On the Bounded Theories of Finite Trees

Sergei Vorobyov

Max-Planck-Institut für Informatik
Im Stadtwald, D-66123, Saarbrücken, Germany (e-mail: sv@mpi-sb.mpg.de)

Abstract. The *theory of finite trees* is the full first-order theory of equality in the Herbrand universum (the set of ground terms) over a functional signature containing non-unary function symbols and constants. Albeit decidable, this theory turns out to be of *non-elementary complexity* [14].

To overcome the intractability of the theory of finite trees, we introduce in this paper the *bounded theory of finite trees*. This theory replaces the usual equality $=$, interpreted as identity, with the infinite family of *approximate equalities* "down to a fixed given depth" $\{=^d\}_{d\in\omega}$, with d written in binary notation, and $s =^d t$ meaning that the ground terms s and t coincide if all their branches longer than d are cut off.

By using a refinement of Ferrante-Rackoff's complexity-tailored Ehrenfeucht-Fraïssé games, we demonstrate that the bounded theory of finite trees can be decided within *linear double exponential space* $2^{2^{cn}}$ (n is the length of input) for some constant $c > 0$.

1 Introduction

Tree-like structures are fundamental for almost all domains of Computer Science, and especially relevant to logic programming, symbolic computation, data types, constraint solving, automated theorem proving, data bases, knowledge representation. Whenever the reasoning about a class of data structures is involved, it is interesting to know what is the *inherent computational complexity* of this reasoning. This may be crucial in practical implementations of theorem provers, constraint solvers, systems of logic programming.

The first-order theory of *finite trees*, also known as the theory of *term algebras*, or *Clark's equational theory*, although *decidable* [9, 7, 8, 4], turns out to be *non-elementary in the sense of Kalmar* [14]. Any nondeterministic decision procedure for the theory takes time exceeding infinitely often any fixed tower of exponents $2^{2^{\cdot^{\cdot^{2^n}}}}$, where n is the length of input.

In this paper we suggest a practical substitute for the theory of finite trees, which we call the *bounded theory of finite trees*. In this theory, instead of the unique usual equality $=$, one has an infinite family of equalities $\{=^d\}_{d\in\omega}$, with $s =^d t$ interpreted as true if and only if the trees s and t coincide to depth d, where d is written in ≥ 2-ary. Thus instead of claiming for the complete equality, one has to specify *explicitly* which precision is needed in every comparison. We demonstrate that the bounded theory is decidable within *elementary space* $2^{2^{cn}}$ for some $c > 0$, and thus can be considered a useful practical alternative to the usual (unbounded) non-elementary recursive theory of finite trees.

Finite trees is one of the basic domains in the Constraint Logic Programming [5]. One can hardly expect to use the full first-order theory of trees to express constraints, because of its non-elementary complexity. Allowing only existential quantification (as is usually done) seems to be a serious restriction of expressiveness. In this respect the *bounded* theory of finite trees, allowing for the full first-order quantification and being elementary, may be considered useful.

Venkataraman in [13] showed that the first-order theory of finite trees with the *subtree predicate* is *undecidable*. By using the machinery of this paper we can show that the bounded theory of trees with the "*to be a subtree at bounded depth*" predicate is decidable in elementary space and time.

Outline. After briefly surveying the standard theory of finite trees we introduce the approximate tree equality and the bounded theory of trees in functional and relational formalizations, and state our Main Theorem in the end of Section 3. In Section 4 we explain Ferrante-Rackoff's complexity-tailored refinement of the Ehrenfeucht-Fraïssé games and extend it for infinite signatures. In Section 5 we settle the upper complexity bounds for the bounded theory of trees.

Preliminaries. We suppose familiarity with standard logical notation. By ω we denote the set of natural numbers. A signature Σ is called *functional* iff it contains no predicate symbols. $Const(\Sigma)$ and $Fun(\Sigma)$ denote the subsets of constant and non-nullary function symbols of Σ respectively. $T(\Sigma)$ denotes the set of all ground (variable-free) terms of signature Σ, usually called the *Herbrand universum over* Σ; $ar(f)$ is the *arity* of $f \in \Sigma$.

