The Isomorphism Problem for ω-Automatic Trees

Dietrich Kuske[1], Jiamou Liu[2], and Markus Lohrey[2, ⋆]

[1] Laboratoire Bordelais de Recherche en Informatique (LaBRI), CNRS and Université Bordeaux I, Bordeaux, France
[2] Universität Leipzig, Institut für Informatik, Germany
kuske@labri.fr, liujiamou@gmail.com,
lohrey@informatik.uni-leipzig.de

Abstract. The main result of this paper is that the isomorphism for ω-automatic trees of finite height is at least has hard as second-order arithmetic and therefore not analytical. This strengthens a recent result by Hjorth, Khoussainov, Montalbán, and Nies [9] showing that the isomorphism problem for ω-automatic structures is not Σ^1_2. Moreover, assuming the continuum hypothesis **CH**, we can show that the isomorphism problem for ω-automatic trees of finite height is recursively equivalent with second-order arithmetic. On the way to our main results, we show lower and upper bounds for the isomorphism problem for ω-automatic trees of every finite height: (i) It is decidable (Π^0_1-complete, resp,) for height 1 (2, resp.), (ii) Π^1_1-hard and in Π^1_2 for height 3, and (iii) Π^1_{n-3}- and Σ^1_{n-3}-hard and in Π^1_{2n-4} (assuming **CH**) for all $n \geq 4$. All proofs are elementary and do not rely on theorems from set theory.

See [19] for a full version of this extended abstract.

1 Introduction

A graph is computable if its domain is a computable set of natural numbers and the edge relation is computable as well. Hence, one can compute effectively in the graph. On the other hand, practically all other properties are undecidable for computable graphs (e.g., reachability, connectedness, and even the existence of isolated nodes). In particular, the isomorphism problem is highly undecidable in the sense that it is complete for Σ^1_1 (the first existential level of the analytical hierarchy [22]); see e.g. [3, 8] for further investigations of the isomorphism problem for computable structures. These algorithmic deficiencies have motivated in computer science the study of more restricted classes of finitely presented infinite graphs. For instance, pushdown graphs, equational graphs, and prefix recognizable graphs have a decidable monadic second-order theory and for the former two the isomorphism problem is known to be decidable [5] (for prefix recognizable graphs the status of the isomorphism problem seems to be open).

Automatic graphs [14] are between prefix recognizable and computable graphs. In essence, a graph is automatic if the elements of the universe can be represented as strings from a regular language and the edge relation can be recognized by a finite state automaton with several heads that proceed synchronously. Automatic graphs (and more

⋆ The second and third author are supported by the DFG research project GELO.

general, automatic structures) received increasing interest over the last years [2, 11, 15, 16, 25]. One of the main motivations for investigating automatic graphs is that their first-order theories can be decided uniformly (i.e., the input is an automatic presentation and a first-order sentence). On the other hand, the isomorphism problem for automatic graphs is Σ_1^1-complete [15] and hence as complex as for computable graphs.

In our recent paper [18], we studied the isomorphism problem for restricted classes of automatic graphs. Among other results, we proved that (i) the isomorphism problem for automatic trees of height at most $n \geq 2$ is complete for the level Π_{2n-3}^0 of the arithmetical hierarchy and (ii) that the isomorphism problem for automatic trees of finite height is recursively equivalent to true arithmetic. In this paper, we extend our techniques from [18] to *ω-automatic trees*. The class of ω-automatic structures was introduced in [1]. It generalizes automatic structures by replacing ordinary finite automata by Büchi automata on ω-words. In this way, uncountable graphs can be specified. Some recent results on ω-automatic structures can be found in [9, 12, 17, 20]. On the logical side, many of the positive results for automatic structures carry over to ω-automatic structures [1, 12]. On the other hand, the isomorphism problem of ω-automatic structures is more complicated than that of automatic structures (which is Σ_1^1-complete). Hjorth et al. [9] constructed two ω-automatic structures for which the existence of an isomorphism depends on the axioms of set theory. Using Schoenfield's absoluteness theorem, they infer that isomorphism of ω-automatic structures does not belong to Σ_2^1. Also using Shoenfield's absoluteness theorem, Finkel and Todorčević [7] recently showed that isomorphism of ω-tree-automatic partial orders[3] (resp. Boolean algebras, rings, non-commutative groups) is not in the class Σ_2^1.

The extension of our elementary techniques from [18] to ω-automatic trees allows us to show directly (without a "detour" through set theory) that the isomorphism problem for ω-automatic trees of finite height is not analytical (i.e., does not belong to any of the levels Σ_n^1). For this, we prove that the isomorphism problem for ω-automatic trees of height $n \geq 4$ is hard for both levels Σ_{n-3}^1 and Π_{n-3}^1 of the analytical hierarchy (our proof is uniform in n). A more precise analysis reveals at which height the complexity jump for ω-automatic trees occurs: For automatic as well as for ω-automatic trees of height 2, the isomorphism problem is Π_1^0-complete and hence arithmetical. But the isomorphism problem for ω-automatic trees of height 3 is hard for Π_1^1 (and therefore outside of the arithmetical hierarchy) while the isomorphism problem for automatic trees of height 3 is Π_3^0-complete [18].

We prove our results by reductions from monadic second-order (fragments of) number theory. The first step in the proof is a normal form for analytical predicates. The basic idea of the reduction then is that a subset $X \subseteq \mathbb{N}$ can be encoded by an ω-word w_X over $\{0, 1\}$, where the i-th symbol is 1 if and only if $i \in X$. The combination of this basic observation with our techniques from [18] allows us to encode monadic second-order formulas over $(\mathbb{N}, +, \times)$ by ω-automatic trees of finite height. This yields the lower bounds mentioned above. We also give an upper bound for the isomorphism problem: for ω-automatic trees of height n, the isomorphism problem belongs to Π_{2n-4}^1. While the lower bound holds in the usual system **ZFC** of set theory, we can prove the

[3] An ω-tree-automatic structures is a structure whose elements are coded as infinite trees and whose universe and relations are accepted by Muller or Rabin tree automata.

upper bound only assuming in addition the continuum hypothesis. The precise recursion theoretic complexity of the isomorphism problem for ω-automatic trees remains open, it might depend on the underlying axioms for set theory.

Related work Results on isomorphism problems for various subclasses of automatic structures can be found in [15, 16, 18, 24]. Some completeness results for low levels of the analytical hierarchy for decision problems on infinitary rational relations were shown in [6].

2 Preliminaries

Let $\mathbb{N}_+ = \{1, 2, 3, \ldots\}$. With $\overline{x}$ we denote a tuple $(x_1, \ldots, x_m)$ of variables, whose length m does not matter.

2.1 The arithmetical and analytical hierarchy

In this paper we follow the definitions of the arithmetical and analytical hierarchy from [22, p. 367]. In order to avoid some technical complications, it is useful to exclude 0 in the following, i.e., to consider subsets of $\mathbb{N}_+$. In the following, f_i ranges over unary functions on $\mathbb{N}_+$, X_i over subsets of $\mathbb{N}_+$, and $u, x, y, z, x_i, \ldots$ over elements of $\mathbb{N}_+$. The class $\Sigma^0_n \subseteq 2^{\mathbb{N}_+}$ is the collection of all sets $A \subseteq \mathbb{N}_+$ of the form

$$A = \{x \in \mathbb{N}_+ \mid (\mathbb{N}, +, \times) \models \exists y_1 \, \forall y_2 \cdots Q y_n : \varphi(x, y_1, \ldots, y_n)\},$$

where $Q = \forall$ (resp. $Q = \exists$) if n is even (resp. odd) and φ is a quantifier-free formula over the signature containing $+$ and $\times$. The class Π^0_n is the class of all complements of Σ^0_n sets. The classes Σ^0_n, Π^0_n ($n \geq 1$) make up the *arithmetical hierarchy*.

