Automatic Structures

Sasha Rubin

University of Auckland
Department of Computer Science

Dagstuhl, 30 October 2007

Finitely presentable structures

How can we compute with infinite structures?

We require the structure have a finite presentation.

Main formalisms:

- Explicit **internal** presentation of the elements, with machines describing the relations and functions (eg. computable structure).
- Structure is **logically interpretable** in a fixed structure (eg. tree-interpretable structure).
- Structure is a least solution of a system of equations in an **algebra** of structures. Similarly, **generating** structure by a deterministic grammar (eg. VR-equational graph).
- Structure is the **transformation** of another structure (eg. graph unfolding)

Outline of Talk

1. Background: S1S, S2S
2. Automatic presentations
   - Fundamental properties
   - Quotient problem
   - Characterisations and Isomorphism problem
3. Canonical presentations (automatic groups, automatic words)
4. Generalisation: Automata with oracles
5. Automatic Model Theory
6. Themes and Questions

ω-string automata

ω-string automata

A deterministic read-only finite-state machine that synchronously

- reads a tuple of infinite strings $(\alpha_1, \cdots, \alpha_k)$, $\alpha_i \in \Sigma^\omega$,
- and accepts if the set of states $\mathrm{Inf}(\rho)$ occuring infinitely often in the run ρ are in $\mathcal{F} = \{F_1, \cdots, F_k\}$ (Muller acceptance).

Reads input synchronously: So view $(\alpha_1, \cdots, \alpha_k) \in (\Sigma^k)^\omega$.

Examples.

- The set of strings $\alpha \in \{0, 1\}^\omega$ containing finitely many 1s.
- The set $\otimes(\alpha_1, \alpha_2) \in (\{0, 1\}^2)^\omega$ such that $\alpha_1 =_{\mathsf{ae}} \alpha_2$ (equal co-ordinate wise almost everywhere).

Languages accepted by ω-string automata are effectively closed under the operations of

- set union, complementation,
- projection, cylindrification, permutation of the co-ordinates.
- instantiation by ultimately-periodic string.

S1S Decidable

S1S is the monadic second-order theory of the structure $(\mathbb{N}, \mathrm{S})$.

Tuples of sets $(A_1, \cdots, A_k)$ of natural numbers encoded by their characteristic string $\otimes(A_1, \cdots, A_k) \in (\{0,1\}^k)^\omega$.

Theorem [Büchi 1962] Formulas $\longleftrightarrow$ Automata

$\mathrm{MSO}(\mathbb{N}, <)$ formulae $\Phi(X_1, \cdots, X_k)$ define the same relations, modulo encoding $\otimes$, as automata $\mathcal{A}(\alpha_1, \cdots, \alpha_k)$ operating on ω-strings; and translations are effective.

Thus S1S reduces to the emptiness problem for ω-automata:

> *Given an ω-automaton as input, does it accept any string whatsoever?*

FO theories

Büchi noticed that certain FO theories are decidable via this technique:

- Code $n \in \mathbb{N}$ by its base 2 representation (lsd first).
- $\mathbb{N}$ corresponds to the regular language $\{0,1\}^\star 0^\omega$.
- The atomic ternary-relation $+$ corresponds to a regular relation $+_2$ over this coding.

Every FO-formula $\phi(\overline{x})$ of $(\{0,1\}^\star 0^\omega, +_2)$ defines a regular relation.

Therefore FO$(\mathbb{N}, +)$ is decidable.

S2S Decidable

S2S is the monadic second-order theory of the structure $(\{0,1\}^\star, \mathrm{s}_0, \mathrm{s}_1)$.

Tuples of sets $(A_1, \cdots, A_k)$ of strings are encoded as a $\{0,1\}^k$-labelled infinite tree $\otimes(A_1, \cdots, A_k) : \{l, r\}^\star \to \{0,1\}^k$.

Theorem [Rabin 1969] Formulas $\longleftrightarrow$ Automata

$\mathrm{MSO}(\{0,1\}^\star, \mathrm{s}_0, \mathrm{s}_1)$ formulae $\Phi(X_1, \cdots, X_k)$ define the same relations, modulo encoding $\otimes$, as automata $\mathcal{A}(T_1, \cdots, T_k)$ operating on ω-trees; and translations are effective.

Corollary The following theories are decidable (via interpretation):

- $\mathrm{MSO}(\mathbb{Q}, <)$ decidable.
- MSOTh of countable linearly ordered sets.
- MSOTh of countable well-ordered sets.