First-order formulas, free and bound occurrences, substitutions are defined as usual. A *sentence* or *closed formula* is a formula without free variables. The *quantifier depth* of a formula ϕ is a maximal number of nested quantifiers in ϕ.

First-order models and their carriers are denoted by A, B. The elements of models are denoted by a, b, possibly with indices; $\overline{a}_k$, $\overline{b}_k$ denote k-tuples of elements $a_1 \dots a_k$, $b_1 \dots b_k$. For example, $\overline{a}_{k+1} = a_1 \dots a_k, a_{k+1} = \overline{a}_k, a_{k+1}$. By $\overline{x}_k$ we denote a k-tuple of distinct variables. By $(A, \overline{a}_k)$ we denote a model A with distinguished elements $\overline{a}_k$. The satisfaction relation $\models$ is defined as usual.

2 Theory of Finite Trees

Global Proviso. Throughout the paper Σ denotes a *finite* functional signature containing *at least one* constant symbol. Hence $T(\Sigma) \neq \emptyset$. □

Definition 1 (Theory of Finite Trees). The *theory of finite trees* is the full first-order theory $Th(T(\Sigma))$ of the Herbrand universum $T(\Sigma)$ in the language of the first-order predicate calculus of signature Σ *with equality.* □

The good well-known news is that the theory is decidable.

Theorem 2 (Mal'cev-Kunen-Maher-Hodges [9, 7, 8, 4]). *Both for finite and infinite signatures the theory of finite trees possesses* complete axiomatizations*; therefore is* decidable*.* □

The quantifier elimination procedures for the theory of finite trees are described in [9, 7, 8, 4]. The bad news is that the decision problem for the theory is *computationally intractable.*

Definition 3 (Iterated Exponentials). For $m, n \in \omega$ let $\exp_0(n) = n$ and $\exp_{m+1}(n) = 2^{\exp_m(n)}$. Define $\exp_\infty(n)$ as $\exp_n(0)$. A decision problem is *elementary in the sense of Kalmar* iff it can be decided within space (or time) bounded by a function $\exp_m(n)$ for some *fixed* $m \in \omega$, where n is the length of input. Otherwise, a problem is called *non-elementary.* □

It turns out that the theory of finite trees is *non-elementary.* This disproves K. Kunen's claim [6] that the theory of finite trees is *PSPACE*-complete:

Theorem 4 ([14]). *The first-order theory of finite trees is* non-elementary *if the signature Σ (finite or infinite) contains non-unary function symbols. Moreover, any decision algorithm for the theory takes time* exceeding infinitely often $\exp_\infty(\lfloor cn \rfloor)$ *for some $c > 0$, where n is the length of input.* □

The same applies to variations of the theory, like the theories of *rational* and *feature* trees (for the definitions of these theories see, e.g., [8, 1, 10]).

3 Approximate Equality and Bounded Theories of Trees

As a partial remedy to overcome the intractability of the theory of finite trees, we introduce the *approximate tree equality* and the *bounded theory of finite trees.*

One of the reasons of the high complexity of the theory of finite trees is as follows: given two pointers to two random constant terms of signature Σ, there is *no upper bound* on the complexity of their comparison. The approximate equality $=^d$ defined below has such a bound (exponential in d).

Definition 5 (Approximate Equality). For $d \in \omega$ define the *approximate equality relations* $=^d$ on $T(\Sigma) \times T(\Sigma)$ inductively as follows:

- $s =^0 t$ iff $s \equiv f(s_1, \ldots, s_m)$, $t \equiv f(t_1, \ldots, t_m)$ for some $f \in \Sigma$;
- $s =^{d+1} t$ iff $s \equiv f(s_1, \ldots, s_m)$, $t \equiv f(t_1, \ldots, t_m)$, and $s_j =^d t_j$ $(1 \le j \le m)$.

