Applications of Kolmogorov Complexity to Computable Model Theory*

Bakhadyr Khoussainov[†], Pavel Semukhin[‡], Frank Stephan[§]

Abstract

In this paper we answer the following well-known open question in computable model theory. Does there exist a computable not $\aleph_0$-categorical saturated structure with a unique computable isomorphism type? Our answer is affirmative and uses a construction based on Kolmogorov complexity. With a variation of this construction, we also provide an example of an $\aleph_1$-categorical but not $\aleph_0$-categorical saturated Σ_1^0-structure with a unique computable isomorphism type. In addition, using the construction we give an example of an $\aleph_1$-categorical but not $\aleph_0$-categorical theory whose only non-computable model is the prime one.

1 Introduction

Our main interest in this paper concerns the existence of a computable not $\aleph_0$-categorical saturated structure with a unique computable isomorphism type. Structures with exactly one computable isomorphism type are called **computably categorical**. All the known standard examples of computably categorical structures are usually prime models (of their own theories) or become prime in expansions by finitely many constants. For example, finitely generated computable algebras, the rational numbers under the natural ordering, finite dimensional vector spaces over computable fields and the ring of integers are computably categorical. There are also pathological examples of computably categorical structures that fail to satisfy certain natural properties (for example, existence of Scott families) exhibited by most computably categorical structures [2, 6, 8]. One notes that these specifically constructed computably categorical structures fail to be prime models in expansions by finitely many constants. In

fact, the theories of these specifically constructed structures do not have saturated models due to the fact that the theories have uncountably many types.

We need some of the main definitions from computable model theory. The languages and structures we consider are all countable unless we tell otherwise. We also assume that we are working in a **recursive** language, that is, that the arity of each relation and function symbol in the language is uniformly recursive. A structure is said to be **computable** if it has universe ω and the open diagram of the structure forms a recursive set of formulas. A structure is **computably presentable** if it has a computable isomorphic copy; such a copy is called a **computable presentation of the structure**.

Two isomorphic computable structures are called **computably isomorphic** if there is a computable isomorphism between these two structures. A computable structure is said to be **computably categorical** if any two computable presentations of the structure are computably isomorphic. As mentioned above, finitely generated computable algebras, the rational numbers under the natural ordering, finite dimensional vector spaces over computable fields are computably categorical. In this paper we are mostly interested in computably categorical structures.

The concepts of computable structure and computable categoricity can naturally be extended as follows. A structure is said to be a Σ_1^0-**structure** if it has universe ω and the open positive diagram of the structure, that is the set of all open formulas without negations true in the structure, forms a recursively enumerable set of formulas. A structure is Σ_1^0-**presentable** if it is isomorphic to a Σ_1^0-structure. We stress that in the definition of a Σ_1^0-structure it is explicit that the domain of the structure is ω. A Σ_1^0-structure is **computably categorical** if any two Σ_1^0-presentations of the structure are computably isomorphic. Clearly, every computable structure is also a Σ_1^0-structure. Therefore if a computable structure is computably categorical when one considers Σ_1^0-presentations, then all Σ_1^0-presentations of the structure must be computable. Similarly, if a non-computable Σ_1^0-structure is computably categorical then the structure does not have a computable presentation. We note, however, that we deal with computably categorical Σ_1^0-structures in Section 4 only.

We briefly recall some basic notions and facts from model theory. Let T be a complete theory. A countable model of T is said to be **prime** if every tuple of the model realizes a principal type. It is well-known that every prime model is elementarily embedded into all models of T. A countable model of T is said to be **saturated** if the model realizes all the types of the theory in all possible expansions of T by finitely many constants. It is well known that T has a saturated model if and only if T has at most countably many countable models. Moreover, every other countable model of T is elementarily embedded into the saturated model (see Hodges' or Marker's textbooks on model theory [4, 5, 10]). An important model-theoretic property of prime and saturated models of a given theory T, in case they exist, is that they are unique up to isomorphisms. We address this uniqueness property of the saturated models from a computability-theoretic point. Finally, we call a structure saturated (prime) if it is the saturated model (the prime model) of its own theory.

Above we have already provided examples of computably categorical prime models. It is also not hard to have natural examples of prime but not computably categorical models. For instance, the natural numbers with their order is the prime and non-computably categorical model. However, in spite the fact that saturated models (of a given theory) form one isomorphism type, all the known computable saturated models are not computably categorical. Roughly, the main reason for them not being computably categorical hides in using the following model-theoretic fact. There exists an infinite sequence $p(x_0) \subset p(x_0, x_1) \subset \ldots \subset p(x_0, x_1, \ldots, x_n) \subset \ldots$ of non-principal types such that each type in the sequence is realized by infinitely many elements of the model. A simple example here is the theory of vector spaces over the field of rational numbers. The non-saturated models of the theory are all finite dimensional vector spaces. They are all computably categorical. The saturated model is the infinite dimensional vector space. The saturated model is, however, not computably categorical. This is because there are two computable copies of the infinite dimensional vector space such that in one copy the dependency problem is computable and in the other is not. Similarly, all the non-saturated models of the theory of one successor are computably categorical while the saturated model of the theory is not. The reason for this is that the saturated model has two computable presentations such that in one the algebraic dependency relation is computable and in the other is not.

Here is the outline of the paper. In the next section we construct a specific uniformly recursively enumerable family $\{B_x\}_{x \in \omega}$ of subsets of natural numbers. The family is defined based on the notion of Kolmogorov complexity. We also prove that some special enumerations of the family are equivalent to each other via computable permutations.

In Section 3, we provide an example of a saturated not $\aleph_0$-categorical structure with exactly one computable isomorphism type. The idea is to code the family $\{B_x\}_{x \in \omega}$ from the previous section into a saturated structure so that the computable copies of the structure induce the special enumerations of the family. The construction of the structure is based on the well-known model theoretic construction of Fraïssé limits.

