The Bulletin of Symbolic Logic
Volume 14, Number 2, June 2008

AUTOMATA PRESENTING STRUCTURES: A SURVEY OF THE FINITE STRING CASE

SASHA RUBIN

Abstract. A structure has a (finite-string) *automatic presentation* if the elements of its domain can be named by finite strings in such a way that the coded domain and the coded atomic operations are recognised by synchronous multitape automata. Consequently, every structure with an automatic presentation has a decidable first-order theory. The problems surveyed here include the classification of classes of structures with automatic presentations, the complexity of the isomorphism problem, and the relationship between definability and recognisability.

Contents

Received February 2, 2004.

1079-8986/08/1402-0001/$5.10

§1. Introduction. Which infinite structures can be stored and manipulated by a computer? At the very least, the structure should be representable in a finite amount of space. Moreover, operations that one would like to perform on the structure, typically logical queries, should be computable.

Finite automata are a robust model of computation that seem well suited to describing infinite structures. Although automata classically recognise sets of strings, they can be generalised to recognise n-ary relations by introducing n input-tapes. A (*finite-string*) *automatic presentation* of a structure is a coding of its domain as a set of finite strings so that the domain and each of the atomic operations is recognised by an automaton operating synchronously on its inputs (Definition 2.5)

Automatic presentations are a refinement of computable presentations, and have two salient properties: 1) given an automatic presentation of a structure $\mathcal{A}$, every first-order definable relation in $\mathcal{A}$, allowing parameters, is computable by a finite automaton (Theorem 3.1); 2) there are automatically presentable structures, such as certain expansions of the standard model of Presburger arithmetic, that interpret every automatically presentable structure (Theorem 3.8).

The effectiveness of using automata to present infinite structures has been amply displayed in the related concept of automatic groups. These are finitely-generated groups whose Cayley graphs (over a certain natural encoding) are recognised by automata. They were introduced by Thurston in 1986 who was motivated by work of Cannon [18] on hyperbolic groups. Automatic groups form a rich collection of finitely presented groups with tractable algorithmic properties that are undecidable in the general case. For instance, in an automatic group the word problem is solvable in quadratic time and the group elements can be efficiently enumerated. Moreover, every automatic group is also finitely presentable (in the group theoretic sense), and a presentation can be extracted from the automata. Software to find and work with automatic groups has been developed, and packaged with the computational algebra tools GAP and MAGMA; see Holt [33]. The standard reference for automatic groups is the book by Cannon, Epstein, Holt, Levy, Paterson, and Thurston [19]. Automatic groups are not covered in this survey.

The connections between automata and logic that are relevant to structures with automatic presentations can be traced to Büchi [16], Elgot [27] and Trahtenbrot [56]. They characterised, in logical terms, the relations over finite strings recognisable by automata (Theorem 3.8 (iv)). In particular they established the decidability of WS1S: the weak monadic second-order theory of the structure $\mathbb{N}$ with the successor function. This technique—relating definability and automata to solve decision problems for certain theories—was generalised to automata working on infinite strings (known as ω-automata) for the monadic second-order theory of one successor S1S [17], to automata on finite trees for WS2S [54, 25], and to automata on infinite trees for S2S [49]. However, structures presentable by these more general automata have received less study, and so this survey focuses solely on the finite string case.

The importance of these particular theories is that many first-order theories can be interpreted in them. For instance, Büchi proved that the first-order theory of $(\mathbb{N}, +)$ is decidable by showing that it is interpretable in WS1S. Taking this lead, Hodgson [30, 31, 32] called the first-order-theory of a structure 'automaton decidable', if after coding elements of the structure as finite or infinite strings, one can effectively construct automata recognising the encodings of its first-order definable relations. He was interested in the connections between two different ways of proving decidability: via automata and via algebraic constructions (such as direct product and direct power). He observed that certain products of (ω-)automatically presentable structures are again (ω-)automatically presentable, and thus decidable.[1]

Khoussainov and Nerode [34], independently of Hodgson and inspired by the success of using automata to describe groups as in [19], introduced structures presentable by automata as part of complexity-theoretic model theory (also called feasible model theory). Here is a brief description of this area.

The general idea is to fix a complexity class $\mathcal{C}$ (such as polynomial time, exponential time, etc.) and study algorithmic properties of $\mathcal{C}$-structures: those that can be represented by machines operating with complexity in $\mathcal{C}$. A typical question is whether a given infinite structure is isomorphic to a $\mathcal{C}$-structure. For instance, in the 1980's Nerode, Remmel, and Cenzer led the development of polynomial-time structures; see the survey [20]. They prove, for instance, that every computable structure is computably isomorphic to a polynomial-time structure over a binary alphabet.

Structures presentable by automata are part of complexity-theoretic model theory—take the complexity class corresponding to synchronous multitape automata.

Khoussainov and Nerode suggested in [34] that the algebraic-, model theoretic-, and complexity theoretic-properties of automatically presentable structures are amenable to systematic investigation. For instance, they characterised the class of automatically presentable structures via a generalisation of the Myhill-Nerode Theorem for regular languages [34, Theorem 3.3].

[1]An ω-automatic presentation, also called Büchi-automatic presentation, is similar to an automatic presentation, except that infinite strings rather than finite strings are used to code the domain of the structure.

https://doi.org/10.2178/bsl/1208442827

Blumensath and Grädel [9, 12] introduced automatic presentations of structures to the logic in computer science community. They established many fundamental results, such as the characterisation of automatically presentable structures via interpretability (Theorem 3.8). They discussed the complexity of various problems associated with evaluating formulas, such as model checking various fragments of first-order logic on automatically presentable structures.[2] They also provided the fundamental results for the classes of structures presentable by ω-automata and (ω)-tree automata.

My co-authors and I have focused on classifying classes of automatically presented structures in terms of classical invariants. For instance, what can be said about the Cantor-normal form of automatically presentable ordinals? (This question was asked by Khoussainov and Nerode [34] and solved by Delhommé [24]). Or what can be said about the Cantor–Bendixson rank of automatically presentable trees or of automatically presentable Boolean algebras? We aim for positive results of the form 'A structure from a certain class is automatically presentable if and only if it has certain algebraic properties', and negative results of the form 'The isomorphism problem for a certain class of automatically presented structures is complete for a certain level of the arithmetic/analytic hierarchy'. These are discussed in detail in Section 4. Since the condition of being automatically presentable is quite strong, it is not surprising that some classes of structures (ordinals and Boolean algebras for instance) are simple, in the sense that they are easy to describe. Surprisingly, some classes turn out to be complex, in the sense that it is hard to detect if two members are isomorphic. Notably, the isomorphism problem for the class of all automatically presented structures, easily seen to be undecidable, is Σ_1^1-complete (Theorem 3.53). Finally, there are many classes (groups, rings, and linear orders for instance) for which it is not yet known whether their automatically presentable members are simple or complex (in the senses described), or somewhere in-between.

A closely related problem is that of providing techniques for showing that a given structure does not have an automatic presentation. For instance, Delhommé [24] provides a necessary condition for a structure to have an automatic presentation, which implies, in particular, that the ordinal ω^ω does not (Corollary 3.52).

The following summary of the rest of this survey should give an indication of the main lines of research in the area.

1.1. Summary. Section 2 (Definitions) includes the definitions of regular relations and automatic presentations. The section ends with some common examples.

Section 3 (Properties) first covers general properties of automatically presentable structures, and later some specific topics.

[2]The model checking problem is, given a presentation of $\mathcal{A}$, a formula $\phi(\overline{x})$, and a tuple of parameters $\overline{a}$ in A, to decide whether or not $\mathcal{A} \models \phi(\overline{a})$.

Subsection 3.1 'Fundamental properties' includes the decidability result (Theorem 3.2), closure under natural operations (Corollary 3.7), and the logical characterisation of automatically presentable structures (Theorem 3.8).

Subsection 3.2 'Extending first-order' refers to extending the basic decidability result by certain additional quantifiers including 'there exists infinitely many' and 'there exist k modulo m many'. This is done in the context of generalised quantifiers and order-invariance.

Subsection 3.3 'Automatic presentations' offers a definition of when two presentations are equivalent, and gives a machine theoretic characterisation.

Subsection 3.4 'Intrinsic regularity' presents the analogue of intrinsically computably enumerable relations (see [1]) for automatically presentable structures.

Subsection 3.5 'Restriction on growth' presents some general ways of proving that a given structure has no automatic presentation. As illustration, the following structures are not automatically presentable: the free semigroup on $k > 1$ generators, Skolem arithmetic $(\mathbb{N}, \times)$, the random graph, and the ordinal ω^ω.

Subsection 3.6 'Isomorphism problem' includes the proof that the isomorphism problem for automatically presented structures is Σ^1_1-complete.

Section 4 (Classifications) is concerned with classifying the automatically presentable members of a given class of structures in relevant algebraic terms. There are many classes for which a complete classification is known for the restricted notion of automaticity requiring that only strings over a unary alphabet are used. In the general (non-unary) case, a complete classification is known for the classes of ordinals, Boolean algebras and fields. However, only partial classifications are known for the following classes: equivalence structures, linear orders, groups, and rings.

Section 5 (Open problems) contains a sample of problems whose solutions will likely require new ideas.

1.2. Scope and related work. Familiarity is assumed with the basics of finite automata, formal languages, and logic. Proofs will generally be sketched. The reader is referred to the literature for details: the original article [34]; the article [12] for an excellent reference of the fundamental results; and the theses [9], [51] and [3] which may serve as introductions to the area.

Automatically presentable structures can be generalised in several directions: for instance, by using finite automata on infinite strings [9, 12, 39], finite ranked trees [9, 5], finite unranked trees [40], or WMSO-interpretations (of dimension one) of trees with decidable WMSO-theory [21] (see Definition 3.4). The general theory goes through: decidability of the first-order theory and characterisations via interpretability in some structure. Problems such as classification and techniques for proving non-automaticity in these more general settings are not dealt with in this survey. Consult [24, 21] for techniques showing that a structure has no finite-tree automatic presentation; and [39, 4] for the addition of generalised quantifiers 'there exists uncountably many', 'there exists countably many' and 'there exists k modulo m many' in ω-automatically presentable structures; and [21] for intrinsic regularity and proving non-automaticity in WMSO-interpreted structures.

Automatic groups [19] are not in the scope of this survey. See [14] for some remarks relating them to automatically presentable structures.

For the complexity of model-checking and related problems, consult [9, 12] for fragments of FO, and [43] for structures of bounded degree.

Definability issues (VC-dimension, quantifier elimination) in universal string- and tree-automatically presentable structures and their reducts can be found in [40, 5, 6].

1.3. Acknowledgement. I thank Valentin Goranko for discussions on intrinsic regularity, Lauri Hella for discussions on generalised quantifiers, André Nies for some corrections, Wolfgang Thomas for clarifying Trahtenbrot's historical contribution, and Vince Bárány, Michael Benedikt and Alexander Raichev for their comments.

§2. Definitions.

2.1. Notation. Countable means finite or countably infinite. A *relational signature* τ is a countable sequence of symbols $(R_i)_i$ and corresponding arities r_i. A τ-structure $\mathcal{A} = (A; (R_i^{\mathcal{A}})_i)$ consists of a countable set A, called the *domain* of $\mathcal{A}$, written $\mathrm{dom}(\mathcal{A})$, and for each i, a relation $R_i^{\mathcal{A}} \subseteq A^{r_i}$, called an *atomic relation* of $\mathcal{A}$. When there is only one structure around, I may drop the superscript with the name of the structure.

Structures are written in script $\mathcal{A}, \mathcal{B}, \dots$ and their corresponding domains are written in capitals $A, B, \dots$. The substructure of relational $\mathcal{A}$ on set $B \subseteq A$ is written $\mathcal{A} \restriction B$ or $\mathcal{B}$ if there can be no confusion.

Signatures are assumed to be computable (the mappings $i \mapsto R_i$ and $i \mapsto r_i$ are computable), and to contain the equality symbol $=$, though this symbol may not be explicitly mentioned in the signature. For convenience, I sometimes write a structure containing functions, such as $(\mathbb{N}, +)$, but am implicitly referring to its *relational variant* obtained by replacing every function $f : B^k \to B$ by its graph $\{(\overline{x}, y) \in B^{k+1} \mid f(\overline{x}) = y\}$.

If unspecified, all formulas $\phi(\overline{x})$ (and associated notions like definability) are first-order and allow parameters. However, we will see extensions $\mathcal{L}$ of first-order logic, particularly monadic second-order logic and extensions by generalised quantifiers.

An $\mathcal{A}$-formula is a formula over the signature of structure $\mathcal{A}$. The *relation in $\mathcal{A}$ defined by $\Phi(\overline{x}, \overline{y})$ with parameters $\overline{b}$*, denoted by $\Phi^{\mathcal{A}}(\cdot, \overline{b})$, is defined as

$$\{(a_1, \dots, a_m) \mid \mathcal{A} \models \Phi(\overline{a}, \overline{b})\},$$

where $\mathcal{A}$ is a τ-structure, $\Phi(\overline{x}, \overline{y})$ is an $\mathcal{A}$-formula with free variables $\overline{x} = (x_1, \dots, x_m)$, and $\overline{b}$ a tuple from A. To ease readability, I often relax the notation and write $\Phi^{\mathcal{A}}(\overline{b})$ or even $\Phi^{\mathcal{A}}$.

Familiarity is assumed with the basics from automata theory. To fix notation: symbols will usually be denoted $a, b, \dots$; a finite alphabet of symbols by Σ; finite strings by $u, v, w, \dots$; the set of finite strings over Σ by $\Sigma^\star$; the set of infinite strings over Σ by Σ^ω; the empty string by λ; concatenation by $\cdot$ as in $w \cdot v$, or simply by juxtaposition, as in wv; the concatenation of a symbol a with itself n times by a^n (a^0 is defined as λ); the length of a string w by $|w|$; the strings in a set $A \subseteq \Sigma^\star$ of length exactly n by $A^{=n}$, and those of length at most n by $A^{\leq n}$. These should not be confused with the set A^n of n-tuples from a set or domain A. A deterministic finite automaton M over alphabet Σ is of the form (Q, ι, Δ, F) where Q is a finite set of states, $\iota \in Q$ is the initial state, $\Delta \colon Q \times \Sigma \to Q$ is the transition function, and $F \subseteq Q$ is the set of accepting states.

The logarithmic and exponential functions are taken with base 2. In particular, the functions $\exp_k \colon \mathbb{N} \to \mathbb{N}$ are defined for $k \in \mathbb{N}$ recursively by $\exp_0(n) = n$ and $\exp_{k+1}(n) = 2^{\exp_k(n)}$. The subscript in $\exp_1$ will be dropped.

2.2. Synchronous finite automata. Consider a finite automaton as a restricted non-deterministic Turing machine: it has a read-only input-tape with one head that only moves in one direction. It has no work tape. By admitting more than one input-tape, say n many, each with its own head moving independently, the language computed by such a machine is an n-ary relation. The resulting relations are called *rational relations*. These multi-tape automata do not share the robustness of one-tape automata: they are not closed under the Boolean operations, and the nondeterministic rational relations strictly contain the deterministic rational relations.

This article deals with particular rational relations that do enjoy strong closure properties (Theorem 2.4), namely those recognisable by *synchronous n-tape automata* (also called letter-to-letter automata). The following informal description follows Eilenberg, Elgot, and Shepherdson [26]. A synchronous n-tape automaton can be thought of as a one-way Turing machine with n input-tapes. Each tape is regarded as semi-infinite having written on it a string over the alphabet Σ followed by an infinite succession of blanks, $\perp$ symbols. The automaton starts in the initial state, reads simultaneously the first symbol of each tape, changes state, reads simultaneously the second symbol of each tape, changes state, etc., until it reads a blank on each tape. The automaton then stops and accepts the n-tuple of strings if it is in an accepting state. The set of all n-tuples accepted by the automaton is the relation recognised by the automaton. Since n–tape automata are simply 1–tape automata over a new alphabet, they can be determinised by the usual subset construction.

Instead of formalising this model of computation, we encode a tuple of strings from $\Sigma^\star$ as a single string over an expanded alphabet. For example, for $\Sigma = \{1\}$ the expanded alphabet is $\{\binom{1}{1}, \binom{1}{\perp}, \binom{\perp}{1}, \binom{\perp}{\perp}\}$ and the string associated with the tuple $(1^m, 1^{m+1})$ is $\binom{1}{1}^m \binom{\perp}{1}$.