Definition 6 ((Functional) Bounded Theories of Finite Trees). Denote by $\Sigma_=$ the signature $\Sigma \cup \{=^d\}_{d \in \omega}$ *without usual equality* $=$. Let $\mathcal{F}^f_{bnd}(\Sigma)$ be the set of all first-order formulas of signature $\Sigma_=$ *without equality* $=$. The *functional bounded theory of finite trees* $Th^f_{bnd}(T(\Sigma))$ is the set of all sentences of $\mathcal{F}^f_{bnd}(\Sigma)$ true in the Herbrand universum $T(\Sigma)$. □

The bounded theory is *different* from the usual one: in the usual theory one has $\forall x \neg(x = t(x))$ for any term $t(x)$ containing x *properly.* In the bounded theory one may have $\neg\forall x \neg(x =^d t(x))$, e.g., $s^{1997}(0) =^{1996} s^{2000}(0)$. In this respect the bounded theory is closer to the theory of *rational* trees.

By a simple reduction to the theory of finite trees we get the following

Proposition 7. *For any finite functional signature Σ the functional bounded theory of trees $Th^f_{bnd}(T(\Sigma))$ is decidable.* □

The reduction to the theory of finite trees suggests only a *very ineffective way* to decide $Th^f_{bnd}(T(\Sigma))$, because the target theory of finite trees is of *non-elementary* complexity. In this paper we describe a much more efficient procedure to decide the theory $Th^f_{bnd}(T(\Sigma))$, which runs in *elementary space (hence time)*.

Since playing Ehrenfeucht-Fraïssé-games is much easier without function symbols it is convenient to get rid of all constant and function symbols.

Definition 8 (Companion Relational Signature). For a signature $\Sigma_= = \Sigma \cup \{=^d\}_{d\in\omega}$, where Σ is a finite functional signature, let the *companion relational signature* $\widehat{\Sigma}_=$ contain:

1. a unary predicate symbol Is_c for every constant symbol $c \in \Sigma$;
2. binary predicate symbols f^d_p for all $d \in \omega$, $f \in \Sigma$, and $1 \le p \le ar(f)$;
3. binary predicate symbols $=^d$ for every $d \in \omega$.

The upper indices ^{d} in the predicate symbols f^d_p and $=^d$ are called *ranks*. □

Definition 9 (Canonical Relational Model of Trees). For a finite functional signature Σ define the *canonical relational model of the bounded theory of trees* $\mathcal{M} \equiv \langle T(\Sigma);\ \widehat{\Sigma}_=\rangle$ with the Herbrand universum $T(\Sigma)$ as a carrier, of signature $\widehat{\Sigma}_=$, the relational companion to $\Sigma_=$, as follows:

- for $d \in \omega$ the meaning of $=^d$ is given by Definition 5;
- for $s \in T(\Sigma)$ one has $\mathcal{M} \models Is_c(s)$ if and only if $T(\Sigma) \models s =^0 c$;
- for $s, t \in T(\Sigma)$ and $1 \le p \le ar(f)$ one has $\mathcal{M} \models f^d_p(s,t)$ if and only if $T(\Sigma) \models \exists x_1 \ldots x_{p-1} x_{p+1} \ldots x_{ar(f)} \Big(s =^d f(x_1, \ldots, x_{p-1}, t, x_{p+1}, \ldots, x_{ar(f)}) \Big)$.

Hence, instead of $y =^d f(x_1 \ldots x_k)$ we may write $\wedge^k_{i=1} f^d(y, x_i)$.

Definition 10 ((Relational) Bounded Theory of Trees). Given a finite functional signature Σ with constants, denote by $\mathcal{F}^p_{bnd}(\Sigma)$ the set of all first-order formulas of the companion relational signature $\widehat{\Sigma}_=$ *without usual equality*. The *relational bounded theory of trees* $Th^p_{bnd}(T(\Sigma))$ is the full first-order theory of the canonical relational model $\mathcal{M} \equiv \langle T(\Sigma);\ \widehat{\Sigma}_=\rangle$ in the first-order language of signature $\widehat{\Sigma}_=$ without equality. □

There is no essential difference between functional and relational theories.

