus \mathcal{W}_k$.

Proof
We leave it to the reader to give a k-lex presentation of s_k for every k.
To show that $s_{k+1} \notin \mathcal{W}_k$ we argue indirectly as follows. Assume that there is a k-lex presentation $(D, <_{k\text{-llex}}, \{P_{a_i}\}_{i \leq k+1})$ of s_{k+1}, and assume it to be in normal form, i.e. $D \subseteq (\{0,1\}^k)^*$. Consider for each $i \leq k+1$ the (regular) relations $S_i(x,y)$ consisting of pairs of consecutive words $x, y \in P_{a_i}$, i.e. such that there are no occurrences of a_i on intermediate positions. Let automata be given for D,P_{a_i}, and S_i for every $i \leq k+1$ and let C be greater than the maximum of the number of states of any of these automata.

Claim 1. *Let x represent the position of the n^{th} occurrence of a_{k+1} in s_{k+1}. Then $(k+1)n < |x| \leq Cn$, i.e. $|x| = \Theta(n)$, and hence $n = \Theta(|x|)$.*
The upper bound $|x| \leq Cn$ is clear, and $(k+1)n \leq |x|$ follows from that there are more than $2^{(k+1)n}$ symbols preceding the n^{th} a_{k+1} in s_{k+1}.

Claim 2. *For every $i = 1, \dots, k$ there is a t_i such that for all $N \in \mathbb{N}$ there are $x = \otimes_k(x^{(1)}, \dots, x^{(k)})$, and $y = \otimes_k(y^{(1)}, \dots, y^{(k)})$ with $|x| = |y| > N$ and such that $S_i(x,y)$ and $x \sim_{k-i} y$ (i.e. $x^{(j)} = y^{(j)}$ for all $j \leq k-i$) and that $x^{(k-i+1)}$ and $y^{(k-i+1)}$ differ only on their last t_i bits.*
For $i = 1$ we immediately get a contradiction since between consecutive a_1's represented by words x and y of length $> N$ there are $2^{\Omega(N)}$ many a_0's but by Claim 2 there are only 2^{t_1} words between x and y in the k-lexicographic ordering.
Proof of Claim 2. We start with $i = k$ and proceed inductively in descending order. Values of the t_i will be implicitly given during the proof.
From Claim 1 we know that $|v| < |u| + C$ for every $S_{k+1}(u,v)$, and that if u represents the position of the n^{th} a_{k+1} then $n = \Theta(|u|)$. Then there are 2^n many a_k's distributed evenly between u and v, therefore there must be some $|u| \leq l \leq |v|$ such that there are still at least $2^n/C$ many $u <_{k\text{-llex}} x <_{k\text{-llex}} v$, $|x| = l$, and $x \in P_{a_k}$. When $n > C\log C$ then $2^n/C > 2^C$, so we have more than 2^C many $|x| = l$, and $x \in P_{a_k}$.
We claim that there are $x = \otimes_k(x^{(1)}, \dots, x^{(k)})$ and $y = \otimes_k(y^{(1)}, \dots, y^{(k)})$ such that $S_k(x,y)$ and $x^{(1)}$ and $y^{(1)}$ agree on their first C symbols. In deed, the first C symbols of the 1^{st} component can be incremented at most 2^C times and by the choice of n and l there are more than 2^C occurrences of a_k on length l.
Let now $t_k = l - C$. By pumping into the initial segment of length kC of the pair (x,y) (note that this involves the first C symbols of each component) we obtain arbitrary long x', y' with $S_k(x',y')$ whose 1^{st} components only differ on their last t_i bits. Thus we have established the case $i = k$.
To advance from $i+1$ to i we do the same as above. By the induction hypothesis we have for arbitrary large n two words $u = \otimes_k(u^{(1)}, \dots, u^{(k)})$ and $v = \otimes_k(v^{(1)}, \dots, v^{(k)})$ both of length n such that $S_{i+1}(u,v)$ and having $u^{(j)} = v^{(j)}$ for all $j < k-i$ and $u^{(k-i)}$ and $v^{(k-i)}$ differing only on their last t_{i+1} bits. By Claim 1 there are $2^{\Theta(n)}$ occurrences of a_i in between these two positions. On the other hand the remaining last t_i bits of the $(k-i)^{th}$ components together with the first C bits of the $(k-i+1)^{th}$ components only allow for 2^{C+t_i} possibilities. Hence for large enough n we must have two consecutive a_i's on positions represented by x and y agreeing on their first $(k-i)$ components and on the first C bits of their $(k-i+1)^{th}$ components. Thus, by pumping into the initial segment of length kC of the pair

(x, y) we obtain arbitrary long x', y' fulfilling the conditions of Claim 2 for i. □

7 Equivalent characterizations

Let Γ be a finite, non-empty stack alphabet. A (level 1) stack is a finite sequence of symbols of Γ, and level $k+1$ stacks are sequences of level k stacks. Additionally, we shall call individual symbols of Γ level 0 stacks. Formally

$$\begin{aligned} \mathsf{Stack}_\Gamma^{(0)} &= \Gamma \\ \mathsf{Stack}_\Gamma^{(k+1)} &= [(\mathsf{Stack}_\Gamma^{(k)})^*] \end{aligned}$$

where '[' and ']' are used to identify the boundaries of lower-level stacks within higher-level ones. Outer most brackets will most often be omitted.

Level k stacks can be viewed as trees of height k having an unbounded number of ordered branches and leaves labelled by elements of Γ. Each leaf, i.e. each level 0 element stored in a k-stack γ can be accessed by a vector of k indices $(i_0, \ldots, i_{k-1})$ leading to it. We denote the sequence of "leaves" of a $k+1$-stack γ, taken in the natural ordering, by $\mathsf{leaves}(\gamma)$. In other words, $\mathsf{leaves}(\gamma)$ is obtained from γ by forgetting the brackets.

The *concatenation* of two $(k+1)$-stacks $\gamma^{(k+1)} = [\gamma_1^{(k)} \ldots \gamma_s^{(k)}]$ and $\xi^{(k+1)} = [\xi_1^{(k)} \ldots \xi_t^{(k)}]$ is the $(k+1)$-stack $\gamma^{(k+1)} \cdot \xi^{(k+1)} = [\gamma_1^{(k)} \ldots \gamma_s^{(k)} \xi_1^{(k)} \ldots \xi_t^{(k)}]$. Concatenation can also be regarded as operations on trees. For $k > 0$ every k-stack $\gamma^{(k)} = [\gamma_0^{(k-1)} \ldots \gamma_{s-1}^{(k-1)}]$ can be written as the concatenation product $\prod_{i=0}^{s-1} [\gamma_i^{(k-1)}]$ and by propagating through all dimensions as

$$\gamma^{(k)} = \prod_{i_0} [\prod_{i_1} [\cdots \prod_{i_{k-1}} [\gamma^{(0)}_{(i_0,\ldots,i_{k-1})}] \cdots]] \tag{1}$$

where the index vector $(i_0, \ldots, i_{k-1})$ runs through all allowed tuples (all branches of length k) in a k-lexicographic fashion.