Definition 2.1. Write $\Sigma_\perp$ for $\Sigma \cup \{\perp\}$, where $\perp$ is a symbol not in Σ. The *convolution of a tuple* $(w_1, \dots, w_n) \in (\Sigma^\star)^n$ is the string $\otimes(w_1, \dots, w_n)$ over alphabet $(\Sigma_\perp)^n$, of length $\max_i |w_i|$, defined as follows. Its kth symbol is $(a_1, \dots, a_n)$ where a_i is the kth symbol of w_i if $k \leq |w_i|$ and $\perp$ otherwise. In particular, $\otimes(\lambda, \dots, \lambda) = \lambda$.

The *convolution of a relation* $R \subseteq (\Sigma^\star)^n$ is the set $\otimes R = \{\otimes \overline{w} \mid \overline{w} \in R\}$; that is, $\otimes R \subseteq (\Sigma_\perp)^{n\star}$ is the set of convolutions of all the tuples in R.

Definition 2.2. A relation $R \subseteq (\Sigma^\star)^n$ is called *synchronous rational*, or simply *regular*, if there is a finite automaton over alphabet $(\Sigma_\perp)^n$ recognising the convolution $\otimes R$.

In case $n = 1$, the convolution $\otimes R$ equals R, and so in this case the definition coincides with the traditional class of regular languages.

Examples 2.3. The following relations on $\Sigma^\star$ are regular.

(i) The prefix relation $\preceq_p$.
(ii) The equal length relation $\mathrm{el}(w, v)$ if $|w| = |v|$.
(iii) The longest common prefix relation, written functionally $x \sqcap y = z$.
(iv) The prefix-lexicographic ordering (induced by a fixed ordering $<$ on Σ) defined by $x \leq_{\mathrm{lex}} y$ if $x \preceq_p y$ or otherwise $za \preceq_p x$ and $zb \preceq_p y$ and $a < b$ where $z = x \sqcap y$.
(v) The length-lexicographic ordering $\leq_{\mathrm{llex}}$ defined by $x \leq_{\mathrm{llex}} y$ if $|x| < |y|$ or otherwise $|x| = |y|$ and $x \leq_{\mathrm{lex}} y$.

The synchronous coding ensures that the regular relations have basic closure properties.

Theorem 2.4. *Let $R, S \subseteq \Sigma^{\star n}$ be regular relations. Then the following relations are also regular*:

(i) *union $R \cup S$, intersection $R \cap S$, relative complementation $R \setminus S$;*
(ii) *projection $\{\overline{y} \mid (\exists x) R(x, \overline{y})\}$;*
(iii) *instantiation $\{\overline{y} \mid R(w, \overline{y})\}$ for fixed $w \in \Sigma^\star$;*
(iv) *cylindrification $\{(x, \overline{y}) \mid x \in \Sigma^\star \wedge R(\overline{y})\}$; and*
(v) *permutation of the co-ordinates of R.*

Moreover, there is an effective procedure that given automata for $\otimes R$ and $\otimes S$, constructs an automaton for the convolution of each of the resulting relations.

2.3. Automatic presentations. We are ready to define what it means for a structure to be described by a collection of automata. Although the definition is only for relational structures, recall the convention that we implicitly replace functions by their graphs.

https://doi.org/10.2178/bsl/1208442827

DEFINITION 2.5. A (*finite-string*) *automatic presentation* of a relational structure $\mathcal{B} = (B; (R_i^{\mathcal{B}})_i)$ consists of a mapping μ and a tuple of automata $\overline{\mathsf{M}} = (\mathsf{M}_A, \mathsf{M}_=, (\mathsf{M}_i)_i)$ so that

(i) M_A recognises a set $A \subseteq \Sigma^\star$,
(ii) the mapping $\mu\colon A \to \mathrm{dom}(\mathcal{B})$ is surjective, and
(iii) for every atomic relation $R_i^{\mathcal{B}}$ of $\mathcal{B}$ (say of arity r_i), the relation

$$\mu^{-1}(R_i^{\mathcal{B}}) := \{(w_1, \dots, w_{r_i}) \in A^{r_i} \mid R_i^{\mathcal{B}}(\mu(w_1), \dots, \mu(w_{r_i}))\},$$

is regular; and its convolution is recognised by the automaton M_i.

Since by our convention there is a symbol for equality, the relation

$$\{(w_1, w_2) \in A^2 \mid \mu(w_1) =^{\mathcal{B}} \mu(w_2)\}$$

need also be regular; and its convolution is recognised by the automaton $\mathsf{M}_=$.

Say that $\mathcal{B}$ is *automatically presentable*. Depending on the focus, the *automatic presentation* may simply refer to the mapping μ or to the tuple of automata $\overline{\mathsf{M}}$.

The induced quotient structure

$$(A; (\mu^{-1}(R_i^{\mathcal{B}}))_i)/\mu^{-1}(=^{\mathcal{B}})$$

is an *automatic copy* of $\mathcal{B}$.

If the signature of $\mathcal{B}$ is infinite, we also require that the function mapping $i \in \mathbb{N}$ to (a code for) the automaton M_i be computable.

If $|\Sigma| = 1$, then we speak of a *unary-automatic presentation*.

The idea is that strings from A code elements of B (via the mapping μ) so that the induced relations $\mu^{-1}(R_i^{\mathcal{B}})$ are regular. In general, an element of B may have more than one code. However, if μ is also an injection, then every element of B is coded by a unique string of A. In this case, μ is called an *injective automatic presentation* of $\mathcal{B}$.

PROPOSITION 2.6. *Every automatically presentable structure $\mathcal{B}$ has an injective automatic presentation ν. Moreover, ν can be chosen over a binary alphabet* (*that is, with* $|\Sigma| = 2$).

PROOF. Let $\mu\colon A \to B$ be an automatic presentation of $\mathcal{B}$ over alphabet Γ. For the first item, it is sufficient to restrict A to a regular set A' so that $\mu\colon A' \to B$ is a bijection. To this end, define A' as the set of strings $v \in A$ such that v is the length-lexicographically least string in the set $\{w \in A \mid \mu(w) = \mu(v)\}$. It is straightforward to check that A' is regular.

For the second item, suppose $k = |\Gamma| > 2$ and let Σ be a binary alphabet. Fix some integer $l \geq \log_2 k$. The idea is to code each symbol $\gamma \in \Gamma$ by a distinct string $\alpha(\gamma) \in \Sigma^\star$ of length l. Then extend α to $\Gamma^\star$ in the natural way (define $\alpha(uv) = \alpha(u)\alpha(v)$). The required presentation is $\nu\colon \alpha(A') \to B$ defined by $\nu(\alpha(v)) = \mu(v)$ for $v \in A'$. ⊣

2.4. Examples.

(i) Every finite structure is automatically presentable.

(ii) For $\Sigma = \{0, \ldots, k-1\}$ ($k \in \mathbb{N}$), define the structure $\mathcal{W}_k$ as

$$(\Sigma^\star; (\sigma_a)_{a\in\Sigma}, \preceq_p, \mathrm{el}).$$

Here $\sigma_a(w) = wa$, $\preceq_p$ is the prefix relation, and $\mathrm{el}(w, v)$ if $|w| = |v|$. It is automatically presentable.

(iii) The structure

$$\mathcal{N}_k = (\mathbb{N}; +, |_k)$$

is automatically presentable where $+$ is the usual addition on the naturals $\mathbb{N}$, and $x|_k y$ means that $x = k^n$ for some $n \in \mathbb{N}$ and $y = mx$ for some $m \in \mathbb{N}$ (that is, x is a power of k and x divides y). One gets an automatic presentation for this structure by coding the natural numbers in base k (least significant digit first). Indeed, a finite automaton can check addition using the usual carry-bit procedure, and it is similarly straightforward to verify that the relation $|_k$ is regular.

(iv) The configuration space of a Turing Machine N, considered as a directed graph, whose edges represent one step transitions of N, is automatically presentable. The idea is that an automaton can compute these transitions simply following a window of fixed size around the current position of the read-head.

(v) The linear ordering of the rationals is automatically presentable. The structure with domain $\{0,1\}^\star$ and the binary relation $x \sqsubseteq_Q y$ if $(x \sqcap y)0$ is a prefix of x or $(x \sqcap y)1$ is a prefix of y (here $x \sqcap y$ denotes their longest common prefix) is an automatic copy.

(vi) Every finitely generated Abelian group $(G; +)$ is automatically presentable. This is straightforward given the classification of these groups as finite sums of $(\mathbb{Z}, +)$ and finite cyclic groups.

(vii) Every ordinal $(L; \leq)$ less than ω^ω is automatically presentable. This is straightforward given the classification of these ordinals as being of the form $\omega^{n_1} + \cdots + \omega^{n_l}$ where $\omega > n_1 \geq n_2 \geq \cdots \geq n_l$.

(viii) The Boolean algebra $(\mathcal{B}_{\mathrm{fin}})^n$ ($n \in \mathbb{N}$) is automatically presentable; here $\mathcal{B}_{\mathrm{fin}} = (B_{\mathrm{fin}}; \wedge, \vee, \neg)$ is the Boolean algebra of finite or co-finite subset of $\mathbb{N}$.

§3. Properties.

3.1. Fundamental properties. As a consequence of Theorem 2.4 one has the following result that may be called the fundamental theorem of automatically presentable structures.

THEOREM 3.1 (Definability). *For every automatic presentation μ, of structure $\mathcal{A}$ say, every first order $\mathcal{A}$-formula (allowing parameters) defines a relation R with $\mu^{-1}(R)$ regular.*

Moreover there is an algorithm that from the automata of an automatic presentation, and a first-order formula allowing parameters, outputs an automaton for the convolution of the relation defined by the formula.

This is simply proved by structural induction on the formula defining the relation. Consequently:

THEOREM 3.2 (Decidability). *The first-order theory (allowing parameters) of each automatically presentable structure is decidable.*

Indeed, from any automatic presentation of $\mathcal{A}$, and sentence of the form $\exists x \Phi(x)$, one can check effectively whether or not the constructed automaton for $\Phi^{\mathcal{A}}(x)$ accepts any string at all.

Remark 3.3. We shall see in subsection 3.2 that the Definability and Decidability Theorems can be extended to include additional quantifiers such as 'there exists infinitely many' and 'there exists k modulo m many'.

Another consequence of the Definability Theorem is that automatically presentable structures are closed under first-order interpretability.

DEFINITION 3.4. An *interpretation* of structure $\mathcal{B} = (B; (R_i^{\mathcal{B}})_i)$ in structure $\mathcal{A}$ consists of the following $\mathcal{A}$-formulas,

(i) a *domain formula* $\Delta(\overline{x})$,
(ii) a *relation formula* $\Phi_{R_i}(\overline{x}_1, \dots, \overline{x}_{r_i})$ for each relation symbol R_i, and
(iii) an *equality formula* $\epsilon(\overline{x}_1, \overline{x}_2)$,

where each $\Phi_{R_i}^{\mathcal{A}}$ is a relation on $\Delta^{\mathcal{A}}$, and $\epsilon^{\mathcal{A}}$ is a congruence on the structure $(\Delta^{\mathcal{A}}; (\Phi_{R_i}^{\mathcal{A}})_i)$, so that $(\Delta^{\mathcal{A}}; (\Phi_{R_i}^{\mathcal{A}})_i)/\epsilon^{\mathcal{A}}$ and $\mathcal{B}$ are isomorphic.

Each of the tuples $\overline{x}_i, \overline{x}$ contain the same number of variables, called the *dimension* of the interpretation.

We say that $\mathcal{B}$ *is interpretable in* $\mathcal{A}$.

Also, say that $\mathcal{B}$ is interpretable in $\mathcal{A}$ *with parameters* $\overline{a}$ if $\mathcal{B}$ is interpretable in $(\mathcal{A}, \overline{a})$.

In case $\mathcal{B}$ has infinite signature, it is also required that the function sending i to Φ_{R_i} be computable.

To stress that all the formulas are first-order, we say that $\mathcal{B}$ is *first-order interpretable* in $\mathcal{A}$. In this case, every first-order $\mathcal{B}$-formula ϕ can be translated into a first-order $\mathcal{A}$-formula ϕ' such that for all $\overline{a} \in \Delta^{\mathcal{A}}$,

$$\mathcal{B} \models \phi(\mu(\overline{a})) \iff \mathcal{A} \models \phi'(\overline{a}),$$

where μ is the isomorphism from the interpretation. The formula ϕ' is formed from ϕ by relativising the quantification to Δ, replacing every relation symbol R_i by its defining formula Φ_{R_i}, and replacing equality by ϵ.

We will have brief occasion to consider the case that the formulas are (weak) monadic second-order, whence we say that $\mathcal{B}$ is (*weak*) *monadic*

https://doi.org/10.2178/bsl/1208442827

second-order interpretable in $\mathcal{A}$.[3] In this case, elements of $\mathcal{B}$ are coded by (finite) subsets of A. As an illustration, $(\mathbb{N}, +)$ is weak monadic second-order interpretable in $(\mathbb{N}, S)$ via the map sending a finite set $X \subset \mathbb{N}$ to the natural number $\Sigma_{i \in X} 2^i$.

EXAMPLE 3.5. The structures $\mathcal{N}_k$ and $\mathcal{W}_l$ (see Example 2.4) are first-order interpretable in each other, for every $k, l \geq 2$.

The following proposition is immediate from the definitions and Theorem 3.1 [Definability].

PROPOSITION 3.6. [9] *If $\mathcal{B}$ is first-order interpretable in $\mathcal{A}$ (possibly with parameters), and $\mathcal{A}$ is automatically presentable, then $\mathcal{B}$ is also automatically presentable.*

The following corollary lists some useful closure properties of automatically presentable structures, some of which were already noted in [34]. It is worth mentioning that, more generally, Blumensath and Grädel [9, 12] show that automatically presentable structures are closed under Feferman-Vaught like products.

Recall that definable means first-order definable allowing parameters.

COROLLARY 3.7. *If structures $\mathcal{A}$ and $\mathcal{B}$ (with the same signature) are automatically presentable, then the following are also automatically presentable:*

(i) *the expansion $(\mathcal{A}, R)$ by any relation definable in $\mathcal{A}$,*
(ii) *the substructure of $\mathcal{A}$ with any definable universe,*
(iii) *the factorisation of $\mathcal{A}$ by any definable congruence,*
(iv) *the direct product $\mathcal{A} \times \mathcal{B}$, the disjoint union $\mathcal{A} \cup \mathcal{B}$,*
(v) *the ω-fold disjoint union of $\mathcal{A}$, written $\cup_\omega \mathcal{A}$,*
(vi) *the ordered sum $\mathcal{A} + \mathcal{B}$ and the ω-fold ordered sum $\Sigma_{i<\omega} \mathcal{A}$ (for this case $\mathcal{A}$ and $\mathcal{B}$ are ordered structures).*

Next is a logical characterisation of the automatically presentable structures.

THEOREM 3.8. *Let $\mathcal{B}$ be a structure. Then the following are equivalent:*

(i) *$\mathcal{B}$ is automatically presentable.*
(ii) *$\mathcal{B}$ is first-order interpretable in $\mathcal{W}_k$ for some, equivalently all, $k \geq 2$.*
(iii) *$\mathcal{B}$ is first-order interpretable in $\mathcal{N}_k$ for some, equivalently all, $k \geq 2$.*
(iv) *$\mathcal{B}$ is weak monadic second-order interpretable in $(\mathbb{N}, S)$, where S is the usual successor function $n \mapsto n + 1$.*

Here is a brief history of this theorem. The equivalences (i)–(iii) are from Blumensath and Grädel [12] who introduced interpretability to characterise the automatically presentable structures. The main ideas arise from older

[3]These are called *(finite) set interpretations* in [21] if the interpretation also has dimension one. The present definition is not to be confused with the case that the formulas are (W)MSO but all free variables are first-order, sometimes also called (W)MSO-interpretations in the literature.

https://doi.org/10.2178/bsl/1208442827

results stated for relations instead of structures: Büchi [16] and Elgot [27] prove that a set of tuples $(A_1, \dots, A_n)$ of finite sets of natural numbers is weak monadic second-order definable in $(\mathbb{N}, \mathrm{S})$ if and only if the corresponding n-ary relation of characteristic strings (a subset of $\{0,1\}^{\star n}$) is synchronous rational. The relationship between weak MSO and finite automata was also realised by Trahtenbrot [56]. Eilenberg, Elgot, and Shepherdson [26] prove that a relation $R \subseteq (\Sigma^\star)^n$ is synchronous rational if and only if R is first-order definable in $\mathcal{W}_k$, where $k = |\Sigma| \geq 2$. The Büchi-Bruyère Theorem (proofs of which can be found in [44] and [57]) states that a relation $R \subseteq \mathbb{N}^n$ (coded in base $k \geq 2$ least significant digit first) is synchronous rational if and only if it is first-order definable in $\mathcal{N}_k$.