Proposition 11. *The functional and the relational bounded theories of trees are* definitionally equivalent*, see [4].* □

The decision complexity of the bounded theory of trees is determined by the number of quantifiers in the prenex form of a formula (see Section 5). It is the same for both theories:

Proposition 12. *An arbitrary formula of length n of $Th^f_{bnd}(T(\Sigma))$ can be transformed into an equivalent prenex formula of $Th^p_{bnd}(T(\Sigma))$ with $O(n)$ quantifiers.*

Main Theorem. *For any finite functional signature Σ, the bounded theory of finite trees over Σ (both functional or relational) can be decided within space $2^{2^{cn}}$ for some constant $c > 0$, where n is the length of input.*

If the signature contains function symbols of arity at most 1, then the bounded theory of trees can be decided within space 2^{cn} for some constant $c > 0$.

If the signature has only constant symbols then the bounded theory of trees can be decided within polynomial space and is PSPACE-complete if Σ contains ≥ 2 constants. □

4 Ferrante-Rackoff's Games for Complexity Analysis

In the next section we prove our Main Theorem by applying Ferrante-Rackoff's games described in Section 2 of [3]. We have to spend additional effort to make these games applicable to *infinite signatures.* This is necessary because companion relational signatures (Definition 8) are *always infinite*, whereas original Ferrante-Rackoff's games apply to finite signatures only. We attain the needed generalization by relativizing Ferrante-Rackoff's boundedness conditions to finite subsignatures and by proving that the games carry over with this modification.

Ferrante-Rackoff's complexity-tailored games [3] refine Ehrenfeucht-Fraïssé-games [2, 4] by additional *boundedness* analysis in the *back-and-forth* conditions. Boundedness means that whenever a formula of the form $\exists x\, \Phi(x)$ is true, one can always find a *small* witness for $\Phi(x)$ from a *finite* subset of a model. Contrapositively, if there are no small witnesses for $\Phi(x)$, one may safely consider $\exists x\, \Phi(x)$ false. Thus, assuming boundedness, to decide $\exists x\, \Phi(x)$, one just needs to check *finitely many small candidates for witnesses.* This forms the basis of the decision method. We carry over this machinery to the case of infinite signatures.

4.1 Modification of Ferrante-Rackoff Games for Infinite Signatures

Although the extension of Ferrante-Rackoff games to infinite signatures can be done in full generality, for the lack of space we develop it here (Theorems 15 and 17) *only* for the bounded theories of trees. We also have to omit proofs. All this appears in the full paper [15].

Definition 13. For $D \in \omega$ denote by $\widehat{\Sigma}_{=}^{D}$ the finite subsignature

$$\{Is_c \ \mid\ Is_c \in \widehat{\Sigma}_{=}\} \ \cup\ \{=^d, f_i^d \ \mid\ =^d, f_i^d \in \widehat{\Sigma}_{=} \text{ and } d \leq D\} \ \subset\ \widehat{\Sigma}_{=}.$$

Obviously, if Σ is finite, then for every $D \in \omega$ the signature $\widehat{\Sigma}_{=}^{D}$ is *finite.* Every formula of $\widehat{\Sigma}_{=}$ is, of course, a formula of signature $\widehat{\Sigma}_{=}^{D}$ for some $D \in \omega$.

For the purposes of decidability and complexity analysis, we need to associate norms to terms. A *norm* of a variable-free term is its height, defined as usual. For such a term a we write $|a| \leq m$ or simply $a \leq m$ to mean that the norm of a does not exceed m. By writing $\overline{a}_k \leq m$ we mean that for every term a_i of the k-tuple $\overline{a}_k$ one has $a_i \leq m$.

Definition 14 (Local Boundedness). Let $\mathcal{M} = \langle T(\Sigma);\ \widehat{\Sigma}_= \rangle$ be the canonical relational model of the bounded theory of trees (see Definition 9) and $H : \omega^4 \to \omega$ be a function. We say that $\mathcal{M}$ is *H-locally bounded* iff for every $n, k, m, D \in \omega$, every $\overline{a}_k \in T(\Sigma)^k$ with $\overline{a}_k \leq m$, and every formula of quantifier depth $\leq n$ with k free variables of signature $\widehat{\Sigma}_=^D$ the following is true:

$$\mathcal{M} \models \exists x_{k+1} \Phi(\overline{a}_k, x_{k+1}) \Rightarrow \mathcal{M} \models \Phi(\overline{a}_k, a_{k+1}) \text{ for some } a_{k+1} \leq H(n, k, m, D).$$