Definition 7.1 (Morphisms of k-stacks). *Morphisms of k-stacks* over Γ are just k-stacks of actions of Γ. That is, $\mathsf{Hom}_\Gamma^{(k)} = \mathsf{Stack}_{\Gamma\to\Gamma}^{(k)}$, i.e. $\mathsf{Hom}_\Gamma^{(0)} = \Gamma \to \Gamma$ and $\mathsf{Hom}_\Gamma^{(k+1)} = [(\mathsf{Hom}_\Gamma^{(k)})^*]$. *Application* is defined inductively as follows.

- $\varphi^{(0)}(\gamma^{(0)})$ is as given
- for $\varphi^{(k+1)} = [\varphi_1^{(k)} \ldots \varphi_s^{(k)}] \in \mathsf{Hom}_\Gamma^{(k+1)}$ and $\gamma^{(k+1)} = [\gamma_1^{(k)} \ldots \gamma_t^{(k)}] \in \mathsf{Stack}_\Gamma^{(k+1)}$ $\varphi^{(k+1)}(\gamma^{(k+1)}) = [\varphi_1^{(k)}(\gamma_1^{(k)}) \ldots \varphi_s^{(k)}(\gamma_1^{(k)}) \cdots \varphi_1^{(k)}(\gamma_t^{(k)}) \ldots \varphi_s^{(k)}(\gamma_t^{(k)})] \in \mathsf{Stack}_\Gamma^{(k+1)}$

Definition 7.2 (k-morphic words).
Let $k \in \mathbb{N}$. An infinite word $w \in \Sigma^\omega$ is *k-morphic* if there is a finite alphabet Γ, an initial k-stack $\gamma^{(k)} = [\cdots[\gamma_0^{(0)}]\cdots] \in \mathsf{Stack}_\Gamma^{(k)}$, a k-morphism $\varphi^{(k)} \in \mathsf{Hom}_\Gamma^{(k)}$ and a terminal homomorphism $h : \Gamma^* \to \Sigma^*$ such that

$$w = h\left(\prod_{n=0}^{\infty} \mathsf{leaves}(\varphi^n(\gamma))\right) .$$

Observe that our morphisms are *uniform* in the sense that each $k+1$-level morphism consists of a fixed number of k-level morphisms which are to be applied to k-stacks in the given order. In particular, $\mathsf{Hom}_\Gamma^{(1)}$ consists of the uniform homomorphisms of Γ^*. Clearly, an infinite word is 0-morphic iff it is *ultimately periodic*, and 1-morphic iff it is *morphic* in the customary sense despite the uniformity restriction on φ, which can be made up for by the choice of h.

Lemma 7.3 (Iteration Lemma). Consider a k-stack $\gamma = [\cdots[\gamma_0]\cdots] \in \mathsf{Stack}_\Gamma^{(k)}$ and a morphism $\varphi = \varphi^{(k)} = \prod_{j_0} \left[\prod_{j_1} \left[\cdots \prod_{j_{k-1}} \left[\varphi^{(0)}_{j_0 \ldots j_{k-1}} \right] \cdots \right]\right] \in \mathsf{Hom}_\Gamma^{(k)}$. Let B be the set of those words $w = j_0 \ldots j_{k-1}$ of length k corresponding to branches of the tree associated to φ, and let $\varphi^{(0)}_u = \varphi^{(0)}_{w_n} \circ \cdots \circ \varphi^{(0)}_{w_1}$ for all words $u = w_1 w_2 \cdots w_n \in B^*$. Then, applying φ n times to γ yields

$$\varphi^n(\gamma) = \underbrace{\prod_{u^{(1)}} [\prod_{u^{(2)}} [\cdots \prod_{u^{(k)}} [}_{u = \otimes_k(u^{(1)}, \ldots, u^{(k)}) \in B^n} \quad \varphi^{(0)}_u(\gamma_0) \quad] \cdots]] \quad .$$

Consider a regular well-ordering $\prec$ of finite binary words and let $u_0 \prec u_1 \prec u_2 \prec \ldots$ be the sequence of words in this ordering. We define the infinite word $w_\prec \in \{0, 1, \#\}^\omega$ as the concatenation of the u_i in ascending order separated by # symbols: $w_\prec = u_0 \# u_1 \# u_2 \# \cdots$. Let $w_{k-\mathrm{llex}}$ be the word thus associated to $<_{k\text{-llex}}$ (restricted to words of length divisible by k). For instance,

$w_{1-\mathrm{llex}} = \#0\#1\#00\#01\#10\#11\#000\#001\#010\#011\#100\# \ldots$
$w_{2-\mathrm{llex}} = \#00\#01\#10\#11\#0000\#0001\#0100\#0101\#0010\#0011\#0110\#0111 \ldots$

Further, let $w_{0-\mathrm{llex}} = \#0\#00\#000\# \ldots$. It is easy to see that $w_{k-\mathrm{llex}} \in \mathcal{W}_{k+1}$ for all $k \in \mathbb{N}$. We say that a sequential transducer S with input alphabet $\{0, 1, \#\}$ and output alphabet Σ is *#-driven* if it is deterministic and in each transition S produces either no output (i.e. the empty string ε) or a single letter output $a \in \Sigma$, but this only on reading a # on the input tape.

Theorem 7.4 (Equivalent Characterizations). *Let Σ be a finite alphabet. For every $k \in \mathbb{N}$ and every ω-word $w \in \Sigma^\omega$ the following are equivalent.*

(1) w is k-morphic

(2) w is k-lexicographic

(3) $w = S(w_{k-\mathrm{llex}})$ for some #-driven sequential transduction S

(4) $W_w \leq^{\mathcal{I}}_{r\mathsf{MSO}} W_{w_{k\text{-}llex}}$ for an $\mathcal{I} = (\varphi_D, <, \{\varphi_a\}_{a \in \Sigma})$ s.t. $\models \forall x(\varphi_D(x) \to P_\#(x))$

Moreover, there are effective translations among these representations.

Proof

(1)⇒(2):(For $k > 0$.) Let $w = h\left(\prod_{n=0}^{\infty} \mathsf{leaves}(\varphi^n(\gamma))\right)$ with $\gamma = [\cdots[\gamma_0]\cdots]$, φ and h as in the definition of k-morphic words. Consider the tree structure of φ, let l be the maximum of the number of children of any of the nodes, and let $B \subseteq [l]^k$ be the set of labels of ordered branches from the root to a leaf, using the natural ordering on $[l]$. We define the *index transition system* of φ as $\mathcal{I}_\varphi = (\Gamma, [l]^k, \delta)$ with $\delta(g, w) = \varphi^{(0)}_w(g)$ for each $g \in \Gamma$ and $w \in B$ and $\delta(g, w)$ undefined otherwise. Note that for uniform morphism of words this definition is identical to that used in the proof of Proposition 4.2. By the Iteration Lemma

$$\mathsf{leaves}(\varphi^n(\gamma)) = \prod_{u \in B^n}^{<_{k\text{-llex}}} \varphi^{(0)}_u(\gamma_0)$$

and, since for each $g \in \Gamma$ the set $P_g = \{u \in B^* \mid \varphi^{(0)}_u(\gamma_0) = g\}$ is obviously accepted by $\mathcal{I}_\varphi$ with initial state γ_0 and single final state g, we can conclude that $(B^*, <_{k\text{-llex}}, \{P_g\}_{g \in \Gamma})$ is a k-lex presentation (in normal form) of $\hat{w} = \prod_{n=0}^{\infty} \mathsf{leaves}(\varphi^n(\gamma)) \in \Gamma^\omega$. By Proposition 4.4, $w = h(\hat{w})$ is also k-lex.