The analytical hierarchy extends the arithmetical hierarchy and is defined analogously using function quantifiers: The class $\Sigma^1_n \subseteq 2^{\mathbb{N}_+}$ is the collection of all sets $A \subseteq \mathbb{N}_+$ of the form

$$A = \{x \in \mathbb{N}_+ \mid (\mathbb{N}, +, \times) \models \exists f_1 \, \forall f_2 \cdots Q f_n : \varphi(x, f_1, \ldots, f_n)\},$$

where $Q = \forall$ (resp. $Q = \exists$) if n is even (resp. odd) and φ is a first-order formula over the signature containing $+$, $\times$, and the functions $f_1, \ldots, f_n$. The class Π^1_n is the class of all complements of Σ^1_n sets. The classes Σ^1_n, Π^1_n ($n \geq 1$) make up the *analytical hierarchy*. The class of *analytical sets*[4] is exactly $\bigcup_{n \geq 1} \Sigma^1_n \cup \Pi^1_n$.

As usual in computability theory, a Gödel numbering of all finite objects of interest allows to quantify over, say, finite automata as well. We will always assume such a numbering without mentioning it explicitly.

[4] Here the notion of *analytical sets* is defined for sets of natural numbers and is not to be confused with the *analytic sets* studied in descriptive set theory [13].

2.2 Büchi automata

For details on Büchi automata, see [23]. Let Γ be a finite alphabet. With Γ^* we denote the set of all finite words over the alphabet Γ. The set of all nonempty finite words is Γ^+. An ω-word over Γ is an infinite sequence $w = a_1 a_2 a_3 \cdots$ with $a_i \in \Gamma$. We set $w[i] = a_i$ for $i \in \mathbb{N}_+$. The set of all ω-words over Γ is denoted by Γ^ω.

A (nondeterministic) Büchi automaton is a tuple $M = (Q, \Gamma, \Delta, I, F)$, where Q is a finite set of states, $I, F \subseteq Q$ are resp. the sets of initial and final states, and $\Delta \subseteq Q \times \Gamma \times Q$ is the transition relation. If $\Gamma = \Sigma^n$ for some alphabet Σ, then we refer to M as an *n-dimensional Büchi automaton over Σ*. A *run* of M on an ω-word $w = a_1 a_2 a_3 \cdots$ is an ω-word $r = (q_1, a_1, q_2)(q_2, a_2, q_3)(q_3, a_3, q_4) \cdots \in \Delta^\omega$ such that $q_1 \in I$. The run r is *accepting* if there exists a final state from F that occurs infinitely often in r. The language $L(M) \subseteq \Gamma^\omega$ defined by M is the set of all ω-words for which there exists an accepting run. An ω-language $L \subseteq \Gamma^\omega$ is *regular* if there exists a Büchi automaton M with $L(M) = L$. The class of all regular ω-languages is effectively closed under Boolean operations and projections.

For ω-words $w_1, \ldots, w_n \in \Gamma^\omega$, the *convolution* $w_1 \otimes w_2 \otimes \cdots \otimes w_n \in (\Gamma^n)^\omega$ is defined by

$$w_1 \otimes w_2 \otimes \cdots \otimes w_n = (w_1[1], \ldots, w_n[1])(w_1[2], \ldots, w_n[2])(w_1[3], \ldots, w_n[3]) \cdots .$$

For $\overline{w} = (w_1, \ldots, w_n)$, we write $\otimes(\overline{w})$ for $w_1 \otimes \cdots \otimes w_n$.

An n-ary relation $R \subseteq (\Gamma^\omega)^n$ is called *ω-automatic* if the language $\otimes R = \{\otimes(\overline{w}) \mid \overline{w} \in R\}$ is a regular ω-language, i.e., it is accepted by some n-dimensional Büchi automaton. We denote with $R(M) \subseteq (\Gamma^\omega)^n$ the relation defined by the n-dimensional Büchi automaton M over the alphabet Γ.

To also define the convolution of finite words (and of finite words with infinite words), we identify a finite word $u \in \Gamma^*$ with the ω-word $u\diamond^\omega$, where $\diamond$ is a new symbol. Then, for $u, v \in \Gamma^*, w \in \Gamma^\omega$, we write $u \otimes v$ for the ω-word $u \diamond^\omega \otimes v\diamond^\omega$ and $u \otimes w$ (resp. $w \otimes u$) for $u \diamond^\omega \otimes w$ (resp. $w \otimes u\diamond^\omega$).

2.3 ω-automatic structures

A *signature* is a finite set τ of relational symbols together with an arity $n_S \in \mathbb{N}_+$ for every relational symbol $S \in \tau$. A *τ-structure* is a tuple $\mathcal{A} = (A, (S^{\mathcal{A}})_{S \in \tau})$, where A is a set (the *universe* of $\mathcal{A}$) and $S^{\mathcal{A}} \subseteq A^{n_S}$. When the context is clear, we simply denote $S^{\mathcal{A}}$ by S, and we write $a \in \mathcal{A}$ for $a \in A$. Let $E \subseteq A^2$ be an equivalence relation on A. Then E is a *congruence* on $\mathcal{A}$ if $(u_1, v_1), \ldots, (u_{n_S}, v_{n_S}) \in E$ and $(u_1, \ldots, u_{n_S}) \in S$ imply $(v_1, \ldots, v_{n_S}) \in S$ for all $S \in \tau$. Then the *quotient structure* $\mathcal{A}/E$ can be defined:

- The universe of $\mathcal{A}/E$ is the set of all E-equivalence classes $[u]$ for $u \in A$.
- The interpretation of $S \in \tau$ is the relation $\{([u_1], \ldots, [u_{n_S}]) \mid (u_1, \ldots, u_{n_S}) \in S\}$.

Definition 2.1. *An* ω-automatic presentation *over the signature τ is a tuple*

$$P = (\Gamma, M, M_\equiv, (M_S)_{S \in \tau})$$

with the following properties:

– Γ *is a finite alphabet*
– M *is a Büchi automaton over the alphabet* Γ.
– *For every* $S \in \tau$, M_S *is an* n_S*-dimensional Büchi automaton over the alphabet* Γ.
– $M_{\equiv}$ *is a 2-dimensional Büchi automaton over the alphabet* Γ *such that* $R(M_{\equiv})$ *is a congruence relation on* $(L(M), (R(M_S))_{S\in\tau})$.

The τ-structure defined *by the* ω*-automatic presentation* P *is the quotient structure*

$$\mathcal{S}(P) = (L(M), (R(M_S))_{S\in\tau})/R(M_{\equiv})\,.$$

If $R(M_{\equiv})$ is the identity relation on Γ^ω, then P is called *injective*. A structure $\mathcal{A}$ is *(injectively)* ω*-automatic* if there is an (injectively) ω-automatic presentation P with $\mathcal{A} \cong \mathcal{S}(P)$. There exist ω-automatic structures that are not injectively ω-automatic [9]. We simplify our statements by saying "given/compute an (injectively) ω-automatic structure $\mathcal{A}$" for "given/compute an (injectively) ω-automatic presentation P of a structure $\mathcal{S}(P) \cong \mathcal{A}$". *Automatic structures* [14] are defined analogously to ω-automatic structures, but instead of Büchi automata ordinary finite automata over finite words are used. For this, one has to pad shorter strings with the padding symbol $\diamond$ when defining the convolution of finite strings. More details on ω-automatic structures can be found in [2,9,12].

Let $\mathsf{FO}[\exists^{\aleph_0}, \exists^{2^{\aleph_0}}]$ be first-order logic extended by the quantifiers $\exists^\kappa x \ldots$ ($\kappa \in \{\aleph_0, 2^{\aleph_0}\}$) saying that there exist exactly κ many x satisfying …. The following theorem lays out the main motivation for investigating ω-automatic structures.