Remark. Notice that the upper bound on the size of a witness a_{k+1} in the above definition may depend on the maximal rank D of a predicate in a formula. This is not taken into account in the original Ferrante-Rackoff games, which apply only to finite signatures; recall that $\widehat{\Sigma}_=$ is *always infinite.*

Notation. For $Q \in \{\exists, \forall\}$ we write $\mathcal{M} \models (Qx_{k+1} \leq H(n, k, m, D))\Phi(\overline{a}_k, x_{k+1})$ to mean that $\mathcal{M} \models \Phi(\overline{a}_k, a_{k+1})$ *for some* (resp. *for all*) $a_{k+1} \leq H(n, k, m, D)$.

Local boundedness yields decidability and provides means to settle upper complexity bounds, quite similar to Theorem 1 from [3] p. 30.

Theorem 15. *Suppose that $\mathcal{M}$ is H-locally bounded and $Q_1x_1Q_2x_2 \dots Q_kx_k\, \Phi(\overline{x}_k)$ is a sentence with $Q_i \in \{\forall, \exists\}$ and a quantifier-free matrix $\Phi(\overline{x}_k)$ of signature $\widehat{\Sigma}_=^D$ for some $D \in \omega$. Suppose $m_0 \leq m_1 \leq m_2 \leq \dots \leq m_k$ is a sequence of natural numbers such that $H(k-i, i-1, m_{i-1}, D) \leq m_i$ for $1 \leq i \leq k$. Then $\mathcal{M} \models\ Q_1x_1Q_2x_2 \dots Q_kx_k\, \Phi(\overline{x}_k) \Leftrightarrow \mathcal{M} \models\ (Q_1x_1 \leq m_1) \dots (Q_kx_k \leq m_k)\, \Phi(\overline{x}_k)$.*

Thus local boundedness reduces the validity of a quantified formula to the validity of a *boundedly quantified* formula. Since Σ is a finite functional signature, the number of terms of bounded height is finite. Therefore, the validity check for the last formula amounts to verification of its matrix over finite number of tuples of terms. Consequently, we have the following simple way to settle the upper complexity bound for the theory. Suppose, an arbitrary element $x_i \leq m_i$ can be written in space at most $S(m_i)$. Then to test the validity of the last formula, it suffices to generate all k-tuples of elements $x_1 \leq m_1, \dots, x_k \leq m_k$ and to check the validity of its quantifier-free matrix $\Phi(\overline{x}_k)$ for each such k-tuple. The latter test does not usually use much additional space. Thus, the space $\sum_{i=1}^{k}\ S(m_i)$ to is sufficient to decide. We return to these calculations in Sections 5.3–5.5.

To prove local boundedness, necessary to apply Theorem 15, we need an auxiliary notion of *indistinguishability* of tuples by formulas of bounded quantifier depth and bounded rank of predicate symbols.

Definition 16 ($\equiv_{n,k}^D$ Relations). For $n, k, D \in \omega$ define the binary relation $\equiv_{n,k}^D$ on the set of k-tuples of constant terms of signature Σ as follows:

$\overline{a}_k \equiv_{n,k}^D \overline{b}_k$ iff $(\mathcal{M}, \overline{a}_k)$ and $(\mathcal{M}, \overline{b}_k)$ satisfy the same formulas of signature $\widehat{\Sigma}_=^D$ with k free variables of quantifier depth at most n. (1)

The following theorem, extending Ferrante-Rackoff's Theorem 3 [3] pp. 34–35 for infinite signatures, simplifies the proof of local boundedness, by reducing it to the proof of two conditions (2) and (3), familiar as the *back-and-forth* conditions in Ehrenfeucht-Fraïssé games [2, 4], but with additional boundedness constraints.

