On generic complexity of theories of finite algebraic structures

Alexander Rybalov
Sobolev Institute of Mathematics, prospekt Koptyuga 4, Novosibirsk, 630090, Russia
E-mail: alexander_rybalov@mail.ru

Abstract. This article is devoted to investigation of generic complexity of universal and elementary theories of finite algebraic structures with universal set with more than one element of finite predicate signature with equality. We prove that there are no polynomial strongly generic algorithm, recognizing any such theory, provided \( P = \text{BPP} \) and \( P \neq \text{NP} \) (\( P \neq \text{PSPACE} \)). The author is supported by Russian Science Foundation, grant 19-11-00209.

1. Introduction
In the XX century, due to the fast development of computer science and applied mathematics, came to the fore in the investigation of finite combinatorial and algebraic objects. For example, finite graphs find numerous applications for solving practical problems related to networks, routes, object classification, etc. Another example is finite fields, without which it is inconceivable modern cryptography and the theory of error-correcting codes. The classical approaches to the investigation of finite objects are algebraic and combinatorial approaches. New approach — logical came from so-called universal algebraic geometry [1]. Within the framework of this approach, the studied objects are considered as algebraic systems in the appropriate language (signature). Many practically important problems about finite objects can be formulated in terms of solution of systems of equations over the corresponding algebraic systems, which leads to the need for the development of algebraic geometry.

Universal and elementary theories are closely related to algebraic geometry over algebraic structures. Decidability and computational complexity of these theories are very important for both practical and theoretical applications. But usually, the problem of recognizing these theories for finite algebraic systems is computationally hard (\( \text{co-NP-complete} \) and \( \text{PSPACE-complete} \)) in the worst case. Considering hardness of this problem in the worst case complexity, we will use the generic approach [3], which considers an algorithmic problem for almost all inputs, and ignores negligible set of the remaining inputs [6, 7, 8, 9, 10, 11, 12, 13, 14]. So the problem of recognizing every such theory is hard and, under the conditions \( P \neq \text{NP} \) and \( P \neq \text{PSPACE} \), there is no polynomial algorithm that recognizes it for all inputs.

This article is devoted to investigation of generic complexity of universal and elementary theories of finite algebraic structures with universal set with more than one element of finite predicate signature with equality. We prove that there are no polynomial strongly generic algorithm, recognizing any such theory, provided \( P = \text{BPP} \) and \( P \neq \text{NP} \) (\( P \neq \text{PSPACE} \)). The

\( ^1 \) Supported by Russian Science Foundation, grant 19-11-00209.
class BPP is the class of algorithmic problems, solvable by probabilistic polynomial algorithms. Although this equality not yet proven, there are strong arguments towards it [2].

2. Preliminaries

Let us recall some definitions from mathematical logic and model theory [1]. Signature is a set σ, consisting of predicate symbols, functional symbols and constant symbols. Algebraic structure of signature σ is a set A = ⟨A, σ⟩, where A is a non-empty set, and to each predicate symbol of σ some predicate is associated on the set A, to each functional symbol of σ some function on A is associated, and to each constant symbol of σ some element from A is associated.

First-order logic formula of signature σ, in which each variable is under some quantifier, is called a sentence. The fact that a sentence Φ of signature σ is true in algebraic structure A = ⟨A, σ⟩ is denoted as A |= Φ. A sentence Φ is called universal (or ∀-sentence), if it has a form

Φ = ∀x1...∀xnϕ(x1,...,xn),

where ϕ is a formula of signature σ without quantifiers. Theory of algebraic structure A is the set Th(A) of all true in A sentences of signature σ. The set of all ∀-sentences of theory Th(A) is called universal theory Th∀(A) of algebraic structure A.

We will use the Catalan numbers C_n, which are defined as

C_n = \frac{1}{n+1} \binom{2n}{n},

where \binom{n}{k} is the binomial coefficient.

Lemma 1. For n > m it holds

\frac{C_{n-m}}{C_n} > \frac{1}{4^m}.

Proof. We can bound:

\frac{C_{n-m}}{C_n} = \frac{n+1}{n-m+1} \times \frac{\binom{2(n-m)}{n-m}}{\binom{2n}{n}} = 

= \frac{n+1}{n-m+1} \times \frac{(2(n-m))!}{(n-m)! (n-m)!} > \frac{(2(n-m))!n!n!}{(n-m)! (n-m)! (2n)!} = 

= \frac{n!}{(n-m)!} \times \frac{2(n-m)\ldots(n-m+1)}{2n\ldots(n+1)} = 

= \frac{n(n-1)\ldots(n-m+1)2(n-m)\ldots(n-m+1)}{2n\ldots(n+1)} = 

= \frac{(n\ldots(n-m+1))^2}{2n\ldots(2(n-m)+1)} > \left(\frac{(n-1)\ldots(n-m)}{2(n-1)\ldots(2n-2m)}\right)^2 > \frac{1}{2^{2m}} = \frac{1}{4^m}.