Summary of definability

Fact. For each type of object $\diamondsuit \in \{\text{string}, \omega\text{-string}, \text{tree}, \omega\text{-tree}\}$ there is a notion of synchronous automaton with robust closure properties.

- WMSO$(\mathbb{N}, \mathrm{S}) \leftrightarrows$ automata on finite words
- WMSO$(\{0,1\}^\star, \mathrm{s}_0, \mathrm{s}_1) \leftrightarrows$ automata on finite trees
- MSO$(\mathbb{N}, \mathrm{S}) \leftrightarrows$ automata on infinite words
- MSO$(\{0,1\}^\star, \mathrm{s}_0, \mathrm{s}_1) \leftrightarrows$ automata on infinite trees

Weak: variables range over finite subsets of domain.

Automatic Presentations

Let $\diamondsuit \in \{\mathsf{string}, \omega\text{-}\mathsf{string}, \mathsf{tree}, \omega\text{-}\mathsf{tree}\}$.

A $\diamondsuit$-automatic presentation of a relational structure $\mathcal{D} = (D, (R_i))$ consists of

1. a tuple of $\diamondsuit$-automata $(\mathsf{M_A}, (\mathsf{M_i}))$,
2. a bijection $\mu : \mathcal{L}(\mathsf{M_A}) \to D$,

so that

$$(\mathcal{L}(\mathsf{M_A}), (\mathcal{L}(\mathsf{M_i}))) \overset{\mu}{\cong} \mathcal{D}.$$

Say that $\mathcal{D}$ is an $\diamondsuit$-automatic structure.

[Hodgson 76] [Khoussainov,Nerode 95] [Blumensath, Grädel 00]

Relationships

$$
\begin{tikzcd}
\omega\texttt{S-AutStr} \arrow[r] & \omega\texttt{T-AutStr} \\
\texttt{S-AutStr} \arrow[r] \arrow[u] & \texttt{T-AutStr} \arrow[u]
\end{tikzcd}
$$

- Structures in `S-AutStr`: $(\mathbb{N}, +)$, ordinals $(<)$ below ω^ω.
- Structures in `T-AutStr` but not in `S-AutStr`: The countable atomless Boolean algebra, $(\mathbb{N}, \times)$, $(\mathbb{Q}, \times)$.
- Uncountable structures in ω`S-AutStr`: $(\mathbb{R}, +)$, $(\mathcal{P}(\mathbb{N}), \cup, \cap, {}^c)$.

Fundamental Properties

Theorem [FO definable $\rightarrow$ regular]
Given an $\diamondsuit$-automatic presentation μ of $\mathcal{D}$, every FO-formula $\Phi(\overline{x})$ defines a $\diamondsuit$-regular relation $\mu^{-1}(\Phi^{\mathcal{D}})$ (and the translation is effective).

Corollary [FO decidability] The following problem is decidable:
Input: The automata forming an automatic presentation of some structure $\mathcal{A}$, and a FO-sentence σ.
Output: Whether or not $\mathcal{A} \models \sigma$.

Parameters: No problem in the finite cases. In the ω-cases, as long as they are ultimately-periodic strings / regular trees.

Fundamental Properties

A *k*-dimensional FO-interpretation is a collection of FO-formulas

- Domain formula $\Delta(\overline{x})$,
- Relation formulas $\Phi_i(\overline{x_1}, \cdots, \overline{x_{r_i}})$.

The interpretation applied to a structure $\mathcal{A}$ produces the structure

$$(\Delta^{\mathcal{A}}, (\Phi_i^{\mathcal{A}})_i).$$

Corollary [FO Interpretability] The class $\Diamond$-`AutStr` is closed under FO-interpretations.

Corollary

1. $\Diamond$-`AutStr` is closed under direct-products (ordered products), ω-fold disjoint unions, definable substructures, definable extensions.
2. ω`T-AutStr` (`T-AutStr`) are closed under (weak) ω-fold direct powers.

The Quotient problem

Let $\mathcal{A}$ be $\diamondsuit$-automatic and let ϵ be a $\diamondsuit$-regular congruence on $\mathcal{A}$.

Then the FO theory of $\mathcal{A}/\epsilon$ is decidable. However,

Is the quotient structure $\diamondsuit$-automatic?

`S-AutStr`: define a regular ϵ set of unique representatives by selecting length-lex least string from each class [Blumensath 99].

Eventhough there is no regular well-order on the set of all finite trees [Shelah,Gurevich] [Carayol,Löding 07]:

`T-AutStr` is closed under quotient [Colcombet,Löding 07].

