Theory Comput Syst (2014) 55:856–880
DOI 10.1007/s00224-013-9508-6

Model-Theoretic Properties of ω-Automatic Structures

Faried Abu Zaid · Erich Grädel · Łukasz Kaiser · Wied Pakusa

Published online: 17 October 2013

Abstract We investigate structural properties of ω-automatic presentations of infinite structures in order to sharpen our methods to determine whether a given structure is ω-automatic. We apply these methods to show that several classes of structures such as pairing functions and infinite integral domains do not have an ω-automatic model.

Keywords Logic · Algorithmic model theory · Automatic structures

1 Introduction

Automatic structures are (in general) infinite structures that admit a finite presentation by automata. Informally, an automatic presentation of a relational structure $\mathfrak{B} = (B, R_1, \dots, R_m)$ consists of a language L, which must be recognizable by an automaton $\mathcal{A}$, and a surjective function $\pi : L \to B$ that associates every word of L with the element of B that it represents. The function π must be surjective (every element of B is named by some word in L) but need not be injective (elements may have more than one name). In addition it must be recognizable by automata, reading

F.A. Zaid (✉) · E. Grädel · W. Pakusa
Mathematische Grundlagen der Informatik, RWTH Aachen University, Aachen, Germany
e-mail: abuzaid@logic.rwth-aachen.de

E. Grädel
e-mail: graedel@logic.rwth-aachen.de

W. Pakusa
e-mail: pakusa@logic.rwth-aachen.de

Ł. Kaiser
CNRS & LIAFA, Université Paris Diderot—Paris 7, Paris, France
e-mail: kaiser@liafa.jussieu.fr

their inputs synchronously, whether two elements of L name the same element of B, and, for each relation R_i, whether a given tuple of words in L names a tuple in R_i. Together, the automata $\mathcal{A}$ and the automata that recognise equality and the relations $R_1, \ldots, R_m$ provide a finite representation of the structure $\mathfrak{B}$.

In principle we can use automata over finite words, infinite words, finite trees, or infinite trees to obtain different classes of automatic structures. All of these automata models are effectively closed under first-order operations (union, intersection, complementation, and projection) and their emptiness problem is decidable. Indeed, these two properties ensure that

– every automatic structure has a decidable first-order theory, and more generally
– given any automatic presentation of $\mathfrak{A}$ and any first-order formula $\phi(x_1, \ldots, x_k)$ one can effectively construct an automaton respresenting the relation $\phi^{\mathfrak{A}} := \{\bar{a} \in A^k : \mathfrak{A} \models \phi(\bar{a})\}$.

Thus, all (first-order) definable properties of automatic structures can be algorithmically investigated using automata-theoretic methods based on appropriate finite presentations. This makes automatic structures a domain of considerable interest for computer science.

While the case of word-automatic structures (with presentations based on automata operating on finite words) is reasonably well understood [3, 5, 8], much less is known for presentations based on other classes of automata. In particular the available methods to analyse whether a given structure admits an automatic presentation on, say, infinite words, or to classify all such structures inside a given domain, are still relatively weak. For countable structures, an important achievement is the result by Kaiser, Rubin, and Bárány [7] that a countable structure is ω-automatic if, and only if, it is word-automatic. Thus the recent result by Tsankov [17] that the additive group of the rationals is not automatic immediately implies that it is not ω-automatic either.

One of the most prominent and important structures with a decidable first-order theory is certainly the field of reals $(\mathbb{R}, +, \cdot)$. The decidability goes back to Tarski [15] and is based on a quantifier elimination argument. Therefore, it is very natural to ask whether the field of reals admits an automatic presentation. Of course, such a presentation cannot be based on automata on finite words (or finite trees) because languages of finite words and trees are countable. However, it might be the case that the field of reals is ω-automatic, i.e., admits a presentation based on automata on infinite words, or that it is ω-tree-automatic, with a presentation based on automata on infinite trees.

The question whether this is the case is closely related to classical problems raised by Büchi and Rabin in the context of decidable theories such as Presburger arithmetic and the theory of the field of reals. The decidability of Presburger arithmetic, the first-order theory of $(\mathbb{N}, +)$, has originally been proven by quantifier elimination, but Büchi's automata based proof of the decidability of S1S (the monadic theory of infinite words) immediately carries over to an automata-theoretic decidability argument for Presburger arithmetic. In Rabin's classical paper [13], where he proved the decidability of S2S (the monadic theory of the infinite binary tree) and several other theories, he explicitly raised the question whether also the decidability of the field of reals could be proved by automata-theoretic methods.

It is quite easy to see that both reducts $(\mathbb{R}, +)$ and $(\mathbb{R}, \cdot)$ of the field of reals are ω-automatic. A presentation for $(\mathbb{R}, +)$ is obtained by using a slightly modified binary encoding of the reals and an implementation of bitwise addition with carry by an automaton. To represent $(\mathbb{R}, \cdot)$ one can use that $(\mathbb{R}, +) \cong (\mathbb{R}_+, \cdot)$ via the exponential function. However, it has so far been open whether two such presentations could be combined to one of the entire field. We shall prove that this is not the case. More generally we investigate structural restrictions on the complexity of ω-automatic relations in order to prove that certain classes of structures do not have ω-automatic models.

In the study of ω-automatic presentations, it is important to distinguish between injective and non-injective presentations. Injective presentations are much easier to work with, since we do not need an automaton to determine whether two words encode the same element. In the case of word-automatic structures it is not hard to see that every such structure admits an injective automatic presentation. However, it had been open for some time whether ω-automatic structures always admit injective presentations, until Hjorth, Khoussainov, Montalbán and Nies [6] described an ω-automatic structure that does not even permit an injective Borel presentation (which is a much more general notion than an injective ω-automatic presentation). Nevertheless, many interesting ω-automatic structures do admit injective presentations such as, for instance, the reducts $(\mathbb{R}, +)$ and $(\mathbb{R}, \cdot)$ of the field of reals. Therefore it is also interesting to ask what kind of structures admit an injective presentation.

We shall study injective ω-automatic presentations in Sect. 3. A central tool for such investigations is the notion of *end-equivalence* of infinite words: two words $\alpha, \beta \in \Sigma^\omega$ are end-equivalent, in short $\alpha \sim_e \beta$, if they are equal from some position onwards, i.e. $\alpha[n, \omega) = \beta[n, \omega)$ for some $n \in \mathbb{N}$, and we shall analyse the size of $\sim_e$-equivalent sets in injectively ω-automatic structures. As a consequence we can derive certain restrictions on the behaviour of definable functions in such structures.

In Sect. 4 we consider ω-automatic structures with not necessarily injective presentations, but with a definable linear order. Inspired by the work of Kaiser, Rubin and Bárány [7] and the work of Kuske [9], we use the algebraic characterization of ω-regular languages by ω-semigroup morphisms to show that every ω-automatic presentation of an uncountable linear order contains a (very simple) encoding of the lexicographic order on all infinite binary strings. As a consequence, we can establish the bounds for definable functions, that we proved in Sect. 3, also for this case.

In Sect. 5 we investigate further properties of functions that can be defined (with parameters) in an ω-automatic structure. Combining the results of Sects. 3 and 5, we shall be able to conclude for several classes of structures, such as infinite integral domains or structures with a definable pairing function, that they do not contain any ω-automatic model. Thus, in particular, the field of reals is not ω-automatic.

2 Preliminaries

2.1 Automata and Presentations

We assume that the reader is familiar with the basic notions of automata theory, especially with automata operating on infinite words. For an introduction see [16].

Here, we just recall some basic definitions in order to fix our notation. For a given alphabet Σ, we denote by Σ^ω the set of all ω-words $\alpha = \alpha[0]\alpha[1]\alpha[2]\cdots$ over Σ. We write $\alpha[m,n)$ as an abbreviation for the infix $\alpha[m]\alpha[n+1]\cdots\alpha[n-1]$ of α. As a convention we use Greek letters $\alpha, \beta, \gamma, \ldots$ to denote infinite words and Roman letters $u, v, w, \ldots$ to denote finite words. An ultimately periodic word is a word that can be written as $\alpha = v(w)^\omega$. The period length of an ultimately periodic word α is the smallest number n such that $\alpha = v(w)^\omega$ for some $w \in \Sigma^n$. Sometimes we use the term a bit more sloppy and say that a word has period length n when the actual period length divides n.

A Büchi-automaton is a tuple $\mathcal{A} = (Q, \Sigma, q_0, \Delta, F)$, where Q is a finite set of states, Σ a finite set of symbols, $q_0 \in Q$ the initial state, $\Delta \subseteq Q \times \Sigma \times Q$ the transition relation and $F \subseteq Q$ the set of accepting states. A run of $\mathcal{A}$ on a word $\alpha \in \Sigma^\omega$ is an ω-word $\rho \in Q^\omega$ where $\rho[0] = q_0$ and for all $i \in \mathbb{N}$ we have $(\rho[i], \alpha[i], \rho[i+1]) \in \Delta$. A run is accepting, if for infinitely many i it holds that $\rho[i] \in F$. A word $\alpha \in \Sigma^\omega$ is accepted by an automaton $\mathcal{A}$ if there is an accepting run of $\mathcal{A}$ on α. The language of all words accepted by a Büchi-automaton $\mathcal{A}$ is denoted by $L(\mathcal{A})$. A language $L \subseteq \Sigma^\omega$ is called ω-automatic (or ω-regular) if $L = L(\mathcal{A})$ for some Büchi-automaton $\mathcal{A}$. For some Büchi-automaton $\mathcal{A} = (Q, \Sigma, q_0, \Delta, F)$ and $w \in \Sigma^*$ the transition profile of w with respect to $\mathcal{A}$ is defined as

$$\begin{aligned}\Delta(w) := \big\{&(p,q,0) : \text{there is a } w\text{-path from } p \text{ to } q \text{ that visits an accepting state}\big\}\\ &\cup \big\{(p,q,1) : \text{there is a } w\text{-path from } p \text{ to } q \text{ that visits no accepting state}\big\}.\end{aligned}$$

If two words have the same transition profile, then they are indistinguishable for the automaton.

Lemma 1 *Let $\mathcal{A} = (Q, \Sigma, q_0, \Delta, F)$ be a Büchi-automaton, $\alpha \in \Sigma^\omega$, and $v, w \in \Sigma^*$ such that $\Delta(v) = \Delta(w)$. Then for every α' obtained by substituting some occurrences of v by w in α it holds that $\alpha \in L(\mathcal{A})$ if, and only if, $\alpha' \in L(\mathcal{A})$.*

Proof Fix some $\alpha, \alpha' \in \Sigma^\omega$ such that α' originates from α by the substitution of v by w at positions $I \subseteq \mathbb{N}$. Let ρ be a run of $\mathcal{A}$ on α. We construct a run ρ' in the following way: for every $i \in I$ replace (simultaneously) the infixes $\rho[i, i+|v|)$ by a sequence $p_0 \cdots p_{|w|-1}$ such that $p_0 \cdots p_{|w|-1}$ is a w-labeled path in $\mathcal{A}$, $p_0 = \rho[i]$, $p_{|w|-1} = \rho[i+|v|-1]$ and $p_j \in F$ for some $0 \le j \le |w|-1$ if, and only if, $\rho[j'] \in F$ for some $i \le j' < i+|v|$. Such a sequence exists since $\Delta(v) = \Delta(w)$. It is easy to check that ρ' is a run of $\mathcal{A}$ on α' and that ρ' is accepting if, and only if, ρ is accepting. □

We pick up on the idea of ω-automatic presentations from the introduction. To give a formal definition, we first need to define the notion of ω-regularity for k-ary relations $R \subseteq (\Sigma^\omega)^k$.

