Open questions in automatic structures

Bakh Khoussainov

Computer Science Department, The University of Auckland,
New Zealand

July 11, 2019

Acknowledgement

I would like to thank the LMS for the invitation and support.
This lectureship tour includes presentations at:

1. Oxford: *Algebraic structures, graphs, and automata*
2. Manchester: *Open questions in automatic structures*
3. Liverpool: *Open questions in automatic structures*
4. LLC: *Finding winners in games played on graphs*
5. LMS meeting: *Finitely presented expansions of groups*
6. Swansea: *Finitely presented expansions of groups*
7. St Andrews: *Algorithmically random structures*
8. Durham: *Effective aspects of differential games*.

Plan

- Brief introduction
- Basic definitions and examples
- Decidability theorem
- Characterisation theorems and algorithmic implications.

Definition

A **structure** $\mathcal{A}$ is a tuple $(A, R_0, \ldots, R_n, F_0, \ldots, F_m)$, where A is the domain of the structure, each R_i is a relation on A, and each F_j is a function on A.

- If no functions exists then the structure is relational.
- Structures can be transformed into relational structures.
- All our structures will be relational.

Brief history

1. Computable structures (Malcev, Rabin, Ershov, Nerode)
2. Feasible structures (Nerode, Remmel)
3. Automatic structures as refinement of feasible structures (Khoussainov-Nerode)
4. Automatic structures as extension of finite model theory (Gradel and Blumensath)

Automata in algebra and logic

- The work of Büchi and Rabin
- Groups defined by automata (Thurston, Holt, Grigorchuk)
- Integer programming and automata (Wolper)
- Theoretical foundation of databases (Libkin, Benedict)
- Verification and model checking
- Automata groups (Aleshin)

Finite automata

Definition

An **automaton** is a machine $\mathcal{M}$ with an initial state and accepting states whose transitions are of the form

$$< state, symbol, state > .$$

An automaton $\mathcal{M}$ accepts or rejects finite words over an alphabet. The **language** of $\mathcal{M}$ is

$$L(\mathcal{M}) = \{w \mid \text{ the word } w \text{ is accepted by} \mathcal{M}\}.$$

Automata recognizing relations

Automata can be used to recognize *n*-tuples of words $(w_1, \ldots, w_n)$. Such an automaton has *n* heads moving synchronously along the words

$$w_1,\ w_2,\ \ldots,\ w_n.$$

The transitions are of the form

$$< state, (symbol_1, \ldots, symbol_n), state > .$$

Definition

An *n*-ary relation *R* is **automatic (regular)** if there exists a synchronous automaton with *n* heads that recognizes *R*.

Definition

A structure $\mathcal{A} = (A; R_0, R_1, \ldots, R_m)$ is **automatic** if its domain A and all relations R_0, R_1, ..., R_m are automata recognizable.

Examples

1. $(1^\star; \leq, S)$ (The successor structure with the order)
2. $((0+1)^\star; \vee, \wedge, \neg)$ (The digit-wise and-or-not algebra)
3. $((0+1)^\star; \preceq;\ L,\ R,\ Eq)$ (The word structure).
4. $((0+1)^\star \cdot 1; +_2,\ S,\ \leq,\ |_2)$ (The weak arithmetic).
5. The configuration space $(Conf(T), E)$ of a TM T.

Automatic presentations

Definition

A structure $\mathcal{A}$ is **automata presentable** if it is isomorphic to an automatic structure $\mathcal{B}$.

The structure $\mathcal{B}$ is usually called an **automatic copy** of $\mathcal{A}$.

Examples:

1. Any finitely generated Abelian group.
2. The group Q_p.
3. The Boolean algebra of finite and co-finite subsets of ω.
4. The linear order $(Q, \leq)$.

Calculus of automata recognisable relations:

The closure properties for the following operations:

1. The union, intersection, and complementation.
2. The projection (also known as $\exists$-operation).
3. The instantiation and rearrangement.
4. The linkage/composition.
5. Cartesian product.

Decidability Theorem 1

Theorem (Khoussainov-Nerode, 1996)

There exists an algorithm that given an automatic structure $\mathcal{A}$ and a first order query $Q(x_1, \ldots, x_n)$ produces an automaton recognizing exactly those tuples $(a_1, \ldots, a_n)$ in the structure that make the query true.