The Quotient problem - ωS-AutStr

Given: $\mathcal{A} \in \omega\texttt{S-AutStr}$ and ϵ a ω-string regular congruence.

- There is no regular set of unique representatives of the equal almost-everywhere $=_{\mathsf{ae}}$ relation [Kuske, Lohrey 06].
- If ϵ has countable index then it has a regular set of unique representatives [Bárány, Kaiser, R 07].

Corollary If $\mathcal{A}/\epsilon$ is countable, then it is in $\omega\texttt{S-AutStr}$, and hence in $\texttt{S-AutStr}$.

Moreover, there exists $(\mathcal{A}, \epsilon) \in \omega\texttt{S-AutStr}$ such that $\mathcal{A}/\epsilon \notin \omega\texttt{S-AutStr}$ (Hjorth, Khoussainov, Montalban, Nies 07).

The Quotient problem

Let $\mathcal{A}$ be $\diamondsuit$-automatic and let ϵ a $\diamondsuit$-regular congruence on $\mathcal{A}$.

Is the quotient structure $\diamondsuit$-automatic?

$\diamondsuit$-AutStr	reg well order	ϵ reg uniq. rep.	Quotient
S-AutStr	Yes	Yes	Yes
T-AutStr	No	?	Yes
ωS-AutStr	No	index $\aleph_0$	index $\aleph_0$
ωT-AutStr	No	?	?

Decidable Extensions of FO

Recall Fundamental Property: FO-definability → regularity

It can be improved by extending FO as in the following cases:

$$\diamondsuit\text{-}\mathtt{AutStr} = \begin{cases} \mathtt{S\text{-}AutStr} & FO + \exists^{\infty} + \exists^{\text{mod}} \\ \mathtt{T\text{-}AutStr} & FO + \exists^{\infty} + \exists^{\text{mod}} \\ \omega\mathtt{S\text{-}AutStr} & FO + \exists^{\infty} + \exists^{\text{mod}} + \exists^{\leq \aleph_0} + \exists^{> \aleph_0} \\ \omega\mathtt{T\text{-}AutStr} & FO + \exists^{\infty} + \exists^{\text{mod}} + \exists^{\leq \aleph_0} + \exists^{> \aleph_0} \end{cases}$$

Approach: Quantifier Q preserves regularity :if $Q\overline{x}\,\Phi(\overline{x}, \overline{y})$ defines regular relation in $\mu^{-1}(\mathcal{D})$.

Unary quantifiers: Exactly the ones listed above.

$\mathtt{S\text{-}AutStr}$: [Blumensath 99], [Khoussainov, R, Stephan 03]
$\mathtt{T\text{-}AutStr}$: [Colcombet 04]
$\omega\mathtt{S\text{-}AutStr}$: [Kuske, Lohrey 06]
$\omega\mathtt{T\text{-}AutStr}$: [Kaiser, Bárány, Rabinovich, R 07]
$(\mathcal{A}, \epsilon) \in \omega\mathtt{S\text{-}AutStr}$: [Kaiser, Bárány, R 07]

Undecidable Extensions

The following extensions of FO are undecidable on automatic structures.

- (W)MSO
  eg. The grid $(\mathbb{N} \times \mathbb{N}, \mathrm{up}, \mathrm{right})$ is in `S-AutStr`.
- FO $+$ certain binary generalised quantifiers (Reachability, ...)
  The one-step configuration graph of an arbitrary Turing machine is in `S-AutStr`.

Characterisation via set interpretations

Recall that $(\mathbb{N}, +)$ is in `S-AutStr`:
Interpret numbers as finite $0, 1$-strings so that $+_2$ is regular.
Interpret numbers as finite subsets of $\mathbb{N}$ so that $+_2$ is WMSO definable in $(\mathbb{N}, \mathrm{S})$.

(finite) set interpretation $(\Delta(X), \Phi_i(X_1, \cdots, X_{r_i}), \epsilon(X_1, X_2))$

- Δ, Φ_i, ϵ are (weak) monadic formulas.
- X_{i_j} are (weak) monadic variables.

(Interpretations can be of dimension k).