Definition 2 We call a relation $R \subseteq (\Sigma^\omega)^k$ ω-regular if the corresponding language $L_R := \{\langle\alpha_1, \ldots, \alpha_k\rangle \colon (\alpha_1, \ldots, \alpha_k) \in R\}$ is ω-regular. Here $\langle\alpha_1, \ldots, \alpha_k\rangle \in (\Sigma^k)^\omega$ denotes the convolution of $\alpha_1, \ldots, \alpha_n$ which is defined as $\langle\alpha_1, \ldots, \alpha_k\rangle[i] = (\alpha_1[i], \ldots, \alpha_k[i]) \in \Sigma^k$.

For finite words $(w_1, \dots, w_k) \in (\Sigma^n)^k$ of length n the convolution $\langle w_1, \dots, w_k \rangle \in (\Sigma^k)^n$ is defined analogously.

Observe that we have defined the convolution of finite words only for words of the same length, and hence, we can avoid the introduction of a padding symbol. In the following we usually identify $(w_1, \dots, w_k)$ and $\langle w_1, \dots, w_k \rangle$.

Having introduced the notion of ω-automatic (or ω-regular) relations, we proceed to define the central notion of ω-automatic presentations of structures.

Definition 3 Let $\tau = \{R_1, \dots, R_n\}$ be a relational vocabulary, where $R_i^{\mathfrak{A}}$ denotes a relation symbol of arity r_i. An ω-automatic presentation (over the alphabet Σ) of a τ-structure $\mathfrak{A} = (A, R_1^{\mathfrak{A}}, \dots, R_n^{\mathfrak{A}})$ is a pair $(\mathcal{L}, \pi)$ consisting of a structure $\mathcal{L} = (L, \approx, R_1^{\mathcal{L}}, \dots, R_n^{\mathcal{L}})$ and a surjective labeling function $\pi : L \to A$ such that the following holds:

- $L \subseteq \Sigma^\omega$;
- $\approx \; = \{(\alpha, \beta) : \pi(\alpha) = \pi(\beta)\}$;
- $R_i^{\mathcal{L}} = \{(\alpha_1, \dots, \alpha_{r_i}) : (\pi(\alpha_i), \dots, \pi(\alpha_{r_i})) \in R_i\}$ for $i \in \{1, \dots, n\}$;
- $L, \approx, R_1^{\mathcal{L}}, \dots, R_n^{\mathcal{L}}$ are ω-regular relations.

We call a presentation injective if π is injective, in which case we omit $\approx$ in the signature of $\mathcal{L}$. A τ-structure is ω-automatic if it has an ω-automatic presentation.

We remark that we can easily obtain a notion of ω-automaticity also for structures which contain function symbols, by just replacing every function f by its graph $G_f := \{(a_1, \dots, a_k, b) : f(a_1, \dots, a_k) = b\}$. Furthermore, note that instead of giving a concrete labeling function π it suffices to require that $(L, R_1^{\mathcal{L}}, \dots, R_n^{\mathcal{L}})/\approx$ is isomorphic to $\mathfrak{A}$.

As mentioned in the introduction, every ω-automatic structure has a decidable first-order theory and moreover, every first-order definable relation on an ω-automatic structure can be described by an effectively computable automaton. Indeed, for ω-automatic structures we can also allow the formulas to contain quantifiers of the form *"there are k mod m many elements"*, *"there are at most countably many elements"* and *"there are uncountably many elements"*. The extension of first-order logic by these counting quantifiers is denoted by FOC. This was first proven for injective ω-automatic presentations by Kuske and Lohrey [10], and later generalised to all ω-automatic presentations by Kaiser, Rubin, and Bárány [7].

Theorem 4 ([7, 10]) *Given an ω-automatic presentation $(\mathcal{L}, \pi)$ of a structure $\mathfrak{A}$ and an* FOC*-formula ϕ one can effectively construct an automaton that recognises $\pi^{-1}(\phi^{\mathfrak{A}})$.*

It follows that also the FOC-theory of every ω-automatic structure is decidable. When we say that a function or relation is definable in a structure $\mathfrak{A}$ we implicitly refer to FOC-definability.

Another important result in [7] answers a question raised by Blumensath in [2].

Theorem 5 ([7]) *Let $\mathfrak{A}$ be a countable structure. Then the following statements are equivalent.*

- *$\mathfrak{A}$ is ω-automatic.*
- *$\mathfrak{A}$ has an injective ω-automatic presentation.*
- *$\mathfrak{A}$ is finite word automatic.*

2.2 ω-Semigroups

There is also an algebraic way to characterise the notion of ω-regularity which is based on ω-semigroups. Here, we give a short overview about the fundamental concepts and known results about this correspondence, since some of our constructions can be more conveniently formulated in the algebraic framework. With our presentation we follow Perrin and Pin [12].

Definition 6 An ω-semigroup is a two-sorted structure $S = (S_f, S_\omega, \cdot, *, \pi)$ with the following properties.

- $(S_f, \cdot)$ is a semigroup.
- $* : S_f \times S_\omega \to S_\omega$ is the mixed product satisfying for all $x, y \in S_f$, $z \in S_\omega$
$$x * (y * z) = (x \cdot y) * z.$$
- $\pi : S_f^\omega \to S_\omega$ is the infinite product that satisfies for all $x_0, x_1, \ldots \in S_f$
$$x_0 * \pi(x_1, x_2, x_3, \ldots) = \pi(x_0, x_1, x_2, \ldots).$$
- Additionally one demands some kind of associativity rule for π, namely that for every strictly increasing sequence of positive integers $(k_i)_{i\in\mathbb{N}}$ it holds that
$$\pi(x_1, x_2, x_3, \ldots) = \pi\big((x_1x_2\cdots x_{k_1}), (x_{k_1+1}x_{k_1+2}\cdots x_{k_2}), \ldots\big).$$

Because of the last two properties we can represent the product $x_0 * \pi(x_1, x_2, x_3, \ldots)$ without ambiguity as $x_0x_1x_2x_3\cdots$ and we also write $(x_i)_{i\in\mathbb{N}}$ to denote this product. Furthermore we write $(x_i)_{[n,m)}$ to abbreviate $x_n \cdots x_{m-1}$.

In the study of finite (ω-)semigroups idempotence and absorption play an important role. An element e of a semigroup $(S, \cdot)$ is idempotent if $e \cdot e = e$, and e absorbs d from the left if $e \cdot d = e$. For every element s of a *finite* semigroup $(S, \cdot)$ there is a $k \in \mathbb{N}$ such that s^k is idempotent. The smallest number k such that s^k is idempotent for all elements $s \in S$ is called the exponent of the semigroup.

Example 7 For an alphabet Σ, the free ω-semigroup over Σ is the structure $\Sigma^\infty = (\Sigma^+, \Sigma^\omega, \cdot, *, \pi)$, where $\cdot, *$ and π are interpreted as the usual concatenation operations.

Observe that even in the case that S_f and S_ω are finite sets, the ω-semigroup $(S_f, S_\omega, \cdot, *, \pi)$ is not a finite object, since the domain of π is still uncountable. Thomas Wilke [18] solved this problem by showing that every ω-semigroup is completely determined by three operations of finite arity. The resulting structure is called Wilke-algebra.

Definition 8 A Wilke-algebra is a two sorted structure $(S_f, S_\omega, \cdot, *, ()^\omega)$ where:

- $(S_f, \cdot)$ is a semigroup.
- $* : S_f \times S_\omega \to S_\omega$ is the mixed product satisfying for all $x, y \in S_f$, $z \in S_\omega$

$$x * (y * z) = (x \cdot y) * z.$$

- $()^\omega : S_f \to S_\omega$ is the power operation with the property that for all $x, y \in S_f$ it holds that

$$x(yx)^\omega = (xy)^\omega \quad \text{and}$$
$$\left(x^n\right)^\omega = x^\omega \quad \text{for all } n \geq 1.$$

From a given ω-semigroup one naturally obtains a Wilke-algebra by restricting the infinite product π to the products of the form $\pi(a, a, a, \ldots)$ for $a \in S_f$. But also the converse is true.

Theorem 9 (Wilke [18]) *Every finite Wilke-algebra can be uniquely extended to a finite ω-semigroup.*

The key to prove this theorem is the Theorem of Ramsey. We state it here since we make extensive use of it, especially in Sect. 4.

Theorem 10 (Ramsey's Theorem [14]) *Let $G = (\mathbb{N}, E)$ be the complete countable undirected graph and $f : E \to C$ a coloring of the edges with some finite set of colors C. Then there is an infinite set $N = \{n_1 < n_2 < n_3 < \cdots\} \subseteq \mathbb{N}$ such that every edge in $E \cap (N \times N)$ has the same color.*

Definition 11 Let S and T be ω-semigroups. An ω-semigroup morphism $g : S \to T$ is a pair $g = (g_f, g_\omega)$ such that

- g_f is a semigroup morphism from $(S_f, \cdot)$ to $(T_f, \cdot)$, and
- g_ω is a function $g_\omega : S_\omega \to T_\omega$ that preserves the mixed and the infinite product, i.e. for every sequence $(x_i)_{i \in \mathbb{N}}$, $x_i \in S_f$ it holds that

$$g_\omega(x_1 x_2 x_3 \cdots) = g_f(x_1) g_f(x_2) g_f(x_3) \cdots$$

and for $x \in S_f$, $y \in S_\omega$

$$g_f(x) * g_\omega(y) = g_\omega(x * y).$$

If we have given an ω-semigroup morphism $g : S \to T$, we usually omit the subscripts of the mappings g_f and g_ω whenever this cannot lead to any confusion. We are now ready to give the central definition of this section.

Definition 12 Let $L \subseteq \Sigma^\omega$ be a language and $g : \Sigma^\infty \to S$ a morphism into some finite ω-semigroup S. We say that L is recognised by S via g if, and only if, $g_\omega^{-1}(g_\omega(L)) = L$.

In other words, for every set $X \subseteq S_\omega$ we say that the language $g^{-1}(X)$ is recognised by S via g. Observe that every morphism from $g : \Sigma^\infty \to S$ is completely determined by the values $g_f(a), a \in \Sigma$. Therefore we can represent every such morphism in a finite way.

Now that we know how to recognise a language by an ω-semigroup, we state that the class of languages which are recognizable by a finite ω-semigroups is precisely the class of ω-regular languages.

Theorem 13 *From a finite ω-semigroup S given as its corresponding Wilke-algebra and a morphism $g : \Sigma^\infty \to S$ that recognises the language L one can effectively compute a Büchi-automaton that recognises L, and vice versa.*

For a proof of this theorem we refer the reader to [12].

3 Injective Presentations

It is well-known that every structure which possesses an automatic presentation that encodes elements by *finite* words or trees, can also be represented by an *injective* automatic presentation. In contrast, the class of injectively presentable ω-automatic structures does not coincide with the class of all ω-automatic structures [6]. Therefore it is interesting to ask which structures allow an injective ω-automatic presentation. At this point, only a few examples of structures are known which have an ω-automatic representation but not an injective one.