The proofs of these results, which are by now standard, usually go as follows. From formulas to automata one proceeds by structural induction on the given formula, using Theorem 2.4 for the logical operations. Conversely, starting with an automaton, one constructs a formula stating the existence of a successful run of the automaton (alternatively, [57] proceeds by induction on regular expressions).

We turn to unary-automatically presentable structures. An important example is the structure $\mathcal{U} = (\mathbb{N}; \leq, (\equiv_n)_{n\in\mathbb{N}})$ where $x \equiv_n y$ if x is congruent to y modulo n.

Theorem 3.9. [9, 45] *A structure is automatically presentable over a unary alphabet if and only if it is first-order interpretable* (*via a* 1*-dimensional interpretation*) *in the structure* $(\mathbb{N}; \leq, (\equiv_n)_{n\in\mathbb{N}})$.

Blumensath [11] calls a structure $\mathcal{A}$ *tree-interpretable* if there is a monadic second-order interpretation of dimension 1 of $\mathcal{A}$ into $\mathcal{T}_2$ with the restriction that all the free variables in the interpreting formulas are first-order. Here $\mathcal{T}_2 = (\{0,1\}^\star; \sigma_0, \sigma_1)$ where $\sigma_\epsilon(w) = w\epsilon$ for all $w \in \{0,1\}^\star$ and $\epsilon \in \{0,1\}$. Since $\mathcal{T}_2$ has decidable monadic second-order theory [49], so does $\mathcal{A}$.

Since $\mathcal{U}$ is tree-interpretable, every unary-automatically presentable structure has decidable monadic second-order theory [10, Prop. 5]. Moreover, Blumensath proves that every tree-interpretable structure is automatically presentable (the converse is false because there are automatically presentable structures, such as the infinite grid, with undecidable monadic-second order theory). The tree-interpretable graphs coincide with other classes of infinite graphs that appear in the literature; see [10] for a discussion.

3.2. Extending first-order. The Definability Theorem can be strengthened to include order-invariantly definable formulas as well as certain additional quantifiers.

Order-invariance.

Definition 3.10. Fix a signature τ and a new symbol $\leq$. A formula $\phi(\overline{x})$ in the signature $\tau \cup \{\leq\}$, is called *ω-order invariant* on a τ-structure $\mathcal{A}$, if for all tuples $\overline{a}$ from A, and all linear orders $\leq_1$ and $\leq_2$ on A of order type ω (or $|A|$ if A is finite), we have that $(\mathcal{A}, \leq_1) \models \phi(\overline{a})$ if and only if $(\mathcal{A}, \leq_2) \models \phi(\overline{a})$.

The *relation in $\mathcal{A}$ defined by the ω-order invariant ϕ* is the set of tuples $\overline{a}$ from A such that $(\mathcal{A}, \leq) \models \phi(\overline{a})$ for some (and hence all) linear orders $\leq$ on A of order type ω (or $|A|$ if A is finite).

Write $\mathcal{L}_{\omega-\mathrm{inv}}(\mathcal{A})$ for the relations that are definable by ω-order invariant $\mathcal{A}$-formulas in logic $\mathcal{L}$.

Remark 3.11. Every automatic presentation μ of structure $\mathcal{A}$ can be expanded to an automatic presentation of an expansion $(\mathcal{A}, \leq)$ where $\leq$ linearly orders A. For instance let $\mu^{-1}(\leq)$ be the regular linear order $<_{\mathrm{llex}}$ restricted to the domain $\mu^{-1}(A)$. Observe that this linear order has order-type ω if A is infinite.

Thus we can replace FO by $\mathrm{FO}_{\omega-\mathrm{inv}}$ in Theorem 3.1 [Definability].

Generalised quantifiers. We briefly recall the definition of generalised quantifiers as introduced by Lindström [41].

DEFINITION 3.12. Fix a finite signature $\tau = (R_i)_{i \leq k}$, where R_i has associated arity r_i. A *quantifier* Q is a class of τ-structures closed under isomorphism. Let σ be another signature. Given σ-formulas $\Psi_i(\overline{x}_i, \overline{z})$ with $|\overline{x}_i| = r_i$ $(i \leq k)$, the syntax $Q\overline{x}_1, \dots, \overline{x}_n(\Psi_1, \dots, \Psi_k)$ has the following meaning on a σ-structure $\mathcal{A}$:

$$(\mathcal{A}, \overline{a}) \models Q\overline{x}_1, \dots, \overline{x}_k(\Psi_1, \dots, \Psi_k) \text{ iff } (A; \Psi_1^{\mathcal{A}}(\cdot, \overline{a}), \dots, \Psi_k^{\mathcal{A}}(\cdot, \overline{a})) \in Q,$$

where $\Psi^{\mathcal{A}}(\cdot, \overline{a})$ is the relation defined in $\mathcal{A}$ by Ψ with parameters $\overline{a}$. The *arity* of a quantifier is the maximum of the r_is. A quantifier is *n-ary* if its arity is at most n.

The *relation defined (in $\mathcal{A}$) by this formula* is the set of tuples $\overline{a}$ from A such that $(\mathcal{A}, \overline{a}) \models Q\overline{x}_1, \dots, \overline{x}_k(\Psi_1, \dots, \Psi_k)$.

A collection of quantifiers is denoted by $\boldsymbol{Q}$, and the extension of first-order logic by the quantifiers in $\boldsymbol{Q}$ is written FO[$\boldsymbol{Q}$].

As an illustration, 'there exists' is given by the unary quantifier

$$\{(A; X) \mid \emptyset \neq X \subseteq A\}.$$

Here are some quantifiers that will feature in this section:

EXAMPLES 3.13. (i) The unary quantifier 'there exists infinitely many', written $\exists^\infty$, is the class of structures $(A; X)$ where X is an infinite subset of A.

(ii) The unary *modulo quantifier* 'there are k modulo m many' (here $0 \leq k < m$), written $\exists^{(k,m)}$, is the class of structures $(A; X)$ where X contains k modulo m many elements. Write $\exists^{\mathrm{mod}}$ for the collection of modulo quantifiers.

(iii) The unary *Härtig quantifier* is the class of structures $(A; P, Q)$ where $P, Q \subseteq A$ and $|P| = |Q|$.

(iv) The *cardinality quantifiers*: Every set $C \subseteq (\mathbb{N} \cup \{\infty\})^n$ induces the unary quantifier $Q_C = \{(A; P_1, \ldots, P_n) \mid (|P_1|, \ldots, |P_n|) \in C\}$. In fact, a given unary quantifier over signature $(R_i)_{i \leq k}$ is identical to some cardinality quantifier with $n = 2^k$.

(v) The binary *reachability quantifier* is the class of structures of the form $(A; E, \{c_s\}, \{c_f\})$ where $E \subseteq A^2$, $c_s, c_f \in A$, and there is a path in the directed graph $(A; E)$ from c_s to c_f.

(vi) The k-ary *Ramsey quantifier* $\exists^{\text{k-ram}}$ is the class of structures $(A; E)$, $E \subseteq A^k$, for which there is an infinite $X \subseteq A$ such that for all pairwise distinct $x_1, \ldots, x_k \in X$, $E(x_1, \ldots, x_k)$.

Of course the generalised quantifiers that interest us most are the ones, like $\forall$ and $\exists$, that preserve regularity; meaning, loosely, that quantifying over regular relations yields regular relations.

Definition 3.14. Let Q be a quantifier with signature $\tau = (R_i)_{i \leq k}$, where R_i has associated arity r_i. Say that the quantifier *preserves regularity* if for every $n \in \mathbb{N}$, and every automatic presentation μ of a structure $\mathcal{A}$, every formula

$$Q\overline{x}_1, \ldots, \overline{x}_k(\Psi_1^{\mathcal{A}}(\overline{x}_1, \overline{z}), \ldots, \Psi_k^{\mathcal{A}}(\overline{x}_k, \overline{z}))$$

defines a relation R in $\mathcal{A}$ so that $\mu^{-1}(R)$ is regular (here $\overline{z} = (z_1, \ldots, z_n)$ and the Ψ_is are first-order $\mathcal{A}$-formulas).

Moreover, say that Q *preserves regularity effectively* if an automaton for $\mu^{-1}(R)$ can effectively be constructed from the automata of the presentation and the formulas Ψ_i.

Remark 3.15. Since every automatic presentation restricts (effectively) to an injective automatic presentation (proof of Proposition 2.6), to show that Q preserves regularity (effectively) it is sufficient to satisfy the above definition with ‘injective automatic presentation’ replacing ‘automatic presentation’.

Example 3.16. [9] The quantifier $\exists^{\infty}$ preserves regularity effectively. Let μ be an injective automatic presentation of structure $\mathcal{A}$, and $\Psi(x, \overline{z})$ a first-order $\mathcal{A}$-formula. Then the relation defined in $\mathcal{A}$ by $(\exists^{\infty} x)\Psi(x, \overline{z})$ is equal to the relation defined by the ω-order invariant formula

$$(\forall w)(\exists x)[w < x \wedge \Psi(x, \overline{z})].$$

Here is a non-example.

Example 3.17. The reachability quantifier is not regularity preserving. For otherwise, by Example 2.4 (4), the set of starting configurations that drive a given Turing Machine to a halting state would be regular, and hence computable.

Remark 3.18. In our new terminology, Theorem 3.1 [Definability] says that $\exists$ and $\forall$ preserve regularity effectively. Moreover if every quantifier in a collection $\boldsymbol{Q}$ preserves regularity effectively, then we can replace FO by FO[$\boldsymbol{Q}$] in Theorem 3.1 [Definability] and Theorem 3.2 [Decidability]. If quantifiers in $\boldsymbol{Q}$ simply preserve regularity, then we can replace FO by FO[$\boldsymbol{Q}$] in Proposition 3.6, Corollary 3.7 and parts two and three of Theorem 3.8.

THEOREM 3.19. [36] *Every quantifier in* $\exists^{mod}$ *preserves regularity effectively.*

PROOF. Let μ be an injective automatic presentation of structure $\mathcal{D}$, and let $\Phi(y_1, \ldots, y_n)$ denote

$$\exists^{(k,m)} x \Psi(x, y_1, \ldots, y_n),$$

for fixed $k, m \in \mathbb{N}$ and first-order $\mathcal{D}$-formula Ψ. Let $\mathcal{A}$, with $A \subseteq \Sigma^\star$ be the automatic copy of $\mathcal{D}$ induced by the presentation. The aim is to show that the relation defined in $\mathcal{A}$ by Φ is regular, and to construct the automaton effectively from the automata of the presentation.

First, the required automaton should not accept those $\overline{y}$ for which there are infinitely many x with $\mathcal{A} \models \Psi(x, \overline{y})$. So, define $\Psi_f(x, \overline{y})$ to be

$$\Psi(x, \overline{y}) \wedge \neg\exists^\infty x \Psi(x, \overline{y}).$$

Now, it is sufficient to construct an automaton, which will be called $\boldsymbol{B}'$, that recognises the relation defined in $\mathcal{A}$ by $\exists^{(k,m)} x \Psi_f(x, \overline{y})$. The following property of Ψ_f will be used: For every tuple $\overline{y}$, there are finitely many x with $\mathcal{A} \models \Psi_f(x, \overline{y})$.

Let $\boldsymbol{B} = (Q, \iota, \delta, F)$ be a deterministic automaton recognising the convolution of $(\Psi_f)^{\mathcal{A}}$. The idea is that on input $\overline{y}$, the required automaton, $\boldsymbol{B}'$, will count, modulo m, the number of accepting paths of $\boldsymbol{B}$ on inputs of the form $(x, \overline{y})$ for $x \in \Sigma^\star$. Here is some notation to help describe the construction.

Notation: For states $q, q' \in Q$, and input $\overline{y} \in \Sigma^{\star n}$ define $\#(q, \otimes\overline{y}, q')$ to be the number of strings $x \in \Sigma^\star$ with $|x| = |\otimes\overline{y}|$, such that there is a path in $\boldsymbol{B}$ labeled $\otimes(x, \overline{y})$ from state q to state q'. For $S \subseteq Q$ define $\#(S, \otimes\overline{y}, q')$ to be the sum $\sum_{q \in S} \#(q, \otimes\overline{y}, q')$.

More to the point then, the states of $\boldsymbol{B}'$ are of the form $(S_1, \ldots, S_m)$ where $S_i \subseteq Q$. If $\boldsymbol{B}'$ is in state $(S_1, \ldots, S_m)$ after reading input $\overline{y}$, then S_i will consists of those states q of $\boldsymbol{B}$ for which $\#(\iota, \otimes\overline{y}, q)$ is congruent to i modulo m. Finally, once $\otimes\overline{y}$ has been completely read, the automaton needs to account for more input to $\boldsymbol{B}$ of the form $\otimes(x', \lambda, \ldots, \lambda)$.

Definition: The required automaton $\boldsymbol{B}' = (Q', \iota', \Delta', F')$ over alphabet $(\Sigma_\perp)^n$ is defined as follows.

(i) The state set Q' is $\prod_{1 \le i \le m} \mathbb{P}(Q)$, where $\mathbb{P}(Q)$ denotes the powerset of Q.
(ii) The initial state ι' is $\{\iota\} \times \prod_{2 \le i \le m} \{\emptyset\}$.

(iii) For a state $T = (T_1, \dots, T_m) \in Q'$ and input symbol $\sigma \in (\Sigma_\perp)^n$ define the transition function $\Delta'(T, \sigma) = (S_1, \dots, S_m)$ as follows. For every set S_i and $q \in Q$, let $q \in S_i$ if and only if

$$\sum\{j \times \#(T_j, \sigma, q) \mid 1 \leq j \leq m\} \equiv i \mod m.$$

(iv) Let $(S_1, \dots, S_m) \in F'$ if and only if

$$\sum_{f \in F} \sum_{r < |Q|} \sum_{1 \leq j \leq m} j \times \#(S_j, (\{\perp\}^n)^r, f) \equiv k \mod m.$$

Here $(\{\perp\}^n)^r$ is the concatenation of r copies of the new symbol $(\perp, \dots, \perp) \in (\Sigma_\perp)^n$.

Note that the automaton is deterministic.

Correctness: Let $w = \otimes\overline{y}$ be the input to $\mathcal{B}'$ and let $(S_1, \dots, S_m)$ be the state $\Delta'(\iota', w)$. The following can be proved by induction on $|w|$:

(i) If $|w| \geq 1$, then for each S_i,

$$q \in S_i \text{ if and only if } \#(\iota, w, q) \equiv i \mod m.$$

(ii) $(S_1, \dots, S_m) \in F'$ if and only if

$$\sum_{f \in F} \sum_{r \in \mathbb{N}} \#(\iota, w \cdot (\{\perp\}^n)^r, f) \equiv k \mod m.$$

Finally, note that the index $r \in \mathbb{N}$ can be restricted to $0 \leq r < |Q|$, since the number of strings x with $\Psi_f(x, \overline{y})$ is finite. $\dashv$

A result in [37] says that the set of extendible nodes of an automatically presentable tree $(T; \preceq)$ is regular. We adapt the proof for the next Theorem.

Theorem 3.20. *Each k-ary Ramsey quantifier preserves regularity effectively.*

Proof. Note that the 1-Ramsey quantifier is identical to ‘there exists infinitely many’. We illustrate the proof for $k = 2$, the general case being similar.