3. Generic computability and complexity

Suppose W is the set of all possible inputs. Denote by W_n the set of all inputs in W of size n. For any set S ⊆ W consider the following limit

\mu(S) = \lim_{n \to \infty} \mu_n(S) = \lim_{n \to \infty} \frac{|S \cap W_n|}{|W_n|}.

We will call the set S generic if it holds \mu(S) = 1 and we will call it negligible if it holds \mu(S) = 0. Also S will be strongly generic (negligible) if the sequence \{\mu_n(S) : n = 1, 2, 3, \ldots\} converges exponentially fast to 1 (to 0).

An algorithm A : W → U ∪ {?} is called (strongly) generic if
(i) \( \mathcal{A} \) terminates \( \forall x \in W \),
(ii) set \( \{ x \in W : \mathcal{A}(x) = ? \} \) is (strongly) negligible.

We say that (strongly) generic algorithm \( \mathcal{A} : W \to \{0, 1, ?\} \) decides a problem of membership to a set \( S \subseteq W \) if
\[
\forall x \in W \; \mathcal{A}(x) = yes \Rightarrow x \in S.
\]
\[
\forall x \in W \; \mathcal{A}(x) = no \Rightarrow x \notin S.
\]

4. Representation of logical formulas

Fix a finite predicate signature
\[
\sigma = \{ P_1^{(a_1)}, \ldots, P_k^{(a_k)}, c_1, \ldots, c_l \},
\]
where \( P_i \) are predicate (including the equality predicate) and \( c_i \) are constants. Define
\[
A = \max_{i=1,...,k} \{a_i\}.
\]

Suppose that an universal sentence \( \Phi \) of \( \sigma \) has a form:
\[
\Phi = \forall x_1 \ldots \forall x_t \phi,
\]
where \( \phi \) is a formula without quantifiers, constructed by disjunctions and conjunctions from atom formulas of type \( P_i(x_1, \ldots, x_{a_i}) \) or its negations. We can represent formula \( \phi \) as a binary tree \( T_\phi \) with internal vertices marked by \( \lor \) and \( \land \), and with leaves labeled by atomic formulas or its negations. Let us assume that all variables of \( T_\phi \) are from the list \( x_1, \ldots, x_{An} \).

We define the size of the sentence \( \Phi \) as the number of leafs in \( T_\phi \). Assume that sentence \( \Phi \) of size \( n \) depends on all variables \( \{x_1, \ldots, x_{An}\} \) and all these variables are under quantifiers.

Arbitrary sentence (not only universal) \( \Phi \)
\[
\Phi = Q_1 x_1 \ldots Q_t x_t \phi,
\]
can be represented in the similar way, its representations is \( T_\phi \) and prefix of quantifiers \( Q_1 \ldots Q_t \).

The set of all universal sentences, represented by aforementioned manner, is denoted by \( \mathcal{U}\mathcal{F} \). The set of all sentences, represented by aforementioned manner, is denoted by \( \mathcal{F} \).

**Lemma 2.** Fix a size \( n \). The number of all universal sentences of size \( n \) is
\[
|\mathcal{U}\mathcal{F}_n| = 2^{n-1}C_{n-1}\left(2 \sum_{i=1}^{k} (An + l)^n\right)^n.
\]

**Proof.** Each universal sentence from \( \mathcal{U}\mathcal{F}_n \) consists of the universal prefix of quantifiers and a tree with \( n-1 \) internal vertices and \( n \) leafs. It is known (see [4]), that there exist \( C_{n-1} \) unmarked binary trees with \( n \) leafs. We can mark by symbols \( \lor \) or \( \land \) each internal veritix. Therefore there are \( 2^{n-1} \) labeling. Each of \( n \) leafs can be labeled by one of \( a_i \)-arity predicates \( P_i, i = 1, \ldots, k \), or by its negation. We can substitute one of variables \( x_j, j = 1, \ldots, An \), or one of constants \( c_j, j = 1, \ldots, l \) in predicate \( P_i, i = 1, \ldots, k \). So we have
\[
|\mathcal{U}\mathcal{F}_n| = 2^{n-1}C_{n-1}\left(2 \sum_{i=1}^{k} (An + l)^n\right)^n.
\]
Lemma 3. For any $n$ the number of sentences of size $n$ is

$$|F_n| = 2^{2n-1}C_{n-1} \left( \sum_{i=1}^{k} (A^n + l)^n \right)^n.$$

Proof. Analogically to the proof of Lemma 2, considering that there are $2^n$ variants to select a prefix of quantifiers for $n$ variables.