In Section 4 of the paper we also address a question of Goncharov that asks if there exists an $\aleph_1$-categorical but not $\aleph_0$-categorical saturated model that has unique computable isomorphism type. Recall that a first order theory T is called κ-**categorical** (for an infinite cardinal κ) if $\mathcal{A} \cong \mathcal{B}$ whenever $\mathcal{A}, \mathcal{B} \models T$ and $|\mathcal{A}| = |\mathcal{B}| = \kappa$. Morley [11] proved that if T is categorical in some uncountably cardinality, then it is categorical in all uncountable cardinalities. Therefore, any $\aleph_1$-categorical theory has a unique model of each uncountable cardinality. We partially answer the question of Goncharov positively by providing a saturated $\aleph_1$-categorical but not $\aleph_0$-categorical saturated Σ_1^0-structure that has unique computable isomorphism type. Unfortunately, the structure is not computable. Our constructions codes the family $\{B_x\}_{x \in \omega}$.

In Section 4, we provide an alternative proof of the main result in [7]. There an $\aleph_1$-categorical but not $\aleph_0$-categorical theory T is constructed such that all models of T but the prime one are computable. Our construction is again based on coding the family $\{B_x\}_{x \in \omega}$ into an $\aleph_1$-categorical but not $\aleph_0$-categorical theory.

2 The Role of Kolmogorov Complexity

The first result is the construction of an auxiliary family of recursively enumerable sets $B_0, B_1, \ldots$ with the following properties:

- the finite members of this family occur only once each;

- all infinite members are equal and occur infinitely often in every recursive enumeration of the family.

The construction goes as follows. Let U be a universal partial-recursive function in the sense that for every further partial-recursive function ψ there is a constant c, such that for all x in the domain of ψ there is a $y \leqslant c(x+1)$ with $U(y) = \psi(x)$. Then the Kolmogorov complexity C (based on U) of any number z is defined as

$$C(z) = \min\{\log(x) : U(x) \downarrow = z\}$$

where in this paper the logarithm $\log(x)$ is defined as the smallest natural number y with $2^y \geqslant x$. The rationale behind this definition is that it should roughly invert exponentiation, have the base 2 and avoid undefined places, proper fractions and irrational numbers. Note that $C(z) \geqslant \log(z)$ for infinitely many z. The reader should consult standard textbooks [1, 9, 12, 13] for more information on Kolmogorov Complexity and Recursion Theory. The family in question is now defined as follows.

Definition 2.1. Let $A = \{x : C(x) < \log(x)\}$ be the set of compressible or non-random numbers. Define

$$B_x = \begin{cases} \{x\} \cup \{y \in A : y < \log(x)\} & \text{if } x \notin A; \\ A & \text{if } x \in A. \end{cases}$$

The family $B_0, B_1, \ldots$ is uniformly recursively enumerable as $\{x\} \cup \{y \in A : y < \log(x)\} \subseteq B_x$ for all x and the set A is recursively enumerable. So a uniform enumeration of the B_x starts with an enumeration of $\{x\} \cup \{y \in A : y < \log(x)\}$ and later adds all other elements of A in the case that x turns out to be an element of A.

Theorem 2.2. If $E_0, E_1, \ldots$ is a recursive enumeration of r.e. sets such that for every x there is y with $E_x = B_y$, and for every $y \notin A$ there is a unique x with $B_y = E_x$, then there is a computable permutation f with $B_y = E_{f(y)}$ for all y.

Proof. For every $y \notin A$ there is a unique x with $y \in E_x$. As every $y \in A$ satisfies $y \in E_x$ for almost all x, one has that there are infinitely many x with $y \in E_x$. Thus there is a computable function g such that $y \in E_{g(y)}$ for all y, this function can be obtained by searching in parallel in all $E_0, E_1, \ldots$ until an x with $y \in E_x$ is found.

If $y \notin A$ then $E_{g(y)} = B_y$ as B_y is the unique set in the enumeration $B_0, B_1, \ldots$ containing y. Thus $E_{g(y)}$ has to be equal to B_y. If $y \in A$ and $E_{g(y)} = A$ then $B_y = E_{g(y)}$. If $y \in A$ but $E_{g(y)} = B_x$ for some $x \notin A$ then $y < \log(x)$ and x is the unique element of $E_{g(y)}$, which

4

is larger than $\log(x)$. Thus one can compute x from y and $\log(x)$. As Cantor's pairing function is invertible, one can compute x also from $(y + \log(x))(y + \log(x) + 1)/2 + \log(x)$. As $y < \log(x)$, the logarithm of this expression is roughly $2 \log \log(x)$. So, on one hand, there is a constant c_1 with $C(x) \leqslant 2 \log \log(x) + c_1$. On the other hand, $\log(x) \leqslant C(x)$. Therefore, $\log(x) \leqslant 2 \log \log(x) + c_1$ and hence there are only finitely many such x. Thus it follows that g, for all but finitely many y, satisfies $E_{g(y)} = B_y$. Since the errors occur on finitely many $y \in A$, the other elements of A are mapped to an E_x with $E_x = A$. By modifying g at finitely many places, one even obtains $\forall y \, (E_{g(y)} = B_y)$. Note that A occurs in the enumeration $E_0, E_1, \ldots$ infinitely often as A is not computable and cannot be of the form $g^{-1}(D)$ for any finite set D.

Now let I be an infinite computable subset of A which includes all y satisfying the condition $g(y) \in \{g(z) : z < y\}$. Furthermore, let x be any index. If $E_x = B_y$ for an $y \notin A$, then one can compute y from x by inverting g and thus $C(y) \leqslant C(x) + c_2$ for some constant c_2. As y is incompressible, $\log(y) \leqslant C(x) + c_2 \leqslant \log(x) + c_3$ for some constant c_3. Thus one knows that whenever $x \neq g(y)$ for all y with $\log(y) \leqslant \log(x) + c_3$ then $E_x = A$. So the set of all x with $E_x = A$ is recursively enumerable: $E_x = A$ iff either $g(y) = x$ for some $y \in A$ or $g(y) \neq x$ for the finitely many y with $\log(y) \leqslant \log(x) + c_3$. In particular there is an infinite computable set J such that $E_x = A$ for all $x \in J$ and $J \cup g(\mathbb{N}) = \mathbb{N}$.