Let μ be an injective automatic presentation of some structure $\mathcal{D}$, and $\Psi(x, y, z_1, \dots, z_n)$ a first-order $\mathcal{D}$-formula. Let $\mathcal{A}$ be an automatic copy of $\mathcal{D}$ with $A \subseteq \Sigma^\star$. The aim is to show constructively that the binary relation defined in $\mathcal{A}$ by the formula

$$\exists^{2\text{-ram}} xy \Psi(x, y, \overline{z}),$$

is regular.

Now, $\mathcal{A} \models \exists^{2\text{-ram}} xy \Psi(x, y, \overline{z})$ if and only if there exists an infinite string $\alpha \in \Sigma^\omega$ and an infinite set $I \subseteq \mathbb{N}$:

(i) for every $i \in I$, there is a string $w_i := p_i t_i$ where p_i is the initial prefix of α of length i, and t_i is a non-empty string;
(ii) and for every $i \neq j \in I$, it holds that $w_i \neq w_j$ and $\Psi(w_i, w_j, \overline{z})$.

Since this expression quantifies over infinite strings, we will use automata operating over infinite strings, so called ω-automata or Büchi-automata. Here is a brief reminder of what this means. A Büchi automaton has the same form (Q, ι, Δ, F) as a non-deterministic finite automaton. A run on an infinite string is accepting if some accepting state occurs infinitely in the run. A relation $R \subseteq (\Sigma^\omega)^m$ is *Büchi recognisable* if there is a Büchi automaton recognising the convolution of R, namely the relation $\otimes R \subseteq (\Sigma^m)^\omega$ where the ith symbol of $\otimes R$ is defined as $(\sigma_1, \dots, \sigma_m)$ where σ_j is the ith symbol of the jth component of R. The languages recognised by Büchi automata satisfy an analogous version of Theorem 2.4. For details, see for instance [55].

We need to mark the boundaries as given by I. Introduce a new symbol $\wr$ for this purpose, and let $\Delta = \Sigma \cup \{\wr\}$. Call a pair (s, t) of infinite strings $s, t \in \Delta^\omega$ *good* if s can be expressed as $s_0 \wr s_1 \wr s_2 \wr \dots$ and t as $t_0 \wr t_1 \wr t_2 \wr \dots$, where $s_i, t_i \in \Sigma^\star$ and $|s_i| = |t_i| > 0$ for each i. In this case, say that every string of the form $s_0 s_1 \dots s_l t_{l+1}$, for $l \in \mathbb{N}$, is *on* the good pair (s, t).

Now define the language $R \subseteq (\Delta^\omega)^{2+n}$ as consisting of tuples of the form $(s, t, u_1, \dots, u_n)$ with the following properties:

(i) (s, t) is a good pair;
(ii) every u_i is of the form $z_i\{\wr\}^\omega$ for some $z_i \in \Sigma^\star$;
(iii) and if x and y are distinct strings on (s, t), then $\Psi(x, y, z_1, \dots, z_n)$ holds.

Then $\mathcal{A} \models \exists^{\text{2-ram}} xy\Psi(x, y, \overline{z})$ if and only if there exists (s, t) so that $R(s, t, z_1\{\wr\}^\omega, \dots, z_n\{\wr\}^\omega)$.

To finish the proof, one can show that R is Büchi recognisable. Moreover, since Büchi automata are closed under projection, there is a Büchi automaton that accepts $(z_1\{\wr\}^\omega, \dots, z_n\{\wr\}^\omega)$ if and only if $\mathcal{A}$ models $\exists^{\text{2-ram}} xy\Psi(x, y, \overline{z})$. This automaton can be transformed into a finite automaton recognising the relation defined in $\mathcal{A}$ by $\exists^{\text{2-ram}} xy\Psi(x, y, \overline{z})$. ⊣

As an aside, we now show that one can extract from this proof an automatic version of Ramsey's Theorem.

Recall that the infinitary version of Ramsey's Theorem (for k-tuples and two colours) says that if D is infinite and $E \subseteq D^k$, then there exists an infinite monochromatic set $X \subseteq D$; namely, either $E(\overline{x})$ for all pairwise distinct $x_1, \dots, x_k \in X$, or $\neg E(\overline{x})$ for all pairwise distinct $x_1, \dots, x_k \in X$.

Proposition 3.21. *Let μ be an automatic presentation of a structure $\mathcal{D} = (D; E)$, with D infinite and $E \subseteq D^k$. Then there exists an infinite monochromatic set $X \subseteq D$ with $\mu^{-1}(X)$ regular.*

Proof. For simplicity we consider the case $k = 2$. By Ramsey's Theorem there is a monochromatic set $X \subseteq D$. Suppose that $E(x, y)$ for all $x \neq y \in X$ (the other case is symmetric).

Let $\mathcal{A} = (A; E)$ be the automatic copy of $\mathcal{D}$ induced by μ. By the previous proof, the set of good pairs (s, t) such that $x \neq y \in A$ on (s, t) implies

$E(x, y)$, is a Büchi recognisable relation, say R. If $R(s, t)$ then the set of strings $w \in A$ that are on (s, t) forms an infinite monochromatic set.

By assumption $R(\cdot, \cdot)$ is non-empty. Since every Büchi automaton accepting some string also accepts some ultimately periodic string, we get that R contains some ultimately periodic $\otimes(s', t') \in (\Delta^2)^\omega$. So both s' and t' are ultimately periodic. This pair (s', t') can be transformed into a finite automaton accepting all the strings $w \in A$ that are on (s', t'). ⊣

Definition 3.22. Write $\boldsymbol{Q}_n^{\text{reg}}$ for the collection of n-ary generalised quantifiers that preserve regularity.

For instance, $\boldsymbol{Q}_1^{\text{reg}}$ contains the usual quantifiers $\exists, \forall$, as well as the modulo quantifiers $\exists^{\text{mod}}$ and $\exists^\infty$; and $\boldsymbol{Q}_n^{\text{reg}}$ contains $\exists^{\text{n-ram}}$.

Problem 3.23. Classify the quantifiers in $\boldsymbol{Q}_n^{\text{reg}}$, for each n.

The following general definition will allow us to compare the expressive strength of quantifiers.

Definition 3.24. Let Q be a quantifier, $\boldsymbol{Q}$ a collection of quantifiers, and τ the signature of Q. Say that Q *is definable in* $\boldsymbol{Q}$ if there is a sentence θ over the signature τ in the logic FO[$\boldsymbol{Q}$] with $Q = \{\mathcal{A} \mid \mathcal{A} \models \theta\}$.

For instance, a structure $(A; X)$ satisfies the sentence $\neg[\exists^{(0,2)} z X(z) \vee \exists^{(1,2)} z X(z)]$ if and only if X is infinite. So, $\exists^\infty$ is definable in $\{\exists^{(0,2)}, \exists^{(1,2)}\}$.

As a first step, we characterise the unary quantifiers that preserve regularity.

Proposition 3.25. *Every quantifier in* $\boldsymbol{Q}_1^{reg}$ *is definable in* $\exists^{mod}$.

Proof. Every unary quantifier is equal to some cardinality quantifier where $C \subseteq (\mathbb{N} \cup \{\infty\})^r$ (Example 3.13 (4)). So suppose that Q_C preserves regularity. Since $\exists^\infty$ is definable in $\exists^{\text{mod}}$, we may assume that $C \subseteq \mathbb{N}^r$. Let $\Sigma = \{1, \ldots, r\}$. Write $\#_i(z)$ for the number of occurrences of the symbol i in the string $z \in \Sigma^\star$. Then the set

$$C' = \{z \in \Sigma^\star \mid (\#_1(z), \ldots, \#_r(z)) \in C\}$$

is regular since it is definable as $Q_C \overline{x}(M_1(x_1, z), \ldots, M_r(x_r, z))$, where for each $i \leq r$, M_i is the regular binary relation of those pairs (x, z) such that $x \in 11^\star$ and z has the symbol i in position $|x|$.

Now if C' is regular, then C is recognisable - meaning that C is the union of classes of a congruence on $(\mathbb{N}^r, +)$ of finite index.

By Mezei's description of the recognisable subsets of a product of monoids (see for instance [8]), C is recognisable if and only if it is a finite union of sets of the form $X_1 \times \cdots \times X_r$, where each $X_i \subseteq \mathbb{N}$ is ultimately periodic. In other words, Q_C is definable in $\exists^{\text{mod}}$. ⊣

Example 3.26. The Härtig quantifier does not preserve regularity. Indeed, the previous proof says that this is because one could use it to define, in a suitable automatically presentable structure, the (non-regular) set of strings $x \in \{0,1\}^\star$ for which x has the same number of 0s as 1s.

What about $\boldsymbol{Q}_n^{\mathrm{reg}}$ for $n \geq 2$? I do not know of a characterisation as for $n = 1$. Here is a simple example separating $\boldsymbol{Q}_2^{\mathrm{reg}}$ from $\boldsymbol{Q}_1^{\mathrm{reg}}$.

Example 3.27. Define the quantifier $\exists^\rho$ stating that '$(A; \psi(\cdot,\cdot,\overline{z}))$ is an equivalence structure with infinitely many infinite classes'.

Then $\exists^\rho$ is in $\boldsymbol{Q}_2^{\mathrm{reg}}$, since it is $\mathrm{FO}_{\omega-\mathrm{inv}}(\exists^\infty)$ definable.

However, it is not expressible in first-order plus all unary quantifiers over the class of automatically presentable structures. This can be seen by checking that for every $r \in \mathbb{N}$, the Duplicator has a winning strategy in the 1-bijective r-round game[4] between the following structures: $\mathcal{A}_r$ comprises exactly r infinite classes, and $\mathcal{B}_r$ comprises infinitely many infinite classes.

3.3. Automatic presentations. An automatically presentable structure $\mathcal{B}$ has many automatic presentations. Here we look at the relationship amongst these presentations, and in the next subsection at their relationship to definability in the structure.

For a given automatic presentation $\mu \colon A \to B$ of $\mathcal{B}$, write $\mu(\mathrm{Reg})$ for the collection of relations

$$\{\mu(R) \mid R \subseteq A^r \text{ is a regular relation}, r \in \mathbb{N}\}.$$

Definition 3.28. [2] Let $\mathcal{B}$ be an automatically presentable structure. Call two automatic presentations μ, ν of $\mathcal{B}$ *equivalent* if $\mu(\mathrm{Reg}) = \nu(\mathrm{Reg})$.

Example 3.29. Fix an automatic presentation μ of an infinite structure $\mathcal{A}$. There are infinitely many equivalent presentations ν so that the binary relation $\{(u,v) \mid \mu(u) = \nu(v)\}$ that translates between the presentations is not regular. For instance, for $c \notin \Sigma$ and $n \in \mathbb{N}$, let ν_n be the presentation sending $a_0 c^n a_1 c^n \ldots a_k c^n$ to $\mu(a_0 a_1 \ldots a_k)$ for $a_0 \ldots a_k \in \mathrm{dom}(\mu)$. Here c^n denotes the concatenation of n many c symbols.

Proposition 2.6 says that we can transform an arbitrary automatic presentation into one that is injective and over a binary alphabet. Its proof yields an equivalent presentation.

Proposition 3.30. *Let μ be an automatic presentation of $\mathcal{B}$. Then there is an* equivalent *presentation ν of $\mathcal{B}$ with the following properties*:

(i) *ν is injective, and*
(ii) *ν is a presentation over a binary alphabet* $|\Sigma| = 2$.

[4]The n-bijective games, introduced by Hella [29], are Ehrenfeucht–Fraïssé like games characterising definability in first-order logic extended by all n-ary generalised quantifiers. Here is a brief description: in round i the Duplicator chooses a bijection $b_i \colon A \to B$, and the spoiler answers with a set $C_i \subseteq A$ where $|C_i| \leq n$. Duplicator wins the r-round game if $\cup_{j \leq r} b_j \restriction C_j$ is a partial isomorphism from $\mathcal{A}$ to $\mathcal{B}$.

https://doi.org/10.2178/bsl/1208442827

Bárány has characterised equivalence of presentations in automata theoretic terms. For the sake of completeness, here is the definition of semi-synchronous rational relation, which can be thought of as a relation recognised by a multi-tape automaton where each read-head advances at a different, but still constant, speed.

DEFINITION 3.31. Fix a finite alphabet Σ and a vector of positive integers $\underline{m} = (m_1, \ldots, m_r)$. Let $\perp$ be a symbol not in Σ. For each component m_i introduce the alphabet $(\Sigma_\perp)^{m_i}$. The *$\underline{m}$-convolution of a tuple* $(w_1, \ldots, w_r) \in (\Sigma^\star)^r$ is formed as follows: First, consider the intermediate string $(w_1 \perp^{a_1}, \ldots, w_r \perp^{a_r})$ where the a_i are minimal such that there is some $k \in \mathbb{N}$ so that for all i, $|w_i| + a_i = km_i$. Second, partition each component $w_i \perp^{a_i}$ into k-many blocks of size m_i, and view each block as an element of $(\Sigma_\perp)^{m_i}$. Thus the string $\otimes_{\underline{m}}(w_1, \ldots, w_r)$ is formed over alphabet $(\Sigma_\perp)^{m_1} \times \cdots \times (\Sigma_\perp)^{m_r}$.

The *$\underline{m}$-convolution of a relation* $R \subseteq (\Sigma^\star)^r$ is the set $\otimes_{\underline{m}} R$ defined as $\{\otimes_{\underline{m}} \overline{w} \mid \overline{w} \in R\}$.

Call R *$\underline{m}$-synchronous rational* if there is a finite automaton recognising $\otimes_{\underline{m}} R$.

Call R *semi-synchronous* if it is $\underline{m}$-synchronous rational for some $\underline{m}$.

In particular, a $(1, \ldots, 1)$-synchronous rational relation R is synchronous rational (Definition 2.2).

EXAMPLE 3.32. In the proof of Proposition 2.6, the coding $\alpha \subseteq \Gamma^\star \times \Sigma^\star$ is semi-synchronous. Similarly, the translation in Example 3.29 is semi-synchronous.

THEOREM 3.33. [2] *Two automatic presentations μ and ν of a structure $\mathcal{B}$ are equivalent if and only if the relation $\{(u, v) \mid \mu(u) = \nu(v)\}$ is semi-synchronous.*

Here are a few words about the proof. The 'if' part follows from the basic property that the composition of a semi-synchronous relation with a regular relation is regular. On the other hand, for the 'only if' part it is sufficient to consider injective presentations, and prove that the translation $\mu^{-1}\nu\colon \mathrm{dom}(\nu) \to \mathrm{dom}(\mu)$, which by assumption preserves the regularity and non-regularity of relations on $\mathrm{dom}(\nu)$, is semi-synchronous. This is achieved by a series of transformations, and is nicely presented in [3].

EXAMPLES 3.34. (i) With a little more work, one can show that each of the structures $\mathcal{W}_k$ and $\mathcal{N}_k$ (for $k \geq 2$) has exactly one presentation up to equivalence [2, 3].

(ii) Presburger arithmetic $(\mathbb{N}; +)$ has infinitely many non-equivalent presentations. This follows from a result of Büchi [16] saying that the set of powers of $n \geq 2$, written in base $m \geq 2$ coding, is regular if and only if $n^p = m^q$ for some positive integers p and q.

3.4. Intrinsic regularity. This subsection looks at the relationship between regularity and definability in automatically presentable structures. The following definition parallels that of the intrinsically computably enumerable relations; see Ash and Nerode [1]. Loosely, a relation $R \subseteq B^n$ (not assumed to be an atomic relation of $\mathcal{B}$) is called *intrinsically regular in* $\mathcal{B}$ if it is regular in every automatic copy of $\mathcal{B}$.

Definition 3.35. Fix an automatically presentable structure $\mathcal{B}$. Call a relation $R \subseteq B^n$ *intrinsically regular* in $\mathcal{B}$ if for every automatic presentation μ of $\mathcal{B}$, the relation $\mu^{-1}(R)$ is regular (or in the terminology of the previous section, that $R \in \mu(\mathrm{Reg})$).

Denote by $\mathrm{IR}(\mathcal{B})$ the set of intrinsically regular relations in $\mathcal{B}$.