In the following we introduce a technique that is particularly useful when injective presentation are considered. Two words $\alpha, \beta \in \Sigma^\omega$ are end-equivalent, in short $\alpha \sim_e \beta$, if they are equal from some position onwards, i.e. $\alpha[n, \omega) = \beta[n, \omega)$ for some $n \in \mathbb{N}$. Making explicit a position m after which the words are equal, we obtain refined relations $\sim_e^m$, i.e. two words are $\sim_e^m$-equivalent (m-end-equivalent) if $\alpha[m, \omega) = \beta[m, \omega)$. Clearly $\alpha \sim_e \beta$ if, and only if, $\alpha \sim_e^m \beta$ for some m. Moreover, the equivalence relation $\sim_e^m$ partitions any language into finite classes, each of size at most $|\Sigma|^m$.

In general, end-equivalence plays a crucial role in the study of ω-regular languages. We first observe that every infinite ω-regular language has an infinite $\sim_e$-class.

Lemma 14 *Let L be an infinite ω-regular language. Then L has an $\sim_e$-equivalence class which is infinite.*

Proof Since L is an ω-regular language, by [4] it has the form $L = \bigcup_{1 \le i \le n} U_i V_i^\omega$ for some (finite-word) regular languages U_i, V_i which are not empty. We consider two cases.

Suppose that $V_i^\omega = \{v_i^\omega\}$ for each i, in which case L is countable. Since L is infinite there is an i such that $U_i v_i^\omega$ is also infinite. In general the words in $U_i v_i^\omega$ are not pairwise $\sim_e$-equivalent. Nevertheless all $w \in U_i v_i^\omega$ fall into one of at most $|v_i|$ many $\sim_e$-classes and therefore $U_i v_i^\omega / \sim_e$ must contain an infinite class.

In the other case, there is a V_i which contains two words w, v such that $w^\omega \neq v^\omega$. Set $U_i' := U_i w^*$ then $U_i' v^\omega \subseteq (U_i V_i^*) V_i^\omega = U_i V_i^\omega$. The language $U_i' v^\omega$ is infinite. Otherwise the language $w^* v^\omega$ would also be finite and therefore $w^i v^\omega = w^j v^\omega$ for some $i \neq j$. But then $w^l = v^k$ for some $k, l \in \mathbb{N}$ and therefore $w^\omega = v^\omega$, a contradiction. Since $U_i' v^\omega$ is infinite we know from the first part of the proof that $U_i' v^\omega / \sim_e$ contains an infinite $\sim_e$-class. □

In the following we examine which elements of a structure can be encoded by words from the same $\sim_e$-class. To this end it is convenient to transfer the notion of end-equivalence from words in a given presentation to elements of the encoded original structure.

Definition 15 Let $\mathfrak{A}$ be a structure with an ω-automatic presentation $(\mathcal{L}, \pi)$ and let $\sim$ be an equivalence relation on the domain L of the presentation. Then we say that a set $B \subseteq A$ is $(\sim, \mathcal{L}, \pi)$-equivalent (or $\sim$-equivalent in $(\mathcal{L}, \pi)$) if, and only if, there is a set of words $X \subseteq L$ which are pairwise $\sim$-equivalent such that $B \subseteq \pi(X)$.

If the presentation $(\mathcal{L}, \pi)$ is clear from the context we just say that a set $B \subseteq A$ is $\sim$-equivalent without mentioning the presentation explicitly.

Observe that an equivalence relation $\sim$ on the domain of the presentation does not need to induce an equivalence relation on the domain of the structure. Indeed, an element of the structure can have several encodings in the presentation and can thus occur in the image of more that one $\sim$-class.

Lemma 16 *Let $\mathfrak{A}$ be an infinite structure. Then for every injective ω-automatic presentation $(\mathcal{L}, \pi)$ of $\mathfrak{A}$ there is an infinite set $M \subseteq A$ that is $\sim_e$-equivalent in $(\mathcal{L}, \pi)$.*

Proof Since A is infinite, L must also be infinite and therefore by Lemma 14 there must be an infinite class $X \in L/\sim_e$. Since π is injective it follows that $\pi(X)$ is an infinite set that is $\sim_e$-equivalent in $(\mathcal{L}, \pi)$. □

The importance of the notion of $\sim_e^m$-equivalent sets stems from the fact that we can, in some sense, control the action of a definable function f on such sets. More precisely, the image $f(B)$ of every $\sim_e^m$-equivalent set $B \subseteq A$ can be partitioned into a *constant* number of $\sim_e^m$-equivalent sets in such a way that the constant does only depend on the underlying ω-automatic presentation of the structure together with the function f, but not on the specific set $B \subseteq A$.

Lemma 17 *Let $\mathfrak{A}$ be a structure with ω-automatic presentation $(\mathcal{L}, \pi)$ and let $f : A^{k+\ell} \to A$ be a function which is* FOC*-definable in $\mathfrak{A}$. Then there is a constant q such that, for every $m \in \mathbb{N}$, every $\sim_e^m$-equivalent subset $B \subseteq A$ and every tuple $\overline{a} \in A^\ell$, the image $f(B^k, \overline{a})$ admits a partition into q-many sets $C_0, \ldots, C_{q-1}$ which are all $\sim_e^m$-equivalent.*

Proof Let $\mathcal{A}$ be a Büchi-automaton with state set $Q = \{0, \ldots, q-1\}$ that recognises f in $(\mathcal{L}, \pi)$. First we choose a tuple of words $v_{\bar{a}}$ that represent $\bar{a}$. Since

B is $\sim_e^m$-equivalent in $(\mathcal{L},\pi)$, there is a set $M=\{v_b : b\in B\}\subseteq L$ of pairwise $\sim_e^m$-equivalent representatives of B. For every tuple $\bar{b}\in B^k$ we denote with $v_{\bar{b}}$ the unique tuple in M^k that represents $\bar{b}$. Now choose for every tuple $v_{\bar{b}}$ a representative $f_{\bar{b}}$ of $f(\bar{b},\bar{a})$. This means that the word $(v_{\bar{b}}, v_{\bar{a}}, f_{\bar{b}})$ is accepted by $\mathcal{A}$ for every tuple $\bar{b}\in B^k$. Let $\rho_{\bar{b}}$ be an accepting run of $\mathcal{A}$ on $(v_{\bar{b}}, v_{\bar{a}}, f_{\bar{b}})$.

We obtain a partition of M^k into sets $M_0,\dots,M_{q-1}$ by setting $M_i := \{v_{\bar{b}} : \rho_{\bar{b}}[m]=i\}$. For every non-empty M_i we fix a tuple $v_{\bar{b}}\in M_i$. We show that for any $\bar{d}$ with $v_{\bar{d}}\in M_i$ there is an encoding of $f(\bar{d},\bar{a})$ that is $\sim_e^m$-equivalent to $f_{\bar{b}}$.

The main observation is that we can simply replace the tail of $(v_{\bar{d}}, v_{\bar{a}}, f_{\bar{d}})$ by the tail of $(v_{\bar{b}}, v_{\bar{a}}, f_{\bar{b}})$ and obtain a new word that is accepted by $\mathcal{A}$. This will give us a new encoding of $f(\bar{d},\bar{a})$ that is $\sim_e^m$-equivalent to $f_{\bar{b}}$. More formally for every such $\bar{d}$ it holds that $\rho_{\bar{d}}[0,m)\rho_{\bar{b}}[m,\omega)$ is an accepting run on

$$\begin{aligned}(v_{\bar{d}}, v_{\bar{a}}, f_{\bar{d}})[0,m)(v_{\bar{b}}, v_{\bar{a}}, f_{\bar{b}})[m,\omega) &= (v_{\bar{d}}, v_{\bar{a}}, f_{\bar{d}})[0,m)(v_{\bar{d}}, v_{\bar{a}}, f_{\bar{b}})[m,\omega)\\ &= \big(v_{\bar{d}}, v_{\bar{a}}, f_{\bar{d}}[0,m\big)f_{\bar{b}}[m,\omega)).\end{aligned}$$

This holds since $v_{\bar{d}}\sim_e^m v_{\bar{b}}$ and $\rho_{\bar{d}}[m]=\rho_{\bar{b}}[m]=i$. Since $\mathcal{A}$ recognises f in $(\mathcal{L},\pi)$ it follows that $\pi(f_{\bar{d}}[0,m)f_{\bar{b}}[m,\omega)) = f(\bar{d},\bar{a})$. So, for every $i\in Q$ there exists a $\sim_e^m$-class such that for every $v_{\bar{d}}\in M_i$ the value $f(\bar{d},\bar{a})$ has an encoding in this class. This implies that all the sets C_i defined by $C_i := \{f(\bar{b},\bar{a}) : v_{\bar{b}}\in M_i\}$ are $\sim_e^m$-equivalent in $(\mathcal{L},\pi)$, and that $\bigcup_{0\le i\le q-1} C_i = f(B^k,\bar{a})$. □

We shall apply this result to obtain certain restrictions on the behaviour of FOC-definable k-ary functions f for structures that have an injective ω-automatic presentation. Intuitively we show that the size of the set $f(B^k)$ cannot be always much larger than the set B itself. In order to formulate this idea precisely, we introduce the notion of the minimal image size of a k-ary function f.

Definition 18 For every function $f : A^k\to A$ over an infinite set A we define the minimal image size $\mathrm{MIS}_f:\mathbb{N}\to\mathbb{N}$ by

$$\mathrm{MIS}_f(n) = \min\big\{\big|f\big(X^k\big)\big| : X\subseteq A, |X|=n\big\}.$$

We now show that for injectively presentable structures the minimal image size of every FOC-definable function grows at most linearly with n.

Lemma 19 *Let $\mathfrak{A}$ be an infinite structure with injective ω-automatic presentation. Then for every* FOC*-definable function f it holds that* $\mathrm{MIS}_f(n)=\mathcal{O}(n)$.

Proof Suppose there is an injective automatic presentation $(\mathcal{L},\pi)$ (over some alphabet Σ) of an infinite structure with FOC-definable function $f : A^k\to A$ such that MIS_f grows super-linearly.

Let q be the constant from Lemma 17 with respect to f and $(\mathcal{L},\pi)$. Now choose n such that $\mathrm{MIS}_f(n) > |\Sigma|\cdot q\cdot n$. This is possible since MIS_f grows super-linearly. By Lemma 16 there is an infinite set $M\subseteq A$ that is $\sim_e$-equivalent in $(\mathcal{L},\pi)$. Therefore we can choose the smallest m such that there is a $(\sim_e^m,\mathcal{L},\pi)$-equivalent set of size at least n. Let N be such a $\sim_e^m$-equivalent set of maximal size. This is possible since the size of every $\sim_e^m$-class is bounded by $|\Sigma|^m$. The size of N is bounded from above $|N| \leq |\Sigma| \cdot n$. This holds since N can be partitioned into $|\Sigma|$ many $\sim_e^{m-1}$-equivalent sets. So, if $|N| > |\Sigma| \cdot n$ then one such set must contain more than n elements, which contradicts the choice of m.