Characterisation via set interpretations

Theorem [Set Interpretations] A structure has a $\Diamond$-presentation if and only if it is interpretable in:

$$\Diamond = \begin{cases} \texttt{string} & \text{finite-set interpretable in } (\mathbb{N}, \mathrm{S}) \\ \omega\texttt{-string} & \text{set interpretable in } (\mathbb{N}, \mathrm{S}) \\ \texttt{tree} & \text{finite-set interpretable in } (\{0,1\}^\star, \mathrm{s}_0, \mathrm{s}_1) \\ \omega\texttt{-tree} & \text{set interpretable in } (\{0,1\}^\star, \mathrm{s}_0, \mathrm{s}_1) \end{cases}$$

Characterisation via FO interpretations

Theorem [FO Interpretations] A structure has a $\Diamond$-presentation if and only if it is FO-interpretable as follows:

$$\Diamond = \begin{cases} \texttt{string} & (\{0,1\}^\star, \preceq_{\texttt{prefix}}, =_{\text{len}}, \mathsf{suc}_0, \mathsf{suc}_1) \\ \omega\texttt{-string} & (\{0,1\}^\omega, \preceq_{\texttt{prefix}}, =_{\text{len}}, \mathsf{suc}_0, \mathsf{suc}_1) \\ \texttt{tree} & (\text{fintrees}, \preceq_{\texttt{ext}}, \equiv_{\mathsf{dom}}, (\mathsf{suc}_a^d)_{d \in \{l,r\}, a \in \{0,1\}}) \\ \omega\texttt{-tree} & (\omega\text{-trees}, \preceq_{\texttt{ext}}, \equiv_{\mathsf{dom}}, (\mathsf{suc}_a^d)_{d \in \{l,r\}, a \in \{0,1\}}) \end{cases}$$

Comments.

- ω cases: $u \in \{0,1\}^\star$ is identified with $u10^\omega$, and similary for trees.
- string: Can take $(\mathbb{N}, +, |_k)$ for fixed $k \in \mathbb{N}^+$, where $n|_k m$ if n is a power of k and n divides m.
- ω-string: Can take $(\mathbb{R}, +, <, |_k, 1)$ where $x|_k y$ if $\exists n, m \in \mathbb{Z}$: $x = k^n$ and $y = xm$.

Characterisations of classes: `S-AutStr`

Full characterisations

- Ordinals $(L, <)$: those below ω^ω (Delhommè 01)
- Boolean algebras: $(B_{\mathrm{fin/co\text{-}fin}})^n$ (KNRS 04)
- Finitely generated groups $(G, +)$: virtually Abelian (Oliver, Thomas 05)
- Fields $(F, +, \times, 0, 1)$: finite (Stephan 04).

Partial characterisations

- Linear orders $(L, <)$ and trees $(T, \prec)$ have finite rank (KRS03)

Individual structures not in `S-AutStr`

- Random Graph (Delhommè 01) (Stephan 02)
- $(\mathbb{N}, \times)$, $(\mathbb{N}, \mathrm{pairing})$ (Blumensath 99)
- $(\mathbb{Q}, \times)$, $\oplus_k \mathbb{Z}[\frac{1}{k}]$, countable atomless BA (KNRS 04)

Characterisations of classes - `T-AutStr`

Delhommè 04 gives methods to prove that certain structures are not in `T-AutStr`. In particular,

- The least ordinal not in `T-AutStr` is $\omega^{\omega^{\omega}}$.
- The random graph is not in `T-AutStr`.

Theorem [Colcombet, Löding 06] If $\mathcal{P}^f(\mathcal{S})$ is finite-set interpretable in a tree t, then $\mathcal{S}$ is WMSO interpretable in t.

Here $\mathcal{P}^f(\mathcal{S})$ is the structure

- domain consists of the finite subsets of S,
- $\subset$ order,
- relations of $\mathcal{S}$ restricted to singletons.

So, the following are not finite-set interpretable in any tree:

- The free monoid $(\{a, b\}^\star, \cdot, a, b)$
- The random graph

Needed: Techniques for ω`S-AutStr` and ω`T-AutStr`.

Isomorphism problem

Fix a class $\mathfrak{C}$ of structures closed under isomorphism (graphs, linear orders, ...)

The isomorphism problem for $\mathfrak{C} \cap \diamondsuit$-`AutStr` is

> *Given automata from $\diamondsuit$-automatic presentations of two structures $\mathcal{A}, \mathcal{B} \in \mathfrak{C}$, is $\mathcal{A}$ isomorphic to $\mathcal{B}$?*

The complexity of the isomorphism problem for `S-AutStr`:

- Decidable: ordinals, Boolean algebras, fields.
- Π^0_1: equivalence structures (decidable?)
- Π^0_3-complete: locally finite graphs.
- Σ^1_1-complete: graphs, sucessor trees, ...

Σ^1_1-completeness (KNR, K Minnes): Reduction from the isomorphism problem for computable structures.