Problem 3.36. Can we capture intrinsic regularity using definability? For instance, by Remarks 3.11 and 3.18, for every automatically presentable structure $\mathcal{B}$ we see that $\mathrm{FO}[\cup_n \boldsymbol{Q}_n^{\mathrm{reg}}]_{\omega-\mathrm{inv}}(\mathcal{B}) \subseteq \mathrm{IR}(\mathcal{B})$. Is there equality here?

IR *in some specific structures.* We can describe $\mathrm{IR}(\mathcal{B})$ for some specific automatically presentable $\mathcal{B}$s.

Proposition 3.37.

(i) $\mathrm{IR}(\mathbb{N}, +, |_m) = \mathrm{FO}(\mathbb{N}, +, |_m)$, *for* $m > 1$.
(ii) $\mathrm{IR}(\mathbb{N}, +) = \mathrm{FO}(\mathbb{N}, +)$.
(iii) $\mathrm{IR}(\mathbb{N}, \leq) = \mathrm{FO}[\exists^{mod}](\mathbb{N}, \leq)$.

The first item follows from the fact that (the base m coding of) a relation is regular only if it is first-order definable in $(\mathbb{N}, +, |_m)$ (see Theorem 3.8).

The second item follows from the Cobham-Semenov Theorem: R is first-order definable in $(\mathbb{N}, +)$ if for some $m, n \in \mathbb{N}$, R in base m is regular (as a subset of $\{0, 1, \dots, m-1\}^\star$) and R in base n is regular, and $m^p \neq n^q$ for all positive integers p, q (see [15] for a discussion).

The third item follows from the characterisation of the structures with unary-automatic presentations (Theorem 3.9).

The proof of the next theorem, due to Frank Stephan, is technical and can be found in [51].

Theorem 3.38. *For every* $k \geq 2$, *there is an automatic copy of* $(\mathbb{N}, S)$ *in which the image of the set* $\{n \in \mathbb{N} \mid k \text{ divides } n\}$ *is not regular.*

Here S is the usual successor function $n \mapsto n+1$. A little work establishes the following corollary.

Corollary 3.39. *A unary relation* $R \subseteq \mathbb{N}$ *is intrinsically regular for the structure* $(\mathbb{N}, S)$ *if and only if* R *is in* $\mathrm{FO}(\mathbb{N}, S)$.

The proof of Theorem 3.38 may be adapted to yield automatic presentations with pathological properties.

A *cut* of the structure $(\mathbb{Z}, S)$ is a set of the form $\{x \in \mathbb{Z} \mid x \geq n\}$ where $n \in \mathbb{Z}$ is fixed.

https://doi.org/10.2178/bsl/1208442827

COROLLARY 3.40. *There is an automatic copy of* $(\mathbb{Z}, S)$ *in which no cut is regular.*

COROLLARY 3.41. *There is an automatic copy of a graph with exactly two connected components, each isomorphic to* $(\mathbb{N}, S)$, *and neither regular.*

3.5. Restriction on growth. How can one prove that a given structure is not automatically presentable? This section provides some properties of automatically presentable structures, that in the contrapositive can be used to establish that a given structure has no automatic presentation. I first mention two methods based on results that we have already seen.

First, Theorem 3.2 says that the first-order theory of an automatically structure is decidable. This settles, for instance, that arithmetic $(\mathbb{N}, +, \times)$ is not automatically presentable. A closer approach would be to note that the proof of the theorem gives an upper bound of $\exp_{k-1}(c^n)$ for the space complexity of deciding the Σ_k (or Π_k) fragment of the first-order theory of an automatically presentable structure (here c is the size of the automata in the presentation, and n the size of the input sentence).

Second, if it has already been established that $\mathcal{A}$ is not automatically presentable, then no structure that interprets $\mathcal{A}$ is automatically presentable (Proposition 3.6).

We now turn to some finer techniques. These share the idea that in an automatically presentable structures, certain functions that count (elements, definable sets, etc.) cannot grow too fast.

A *locally-finite* relation R parameterised by $k, l \in N$, is one such that $R \subseteq A^{k+l}$ and for every tuple $\overline{x}$ of size k there are at most a finite number of tuples $\overline{y}$ of size l such that $(\overline{x}, \overline{y}) \in R$. The canonical example of a locally-finite relation is the graph of a function.

The following proposition says, in a simple case, that if (the graph of) a function f is regular, then for every x, the value $|f(x)| - |x|$ is bounded above by the number of states of an automaton for f. Although the proposition has an easy proof, it is an important tool.

PROPOSITION 3.42. *Suppose that locally-finite* $R \subseteq A^{k+l}$ *is regular. There exists a constant* p, *that depends only on the automaton for* R, *such*

$$\max\{|y| \mid y \in \overline{y}\} - \max\{|x| \mid x \in \overline{x}\} \leq p$$

for every $(\overline{x}, \overline{y}) \in R$.

PROOF. Take the simple case that R is the graph of a (partial) function $f : A \to A$. Suppose an automaton recognising $\otimes R$ has p states. Assume, aiming for a contradiction, that $|f(x)| > |x| + p$ for some x in the domain of f. Then the run of $\otimes(x, f(x))$, after reading the first $|x|$ symbols, must repeat a state. This implies that the automaton also accepts infinitely many strings of the form $\otimes(x, \cdot)$. ⊣

We will now see a useful way of iterating Proposition 3.42.

DEFINITION 3.43. Start with an automatic copy $\mathcal{A}$ of a structure that includes functions $f_1, \ldots, f_k$ of arities $r_1, \ldots, r_k$ respectively. Let $D \subseteq A$ be regular, listed as d_i so that $|d_i| \leq |d_{i+1}|$.

Define the *nth growth level*, written $G_n(D)$, inductively by $G_1(D) = \{d_1\}$ and

$$G_{n+1}(D) = \{d_{n+1}\} \cup G_n(D) \cup \\ \cup_{i \leq k} \{f_i(x_1, \ldots, x_{r_i}) \mid x_j \in G_n(D) \text{ for } 1 \leq j \leq r_i\}.$$

We are interested in how fast $|G_n(D)|$ grows as a function of n. For example, consider the free semigroup $(\Sigma^\star, \cdot)$ with generating set $\Sigma = \{d_1, \ldots, d_m\}$. For $m \geq 2$, since $G_m(\Sigma) \supset \Sigma$, the set $G_{m+n}(\Sigma)$ includes all strings over Σ of length at most 2^n; thus the cardinality of $G_{m+n}(\Sigma)$ is at least a double exponential, namely $\exp_2(n)$.

PROPOSITION 3.44 ([12], cf. [34]). *Let $\mathcal{A}$ and $G_n(D)$ be as in Definition* 3.43. *Then there is a linear function $t\colon \mathbb{N} \to \mathbb{N}$ so that every string of $G_n(D)$ has length at most $t(n)$.*

In other words, if $|\Sigma| = 1$ then $|G_n(D)| = O(n)$, else $|G_n(D)| = |\Sigma|^{O(n)}$.

COROLLARY 3.45. *The following structures do not have automatic presentations.*

(i) *The free semigroup $(\Sigma^\star, \cdot)$ on $|\Sigma| > 1$ generators* [34].
(ii) *Any term algebra generated by finitely many constants and at least one non-unary atomic function* [34].
(iii) $(\mathbb{N}, \times)$, *multiplication of natural numbers* [12].
(iv) $(\mathbb{N}, \langle\cdot,\cdot\rangle)$, *where $\langle\cdot,\cdot\rangle\colon \mathbb{N}^2 \to \mathbb{N}$ is a bijection* [12].

A straightforward application of Proposition 3.42 is the following result. It can be proved by induction on m.

PROPOSITION 3.46. *Let $(M;\cdot)$ be an automatic copy of a semigroup. Then there is a constant e, that depends only on an automaton for $\cdot$, so that for every $x_1, \ldots, x_m \in M$,*

$$|\prod_{i=1}^{m} x_i| \leq \max\{|x_i|\colon 1 \leq i \leq m\} + e \log m.$$

This proposition will be used in Section 4 in the classification of the automatically presentable Boolean algebras and for proving certain groups do not have automatic presentations.

A different technique, independently reported by Delhommé [23] and Stephan [52], is used to prove that certain universal homogeneous structures have no automatic presentation. The basic ideas are presented below, following [35], although it is worth mentioning that Delhommé [24] presents a more general result, as well stating a corresponding condition for structures with tree-automatic presentations.

Recall $A^{\leq n}$ denotes the strings of A of length at most n.

Definition 3.47. Suppose that the structure $\mathcal{A}$ contains an atomic binary relation E and that $A \subseteq \Sigma^\star$. For $n \in \mathbb{N}$ and $y \in A$ define the set $\mathrm{tp}_n(y) \subseteq A^{\leq n}$ as those $x \in A^{\leq n}$ for which $\mathcal{A} \models E(x, y)$.

Write $\#\mathrm{tp}_n$ for the cardinality of the set $\{\mathrm{tp}_n(y) \mid y \in A\}$. Note that the set tp_n and the number $\#\mathrm{tp}_n$ are defined with respect to A and E. I may write $\mathrm{tp}_{n,A,E}$ to stress this.

In general $\#\mathrm{tp}_n \leq |\Sigma|^{|A^{\leq n}|}$, however if $\mathcal{A}$ is an automatic copy of some structure then we can say more.

Theorem 3.48. *If $\mathcal{A}$ is an automatic copy of a structure containing an atomic binary relation E, then $\#\mathrm{tp}_{n,A,E} \leq k|A^{\leq n}|$ for some constant $k \in \mathbb{N}$ that depends on the automata for A and E.*

Proof. One proves that there is a constant c, depending on the number of states of the automata for A and E, so that for every $n \in \mathbb{N}$ and $y \in A$, there is a $y' \in A^{\leq n+c}$ with $\mathrm{tp}_n(y) = \mathrm{tp}_n(y')$. Now apply the fact that $|A^{\leq n+c}| \leq k(|A^{\leq n}|)$ where the constant $k \in \mathbb{N}$ depends on the number of states of the automaton for A. ⊣

Corollary 3.49. *The following structures do not have automatic presentations.*

(i) *The random graph.*
(ii) *The universal, homogeneous partial order.*
(iii) *The K_p-free random graph for every $p > 2$ (K_p is the complete graph on p vertices).*

Proof. The first case is illustrated. Suppose $(A; E)$ is an automatic copy of the random graph over a binary alphabet. The random graph has the following property. For every finite subset F of A, and every partition X_1, X_2 of F, there is a vertex $x \in A$ such that $(x, x_1) \in E$ for every $x_1 \in X_1$ and $(x, x_2) \notin E$ for every $x_2 \in X_2$. Hence taking F to be $A^{\leq n}$ it holds that $\#\mathrm{tp}_{n,A,E} = \exp(|A^{\leq n}|)$ contradicting the theorem. ⊣

We finish with another necessary condition of having an automatic presentation, due to Delhommé [24].

Definition 3.50. Say that a structure $\mathcal{B}$ is a *sum-augmentation* of a set of structures $\mathcal{S}$ (each having the same signature as $\mathcal{B}$) if there is a finite partition of $B = B_1 \cup \cdots \cup B_n$ such that for each i the substructure $\mathcal{B} \restriction B_i$ is isomorphic to some structure in $\mathcal{S}$.

Theorem 3.51. *Suppose $\mathcal{A}$ has finite signature. If $\mathcal{A}$ is an automatically presentable then for every $\mathcal{A}$-formula $\phi(x, \overline{y})$, there is a finite set of structures $\mathcal{S}$ so that for every tuple of elements $\overline{b}$ from A, the substructure $\mathcal{A} \restriction \phi^{\mathcal{A}}(\cdot, \overline{b})$ is a sum-augmentation of $\mathcal{S}$.*

Proof. Say $\mathcal{A} = (A; R_1^{\mathcal{A}}, \dots, R_r^{\mathcal{A}})$, $A \subseteq \Sigma^\star$, is an automatic copy of a structure. For any given $\mathcal{A}$-formula ψ, fix a deterministic automaton

$(Q_\psi, \iota_\psi, \Delta_\psi, F_\psi)$ recognising $\psi^{\mathcal{A}}$, and write $\Gamma_\psi(w)$ for $\Delta_\psi(\iota_\psi, w)$. We will use the following property (P_ψ): For all strings c_i, d_i with the c_is all the same length,

$$\Delta_\psi(\Gamma_\psi(\otimes(c_1, \ldots, c_k)), \otimes(d_1, \ldots, d_k)) \in F_\psi.$$

if and only if $\psi(c_1 d_1, \ldots, c_k d_k)$ holds in $\mathcal{A}$.

Now, given an $\mathcal{A}$-formula $\phi(x, y_1, \ldots, y_l)$ as in the hypothesis, and tuple $\overline{b}$, observe that for $m = \max\{|b_i|\}$, we can partition $\phi^{\mathcal{A}}(\cdot, \overline{b})$ into the finitely many singletons $\{c\}$ for $\phi(c, \overline{b})$ with $|c| < m$, and the finitely many sets $\phi^{a\Sigma^\star}(\cdot, \overline{b}) := \{aw \in A \mid \mathcal{A} \models \phi(aw, \overline{b}), w \in \Sigma^\star\}$ for $|a| = m$. Since the signature is assumed to be finite, there are finitely many isomorphism types amongst substructures of the form $\mathcal{A} \restriction \{a\}$, for $a \in A$. So, it is sufficient to show, that as we vary $(a, \overline{b})$ subject to $|a| = \max\{|b_i|\}$, there are finitely many isomorphism types amongst substructures of the form $\mathcal{A} \restriction \phi^{a\Sigma^\star}(\cdot, \overline{b})$.

The idea is to bound this number of isomorphism types in terms of the number of states of the automata involved. To this end, define a function f_ϕ as follows. Its domain consists of tuples $(a, \overline{b})$ where $|a| = \max\{|b_i|\}$; and f_ϕ sends this tuple to the tuple of states

$$(\Gamma_\phi(\otimes(a, b_1, \ldots, b_l)), \Gamma_A(a), (\Gamma_{R_i^{\mathcal{A}}}(\otimes(a, \ldots, a)))_{i \le r}).$$

The range of f_ϕ is bounded by $|Q_\phi| \times |Q_A| \times \prod_{i \le r} |Q_{R_i^{\mathcal{A}}}|$. In particular, the range is finite.

To finish the proof, one needs to argue that the isomorphism type of the substructure $\mathcal{A} \restriction \phi^{a\Sigma^\star}(\cdot, \overline{b})$ depends only on the value $f_\phi(a, \overline{b})$. This follows from the fact that if $f_\phi(a, \overline{b}) = f_\phi(a', \overline{b'})$, then the corresponding substructures are isomorphic via the mapping $I \colon aw \mapsto a'w$ $(w \in \Sigma^\star)$. Indeed, by property $(P_{x \in A})$ we get that $aw \in A$ if and only if $a'w \in A$, for every w. This means that I is a bijection between the sets $A \cap a\Sigma^\star$ and $A \cap a'\Sigma^\star$. Similarly, by the properties (P_ψ) where ψ is taken to be $\overline{x} \in R_i^{\mathcal{A}}$, we get that I is an isomorphism between the substructures $\mathcal{A} \restriction a\Sigma^\star$ and $\mathcal{A} \restriction a'\Sigma^\star$. Finally, by (P_ϕ) we get that I is an isomorphism between the substructures $\mathcal{A} \restriction \phi^{a\Sigma^\star}(\cdot, \overline{b})$ and $\mathcal{A} \restriction \phi^{a'\Sigma^\star}(\cdot, \overline{b})$. ⊣

Here is an illustration of the theorem.

Corollary 3.52. [22] *The ordinal* $(\omega^\omega, \le)$ *is not automatically presentable.*

Proof. Suppose for a contradiction that $(\omega^\omega, \le)$ has an automatic presentation. In Theorem 3.51, take $\phi(x, y)$ to be $x < y$ and consider the following fact (proved by induction): If the domain of a well-order, isomorphic to some ordinal of the form ω^n for $n \in \mathbb{N}$, is partitioned into finitely many pieces $\{B_i\}_i$, then there is some i so that the substructure on domain B_i is isomorphic to ω^n.

This means that the set of structures $\mathcal{S}$ must contain $(\omega^n, <)$ for every $n \in \mathbb{N}$, contradicting the finiteness of $\mathcal{S}$. ⊣

3.6. Isomorphism problem. Let C be a class of structures (over a finite signature) closed under isomorphism. Write $\mathsf{C}^{\mathrm{aut}}$ for the (codes of the) tuples of automata $\overline{\mathsf{M}} = (\mathsf{M}_A, \mathsf{M}_=, (\mathsf{M}_i)_i)$ in automatic presentations of members of C.