Now define $f(y)$ to be the n-th element of J whenever y is the n-th element of the computable set $I \cup g^{-1}(J)$ and let $f(y) = g(y)$ otherwise. Again $E_{f(y)} = B_y$ for all y as f coincides with g on those y where B_y is finite while f is modified from one index of A to another one in the case that $y \in I \cup g^{-1}(J)$. Furthermore, by construction, f is a permutation. It is also easy to see that f is computable. $\square$

3 The first application

Our main result in this section is the following theorem.

Theorem 3.1. *There exists a saturated not $\aleph_0$-categorical model that has a unique computable isomorphism type.*

We need to recall the construction of Fraïssé limits. Let $\mathcal{K}$ be a class of finite structures closed under isomorphisms. We assume the language of structures is finite and contains only relational symbols. Assume that the class $\mathcal{K}$ has the following properties:

1. Hereditary property (*HP*): for all $A \in \mathcal{K}$, if B is a substructure of A then $B \in \mathcal{K}$.

2. Joint embedding property (*JEP*): for all $A, B \in \mathcal{K}$ there exists a $C \in \mathcal{K}$, such that A and B can be embedded into C.

3. Amalgamation property (*AP*): for all $A, B, C \in \mathcal{K}$, if $f \colon A \to C$ and $g \colon A \to B$ are embeddings then there exists a structure $D \in \mathcal{K}$ and embeddings $h \colon B \to D$ and $k \colon C \to D$ such that $hf = kg$ on A.

A structure D is called **weakly homogeneous** if it has the property

if A, B are finite substructures of D, $A \subseteq B$ and $f \colon A \to D$ is an embedding, then there is an embedding $g \colon B \to D$ which extends f.

A structure is called **ultrahomogeneous** if any finite partial isomorphism of the structure into itself can be extended to an automorphism. A finite or countable structure is ultrahomogeneous if and only if it is weakly homogeneous (see Lemma 7.1.4(b) in [4]).

The age of a structure D is the class of all finite structures embeddable in D. There is a well known result in model theory that connects the ultrahomogeneous structures with classes $\mathcal{K}$ that possess properties HP, JEP and AP. It is stated in the following theorem:

Theorem 3.2. *For any class $\mathcal{K}$ that has HP, JEP and AP there exists a unique at most countable ultrahomogeneous structure $\lim(\mathcal{K})$ whose age coincides with $\mathcal{K}$. Moreover, the structure $\lim(\mathcal{K})$ is $\aleph_0$-categorical.*

The structure $\lim(\mathcal{K})$ is called the **Fraïssé limit** of the class $\mathcal{K}$. We restate this theorem with an eye towards computable categoricity:

Theorem 3.3. *Let $\mathcal{K}$ be a class finite structures with the following properties.*

1. *$\mathcal{K}$ has properties HP, JEP and AP.*

2. *There exists a sequence $H_0, H_1, \ldots$ such that*

   - *$\{H_n \mid n \in \omega\} \subseteq \mathcal{K}$ and each $A \in \mathcal{K}$ is isomorphic to some H_n;*
   - *the domain and the atomic diagram of H_n is computable uniformly in n;*
   - *the function $n \mapsto |H_n|$ is computable.*

Then the Fraïssé limit of $\mathcal{K}$ is a computably categorical structure.

Proof. First, show that the Fraïssé limit of $\mathcal{K}$ is computable. Let $\{(A_i, B_i)\}_{i \in \omega}$ be a recursive enumeration of pairs of structures from $\mathcal{K}$ such that $A_i \subseteq B_i$, and for every pair $A, B \in \mathcal{K}$ such that $A \subseteq B$ there exist i and an isomorphism $f \colon B \to B_i$ such that $f(A) = A_i$. We now construct a computable structure D as follows. Let $D_0 = H_0$. Suppose that D_k has been constructed. Applying the AP property the required number of times one can show that there is an isomorphic copy H_n' of some $H_n \in \mathcal{K}$ such that $D_k \subseteq H_n'$ and for all $i \leqslant k$, if A_i can be embedded in D_k then for every embedding $f \colon A_i \to D_k$, there is an embedding $g \colon B_i \to H_n'$ extending f. For every n, we can effectively check whether there is a copy of H_n satisfying the condition above. So let D_{k+1} be an isomorphic copy of H_n with minimal index n satisfying that condition.

Now consider a computable structure $D = \bigcup_{k < \omega} D_k$. Since each D_k is in $\mathcal{K}$ and $\mathcal{K}$ possesses HP property, the age of D is included in $\mathcal{K}$. Suppose A is in $\mathcal{K}$; then by JEP there are $B \in \mathcal{K}$, such that $D_0 \subseteq B$, and an embedding $h \colon A \to B$. Now let a pair (A_i, B_i) be such that there is an isomorphism $f \colon B_i \to B$ with $f(A_i) = D_0$. By construction, the embedding $f \restriction A_i \colon A_i \to D_0$ extends to an embedding $g \colon B_i \to D$. Hence both B and A are in the age of D. Therefore, the age of D is exactly $\mathcal{K}$.

Let $A \subseteq B$ be finite substructures of D and $h: A \to D$ be an embedding. Since $A, B \in \mathcal{K}$, there are a pair (A_i, B_i) and an isomorphism $f: B_i \to B$ with $f(A_i) = A$. Furthermore, there is $k \geqslant i$ such that $hf \upharpoonright A_i$ is an embedding of A_i into D_k. By construction, $hf \upharpoonright A_i$ extends to an embedding $g: B_i \to D_{k+1}$. Now $gf^{-1}: B \to D$ is an embedding that extends h. This proves that D is weakly homogeneous and, therefore, ultrahomogeneous. Thus D is the Fraïssé limit of the class $\mathcal{K}$.