(2)⇒(1): (For $k > 0$.) By the Normal Form Lemma w has a k-lex presentation $(D, <_{k\text{-llex}}, \{P_a\}_{a \in \Sigma})$ in normal form over $\{0, 1\}$, i.e. with D and each P_a being a regular subset of $(\{0, 1\}^k)^*$. Recall $\mathcal{A}_\mathfrak{d}$, $\mathcal{T}\mathfrak{d}$,

$\sigma_{\mathfrak{d}}$, etc. from Section 5. To provide a proof, we only need to adapt the notion of transition morphism to one over k-stacks. The stack alphabet will, of course, be $\Gamma = Q_{\mathfrak{d}}$. We define for each $l \leq k$ and for every $u \in \{0,1\}^{k-l}$ a morphism $\tau_u^{(l)} \in \mathsf{Hom}_\Gamma^{(l)}$ recursively by setting $\tau_u^{(l+1)} = [\tau_{u0}^{(l)}\tau_{u1}^{(l)}]$ for each u of length $k-l-1$, $l < k$, and by setting $\tau_u^{(0)}(\vec{q}) = \delta_{\mathfrak{d}}^*(\vec{q}, u)$ for every $u \in \{0,1\}^k$. Finally, let $\varphi = \tau_\varepsilon^{(k)} = \prod_{j_0=0}^1 \big[\prod_{j_1=0}^1 \big[\cdots \prod_{j_{k-1}=0}^1 \big[\tau_{j_0\cdots j_{k-1}}^{(0)}\big]\cdots\big]\big]$ and $\gamma = [..[\vec{q}_0]..] \in \mathsf{Stack}_\Gamma^{(k)}$. Observe that the structure of φ is the complete binary tree of depth k. Noting that $\tau_{w_n}^{(0)}(\ldots\tau_{w_2}^{(0)}(\tau_{w_1}^{(0)}(\vec{q}))\ldots) = \delta^*(\vec{q}, w_1w_2\ldots w_n)$ the Iteration Lemma yields

$$\varphi^n(\gamma) = \prod_{u^{(1)}=0^n}^{1^n} [\prod_{u^{(2)}=0^n}^{1^n} [\cdots \prod_{u^{(k)}=0^n}^{1^n} [\delta^*(\vec{q}_0, \otimes_k(u^{(1)},\ldots,u^{(k)}))]\cdots]]$$

and we can conclude that $w = \sigma_{\mathfrak{d}}(\prod_{n=0}^\infty \mathsf{leaves}(\varphi^n(\gamma)))$.

(2)⇒(3): (Hint) $\mathcal{S}$ simulates $\mathcal{A}_{\mathfrak{d}}$, restarting on every #.
(3)⇒(4): (Hint) The run of $\mathcal{S}$ is obviously restricted MSO-interpretable.
(4)⇒(2): There is a $k+1$-lex presentation $(\mathfrak{d}, \nu)$ of $w_{k\text{-llex}}$, similar to that given in Example 4.3, such that each maximal factor $u\#$ with $u \in \{0,1\}^*$ is represented on words $x \in D$ satisfying $x' = u$ and with the $k+1^{st}$ component telling the position within $u\#$. Let $\mathcal{I} = (\varphi_D, <, \{\varphi_a\}_{a\in\Sigma})$ be a restricted MSO-interpretation as in (4). By Theorem 3.7 each color-formula φ_a can be transformed into an equivalent automaton $\mathcal{A}_a$. Finally, to obtain a a k-lex presentation of $\mathcal{I}(W_{w_{k\text{-llex}}})$, we construct automata $\mathcal{A}'_a$ accepting those x' such that $x \in L(\mathcal{A}_a)$. □

Questions

(1) Is isomorphism of k-lexicographic words decidable?

(2) Let $k > k'$. Is it decidable whether a k-lex word is k'-lexicographic?
In particular, is eventual periodicity of k-lex word decidable?

On both of these problems, for $k = 1$, see [13] and the references therein.

References

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

[2] V. B. Invariants of automatic presentations and semi-synchronous transductions. In *STACS '06*, volume 3884 of *LNCS*, pages 289–300, 2006.

[3] J. Berstel. *Transductions and Context-Free Languages*. Teubner, Stuttgart, 1979.

[4] A. Blumensath. Automatic structures. Diploma thesis, RWTH-Aachen, 1999.

[5] A. Blumensath and E. Grädel. Finite presentations of infinite structures: Automata and interpretations. *Theory of Computing Systems*, 37:641 – 674, 2004.

[6] V. Bruyère and G. Hansel. Bertrand numeration systems and recognizability. *Theoretical Computer Science*, 181(1):17–43, 1997.

[7] V. Bruyère, G. Hansel, Ch. Michaux, and R. Villemaire. Logic and p-recognizable sets of integers. *Bull. Belg. Math. Soc.*, 1:191 – 238, 1994.

[8] J.W. Cannon, D.B.A. Epstein, D.F. Holt, S.V.F. Levy, M.S. Paterson, and W.P. Thurston. *Word processing in groups*. Jones and Barlett Publ., Boston, MA, 1992.

[9] O. Carton and W. Thomas. The monadic theory of morphic infinite words and generalizations. *Inf. Comput.*, 176(1):51–65, 2002.

[10] C. C. Elgot and M. O. Rabin. Decidability and undecidability of extensions of second (first) order theory of (generalized) successor. *Journal of Symbolic Logic*, 31(2):169–181, 1966.

[11] Ch. Frougny and J. Sakarovitch. Synchronized rational relations of finite and infinite words. *Theoretical Computer Science*, 108:45–82, 1993.

[12] B.R. Hodgson. Décidabilité par automate fini. *Ann. sc. math. Québec*, 7(1):39–57, 1983.

[13] J. Honkala and M. Rigo. A note on decidability questions related to abstract numeration systems. *Discrete Math.*, 285:329–333, 2004.

[14] B. Khoussainov and A. Nerode. Automatic presentations of structures. In *LCC '94*, volume 960 of *LNCS*, pages 367–392. Springer-Verlag, 1995.

[15] B. Khoussainov and S. Rubin. Automatic structures: Overview and future directions. *Journal of Automata, Languages and Combinatorics*, 8(2):287–301, 2003.