The *isomorphism problem for the automatically presentable members of* C is the set

$$\{\langle \overline{\mathsf{M}}, \overline{\mathsf{M}'}\rangle \mid \overline{\mathsf{M}}, \overline{\mathsf{M}'} \in \mathsf{C}^{\mathrm{aut}} \text{ present isomorphic structures}\}.$$

Note that the $\mathsf{C}^{\mathrm{aut}}$ may be computable (for instance, if C are the Boolean algebras, linear orderings, or finite graphs). However this does not mean that we can describe the isomorphism types of the automatically presentable members of $\mathsf{C}^{\mathrm{aut}}$. Indeed, the isomorphism problem for the automatically presentable directed graphs is undecidable [13]:

Proof. We encode the halting problem into this isomorphism problems as follows. For a given a Turing machine N, construct an equivalent reversible Turing machine N_r. In particular, this means that its configuration space is a disjoint union of *chains* (these are graphs isomorphic to initial segments of $(\mathbb{N}, S)$ where $S\colon n \mapsto n+1$). Such a machine can be constructed by introducing to N a tape that records the sequence of transitions that it takes, see Bennett [7].

We may assume that N (and hence also N_r) loops indefinitely instead of halting in a reject state. So, the configuration space of N_r consists of a finite chain for every accepting computation of N_r, an infinite chain for every rejecting computation of N_r, and possibly some other chains that do not correspond to valid computations of N_r (the *junk*).

The configuration space of a Turing machine N_r is automatically presentable (Example 2.4 (4)). It can be massaged so that its only finite chains are those that correspond to *valid* accepting computations of N_r. Consider the set of configurations I of N_r which have no predecessor in N_r, but are not valid initial configurations (in other words, consider the roots of chains from the junk). Write $(\mathbb{N}, P)$ for the graph with domain $\mathbb{N}$ and edge relation $(n+1, n)$ for every $n \in \mathbb{N}$. For each element $i \in I$, attach the chain with root i to a copy of $(\mathbb{N}, P)$ at 0. Here *attach* $\mathcal{H}$ *to* $\mathcal{G}$ means take the disjoint union of $\mathcal{H}$ and $\mathcal{G}$ and add an edge from the distinguished node of g to that of h. Call the resulting graph $\mathcal{N}'$. It has the property that every invalid computation is isomorphic to either $(\mathbb{N}, P)$ (if it were a terminating computation) or to $(\mathbb{Z}, S)$ (if it were non-terminating).

Denote by $\mathcal{J}$ an automatic copy of the graph consisting of infinitely many disjoint copies of each of the following structures: $(\mathbb{N}, S)$, $(\mathbb{N}, P)$ and $(\mathbb{Z}, S)$. Denote by $\mathcal{N}''$ the disjoint union of the graph $\mathcal{N}'$ and $\mathcal{J}$. Then N rejects every input string if and only if $\mathcal{N}''$ is isomorphic to $\mathcal{J}$.

It is straightforward to check that $\mathcal{N}''$ is automatically presentable, and that automata presenting it can be computed from N. ⊣

In the previous proof the configuration space of N_r is *locally-finite*, namely the degree of every vertex in the underlying undirected graph is finite. The isomorphism problem for the locally-finite automatically presentable directed graphs is Π^0_3-complete [51]. So, at first sight, the following result seems surprising.

THEOREM 3.53. [35] *The isomorphism problem for the class of automatically presentable directed graphs is* Σ^1_1*-complete.*

PROOF. For hardness, we encode the isomorphism problem for computable subtrees of $\omega^{<\omega}$. A proof of its Σ^1_1-completeness can be found in Goncharov and Knight [28, Theorem 4.4(b)] who attribute it to the folklore. Let us briefly describe this problem.

Here $\omega^{<\omega}$ is the set of all finite sequences from the set of natural numbers ω. Implicitly, there is an edge from x to $x \cdot n$ where $x \in \omega^{<\omega}$ and $n \in \omega$; this is called the *immediate successor* relation. A *subtree of* $\omega^{<\omega}$ is a subset T that is downward closed with respect to the immediate successor relation: for $x \in \omega^{<\omega}$ and $n \in \omega$, if $x \cdot n \in T$ then $x \in T$. A subtree T is *computable* if there is an algorithm that on input $x \in \omega^{<\omega}$ decides whether or not x is in T. Two subtrees are isomorphic if there is a bijection respecting the immediate successor relation. The isomorphism problem for computable subtrees of $\omega^{<\omega}$ is the set of pairs $\langle n, m\rangle$ for which n and m are indices of computable subtrees, say T_n and T_m respectively, and T_n is isomorphic to T_m.

Our aim is to exhibit, for each computable subtree T, an automatically presentable directed graph $\mathcal{G}_T$, so that computable subtrees T_1 and T_2 are isomorphic if and only if the directed graphs $\mathcal{G}_{T_1}$ and $\mathcal{G}_{T_2}$ are isomorphic. Moreover, given an index for T, we effectively construct automata presenting $\mathcal{G}_T$.

To begin, $\omega^{<\omega}$ has an automatic copy $\mathcal{W}$: the domain W is $\{\lambda\} \cup \{0,1\}^\star 1$ and immediate successor is given by $x \prec_p y \wedge \neg\exists z x \prec_p z \prec_p y$. So replacing $\omega^{<\omega}$ with $\mathcal{W}$, we talk instead about (*computable*) *subtrees of* $\mathcal{W}$. Now, from a computable subtree T of $\mathcal{W}$, construct a reversible Turing machine N_T with the property that $w \in T$ if and only if the computation of N_T starting on input w converges (compare with the previous proof). Writing $\mathcal{G}'_T$ for the configuration space of N_T, note that $\mathcal{G}'_T$ consists of disjoint unions of chains. Abuse notation and identify the initial configuration of $\mathcal{G}'_T$ on input $w \in W$ with the string w.

We need an auxilliary graph $\mathcal{O}$ formed from a copy of $\mathcal{W}$ by attaching to every $w \in W$, infinitely many finite chains of every size and exactly one infinite chain.

Form $\mathcal{G}''_T$ from $\mathcal{G}'_T$ by attaching infinitely many copies of $\mathcal{O}$ to every $w \in W \subseteq G'_T$.

Finally, form the graph $\mathcal{G}_T$ by taking the disjoint union of $\mathcal{G}''_T$ and infinitely many chains of every size (finite and infinite). This last step takes care of the non-valid computations of N_T. It is straightforward to see that $\mathcal{G}_T$ is automatically presentable.

For every $w \in W \subseteq G_T$,

(i) $w \in T$ if and only if there is not an isolated infinite chain in $\mathcal{G}_T$ whose root is an immediate successor of w (*isolated* means that every element of this chain has at most one immediate successor in $\mathcal{G}_T$);
(ii) if $w \notin T$ then the tree in $\mathcal{G}_T$ rooted at w is isomorphic to $\mathcal{O}$ (this is because T is downward closed).

The first property ensures that every isomorphism $f : \mathcal{G}_{T_1} \cong \mathcal{G}_{T_2}$ restricts to an isomorphism between T_1 and T_2, since f maps an isolated chain of size $\kappa \leq \omega$ to an isolated chain of the same size.

Conversely, in order to extend an isomorphism $g : T_1 \cong T_2$ to $\mathcal{G}_{T_1} \cong \mathcal{G}_{T_2}$, note that every $w \in W \subseteq G_T$ has infinitely many immediate successors $v \in W$ that are the roots in $\mathcal{G}_T$ of trees isomorphic to $\mathcal{O}$. This mitigates against the case that the cardinality of $S(w) \setminus T_1$ is different from the cardinality of $S(g(w)) \setminus T_2$, where $S(\cdot) \subseteq W$ is the function denoting the set of immediate successors in the tree $\mathcal{W}$. ⊣

This theorem can be extended to other classes of automatically presentable structures.

If $(T; \preceq)$ is a tree, then call $(T; S_{\preceq})$ a *successor tree*, where $S_{\preceq}(x)$ is defined as the set of immediate $\preceq$-successors of $x \in T$. The proof of Theorem 3.53 can easily be adapted to show that the isomorphism problem for the class of automatically presentable successor trees is Σ^1_1-complete. We can get the result for undirected graphs by attaching a gadget (such as a cycle of length 3) to identify the root of the successor tree, and then considering the underlying graph.

Nies [46] observes that since the class of undirected graphs is bi-interpretable in a variety of other classes, one gets Σ^1_1-completeness for these classes as well.

Corollary 3.54. *The isomorphism problem for each of the following classes is Σ^1_1-complete.*

(i) *The automatically presentable successor trees.*
(ii) *The automatically presentable undirected graphs.*
(iii) *The automatically presentable commutative monoids.*
(iv) *The automatically presentable partial orders.*
(v) *The automatically presentable lattices of height* 4.
(vi) *The automatically presentable algebras consisting of two* 1*-ary functions.*

§4. Classifications.

Fix a class C of structures (say the class of Boolean algebras, or linear orders). Which members of C are automatically presentable? This section is concerned with describing, in terms of classical invariants, the isomorphism types of the automatic members of C.

Each subsection begins with the classification of the unary-automatically presentable structures in the given class. These are usually based on the

https://doi.org/10.2178/bsl/1208442827

following proposition, whose proof follows from an analysis of the structure of the corresponding automata.

PROPOSITION 4.1. [9, 51] *If a graph $\mathcal{G} = (G; E)$ is a* unary-*automatically presentable, then it contains a finite number of infinite connected components, and a finite bound on the sizes of the finite connected components.*

For the non-unary case, Proposition 2.6 says that it is sufficient to consider $|\Sigma| = 2$.

Remark 4.2. Each of the following subsections indicate that the classification of classes of unary-automatically presentable structures is considerably simpler than the classification of classes of automatically presentable structures over a binary alphabet.

4.1. Equivalence structures. An *equivalence structure* $(E; \rho)$ is one where ρ is an equivalence relation on the set E. Each equivalence structure is characterised up to isomorphism by the number of equivalence classes of every size. To this end, define the *height function* $h_{\mathcal{E}}\colon \mathbb{N} \cup \{\infty\} \to \mathbb{N} \cup \{\infty\}$ of an equivalence structure $\mathcal{E}$ as $h(n)$ being the number (possibly infinite) of equivalence classes of size n (possibly infinite). Then $\mathcal{E}_1$ is isomorphic to $\mathcal{E}_2$ if and only if $h_{\mathcal{E}_1} = h_{\mathcal{E}_2}$.

The unary case follows immediately from Proposition 4.1.

THEOREM 4.3. [9, 51] *An equivalence structure $(E; \rho)$ has a* unary-*automatic presentation if and only if $h_{\mathcal{E}}(\infty)$ is finite and $h_{\mathcal{E}}(n) = 0$ for all but finitely many $n \in \mathbb{N}$.*

We now turn to the non-unary case. Consider an automatically presentable equivalence structure $\mathcal{E}$. The set of elements of E in infinite classes is definable using $\exists^\infty$, and so we can compute how many infinite classes there are. So to characterise the automatically presentable equivalence structures it is sufficient to consider those with no infinite classes. To this end, write $\mathcal{H}$ for the set of height functions of automatically presentable equivalence structures with no infinite classes. Adopt the convention that $m + n = m \times n = \infty$ if at least one of m or n is ∞.

PROPOSITION 4.4. [51]

(i) *$\mathcal{H}$ contains all functions of the form $h_f\colon \mathbb{N} \to \mathbb{N}$ defined by $h(f(n)) = 1$, where $f : \mathbb{N} \to \mathbb{N}$ is either a polynomial (with coefficients in $\mathbb{N}$) or an exponential k^{an+b} (for some $k, a, b \in \mathbb{N}$).*

(ii) *If $h, h' \in \mathcal{H}$ then so is their*
  (a) *sum $(h + h')(n) = h(n) + h'(n)$,*
  (b) *Dirichlet convolution $(h \star h')(n) = \Sigma_{ab=n} h(a)h'(b)$, and*
  (c) *Cauchy product $(h\#h')(n) = \Sigma_{a+b=n} h(a)h'(b)$.*

PROOF. For the first item, consider automatic copy $(R; \mathrm{el})$ of an equivalence structure where $R \subseteq \Sigma^\star$ and el is the equal length predicate. Thus the equivalence classes are of the form $R^{=n}$ for $n \in \mathbb{N}$. But for every polynomial

https://doi.org/10.2178/bsl/1208442827

or exponential function f as in the statement of the proposition, there is a regular set $R_f \subseteq \Sigma^\star$ with $f(n) = |(R_f)^{=n}|$ (see [53] or [51, Lemma D.2.3]).

For the second item, sum corresponds to disjoint union of equivalence structures, and Dirichlet convolution corresponds to direct product. Cauchy product is only slightly more involved. $\dashv$

So in particular there is an automatically presentable equivalence structure whose height function has unbounded range. As far as I know, neither is there a classification of $\mathcal{H}$, nor is it known whether the isomorphism problem for automatically presentable equivalence relations, seen to be Π^0_1, is decidable.

4.2. Linear orders. An excellent reference for linear orders is [50]. The classical ranking of linear orders $\mathcal{L} = (L; \leq)$ is based on iteratively factoring $\mathcal{L}$ by the equivalence relation c stating that x is equivalent to y if the number of elements between x and y is finite.

For ease of reference, here is the formal definition. For every ordinal α define an equivalence relation $\sim_\alpha$ on $\mathcal{L}$ inductively: $x \sim_\alpha y$ if $x = y$ or for some $\beta < \alpha$, the number of elements in $\mathcal{L}/\sim_\beta$ between $[x]_{\sim_\beta}$ and $[y]_{\sim_\beta}$ is finite. Here $\mathcal{L}/\sim_\beta$ is the linear ordering defined as follows: its domain is the collection of non-empty $\sim_\beta$-equivalence classes $[x]_{\sim_\beta}$ $(x \in L)$. The order is defined by $[x]_{\sim_\beta}$ less than $[y]_{\sim_\beta}$ if $a <_L b$ for every $a \in [x]_{\sim_\beta}, b \in [y]_{\sim_\beta}$.

The FC-*rank of* $\mathcal{L}$ is defined as the least ordinal α so that for every $\beta < \alpha$ and every $x \in L$, we have $[x]_{\sim_\beta} = [x]_{\sim_\alpha}$. For example, the FC-rank of the ordinal ω^α is α.

Here FC is an acronym for 'finite condensation', hinting that elements a finite distance apart are condensed together.

A linear order $\mathcal{L}$ is *dense* if for every distinct $a, b \in L$, there is a $z \in L$ with $a < z < b$. Note that the linear order with exactly one element is dense. A linear order is *scattered* if none of its suborderings are both dense and infinite. The *ordered sum* of the orderings $\mathcal{A}_b$ indexed by the linear order $\mathcal{B}$ is the result of replacing each $b \in B$ by a copy of $\mathcal{A}_b$, and is written $\Sigma_\mathcal{B}\mathcal{A}_b$.

Fact 4.5. Every linear order $\mathcal{L}$ is a dense sum of scattered linear orders. That is, $\mathcal{L}$ can be expressed as $\Sigma_\mathcal{D}\mathcal{L}_d$ for some dense $\mathcal{D}$ and scattered $\mathcal{L}_d$s, $d \in D$.

I do not know of a non-machine theoretic classification of the automatically presentable linear orders, or even the FC-rank 1 linear orders (it is straightforward to exhibit a linear order of FC-rank 1 with undecidable first-order theory). The next result describes the unary case.

THEOREM 4.6. [9, 51] *A linear order* $(L; \leq)$ *has a* unary-*automatic presentation if and only if it is a finite sum of scattered orders of* FC-*rank at most* 1.

In other words, the linear orders with unary-automatic presentations are finite sums of linear orders amongst the set ω, ω^*, and $\boldsymbol{n}$, for $n < \omega$. In particular, the order type of the rationals is not automatically presentable over a unary alphabet, and the least ordinal without a unary automatic presentation is ω^2.