We now show that D is computably categorical. Let D' be a computable structure isomorphic to D; then there is a computable chain $\{D'_k\}_{k<\omega}$ of finite structures such that $D' = \bigcup_{k<\omega} D'_k$. We construct a computable isomorphism from D to D' as follows. Let f_0 be an embedding of D_0 into D'. Suppose that a finite partial embedding f_n has been constructed. If $n = 2m$ then look for the smallest $k \geqslant m$, such that $\mathrm{Dom}(f_n) \subseteq D_k$. Since D' is weakly homogeneous, there is an embedding $g: D_k \to D'$ that extends f_n and that can be found effectively. So let $f_{n+1} = g$. If $n = 2m+1$ then look for the smallest $k \geqslant m$, such that $\mathrm{Im}(f_n) \subseteq D'_k$. Since D is weakly homogeneous, there is an embedding $g: D'_k \to D$ that extends f_n^{-1}. So let $f_{n+1} = g^{-1}$. Thus $f = \bigcup_{n<\omega} f_n: D \to D'$ is a computable isomorphism. $\qquad\square$

Now we define special classes of finite structures that have properties *HP*, *JEP* and *AP*. A **cycle** of length $n \geqslant 3$ is the graph $C_n = (\{1, \ldots, n\}, E)$ with $E = \{(1,2), (2,1), (2,3), (3,2),$ $\ldots, (n-1, n), (n, n-1), (n, 1), (1, n)\}$. We say that a graph **contains** a cycle of length n if there exists an embedding from C_n into the graph.

Let Y be a non-empty subset of natural numbers. Consider the following class of finite directed graphs:

$$\mathcal{K}(Y) = \{(V, E) \mid \text{ If } (V, E) \text{ contains a cycle of length } n+3 \text{ then } n \in Y\}.$$

Lemma 3.4. *The class $\mathcal{K}(Y)$ possesses properties HP, JEP and AP.*

Proof. It is easy to see that $\mathcal{K}(Y)$ satisfies properties *HP* and *JEP*. We prove that $\mathcal{K}(Y)$ satisfies *AP*. Let A, B, C be graphs in $\mathcal{K}(Y)$ such that A is a subgraph of B and C and the domain of A is the intersection of the domains of B and C. Define the graph D as follows. The domain of D is the union of domains of B and C. The graph D contains all the edges of the graphs B and C. In addition, D contains all the edges of the form (b, c), where $b \in B \setminus A$ and $c \in C \setminus A$. It is not hard to see that the graph D built in this way belongs to $\mathcal{K}(Y)$. $\qquad\square$

Now we construct the desired structure C_ω as follows. To do this we use the family $\{B_x\}_{x\in\omega}$ from the previous section. For each B_x consider the limit structure $\lim \mathcal{K}(B_x)$. One can construct a sequence

$$\lim \mathcal{K}(B_0), \lim \mathcal{K}(B_1), \lim \mathcal{K}(B_2), \ldots$$

of these structures so that the following properties hold:

1. the graphs in this sequence are all pairwise disjoint;

2. the union of domains of these graphs is ω;

3. the sequence is uniformly computable meaning that the set $\{(n, m) \mid m \in \lim \mathcal{K}(B_n)\}$ is computable.

The signature of C_ω consists of two binary relational symbols R and S. The domain of C_ω is ω. The relation R is the union of all edges of graphs that appear in the sequence above. The relation S consists of all pairs (n, m) such that n, m belong to the same graph $\lim \mathcal{K}(B_x)$ for some x. Clearly S is a computable equivalence relation. Thus, the structure C_ω constructed is computable. Our goal now is to show that C_ω satisfies the theorem stated in the beginning of this section.

Lemma 3.5. *The structure C_ω is computably categorical.*

Proof. Let D be any computable structure isomorphic to C_ω. Since the equivalence relation S in D is computable, there is a computable sequence $\{x_i\}_{i \in \omega}$, which consists of exactly one representative for each S-equivalence-class. Let E_i be a set such that the substructure of D with domain $[x_i]_S$, S-equivalence-class of x_i, is isomorphic to $\lim \mathcal{K}(E_i)$. Using the fact that D is computable, one can show that the sequence $\{E_i\}_{i \in \omega}$ is uniformly recursively enumerable. Furthermore, for every x there is y with $E_x = B_y$, and for every $y \notin A$ there is a unique x with $B_y = E_x$. Thus, by Theorem 2.2, there is a computable permutation f such that $B_i = E_{f(i)}$ for all i. By Theorem 3.3, $\lim \mathcal{K}(B_i)$ is a computably categorical structure. Note that the construction of the computable isomorphism between $\lim \mathcal{K}(B_i)$ and $\lim \mathcal{K}(E_{f(i)})$ can be done uniformly in i. Therefore, D is computably isomorphic to C_ω. $\square$

Let $T = \text{Th}(C_\omega)$ be the first-order theory of C_ω. Our goal is to show that C_ω is the saturated model of T. This is proved in the Lemma 3.6 below that also characterizes the isomorphism types of the models of T. Call an S-equivalence-class **non-standard** if the restriction of R to this class is isomorphic to the Fraïssé limit $\lim \mathcal{K}(A)$. Consider a subsequence

$$\lim \mathcal{K}(B_{n_0}), \lim \mathcal{K}(B_{n_1}), \lim \mathcal{K}(B_{n_2}), \ldots$$

of the sequence

$$\lim \mathcal{K}(B_0), \lim \mathcal{K}(B_1), \lim \mathcal{K}(B_2), \ldots,$$

where n_1, n_2, $\ldots$ is the list of all numbers outside of A in the increasing order. Consider the substructure of C_ω restricted to the subsequence above and denote it by C_0. Let C_n be the structure obtained by adjoining to C_0 exactly n copies of non-standard S-equivalence-classes.

Lemma 3.6. *The theory T satisfies the following properties.*

1. *C_0 is the prime model of T.*

2. *The class of all countable models of T is $\{C_0, C_1, C_2, \ldots, C_\omega\}$.*

3. *C_ω is the saturated model of T.*

Proof. Let us write down the axioms for T. First, note that the fact that x and y lie in the same component $\lim \mathcal{K}(B_n)$ of C_ω can be expressed by a first-order formula. Indeed, let $x, y \in \lim \mathcal{K}(B_n)$ for some n. Suppose that there is no edge from x to y and from y to x. Let B be the substructure of $\lim \mathcal{K}(B_n)$ with domain $\{x, y\}$. Let D be a graph with domain $\{x, y, z\}$ that extends B and contains additional edges (x, z), (z, y). Note that D is in $\mathcal{K}(B_n)$ since it does not contain any cycle. By the weakly homogeneity of $\lim \mathcal{K}(B_n)$, there is an embedding of D into $\lim \mathcal{K}(B_n)$ that extends the identity map on B. Therefore, we can express the fact that x, y belong to the same $\lim \mathcal{K}(B_n)$ by the formula

$$\varphi(x, y) = R(x, y) \vee R(y, x) \vee \exists z (R(x, z) \wedge R(z, y)).$$