By Lemma 17, $f(N^k)$ can be partitioned into q many $\sim_e^m$-equivalent sets. One of these sets has size at least

$$\frac{|f(N^k)|}{q} > \frac{|\Sigma| \cdot q \cdot n}{q} = |\Sigma| \cdot n \geq |N|.$$

But this contradicts the maximality of N among all $\sim_e^m$-equivalent sets. □

Corollary 20 *No infinite structure with an* FOC*-definable pairing function admits an injective* ω*-automatic presentation.*

Proof Note that, for a pairing function f, $\mathrm{MIS}_f(n) = n^2$. □

4 Non-injective Presentations with a Linear Order

For ω-automatic structures with a definable linear order, it is possible to transfer the results of the previous section from injective presentations to general ω-automatic presentations. The problem that we face when we consider non-injective presentations $(\mathcal{L}, \pi)$ is that, in general, infinite $\sim_e$-equivalent sets do not need to exist. For example, $\sim_e$ is an ω-automatic equivalence relation and thus the presentation $(\mathcal{L}, \pi)$ might indeed identify all end-equivalent words.

In this section we show that this cannot happen for ω-automatic presentations of uncountable linear orders. In [9] Kuske has already shown that $(\{0,1\}^\omega, <_{lex})$ is embeddable into any ω-automatic uncountable linear order. More specifically, he constructs from a given ω-automatic presentation of such an order a sub-presentation that is a presentation of $(\{0,1\}^\omega, <_{lex})$. This sub-presentation is not ω-automatic but its domain is the complement of a language $\bigcup_{i \leq n} V_i U_i^\omega$ where the V_i are context free and the U_i are regular. In particular his presentation does not contain any two $\sim_e$-equivalent words.

We present here a strengthening of Kuske's result. We show that every automatic presentation of an uncountable linear order contains an injective automatic presentation of $(\{0,1\}^\omega, <_{lex})$. The main techniques originate from [7]. For a given ω-automatic presentation $(L, \approx, <)$ we construct finite words u, v_0, v_1 such that $u\{v_0, v_1\}^\omega \subseteq L$ and for any two words $\alpha, \beta \in \{0,1\}^\omega$ it holds that $uv_{\alpha[0]}v_{\alpha[1]}v_{\alpha[2]} \cdots < uv_{\beta[0]}v_{\beta[1]}v_{\beta[2]} \cdots$ if, and only if, $\alpha <_{lex} \beta$. In other words this shows that if one indentifies v_0 with 0 and v_1 with 1, then the natural encoding of $\{0,1\}^\omega$ by the language $u\{v_0, v_1\}^\omega$ is compatible with the lexicographical ordering $<_{lex}$. To construct these words we make use of the characterization of ω-regular languages by morphisms to ω-semigroups. In this framework we can apply Ramsey's Theorem to obtain a suitable factorization of some given words which we can use to construct u, v_0 and v_1. Finally, the algebraic structure of the underling ω-semigroups can be used to ensure that the elements encoded by the newly constructed words are ordered as claimed.

Theorem 21 *For any ω-automatic presentation* $(L, \approx, <)$ *of an uncountable linear order there is a subset* L' *of* L *such that* $(L', <')$, *where* $<'$ *is the restriction of* $<$ *to* L', *is an injective ω-automatic presentation of* $(\{0,1\}^\omega, <_{lex})$.

Proof Let $\mathcal{L} = (L, \approx, <)$ be an ω-automatic presentation of an uncountable linear order. Since $\mathcal{L}$ is automatic there are ω-semigroup morphisms to finite ω-semigroups $S_\delta = (S^\delta_f, S^\delta_\omega)$, for $\delta \in \{L, \approx, <\}$,

$$()^L : \Sigma^\infty \to S_L,$$
$$()^\approx : (\Sigma \times \Sigma)^\infty \to S_\approx \quad \text{and}$$
$$()^< : (\Sigma \times \Sigma)^\infty \to S_<$$

that recognise the corresponding relations. For $\delta \in \{L, \approx, <\}$ we set $F_\delta := (\delta)^\delta \subseteq S^\delta_\omega$.

So an ω-word $\alpha \in \Sigma^\omega$ ($\alpha \in \Sigma^\omega \times \Sigma^\omega$, respectively) is in δ if, and only if, $(\alpha)^\delta \in F_\delta$ We define $C := |S_L \times S_\approx \times S_<|$ and k as the least common multiple of the exponents of the semigroups $S_L, S_\approx, S_<$.

We start our construction by translating Ramsey's Theorem into the language of ω-semigroup morphisms.

Definition 22 Let $h : \Sigma^\infty \to S = (S_f, S_\omega)$ be an ω-semigroup morphism, $\alpha \in \Sigma^\omega$, $G = \{g_1 < g_2 < g_3 < \cdots\} \subseteq \mathbb{N}$ and $e \in S_f$. We say that G is a h, e-homogeneous factorization of α if, and only if, for all $i < j \in \mathbb{N}$ it holds that $h(\alpha[g_i, g_j)) = e$.

Lemma 23 *For* $1 \le i \le n$ *let* $h_i : \Sigma_i^\infty \to S_i$ *be morphisms into finite ω-semigroups* S_i *and* $\alpha_i \in \Sigma_i^\omega$ *words over the corresponding alphabets. Then there is a* $G = \{g_1 < g_2 < g_3 < \cdots\} \subseteq \mathbb{N}$ *such that for every* i *it holds that* G *is an* h_i,e_i*-homogeneous factorization of* α_i *for some idempotent semigroup element* $e_i \in S_i$.

Proof We color every $\{k, l\} \subset \mathbb{N}$, $k < l$ with the tuple of semigroup elements $h_i(\alpha_i[k, l))_{1 \le i \le n}$. With Ramsey's Theorem we obtain a $G = \{g_1 < g_2 < g_3 < \cdots\}$ such that all $\{g_i, g_j\}$, $i \neq j$ have the same color. Having set $e_i := h_i(\alpha_i[g_1, g_2))$, G obviously is a h_i,e_i-homogeneous factorization of α_i and the element e_i is idempotent since $e_i e_i = h_i(\alpha_i[g_1, g_2))h_i(\alpha_i[g_2, g_3)) = h_i(\alpha_i[g_1, g_3)) = e_i$. □

Next, we show that there are two words α_0 and α_1 and a factorization H such that, with respect to H, α_0 and α_1 are mapped to the same elements under every mentioned morphism. Later on we use these words to obtain suitable candidates for u, v_0 and v_1.

Lemma 24 *There are* $\alpha_0, \alpha_1 \in L$ *with* $[\alpha_i]_{\sim_e} \cap [\alpha_{1-i}]_\approx = \emptyset$ *and a factorization* $G^* = \{g^*_1 < g^*_2 < g^*_3 < \cdots\}$ *such that for some idempotent* $e_L \in S_L$, G^* *is a* $()^L$, e_L*-homogeneous factorization of* α_0 *and* α_1. *For* $\delta \in \{\approx, <\}$ *there are idempotents* e_δ, e^{01}_δ *and* e^{10}_δ *in* S_δ *such that*

– *G^* is a $()^\delta, e_\delta$-homogeneous factorization of* (α_0, α_0) *and* (α_1, α_1), *and*
– *G^* is a $()^\delta, e_\delta^{01}$-homogeneous factorization of* (α_0, α_1), *and*
– *G^* is a $()^\delta, e_\delta^{10}$-homogeneous factorization of* (α_1, α_0).

Proof Since $\mathcal{L}$ is a presentation of an uncountable structure, there is an infinite set X such that for all $\alpha \neq \beta \in X$ it holds that $[\alpha]_{\sim_e} \cap [\beta]_{\approx} = \emptyset$. To see this, note that every $\sim_e$-class contains only countably many elements. We choose distinct elements $\beta_0, \beta_1, \dots, \beta_C$ from X and take a look at their images under the given morphisms. For that we apply Lemma 23 simultaneously to the pairs

– $(()^L, \beta_i)$ for $0 \leq i \leq C$,
– $(()^\approx, (\beta_i, \beta_j))$ for $0 \leq i, j \leq C$, and
– $(()^<, (\beta_i, \beta_j))$ for $0 \leq i, j \leq C$.

Then we obtain $G^* = \{g_1^* < g_2^* < g_3^* < \cdots\} \subseteq \mathbb{N}$ such that for $0 \leq i, j \leq C$ there are idempotents $e_L^i, e_\approx^{ij}, e_<^{ij}$ for which G^* is

– a $()^L, e_L^i$-homogeneous factorization of β_i,
– a $()^\approx, e_\approx^{ij}$-homogeneous factorization of (β_i, β_j), and
– a $()^<, e_<^{ij}$-homogeneous factorization of (β_i, β_j).

If we look at the tuples $(e_L^i, e_\approx^{ii}, e_<^{ii}) \in (S_f^L \times S_f^\approx \times S_f^<), 0 \leq i \leq C$ we find that since $|S_f^L \times S_f^\approx \times S_f^<| = C$ there are some $i \neq j$ with $(e_L^i, e_\approx^{ii}, e_<^{ii}) = (e_L^j, e_\approx^{jj}, e_<^{jj})$. This means β_i, β_j and G^* fulfill the properties that we were looking for. □

Since $\alpha_0 \not\sim_e \alpha_1$, we may also assume that G^* is coarse enough such that $\alpha_0[g_\ell^*, g_{\ell+1}^*) \neq \alpha_1[g_\ell^*, g_{\ell+1}^*)$ for all $\ell \in \mathbb{N}$. We need to modify α_0, α_1 a bit to ensure all the properties we need in the following.

Lemma 25 *There are $\beta_0, \beta_1 \in L$ with $\beta_0 \not\approx \beta_1$ and $G = \{g_1 < g_2 < g_3 < \cdots\} \subseteq \mathbb{N}$ with the following properties*:

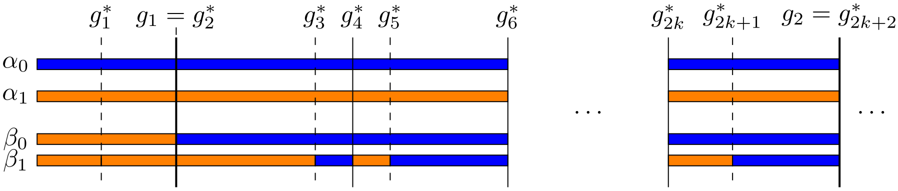
– $\beta_0[0, g_1) = \beta_1[0, g_1)$ *and* $\beta_0[g_i, g_{i+1}) \neq \beta_1[g_i, g_{i+1})$ *for all* $i \in \mathbb{N}$.
– *for $\delta \in \{\approx, <\}$ there is an element $\rightarrow_\delta \in S_f^\delta$ and idempotent elements $\Box_\delta, \uparrow_\delta, \downarrow_\delta \in S_f^\delta$ such that*
  – $(\beta_0, \beta_0)[0, g_1)^\delta = \rightarrow_\delta$,
  – $(\beta_0, \beta_0)[g_i, g_{i+1})^\delta = (\beta_1, \beta_1)[g_i, g_{i+1})^\delta = \Box_\delta$,
  – $(\beta_0, \beta_1)[g_i, g_{i+1})^\delta = \uparrow_\delta$,
  – $(\beta_1, \beta_0)[g_i, g_{i+1})^\delta = \downarrow_\delta$, *and*
  – $\rightarrow_\delta, \uparrow_\delta$ *and* $\downarrow_\delta$ *absorb* $\Box_\delta$ *from the right*.