Regarding the general (non-unary) case, Khoussainov and Nerode asked for the least ordinal without an automatic presentation [34] (it is easy to see that the automatically presentable ordinals form an initial segment of the countable ordinals). In Corollary 3.52, we saw that the condition of Theorem 3.51 implies that the answer is ω^ω.

Similarly, we can use the condition of Theorem 3.51 to get a generalisation to linear orders [37].

THEOREM 4.7. *The* FC*-rank of every automatically presentable linear order is finite.*

PROOF. In Theorem 3.51, let $\mathcal{L}$ be an automatically presentable linear order, and $\phi(x, y_1, y_2)$ be the relation $y_1 \leq x \leq y_2$. Then there is a finite set of structures $\mathcal{S}$, so that for every $a_1 \leq a_2 \in L$, there is a partition of the domain of $[a_1, a_2]$ into finitely many pieces $\{A_i\}$, so that every $\mathcal{L} \upharpoonright A_i$ is isomorphic to some structure in $\mathcal{S}$.

Just as in Corollary 3.52, we look for a decomposition property such as: If the domain of an automatically presentable linear order $\mathcal{L}$ is partitioned into finitely many pieces $\{B_i\}_i$, then some $\mathcal{L} \upharpoonright B_i$ has the same FC-rank as $\mathcal{L}$. Of course this condition fails horribly if $\mathcal{L}$ is not scattered, because $\mathcal{L}$ then embeds orders of arbitrary countable FC-rank. However, in the scattered case it just falls short; for instance, partition $\mathcal{L} = \omega + \omega$ into the first copy of ω ($B_1 = \{0, 1, \dots\}$) and the second copy ($B_2 = \{\omega, \omega + 1, \dots\}$). Then $\mathrm{FC}(\mathcal{L}) = 2$, but $\mathrm{FC}(\mathcal{B}_i) = 1$.

By slightly altering the notion of rank we can achieve the required decomposition result for scattered linear orders. Define the FC_*-rank of a scattered linear order $\mathcal{L}$ as the least ordinal α such that $\mathcal{L}$ can be expressed as a finite sum of orders of FC-rank at most α. Then $\mathrm{FC}_*(\mathcal{L}) \leq \mathrm{FC}(\mathcal{L}) \leq \mathrm{FC}_*(\mathcal{L})+1$.

The following decomposition property can be proved by induction on FC_*-rank: If the domain of a scattered linear order $\mathcal{L}$ is partitioned into finitely many pieces $\{B_i\}_i$, then there is some i with $\mathrm{FC}_*(\mathcal{L} \upharpoonright B_i) = \mathrm{FC}_*(\mathcal{L})$.

Combining this with the statement in the first paragraph, we see that for every $a_1 \leq a_2 \in L$, if $\mathcal{L} \upharpoonright [a_1, a_2]$ is scattered, say with FC_*-rank α, then $\mathcal{S}$ contains a scattered linear order of FC_*-rank α. Moreover, α is finite; for otherwise, using elementary properties of rank, $\mathcal{L} \upharpoonright [a_1, a_2]$ would contain infinitely many closed scattered subintervals having pairwise distinct FC_*-ranks, contradicting the finiteness of $\mathcal{S}$. Thus there is a uniform finite bound, say k, on the FC-rank of every closed scattered subinterval of $\mathcal{L}$.

Now if $\mathcal{L}$ is scattered, then every two elements $a_1, a_2 \in L$ are condensed within k steps; thus $\mathrm{FC}(\mathcal{L}) \leq k$. Incase $\mathcal{L}$ is not scattered, by Fact 4.5, $\mathcal{L}$ is the sum of scattered orders $\mathcal{L}_d$ ($d \in \mathcal{D}$, $\mathcal{D}$ dense). Consequently, each $\mathcal{L}_d$ has FC-rank at most k, and so $\mathcal{L}$ has FC-rank at most k. ⊣

Some decidability results for automatically presentable linear orders now follow. Since the axioms stating that $\leq$ linearly orders L are first-order, it is decidable whether a given automatically presentable structure $(L; \leq)$ is a linear order or not.

Corollary 4.8. [37] *Let $\mathcal{L} = (L; \leq)$ be an automatically presentable linear order.*

(i) *It is decidable whether or not $\mathcal{L}$ is scattered. In case $\mathcal{L}$ is not scattered, we can effectively compute an automatically presentable dense subordering $\mathcal{D}$ and automatically presentable scattered suborderings $\mathcal{L}_d$ for which $\mathcal{L} = \Sigma_{\mathcal{D}}\mathcal{L}_d$.*
(ii) *It is decidable whether or not $\mathcal{L}$ is isomorphic to an ordinal.*
(iii) *The isomorphism problem for automatically presentable ordinals is decidable. In fact the Cantor-normal-form may be extracted from an automatic copy of an ordinal.*

It is not known whether the isomorphism problem for automatically presentable linear orderings is decidable.

Cantor's theorems. One of Cantor's theorems says that every countable linear ordering embeds in the rational ordering $\mathbb{Q}$.

There are, potentially, a variety of possible automatic versions. The following proposition is the best known.

Proposition 4.9. [38] *Every automatic copy $\mathcal{M}$ of a linear order can be embedded into some automatic copy of $\mathbb{Q}$ by a function $f : \mathcal{M} \to \mathbb{Q}$ with the following properties*:

(i) *The function f is continuous with respect to the order topology.*
(ii) *The graph of f is regular.*

Proof. We mainly present the definition of the required automatic copy of $\mathbb{Q}$. Let $\mathcal{M} = (M; \leq_M)$ be an automatic copy of a linear order, where $M \subseteq \Delta^\star$ and $0, 1 \notin \Delta$. Define $M_L \subseteq M$ as the set of elements of $\mathcal{M}$ that cannot be approximated from the left. That is, $w \in M_L$ if and only if

$$(\exists u)[u <_M w \wedge \neg(\exists z)(u <_M z <_M w)] \vee \neg(\exists u)[u <_M w].$$

Similarly define $M_R \subseteq M$ as the set of elements that cannot be approximated from the right. The required automatic copy of $\mathbb{Q}$ has domain

$$M \cup M_L \cdot 0\{0,1\}^\star \cup M_R \cdot 1\{0,1\}^\star$$

and relation $uw \leq u'w'$ (here $u, u' \in M$ and $w, w' \in \{0,1\}^\star$) if $u \leq_M u'$ or otherwise $u = u'$ and $w \sqsubseteq_Q w'$. Here $(\{0,1\}^\star, \sqsubseteq_Q)$ is the automatic copy of $\mathbb{Q}$ from Example 2.4 (5). The required function f maps $u \mapsto u$ for $u \in M$. ⊣

It is not known whether there is a single automatic copy of $\mathbb{Q}$ that embeds, in the sense above, all automatic copies of all automatically presentable linear orders $\mathcal{M}$.

Cantor also proved that $\mathbb{Q}$ is homogeneous: For every two tuples $x_1 < \cdots < x_m$ and $y_1 < \cdots < y_m$ there is an automorphism $f : \mathbb{Q} \to \mathbb{Q}$ with $f(x_i) = y_i$ for $i \leq m$. Again there might be a number of automatic variations. Call an automatic copy of $\mathbb{Q}$ *automatically homogeneous* if for every two tuples there is an automorphism as above that is also regular.

PROPOSITION 4.10. [38] *The automatic copy of* $\mathbb{Q}$ *in Example* 2.4(5) *is automatically homogeneous. There exists an automatic copy of* $\mathbb{Q}$ *that is not automatically homogeneous.*

4.3. Trees. A *tree* $(T; \preceq)$ is a partial order with a smallest element and the property that for every element $v \in T$, the set $\{w \in T \mid w \preceq v\}$ is finite and linearly ordered by $\preceq$.

For $u \in T$, write $S(u)$ for the set of $\prec$-immediate successors of u; namely,

$$S(u) = \{v \in T \mid u \prec v \wedge \forall z[u \prec z \preceq v \to z = v]\}.$$

A tree is *finitely-branching* if $S(u)$ is finite for every $u \in T$. An *infinite path* in a tree is a sequence of elements $(w_i)_{i \in \mathbb{N}}$ with $w_{i+1} \in S(w_i)$ for every i.

The complexity of the isomorphism problem for automatically presentable trees, over the signature of partial orders, is not known. However, another measure of the complexity of a tree is its Cantor–Bendixson rank. Before giving the definition, we need some preliminary concepts. Define the extendible part $E(\mathcal{T})$ of $\mathcal{T}$ as those $x \in T$ that are on some infinite path of T. Define $d(\mathcal{T})$, the *derivative of* $\mathcal{T}$, as the subtree of $\mathcal{T}$ restricted to those elements that are on two distinct infinite paths, namely:

$$\{x \in T \mid \exists z, z' \in E(\mathcal{T}), z, z' \succ x \text{ and neither } z \preceq z' \text{ nor } z' \preceq z\}.$$

For each ordinal α define the iterated operation $d^\alpha(\mathcal{T})$ inductively by

$$d^\alpha(\mathcal{T}) = \{x \in T : \forall \beta < \alpha[x \text{ is an element of the tree } d(d^\beta(\mathcal{T}))]\}.$$

Note that for $\alpha = 0$ the range of the universal quantifier is void, and therefore $d^0(\mathcal{T})$ is just T.

DEFINITION 4.11. The *Cantor–Bendixson rank* $\mathrm{CB}(\mathcal{T})$ of a tree $\mathcal{T}$ is the least ordinal α such that $d^\alpha(\mathcal{T}) = d^{\alpha+1}(\mathcal{T})$.

Since $\mathcal{T}$ is countable, $\mathrm{CB}(\mathcal{T})$ is a countable ordinal. Also, if $\mathcal{T}$ contains countably many infinite paths, then $d^{\mathrm{CB}(\mathcal{T})}$ is the empty tree.

THEOREM 4.12. [37] *The* CB*-rank of every automatically presentable tree* $\mathcal{T} = (T; \preceq)$ *is finite and computable from* $\mathcal{T}$.

The main idea of the proof is to associate to each automatically presentable tree $\mathcal{T}$ an automatically presentable linear ordering, the Kleene–Brouwer ordering for instance, whose FC-rank can be used to bound the CB-rank of $\mathcal{T}$. By Theorem 4.7 this FC-rank is finite.

Kőnig's Lemma. Kőnig's Lemma says that an infinite finitely-branching tree has an infinite path. Here is an automatic analogue.

THEOREM 4.13. *If $\mathcal{T} = (T; \preceq)$ is an automatic copy of an infinite finitely-branching tree, then $\mathcal{T}$ has a regular infinite path. That is, there exists a regular set $P \subseteq T$ where P is an infinite path of $\mathcal{T}$.*

PROOF. Define a set P as those elements x such that $\exists^\infty w[x \prec w]$ and for which every $y \prec x$ satisfies that

$$\forall z, z' \in S(y)[z \preceq x \Rightarrow z \leq_{\text{llex}} z'].$$

Then P is the length-lexicographically least infinite path of $\mathcal{T}$ (in the ordering induced by the finite strings presenting the tree). ⊣

However, using the 2-Ramsey quantifier we can do more.

THEOREM 4.14. *If $\mathcal{T}$ is an automatic copy of a tree with countably many infinite paths, then every infinite path is regular.*

PROOF. Denote by $E(\mathcal{T}) \subseteq T$ the set of elements of a tree $\mathcal{T}$ that are on infinite paths. It is definable in $\mathcal{T}$ using the 2-Ramsey quantifier, so Theorem 3.20 gives that $E(\mathcal{T})$ is regular. Then every isolated path of $\mathcal{T}$ is regular, since it is definable as $\{x \in E(\mathcal{T}) \mid p \preceq x\} \cup \{x \in E(\mathcal{T}) \mid x \prec p\}$, for suitable $p \in E(\mathcal{T})$. Replace $\mathcal{T}$ by its derivative $d(\mathcal{T})$, which is also automatically presentable. Since the CB-rank of $\mathcal{T}$ is finite (Theorem 4.12), and $d^{\text{CB}(\mathcal{T})}$ is the empty tree, every infinite path is defined in this way. ⊣

The trees considered above are partial orders. Instead we may consider successor trees: structures of the form $(T; S, r)$ where S is the immediate successor relation and r is the root. These behave quite differently from trees in the signature of partial orders. By Theorem 3.53, the isomorphism problem for successor trees is Σ^1_1-complete. A slight modification of Corollary 3.41 yields the following pathology.

PROPOSITION 4.15. *There is an automatic copy $(T; S, r)$ of a successor tree with exactly two infinite paths, neither of which is a regular subset of T.*

4.4. Boolean algebras. Boolean algebras are considered in the signature $(\vee, \wedge, \neg)$. Write $\mathbf{0}$ for the bottom element, $\mathbf{1}$ for the top and $\subseteq$ for the associated partial order.

The following lemma, which will be useful for classifying the automatically presentable Boolean algebras, is a direct application of Proposition 3.46.

LEMMA 4.16. *Suppose $\mathcal{B}$ is an automatic copy of a Boolean algebra. There exists a constant $e \in \mathbb{N}$, that depends only on the automaton recognising $\vee$, such that for every set S of n elements of B, the length of the element $\bigvee S$ is at most*

$$\max_{s \in S}\{|s|\} + e \log n.$$

PROPOSITION 4.17. *The countable atomless Boolean algebra is not automatically presentable.*

PROOF. Suppose $\mathcal{B} = (B; \vee, \wedge, \neg)$ is an automatic copy of the countable atomless Boolean algebra. The idea is that, starting with the top element $\boldsymbol{1}$, we can generate too many pairwise disjoint elements by repeated splitting.

For non-zero $x \in B$, there is exists a non-zero element $a < x$, because x is not an atom. Thus there are two disjoint non-zero elements $(x \wedge a)$ and $(x \wedge \neg a)$ below x. Define $a(x)$ as the length-lexicographically least element a with this property. Expand the presentation of $\mathcal{B}$ to include the functions $f_l \colon x \mapsto x \wedge a(x)$ and $f_r \colon x \mapsto x \wedge \neg a(x)$. Apply Proposition 3.44 with respect to the functions f_l, f_r and with $D = \{\boldsymbol{1}\}$: There exists a linear function t so that every string in $G_n(\boldsymbol{1})$ has length at most $t(n)$.

The subalgebra $\mathcal{A}$ of $\mathcal{B}$ generated by $G_n(\{\boldsymbol{1}\})$ has $\exp(n)$ atoms. By Lemma 4.16, there is a constant e, so that every element of $\mathcal{A}$ has length at most $t(n) + en$. So, there are only $|\Sigma|^{t(n)+en}$ available strings to code for elements of $\mathcal{A}$; contradicting the fact that $\mathcal{A}$ has $\exp_2(n)$ elements. ⊣

Consequently, no automatically presentable Boolean algebra has the property that the countable atomless boolean algebra is definable in it. In particular, if an automatically presentable Boolean algebra has finitely many atoms, then it is finite.

COROLLARY 4.18. *There are no infinite* unary-*automatically presentable Boolean algebras.*

PROOF. Suppose $\mathcal{B}$ is a unary-automatic copy of a Boolean algebra with infinitely many atoms. The set of atoms B_A, being first-order definable in $\mathcal{B}$, is regular. Apply Proposition 3.44 to the functions $\vee$, $\wedge$, and $\neg$, and the set B_A. For $n \geq 3$, the set $G_{n+2}(B_A)$ contains every element generated by the first n atoms. So $|G_{n+2}(B_A)| \geq 2^n$, which exceeds the linear bound. ⊣

Recall that $\mathcal{B}_{\text{fin}}$ is the Boolean algebra of finite or co-finite subsets of $\mathbb{N}$. Write $\mathcal{B}_{\text{fin}}^n$ for the n-fold power. A refinement of the proof of Proposition 4.17 is used to classify the automatically presentable Boolean algebras in the non-unary case.