We use the notation $\{\bar{c}\}$ for the set consisting of elements of the tuple $\bar{c}$. Let $\psi^n(x_0, \ldots, x_{n-1})$ be a formula such that for any graph B

$$B \models \psi^n(\bar{b}) \quad \text{if and only if} \quad \{\bar{b}\} \text{ is a cycle of length } n \text{ in } B.$$

For any graph B and n-tuple of distinct elements $\bar{b}$ such that $B = \{\bar{b}\}$ let $\psi_{B,\bar{b}}(x_0, \ldots, x_{n-1})$ be a conjunction of formulas $R(x_i, x_j)$ or $\neg R(x_i, x_j)$ satisfied by $\bar{b}$ in B. Thus for any graph D and a tuple $\bar{d} \in D$ of the same length as $\bar{b}$,

$$D \models \psi_{B,\bar{b}}(\bar{d}) \quad \text{iff} \quad \text{there is an isomorphism from } B \text{ to } \{\bar{d}\} \text{ which takes } \bar{b} \text{ to } \bar{d}.$$

Let $S_n(x)$ be a formula that says the S-equivalence-class of x contains a cycle of length $n + 3$, that is

$$S_n(x) = \exists \bar{y} \left(\psi^{n+3}(\bar{y}) \wedge \bigwedge_{i \leqslant n-1} S(x, y_i) \right).$$

We also use an abbreviation $\bar{x} \in S_n$ for a formula $\bigwedge_{i \leqslant n-1} S_n(x_i)$ and $\bar{x} \in [z]$ for a formula $\bigwedge_{i \leqslant n-1} S(x_i, z)$, where $\bar{x} = x_0, \ldots, x_{n-1}$. Let U be the following list of axioms.

(Ax^0) S is an equivalence relation.

(Ax^1) $S(x, y) \rightarrow R(x, y) \vee R(y, x) \vee \exists z (R(x, z) \wedge R(z, y)).$

For every n:

(Ax_n^2) $\neg S(x, y) \rightarrow \neg \exists x_0, \ldots, x_{n+1} \left(x_0 = x \wedge x_{n+1} = y \wedge \bigwedge_{i \leqslant n} (R(x_i, x_{i+1}) \vee R(x_{i+1}, x_i)) \right).$

For every $n \notin A$:

(Ax_n^3) $\exists x (S_n(x) \wedge \forall y (S_n(y) \rightarrow S(x, y))).$

For every $n \notin A$, every $B, D \in \mathcal{K}(B_n)$ and every tuple $\bar{b}d$ of distinct elements such that $B = \{\bar{b}\}$ and $D = \{\bar{b}d\}$:

$(Ax_{n,B,D,\bar{b}d}^4)$ $(\forall \bar{x} \in S_n) (\psi_{B,\bar{b}}(\bar{x}) \rightarrow (\exists y \in S_n) \, \psi_{D,\bar{b}d}(\bar{x}, y)).$

If $\bar{b}$ is empty then this sentence reduces to $(\exists y \in S_n) \, \psi_{D,d}(y).$

For every $n \notin A$:

9

$(\mathrm{A}x_n^5)$ $(\forall \bar{x} \in S_n) \bigvee\limits_{B,\bar{b}} \psi_{B,\bar{b}}(\bar{x}),$

where the disjunction ranges over all pairs $B, \bar{b}$ such that $B \in \mathcal{K}(B_n)$ and $\bar{b}$ is a tuple of the same length as $\bar{x}$ with $B = \{\bar{b}\}$. Note that this disjunction is finite.

For every $B, D \in \mathcal{K}(A)$ and every tuple $\bar{b}d$ of distinct elements such that $B = \{\bar{b}\}$ and $D = \{\bar{b}d\}$:

$(\mathrm{A}x_{B,D,\bar{b}d}^6)$ $\forall z \left(\left(\bigwedge\limits_{i \leqslant k-1} \neg S_{n_i}(z) \right) \rightarrow (\forall \bar{x} \in [z]) \, (\psi_{B,\bar{b}}(\bar{x}) \rightarrow (\exists y \in [z]) \, \psi_{D,\bar{b}d}(\bar{x}, y)) \right),$

where $n_0, \ldots, n_{k-1}$ are the indices of all components $\lim \mathcal{K}(B_n)$ of C_ω into which D can not be embedded. Note that $n_i \notin A$ for all $i \leqslant k-1$. If $\bar{b}$ is empty then this sentence reduces to

$$\forall z \left(\left(\bigwedge\limits_{i \leqslant k-1} \neg S_{n_i}(z) \right) \rightarrow (\exists y \in [z]) \, \psi_{D,d}(y) \right).$$

Now let M be a countable model of U. Axioms $\mathrm{A}x^0$, $\mathrm{A}x^1$ and $\mathrm{A}x_n^2$ imply that S is an equivalence relation and that every S-equivalence-class is a connected component of M. For every $n \notin A$, $\mathrm{A}x_n^3$ states that there is a unique component that contains a cycle of length n. Denote this component by M_n.

When $\bar{b}$ is empty, $\mathrm{A}x_n^4$ says that every one-element structure in $\mathcal{K}(B_n)$ is embeddable in M_n. In general, $\mathrm{A}x_n^4$ says that

> if B, D are finite structures in $\mathcal{K}(B_n)$, D comes from B by adding one more element and $f \colon B \rightarrow M_n$ is an embedding, then there is an embedding $g \colon D \rightarrow M_n$ which extends f.

Now, using induction on the number of elements, it is not hard to see that every structure in $\mathcal{K}(B_n)$ is embeddable in M_n. On the other hand, $\mathrm{A}x_n^5$ implies that any finite substructure of M_n is in $\mathcal{K}(B_n)$. Thus, the age of M_n is exactly $\mathcal{K}(B_n)$.