[16] B. Khoussainov, S. Rubin, and F. Stephan. Definability and regularity in automatic structures. In *STACS '04*, volume 2996 of *LNCS*, pages 440–451, 2004.

[17] A. Maes. An automata theoretic decidability proof for first-order theory of $\langle \mathbb{N}, <, p \rangle$ with morphic predicate p. *J. Automata, Languages and Comb.*, 4:229–245, 1999.

[18] J.-E. Pin and P. V. Silva. A topological approach to transductions. *Theoretical Computer Science*, 340(1):443–456, 2005.

[19] A. Rabinovich. On decidability of monadic logic of order over the naturals extended by monadic predicates. Unpublished note, 2005.

[20] A. Rabinovich and W. Thomas. Decidable theories of the ordering of natural numbers with unary predicates. Proceedings of Computer Science Logic (CSL '06), Springer, 2006.

[21] M. Rigo and A. Maes. More on generalized automatic sequences. *J. of Automata, Languages and Combinatorics*, 7(3):351–376, 2002.

[22] S. Rubin. Automatic structures. Ph.D. thesis, University of Auckland, NZ, 2004.

[23] W. Thomas. Languages, automata, and logic. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages, volume III*, pages 389–455. Springer, New York, 1997.

Appendix

A Section 3 – Word structures

Lemma 3.4 *Let $\mathfrak{d} = (D, <, \{P_a\}_{a \in \Sigma})$ and ν constitute a canonical presentation of $w \in \Sigma^\omega$. Then for every deterministic Muller automaton $\mathcal{A}$ an automaton recognizing the following set can be effectively constructed.*

$$E_{\mathcal{A}} = \{x \in D \mid w[\nu(x), \infty) \in L(\mathcal{A})\}$$

Proof
Consider $\mathcal{A}$ as a pair (ψ, M) with $M = (Q \to Q, \circ)$ and $\psi \in \mathsf{Hom}(\Sigma^*, M)$. Canonicity of $\mathfrak{d}$ yields automata recognizing $X_q = \{(x,y) \in D^2 \mid x < y \wedge \psi(w[x,y])(q_0) = q\}$ for each $q \in Q$. Using Theorem 2.2 we can construct automata recognizing $Y_F = \{x \in D \mid \bigwedge_{q \in F} \exists^\infty y\, X_q(x,y) \wedge \bigwedge_{q \notin F} \neg\exists^\infty y\, X_q(x,y)\}$ for all $F \subseteq Q$. Finally, $E_{\mathcal{A}}$ is the union of those Y_F such that a run of $\mathcal{A}$ is accepting with F being the set of infinitely often occurring states. The claim follows. □

Lemma 3.6 [Normal Form of MSO formuli over word structures]
Every MSO formula $\varphi(\vec{x})$ having free first-order variables $x_0, \ldots, x_{n-1}$ and no free second-order variables is equivalent to a boolean combination of formuli $x_i < x_j$ and relativized MSO sentences $\vartheta^{[0,x_i]}$, $\vartheta^{[x_i,x_j]}$, and $\vartheta^{[x_i,\infty)}$ with $i, j \in [n]$.
Proof
We present a proof through automata. Via standard construction, there is a deterministic Muller automaton $\mathcal{A}$ over the alphabet $\Sigma \times \{0,1\}^n$ such that $W_w \models \varphi(\vec{k})$ iff $w \otimes \xi_{\vec{k}} \in L(\mathcal{A})$ for all $\vec{k} \in \mathbb{N}^n$, where $\xi_{\vec{k}} \in (\{0,1\}^n)^\omega$ is the characteristic word of the tuple $\vec{k}$, i.e. $\xi_{\vec{k}}[i]_j = 1$ iff $k_j = i$. We collect for each pair of states (p,q) of $\mathcal{A}$ the regular language $L_{p,q} = \{u \in \Sigma^* \mid \delta^*(p, u \otimes (0^n)^{|u|}) = q\}$. Additionally, we let $L_q = \{u \in \Sigma^\omega \mid \mathcal{A} \text{ accepts } u \otimes (0^n)^\omega \text{ from state } q\}$. Again, by standard constructions, we find MSO sentences $\vartheta_{p,q}$ respectively ϑ_q defining these languages.
Each infinite word $w \otimes \xi_{\vec{k}}$ is naturally factored into segments in between consecutive k_i's, some of which can be equal. Accordingly, each run of $\mathcal{A}$ can be factored into finite number of finite segments and an infinite segment by those positions where in at least one of the last n components of the symbol read a 1 is encountered. The intermediate segments and the last infinite segment are models of the appropriate sentences $\vartheta_{p,q}$ and of ϑ_q respectively.
By summing up all possible factorizations of accepting runs we obtain in a first attempt a boolean combination of formuli of type $x_i < x_j$, $x_i = x_j$, $P_a x_i$ and of relativized sentences of the form $\vartheta_{q_0,q}^{[0,x_i)}$, $\vartheta_{p,q}^{(x_i,x_j)}$ and $\vartheta_q^{(x_i,\infty)}$. Equality can be expressed using $<$, and integrating the $P_a x_i$ into the neighboring openly relativized segment formulae we finally arrive at a normal form as promised. □

B Section 4 – k-lexicographic presentations

In preparation to proving Proposition 4.2 we introduce a few definitions.

Automata and morphisms

To each morphism $\varphi \in \mathsf{Hom}(\Sigma^*, \Sigma^*)$ with $|\varphi| = l$ we associate its *index transition system* $\mathcal{I}_\varphi = (\Sigma, [l], \delta)$ where $\delta(a,i) = \varphi(a)[i]$ for every $i < |\varphi(a)|$ and undefined otherwise. For each $a \in \Sigma$ considered as the initial state, the DFA $(\mathcal{I}_\varphi, a, \Sigma)$ accepts the set $I(a) = I_\varphi(a)$ of valid sequences of indices starting from a. Applying φ n times to a gives the word

$$\varphi^n(a) = \prod_{x \in I(a) \cap [l]^n}^{lex} \delta^*(a, x) \tag{2}$$

where x is meant to run through all valid sequences of indices of length n in lexicographic order. Thus $\varphi^n(a)$ is the sequence of labels of the n^{th} level of the tree unfolding of $\mathcal{I}_\varphi$ from a

Conversely, given a linear ordering $a_0 < a_1 < \ldots < a_s$ of Σ we associate to each DTS $\mathcal{T} = (Q, \Sigma, \delta)$ its *transition morphism* $\tau = \tau_\mathcal{T} \in \mathsf{Hom}(Q^*, Q^*)$ defined as $\tau(q) = \delta(q, a_{i_1})\delta(q, a_{i_2})\ldots\delta(q, a_{i_k})$ where $a_{i_1} < a_{i_2} < \ldots < a_{i_k}$ are precisely those symbols for which a transition from q is defined. Just as in (2) applying τ n times to some q results in $\tau^n(q) = \prod_{w \in L(\mathcal{T},q,Q) \cap \Sigma^n}^{lex} \delta^*(q, w)$, where w runs through, in lexicographic order, all words of length n, which are labels of some path in $\mathcal{T}$ starting from q. Thus $\tau^n_\mathcal{T}(q)$ is the sequence of labels of the n^{th} level of the tree unfolding of $\mathcal{T}$ from q.