Theorem 2.2 ([1,12]). *From an* ω*-automatic presentation* $P = (\Gamma, M, M_{\equiv}, (M_S)_{S\in\tau})$ *and a formula* $\varphi(\overline{x}) \in \mathsf{FO}[\exists^{\aleph_0}, \exists^{2^{\aleph_0}}]$ *in the signature* τ *with* n *free variables, one can compute a Büchi automaton for the relation*

$$\{\overline{a} \in L(M)^n \mid \mathcal{S}(P) \models \varphi([a_1], [a_2], \ldots, [a_n])\}\,.$$

In particular, the $\mathsf{FO}[\exists^{\aleph_0}, \exists^{2^{\aleph_0}}]$ *theory of any* ω*-automatic structure* $\mathcal{A}$ *is (uniformly) decidable.*

Definition 2.3. *Let* $\mathcal{K}$ *be a class of* ω*-automatic presentations. The* isomorphism problem $\mathsf{Iso}(\mathcal{K})$ *is the set of pairs* $(P_1, P_2) \in \mathcal{K}^2$ *of* ω*-automatic presentations from* $\mathcal{K}$ *with* $\mathcal{S}(P_1) \cong \mathcal{S}(P_2)$.

If $\mathcal{S}_1$ and $\mathcal{S}_2$ are two structures over the same signature, we write $\mathcal{S}_1 \uplus \mathcal{S}_2$ for the disjoint union of the two structures. We use $\mathcal{S}^\kappa$ to denote the disjoint union of κ many copies of the structure $\mathcal{S}$ (where κ is any cardinal).

2.4 Trees and dags

A *forest* is a partial order $F = (V, \leq)$ such that for every $x \in V$, the set $\{y \mid y \leq x\}$ of ancestors of x is finite and linearly ordered by $\leq$. The *level* of a node $x \in V$ is $|\{y \mid y < x\}| \in \mathbb{N}$. The *height* of F is the supremum of the levels of all nodes in V; it may be infinite, but this paper deals with forests of finite height only. For all $u \in V$,

$F(u)$ denotes the restriction of F to the set $\{v \in V \mid u \leq v\}$ of successors of u. We will speak of the *subtree rooted at* u. A *tree* is a forest that has a minimal element, called the *root*. For a forest F and r not belonging to the domain of F, we denote with $r \circ F$ the tree that results from adding r to F as a new root. The *edge relation* E of the forest F is the set of pairs $(u, v) \in V^2$ such that u is the largest element in $\{x \mid x < v\}$. For any node $u \in V$, we use $E(u)$ to denote the set of children (or immediate successors) of u.

We use $\mathcal{T}_n$ (resp. $\mathcal{T}_n^{\mathrm{i}}$) to denote the class of (injectively) ω-automatic presentations of trees of height at most n. Note that it is decidable whether a given ω-automatic presentation P belongs to $\mathcal{T}_n$ and $\mathcal{T}_n^{\mathrm{i}}$, resp. (since the class of trees of height at most n can be axiomatized in first-order logic).

3 ω-automatic trees of height 1 and 2

Two trees of height 1 are isomorphic if and only if they have the same size. Since the size of an ω-automatic structure is computable from any presentation [12], the isomorphism of ω-automatic trees of height 1 is decidable.

For ω-automatic trees of height 2 we need the following result:

Theorem 3.1 ([12]). *Let $\mathcal{A}$ be an ω-automatic structure and let $\varphi(x_1, \ldots, x_n, y)$ be a formula of $\mathsf{FO}[\exists^{\aleph_0}, \exists^{2^{\aleph_0}}]$. Then, for all $a_1, \ldots, a_n \in \mathcal{A}$, the cardinality of the set $\{b \in \mathcal{A} \mid \mathcal{A} \models \varphi(a_1, \ldots, a_n, b)\}$ belongs to $\mathbb{N} \cup \{\aleph_0, 2^{\aleph_0}\}$.*

Now, let us take two trees T_1 and T_2 of height 2 and let E_i be the edge relation of T_i and r_i its root. For $i \in \{1, 2\}$ and a cardinal λ let $\kappa_{\lambda,i}$ be the cardinality of the set of all $u \in E_i(r_i)$ such that $|E_i(u)| = \lambda$. Then $T_1 \cong T_2$ if and only if $\kappa_{\lambda,1} = \kappa_{\lambda,2}$ for any cardinal λ. Now assume that T_1 and T_2 are both ω-automatic. By Theorem 3.1, for all $i \in \{1, 2\}$ and every $u \in E_i(r_i)$ we have $|E_i(u)| \in \mathbb{N} \cup \{\aleph_0, 2^{\aleph_0}\}$. Moreover, again by Theorem 3.1, every cardinal $\kappa_{\lambda,1}$ ($\lambda \in \mathbb{N} \cup \{\aleph_0, 2^{\aleph_0}\}$) belongs to $\mathbb{N} \cup \{\aleph_0, 2^{\aleph_0}\}$ as well. Hence, $T_1 \cong T_2$ if and only if: $\forall \lambda, \kappa \in \mathbb{N} \cup \{\aleph_0, 2^{\aleph_0}\} : \kappa_{\lambda,1} = \kappa \Leftrightarrow \kappa_{\lambda,2} = \kappa$. By Theorem 2.2, the statement $\kappa_{\lambda,1} = \kappa \Leftrightarrow \kappa_{\lambda,2} = \kappa$ is decidable, so the whole statement belongs to Π_1^0. Hardness for Π_1^0 follows from the corresponding result on automatic trees of height 2 [18].

Theorem 3.2. *The following holds:*

- *The isomorphism problem $\mathsf{Iso}(\mathcal{T}_1)$ for ω-automatic trees of height 1 is decidable.*
- *There exists a tree U such that $\{P \in \mathcal{T}_2^{\mathrm{i}} \mid \mathcal{S}(P) \cong U\}$ is Π_1^0-hard. The isomorphism problems $\mathsf{Iso}(\mathcal{T}_2)$ and $\mathsf{Iso}(\mathcal{T}_2^{\mathrm{i}})$ for (injectively) ω-automatic trees of height 2 are Π_1^0-complete.*

4 A normal form for analytical sets

To prove our lower bound for the isomorphism problem of ω-automatic trees of height $n \geq 3$, we will use the following normal form of analytical sets. A formula of the form $x \in X$ or $x \notin X$ is called a *set constraint*.

Proposition 4.1. *For every odd (resp. even) $n \in \mathbb{N}_+$ and every Π^1_n (resp. Σ^1_n) relation $A \subseteq \mathbb{N}^r_+$, there exist polynomials $p_i, q_i \in \mathbb{N}[\overline{x}, y, \overline{z}]$ and disjunctions ψ_i $(1 \leq i \leq \ell)$ of set constraints (on the set variables $X_1, \dots, X_n$ and individual variables $\overline{x}, y, \overline{z}$) such that $\overline{x} \in A$ if and only if*

$$Q_1 X_1\ Q_2 X_2 \cdots Q_n X_n\ \exists y\ \forall \overline{z} : \bigwedge_{i=1}^{\ell} p_i(\overline{x}, y, \overline{z}) \neq q_i(\overline{x}, y, \overline{z}) \vee \psi_i(\overline{x}, y, \overline{z}, X_1, \dots, X_n)$$

where $Q_1, Q_2, \dots, Q_n$ are alternating quantifiers with $Q_n = \forall$.

The proof of this proposition uses standard arguments and Matiyasevitch's theorem on the equivalence of recursively enumerable and Diophantine sets [21], but we could not find it stated in precisely this form anywhere in the literature. It is known that the first-order quantifier block $\exists y \forall \overline{z}$ in Proposition 4.1 cannot be replaced by a block with only one type of first-order quantifiers, see e.g. [22].

5 ω-automatic trees of height at least 4

We prove the following theorem for injectively ω-automatic trees of height at least 4.