Corollary

The first order theory of any automatic structure is decidable.

Corollary

If a structure has undecidable first order theory then it is not automatic.

Corollaries

The FO-theories of the following structures are decidable:

- The Presburger arithmetic.
- Any finitely generated Abelain group.
- Dense linear order.
- The weak arithmetic.
- The configuration graph of any Turing machine.
- etc.

Decidability Theorem 2

Consider the logic $(FO + \exists^{\infty} + \exists^{n,m})$.

Theorem (Khoussainov, Rubin, Stephan; 2003)

If $\mathcal{A}$ is automatic then there exists an algorithm that, applied to a $(FO + \exists^{\infty} + \exists^{n,m})$-definition of any relation R, produces an automaton that recognizes the relation.

In particular, the $(FO + \exists^{\infty} + \exists^{n,m})$-theory of $\mathcal{A}$ is decidable.

Kuske, Lohrey, Liu, Rubin extended this decidability theorem to other logics, e.g. logics that include Ramsey's quantifier.

Open Question 1

Extensions of the decidability theorem are about intrinsically regular relations:

Definition

A relation R on automatic structure is **intrinsically regular** if R is regular under all automatic presentations of the structure.

Question 1:

Is the natural order on $(Z; +)$ *intrinsically regular?*

Research themes:

- Find isomorphism invariants of automatic structures
- Study complexity of automatic structures
- Study the isomorphism problem for automatic structures

The Constant Growth Lemma

Lemma (Khoussainov, Nerode 1994)

Let $f : D^n \to D$ be a function such that the graph(f) is regular. There exists a constant C such that for all $x_1, \ldots, x_n \in D$:

$$|f(x_1, \ldots, x_n)| \leq max\{|x_1|, \ldots, |x_n|\} + C.$$

Proof. The Pumping lemma does the job. □

Generating sets

Let $\mathcal{A} = (A; \ F_0, F_1, \ldots, F_n)$ be automatic structure and $X \subset A$. Let us list the elements of X in length-lex-order:

$$x_1, x_2, x_3, \ldots.$$

Let C' be a constant such that $|x_n| \leq C' \cdot n$ for all $n \geq 1$.

Define $G_n(X)$:

1. $G_1(X) = \{x_1\}$.
2. $G_{n+1}(X) = G_n(X) \cup \{F_i(\bar{a}) \mid \bar{a} \in G_n(X)\} \cup \{x_{n+1}\}$.

The growth of generation theorem

Theorem (Khoussainov/Nerode; Blumensath/Gradel)

There exists a constant C such that for all $a \in G_n(X)$

$$|a| \leq C \cdot n.$$

In particular, $G_n(X) \subseteq \Sigma^{\leq C \cdot n}$ when $|\Sigma| > 1$, and $|G_n(X)| \leq C \cdot n$ when $|\Sigma| = 1$. □

Implications

Corollary

The following structures are not automatic:

- *The free semigroup* $(\Sigma^\star; \cdot)$.
- $(\omega; f)$, *where* $f : \omega^2 \to \omega$ *is a bijection.*
- *The free group* $F(n)$ *with* $n > 1$ *generators.*
- $(\omega; \times)$.
- $(\omega;\ Div(x, y))$.
- $(\omega; \leq, \{n! \mid n \in \omega\})$.

Automatic Boolean algebras

Examples:

1. The Boolean algebra $\mathcal{B}_\omega$, the collection of all finite or co-finite subsets of ω.
2. The Boolean algebra $\mathcal{B}_\omega^n$, where $n \geq 1$.

The Characterization Theorem for Boolean algebras

Theorem (Khoussainov, Nies, Rubin, Stephan)

A Boolean algebra is automatic if and only if it is isomorphic to $\mathcal{B}_\omega^n$ for some $n \geq 1$.

An application

Corollary

The isomorphism problem for automatic Boolean algebras is decidable.

Proof. Elements $a, b \in B$ are $\equiv_F$**-equivalent** if their symmetric difference $(a \cap \bar{b}) \cup (\bar{a} \cap b)$ is a finite union of atoms.