Proof We first construct β_0 and β_1 and then show that these have the desired properties. We define β_0 as

$$\beta_0 := \alpha_1[0, g_2^*)\alpha_0[g_2^*, \omega)$$

and β_1 by

$$\beta_1[0, g_2^*) := \alpha_1[0, g_2^*) \quad \text{and}$$
$$\beta_1[g_{2\ell}^*, g_{2\ell+2}^*) := \alpha_1[g_{2\ell}^*, g_{2\ell+1}^*)\alpha_0[g_{2\ell+1}^*, g_{2\ell+2}^*) \quad \text{for } \ell \geq 1.$$



Construction of β_0, β_1 and G

Due to our assumptions about G^* we have $\beta_0 \not\sim_e \beta_1$. We set $G = \{g_{2k\ell+2}^* =: g_{\ell+1} : \ell \in \mathbb{N}\}$ (remember, k is the least common multiple of the exponents of the involved semigroups). Then $\beta_0[0, g_1) = \beta_1[0, g_1)$ and $\beta_0[g_i, g_{i+1}) \neq \beta_1[g_i, g_{i+1})$ as postulated. As e_L is idempotent, G is a $()^L, e_L$-homogeneous factorization of β_0 and β_1 and therefore it holds for $i \in \{0, 1\}$ that $(\beta_i)^L = \alpha_1[0, g_2)^L (e_L)^\omega = (\alpha_1)^L$ which implies that $\beta_i \in L$. Next we show $\beta_0 \not\approx \beta_1$:

$$\begin{aligned}
(\beta_0, \beta_1)^{\approx} &= (\beta_0, \beta_1)[0, g_1)^{\approx} * \big((\beta_0, \beta_1)[g_i, g_{i+1})^{\approx}\big)_{i\geq 1} \\
&= \underbrace{(\alpha_1, \alpha_1)[0, g_1^*)^{\approx}}_{:=s}(\alpha_1, \alpha_1)[g_1^*, g_2^*)^{\approx} * \big((\beta_0, \beta_1)[g_i, g_{i+1})^{\approx}\big)_{i\geq 1} \\
&= s e_{\approx} * \big((\alpha_0, \alpha_1)[g_{2i}^*, g_{2i+1}^*)^{\approx} \cdot (\alpha_0, \alpha_0)[g_{2i+1}^*, g_{2i+2}^*)^{\approx}\big)_{i\geq 1} \\
&= s e_{\approx}\big(e_{\approx}^{01} e_{\approx}\big)^\omega \\
&= s e_{\approx} e_{\approx}\big(e_{\approx}^{01} e_{\approx}\big)^\omega \\
&= s e_{\approx}\big(e_{\approx} e_{\approx}^{01}\big)^\omega \\
&= (\alpha_1, \alpha_1)[0, g_2^*)^{\approx} * \big((\alpha_1, \alpha_1)[g_{2i}^*, g_{2i+1}^*)^{\approx} \cdot (\alpha_0, \alpha_1)[g_{2i+1}^*, g_{2i+2}^*)^{\approx}\big)\big)_{i\geq 1} \\
&= (\beta_1, \alpha_1)^{\approx}
\end{aligned}$$

So if $\beta_0 \approx \beta_1$ then also $\beta_1 \approx \alpha_1$ and therefore by transitivity $\beta_0 \approx \alpha_1$. But $\beta_0 \sim_e \alpha_0$ and therefore $\beta_0 \approx \alpha_1$ implies $[\alpha_0]_{\sim_e} \cap [\alpha_1]_{\approx} \neq \emptyset$, which is contradicting the initial choice of α_0 and α_1.

We will now compute the values $\rightarrow_\delta, \uparrow_\delta, \downarrow_\delta$ and $\square_\delta$ for $\delta \in \{\approx, <\}$:

$$\begin{aligned}
\rightarrow_\delta &= \big(\beta_0[0, g_1), \beta_0[0, g_1)\big)^\delta \\
&= \big(\alpha_1[0, g_2^*), \alpha_1[0, g_2^*)\big)^\delta \\
&= \big(\alpha_1[0, g_1^*), \alpha_1[0, g_1^*)\big)^\delta e_\delta \\
\square_\delta &= \big(\beta_0[g_i, g_{i+1}), \beta_0[g_i, g_{i+1})\big)^\delta \\
&= \big(\alpha_0[g_{2ki+2}^*, g_{2ki+2k+2}^*), \alpha_0[g_{2ki+2}^*, g_{2ki+2k+2}^*)\big)^\delta
\end{aligned}$$

$$= (e_\delta)^{2k}$$
$$= e_\delta.$$

By similar computations we get $\uparrow_\delta = (e_\delta^{01} e_\delta)^k$ and $\downarrow_\delta = (e_\delta^{10} e_\delta)^k$. Since e_δ is idempotent and k is a multiple of the exponent of S_f^δ, the elements $\square_\delta$, $\uparrow_\delta$ and $\downarrow_\delta$ are idempotent and $\rightarrow_\delta$, $\uparrow_\delta$ and $\downarrow_\delta$ absorb $\square_\delta$.

It remains to show that $(\beta_1[g_i, g_{i+1}), \beta_1[g_i, g_{i+1}))^\delta = \square_\delta$.

$$\begin{aligned}
&\big(\beta_1[g_i, g_{i+1}\big), \beta_1[g_i, g_{i+1}))^\delta \\
&\quad = (\beta_0, \beta_0)[g^*_{2ki+2}, g^*_{2k(i+1)+2})^\delta \\
&\quad = (\alpha_1, \alpha_1)[g^*_{2ki+2}, g^*_{2ki+3})^\delta (\alpha_0, \alpha_0)[g^*_{2ki+2}, g^*_{2ki+4})^\delta \\
&\qquad \ldots \\
&\qquad (\alpha_1, \alpha_1)[g^*_{2k(i+1)}, g^*_{2k(i+1)+1})^\delta (\alpha_0, \alpha_0)[g^*_{2k(i+1)+1}, g^*_{2k(i+1)+2})^\delta \\
&\quad = (e_\delta)^{2k} \\
&\quad = \square_\delta
\end{aligned}$$

□

Now that we have β_0 and β_1 we are ready to construct u, v_0 and v_1. We set $u := \beta_1[0, g_1)$, $v_0 := \beta_0[g_1, g_2)$ and $v_1 := \beta_1[g_1, g_2)$. From this definition we immediately get for $\delta \in \{\approx, <\}$:

$$\begin{aligned}
(u, u)^\delta &= \rightarrow_\delta, \\
(v_0, v_0)^\delta &= (v_1, v_1)^\delta = \square_\delta, \\
(v_0, v_1)^\delta &= \uparrow_\delta \quad \text{and} \\
(v_1, v_0)^\delta &= \downarrow_\delta .
\end{aligned}$$

In the following we omit the subscripts and just write $\rightarrow$, $\uparrow$, $\downarrow$ and $\square$ since it will be clear from the context which one is meant. Further we write v_α with $\alpha \in \{0, 1\}^\omega$ for the word $v_{\alpha[1]} v_{\alpha[2]} v_{\alpha[3]} \cdots \in \Sigma^\omega$.

We proceed by showing that $u\{v_0, v_1\}^\omega$ has all the properties that were announced at the beginning of the proof.

Lemma 26 $u\{v_0, v_1\}^\omega \subseteq L$

Proof Let α be any sequence from $\{0, 1\}^\omega$.

$$\begin{aligned}
(uv_\alpha)^L &= \beta_1[0, g_1)^L \big(\beta_{\alpha[i]}[g_1, g_2)^L\big)_{i \in \mathbb{N}} \\
&= \alpha_1[0, g_1)^L (e_L)^\omega \\
&= (\alpha_1)^L \in F_L
\end{aligned}$$

This means every uv_α is in L and therefore $u\{v_0, v_1\}^\omega \subseteq L$. □

Next we show that at least some words from $u\{v_0, v_1\}^\omega$ do encode distinct elements.

Lemma 27 $\rightarrow (\uparrow\downarrow)^\omega \notin F_\approx$.

Proof In a first step we see that $\rightarrow\uparrow^\omega \notin F_\approx$ since $(\beta_0, \beta_1)^\approx = \rightarrow\uparrow^\omega$ and $\beta_0 \not\approx \beta_1$. We will make use of the transitivity of $\approx$ to show that also $\rightarrow (\uparrow\downarrow)^\omega \notin F_\approx$. Suppose $\rightarrow (\uparrow\downarrow)^\omega \in F_\approx$, then consider the words $u(v_0v_1v_0)^\omega$,$u(v_1v_0v_1)^\omega$ and $u(v_1v_1v_0)^\omega$. We have

$$\big(u(v_0v_1v_0)^\omega, u(v_1v_0v_1)^\omega\big)^\approx = \rightarrow (\uparrow\downarrow\uparrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega \quad \text{and}$$
$$\big(u(v_1v_0v_1)^\omega, u(v_1v_1v_0)^\omega\big)^\approx = \rightarrow (\Box\uparrow\downarrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega.$$

Hence $u(v_0v_1v_0)^\omega \approx u(v_1v_0v_1)^\omega$ and $u(v_1v_0v_1)^\omega \approx u(v_1v_1v_0)^\omega$ and therefore, by transitivity, $u(v_0v_1v_0)^\omega \approx u(v_1v_1v_0)^\omega$, but

$$\big(u(v_0v_1v_0)^\omega, u(v_1v_1v_0)^\omega\big)^\approx = \rightarrow (\uparrow\Box\Box)^\omega = \rightarrow\uparrow^\omega \notin F_\approx,$$

which yields a contradiction. □

We conclude our proof by showing that $<$ is indeed the lexicographic order on $u\{v_0, v_1\}^\omega$.

Lemma 28 *Either it holds for every $\alpha \neq \beta \in \{0,1\}^\omega$ $uv_\alpha < uv_\beta$ if, and only if, $\alpha <_{lex} \beta$ or it holds for every $\alpha \neq \beta \in \{0,1\}^\omega$ $uv_\alpha < uv_\beta$ if, and only if, $\beta <_{lex} \alpha$.*

Proof First observe that $u(v_0v_1)^\omega \not\approx u(v_1v_0)^\omega$ since $(u(v_0v_1)^\omega, u(v_1v_0)^\omega)^\approx = \rightarrow (\uparrow\downarrow)^\omega \notin F_\approx$. With this fact we know that either $u(v_0v_1)^\omega < u(v_1v_0)^\omega$ or $u(v_1v_0)^\omega < u(v_0v_1)^\omega$ holds. We carry out the proof for the case that $u(v_0v_1)^\omega < u(v_1v_0)^\omega$ holds, which means $(u(v_0v_1)^\omega, u(v_1v_0)^\omega)^< = \rightarrow (\uparrow\downarrow)^\omega \in F_<$. In this case, we show that for every $\alpha \neq \beta \in \{0,1\}^\omega$ $uv_\alpha < uv_\beta$ if, and only if, $\alpha <_{lex} \beta$. For the other case it can analogously be shown that $uv_\alpha < uv_\beta$ if, and only if, $\beta <_{lex} \alpha$. One only needs to interchange the roles of v_0 and v_1 as well as the roles of $\uparrow$ and $\downarrow$ in the following.

Take a look at the images of (uv_α, uv_β), $\alpha \neq \beta \in \{0,1\}^\omega$ under $()^<$. If we take the infinite product $(u,u)^<(v_{\alpha[0]}, v_{\beta[0]})^<(v_{\alpha[1]}, v_{\beta[1]})^<(v_{\alpha[2]}, v_{\beta[2]})^<, \ldots$ and use idempotence and absorption to eliminate multiple successive occurrences of $\uparrow$, $\downarrow$ and $\Box$ (except for occurrences in $\uparrow^\omega$, $\downarrow^\omega$ and $\Box^\omega$ at the end) we get a product of one of the following forms.