Theorem 5.1. *Let $n \geq 1$ and $\Theta \in \{\Sigma, \Pi\}$. There exists a tree $U_{n,\Theta}$ of height $n+3$ such that $\{P \in \mathcal{T}^{\mathrm{i}}_{n+3} \mid \mathcal{S}(P) \cong U_{n,\Theta}\}$ is hard for Θ^1_n. Hence,*

- *the isomorphism problem $\mathsf{Iso}(\mathcal{T}^{\mathrm{i}}_{n+3})$ for the class of injectively ω-automatic trees of height $n+3$ is hard for both the classes Π^1_n and Σ^1_n,*
- *and the isomorphism problem $\mathsf{Iso}(\mathcal{T}^{\mathrm{i}})$ for the class of injectively ω-automatic trees of finite height is not analytical.*

Theorem 5.1 will be derived from the following proposition whose proof occupies Sections 5.1 and 5.2.

Proposition 5.2. *Let $n \geq 1$. There are trees $U[0]$ and $U[1]$ of height $n+3$ such that for any set A that is Π^1_n if n is odd and Σ^1_n if n is even, one can compute from $x \in \mathbb{N}_+$ an injectively ω-automatic tree $T[x]$ of height $n+3$ with $T[x] \cong U[0]$ if and only if $x \in A$ and $T[x] \cong U[1]$ otherwise.*

Note that this implies in particular that $U[0]$ and $U[1]$ are injectively ω-automatic.

Let $n \geq 1$ and set $U_{n,\Sigma} = U[n \bmod 2]$ and $U_{n,\Pi} = U[(n+1) \bmod 2]$. Then Proposition 5.2 implies the first statement of Theorem 5.1. The remaining statements are consequences of this first statement.

The construction of the trees $T[x]$, $U[0]$, and $U[1]$ is uniform in n and the formula defining A. Hence the second-order theory of $(\mathbb{N}, +, \times)$ can be reduced to $\bigcup_{n \in \mathbb{N}_+} \{n\} \times \mathsf{Iso}(\mathcal{T}^{\mathrm{i}}_n)$ and therefore to the isomorphism problem $\mathsf{Iso}(\bigcup_{n \in \mathbb{N}_+} \mathcal{T}^{\mathrm{i}}_n)$. This proves:

Corollary 5.3. *The second-order theory of $(\mathbb{N}, +, \times)$ can be reduced to the isomorphism problem $\mathsf{Iso}(\bigcup_{n \in \mathbb{N}_+} \mathcal{T}^{\mathrm{i}}_n)$ for the class of all injectively ω-automatic trees of finite height.*

We now start to prove Proposition 5.2. Let A be a set that is Π^1_n if n is odd and Σ^1_n otherwise. By Proposition 4.1 it can be written in the form

$$A = \{x \in \mathbb{N}_+ \mid Q_1 X_1 \dots Q_n X_n \exists y \, \forall \overline{z} : \bigwedge_{i=1}^{\ell} p_i(x, y, \overline{z}) \neq q_i(x, y, \overline{z}) \vee \psi_i(x, y, \overline{z}, \overline{X})\}$$

where

- $Q_1, Q_2, \dots, Q_n$ are alternating quantifiers with $Q_n = \forall$,
- p_i, q_i $(1 \leq i \leq \ell)$ are polynomials in $\mathbb{N}[x, y, \overline{z}]$ where $\overline{z}$ has length k, and
- every ψ_i is a disjunction of set constraints on the set variables $X_1, \dots, X_n$ and the individual variables $x, y, \overline{z}$.

For $0 \leq m \leq n$, we will consider the formula $\varphi_m(x, X_1, \dots, X_{n-m})$ defined by

$$Q_{n+1-m} X_{n+1-m} \dots Q_n X_n \; \exists y \; \forall \overline{z} : \bigwedge_{i=1}^{\ell} p_i(x, y, \overline{z}) \neq q_i(x, y, \overline{z}) \vee \psi_i(x, y, \overline{z}, \overline{X})$$

such that $\varphi_0(x, X_1, \dots, X_n)$ is a first-order formula and $\varphi_n(x)$ holds if and only if $x \in A$. In addition, let $\varphi_{-1}(x, y, X_1, \dots, X_n)$ be the subformula starting with $\forall \overline{z}$.

To prove Proposition 5.2, we construct by induction on $0 \leq m \leq n$ height-$(m+3)$ trees $T_m[X_1, \dots, X_{n-m}, x]$ and $U_m[i]$ where $X_1, \dots, X_{n-m} \subseteq \mathbb{N}_+$, $x \in \mathbb{N}_+$, and $i \in \{0, 1\}$ such that the following holds:

$$\forall \overline{X} \in (2^{\mathbb{N}_+})^{n-m} \; \forall x \in \mathbb{N}_+ : T_m[\overline{X}, x] \cong \begin{cases} U_m[0] & \text{if } \varphi_m(x, \overline{X}) \text{ holds} \\ U_m[1] & \text{otherwise} \end{cases} \tag{1}$$

Setting $T[x] = T_n[x]$, $U[0] = U_n[0]$, and $U[1] = U_n[1]$ and constructing from x an injectively ω-automatic presentation of $T[x]$ then proves Proposition 5.2.

5.1 Construction of trees

Note that $C : \mathbb{N}_+^2 \to \mathbb{N}_+$ with $C(x, y) = (x+y)^2 + 3x + y$ is an injective polynomial function [**?**] ($C(x, y)/2$ defines a pairing function, see e.g. [10]). For two numbers $x, y \in \mathbb{N}_+$, let $S[x, y]$ denote the height-1 tree with $C(x, y)$ many leaves.

The trees are constructed by induction on m, $m = 0$ being the base case: For all $\overline{X} \in (2^{\mathbb{N}_+})^n$, $\overline{z} \in \mathbb{N}_+^k$, $x, y, z_{k+1} \in \mathbb{N}_+$, $1 \leq i \leq \ell$, and $\kappa \in \mathbb{N}_+ \cup \{\omega\}$ define the trees[5]

$$T'[\overline{X}, x, y, \overline{z}, z_{k+1}, i] = \begin{cases} S[1, 2] & \text{if } \psi_i(x, y, \overline{z}, \overline{X}) \\ S[p_i(x, y, \overline{z}) + z_{k+1}, q_i(x, y, \overline{z}) + z_{k+1}] & \text{otherwise} \end{cases}$$

and

$$T''[\overline{X}, x, y] = r \circ \begin{pmatrix} \biguplus \{S[e_1, e_2] \mid e_1 \neq e_2\} \uplus \\ \biguplus \{T'[\overline{X}, x, y, \overline{z}, z_{k+1}, i] \mid \overline{z} \in \mathbb{N}_+^k, z_{k+1} \in \mathbb{N}_+, 1 \leq i \leq \ell\} \end{pmatrix}^{\aleph_0}$$

$$U''[\kappa] = r \circ \left(\biguplus \{S[e_1, e_2] \mid e_1 \neq e_2\} \uplus \biguplus \{S[e, e] \mid \kappa \leq e < \omega\} \right)^{\aleph_0}.$$

[5] The choice of $S[1, 2]$ in the definition of $T'[\overline{X}, x, y, \overline{z}, z_{k+1}, i]$ is arbitrary. Any $S[a, b]$ with $a \neq b$ would be acceptable.