Proposition 4.2 ([21]) *$\mathcal{W}_1$ is the class of morphic words.*

Proof

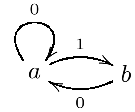
The proof is essentially that of [21], only slightly more compact.

Let $\tau \in \mathsf{Hom}(\Gamma^*, \Gamma^*)$ be prolongable on a and consider its index transition system $\mathcal{I} = \mathcal{I}_\tau$. It is clear from our previous observations that the language $L(\mathcal{I}, a, \Gamma)$ recognized by $\mathcal{I}$ with all states final and a as its initial state provides, equipped with the prefix-ordering, an automatic presentation of the tree unfolding $\mathcal{T} = \mathcal{T}_{\mathcal{I},a}$ of $\mathcal{I}$ from the initial state a. As also remarked, $\tau^n(a)$ is precisely the word one obtains by reading the n^{th} level of $\mathcal{T}$ from "left to right", i.e. in lexicographic order. Also note that τ being prolongable on a, $\mathcal{I}_\tau$ contains a transition $a \overset{0}{\mapsto} a$, therefore the subtree of $\mathcal{T}$ rooted at 0 is isomorphic to the whole tree. Let $\tau(a) = au$ for some $u = u_1 \ldots u_t \in \Gamma^*$ and let $\mathcal{U}_i$ be the subtree rooted at $0 < i \leq t$. Then $\tau^{n+1}(a) = au\tau(u)\cdots\tau^n(u) = \tau^n(a)\cdot\tau^n(u)$ and $\mathcal{T} \cong a(\mathcal{T}, \mathcal{U}_1, \ldots, \mathcal{U}_t)$. To obtain a length-lexicographic presentation of $\tau^\omega(a)$ we dispense with the subtree rooted at 0 so that the levels of the remaining regular tree $a(\mathcal{U}_1, \ldots, \mathcal{U}_t)$ correspond to the increments $\tau^n(u)$ between iterations of τ. We have thus shown that $D = L(\mathcal{I}_\tau, a, \Gamma)\setminus 0[|\tau|]^*$ and $P_c = L(\mathcal{I}_\tau, a, c)\setminus 0[|\tau|]^*$ for each $c \in \Gamma$ together with the natural length-lexicographic ordering provide an automatic presentation of $\tau^\omega(a)$. To give a lexicographic presentation of $w = \sigma(\tau^\omega(a))$ where $\sigma \in \mathsf{Hom}(\Gamma^*, \Sigma^*)$ we set $D' = \{xi \mid c \in \Gamma, x \in P_c, i < |\sigma(c)|\}$ and $P_b = \{xi \mid c \in \Gamma, x \in P_c, \sigma(c)[i] = b\}$ for each $b \in \Sigma$.

Conversely, given a lexicographic presentation $(\mathcal{A}_D, <_{\text{lex}}, \{\mathcal{A}_{P_a}\}_{a\in\Sigma})$ of some w consider the product automaton $\mathcal{A} = \prod_{a\in\Sigma}\mathcal{A}_{P_a}$. Let $\tau = \tau_\mathcal{A}$ be its transition morphism, and let us define $\sigma \in \mathsf{Hom}(Q(\mathcal{A})^*, \Sigma^*)$ by stipulating that $\sigma(\vec{q}) = a$ whenever the a^{th} component of $\vec{q}$ is an accepting state of $\mathcal{A}_{P_a}$ (clearly, a is then uniquely determined) and $\sigma(\vec{q}) = \varepsilon$ when no such a exists. To ensure that τ is prolongable, we introduce a new symbol $\vec{q}_0' \notin Q(\mathcal{A})$ and set $\tau(\vec{q}_0') = \vec{q}_0'\tau(\vec{q}_0)$ and $\sigma(\vec{q}_0') = \sigma(\vec{q}_0)$, where $\vec{q}_0$ is the initial state of $\mathcal{A}$. We leave it to the reader to check that $w = \sigma(\tau^\omega(\vec{q}_0'))$. □

Example B.1. Recall the Fibonacci word generated by the morphism $\phi : a \mapsto ab, b \mapsto a$ of Example 3.2. The index transition system of ϕ,

$\mathcal{I}$:



accepts, with a being initial and both states being final, the language $\{0,1\}^*\setminus\{0,1\}^*11\{0,1\}^*$ of Fibonacci numerals *with* leading zeros. The construction of the proof of Proposition 4.2 dispenses precisely with those numerals starting with a zero, thus producing an injective presentation. Length-lexicographically ordered, the first few numerals are $\varepsilon, 1, 10, 100, 101, 1000, 1001, 1010, 10000, 10001, \ldots$ with 10^n representing the n^{th} Fibonacci number.

Proposition 4.4 [Closure under homomorphic mappings] *The class of automatically presentable ω-words is closed under homomorphic mappings. In particular, if w is k-lexicographic, then so is $h(w)$ for every homomorphism h.*

Proof

The idea is to append each word $x \in P_a$ of a given presentation of w indexing a symbol a by $|h(a)|$ many appropriately chosen suffixes $u_{a,i}$ with $i < |h(a)|$. For k-lexicographic presentations we choose $|u_{a,i}| = k$ and take care that differences fall within the k^{th} component of the k-split of $xu_{a,i}$. □

Lemma 4.5 [Normal Form Lemma] *Let $1 < k \in \mathbb{N}$. Each k-lexicographic presentation $\mathfrak{d} = (D, <_{k\text{-}llex})$ of $(\mathbb{N}, <)$ over an alphabet Σ is equivalent to one $\mathfrak{d}' = (D', <_{k\text{-}llex})$ over some Γ such that $D' \subseteq (\Gamma^k)^*$. In fact, one can choose $\Gamma = \{0, 1\}$ in the above.*

Proof

Let first $\Gamma = \Sigma \uplus \{\widehat{0}, \ldots, \widehat{k-1}, \diamond\}$ endowed with the ordering $\diamond < \widehat{0} < \ldots < \widehat{k-1} < a_1 < \ldots < a_s$ where $a_1 < \ldots < a_s$ is the ordering of Σ used in the presentation $\mathfrak{d}$. We define the translation $t : \Sigma^* \to (\Gamma^k)^*$ padding each word x to $t(x) = \widehat{l}\,\diamond^{k-1}\, x \diamond^{k-l}$ where $l = |x| \bmod k$. Observe that the moduli of the positions of symbols of x are preserved in the process of this coding, i.e. $t(x)^{(i)} = \alpha x^{(i)} \diamond$ with α being $\widehat{l}$ for $i = 0$ and $\diamond$ otherwise. Consequently $x <_{k\text{-llex}} y$ iff $t(x) <_{k\text{-llex}} t(y)$ in the orderings induced by that of the symbols. Since t is a synchronized rational bijection $\mathfrak{d}' = (t(D), <_{k\text{-llex}})$ is equivalent to $\mathfrak{d}$.