1. $\rightarrow (\uparrow\downarrow)^\omega$
2. $\rightarrow (\uparrow\downarrow)^n \uparrow^\omega, n \geq 0$
3. $\rightarrow (\uparrow\downarrow)^n \Box^\omega, n > 0$
4. $\rightarrow (\uparrow\downarrow)^n \uparrow \cdot\{\Box^\omega, \downarrow^\omega\}, n \geq 0$
5. $\rightarrow (\downarrow\uparrow)^\omega$
6. $\rightarrow (\downarrow\uparrow)^n \downarrow^\omega, n \geq 0$

7. $\rightarrow (\downarrow\uparrow)^n \Box^\omega, n > 0$
8. $\rightarrow (\downarrow\uparrow)^n \downarrow \cdot\{\Box^\omega, \uparrow^\omega\}, n \geq 0$

The product we obtain has one of the first four forms if, and only if, on the first position i where α and β differ we have $\alpha[i] = 0$ and $\beta[i] = 1$ i.e. $\alpha <_{lex} \beta$. It has one of the last four forms if, and only if, the "reversed" product obtained from $(uv_\beta, uv_\alpha)^<$ has one of the first four forms. Since $<$ is a linear order it follows that if we can show that every product of the form 1, 2, 3 and 4 is in $F_<$, it follows that the products of the form 5 to 8 are not in $F_<$ and we get that $u\{v_0, v_1\}^\omega$ is ordered as desired by $<$.

We already know that $\rightarrow (\uparrow\downarrow)^\omega \in F_<$. Once again we will use a transitivity argument to show that the products of the forms 2, 3 and 4 also belong to $F_<$. More precisely we will show that for every product ρ of the forms 2, 3 and 4 there are words w_1, w_2 and w_3 such that $(w_1, w_2)^< = (w_2, w_3)^< = \rightarrow (\uparrow\downarrow)^\omega$ and $(w_1, w_3)^< = \rho$. Since $\rightarrow (\uparrow\downarrow)^\omega \in F_<$ it holds that $w_1 < w_2$ and $w_2 < w_3$ and so by transitivity $w_1 < w_3$, but this means $\rho \in F_<$.

– $\rightarrow (\uparrow\downarrow)^n \uparrow^\omega, n \geq 0$:

$$w_1 := u(v_0v_1)^n(v_0v_0v_1)^\omega$$
$$w_2 := u(v_1v_0)^n(v_0v_1v_0)^\omega$$
$$w_3 := u(v_1v_0)^n(v_1v_0v_1)^\omega$$
$$(w_1, w_2)^< = \rightarrow (\uparrow\downarrow)^n(\Box \uparrow\downarrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega$$
$$(w_2, w_3)^< = \rightarrow \Box^{2n}(\uparrow\downarrow\uparrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega$$
$$(w_1, w_3)^< = \rightarrow (\uparrow\downarrow)^n(\uparrow \Box\Box)^\omega = \rightarrow (\uparrow\downarrow)^n \uparrow^\omega$$

– $\rightarrow (\uparrow\downarrow)^n\Box^\omega, n > 0$:

$$w_1 := u(v_0v_1)^n(v_0v_1)^\omega$$
$$w_2 := u(v_0v_1)^n(v_1v_0)^\omega$$
$$w_3 := u(v_1v_0)^n(v_0v_1)^\omega$$
$$(w_1, w_2)^< = \rightarrow \Box^{2n}(\uparrow\downarrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega$$
$$(w_2, w_3)^< = \rightarrow (\uparrow\downarrow)^n(\downarrow\uparrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega$$
$$(w_1, w_3)^< = \rightarrow (\uparrow\downarrow)^n(\Box\Box)^\omega = \rightarrow (\uparrow\downarrow)^n\Box^\omega$$

– $\rightarrow (\uparrow\downarrow)^n \uparrow \Box^\omega, n \geq 0$:

$$w_1 := u(v_0v_1)^n v_0(v_0v_1)^\omega$$
$$w_2 := u(v_0v_1)^n v_0(v_1v_0)^\omega$$
$$w_3 := u(v_1v_0)^n v_1(v_0v_1)^\omega$$
$$(w_1, w_2)^< = \rightarrow \Box^{2n+1}(\uparrow\downarrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega$$

$$(w_2, w_3)^< = \rightarrow (\uparrow\downarrow)^n \uparrow (\downarrow\uparrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega$$
$$(w_1, w_3)^< = \rightarrow (\uparrow\downarrow)^n \uparrow (\Box\Box)^\omega = \rightarrow (\uparrow\downarrow)^n \uparrow \Box^\omega$$

– $\rightarrow (\uparrow\downarrow)^n \uparrow\downarrow^\omega, n \geq 0$:

$$w_1 := u(v_0v_1)^n v_0(v_1v_0v_1)^\omega$$
$$w_2 := u(v_0v_1)^n v_0(v_1v_1v_0)^\omega$$
$$w_3 := u(v_1v_0)^n v_1(v_0v_0v_1)^\omega$$
$$(w_1, w_2)^< = \rightarrow \Box^{2n+1}(\Box \uparrow\downarrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega$$
$$(w_2, w_3)^< = \rightarrow (\uparrow\downarrow)^n \uparrow (\downarrow\downarrow\uparrow)^\omega = \rightarrow (\uparrow\downarrow)^\omega$$
$$(w_1, w_3)^< = \rightarrow (\uparrow\downarrow)^n \uparrow (\downarrow \Box\Box)^\omega = \rightarrow (\uparrow\downarrow)^n \uparrow\downarrow^\omega$$ □

Taking all together we get that $\mathcal{L}$ restricted to $u\{v_0, v_1\}^\omega$ is an injective ω-automatic-presentation of $(\{0, 1\}^\omega, <_{lex})$. □

Corollary 29 (Kuske [9]) $(\{0, 1\}^\omega, <_{lex})$ *is embeddable into every uncountable ω-automatic linear order.*

Note that for the construction of u, v_0, v_1 we did not make use of the linear order, but only of the equivalence relation of the structure. Indeed, the semigroup elements found in the previous Lemmas 24, 25, 26 and 27 for $\approx$ still have the stated properties indedependet of the presence of a linear order, which allows us to re-prove the following theorem, already mentioned in [7].

Theorem 30 *Let $\mathcal{L} = (L, \approx, \ldots)$ be an ω-automatic presentation of an uncountable structure. Then there are $u, v_0, v_1 \in \Sigma^+$ with $|v_0| = |v_1|$, $v_0 \neq v_1$ and $u\{v_0, v_1\}^\omega \subseteq L$ such that for all $\alpha, \beta \in u\{v_0, v_1\}^\omega$ it holds that $\alpha \not\sim_e \beta \Rightarrow \alpha \not\approx \beta$.*

Proof By Lemma 27 $\rightarrow (\uparrow)^\omega \notin F_\approx$ and $\rightarrow (\uparrow\downarrow)^\omega \notin F_\approx$. Now for $\alpha \not\sim_e \beta \in \{0, 1\}^\omega$ consider the product $(u, u)^\approx * ((v_{\alpha[1]}, v_{\beta[1]})^\approx)_{i\geq 0}$. Because of the symmetry of $\approx$ we might assume that the first position i with $\alpha[i] \neq \beta[i]$ we have $\alpha[i] = 0$ and $\beta[i] = 1$. As described in the proof of Lemma 28 we can use the idempotence and absorption properties to cancel out multiple successive occurrences of $\Box, \uparrow, \downarrow$. We end up with a sequence that has one the following forms $\rightarrow (\uparrow\downarrow)^\omega$ or $\rightarrow (\uparrow\downarrow)^n \uparrow^\omega$ or $\rightarrow (\uparrow\downarrow)^n \downarrow^\omega$. We use the same strategy as in Lemma 27 to show that neither one of these sequences is in $F_\approx$. For sequences of the form $\rightarrow (\uparrow\downarrow)^n \uparrow^\omega$ consider the words $w_1 = u(v_0v_1)^n(v_0)^\omega$, $w_2 = u(v_1v_0)^n(v_1v_0)^\omega$, $w_3 = uv_1(v_1v_0)^n(v_1)^\omega$. Then

$$(w_1, w_2)^\approx = \rightarrow (\uparrow\downarrow)^n(\uparrow \Box)^\omega = \rightarrow (\uparrow\downarrow)^n(\uparrow)^\omega \quad \text{and}$$
$$(w_2, w_3)^\approx = \rightarrow \Box(\uparrow\downarrow)^n(\uparrow \Box)^\omega = \rightarrow (\uparrow\downarrow)^n(\uparrow)^\omega.$$

If $\rightarrow (\uparrow\downarrow)^n(\uparrow)^\omega \in F_\approx$ then by transitivity also $(w_1, w_3)^\approx \in F_\approx$. But this is impossible, since $(w_1, w_3)^\approx = \rightarrow\uparrow \Box^{2n} \uparrow^\omega = \rightarrow\uparrow^\omega \notin F_\approx$. For the product $\rightarrow (\downarrow\uparrow)^n(\uparrow)^\omega$

consider the words $w_1 = u(v_1v_0)^n(v_0)^\omega$, $w_2 = u(v_0v_1)^n(v_1)^\omega$, $w_3 = (v_1v_0)^n(v_0v_1)^\omega$. We have

$$(w_1, w_2)^{\approx} = \rightarrow (\downarrow\uparrow)^n(\uparrow)^\omega \quad \text{and}$$
$$(w_2, w_3)^{\approx} = \rightarrow (\uparrow\downarrow)^n(\downarrow\Box)^\omega = \rightarrow (\uparrow\downarrow)^n(\downarrow)^\omega.$$

By symmetry of $\approx$, if $\rightarrow (\uparrow\downarrow)^n \downarrow^\omega \in F_\approx$ then also $\rightarrow (\downarrow\uparrow)^n \uparrow^\omega \in F_\approx$, and therefore $w_1 \approx w_2 \approx w_3$. But this is not possible since $(w_1, w_3)^{\approx} = \rightarrow \Box^{2n}(\Box\uparrow)^\omega = \rightarrow\uparrow^\omega \notin F_\approx$. □

As mentioned before, in the presence of a linear order, the results of the last section carry over to the class of all ω-automatic structures.

Corollary 31 *Let f be an* FOC*-definable function on a linearly ordered ω-automatic structure. Then* $\mathrm{MIS}_f(n) = \mathcal{O}(n)$.

Proof For countable structures this follows from Theorem 5, since ω-automatic countable structures have injective presentations. For uncountable structures Theorem 21 implies the existence of an infinite $\sim_e$-equivalent set in every ω-automatic presentation. Therefore the proof of Lemma 19 can be reused. □

In particular no linearly ordered ω-automatic structure admits an FOC-definable pairing function.

5 Parameterised Functions

We shall now extend our techniques by considering definable functions (with parameters) on uncountable ω-automatic structures. Our main technical result shows that there is no ω-automatic structure with FOC-definable parameterised functions of unbounded arity such that all pairs of different functions agree on at most countably many inputs.