Note that all the trees $T''[\overline{X}, x, y]$ and $U''[\kappa]$ are build from trees of the form $S[e_1, e_2]$. Furthermore, if $S[e, e]$ appears as a building block, then $S[e + a, e + a]$ also appears as one for all $a \in \mathbb{N}$. In addition, any building block $S[e_1, e_2]$ appears either infinitely often or not at all. These observations allow to prove the following:

(a) $T''[\overline{X}, x, y] \cong U''[\kappa]$ for some $\kappa \in \mathbb{N}_+ \cup \{\omega\}$
(b) $T''[\overline{X}, x, y] \cong U''[\omega]$ if and only if $\varphi_{-1}(x, y, \overline{X})$ holds

In a next step, we collect the trees $T''[\overline{X}, x, y]$ and $U''[\kappa]$ into the trees $T_0[\overline{X}, x]$, $U_0[0]$, and $U_0[1]$ as follows:

$$\begin{aligned} T_0[\overline{X}, x] &= r \circ \left(\biguplus \{U''[m] \mid m \in \mathbb{N}_+\} \uplus \biguplus \{T''[\overline{X}, x, y] \mid y \in \mathbb{N}_+\} \right)^{\aleph_0} \\ U_0[0] &= r \circ \left(\biguplus \{U''[m] \mid m \in \mathbb{N}_+ \cup \{\omega\}\} \right)^{\aleph_0} \\ U_0[1] &= r \circ \left(\biguplus \{U''[m] \mid m \in \mathbb{N}_+\} \right)^{\aleph_0} \end{aligned}$$

By (a), these trees are build from copies of the trees $U''[\kappa]$ (and are therefore of height 3), each appearing infinitely often or not at all. Hence $T_0[\overline{X}, x]$ is isomorphic to $U_0[0]$ or to $U_0[1]$ (and these two trees are not isomorphic). Note that $T_0[\overline{X}, x] \cong U_0[0]$ if and only if there exists some $y \in \mathbb{N}_+$ with $T''[\overline{X}, x, y] \cong U''[\omega]$. By (b) and the definition of $\varphi_0(\overline{X}, x)$, this is the case if and only if $\varphi_0(\overline{X}, x)$ holds. Hence (1) holds for $m = 0$.

Suppose for some number $0 \leq m < n$ we have trees $T_m[X_1, \ldots, X_{n-m}, x]$, $U_m[0]$ and $U_m[1]$ satisfying (1). Let $\overline{X}$ stand for $(X_1, \ldots, X_{n-m-1})$ and let $\alpha = m \bmod 2$. We define the following height-$(m + 4)$ trees:

$$\begin{aligned} T_{m+1}[\overline{X}, x] &= r \circ \left(U_m[\alpha] \uplus \biguplus \left\{ T_m[\overline{X}, X_{n-m}, x] \mid X_{n-m} \subseteq \mathbb{N}_+ \right\} \right)^{2^{\aleph_0}} \\ U_{m+1}[i] &= r \circ (U_m[\alpha] \uplus U_m[i])^{2^{\aleph_0}} \quad \text{for } i \in \{0, 1\} \end{aligned}$$

Note that the trees $T_{m+1}[\overline{X}, x]$, $U_{m+1}[0]$, and $U_{m+1}[1]$ consist of $2^{\aleph_0}$ many copies of $U_m[\alpha]$ and possibly $2^{\aleph_0}$ many copies of $U_m[1 - \alpha]$. Hence $T_{m+1}[\overline{X}, x]$ is isomorphic to one of the trees $U_{m+1}[0]$ or $U_{m+1}[1]$. We show that $T_{m+1}[\overline{X}, x] \cong U_{m+1}[0]$ if and only if $\varphi_{m+1}(x, \overline{X})$ for the case that m even, i.e., $\alpha = 0$ (the case m odd is similar): The two trees are isomorphic if and only if $T_m[\overline{X}, X_{n-m}, x] \cong U_m[0]$ for all sets $X_{n-m} \subseteq \mathbb{N}_+$. By the induction hypothesis, this is equivalent to saying that the formula $\varphi_m(x, \overline{X}, X_{n-m})$ holds for all sets $X_{n-m} \subseteq \mathbb{N}_+$. But since m is even, the formula $\varphi_{m+1}(x, \overline{X})$ equals $\forall X_{n-m} : \varphi_m(x, \overline{X}, X_{n-m})$.

This finishes the construction of the trees $T_m[\overline{X}, x]$, $U_m[0]$, and $U_m[1]$ as well as the verification of (1). For $m = n$ we get:

Lemma 5.4. *For all $x \in \mathbb{N}_+$, we have $T_n[x] \cong U_n[0]$ if $x \in A$ and $T_n[x] \cong U_n[1]$ otherwise.*

5.2 Injective ω-automaticity

Injectively ω-automatic presentations of the trees $T_m[\overline{X}, x]$, $U_m[0]$, and $U_m[1]$ will be constructed inductively. Note that the construction of $T_{m+1}[\overline{X}, x]$ involves all the trees

$T_m[\overline{X}, X_{n-m}, x]$ for $X_{n-m} \subseteq \mathbb{N}_+$. Hence we need *one single injectively ω-automatic presentation* for the forest consisting of all these trees. Therefore, we will deal with forests. To move from one forest to the next, we will always proceed as follows: add a set of new roots and connect them to some of the old roots *which results in a directed acyclic graph* (or dag) and not necessarily in a forest. The next forest will then be the unfolding of this dag.

The *height* of a dag D is the length (number of edges) of a longest directed path in D. We only consider dags of finite height. A *root* of a dag is a node without incoming edges. A dag $D = (V, E)$ can be unfold into a forest $\mathrm{unfold}(D)$ in the usual way: Nodes of $\mathrm{unfold}(D)$ are directed paths in D that start in a root and the order relation is the prefix relation between these paths. For a root $v \in V$ of D, we define the tree $\mathrm{unfold}(D, v)$ as the restriction of $\mathrm{unfold}(D)$ to those paths that start in v. We will make use of the following lemma whose proof is based on the immediate observation that the set of convolutions of paths in D is again a regular language

Lemma 5.5. *From a given $k \in \mathbb{N}$ and an injectively ω-automatic presentation for a dag D of height at most k, one can construct effectively an injectively ω-automatic presentation for $\mathrm{unfold}(D)$ such that the roots of $\mathrm{unfold}(D)$ coincide with the roots of D and $\mathrm{unfold}(D, r) = (\mathrm{unfold}(D))(r)$ for any root r.*

For a symbol a and a tuple $\overline{e} = (e_1, \ldots, e_k) \in \mathbb{N}_+^k$, we write $a^{\overline{e}}$ for the ω-word

$$a^{e_1} \otimes a^{e_2} \otimes \cdots \otimes a^{e_k} = (a^{e_1} \diamond^\omega) \otimes (a^{e_2} \diamond^\omega) \otimes \cdots \otimes (a^{e_k} \diamond^\omega) \, .$$

For an ω-language L, we write $\otimes_k(L)$ for $\otimes(L^k)$. For $X \subseteq \mathbb{N}_+$, let $w_X \in \{0,1\}^*$ be the characteristic word (i.e., $w_X[i] = 1$ if and only if $i \in X$) and, for $\overline{X} = (X_1, \ldots, X_n) \in (2^{\mathbb{N}_+})^n$, write $w_{\overline{X}}$ for the convolution of the words w_{X_i}. The following lemma is key to the construction of ω-automatic presentations for $T_n[x]$, $U_n[0]$, and $U_n[1]$. We refer to the definition of the set A from Section 5.

Lemma 5.6. *For $1 \leq i \leq \ell$, there exists a Büchi-automaton $\mathcal{A}_i$ with the following property: For all $\overline{X} \in (2^{\mathbb{N}_+})^n$, $\overline{z} \in \mathbb{N}_+^k$, and $x, y, z_{k+1} \in \mathbb{N}_+$, the number of accepting runs of $\mathcal{A}_i$ on the word $w_{\overline{X}} \otimes a^{(x,y,\overline{z},z_{k+1})}$ equals $C(1,2)$ if $\psi_i(x, y, \overline{z}, X_1, \ldots, X_n)$ holds and $C(p_i(x, y, \overline{z}) + z_{k+1}, q_i(x, y, \overline{z}) + z_{k+1})$ otherwise.*

Proof sketch. One first builds a Büchi automaton that, on the ω-word $a^{(x,y,\overline{z},z_{k+1})}$, has precisely $C(p_i(x,y,\overline{z})+z_{k+1}, q_i(x,y,\overline{z})+z_{k+1})$ many accepting runs. This is possible using disjoint union of automata and the flag construction (cf. [4, 23, 26]) for addition and multiplication of polynomials since $C(p_i(x,y,\overline{z}) + z_{k+1}, q_i(x,y,\overline{z}) + z_{k+1})$ is a polynomial over $\mathbb{N}$. Secondly, one builds deterministic Büchi automata accepting a word $w_{\overline{X}} \otimes a^{(x,y,\overline{z},z_{k+1})}$ if and only if the disjunction $\psi_i(x, y, \overline{z}, \overline{X})$ of set constraints is satisfied (not satisfied, resp.) A straightforward combination of (several copies of) the automata obtained in this way has the desired properties. □

For a Büchi-automaton $\mathcal{A}$, let $\mathrm{Run}_{\mathcal{A}}$ denote the set of accepting runs. Note that this set is a regular ω-language over the alphabet of transitions of $\mathcal{A}$.