Finally, to obtain an equivalent presentation over $\{0, 1\}$ take any binary coding $a \mapsto b_0 \ldots b_{l-1}$ of the symbols $a \in \Gamma$ uniformly of length l and such that $a < a'$ iff $b_0 \ldots b_{l-1} <_{1\text{-llex}} b'_0 \ldots b'_{l-1}$. Extend this into a coding of blocks of k consecutive symbols as $a^0 \ldots a^{k-1} \mapsto b_0^0 \ldots b_0^{k-1} \ldots b_{l-1}^0 \ldots b_{l-1}^{k-1}$, and extend this homomorphically to $(\Gamma^k)^*$. Due to the uniformity requirement, this translation is semi-synchronous, further it respects the k-lexicographic ordering, thus providing an equivalent k-lexicographic presentation. □

C Section 5 – Canonicity, Closure and Decidability

C.1 Technical tools: automata transformations

Consider a finite deterministic transition system $\mathcal{T} = (Q, \Sigma, \delta)$ and the associated pair (M, φ) consisting of the finite monoid $M = (Q \to Q, \circ)$ and the homomorphism $\varphi \in \mathsf{Hom}(\Sigma^*, M)$ induced by δ. We call $\mathsf{Hom}(\Sigma^*, M)$ the *derived state space* and denote it by $Q^{(\Sigma)}$. Furthermore, we call $M^{(\Sigma)} = Q^{(\Sigma)} \to Q^{(\Sigma)}$ the monoid of *automata transformations*. Note that both $Q^{(\Sigma)}$ and $M^{(\Sigma)}$ are finite. This terminology is justified by the fact that $Q^{(\Sigma)} = \mathsf{Hom}(\Sigma^*, M)$ is in essence the set of all Σ-labelled DTS's over the state space Q, hence $M^{(\Sigma)}$ is indeed the monoid of all transformations of such transition systems.

A particular submonoid of $M^{(\Sigma)}$ that interests us is that of *inverse homomorphic transformations* $H^{(\Sigma)}$ defined as follows. Consider a homomorphism $h \in \mathsf{Hom}(\Sigma^*, \Sigma^*)$. We can associate to h the element $\Phi(h)$ of $M^{(\Sigma)}$ defined as $(Q^{(\Sigma)} \ni \chi \mapsto \chi \circ h)$. It can be readily seen that Φ is a monoid homomorphism from $\mathsf{Hom}(\Sigma^*, \Sigma^*)$ to $M^{(\Sigma)}$, therefore $H^{(\Sigma)} \stackrel{\text{def}}{=} \Phi(\mathsf{Hom}(\Sigma^*, \Sigma^*))$ is a submonoid of $M^{(\Sigma)}$. In terms of automata transformations this amounts to mapping a transition function δ to δ' such that $\delta'(q, a) = q'$ whenever $\delta^*(q, h(a))$, where δ^* denotes as usual the extension of δ to all words over Σ. We let $h^{-1}(\mathcal{T})$ denote the transition system (Q, Σ, δ'). Thus, for every $q, q' \in Q$ and $w \in \Sigma^*$ there is a path in $h^{-1}(\mathcal{T})$ labelled w from q to q' iff there is a path in $\mathcal{T}$ labelled $h(w)$ from q to q'.

Consider a finite alphabet Θ and a mapping $\vartheta : \Theta \to \mathsf{Hom}(\Sigma^*, \Sigma^*)$. We extend ϑ to Θ^* according to the rule

$$\vartheta(x \cdot x') = \vartheta(x') \circ \vartheta(x) \tag{3}$$

which ensures that $\Phi_\vartheta = \Phi \circ \vartheta$ is a homomorphism from Θ^* to $H^{(\Sigma)}$.

$$\begin{aligned} \Phi(\vartheta(x \cdot x'))(\chi) &= \Phi(\vartheta(x') \circ \vartheta(x))(\chi) = \chi \circ \vartheta(x') \circ \vartheta(x) \\ &= \Phi(\vartheta(x'))(\chi) \circ \vartheta(x) = \Phi(\vartheta(x))(\Phi(\vartheta(x'))(\chi)) \\ &= (\Phi(\vartheta(x)) \circ \Phi(\vartheta(x')))(\chi) \quad . \end{aligned}$$

Therefore, the pair $(H^{(\Sigma)}, \Phi_\vartheta)$ represents, in accordance with our initial correspondence, a Θ-labelled finite transition system with state space $Q^{(\Sigma)}$. Elements of Θ^* can thus be seen as words over Θ, or, via ϑ as homomorphisms from Σ^* to Σ^*, or, via Φ_ϑ, as transformations of Σ-labelled transition systems. Given a word $w \in \Sigma^*$ and a monoid element $m \in M$, we are interested in the following subset of Θ^*.

$$L_{\mathcal{T},w,m,\vartheta} = \{x \in \Theta^* \mid \text{the state transformation induced by } w \text{ in } \vartheta(x)^{-1}(\mathcal{T}) \text{ is } m\}$$

Let $n = |Q|$. Since $Q^{(\Sigma)}$ is finite, and the kernel $\mathsf{Ker}(\varphi)$ of every homomorphism $\varphi \in Q^{(\Sigma)}$ is a congruence (wrt. concatenation) of finite index, their intersection

$$\sim_n \stackrel{\text{def}}{=} \bigcap_{\varphi \in Q^{(\Sigma)}} \mathsf{Ker}(\varphi) = \{(u, u') \mid \forall \varphi \in \mathsf{Hom}(\Sigma^*, M)\ \varphi(u) = \varphi(u')\}$$

is again a congruence of finite index, i.e. the factor monoid $\widetilde{Q} = \Sigma^* / \sim_n$ is finite. Note that this equivalence depends only on the size of Q, hence the notation. Intuitively, $u \sim_n u'$ iff there is no automaton having at most n states that could distinguish u from u'. This equivalence can be used to define the Hall metric on Σ^* giving rise to a compact Hausdorff topology (cf. [18]), which is essentially what one obtains from the analogous equivalences wrt. MSO formulae of restricted quantifier ranks (cf. [19],[20]).