Theorem 17. *Let $\mathcal{M}$ be the canonical relational model of the bounded theory of trees. Suppose $H : \omega^4 \to \omega$ is a function and there exist binary relations $E^D_{n,k}$ satisfying properties (2), (3) for all $n, k, m, D \in \omega$, and $\overline{a}_k, \overline{b}_k \in T(\Sigma)^k$:*

- $\overline{a}_k \; E^D_{0,k} \; \overline{b}_k \;\Rightarrow\; \overline{a}_k \; \equiv^D_{0,k} \; \overline{b}_k.$ (2)
- *If $\overline{a}_k \; E^D_{n+1,k} \; \overline{b}_k$ and $\overline{b}_k \leq m$, then for every $a_{k+1} \in T(\Sigma)$ there exists $b_{k+1} \in T(\Sigma)$ such that $b_{k+1} \leq H(n,k,m,D)$ and $\overline{a}_{k+1} \; E^D_{n,k+1} \; \overline{b}_{k+1}$.* (3)

THEN:
- $\overline{a}_k \; E^D_{n,k} \; \overline{b}_k \Rightarrow \overline{a}_k \; \equiv^D_{n,k} \; \overline{b}_k$ *for all $n, k, D \in \omega$.* (4)
- *The model $\mathcal{M}$ is H-locally bounded.* (5)

5 Upper Bounds for the Bounded Theories of Trees

5.1 $E^D_{n,k}$ Relations

Now we apply Theorem 17 to prove the local boundedness of the bounded theory of finite trees, and then use Theorem 15 to conclude its decidability and to settle the upper complexity bounds. The crucial point in application of Theorem 17 is the invention of appropriate refinement relations $E^D_{n,k}$. We first need a simple auxiliary definition.

Definition 18 (Truncation). Let t be a ground term of signature Σ and $h \in \omega$. The *h-truncation* of t results from t by replacing all the subterms of t at depth $h+1$ with an *arbitrary but fixed* constant symbol from Σ. Define the h-truncation of a k-tuple of ground terms *componentwise.* □

Proposition 19. *Let for some $D \in \omega$ the D-truncations of $\overline{a}_k$ and $\overline{b}_k$ coincide ($k \in \omega$). Then for any $d \in \{0, \ldots, D\}$ and any $i, j \in \{1, \ldots, k\}$ one has:*

1) $a_i =^d a_j \Leftrightarrow b_i =^d b_j$; 2) $f^d_p(a_i, a_j) \Leftrightarrow f^d_p(b_i, b_j)$; 3) $Is_c(a_i) \Leftrightarrow Is_c(b_i)$.

The proof is immediate from definitions. Here comes the principal

Definition 20 ($E^D_{n,k}$ Relations). For $D, n, k \in \omega$ define the binary relation $E^D_{n,k}$ on the set of k-tuples of constant terms of signature Σ as follows:

$\overline{a}_k \; E^D_{n,k} \; \overline{b}_k$ *if and only if the $2^n + D$-truncations of $\overline{a}_k$ and $\overline{b}_k$ coincide.* (6)

We now prove that $E^D_{n,k}$ satisfy conditions (2), (3) of Theorem 17.

5.2 Basis: Condition (2) of Theorem 17

We must prove $\overline{a}_k\ E^D_{0,k}\ \overline{b}_k \ \Rightarrow\ \overline{a}_k\ \equiv^D_{0,k}\ \overline{b}_k$. (7)

By (6), $\overline{a}_k\ E^D_{0,k}\ \overline{b}_k$ means that $1+D$-truncations of $\overline{a}_k$ and $\overline{b}_k$ coincide.

By (1), $\overline{a}_k\ \equiv^D_{0,k}\ \overline{b}_k$ means that $(\mathcal{M}, \overline{a}_k)$ and $(\mathcal{M}, \overline{b}_k)$ satisfy the same *atomic* formulas of signature $\widehat{\Sigma}^D_=$. Such an atomic formula is either $x =^d y$ or $f^d_p(x,y)$, or $Is_c(x)$ for some $d \le D$, $f \in Fun(\Sigma)$, $p \in \{1,\ldots,ar(f)\}$, and $c \in Const(\Sigma)$.

Thus (7) is true by Proposition 19.