Lemma 32 (cf. [6, Lemma 3.10]) *Let $R \subseteq \Sigma^\omega \times \Gamma^\omega$ be an ω-regular relation recognised by a Büchi-automaton $\mathcal{A} = (Q, \Sigma \times \Gamma, q_0, \Delta, F)$. Further let $\alpha \in \Sigma^\omega$ be some ultimately periodic word with period length $p \in \mathbb{N}$. Then if $\alpha R \neq \emptyset$ then there is a word $\beta \in \alpha R$ with period length at most $|Q| \cdot p$.*

Theorem 33 *Let $\mathfrak{A} = (A, R_1, \ldots, R_n)$ be an uncountable ω-automatic structure. Then there is a $k \in \mathbb{N}$ such that for every definable $(k+1)$-ary function $f(\bar{x}, y)$ there exist uncountable sets $M \subseteq A^k$ and $N \subseteq A$ with $f(\bar{a}, b) = f(\bar{a}', b)$ for all $\bar{a}, \bar{a}' \in M, b \in N$.*

Proof Fix an ω-automatic presentation $(\mathcal{L}, \pi)$ over the alphabet $\{0, 1\}$. Since $\mathfrak{A}$ is uncountable, we can apply Theorem 30 and obtain a language $L' = w\{v_0, v_1\}^\omega \subseteq L$ such that $|v_0| = |v_1|$ and for all $\alpha, \beta \in L'$ it holds that $\alpha \not\sim_e \beta$ implies $\alpha \not\approx \beta$.

Set $k = |v_0| + 1$ and let $f(\bar{x}, y)$ be a definable $(k+1)$-ary function. Let $\mathcal{A}_f = (Q, \{0,1\}^{k+1}, q_0, \Delta, F)$ be an automaton that recognises f in $(\mathcal{L}, \pi)$ and let p denote the number of possible transition profiles of $\mathcal{A}_f$ i.e. $p = |\{\Delta(w) : w \in (\{0,1\}^{k+2})^*\}|$.

We proceed as follows: first we define two languages L_p and L_{id} of ultimately periodic words. In L_p and L_{id} respectively, we will find suitable encodings of pairs of distinct tuples which can be combined to encodings of uncountably many pairwise distinct tuples. These will then be the encodings for our sets M and N. To ensure that the combinations of tuples from L_p and the combinations of tuples from L_{id} do not interfere with each other, we define these languages in such a way that the set of positions where the words of L_p "encode their information" is disjoint from the set of positions where this is the case for the words in L_{id}. More precisely we are going to guarantee that whenever two words of L_p differ at a given position $i \in \mathbb{N}$, then all pairs of words in L_{id} have the same "dummy letter" at position i, and vice versa. We set

$$L_p := \left\{ wv^\omega : v \in \{v_0, v_1\}^p \overline{\Box} \right\} \quad \text{and} \quad L_{id} := \left\{ wv^\omega : v \in \underline{\Box}\{v_0, v_1\}^{2kp+p} \right\}$$

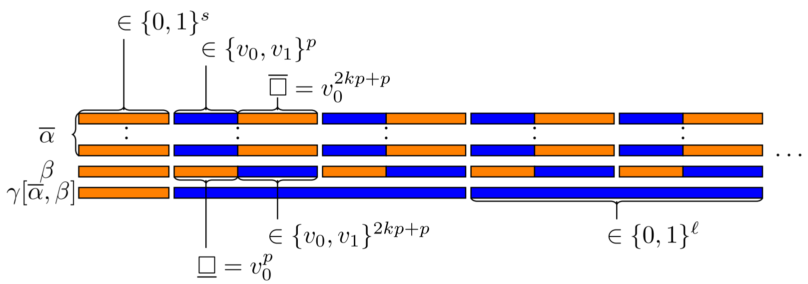
where $\overline{\Box} := (v_0)^{2kp+p}$ and $\underline{\Box} := (v_0)^p$. Then $L_p, L_{id} \subseteq L'$ and every word in $L_p \cup L_{id}$ represents a distinct element from the domain A, i.e. for every pair of words $\alpha, \beta \in L_p \cup L_{id}$ we have $\alpha \not\approx \beta$ if $\alpha \neq \beta$. Furthermore, every word $\alpha \in L_p \cup L_{id}$ is completely determined by the unique finite word $v[\alpha] \in \{v_0, v_1\}^p \cup \{v_0, v_1\}^{2kp+p}$ with $\alpha = w(v[\alpha]\overline{\Box})^\omega$ or $\alpha = w(\underline{\Box} v[\alpha])^\omega$, respectively. We extend this notation to tuples of words $\overline{\alpha} \in (L_p)^k$ as words $\bar{v}[\overline{\alpha}] \in (\{v_0, v_1\}^{2kp+p})^k$ in the obvious way. In particular we have $|L_p| = 2^p$ and $|L_{id}| = 2^{2kp+p}$.

Let us now consider $(k+1)$-tuples of the form $(\overline{\alpha}, \beta)$ with $\overline{\alpha} \in (L_p)^k$ and $\beta \in L_{id}$ as inputs of the automaton $\mathcal{A}_f$. We note that all words in $(L_p)^k \times L_{id}$ are periodic from position $|w|$ onwards and that the length of their periods divides $r := (2k+1)p|v_0|$. Both values $|w|$ and r are independent of the particular word in $(L_p)^k \times L_{id}$.

By Lemma 32 we can infer that for every $\overline{\alpha} \in (L_p)^k, \beta \in L_{id}$ the element $f(\pi(\overline{\alpha}), \pi(\beta))$ has an ultimately periodic encoding $\gamma[\overline{\alpha}, \beta]$ with the following properties:

- the length of the non-periodic prefix of $\gamma[\overline{\alpha}, \beta]$ is $s = |w| + c \cdot r$ for $c \in \mathbb{N}$, and
- the length of the period of $\gamma[\overline{\alpha}, \beta]$ is $\ell = d \cdot r$ for some $d \in \mathbb{N}$, and
- both constants c and d are independent of the particular choice of $\overline{\alpha}, \beta$.

We illustrate the situation in figure below.



Schematic drawing of $\overline{\alpha}, \beta$ and $\gamma[\overline{\alpha}, \beta]$

By the choice of k, the number of tuples in $(L_p)^k$ is $2^{pk} \geq 2^{p|v_0|} \cdot 2^p > 2^{p|v_0|} \cdot p$, hence the number of tuples in $(L_p)^k$ exceeds the number of tuples of words in $\Sigma^{p|v_0|}$ and transition profiles of the automaton $\mathcal{A}_f$. Therefore, for every $\beta \in L_{id}$ there exist two distinct words $\overline{\alpha}, \overline{\alpha}' \in (L_p)^k$ such that for some $\delta \in \Sigma^{p|v_0|}$

- the *finite* words $(\overline{v}[\overline{\alpha}], \underline{\Box}, \delta)$ and $(\overline{v}[\overline{\alpha}'], \underline{\Box}, \delta)$ occur at positions $s + i\ell, i \in \mathbb{N}$ in $(\overline{\alpha}, \beta, \gamma[\overline{\alpha}, \beta])$ and $(\overline{\alpha}', \beta, \gamma[\overline{\alpha}', \beta])$, respectively, and
- these infixes $(\overline{v}[\overline{\alpha}], \underline{\Box}, \delta)$ and $(\overline{v}[\overline{\alpha}'], \underline{\Box}, \delta)$ have the same $\mathcal{A}_f$-transition profiles.

For every $\beta \in L_{id}$ we fix such a pair $(\overline{\alpha}, \overline{\alpha}')$. Furthermore, we recall that the number of words in L_{id} is 2^{2kp+p}. Since the number of pairs of $(L_p)^k$-tuples is 2^{2pk} there are at least $2^{2kp+p}/2^{2pk} = 2^p > p$ words $\beta \in L_{id}$ to which the same pair $(\overline{\alpha}, \overline{\alpha}')$ is assigned. Hence, we can also find two distinct $\beta, \beta' \in L_{id}$ with this property such that for some $\lambda, \lambda' \in \Sigma^{(2kp+p)|v_0|}$ the *finite* words $(\overline{\Box}^k, v[\beta], \lambda)$ and $(\overline{\Box}^k, v[\beta'], \lambda')$ have the same $\mathcal{A}_f$-transition profile, and occur at positions $h_i = (s + |v_0|p) + i\ell, i \in \mathbb{N}$ in $(\overline{\alpha}, \beta, \gamma[\overline{\alpha}, \beta])$ and $(\overline{\alpha}, \beta', \gamma[\overline{\alpha}, \beta'])$, respectively.

The proof now follows by examining the properties of $\overline{\alpha}, \overline{\alpha}' \in (L_p)^k$ and $\beta, \beta' \in L_{id}$. Let us consider the input $(\overline{\alpha}, \beta, \gamma[\overline{\alpha}, \beta])$ which is accepted by $\mathcal{A}_f$. By the properties stated above we can replace the infix $(\overline{v}[\overline{\alpha}], \underline{\Box}, \delta)$ in $(\overline{\alpha}, \beta, \gamma[\overline{\alpha}, \beta])$ by $(\overline{v}[\overline{\alpha}'], \underline{\Box}, \delta)$ at infinitely many positions. In this way we can obtain an uncountably infinite set $L_M \subseteq (L')^k$ with

- $\overline{\alpha}_0 \not\approx \overline{\alpha}_1$ if $\overline{\alpha}_0 \neq \overline{\alpha}_1$ (i.e. $\overline{\alpha}_0$ and $\overline{\alpha}_1$ encode different elements from A^k) for all $\overline{\alpha}_0, \overline{\alpha}_1 \in L_M$, and
- $(\overline{\alpha}_0, \beta, \gamma[\overline{\alpha}, \beta])$ is accepted by the automaton $\mathcal{A}_f$ for all $\overline{\alpha}_0 \in L_M$.

This is done in the following way: choose a set $X \subseteq \{0,1\}^\omega$ of the size 2^ω such that $x \not\sim_e y$ for all $x, y \in X$. This is possible since $\sim_e$ partitions $\{0,1\}^\omega$ into countable equivalence classes and $\{0,1\}^\omega$ has continuum many elements. For every $x \in X$ we define $\overline{\alpha}_x$ by

$$\overline{\alpha}_x[0, s) := \overline{\alpha}[0, s)$$

$$\overline{\alpha}_x[h_i, h_{i+1}) := \begin{cases} \overline{\alpha}[h_i, h_{i+1}) & \text{if } x[i] = 0 \\ \overline{\alpha}'[h_i, h_{i+1}) & \text{if } x[i] = 1. \end{cases}$$

We then define $L_M := \{\overline{\alpha}_x : x \in X\}$. It is easy to check that L_M has the claimed properties: $\overline{\alpha}_x, \overline{\alpha}_{x'} \subseteq L'$ and $\overline{\alpha}_x \not\sim_e \overline{\alpha}_{x'}$ for all $\overline{\alpha}_x \neq \overline{\alpha}_{x'} \in L_M$ and therefore $\overline{\alpha}_x \not\approx \overline{\alpha}_{x'}$. Since $\Delta((\overline{v}[\overline{\alpha}], \underline{\Box}, \delta)) = \Delta((\overline{v}[\overline{\alpha}'], \underline{\Box}, \delta))$, by Lemma 1, $(\overline{\alpha}_x, \beta, \gamma[\overline{\alpha}, \beta])$ is accepted by $\mathcal{A}_f$. In particular, we can still interchange the infixes $(\overline{\Box}^k, v[\beta], \lambda)$ and $(\overline{\Box}^k, v[\beta'], \lambda')$ in every input $(\overline{\alpha}_x, \beta, \gamma[\overline{\alpha}, \beta])$ in any way without affecting the acceptance behaviour of $\mathcal{A}_f$. We obtain a set $L_N \subseteq L'$ of uncountably many different tuples with the following properties

- $\beta \not\approx \beta'$ if $\beta \neq \beta'$ (i.e. β and β' encode different elements from A), and
- for every $\beta \in L_N$ there exists $\gamma \in \Sigma^\omega$ such that $(\overline{\alpha}, \beta, \gamma)$ is accepted by the automaton $\mathcal{A}_f$ for all for all $\overline{\alpha} \in L_M$.