The factor algebra $\mathcal{B}/F$ is finite. Thus, $\mathcal{B}$ and $\mathcal{B}'$ are isomorphic iff $\mathcal{B}/F$ and $\mathcal{B}'/F'$ are isomorphic. □

Automatic linearly ordered sets

Example

- The ordinals ω, ω^2, ..., ω^n, ... are automatic.
- The linear order of rational numbers $(Q; \leq)$ is automatic.
- The order $Z + 1 + Z + 2 + Z + 4 + Z + 8 + \ldots$ is automatic.

Let $\mathcal{L} = (L; \leq)$ be a linear order. Define

$a \sim b$ if there are finitely many elements between a and b

By ordinal induction define:

$$\mathcal{L}_1 = \mathcal{L}/\sim, \mathcal{L}_{n+1} = \mathcal{L}_n/\sim, \ldots$$

CB-rank of linear orders

Definition

The least ordinal α such that $\mathcal{L}_\alpha = \mathcal{L}_{\alpha+1}$ is the **CB-rank** of $\mathcal{L}$.

Theorem (Khoussainov, Rubin, Stephan)

The CB-rank of any automatic linear order is finite.

Applications

Corollary

Given an automatic lo $\mathcal{L}$, we can compute the following:

1. *The CB-rank of $\mathcal{L}$,*
2. *If $\mathcal{L}$ embeds the order of rationals,*
3. *If $\mathcal{L}$ is a well-order,*
4. *The cantor normal form of $\mathcal{L}$ if $\mathcal{L}$ is an ordinal.*

Corollary

The isomorphism problem for automatic ordinals is decidable.

Open Question 2

Definition

A linear order is **scattered** if it has no dense sub-order.

Question 2:

Is the isomorphism problem for scattered automatic linear orders decidable?

Kuske, Lohrey, and Liu proved that the isomorphism problem for automatic linear orders is undecidable. Their proof uses non strongly discrete linear orders in an essential way.

FA presentable groups

Theorem

If a group G has FA presentation then all of its finitely generated subgroups are virtually abelian. In particular, a f.g. group has FA presentation iff it is virtually abelian.

The proof uses Gromov's theorem that characterises f.g. groups of polynomial growth.

This characterization theorem does not imply decidability of the isomorphism problem for f.g. FA presentable groups.

Open Questions 3

Question 3:

Is the isomorphism problem for f.g. automata presentable groups decidable?

Question 3a:

Is the isomorphism problem for automata presentable f.g. abelian groups decidable?

Cayley graphs

The *Cayley graph of G*, denoted by $\Gamma(G, A)$, is this:

- The vertices of the graph are the elements of the group.
- Put edges between vertices g and ga, $a \in A$.

Group G has a decidable word problem if and only if $\Gamma(G, A)$ is a computable graph.

Thurston automatic groups

Definition (Cannon, Epstein, Gillman, Holt, Thurston)

The group G with generator set A is **Thurston-automatic** if

1. There is a regular set $L \subseteq A^\star$ such that $\pi : L \to G$ is onto.
2. The word problem (on L) is regular.
3. For all $a \in A$, there is an automaton M_a recognising

$$\{(u, v) \mid u, v \in L \text{ and } u = va \text{ in } G\}.$$

The automata M and M_a, $a \in A$, are called **automatic structure** for G.

Properties of Thurston automatic groups:

- Generator set independent.
- Have decidable word problem (in quadratic time).
- Finitely presented.
- Closed under:
  - finite free products,
  - finite direct products,
  - finite extensions.

Examples

Thurston-automatic groups:

- Free abelian groups Z^n.
- Hyperbolic groups, e.g. free groups.
- Braid groups.
- Fundamental groups of many natural manifolds.
- Finitely generated FA presentable groups.

Non-Thurston-automatic groups:

- $SL_n(\mathbb{Z})$ and $H_3(\mathbb{Z})$.
- The wreath product of $\mathbb{Z}_2$ with $\mathbb{Z}$.
- Non-finitely presented groups.
- Baumslag-Solitar groups.

Automatic graphs

Definition

A graph $\Gamma = (V, E)$ is automatic of both V and E are FA recognizable sets.

Example

For a Turing machine T, consider the graph $(Conf(T), E_T)$:

1. $Conf(T)$ = all configurations of T, and
2. E_T = transitions of T.

The structure $(Conf(T), E_T)$ is an automatic graph.

Example

The n-dimensional grid $\mathbb{Z}^n$ is an automatic graph.

Definition

Let G be a group generated by a finite set X of generators. The group G is **Cayley automatic** if the graph $\Gamma(G, X)$ is automatic.