Define for every universal sentence $\Phi = \forall x_1 \ldots \forall x_t \phi$ the following set

$$cl(\Phi) = \{ \forall x_1 \ldots \forall x_n (\phi \lor ((x_1 \neq x_1) \land \psi)) \},$$

where $n > t$ and $\psi$ is an arbitrary quantifier-free formula on variables $x_1, \ldots, x_n$. Observe that all sentences from $cl(\Phi)$ are equivalent $\Phi$.

Lemma 4. For every universal sentence $\Phi$ we have

$$\frac{|cl(\Phi) \cap UF_n|}{|UF_n|} > \frac{1}{(16k(A^n + l)^A)^{m+2}}$$

for all $n > m + 2$, where sentence $\Phi$ has size $m$.

Proof. Assume that the size of sentence $\Psi$ is equal to $m$. Consider an arbitrary sentence from the set $cl(\Phi)$ of size $n > m + 2$. Note that the number of vertices in the tree $T_\Psi$ of $\Psi$ is $n - m - 2$. Similarly to the proof of Lemma 2, we may count

$$|cl(\Phi) \cap UF_n| = 2^{n-m-3}C_{n-m-3} \left( \sum_{i=1}^{k} (A^n + l)^n \right)^{n-m-2}.$$

Now

$$\frac{|cl(\Phi) \cap UF_n|}{|UF_n|} = \frac{2^{n-m-3}C_{n-m-3} \left( \sum_{i=1}^{k} (A^n + l)^n \right)^{n-m-2}}{2^{n-1}C_{n-1} \left( \sum_{i=1}^{k} (A^n + l)^n \right)^n} = \frac{1}{2^{m+2} \left( \sum_{i=1}^{k} (A^n + l)^n \right)^{m+2} \frac{C_{n-m-3}}{C_{n-1}}} > \frac{1}{(16k(A^n + l)^A)^{m+2}}.$$

Here we used Lemma 1 to bound the relation of Catalan numbers.

For each sentence $\Phi = Q_1x_1 \ldots Q_tx_t\phi$ define the following set

$$cl(\Phi) = \{ Q_1x_1 \ldots Q_t x_t Q_{t+1} x_{t+1} \ldots Q_n x_n (\phi \lor ((x_1 \neq x_1) \land \psi)) \},$$

where $n > t$, $\psi$ is an arbitrary quantifier-free formula on variables $x_1, \ldots, x_n$, and $Q_{t+1}, \ldots, Q_n$ are arbitrary quantifiers. Again all sentences from $cl(\Phi)$ are equivalent $\Phi$.

Lemma 5. For every sentence $\Phi$ we have

$$\frac{|cl(\Phi) \cap F_n|}{|F_n|} > \frac{1}{(32k(A^n + l)^A)^{m+2}}$$

for all $n > m + 2$, where sentence $\Phi$ has size $m$.

Proof. Analogically to the proof of Lemma 4, using Lemma 2.
5. Main results

**Theorem 1.** Let $\mathfrak{A}$ be a finite algebraic structure of finite predicate signature with more one element in universe set. Suppose there is a strongly generic polynomial algorithm, recognizing $Th(V(\mathfrak{A}))$. There exists a polynomial probabilistic algorithm, recognizing $Th(V(\mathfrak{A}))$ on the whole set of sentences.

*Proof.* Assume that there is a polynomial strongly generic algorithm $A$, recognizing $Th(V(\mathfrak{A}))$.

Then the set

$$G(A) = \{ \Phi \in \mathcal{U}F : A(\Phi) \neq ? \}$$

is strongly generic.

Design a probabilistic algorithm $B$, working in polynomial time, and deciding every universal sentence $\Phi$. It runs on any sentence $\Phi$ of size $n$ in the following manner

(i) Generate a random universal sentence $\Psi$ from the set $cl(\Phi)$ of size $n^2$.

(ii) Run $A$ on $(\Psi)$.

(iii) If $A(\Psi) \neq ?$, then determinate the truth of $\Phi$.

(iv) If $A(\Psi) = ?$, then gives the answer NO.

Observe that algorithm $B$ produces a correct answer at the step 3, but can make a mistake at step 4. We want to prove that the probability of output at the step 4 is $< 1/2$.