We now build a first injectively ω-automatic forest $\mathcal{H}' = (L', E')$: its underlying ω-language is (every $1 \leq i \leq \ell$ is a new symbol in L')

$$L' = \bigcup_{1 \leq i \leq \ell} i \otimes (L(\mathcal{A}_i) \cup \mathrm{Run}_{\mathcal{A}_i})$$

and $(i \otimes v, j \otimes w)$ forms an edge if and only if $i = j$ and w is an accepting run of $\mathcal{A}_i$ on v. Then $\mathcal{H}'$ is a forest of height 1, the roots of $\mathcal{H}'$ are the words from $\{1, \ldots, \ell\} \otimes (\otimes_n(\{0,1\}^\omega)) \otimes (\otimes_{k+3}(a^+))$, and for any root $r = i \otimes w_{\overline{X}} \otimes a^{(x,y,\overline{z},z_{k+1})}$, we have $\mathcal{H}'(r) \cong T'[\overline{X}, x, y, \overline{z}, z_{k+1}, i]$ by Lemma 5.6.

In a similar way, one can build an injectively ω-automatic forest $\mathcal{F} = (L_\mathcal{F}, E_\mathcal{F})$ whose roots are the words from $b^+ \otimes b^+$ such that $\mathcal{F}(b^{(e_1,e_2)}) \cong S[e_1, e_2]$. From $\mathcal{H}'$ and $\mathcal{F}$, we build an injectively ω-automatic dag $\mathcal{D}$ as follows:

– The domain of $\mathcal{D}$ is the set $(\otimes_n(\{0,1\}^\omega) \otimes a^+ \otimes a^+) \,\cup\, b^* \,\cup\, \big(\$^* \otimes (L' \cup L_\mathcal{F})\big)$.
– For $u, v \in L' \cup L_\mathcal{F}$, the words $\$^i \otimes u$ and $\$^j \otimes v$ are connected if and only if $i = j$ and $(u, v) \in E' \cup E_\mathcal{F}$. In other words, the restriction of $\mathcal{D}$ to $\$^* \otimes (L' \cup L_\mathcal{F})$ is isomorphic to $(\mathcal{H}' \uplus \mathcal{F})^{\aleph_0}$.
– For all $\overline{X} \in (2^{\mathbb{N}}_+)^n$, $x, y \in \mathbb{N}_+$, the new root $w_{\overline{X}} \otimes a^{(x,y)}$ is connected to all nodes in

$$\$^* \otimes \Big((\{1, \ldots, \ell\} \otimes w_{\overline{X}} \otimes a^{(x,y)} \otimes (\otimes_{k+1}(a^+))) \cup \{b^{(e_1,e_2)} \mid e_1 \neq e_2\} \Big) .$$

– The new root ε is connected to all nodes in $\$^* \otimes \{b^{(e_1,e_2)} \mid e_1 \neq e_2\}$.
– For all $m \in \mathbb{N}_+$, the new root b^m is connected to all nodes in

$$\$^* \otimes \{b^{(e_1,e_2)} \mid e_1 \neq e_2 \vee e_1 = e_2 \geq m\}.$$

It is easily seen that $\mathcal{D}$ is an injectively ω-automatic dag. Setting $\mathcal{H}'' = \mathrm{unfold}(\mathcal{D})$, one can then verify $\mathcal{H}''(w_X \otimes a^{(x,y)}) \cong T''[\overline{X}, x, y]$, $\mathcal{H}''(\varepsilon) \cong U''[\omega]$, and $\mathcal{H}''(b^m) \cong U''[m]$ for all $m \in \mathbb{N}_+$. Note that also $\mathcal{H}''$ is injectively ω-automatic by Lemma 5.5.

We now construct a new forest $\mathcal{H}_0$ from $\$^* \otimes \mathcal{H}''$ by adding new roots:

– For $\overline{X} \in (2^{\mathbb{N}_+})^n$, $x \in \mathbb{N}_+$, connect a new root $w_{\overline{X}} \otimes a^x$ to all nodes in

$$\$^* \otimes \left(w_{\overline{X}} \otimes a^x \otimes a^+ \,\cup\, b^+\right).$$

– Connect a new root ε to all nodes in $\$^* \otimes b^*$.
– Connect a new root b to all nodes in $\$^* \otimes b^+$.

The result is an injectively ω-automatic dag of height 3 whose unfolding we denote by $\mathcal{H}_0$. This forest $\mathcal{H}_0$ is actually the base case of the following lemma. The induction is done similarly: one adds to $\{\$_1\$_2\}^\omega \otimes \mathcal{H}_m$ new roots $w_{\overline{X}} \otimes a^x$ for $x \in \mathbb{N}_+$ and $\overline{X} \in (2^{\mathbb{N}_+})^{n-m}$, ε, and b and one connects them to the appropriate words $u \otimes v$ with $u \in \{\$_1, \$_2\}^\omega$ and $v \in \mathcal{H}_m$ (cf. the definition of the trees $T_{m+1}[\overline{X}, x]$ and $U_{m+1}[i]$ for $i \in \{0, 1\}$).

Lemma 5.7. *From each $0 \leq m \leq n$, one can effectively construct an injectively ω-automatic forest $\mathcal{H}_m$ such that*

– *the set of roots of* $\mathcal{H}_m$ *is* $\big(\otimes_{n-m}(\{0,1\}^\omega)\otimes a^+\big)\cup\{\varepsilon,b\}$,
– $\mathcal{H}_m(w_{\overline{X}}\otimes a^x)\cong T_m[\overline{X},x]$ *for all* $\overline{X}\in(2^{\mathbb{N}_+})^{n-m}$ *and* $x\in\mathbb{N}_+$,
– $\mathcal{H}_m(\varepsilon)\cong U_m[0]$, *and*
– $\mathcal{H}_m(b)\cong U_m[1]$.

Note that $T_n[x]$ is the tree in $\mathcal{H}_n$ rooted at a^x. Hence $T_n[x]$ is (effectively) an injectively ω-automatic tree. Now Lemma 5.4 finishes the proof of Proposition 5.2 and therefore of Theorem 5.1.