Example

- Finitely generated abelian groups are Cayley automatic.
- Thurston automatic groups are Cayley automatic.
- FA presentable groups.

Heisenberg group

The Heisenberg group $H_3(\mathbb{Z})$ consists of matrices X over $\mathbb{Z}$:

$$X = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

The group has 3 generators which are

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Multiplication by the generators in $H_3(\mathbb{Z})$

The multiplication of X by A, B, and C can be presented as:

$$X \cdot A = \begin{pmatrix} 1 & a+1 & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, \quad X \cdot B = \begin{pmatrix} 1 & a & b+1 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, \textit{ and}$$

$$X \cdot C = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c+1 \\ 0 & 0 & 1 \end{pmatrix}.$$

respectively. These are all automata recognizable events. Thus, $H_3(\mathbb{Z})$ is graph-automatic.

Nilpotent groups of class 2

Theorem (Kharlampovich, Khoussainov, Miasnikov)

Every finitely generated group G of nilpotency class at most two is Cayley graph automatic.

The proof uses special bases of nilpotent groups.

Baumslag-Solitar groups

For each $n, m \in \mathbb{N}$ the presentation of the Baumslag-Solitar group $B(m, n)$ is given by the following relation:

$$a^{-1} b^m a = b^n.$$

The groups $B(m, n)$ are not Thurston-automatic iff $m \neq n$.

Theorem

The groups $B(m, n)$ are Cayley automatic.

The case $B(1, n)$ is by Miasnikov/Khoussainov/Kharlampovich.

The general case is by Berdinsky and Khoussainov.

Closure properties of Cayley automatic groups

Theorem (Kharlampovich, Khoussainov, Miasnikov)

The class is closed under the following operations:

1. *Direct sum*
2. *Free product*
3. *Finite extensions*
4. *Amalgamated product*
5. *Semidirect product*
6. *The wreath-product of finite groups with the group $\mathbb{Z}$.*

Items (4) and (5) require natural regularity conditions.

Open Questions 4 and 5

Question 4:

Is the free group of nilpotency class $k \geq 3$ *Cayley automatic?*

Question 5:

Which wreath-products are Cayley automatic? In particular, is the wreath product of a finite group with Z^2 *Cayley automatic?*

Trees and Σ-trees

Definition

A tree is a finite subset X of $\{0,1\}^\star$ such that (1) X is closed under the prefix relation, and (2) For every $x \in X$ either no $y \in X$ properly extends x or both $x0$ and $x1$ belong to X.

Definition

A Σ-tree is a function $t : X \to \Sigma$ where X is a tree and Σ is a finite alphabet.

Let T_Σ be the set of all Σ-trees.

Definition

A Σ-tree language (or simply a language) is a subset of T_Σ.

Tree-automata

Definition

A **tree automaton** is a machine $\mathcal{M}$ with an initial state and accepting states whose transitions are of the form

$$< \textit{state}, \textit{symbol}, (\textit{state}, \textit{state}) > .$$

Runs of tree automata

Given a tree automaton $\mathcal{M}$ and a Σ-tree t, a run of $\mathcal{M}$ on the tree t is a function $r : dom(t) \to S$ such that:

1. The run starts with an initial state: $r(\lambda) \in I$.
2. The run is consistent with the transition table:

   For all internal nodes $x \in dom(t)$,
   if $r(x) = s$ and $t(x) = \sigma$ then $(r(x0), r(x1)) \in \delta(s, \sigma)$.

If $r(x) \in F$ for all leaves of $dom(t)$ then r is an accepting run.

Define $L(\mathcal{M}) = \{t \mid \text{the automaton } \mathcal{M} \text{ accepts } t\}$.

Calculus of tree-automata languages

Definition

A Σ-tree language L is regular if there is an automaton $\mathcal{M}$ such that L is the language of the automaton $\mathcal{M}$, that is, $L = L(\mathcal{M})$.

Theorem (Calculus)

The class of regular Σ-tree languages forms a Boolean algebra under the set-theoretic boolean operations.

Theorem (Deciding the emptiness problem)

There exists an algorithm that, given an automaton $\mathcal{M}$, decides if $\mathcal{M}$ accepts at least one Σ-tree.