The probability that a random universal $\Psi$ in $cl(\Phi)$ does not hit in set $G(A)$ is not greater than

$$\frac{|(\mathcal{U}F \setminus G(A))_{n^2}|}{|cl(\Phi)_{n^2}|} \leq \frac{|(\mathcal{U}F \setminus G(A))_{n^2}|}{|\mathcal{U}F_{n^2}|} \times \frac{|\mathcal{U}F_{n^2}|}{|cl(\Phi)_{n^2}|}.$$

The set $G(A)$ is strongly generic, so there exists a constant $\beta > 0$ for which

$$\frac{|(\mathcal{U}F \setminus G(A))_{n^2}|}{|\mathcal{U}F_{n^2}|} < \frac{1}{2^{\beta n^2}}$$

for each $n$. By Lemma 4

$$\frac{|\mathcal{U}F_{n^2}|}{|cl(\Phi)_{n^2}|} < (16k(A n^2 + l)^A)^{n+2}.$$

Therefore, the desired probability is not greater than

$$\frac{(16k(A n^2 + l)^A)^{n+2}}{2^{3n^2}} = \frac{2^{(n+2) \log (16k(A n^2 + l)^A)}}{2^{3n^2}}$$

and is less than $1/2$ for large numbers $n$. This means that the probability 4 is $< 1/2$.

The polynomiality of the described algorithm follows from the existence of procedure for generation of random binary tree of size $N$ within time, bounded by a polynomial in $N$. This procedure is described in [5].

This Theorem implies:

**Corollary 1.** Let $\mathfrak{A}$ be a finite algebraic structure of finite predicate signature with more one element in universe set. Suppose $P \neq NP$ and $P = BPP$. Then there is no polynomial strongly generic algorithm, recognizing $Th(V(\mathfrak{A}))$.

Analogically to Theorem 1, using Lemma 5, we can prove

**Theorem 2.** Let $\mathfrak{A}$ be a finite algebraic structure of finite predicate signature with more one element in universe set. Assume, that there is a strongly generic polynomial algorithm, recognizing $Th(\mathfrak{A})$. Then there is a probabilistic polynomial algorithm, recognizing $Th(\mathfrak{A})$ on the whole set of sentences.
From Theorem 2 it follows

**Corollary 2.** Let $\mathfrak{A}$ be a finite algebraic structure of finite predicate signature with more one element in universe set. Suppose, that $P = \text{BPP}$ and $P \neq \text{PSPACE}$. Then there is no polynomial strongly generic algorithm, recognizing $Th(\mathfrak{A})$.

**References**

[1] Daniyarova E Y and Myasnikov A G and Remeslennikov V N 2016 Algebraic geometry over algebraic structures Novosibirsk: SB RAS 288 p

[2] Impagliazzo R and Wigderson A 1997 P=BPP unless E has subexponential circuits: derandomizing the XOR Lemma Proceedings of the 29th STOC pp 220–229

[3] Kapovich I and Myasnikov A and Schupp P and Shpilrain V 2003 Generic-case complexity, decision problems in group theory and random walks Journal of Algebra 264 (2) pp 665–694

[4] Knuth D 1997 The Art of Computer Programming, Addison-Wesley 725 p

[5] Rybalov A 2016 O genericheskoi slozhnosti problemy proverki istinnosti bulevyh formul Prikladnaya Diskretnaya Matematika 32 (2) pp 119–126

[6] Rybalov A 2019 O genericheskoi nerazreshimosti desyatoi problemy Gilberta dlya polinomialnyh dereviev Prikladnaya Diskretnaya Matematika 44 107–112

[7] Rybalov A 2019 O genericheskoi slozhnosti problemy klerstizacii grafov Prikladnaya Diskretnaya Matematika 46 72–77

[8] Rybalov A 2020 O genericheskoi NP-polnote problemy vypolnimosti bulevых shem Prikladnaya Diskretnaya Matematika 47 101–107

[9] Rybalov A 2020 O genericheskoi slozhnosti problemy predstavimosti naturalnych chisel summov dvuh kvadratov Prikladnaya Diskretnaya Matematika 48 93–99

[10] Rybalov A 2020 O genericheskoi slozhnosti ekzistencialnyh teorii Prikladnaya Diskretnaya Matematika 49 120–126

[11] Rybalov A 2020 O genericheskoi slozhnosti poblemu o summe podmnozhestv dlya polugrupp celyh matric Prikladnaya Diskretnaya Matematika 50 118–126

[12] Rybalov A 2020 On the generic existential theory of finite graphs Siberian Electronic Mathematical Reports 17 1710–1714

[13] Rybalov A 2020 Genericheskii algoritm dlya problemy vhozhdeniya v polugruppah celochislennyh matric Vestnik Omskovo Gosudarstvennogo Universiteta 25 (3) 8–12

[14] Rybalov A 2020 O genericheskoi slozhnosti problemy o summe podmnozhestv dlya moindov i grupp celochislennyh matric vtorogo poryadka Vestnik Omskovo Gosudarstvennogo Universiteta 25 (4) 10–15