THEOREM 4.19. [35] *An infinite Boolean algebra is automatically presentable if and only if is isomorphic to $\mathcal{B}_{\text{fin}}^n$ for some $n \in \mathbb{N}$.*

COROLLARY 4.20. *The isomorphism problem for automatically presentable Boolean algebras is decidable.*

PROOF. Start with an automatic copy of an infinite Boolean algebra $\mathcal{B}$. Recall that the Frechét congruence on $\mathcal{B}$ consists of those pairs whose symmetric difference is either $\boldsymbol{0}$ or a meet of a finite number of atoms of $\mathcal{B}$. This congruence is first-order definable in $\mathcal{B}$ with the additional quantifier $\exists^\infty$, and so is regular. Hence the quotient of $\mathcal{B}$ by its Frechét congruence is automatically presentable. Compute the least number of times one has to factor $\mathcal{B}$ until one reaches the two element algebra, say n times. This value defines $\mathcal{B}$ up to isomorphism; indeed, $\mathcal{B}$ is isomorphic to $\mathcal{B}_{\text{fin}}^n$. ⊣

4.5. Groups, rings and fields. By algebraically characterising the relations recognised by automata over a unary alphabet, Blumensath establishes the following theorem.

THEOREM 4.21. [9] *There are no infinite* unary-*automatically presentable groups* $(G; \cdot)$, *rings, or fields.*

Let us consider the non-unary cases. Recall that an integral domain $(D; +, \mathbf{0}, \cdot, \mathbf{1})$ is a commutative ring with identity with the property that if $x \cdot y = \mathbf{0}$ then $x = \mathbf{0}$ or $y = \mathbf{0}$. With a little work, the ideas in Section 3.5 (Restriction on growth) can be used to prove the next result.

THEOREM 4.22. [35] *There are no infinite automatically presentable integral domains. In particular, there are no infinite automatically presentable fields.*

However, the cases of groups and rings are open. The current state of affairs is nicely detailed in Nies [46]. I only mention a sample here.

For $k \geq 2$, write $\mathbb{Z}[1/k]$ for the additive group of rationals of the form z/k^i, where $z \in \mathbb{Z}$ and $i \in \mathbb{N}$. It is straightforward to show that $\mathbb{Z}[1/k]$ and $\mathbb{Z}[1/k]/\mathbb{Z}$ have automatic presentations.

PROPOSITION 4.23. [35] *The following groups do not have automatic presentations*:

(i) *The positive rationals under multiplication.*
(ii) *The direct sum of infinitely many copies of* $\mathbb{Z}[1/k]$, *for a fixed* k.
(iii) *The direct sum of infinitely many copies of the Prüfer group* $\mathbb{Z}[1/k]/\mathbb{Z}$, *for a fixed* k.

PROOF. The idea, illustrated here in the proof of the first item, is to define sets F_n of lots of distinct elements using the primes of length at most n, and then show, as before using Proposition 3.46, that the lengths of these elements are too 'short'.

Suppose that $(Q^+, \times)$ is automatically presentable; say an automatic copy is $(D; \times)$. Let P be the elements of D that correspond to primes. Recall $P^{\leq n}$ are the elements of P of length at most n. Let

$$F_n = \{ \prod_{p \in P^{\leq n}} p^{\beta_p} : 0 \leq \beta_p \leq 2^n \}.$$

Then F_n contains $\exp(nr_n)$ distinct elements, where $r_n = |P^{\leq n}|$.

On the other hand, we compute an upperbound on the size of F_n. By successive applications of Proposition 3.46, there is a constant e, so that $|p^\beta| \leq n + en \leq n(e+1)$ and so each string has length at most $n(e+1) + e \log r_n$. This places an upper bound of $\exp(n(e+1) + e \log r_n))$ on the size of F_n, contradicting that r_n goes to infinity. ⊣

Using a mixture of algebra, coding techniques, and growth arguments, Nies and Thomas [47] prove the next result (the first item was already established for finitely generated groups $\mathcal{G}$ by Oliver and Thomas [48]).

THEOREM 4.24. (i) *Let $\mathcal{G}$ be an automatically presentable infinite group. Then every finitely generated subgroup of $\mathcal{G}$ has an Abelian subgroup of finite index.*
(ii) *Every finitely generated subring of an automatically presentable ring is finite.*

§**5. Open problems.** This work has only dealt with automata operating on finite strings. The fundamental result that makes the theory work is Theorem 3.1 [Definability]. Analogues of this theorem hold for other models of finite automata: automata operating on finite trees (ranked or unranked), infinite strings, and infinite trees. Each model yields a class of 'automatically presentable structures' where the basic theory goes through: particularly, a logical characterisation via interpretability similar to Theorem 3.8. Of course, the problems considered in this survey—such as classification, intrinsic regularity, and proving non-automaticity—can be asked in these more general settings, see for instance [24, 39, 21].

However, there are still many problems in the finite string case. Here are some problems whose solutions will likely require new techniques.

Problem 5.1. Find new ways to prove that a structure does not have an automatic presentation. For instance, do the following structures have automatic presentations?

(i) The additive group of rationals $(\mathbb{Q};+)$. This structure is WMSO-interpretable in a certain infinite string (compare [46]). This limits the kind of proof that could be used to show that $(\mathbb{Q};+)$ has no finite-string automatic presentation.
(ii) The graph generated by the ground term rewriting system with one constant a, one binary function f, initial term $f(a,a)$ and rewriting rule $a \mapsto f(a,a)$. In other words, the structure whose domain consists of all finite binary branching trees t, and for which there is an edge from t to t' if t' extends t by exactly one node. This is a candidate for separating the class of graphs generated by ground term rewriting systems from the finite-string automatic graphs (see [42]).

Problem 5.2. Investigate the complexity in the arithmetic/analytic hierarchy of the isomorphism problem for classes of automatically presentable structures. For instance, what is the complexity of the isomorphism problem for finite-string automatically presentable linear orders, or equivalence structures?

Problem 5.3. Can we capture intrinsic regularity using definability?

For instance, is it the case that for every finite-string automatically presentable structure $\mathcal{A}$,

$$\mathrm{FO}[\cup_n \boldsymbol{Q}_n^{\mathrm{reg}}]_{\omega-\mathrm{inv}}(\mathcal{A}) = \mathrm{IR}(\mathcal{A})?$$

REFERENCES

[1] C. J. Ash and A. Nerode, *Intrinsically recursive relations*, ***Aspects of effective algebra (clayton, 1979)***, Upside Down A Book Co., Yarra Glen, 1981, pp. 26–41.

[2] V. Bárány, *Invariants of automatic presentations and semi-synchronous transductions*, ***STACS 2006*** (B. Durand and W. Thomas, editors), Lecture Notes in Computer Science, vol. 3884, Springer, 2006, pp. 289–300.

[3] ———, ***Automatic presentations of infinite structures***, Ph.D. thesis, RWTH Aachen, 2007.

[4] V. Bárány, L. Kaiser, and S. Rubin, *Countable omega-automatic structures and their presentations*, ***STACS 2008*** (S. Albers, P. Weil, and C. Rochange, editors), Dagstuhl Seminar Proceedings, vol. 08001, Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2008.

[5] M. Benedikt and L. Libkin, *Tree extension algebras: logics, automata, and query languages*, ***LICS 2002***, IEEE Computer Society, 2002, pp. 203–212.

[6] M. Benedikt, L. Libkin, T. Schwentick, and L. Segoufin, *A model-theoretic approach to regular string relations*, ***LICS 2001***, IEEE Computer Society, 2001, pp. 431–440.

[7] C. H. Bennett, *Logical reversibility of computation*, ***IBM Journal of Research and Development***, vol. 17 (1973), no. 6, pp. 525–532.

[8] Jean Berstel, ***Transductions and context-free languages***, Leitfäden der Angewandten Mathematik und Mechanik [Guides to Applied Mathematics and Mechanics], vol. 38, B. G. Teubner, Stuttgart, 1979.

[9] A. Blumensath, ***Automatic structures***, Diploma thesis, RWTH Aachen, 1999.

[10] ———, *Prefix-recognisable graphs and monadic second-order logic*, Technical report, RWTH Aachen, 2001.

[11] ———, *Axiomatising tree-interpretable structures*, ***STACS 2002*** (H. Alt and A. Ferreira, editors), Lecture Notes in Computer Science, vol. 2285, Springer, 2002, pp. 596–607.

[12] A. Blumensath and E. Grädel, *Automatic structures*, ***LICS 2000***, IEEE Computer Society, 2000, pp. 51–62.

[13] ———, *Finite presentations of infinite structures: Automata and interpretations*, ***Proceedings of the 2nd International Workshop on Complexity in Automated Deduction, CiAD 2002***, 2002.

[14] ———, *Finite presentations of infinite structures: Automata and interpretations*, ***Theory of Computing Systems***, vol. 37 (2004), pp. 641–674.

[15] V. Bruyère, G. Hansel, C. Michaux Christian, and R. Villemaire, *Logic and p-recognizable sets of integers*, ***Bulletin of the Belgian Mathematical Society. Simon Stevin***, vol. 1 (1994), no. 2, pp. 191–238, Journées Montoises (Mons, 1992).

[16] J. R. Büchi, *Weak second-order arithmetic and finite automata*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 6 (1960), pp. 66–92.

[17] ———, *On a decision method in restricted second order arithmetic*, ***Logic, methodology and philosophy of science***, Stanford University Press, 1962, pp. 1–11.

[18] J. W. Cannon, *The combinatorial structure of cocompact discrete hyperbolic groups*, ***Geometriae Dedicata (Historical Archive)***, vol. 16 (1984), no. 2, pp. 123–148.

[19] J. W. Cannon, D. B. H. Epstein, D. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston, ***Word processing in groups***, Jones and Bartlett, 1992.

[20] D. Cenzer and J. B. Remmel, *Complexity-theoretic model theory and algebra*, ***Handbook of recursive mathematics, vol. 1*** (Yu. L. Ershov, S. S. Goncharov, A. Nerode, and J. B. Remmel, editors), Studies in Logic and the Foundations of Mathematics, vol. 138, North-Holland, Amsterdam, 1998, pp. 381–513.

[21] T. Colcombet and C. Löding, *Transforming structures by set interpretations*, Technical Report AIB-2006-07, RWTH Aachen, 2006.

[22] C. Delhommé, *Non-automaticity of ω^ω*, 2001, manuscript.

[23] ———, *The Rado graph is not automatic*, 2001, manuscript.

[24] ———, *Automaticité des ordinaux et des graphes homogènes*, ***Comptes Rendus Mathematique***, vol. 339 (2004), no. 1, pp. 5–10.

[25] J. Doner, *Tree acceptors and some of their applications*, ***Journal of Computer and System Sciences***, vol. 4 (1970), pp. 406–451.

[26] S. Eilenberg, C. C. Elgot, and J. C. Shepherdson, *Sets recognised by n–tape automata*, ***Journal of Algebra***, vol. 13 (1969), no. 4, pp. 447–464.

[27] C. C. Elgot, *Decision problems of finite automata design and related arithmetics*, ***Transactions of the American Mathematical Society***, vol. 98 (1961), pp. 21–51.

[28] S. S. Goncharov and J. F. Knight, *Computable structure and non-structure theorems*, ***Algebra and Logic***, vol. 41 (2002), no. 6, pp. 351–373.

[29] L. Hella, *Definability hierarchies of generalized quantifiers*, ***Annals of Pure and Applied Logic***, vol. 43 (1989), no. 3, pp. 235–271.

[30] B. R. Hodgson, ***Théories décidables par automate fini***, Ph.D. thesis, University of Montréal, 1976.

[31] ———, *On direct products of automaton decidable theories*, ***Theoretical Computer Science***, vol. 19 (1982), pp. 331–335.

[32] ———, *Decidabilite par automate fini*, ***Annales de Sciences Mathématiques du Québec***, vol. 7 (1983), pp. 39–57.

[33] D. F. Holt, *The Warwick automatic groups software*, ***Geometric and computational perspectives on infinite groups (Minneapolis, MN and New Brunswick, NJ, 1994)***, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 25, American Mathematical Society, 1996, pp. 69–82.

[34] B. Khoussainov and A. Nerode, ***Automatic presentations of structures***, Lecture Notes in Computer Science, vol. 960, 1995.

[35] B. Khoussainov, A. Nies, S. Rubin, and F. Stephan, *Automatic structures: Richness and limitations*, ***Logical Methods in Computer Science***, vol. 3 (2007), no. 2, pp. 0–0.

[36] B. Khoussainov, S. Rubin, and F. Stephan, *Definability and regularity in automatic structures*, ***STACS 2004*** (V. Diekert and M. Habib, editors), Lecture Notes in Computer Science, vol. 2996, Springer, 2004, pp. 440–451.

[37] ———, *Automatic linear orders and trees*, ***ACM Transactions on Computational Logic***, vol. 6 (2005), no. 4, pp. 675–700.

[38] D. Kuske, *Is Cantor's theorem automatic?*, ***LPAR 2003*** (M. Y. Vardi and A. Voronkov, editors), Lecture Notes in Artificial Intelligence, vol. 2850, Springer, 2003, pp. 332–345.

[39] D. Kuske and M. Lohrey, *First-order and counting theories of omega-automatic structures*, Fakultätsbericht Nr. 2005/07, Universität Stuttgart, Fakultät Informatik, Elektrotechnik und Informationstechnik, September 2005.

[40] L. Libkin and F. Neven, *Logical definability and query languages over unranked trees*, ***LICS 2003***, IEEE Computer Society, 2003, pp. 178–187.

[41] Per Lindström, *First order predicate logic with generalized quantifiers*, ***Theoria***, vol. 32 (1966), pp. 186–195.

[42] Christof Löding, ***Infinite graphs generated by tree rewriting***, Ph.D. thesis, RWTH Aachen, 2003.

[43] M. Lohrey, *Automatic structures of bounded degree*, ***LPAR 2003*** (M. Y. Vardi and A. Voronkov, editors), Lecture Notes in Artificial Intelligence, vol. 2850, Springer, 2003, pp. 344–358.

[44] C. Michaux and F. Point, *Les ensembles k-reconnaissables sont définissables dans $\langle N, +, V_k\rangle$*, ***Comptes Rendus des Séances de l'Académie des Sciences. Série I. Mathématique***, vol. 303 (1986), no. 19, pp. 939–942.

[45] A. A. NABEBIN, *Multitape automata in a unary alphabet*, ***Trudy Moskovskogo Ordena Lenina Ènergeticheskogo Instituta***, vol. 292 (1976), pp. 7–11.

[46] A. NIES, *Describing Groups*, this BULLETIN, vol. 13 (2007), no. 3, pp. 305–339.

[47] A. NIES and R. THOMAS, *FA-presentable groups and rings*, 2005, to appear.

[48] G. OLIVER and R. THOMAS, *Automatic presentations of finitely generated groups*, ***STACS 2005*** (V. Diekert and B. Durand, editors), Lecture Notes in Computer Science, vol. 3404, Springer, 2005.

[49] M. O. RABIN, *Decidability of second-order theories and automata on infinite trees*, ***Transactions of the American Mathematical Society***, vol. 141 (1969), pp. 1–35.

[50] J. G. ROSENSTEIN, ***Linear orderings***, Academic Press, 1982.

[51] S. RUBIN, ***Automatic structures***, Ph.D. thesis, University of Auckland, 2004.

[52] F. STEPHAN, *The random graph is not automatically presentable*, 2002, manuscript.

[53] A. SZILARD, S. YU, K. ZHANG, and J. SHALLIT, *Characterizing regular languages with polynomial densities*, ***MFCS '92*** (I. M. Havel and V. Koubek, editors), Lecture Notes in Computer Science, vol. 629, Springer, 1992, pp. 494–503.

[54] J. W. THATCHER and J. B. WRIGHT, *Generalized finite automata theory with an application to a decision problem of second-order logic*, ***Mathematical Systems Theory***, vol. 2 (1968), pp. 57–81.

[55] W. THOMAS, *Automata on infinite objects*, ***Handbook of theoretical computer science*** (J. van Leeuwen, editor), vol. B: Formal models and semantics, Elsevier, 1990, pp. 133–191.

[56] B. A. TRAHTENBROT, *Finite automata and the logic of one-place predicates*. Russian, ***Siberian Mathematical Journal***, vol. 3 (1962), pp. 103–131, English translation: ***American Mathematical Society Translations, Series 2***, vol. 59 (1966), pp. 23–55.

[57] R. VILLEMAIRE, *The theory of* $\langle \boldsymbol{N}, +, V_k, V_l \rangle$ *is undecidable*, ***Theoretical Computer Science***, vol. 106 (1992), pp. 337–349.

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF AUCKLAND, NEW ZEALAND
E-mail: rubin@cs.auckland.ac.nz