Tree automata recognising n-ary relations

Just like for finite automata, we can define tree automata that read n-tuples of Σ-trees

$$(t_1, \ldots, t_n).$$

Such automata recognise n-ary relations on the set $T(\Sigma)$ of all Σ-trees.

Automatic structure: Tree automata case

Definition

A structure $\mathcal{A} = (A; R_0, R_1, \ldots, R_m)$ is **tree-automatic** over Σ if its domain A and all relations R_0, R_1, …, R_m are recognised by tree automata.

A **tree-automata presentable structure** is one isomorphic to a tree-automatic structure.

Examples of tree-automatic structures:

1. $(\omega; \times)$
2. The ordinals ω^{ω^n}.
3. The atomless Boolean algebra.

Some familiar ordinals:

$$1, 2, \ldots, \omega, \ \ \omega^2, \ldots, \ \omega^n, \ldots$$

Also:

$$\omega^\omega = 1 + \omega + \omega^2 + \omega^3 + \ldots$$

The ordinal ω^{ω^n} is the supremum of the sequence:

$$\omega^{\omega^{n-1}}, \ \ \omega^{\omega^{n-1}} \cdot \omega^{\omega^{n-1}}, \ \ \omega^{\omega^{n-1}} \cdot \omega^{\omega^{n-1}} \cdot \omega^{\omega^{n-1}}, \quad \ldots.$$

So,

$$\omega^{\omega^\omega} = \omega^\omega + \omega^{\omega^2} + \omega^{\omega^3} + \ldots + \omega^{\omega^n} + \ldots.$$

Ordinals less than $\omega^{\omega^{\omega}}$

For $\alpha < \omega^{\omega^{\omega}}$, there are polynomials $p_0(X)$, ..., $p_k(X)$ and integer coefficients c_0, ..., c_k with $c_k > 0$ such that

- $\alpha = \omega^{p_0(\omega)} c_0 + \omega^{p_1(\omega)} c_1 + \ldots + \omega^{p_{k-1}(\omega)} c_{k-1} + \omega^{p_k(\omega)} c_k$ and
- $p_0(\omega) > p_1(\omega) > \ldots > p_k(\omega)$.

When adding these types of ordinals, we use equalities:

$$\omega^{\alpha} m + \omega^{\alpha} n = \omega^{\alpha}(m+n), \text{ and } \omega^{\alpha} + \omega^{\beta} = \omega^{\beta},$$

where m,n are natural numbers and $\alpha < \beta$. For instance,

$$(\omega^{\omega^3}4 + \omega^{\omega^2}7 + \omega^6 3 + \omega^2 + 1) + (\omega^{\omega^2}2 + \omega^6 3 + \omega^5 + 5) =$$
$$= \omega^{\omega^3}4 + \omega^{\omega^2}9 + \omega^6 3 + \omega^5 + 5.$$

Tree-automatic ordinals with addition

Theorem (Jain, Khoussainov, Stephan (2018))

An ordinal structure $(\alpha; \leq, +)$ *is tree automatic iff* $\alpha < \omega^{\omega^\omega}$.

Proof is by induction on *n* showing that ω^{ω^n} with the addition operation is tree-automatic.

Algorithmic implication

Corollary

It is decidable if two tree automatic ordinals with the addition operation are isomorphic.

Proof. Let α be tree-automatic. Here are several facts:

1. Ordinal $\beta < \omega^{\omega^\omega}$ is closed under the addition operation $+$ iff β is a power of ω.
2. For ordinal α consider P_α:

   $$P_\alpha = \{\beta \mid \beta \text{ is closed under } +\}.$$

   The ordinal P_α is tree automatic and is less than ω^ω.
3. We can effectively compute Cantor normal form for P_α.

Proof: continuation

Now we can write α as

$$\alpha = \omega^{P_\alpha} + \alpha',$$

where $\alpha' < \omega^{P_\alpha}$. Continue this on we produce the Cantor normal form for α:

$$\alpha = \omega^{P_{\alpha_1}} + \omega^{P_{\alpha_2}} + \ldots + \omega^{P_{\alpha_m}}.$$

So, we produce the Cantor normal form for the ordinal α. This determine the isomorphism type of α.

Open Questions 6 and 7:

Question 6:

Is the isomorphism problem for tree-automatic ordinals decidable?

Question 7:

Is the isomorphism problem for tree-automatic Boolean algebras decidable?