Canonical Presentations - groups

Automatic Groups (Cannon 84, Thurston 86)

A finitely generated group, say with semigroup generators S, is automatic if

$$(D, \sigma_1, \cdots, \sigma_{|S|}, =)$$

is an automatic presentation of its Cayley graph, where $D \subset S^\star$.

Facts

- Algebraic notion: independent of generators.
- Such groups are finitely-presentable.

Canonical Presentations - sequences and numerations

Automatic sequences $(\mathbb{N}, <, C_1, \cdots, C_m)$

- Morphic words $\equiv$ string-automatic $(D, <_{\text{llex}}, \overline{C})$
  eg. Thue-Morse
- k-llex words $(D, <_{k\text{-llex}}, \overline{C})$
  eg. Champernowne 01234567891011121314...

Theorem [Bárány 2006] MSO definable relations on $<_{k\text{-llex}}$ words define synchronous regular relations.

Automatic numeration systems $(\mathbb{N}, +, <)$ (eg. -2, fibonacci)

Question
Is every finite-string automatic presentation of $(\mathbb{N}, +, <)$ equivalent to one in which $<$ is length reverse-lexicographic order?

Problem Is $(\mathbb{Q}, +)$ $\diamondsuit$-automatic ?

Generalisations via set interpretations

Fact. If $\mathcal{A}$ is set-interpretable in a structure with decidable MSO, then $\mathsf{FO}(\mathcal{A})$ is decidable.

- Natural extensions of $(\mathbb{N}, \mathrm{S})$ by functions are usually undecidable (eg. $n \mapsto 2n$).
- Many decidable extensions by natural unary predicates (eg. $\{n!\}$, $\{2^n\}$, $\{n^2\}$, $<_{k\text{-llex}}$, pushdown hierarchy).

Example.

- $(\mathbb{Q}, +)$ is finite-set interpretable in $(\mathbb{N}, \mathrm{S}, P)$ where P is the word $1\#2\#3\#4\cdots$ (Miller, Stephan).

Model Theory on S-AutStr

Negative results.

- Compactness, Beth-property, interpolation, Łoś-Tarski fail on the class S-AutStr (Blumensath).

Positive results.

- Automatic KHonig's Lemma.
  1. Every finitely-branching infinite tree $(T, \prec) \in$ S-AutStr can be expanded to $(T, \prec, P) \in$ S-AutStr where P is an infinite path.
  2. If $(T, \prec)$ has countably many infinite paths (not nec. fb.), this can be done for every path [KRS 03].
- Automatic Ramsey's Theorem. Every presentation of an infinite graph $(G, E) \in$ S-AutStr can be expanded to $(G, E, H) \in$ S-AutStr where $H \subset G$ is infinite and either a clique or an independent set [KRS 03].
- Cantor's Theorem. Every presentation of a linear order $(L, <) \in$ S-AutStr can be extended to one of $(\mathbb{Q}, <)$ so that the embedding is continuous and regular [Kuske 03].

Model Theory on `AutStr`

- Downward Lowenheim-Skolem. Every ω`T-AutStr` (ω`S-AutStr`) structure has a countable elementarily-equivalent substructure consisting of the regular trees (u.p. strings) of the original presentation.

In particular, we get an automata theoretic proof that FO($\mathbb{Q}, +$) is decidable.

Summary of Themes

1. Automatic structures have decidable FO[...] theories.
2. Some classes are easy (eg. automatic ordinals), others are complicated (eg. automatic graphs).
3. Automatic structures can be characterised in a number of ways (internally, logically, equationally).
4. For structures in a given class (f.g. groups, numeration systems, ...) canonical presentations are appropriate.

Some Questions

Specific

1. For $\mathcal{C}$ the class of linear orders, equivalence relations, f.g groups ...
   - Is the isomorphism problem for $\mathcal{C} \cap \texttt{S-AutStr}$ decidable?
   - Describe isomorphism types of $\mathcal{C} \cap \texttt{S-AutStr}$.
2. Is $\omega\texttt{T-AutStr}$ closed by ωT-regular quotient?
3. If $(\mathcal{A}, \epsilon) \in \omega\texttt{T-AutStr}$ and $\mathcal{A}/\epsilon$ is countable, is $\mathcal{A}/\epsilon \in \texttt{T-AutStr}$?
4. Which binary quantifiers preserve regularity?

General

1. Identify useful canonical automatic-presentations of other classes of structures.
2. Develop model theory on the class of $\diamondsuit$-automatic structures. What is a natural logic?
3. Isolate well behaved subclasses of $\texttt{AutStr}$: decidable MSO, working model-theory.