Clearly, every homomorphism $h \in \mathsf{Hom}(\Sigma^*, \Sigma^*)$ preserves $\sim_n$-classes, and thus determines a function $\widetilde{h} : \widetilde{Q} \to \widetilde{Q}$. It is again routine to check that $\Psi : h \mapsto \widetilde{h}$ thus defined is a homomorphism from $(\mathsf{Hom}(\Sigma^*, \Sigma^*), \circ)$ into $\widetilde{M} = (\widetilde{Q} \to \widetilde{Q}, \circ)$. Furthermore, each $\sim_n$ determines an equivalence of homomorphisms $h, h' \in \mathsf{Hom}(\Sigma^*, \Sigma^*)$ defined as follows.

$$\begin{aligned} h \sim_n h' \quad &\stackrel{\text{def}}{\iff} \forall w \in \Sigma^* \quad h(w) \sim_n h'(w) \\ &\iff \forall a \in \Sigma \quad h(a) \sim_n h'(a) \end{aligned} \tag{4}$$

The Hall metric on Σ^* thus induces a similar metric, thereby determining a compact Hausdorff topology, on $\mathsf{Hom}(\Sigma^*, \Sigma^*)$. Moreover, the following equivalence

$$\Phi(h_1) = \Phi(h_2) \iff h_1 \sim_n h_2 \iff \widetilde{h_1} = \widetilde{h_2} \tag{5}$$

can easily be checked to hold:

$$\begin{aligned} & \Phi(h_1) = \Phi(h_2) \\ \iff\ & \forall \chi \in \mathsf{Hom}(\Sigma^*, M) :\ \chi \circ h_1 = \chi \circ h_2 \\ \iff\ & \forall w \in \Sigma^*\, \forall \chi \in \mathsf{Hom}(\Sigma^*, M) :\ \chi(h_1(w)) = \chi(h_2(w)) \\ \iff\ & \forall w \in \Sigma^* :\ h_1(w) \sim_n h_2(w) \\ \iff\ & h_1 \sim_n h_2 \quad . \end{aligned}$$

Lemma C.1 (Higher-Order Regularity (HOR) Lemma).
For every $\mathcal{T} = (Q, \Sigma, \delta)$ with associated (M, φ) and for every $w \in \Sigma^*$, $m \in M$, and every Θ and ϑ as above we can construct an automaton recognizing $L_{\mathcal{T},w,m,\vartheta}$.

Proof
Observe that we can write $L_{\mathcal{T},w,m,\vartheta}$ equivalently as

$$\begin{aligned}
L_{\mathcal{T},w,m,\vartheta} &\stackrel{\text{def}}{=} \{x \in \Theta^* \mid \text{the state transformation induced by } w \text{ in } \vartheta(x)^{-1}(\mathcal{T}) \text{ is } m\} \\
&= \{x \in \Theta^* \mid \text{the state transformation induced by } \vartheta(x)(w) \text{ in } \mathcal{T} \text{ is } m\} \\
&= \{x \in \Theta^* \mid \varphi(\vartheta(x)(w)) = m\} \\
&= \{x \in \Theta^* \mid \Phi(\vartheta(x))(\varphi)(w) = m\} \\
&= \{x \in \Theta^* \mid \Phi(\vartheta(x)) \in H_{m,\varphi,w}\} \\
&= \Phi_\vartheta^{-1}(H_{m,\varphi,w})
\end{aligned}$$

where $H_{m,\varphi,w} = \{\xi = \Phi(h) \in H^{(\Sigma)} \mid \xi(\varphi)(w) = \varphi(h(w)) = m\}$. Hence, $L_{\mathcal{T},w,m,\vartheta}$ is recognized by the subset $H_{m,\varphi,w}$ of the finite monoid $H^{(\Sigma)}$ under the morphism Φ_ϑ. Moreover, $H_{m,\varphi,w}$ can be determined, according to (5), by enumerating all $\sim_n$-classes of homomorphisms. Using the correspondence (4) this can be reduced to enumerating $\sim_n$-classes of words over Σ. □

C.2 Canonicity of k-lexicographic presentations

Our objective is to prove the Contraction Lemma using the HOR Lemma. To this end we generalize the notion of transition morphisms. Wlog. the ordered alphabet Γ of the presentation $\mathfrak{d}$ is $[t] = 0 < 1 < \ldots < t-1$. Let $Q = Q_\mathfrak{d} \times \{l, r, b, n\}$ (standing for left, right, both and none respectively) and $\pi : ((q, x) \mapsto q)$ be the projection onto the first component. We define the mapping $\beta : ([t]^k([t] \times [t]))^* \to \mathsf{Hom}(Q^*, Q^*)$ via homomorphic extension as in (3) while stipulating that

$$\begin{aligned}
\beta_{u(i,j)}(q, n) &= (\delta_\mathfrak{d}^*(q, u0), n)(\delta_\mathfrak{d}^*(q, u1), n) \ldots (\delta_\mathfrak{d}^*(q, u(t-1)), n) \\
\beta_{u(i,j)}(q, l) &= (\delta_\mathfrak{d}^*(q, ui), l)(\delta_\mathfrak{d}^*(q, u(i+1)), n) \ldots (\delta_\mathfrak{d}^*(q, u(t-1)), n) \\
\beta_{u(i,j)}(q, r) &= (\delta_\mathfrak{d}^*(q, u0), n) \ldots (\delta_\mathfrak{d}^*(q, u(j-1)), n)(\delta_\mathfrak{d}^*(q, uj), r) \\
\beta_{u(i,j)}(q, b) &= \begin{cases} \varepsilon & \text{for } i > j \\ (\delta_\mathfrak{d}^*(q, ui), b) & \text{for } i = j \\ (\delta_\mathfrak{d}^*(q, ui), l)(\delta_\mathfrak{d}^*(q, u(i+1)), n) \ldots (\delta_\mathfrak{d}^*(q, u(j-1)), n)(\delta_\mathfrak{d}^*(q, uj), r) \end{cases}
\end{aligned}$$

where u ranges over Γ^k and $i, j < t$. We regard β as a mapping from pairs of $=_k$-equivalent words $x, y \in D$. Indeed, each pair (x, y) of words with $x' = y'$ determines a sequence $u_1(i_1, j_1) \ldots u_n(i_n, j_n)$, and vice versa, such that $x^{(k+1)} = i_1 \ldots i_n$, $y^{(k+1)} = j_1 \ldots j_n$ and $x' = y' = u_1 \ldots u_n$. In accordance with (3) we can thus define $\beta_{x,y}$ as the composition $\beta_{u_n(i_n,j_n)} \circ \cdots \circ \beta_{u_1(i_1,j_1)}$. We further let $\tau_u = \beta_{u(0,t-1)}$. Note that, for $k = 0$, τ_ε is essentially the transition morphism τ associated to $\mathcal{T}_\mathfrak{d}$ as defined above. To allow for uniform treatment we set $\tau_{1^n} = {\tau_\varepsilon}^n$ when $k = 0$.

Lemma C.2. For all $k \in \mathbb{N}$ and $x, y \in (\Gamma^{k+1})^*$ such that $x' = y'$ and $x \leq_{k+1\text{-llex}} y$:

$$\pi(\beta_{x,y}(\vec{q}, b)) = \prod_{z=x}^{y} \delta_\mathfrak{d}^*(\vec{q}, z)$$

where the concatenation product is taken over the values of z in the $(k+1)$-lexicographic ordering. Consequently, when in addition $x, y \in D$ then we have

$$\begin{aligned}
\sigma_\mathfrak{d}(\pi(\beta_{x,y}(\vec{q}_0, b))) &= w[\nu(x), \nu(y)] \\
\sigma_\mathfrak{d}(\pi(\tau_{x'}(\vec{q}_0, b))) &= w[\nu(c(x) \cap D)]
\end{aligned}$$

Using the machinery introduced in Section C.1 and the HOR Lemma, the Contraction Lemma is now easily established.