Using $\mathrm{A}x_n^4$ again and an induction on the size of $D \setminus B$, we can show that

> if $B, D \in \mathcal{K}(B_n)$, $B \subseteq D$ and $f \colon B \rightarrow M_n$ is an embedding, then there is an embedding $g \colon D \rightarrow M_n$ which extends f.

Thus M_n is a weakly homogeneous (hence ultrahomogeneous) model of age $\mathcal{K}(B_n)$. Therefore, M_n is isomorphic to $\lim \mathcal{K}(B_n)$. Note that, in particular, it means that $M_n \neq M_k$ whenever $n \neq k$.

Let M^* be a connected component of M that is different from all M_n's. As shown above, any cycle of length n, for $n \notin A$, can appear only in M_n. So the age of M^* is included in $\mathcal{K}(A)$. Let D be one-element structure in $\mathcal{K}(A)$; then $\mathrm{A}x^6$ implies that D is embeddable in M^*. Now let $B, D \in \mathcal{K}(A)$, D comes from B by adding one more element and $f \colon B \rightarrow M^*$ is an embedding. Suppose that D is not embeddable in $\lim \mathcal{K}(B_{n_0}), \ldots, \lim \mathcal{K}(B_{n_{k-1}})$, where $n_i \notin A$ for all $i \leqslant k-1$. In this case $\mathrm{A}x^6$ states that the embedding of B into any component of M

other than $M_{n_0}, \ldots, M_{n_{k-1}}$ can be extended to an embedding of D into the same component. In particular, f can be extended to an embedding $g \colon D \to M^*$. Now, using induction on the size of D, it is not hard to show that every $D \in \mathcal{K}(A)$ is embeddable in M^*. Thus the age of M^* is exactly $\mathcal{K}(A)$.

Again, an induction on the size of $D \setminus B$ tells us that if $B, D \in \mathcal{K}(A)$, $B \subseteq D$ and $f \colon B \to M^*$ is an embedding, then there is an embedding $g \colon D \to M^*$ which extends f. Thus M^* is a weakly homogeneous structure of the age $\mathcal{K}(A)$ and, therefore, is isomorphic to $\lim \mathcal{K}(A)$.

So any countable model of U consists of exactly one component isomorphic to $\lim \mathcal{K}(B_n)$, for $n \notin A$, and finite or infinite number of components isomorphic to $\lim \mathcal{K}(A)$. In other words, the class of all countable models of U is $\{C_0, C_1, C_2, \ldots, C_\omega\}$.

We now show that, for every $i \in \omega$, C_i is elementary equivalent to C_ω. To do this, we will use the method of Ehrenfeucht–Fraïssé games.

Definition 3.7. Let A, B be the structures of the same language L and let γ be an ordinal. Then $EF_\gamma[A, B]$, the **unnested Ehrenfeucht–Fraïssé game of length γ on A and B**, is defined as follows. There are two players $\forall$ and $\exists$. The game is played in γ steps. At the ith step of the play, player $\forall$ takes one of the structures A, B and chooses an element of this structure; then player $\exists$ chooses an element of the other structure. Each player is allowed to see and remember all previous moves in the play. At the end of the play, sequences $\bar{a} = (a_i : i < \gamma) \in A$ and $\bar{b} = (b_i : i < \gamma) \in B$ have been chosen. The pair $(\bar{a}, \bar{b})$ is known as the **play**. We say that player $\exists$ **wins the play** $(\bar{a}, \bar{b})$ iff

for every unnested atomic formula φ of L, $A \models \varphi(\bar{a}) \Leftrightarrow B \models \varphi(\bar{b})$.

Note that if the language L contains no function symbols or constants, as in our case, then every formula of L is unnested. We write $A \approx_\gamma B$ to mean that player $\exists$ has a winning strategy for the game $EF_\gamma[A, B]$.

Theorem 3.8. *Let L be a finite first-order language. Then for any two L-structures A and B the following are equivalent.*

(I) *$A \equiv B$.*

(II) *For every $k < \omega$, $A \approx_k B$.*

For the proof and more details, see chapters 3.2 and 3.3 in Hodges [4]. So let us fix any i and k and show that player $\exists$ has a winning strategy for the game $EF_k[C_i, C_\omega]$.

The strategy for player $\exists$ that is described below has the following property. In the beginning of every step s, the sequences $\bar{a}_{s-1} = (a_0, \ldots, a_{s-1})$, $\bar{b}_{s-1} = (b_0, \ldots, b_{s-1})$ have been chosen by players $\forall$ and $\exists$ such that the substructures $\{\bar{a}_{s-1}\}$ and $\{\bar{b}_{s-1}\}$ of C_i and C_ω respectively are isomorphic via isomorphism that maps $\bar{a}_{s-1}$ to $\bar{b}_{s-1}$. Moreover, for every $t < s$, if $a_t \in \lim \mathcal{K}(B_n)$ and $b_t \in \lim \mathcal{K}(B_m)$, then either

(a) $n \notin A$, $\log(n) \leqslant k - 3$ and $n = m$,   or

11

(b) $n \notin A \Rightarrow k - 3 < \log(n)$ and $m \notin A \Rightarrow k - 3 < \log(m)$.

In both cases, we have that $B_n \cap [0, k-3] = B_m \cap [0, k-3]$ and, therefore, any substructure of $\lim \mathcal{K}(B_n)$ with at most k elements is embeddable in $\lim \mathcal{K}(B_m)$ and vice versa.

Now suppose that at step s player $\forall$ has chosen $a_s \in C_i$. Then player $\exists$ chooses $b_s \in C_\omega$ such that the substructures $\{\bar{a}_s\}$ and $\{\bar{b}_s\}$ are isomorphic via isomorphism that maps $\bar{a}_s$ to $\bar{b}_s$. Also, if a_s lies in the same component as a_t for some $t < s$, then b_s has to be in the same component as b_t. The fact that $|\bar{a}_s| \leqslant k$ ensures that player $\exists$ can always find such b_s. If the component of a_s does not contain any a_t for $t < s$, then player $\exists$ chooses b_s such that b_s is not in any component containing any b_t for $t < s$ and that property (a) or (b) given above holds for the pair a_s, b_s. Obviously, player $\exists$ can always find such a component, because there are infinitely many components in both C_i and C_ω that are isomorphic either to $\lim \mathcal{K}(A)$ or to $\lim \mathcal{K}(B_n)$ for $n \notin A$ and $k < \log(n)$.