5.3 Inductive Step: Condition (3) of Theorem 17

Suppose $\overline{a}_k\ E^D_{n+1,k}\ \overline{b}_k$, $\overline{b}_k \le m$, and a_{k+1} is *an arbitrary ground term*. We must prove that for an *appropriate* bounding function H one can *always* choose $b_{k+1} \le H(n,k,m,D)$ in such a way that $\overline{a}_{k+1}\ E^D_{n,k+1}\ \overline{b}_{k+1}$ is satisfied. It suffices to select b_{k+1} to be equal the $2^n + D$-truncation of a_{k+1}. With this choice of b_{k+1} we obviously have $\overline{a}_{k+1}\ E^D_{n,k+1}\ \overline{b}_{k+1}$, because (cf., Definition 20):

- $\overline{a}_k\ E^D_{n+1,k}\ \overline{b}_k$ implies $\overline{a}_k\ E^D_{n,k}\ \overline{b}_k$,
- the $2^n + D$-truncation of a_{k+1} and b_{k+1} coincide.

It follows that the appropriate bounding function we need is

$$H(n,k,m,D) \;=\; 2^n + D, \tag{8}$$

because the $2^n + D$-truncation of a_{k+1} is of the norm $2^n + D$. Notice that the value of H *does not depend* neither on the number k of elements in a k-tuple, nor on their size m.

Therefore, the canonical model $\mathcal{M}$ of the bounded theory of trees is H-locally bounded for H defined by (8). This finishes the proof of the Theorem 17. □

5.4 Decidability

We now apply Theorem 15 to derive decidability of the bounded theory of finite trees from the H-local boundedness of its canonical model. We have to find a sequence of natural numbers $m_0 \le m_1 \le m_2 \le \ldots \le m_k$ such that $H(k-i, i-1, m_{i-1}) \le m_i$ for $1 \le i \le k$, where H is the bounding function defined by (8). As our function does not depend on its third argument, we simply let $m_0 = 0$, and $m_i = H(k-i, i-1, *) = 2^{k-i} + D$ for $i \in \{1,\ldots,k\}$. Therefore, to decide $Q_1 x_1 Q_2 x_2 \ldots Q_k x_k\, \Phi(\overline{x}_k)$ or, equivalently, $(Q_1 x_1 \le m_1) \ldots (Q_k x_k \le m_k)\, \Phi(\overline{x}_k)$ (by Theorem 15), we never need to consider trees higher than $2^k + D$. Since for a finite signature Σ the number of such trees is *finite* (finiteness of the signature is *crucial* here!), the bounded theory of finite trees over finite signature is decidable.

5.5 Complexity

We now turn to the upper complexity bound of the bounded theory of finite trees. It follows from Theorem 15 that the principal measure of complexity is the number of quantifiers in the prenex form of a formula. For an arbitrary formula ϕ of length l of signature $\widehat{\Sigma}_=$:

– the number of quantifiers k in ϕ is $O(l)$, and
– the maximal rank D of a predicate symbol in ϕ is $2^{O(l)}$, i.e., is *exponential* in its length; recall that we write the ranks of predicates $=^d$, f_p^d in *binary*.

Since the transformation of an arbitrary formula of the bounded theory of trees in the *functional signature* to an equivalent formula of the companion relational signature in prenex form results in a formula with $O(l)$ quantifiers (see Proposition 12) and of the *same rank*, to decide a formula of length l, we never need to consider trees higher than $2^{O(l)}$ (recall $2^k + D$).

An arbitrary tree of height $2^{O(l)}$ (we need to cycle through the k-tuples of such trees) may have up to $2^{2^{O(l)}}$ vertices and can be represented by an incidence matrix in space $2^{2^{O(l)}}$.

Therefore, an arbitrary formula of length l in the bounded theory of trees can be decided within space at most $2^{2^{O(l)}}$; hence, within deterministic time $2^{2^{2^{O(l)}}}$.

We thus established that the decision problem for the bounded theory of finite trees in a finite functional signature (or its relational companion) belongs to the complexity classes $SPACE(2^{2^{O(l)}}) \subseteq DTIME(2^{2^{2^{O(l)}}})$.

This estimate is true in general, when a signature Σ contains function symbols of arbitrary arities. In the particular case, when Σ has no function symbols of arity > 1, the above upper bound can be decreased. In fact, with monadic function symbols only, an arbitrary tree of height $2^{O(l)}$ may have only up to $2^{O(l)}$ vertices and can be represented in space $2^{O(l)}$. Thus the whole decision procedure runs within space $2^{O(l)}$ in this case.