Lemma 5.1 [Contraction Lemma]
Let $\mathfrak{d} = (D, <_{k+1\text{-}llex}, \{\mathcal{A}_a\}_{a\in\Sigma})$ be a $(k+1)$-lex presentation with coordinate function ν of the word structure of an ω-word $w \in \Sigma^\omega$. Then for every finite monoid M, every $\psi \in \mathsf{Hom}(\Sigma^*, M)$ and for each $m \in M$ the following relations are regular.

$$\begin{array}{rcl} B'_m & = & \{(x,y) \in D^2 \mid x \leq_{k+1\text{-}llex} y \wedge x =_k y \wedge \psi(w[\nu(x),\nu(y)]) = m\} \\ P'_m & = & \{x' \in D' \mid \psi(w[\nu(c(x) \cap D)]) = m\} \end{array}$$

Whence, $(D', <_{k\text{-}llex}, \{P'_m\}_{m\in M})$ is a k-lexicographic presentation of $c^\psi_{\mathfrak{d}}(w)$.
Proof
From Lemma C.2 we know that $\psi(w[\nu(x),\nu(y)]) = \psi(\sigma_{\mathfrak{d}}(\pi(\beta_{x,y}(\vec{q}_0, b))))$ and that $\psi(w[\nu(c(x))]) = \psi(\sigma_{\mathfrak{d}}(\pi(\tau_{x'}(\vec{q}_0, b))))$. Recall that $\beta_{x,y}$ was defined as $\beta_{u_n(i_n,j_n)} \circ \cdots \circ \beta_{u_1(i_1,j_1)}$ for all $x' = y' = u_1 \dots u_n$ with $u_i \in [t]^k$ and $x^{(k+1)} = i_1 \dots i_n$, $y^{(k+1)} = j_1 \dots j_n$. Similarly, $\tau_{x'} = \tau_{u_n} \circ \cdots \circ \tau_{u_1}$. The results are established by applying the HOR Lemma with $\varphi = \psi \circ \sigma_{\mathfrak{d}} \circ \pi$ and $\Theta = [t]^k([t] \times [t])$, $\vartheta_{x\otimes y} = \beta_{x,y}$ in the first case, respectively with $\Theta = [t]^k$, $\vartheta_{x'} = \tau_{x'}$ in the second case. □

Corollary 5.5 [Closure under d.g.s.m. mappings] *For each $k \in \mathbb{N}$ the class $\mathcal{W}_k$ is closed under deterministic generalized sequential mappings.*
Proof
Every deterministic sequential transduction $S(w)$ of a word w can be obtained by a homomorphic mapping of the run of $S \times \mathcal{B}^1_\Sigma$ over w, where $\mathcal{B}^1_\Sigma$ is the De Bruin transition system with memory of the single last symbol of Σ read. The homomorphism is just the output function of the sequential transducer S. The run of S on w is of course rMSO interpretable in W_w, is thus in $\mathcal{W}_k$, and closure under homomorphic mappings holds by Proposition 4.4. □

Theorem 5.6 [Automatic equivalence structures] *Consider $\mathfrak{A} = (A, E)$ with E an equivalence relation on a countably infinite set A having only finite equivalence classes. Assume further that for each n there are $f(n) \in \mathbb{N}$ many equivalence classes of size n.*
Then $\mathfrak{A} \in$ AutStr if and only if there is a 2-lex word $w = 0^{m_0}10^{m_1}10^{m_2}1\dots$ such that $f(n) = |\{i \mid m_i = n\}|$, in which case $\mathfrak{A} \leq^{\mathcal{I}}_{\mathsf{FO}} W_w$ for a fixed one-dimensional FO-interpretation $\mathcal{I}$, also implying that $\mathrm{Th}_{\mathsf{MSO}}(\mathfrak{A})$ is decidable.
Proof
For the "if" direction, the interpretation in question is $\mathcal{I} = (\varphi_A(x), \varphi_E(x,y))$ with $\varphi_A(x) = P_0(x)$ and $\varphi_E(x,y) = \varphi_A(x) \wedge \varphi_A(y) \wedge \forall z(x < z < y \vee y < z < x \rightarrow P_0(z))$. It is now easy to check that $\mathcal{I}(W_w)$ is indeed isomorphic to $\mathfrak{A}$ and is thus, by Theorem 2.2 or by Corollary 5.4, automatic.
For the "only if" direction we construct, given an a.p. (L_A, L_E) of $\mathfrak{A}$, an a.p. of a binary word with the claimed property. First observe that since all equivalence classes of $\mathfrak{A}$ are finite, there is a constant C such that for all $x, y \in L_A$ with $(x,y) \in L_E$ $||x| - |y|| < C$. We can therefore easily construct by padding an equivalent presentation of $\mathfrak{A}$ in which $|x| = |y|$ holds for all x,y representing equivalent elements. We shall now assume this holds. Let Γ be the alphabet of the presentation of $\mathfrak{A}$. Wlog. $\Gamma = \{0, \dots, s-1\}$. The alphabet of the presentation of w will be $\Gamma' = \{0, \dots, s-1, s\}$ ordered naturally. We set $P_0 = \{\otimes_2(x,y) \mid (x,y) \in L_E \wedge \forall (x,z) \in L_E\ x \leq_{\text{lex}} z\}$, $P_1 = \{\otimes_2(x, s^{|x|}) \mid \forall (x,z) \in L_E\ x <_{\text{lex}} z\}$, and $D = P_0 \cup P_1$. It is now clear that $(D, <_{\text{2-llex}}, P_0, P_1)$ is an a.p. as promised. □

D Section 7 – Equivalent characterizations

The Iteration Lemma, used in the proof of Theorem 7.4 is a generalization of (2) and can be proved by a technical but straightforward induction on the number of iterations. Let us finish by making two final remarks.

Remark D.1 (On the irrelevance of uniformity). Observe that the transformation of a level k morphism into a k-lexicographic presentation in the proof of Theorem 7.4 does not rely on uniformity of the morphism. One can define non-uniform morphisms of *tagged k-stacks*, that is a kind of deterministic k-level indexed grammars commuting with concatenation on each level. The same transformation can be used to show that these systems give rise to the same class of ω-words. Thus, uniformity is really no restriction in terms of generating power as long as we allow ourselves to apply an arbitrary homomorphism in the final step.

Remark D.2 (On morphic predicates, cf. [17]). Our definition of morphisms of k-stacks not only resembles that of k-dimensional "pictures", but is essentially identical with that, up to a natural coding. Indeed, k-dimensional pictures are k-stacks satisfying the uniformity condition that every level $l+1$ substack consists of exactly the same number n_{l+1} of l-stacks, where $(n_1, \dots, n_k)$ are the dimensions of the picture. Due to their above mentioned uniformity our morphisms preserve uniformity of stacks. Hence, morphisms of k-stacks and morphisms of k-dimensional pictures are easily seen to be one and the same, up to this coding.