The case when player $\forall$ has chosen $b_s \in C_\omega$ is similar to the above.

Now it is not hard to see that this strategy is indeed a winning strategy for player $\exists$. Thus all the structures $C_0, C_1, \ldots, C_\omega$ are elementary equivalent to each other and U is the set of axioms for $T = \mathrm{Th}(C_\omega)$. Since T has countably many countable models, T has countable saturated model and a prime model (see e.g. Corollary 4.3.8 in Marker [10]). None of $\{C_i : i < \omega\}$ can be a saturated model, because C_ω is not embeddable in any C_i. So we can conclude that C_ω is in fact the saturated model of T. $\qquad\square$

4 The second application

Here we partially answer the question of Goncharov about the existence of an $\aleph_1$-categorical but not $\aleph_0$-categorical structure with a unique computable isomorphism type. Here our answer is affirmative if one considers Σ_1^0-structures rather than computable structures. Here is the result.

Theorem 4.1. *There is a $\aleph_1$-categorical but not $\aleph_0$-categorical theory whose saturated model is a computably categorical Σ_1^0-structure.*

Proof. We use the family $\{B_x\}_{x \in \omega}$ constructed in Section 2. The language of the desired structure is given by the family $P_0, P_1, \ldots$ of unary predicates. Define the following structure M. The domain of the structure is the set of natural numbers. For each x, $P_y(x)$ holds iff $y \in B_x$. Obviously, the structure is Σ_1^0.

Let T be the first-order theory of M. It can be described by the following set of axioms.

For every $i \notin A$:

$(\mathrm{A}x_i^1)$ There is a unique z such that $P_i(z)$.

For $i \notin A$, let c_i be a new constant interpreted as an element on which P_i holds. The axioms $\mathrm{A}x_i^1$ imply that these constants are definable in the original language.

For every $i \in A$:

$$(\mathrm{A}x_i^2) \quad \bigwedge_{\{j:\ i \notin B_j\}} \neg P_i(c_j) \ \wedge \ \forall z\Big(\Big(\bigwedge_{\{j:\ i \notin B_j\}} z \neq c_j\Big) \to P_i(z)\Big),$$

that is P_i holds almost everywhere and does not hold only on constants c_j for $i \notin B_j$. Note that if $i \in A$, then the set $\{j : i \notin B_j\}$ is finite; thus the conjunctions in $\mathrm{A}x_i^2$ are finite, and this is the first order formula.

The theory T is not $\aleph_0$-categorical since the prime model is given by the substructure of M with domain $\mathbb{N} - A$. The theory T is $\aleph_1$-categorical since any model of cardinality $\kappa > \aleph_0$ consists of the following elements:

- one element x with $B = \{n \mid P_n(x)\}$, for every finite set B in the enumeration $\{B_y\}_{y \in \omega}$;

- κ many elements x with $A = \{n \mid P_n(x)\}$.

The statement about Σ_1^0-categoricity can be derived by considering any further Σ_1^0-model M' with domain $\mathbb{N}$ that is isomorphic to M; such a model defines a computable enumeration $E_0, E_1, \dots$ with $n \in E_y \Leftrightarrow P_n(y)$. It is easy to see that every B_x with $x \notin A$ appears in this enumeration only once and that every E_y equals some B_x. By Theorem 2.2, there exists a computable permutation f such that $E_{f(x)} = B_x$ for all x. This computable permutation clearly induces a computable isomorphism between two Σ_1^0-structures M and M'. $\qquad \square$

5   The third application

Khoussainov, Nies and Shore [7] give an example of $\aleph_1$-categorical but not $\aleph_0$-categorical theory such that all models of the theory but the prime model are computable. In this section we provide an alternative proof of this result using the family $\{B_x\}$ constructed in the second section.

To code this type of family, we generalize the notion of "cubes" introduced by Khoussainov, Nies and Shore [7]. Fix a language L consisting of binary relation symbols F_n (for $n \in \omega$), all of which we will assume to be symmetric and irreflexive relation coding the edges of a hypercube. These cubes are constructed from given r.e. parameters sets X and have dimension $|X|$; for each element n of the set X, the edges along one dimension are realized by the relation F_n. More formally, this is done as follows.

For any subset X of ω, let the domain of the default presentation of an X-cube be a set $\{\sum_{m \in Y} 2^m : Y \subseteq X \wedge Y \text{ is finite}\}$ with $\sum_{m \in \emptyset} 2^m = 0$. On this domain, define the relation $F_n(x, y)$ to be true iff there is a finite subset $Y \subseteq X - \{n\}$ such that

$$\{x, y\} = \{\textstyle\sum_{m \in Y} 2^m, \sum_{m \in Y \cup \{n\}} 2^m\}.$$

An X-cube is then a structure isomorphic to the default presentation that we have just defined. Note that every default presentation of an X-cube is uniformly Σ_1^0 in X.

For example, a $\{0, 1, 3\}$-cube is an isomorphic copy of $\{0, 1, 2, 3, 8, 9, 10, 11\}$ together with the relations $F_0(0, 1)$, $F_0(2, 3)$, $F_0(8, 9)$, $F_0(10, 11)$, $F_1(0, 2)$, $F_1(1, 3)$, $F_1(8, 10)$, $F_1(9, 11)$, $F_3(0, 8)$, $F_3(1, 9)$, $F_3(2, 10)$, $F_3(3, 11)$. Other relations than these do not hold between the members of the $\{0, 1, 3\}$-cube.

Alternatively, an X-cube can be defined as follows (for example, see [3]). Consider $A = \mathbb{Z}_2^{|X|}$ as a vector space over $\mathbb{Z}_2$, with basis $\{a_0, \ldots, a_{|X|-1}\}$. Let $f \colon X \to |X|$ be a bijection. If, for every $n \in X$ and every $x, y \in A$, we define $F_n(x, y) \Leftrightarrow x + a_{f(n)} = y$, then every structure isomorphic to A will be an X-cube.