6 ω-automatic trees of height 3

Recall that the isomorphism problem $\mathsf{Iso}(\mathcal{T}_2^{\mathrm{i}})$ is arithmetical by Theorem 3.2 and that $\mathsf{Iso}(\mathcal{T}_4^{\mathrm{i}})$ is not by Theorem 5.1. In this section, we modify the proof of Theorem 5.1 in order to show:

Theorem 6.1. *There exists a tree* U *such that* $\{P\in\mathcal{T}_3^{\mathrm{i}}\mid\mathcal{S}(P)\cong U\}$ *is* Π_1^1*-hard. Hence the isomorphism problem* $\mathsf{Iso}(\mathcal{T}_3^{\mathrm{i}})$ *for injectively* ω*-automatic trees of height 3 is* Π_1^1*-hard.*

So let $A\subseteq\mathbb{N}_+$ be some set from Π_1^1. By Proposition 4.1, it can be written as

$$A=\{x\in\mathbb{N}_+:\forall X\;\exists y\;\forall\overline{z}:\bigwedge_{i=1}^{\ell}p_i(x,y,\overline{z})\neq q_i(x,y,\overline{z})\vee\psi_i(x,y,\overline{z},X)\},$$

where p_i and q_i are polynomials with coefficients in $\mathbb{N}$ and ψ_i is a disjunction of set constraints. As in Section 5, let $\varphi_0(x,X)$ denote the first-order kernel of this formula (starting with $\exists y$) and let $\varphi_{-1}(x,y,X)$ denote the subformula starting with $\forall\overline{z}$. We reuse the trees $T'[X,x,y,\overline{z},z_{k+1},i]$ of height 1. Recall that they are all of the form $S[e_1,e_2]$ and therefore have an even number of leaves (since the range of the polynomial $C:\mathbb{N}^2\to\mathbb{N}$ consists of even numbers). For $e\in\mathbb{N}_+$, let $S[e]$ denote the height-1 tree with $2e+1$ leaves.

Recall that the tree $T''[X,x,y]$ encodes the set of pairs $(e_1,e_2)\in\mathbb{N}_+^2$ such that $e_1\neq e_2$ or there exist $\overline{z}$, z_{k+1}, and i with $e_1=p_i(x,y,\overline{z})+z_{k+1}$ and $e_2=q_i(x,y,\overline{z})+z_{k+1}$. We now modify the construction of this tree such that it, in addition, also encodes the set $X\subseteq\mathbb{N}_+$:

$$\widehat{T}[X,x,y]=r\circ\begin{pmatrix}\biguplus\{S[e]\mid e\in X\}\uplus\biguplus\{S[e_1,e_2]\mid e_1\neq e_2\}\uplus\\ \biguplus\{T'[\overline{X},x,y,\overline{z},z_{k+1}i]\mid\overline{z}\in\mathbb{N}_+^k,z_{k+1}\in\mathbb{N}_+,1\leq i\leq\ell\}\end{pmatrix}^{\aleph_0}$$

In a similar spirit, we define $\widehat{U}[\kappa,X]$ for $X\subseteq\mathbb{N}_+$ and $\kappa\in\mathbb{N}_+\cup\{\omega\}$:

$$\widehat{U}[\kappa,X]=r\circ\begin{pmatrix}\biguplus\{S[e]\mid e\in X\}\uplus\biguplus\{S[e_1,e_2]\mid e_1\neq e_2\}\uplus\\ \biguplus\{S[e,e]\mid\kappa\leq e<\omega\}\end{pmatrix}^{\aleph_0}$$

Then $\widehat{T}[X,x,y] \cong \widehat{U}[\omega,Y]$ if and only if $X = Y$ and $T''[X,x,y] \cong U''[\omega]$, i.e., (by (b) from page 9) if and only if $X = Y$ and $\varphi_{-1}(x,y,X)$ holds. Finally, we set

$$T[x] = r \circ \left(\biguplus \{\widehat{U}[\kappa,X] \mid X \subseteq \mathbb{N}_+, \kappa \in \mathbb{N}_+\} \uplus \biguplus \{\widehat{T}[X,x,y] \mid X \subseteq \mathbb{N}_+, y \in \mathbb{N}_+\} \right)^{\aleph_0}$$

$$U = r \circ \left(\biguplus \{\widehat{U}[\kappa,X] \mid X \subseteq \mathbb{N}_+, \kappa \in \mathbb{N}_+ \cup \{\omega\}\} \right)^{\aleph_0} .$$

Then $T[x] \cong U$ if and only if they have the same height-2 subtrees. The crucial point is that, for any $X \subseteq \mathbb{N}_+$, $\widehat{U}[X,\omega]$ appears in U, i.e., there must be some $Y \subseteq \mathbb{N}_+$ and $y \in \mathbb{N}_+$ such that $\widehat{T}[Y,x,y] \cong \widehat{U}[\omega,X]$ (implying $X = Y$) and therefore $T''[X,x,y] \cong U''[\omega]$. But this was equivalent to saying that $\varphi_{-1}(x,y,X)$ holds. Since $X \subseteq \mathbb{N}_+$ was arbitrary, we showed $x \in A$. The other implication can be shown similarly. Injective ω-automaticity can be shown similar to Section 5. This finishes our proof sketch of Theorem 6.1.

7 Upper bounds assuming CH

We denote with **CH** the continuum hypothesis: Every infinite subset of $2^{\mathbb{N}}$ has either cardinality $\aleph_0$ or cardinality $2^{\aleph_0}$. By the seminal work of Cohen and Gödel, **CH** is independent of the axiom system **ZFC**.

In the following, we will identify an ω-word $w \in \Gamma^\omega$ with the function $w : \mathbb{N}_+ \to \Gamma$ (and hence with a second-order object) where $w(i) = w[i]$. We need the following lemma:

Lemma 7.1. *From a given Büchi automaton M over an alphabet Γ one can construct an arithmetical predicate $\mathsf{acc}_M(u)$ (where $u : \mathbb{N}_+ \to \Gamma$) such that $u \in L(M)$ if and only if $\mathsf{acc}_M(u)$ holds.*

Proof sketch. The idea is to transform M into an equivalent (deterministic and complete) Muller automaton. Determinism then allows to express acceptance using the arithmetical predicate "the prefix of length n results in state q". □

Theorem 7.2. *Assuming* **CH***, the isomorphism problem $\mathsf{Iso}(\mathcal{T}_n)$ belongs to Π^1_{2n-4} for $n \geq 3$.*

Proof sketch. Consider trees $T_i = \mathcal{S}(P_i)$ for $P_1, P_2 \in \mathcal{T}_n$. Define the forest $F = (V, \leq)$ as $F = T_1 \uplus T_2$. Let us fix an ω-automatic presentation $P = (\Sigma, M, M_\equiv, M_\leq)$ for F where $M_\leq$ recognizes the order relation $\leq$. In the following, for $u \in L(M)$ we write $F(u)$ for the subtree $F([u])$ rooted in the F-node $[u] = [u]_{R(M_\equiv)}$ represented by the ω-word u. Similarly, we write $E(u)$ for the set of children of $[u]$. We will define a $\Pi^1_{2n-2k-4}$-predicate $\mathsf{iso}_k(u_1,u_2)$, where $u_1, u_2 \in L(M)$ are on level k in F. This predicate expresses that $F(u_1) \cong F(u_2)$.

As induction base, let $k = n-2$. Then the trees $F(u_1)$ and $F(u_2)$ have height at most 2. Then $F(u_1) \cong F(u_2)$ if and only if the following holds for all $\kappa, \lambda \in$

$\mathbb{N} \cup \{\aleph_0, 2^{\aleph_0}\}$ (see Section 3):

$$F \models \Big(\exists^\kappa x \in V : (([u_1], x) \in E \wedge \exists^\lambda y \in V : (x, y) \in E)\Big) \leftrightarrow$$
$$\Big(\exists^\kappa x \in V : (([u_2], x) \in E \wedge \exists^\lambda y \in V : (x, y) \in E)\Big).$$

Note that by Theorem 2.2, one can compute from $\kappa, \lambda \in \mathbb{N} \cup \{\aleph_0, 2^{\aleph_0}\}$ a Büchi automaton $M_{\kappa,\lambda}$ accepting the set of convolutions of pairs of ω-words (u_1, u_2) satisfying the above formula. Hence $F(u_1) \cong F(u_2)$ if and only if the following arithmetical predicate holds: $\forall \kappa, \lambda \in \mathbb{N} \cup \{\aleph_0, 2^{\aleph_0}\} : \mathsf{acc}_{M_{\kappa,\lambda}}(u_1, u_2)$.