Altogether for $M = \pi(L_M) \subseteq A^k$ and $N = \pi(L_N) \subseteq A$, we have $f(m,n) = f(m',n)$ for all $m, m' \in M$ and $n \in N$. Since both sets M and N are uncountable, the claim follows. □

To illustrate the applications of this result, consider the case of an integral domain. Recall that an integral domain is a commutative ring that has no zero divisors. It turns out that there exist no infinite ω-automatic integral domains.

Theorem 34 *An integral domain is ω-automatic if, and only if, it is finite.*

Proof One direction is trivial since all finite structures are ω-automatic. For the other direction, we recall from [8] that the (finite word) automatic integral domains are exactly the finite ones. Hence, by Theorem 5, there exist no countably infinite ω-automatic integral domains. Suppose now that $\mathfrak{A} = (A, +, \cdot)$ is an uncountable ω-automatic integral domain. Fix a presentation of $\mathfrak{A}$ and let k be the constant from Theorem 33 with respect to this presentation. Consider the family of polynomials of degree $k-1$, i.e. the family of functions of the form $x \mapsto \sum_{i=0}^{k} a_i x^i$ with k parameters $a_0, \dots, a_{k-1} \in A$ and input x. It is obvious that this family of functions can be defined in FOC by using the k coefficients $a_0, \dots, a_{k-1}$ as parameters.

On one hand, it is a well-known fact from algebra that, on an integral domain, two different polynomials of degree at most $k-1$ agree on at most $k-1$ inputs. On the other hand, $\mathfrak{A}$ is uncountable and therefore Theorem 33 implies that there are $\bar{a} \neq \bar{b} \in A^k$ such that $\sum_{i=0}^{k-1} a_i x^i = \sum_{i=0}^{k-1} b_i x^i$ for even uncountably many $x \in A$. □

Corollary 35 ([1]) *The field of reals is not ω-automatic.*

Sometimes it is convenient to apply Theorem 33 in the following simplified version.

Lemma 36 *Let $\mathfrak{A} = (A, R_1, \dots, R_n)$ be an uncountable ω-automatic structure. Then there is an $\ell \in \mathbb{N}$ such that for every definable ℓ-ary function $f(\bar{x})$ there is an uncountable set $M \subseteq A^\ell$ with $f(\bar{a}) = f(\bar{a}')$ for all $\bar{a}, \bar{a}' \in M$.*

Proof Let k be the constant from Theorem 33. We set $\ell = k+1$. Let f be an ℓ-ary function that is definable in $\mathfrak{A}$. Then by Theorem 33 there exist uncountably infinite sets $M' \subseteq A^k, N' \subseteq A$ with $f(\bar{a}, b) = f(\bar{a}, b')$. Hence, we can simply choose the uncountably infinite set $M = \{(\bar{a}, b) : \bar{a} \in M', b \in N'\}$ to satisfy the claim. □

Theorem 37 *There is no infinite ω-automatic structure with an* FOC*-definable pairing function.*

Proof Towards a contradiction, suppose there is an ω-automatic structure $\mathfrak{A}$ in which a pairing function f is definable. First we note that $\mathfrak{A}$ cannot be countable. Otherwise, by Theorem 5, $\mathfrak{A}$ would have an injective presentation. But $\mathrm{MIS}_f(n) = n^2$ in contradiction to Lemma 19. Therefore $\mathfrak{A}$ must be uncountable. In this case we obtain a contradiction to Corollary 36 by constructing a family of definable injective functions of unbounded arity. We let

$$f_1(x, y) := f(x, y) \quad \text{and}$$
$$f_{n+1}(x_1, \dots, x_{2^n}, y_1, \dots, y_{2^n}) := f\big(f_n(x_1, \dots, x_{2^n}), f_n(y_1, \dots, y_{2^n})\big).$$

It is easy to see that f_n is injective and FOC-definable in $\mathfrak{A}$ for fixed $n \geq 1$. □

Another interesting example to which we can apply our techniques are lattices.

Lemma 38 *Every uncountable ω-automatic lattice contains an element such that uncountably many elements are smaller than this element, and contains an element such that uncountably many elements are greater than this element.*

Proof Let $\mathfrak{A} = (A, <, \wedge, \vee)$ be an uncountable ω-automatic lattice. Fix an ω-automatic presentation and let c be the constant from Lemma 36. We consider the definable functions $f(x_1, \dots, x_k) := \bigwedge_{1 \leq i \leq k} x_i$ and $g(x_1, \dots, x_k) := \bigvee_{1 \leq i \leq k} x_i$. By Corollary 36 there must be elements $a, b \in A$ such that $f^{-1}(a)$ and $g^{-1}(b)$ are uncountable. But this is only possible if the sets $\{x : x \text{ appears in some } (x_1, \dots, x_c) \in f^{-1}(a)\} \subseteq \{x : x \leq a\}$ and $\{x : x \text{ appears in some } (x_1, \dots, x_c) \in g^{-1}(b)\} \subseteq \{x : x \geq b\}$ are uncountable. □

As a consequence we can reprove Kuske's theorem that no uncountable ordinal is ω-automatic.

Theorem 39 (Kuske [9]) *There is no uncountable ω-automatic ordinal.*

Proof First note that Lemma 38 directly implies that ω_1, the first uncountable ordinal, is not ω-automatic. Every ordinal is a lattice ($\wedge$ and $\vee$ can be defined) and for every element of ω_1 the number of elements below it is countable. But this implies that no larger ordinal α can be automatic either, since ω_1 is definable in all of these ordinals by the formula $\varphi(x) := \exists^{\leq \aleph_0} y (y < x)$. □

Note that this result also follows from Theorem 21, which is more closely related to the original proof of Kuske.

6 Future Work

A naturally arising question is whether our techniques, or similar ones, can be applied to characterise ω-automatic structures in other algebraic classes. One candidate could be the class of Boolean algebras, since Boolean algebras can be understood as special rings of characteristic 2. There are ω-automatic Boolean algebras that are not automatic (for cardinality reasons) and even theories, such as the theory of all atomless Boolean algebras, that have no automatic models, but an ω-automatic model. So far characterizations of all ω-automatic models are known only for classes where the ω-automatic model coincide with the finite word automatic ones.

Another area of application for the results of Sect. 3 is the notion of automaticity with advice. In this setting, the automata that recognise the domain and the relations of the structure are allowed to access a fixed infinite advice string ρ. Solving the theory of a structure represented in this way reduces to the ρ-acceptance problem, that is given a Büchi-automata $\mathcal{A}$, decide whether $\mathcal{A}$ accepts ρ. A prominent example is given by the rational numbers with addition, which are known to be not automatic [17], but automatic with a certain advice string [11]. This shows that in the presence of an advice string, non-automatic structures might become representable.

However, our techniques can also be applied with additional fixed parameters, and therefore they seem suitable to be used in this setting. Some first investigations let us hope that for many of the structures known to be not automatic it is possible to show that they are not automatic with advice either. Also the results of Sects. 4 and 5 might (at least to some extent) be lifted to this new setting, but here a closer inspection of the proofs, especially of Theorem 21 and of Theorem 33, will be necessary.

Another important direction is the consideration of ω-tree automatic presentations. Especially the question whether the field of reals is ω-tree-automatic is of considerable interest. However, our current techniques do not seem to be powerful enough to deal with this case. For ω-tree-automatic presentations, a technical result similar to Theorem 33 is very unlikely to hold. Indeed, in contrast to the case of ω-automatic structures, there is a tree-automatic (and therefore also ω-tree-automatic) structure with a pairing function, e.g. the free term algebra over countably many generators.

References

1. Zaid, F.A., Grädel, E., Kaiser, Ł.: The field of reals is not omega-automatic. In: Dürr, C., Wilke, T. (eds.) Proceedings of the International Symposium on Theoretical Aspects of Computer Science, STACS 2012, pp. 577–588 (2012)
2. Blumensath, A.: Automatic structures. Diploma thesis, RWTH-Aachen (1999)
3. Blumensath, A., Grädel, E.: Finite presentations of infinite structures: automata and interpretations. Theory Comput. Syst. **37**, 641–674 (2004)
4. Büchi, J.R.: On decision method in restricted second order arithmetic. In: Proceedings of the International Congress on Logic, Methodology and Philosophy of Science, pp. 1–11 (1962)
5. Delhommé, C.: Automaticité des ordinaux et des graphes homogènes. C. R. Acad. Sci., Sér. 1 Math. **339**, 5–10 (2004)
6. Hjorth, G., Khoussainov, B., Montalbán, A., Nies, A.: From automatic structures to Borel structures. In: Proceedings of the IEEE Symposium on Logic in Computer Science, LICS 2008, pp. 431–441 (2008)
7. Kaiser, Ł., Rubin, S., Bárány, V.: Cardinality and counting quantifiers on ω-automatic structures. In: Proceedings of the International Symposium on Theoretical Aspects of Computer Science, STACS 2008, pp. 385–396. Schloss Dagstuhl—Leibniz-Zentrum für Informatik, Germany (2008)
8. Khoussainov, B., Nies, A., Rubin, S., Stephan, F.: Automatic structures: Richness and limitations. Log. Methods Comput. Sci. **3**(2) (2007)
9. Kuske, D.: Is Ramsey's Theorem ω-automatic? In: Proceedings of the International Symposium on Theoretical Aspects of Computer Science, STACS 2010, pp. 537–548 (2010)
10. Kuske, D., Lohrey, M.: First-order and counting theories of ω-automatic structures. J. Symb. Log. **73**, 129–150 (2008)
11. Nies, A.: Describing groups. Bull. Symb. Log. **13**(3), 305–339 (2007)
12. Perrin, D., Pin, J.-E.: Semigroups and automata on infinite words. In: Semigroups, Formal Languages and Groups, York, 1993, pp. 49–72. Kluwer Academic, Dordrecht (1995)
13. Rabin, M.O.: Decidability of second-order theories and automata on infinite trees. Trans. Am. Math. Soc. **41**, 1–35 (1969)

14. Ramsey, F.P.: On a problem in formal logic. Proc. Lond. Math. Soc. **3**(30), 264–286 (1930)
15. Tarski, A.: A Decision Method for Elementary Algebra and Geometry. University of California Press, Berkeley (1951)
16. Thomas, W.: Languages, automata, and logic. In: Rozenberg, G., Salomaa, A. (eds.) Handbook of Formal Languages, vol. 3, pp. 389–455. Springer, New York (1997)
17. Todor, T.: The additive group of the rationals does not have an automatic presentation. J. Symb. Log. **76**(4), 1341–1351 (2011)
18. Wilke, T.: An algebraic theory for regular languages of finite and infinite words. Int. J. Algebra Comput. **3**(4), 447–489 (1993)