Definition 5.1. Let the set A and the sequence $B_0, B_1, \ldots$ of sets as in Definition 2.1. Let C_0 be the disjoint union of all those B_x-cubes with $x \notin A$. Furthermore, let C_n be the disjoint union of C_0 and n many A-cubes for $n \in \{1, 2, 3, \ldots, \omega\}$.

Note that C_ω is isomorphic to the disjoint union over all B_x-cubes for $x \in \omega$.

Proposition 5.2. *The structures $C_0, C_1, C_2, \ldots, C_\omega$ have all the same theory T; this theory is $\aleph_1$-categorical, C_0 is its prime model and C_ω is its saturated model. The models $C_0, C_1, C_2, \ldots, C_\omega$ are the only countable models of T up to isomorphism.*

To prove this proposition we can consider the theory $T = \mathrm{Th}(C_\omega)$ and formulate the list U of its axioms. One group of axioms should say that the models of T consist of cubes. Furthermore, for every $x \in A$, we need an axiom saying that F_x is adjacent to all but n_x nodes, where n_x is the number of nodes in C_ω that are not adjacent to F_x. Similarly, for every $x \notin A$, we need an axiom saying that F_x is adjacent only to n_x many nodes, where n_x is the number of nodes in C_ω that are adjacent to F_x. Now one can show that $C_0, C_1, C_2, \ldots, C_\omega$ are all countable models of U and all models of U of the same uncountable cardinality are isomorphic. This implies that the theory determined by U is complete and U is indeed the set of axioms for T. More details can be found in the proof of Theorem 4.1 in [3].

Theorem 5.3. *The models $C_1, C_2, \ldots, C_\omega$ have all computable presentations but C_0 does not have a computable presentation.*

Proof. This is clear for C_ω as there is a recursive one-one enumeration $(a_0, b_0), (a_1, b_1), \ldots$ of pairs such that for each x the set $\{a : \exists s \, [a_s = a \wedge b_s = x]\}$ is the domain of the default presentation of the B_x-cube from above. Then one defines for the given presentation that $F_n(s, t)$ holds iff $b_s = b_t$ and $F_n(a_s, a_t)$ in the default presentation of the B_{b_s}-cube. It is easy to see that the given model is computable and isomorphic to C_ω.

We now describe how to construct a computable presentation for C_1. Fix some $x_0 \in A$ and start the construction by enumerating all B_x-cubes in some effective way. Also we start enumerating the set A. When at some stage s a number x is enumerated into A_s, we expand the finite part of B_{x_0}-cube constructed so far and merge the finite part of B_x-cube into it. To do this we might need to use new edges F_t with $t \in A$ such that up to stage s it has not been decided for any two nodes a, b whether $F_t(a, b)$ holds or not. So we keep on enumerating A until we have enough of such edges. To build a computable presentation for C_n with $2 \leqslant n < \omega$ one need to add $n - 1$ computable copies of an A-cube to the computable presentation of C_1.

Now assume by contradiction that the prime model C_0 also has a computable presentation. Then there is a computable function mapping every n to a triple $(a(n), b(n), y(n))$ such that $2^n < y(n) \wedge F_{y(n)}(a(n), b(n))$. This function is total as there are infinitely many $x > 2^n$ such that $x \notin A$ and a copy of B_x is merged into C_0. Let $x(n)$ denote the index of the $B_{x(n)}$-cube to which $a(n)$ belongs. Note that $B_{x(n)}$ is finite as C_0 is the prime model. Note that $x(n) \notin A$, $x(n) \geqslant y(n)$ and $C(x(n)) \geqslant \log(x(n))$. Now $x(n)$ can be computed from n and $\log(x(n))$ as $x(n)$ is the only number z larger than $\log(x(n))$ such that there is an d with $F_z(a(n), d)$; so $x(n)$ can be found by exhausting search once $\log(x(n))$ and n are given. As $n \leqslant \log(x(n))$, $C(x(n)) \leqslant C(\log(x(n)), n) + c \leqslant 2 \log \log(x(n)) + c'$ for constants c, c' and all n. Combining the two conditions, one has $\log(x(n)) \leqslant 2 \log \log(x(n)) + c'$ for all n. But this is impossible as $x(n) > 2^n$ for all n. Thus C_0 cannot have a computable presentation. $\square$

References

[1] Cristian Calude, *Information and randomness: an algorithmic perspective*, Springer, Berlin, 2002.

[2] Peter Cholak, Richard Shore and Reed Solomon, A computably stable structure with no Scott family of finitary formulas, *Archive of Mathematical Logic*, to appear.

[3] Denis Hirschfeldt, Bakhadyr Khoussainov and Pavel Semukhin, An uncountably categorical theory whose only computably presentable model is saturated, *Notre Dame Journal of Formal Logic*, 47:63–71, 2006.

[4] Wilfrid Hodges, *Model Theory*, Cambridge University Press, 1993.

[5] Wilfrid Hodges, *A Shorter Model Theory*, Cambridge University Press, 1997.

[6] Oleg V. Kudinov, An autostable 1-decidable model without a computable Scott family of $\exists$-formulas, Algebra and Logic 35:255-260, 1996 (English translation).

[7] Bakhadyr Khoussainov, André Nies and Richard Shore, Computable models of theories with few models, *Notre Dame Journal of Formal Logic*, 38:165-178, 1997.

[8] Bakhadyr Khoussainov and Richard Shore, Computable isomorphisms, degree spectra of relations and Scott families, *Annals of Pure and Applied Logic*, 91:1–15, 1998 and 98:297–298, 1999.

[9] Ming Li and Paul Vitányi, *An introduction to Kolmogorov complexity and its applications*, Texts and Monographs in Computer Science, Springer, New York, 1993.

[10] David Marker, *Model theory: an introduction*, Springer, New York, 2002.

[11] Michael Morley, Categoricity in power, *Transactions of the American Mathematical Society*, 114:514–538, 1965.

[12] Piergiorgio Odifreddi, *Classical Recursion Theory*, North-Holland, Amsterdam, 1989.

[13] Robert I. Soare, *Recursively enumerable sets and degrees*, Springer, Heidelberg, 1987.