Now let $0 \leq k < n-2$. For a set A, let $\mathsf{count}(A)$ denote the set of all countable (possibly finite) subsets of A. On an abstract level, the formula $\mathsf{iso}_k(u_1, u_2)$ is:

$$\big(\forall x \in E(u_1) \exists y \in E(u_2) : \mathsf{iso}_{k+1}(x, y)\big) \ \wedge \ \big(\forall x \in E(u_2) \exists y \in E(u_1) : \mathsf{iso}_{k+1}(x, y)\big)$$
$$\wedge \ \forall X_1 \in \mathsf{count}(E(u_1)) \, \forall X_2 \in \mathsf{count}(E(u_2)) :$$
$$|X_1| = |X_2| \ \vee \ (\exists x, y \in X_1 \cup X_2 : \neg \mathsf{iso}_{k+1}(x, y)) \ \vee$$
$$(\exists x \in X_1 \cup X_2 \ \exists y \in (E(u_1) \cup E(u_2)) \setminus (X_1 \cup X_2) : \mathsf{iso}_{k+1}(x, y))$$

The first line expresses that the children of u_1 and u_2 realize the same isomorphism types of trees of height $n-k-1$. The rest of the formula expresses that if a certain isomorphism type τ of height-$(n-k-1)$ trees appears countably many times below u_1 then it appears with the same multiplicity below u_2 and vice versa. Assuming **CH** and the correctness of iso_{k+1}, one gets $\mathsf{iso}_k(u_1, u_2)$ if and only if $F(u_1) \cong F(u_2)$.

The sets X_i in the above formula can be coded as mappings $f_i : \mathbb{N}_+^2 \to \Sigma$. Then the elements of X_i correspond to natural numbers j coding the word $k \mapsto f_i(j, k)$. But the ω-word $y \notin X_1 \cup X_2$ is another second-order object. This results in two additional second-order quantifier blocks in iso_k. Hence the formula iso_0 belongs to Π^1_{2n-4}. In order to express that e.g. $x \in E(u_i)$ we use Lemma 7.1 with the automaton $M_{\leq}$. $\square$

Corollary 5.3 and 7.2 imply:

Corollary 7.3. *Assuming* **CH***, the isomorphism problem for (injectively) ω-automatic trees of finite height is recursively equivalent to the second-order theory of* $(\mathbb{N}, +, \times)$.

Remark 7.4. For the case $n = 3$ we can avoid the use of **CH** in Theorem 7.2: Let us consider the proof of Theorem 7.2 for $n = 3$. Then, the binary relation iso_1 (which holds between two ω-words u, v in F if and only if $[u]$ and $[v]$ are on level 1 and $F(u) \cong F(v)$) is a Π^0_1-predicate. It follows that this relation is Borel (see e.g. [13] for background on Borel sets). Now let u be an ω-word on level 1 in F. It follows that the set of all ω-words v on level 1 with $\mathsf{iso}_1(u, v)$ is again Borel. Now, every uncountable Borel set has cardinality $2^{\aleph_0}$ (this holds even for analytic sets [13]). It follows that the definition of iso_0 in the proof of Theorem 7.2 is correct even without assuming **CH**. Hence, $\mathsf{Iso}(\mathcal{T}_3)$ belongs to Π^1_2 (recall that we proved Π^1_1-hardness for this problem in Section 6), this can be shown in **ZFC**.

8 Open problems

The main open problem concerns upper bounds in case we assume the negation of the continuum hypothesis. Assuming $\neg\mathbf{CH}$, is the isomorphism problem for (injectively) ω-automatic trees of height n still analytical? In our paper [18] we also proved that the isomorphism problem for automatic linear orders is not arithmetical. This leads to the question whether our techniques for ω-automatic trees can be also used for proving lower bounds on the isomorphism problem for ω-automatic linear orders. More specifically, one might ask whether the isomorphism problem for ω-automatic linear orders is analytical. A more general question asks for the complexity of the isomorphism problem for ω-automatic structures in general. On the face of it, it is an existential third-order property (since any isomorphism has to map second-order objects to second-order objects). But it is not clear whether it is complete for this class.

References

1. A. Blumensath. Automatic structures. Diploma thesis, RWTH Aachen, 1999.
2. A. Blumensath and E. Grädel. Finite presentations of infinite structures: Automata and interpretations. *Theory of Computing Systems*, 37(6):641–674, 2004.
3. W. Calvert and J F. Knight. Classification from a computable viewpoint. *Bull. Symbolic Logic*, 12(2):191–218, 2006.
4. Y. Choueka. Theories of automata on ω-tapes: a simplified approach. *Journal of Computer and System Sciences*, 8:117–141, 1974.
5. B. Courcelle. The definability of equational graphs in monadic second-order logic. In *Proceedings of ICALP 1989*, LNCS 372, pages 207–221. Springer, 1989.
6. O. Finkel. Highly undecidable problems for infinite computations. *R.A.I.R.O. — Informatique Théorique et Applications*, 43(2):339–364, 2009.
7. O. Finkel and S. Todorčević. The Isomorphism Relation Between Tree-Autoamtic Structures. *Central European Journal of Mathematics*, 8(2):299–313, 2010.
8. S. S. Goncharov and J. F. Knight. Computable structure and antistructure theorems. *Algebra Logika*, 41(6):639–681, 2002.
9. G. Hjorth, B. Khoussainov, A. Montalbán, and A. Nies. From automatic structures to Borel structures. In *Proceedings of LICS 2008*, pages 431–441. IEEE Computer Society, 2008.
10. J. Honkala. On the problem whether the image of an N-rational series equals N. *Fund. Inform.*, 73(1-2):127–132, 2006.
11. H. Ishihara, B. Khoussainov, and S. Rubin. Some results on automatic structures. In *Proceedings of LICS 2002*, pages 235–244. IEEE Computer Society Press, 2002.
12. L. Kaiser, S. Rubin, and V. Bárány. Cardinality and counting quantifiers on omega-automatic structures. In *Proceedings of STACS 2008*, volume 08001 of *Dagstuhl Seminar Proceedings*, pages 385–396. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2008.
13. A. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.
14. B. Khoussainov and A. Nerode. Automatic presentations of structures. In *LCC: International Workshop on Logic and Computational Complexity*, LNCS 960, pages 367–392, 1995.
15. B.Khoussainov, A. Nies, S. Rubin, and F. Stephan. Automatic structures: richness and limitations. *Logical Methods in Computer Science*, 3(2):2:2, 18 pp. (electronic), 2007.
16. B. Khoussainov, S. Rubin, and F. Stephan. Automatic linear orders and trees. *ACM Transactions on Computational Logic*, 6(4):675–700, 2005.

17. D. Kuske. Is Ramsey's theorem ω-automatic? In *Proceedings of STACS 2010*, Leibniz International Proceedings in Informatics (LIPIcs) vol. 5, pages 537–548. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2010.

18. D. Kuske, J. Liu, and M. Lohrey. The isomorphism problem on classes of automatic structures. In *Proceedings of LICS 2010*. IEEE Computer Society, 2010. Accepted for publication.

19. D. Kuske, J. Liu, and M. Lohrey. The isomorphism problem for ω-automatic trees. available at http://arxiv.org/abs/1004.0610

20. D. Kuske and M. Lohrey. First-order and counting theories of omega-automatic structures. *Journal of Symbolic Logic*, 73:129–150, 2008.

21. Y. V. Matiyasevich. *Hilbert's Tenth Problem*. MIT Press, Cambridge, Massachusetts, 1993.

22. P. Odifreddi. *Classical recursion theory*, volume 125 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., 1989.

23. D. Perrin and J.-E. Pin. *Infinite Words*. Pure and Applied Mathematics vol. 141. Elsevier, 2004.

24. S. Rubin. *Automatic Structures*. PhD thesis, University of Auckland, 2004.

25. S. Rubin. Automata presenting structures: A survey of the finite string case. *Bulletin of Symbolic Logic*, 14:169–209, 2008.

26. W. Thomas. Automata on infinite objects. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, chapter 4, pages 133–191. Elsevier Science Publishers B. V., 1990.