Finally, consider a functional signature Σ containing ≥ 2 constant symbols only. In this case the bounded theory of finite trees is equivalent to the first-order theory of pure equality in a ≥ 2-element structure, known to be *PSPACE*-complete [12, 11].

6 Conclusion and Future Research

We introduced the bounded theory of finite trees and proved that it can be decided within elementary space (hence time), as contrasted to the usual theory of finite trees, which is of non-elementary decision complexity [14]. We thus demonstrated that the bounded theory of finite trees with its approximate equality may be used as a good practical substitute for the theory of finite trees.

In a subsequent publication we will demonstrate that the *lower bound* for the bounded theory of trees is as follows. For some constant $c > 0$ the theory *does not belong* to the complexity class $SPACE(2^{cn})$; consequently, requires nondeterministic exponential time to decide.

Venkataraman in [13] demonstrated that the first-order theory of finite trees with the *subtree predicate* is *undecidable*. By using the same machinery as we used in the paper it is possible to show that the bounded theory of trees with the "*to be a subtree at bounded depth*" predicate is decidable within elementary space and time. We will do it elsewhere.

As we see, the bounded theories may be useful when their unbounded counterparts are undecidable or intractable. It would be interesting to investigate practical applications of the bounded theories in, say, constraint logic programming schemes. This is, however, the topic of the future research.

References

1. H. Aït-Kaci, A. Podelski, and G. Smolka. A feature constraint system for logic programming with entailment. *Theor. Comput. Sci.*, 122:263–283, 1994. Preliminary version: 5th Intern. Conf. Fifth Generation Computer Systems, June 1992.
2. A. Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fund. Mathematicæ*, 49:129–141, 1961.
3. J. Ferrante and C. W. Rackoff. *The computational complexity of logical theories*, volume 718 of *Lect. Notes Math.* Springer-Verlag, 1979.
4. W. Hodges. *Model Theory*, volume 42 of *Encyclopedia of Mathematics and its Applications*. Cambridge Univ. Press, 1993.
5. J. Jaffar and M. J. Maher. Constraint logic programming: A survey. *J. Logic Programming*, 19 & 20:503–581, 1994.
6. K. Kunen. Answer sets and negation as failure. In *4th International Conference on Logic Programming*, volume 1, pages 219–228. MIT Press, 1987.
7. K. Kunen. Negation in logic programming. *J. Logic Programming*, 4:289–308, 1987.
8. M. J. Maher. Complete axiomatizations of the algebras of finite, rational, and infinite trees. In *3rd Annual IEEE Symp. on Logic in Computer Science LICS'88*, pages 348–357, 1988.
9. A. I. Mal'cev. Axiomatizable classes of locally free algebras. In B. F. Wells, editor, *The Metamathematics of Algebraic Systems (Collected Papers: 1936–1967)*, volume 66 of *Studies in Logic and the Foundations of Mathematics*, chapter 23, pages 262–281. North-Holland Pub. Co., 1971.
10. G. Smolka. Feature constraint logics for unification grammars. *J. Logic Programming*, 12:51–87, 1992.
11. L. J. Stockmeyer. The polynomial-time hierarchy. *Theor. Comput. Sci.*, 3:1–22, 1977.
12. L. J. Stockmeyer and A. R. Meyer. Word problems requiring exponential time: preliminary report. In *5th Symp. on Theory of Computing*, pages 1–9, 1973.
13. K. N. Venkataraman. Decidability of the purely existential fragment of the theory of term algebras. *J. ACM*, 34(2):492–510, 1987.
14. S. Vorobyov. An improved lower bound for the elementary theories of trees. In *Automated Deduction – CADE-13*, volume 1104 of *Lect. Notes Artificial Intelligence*, pages 275–287. Springer-Verlag, 1996.
15. S. Vorobyov. On the decision complexity of the bounded theories of trees. Technical Report MPI-96–2–008, Max-Planck Institut für Informatik, October 1996.

...