MAX-PLANCK-INSTITUT FÜR INFORMATIK

On the Decision Complexity of
the Bounded Theories of Trees

Sergei Vorobyov

MPI–I–96–2-008 November 1996

I N F O R M A T I K

Im Stadtwald
D 66123 Saarbrücken
Germany

Author's Address

Programming Logic Group AG2, Max-Planck Institut für Informatik,
Im Stadtwald, 66123 Saarbrücken, Germany,
e-mail: sv@mpi-sb.mpg.de

Publication Notes

As of November 28, 1996.

The short version of this report appeared in the Proceedings of the Asian'96 Computing Science Conference, Singapore, December 2–5, 1996, Joxan Jaffar, Editor, Springer-Verlag Lecture Notes in Computer Science [Vor96b].

Abstract

The *theory of finite trees* is the full first-order theory of equality in the Herbrand universe (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* [Vor96a].

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, 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 $c > 0$.

Keywords

Decision complexity of logical theories, elementary and non-elementary decision problems, lower and upper bounds for decision complexity, Ehrenfeucht-Fraïssé games

Contents

1 Introduction

Tree-like structures are fundamental for almost all domains of Computer Science, and are especially relevant to logic programming, symbolic computation, data types, constraint solving, automated theorem proving, data bases, knowledge representation, etc. 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 and functional programming.

The first-order theory of *finite trees*, also known as the theory of *term algebras*, or *Clark's equational theory*, although *decidable* [Mal71, Kun87b, Mah88, Hod93], turns out to be *non-elementary in the sense of Kalmar* [Vor96a]. Any nondeterministic decision procedure for the theory takes time exceeding infinitely often any fixed finite-story iterated exponential function $2^{2^{\cdot^{\cdot^{\cdot^{2^n}}}}}$, where n is the length of input. Even worse, every such decision procedure requires nondeterministic time (or space) $2^{\left.2^{\cdot^{\cdot^{\cdot^{2}}}}\right\}c\cdot n}$ (with the height of the tower growing linearly with n) to decide formulas of length n, for some constant $c > 0$ and infitely many $n \in \omega$.

Finite trees is one of the basic domains in the Constraint Logic Programming [JM94]. 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 and conjunctions (as is usually done) seems to be a serious restriction of expressiveness. In this respect the *bounded* theory of finite trees, considered in this paper, allowing for the full first-order quantification and being elementary, may be considered useful.

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 k-ary notation (with $k \geq 2$). Thus instead of stipulating 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.

Our basic decision and complexity analysis techniques are model-theoretic games. More specifically, we use Ferrante-Rackoff's complexity-tailored games [FR79], which refine Ehrenfeucht-Fraïssé-games [Ehr61, Hod93] 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.* By analyzing the size of these candidates for witnesses it is possible to obtain the upper space complexity bounds. This forms the basis of our decision and complexity analysis method. We carry over this machinery to the case of infinite signatures.

Although the analogy is not complete here, we would like to recall a rather similar situation with the full first-order theory of *binary concatenation*[1], which is *undecidable* [Qui46, Smu61], and the theory of *t-bounded concatenation*[2] [BM80, Ber80], which is decidable within elementary space and time if the function t is computable in elementary space.

Venkataraman in [Ven87] showed that the first-order theory of finite trees with the *subtree predicate* ($s \leq t$ meaning that s is a subtree of t) 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 ($s \leq_d t$) is decidable in elementary space and time.

A short version of this paper appeared in [Vor96b].

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

[1] i.e., the first-order theory of the structure $\langle \{0,1\}^*; conc(x,y,z)\rangle$, with the set of binary words as a carrier and the predicate $conc(x,y,z)$ interpreted as true iff the word z results from concatenation of words x, y.

[2] Let A be a finite alphabet. Consider the first-order language $L(A)$ with equality, containing a constant for every $a \in A$, and whose only atomic formulas are $bcat_n(x,y,z)$, where n is a unary numeral. Then for any function $t : \omega \to \omega$ define t-bounded concatenation theory as the set of of true sentences of $L(A)$ under the following interpretation: the underlying domain is A^*, the set of finite words over A; for $a,b,c \in A^*$ the formula $bcat_n(a,b,c)$ is true iff c is concatenation of a and b, and the length of c is at most $t(n)$.

and in Section 6 give the necessary generalization of these games for infinite signatures. Section 7 contains the application of games for the decidability proof and establishing the upper space complexity bounds for the bounded theory of trees.

2 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 universe 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.

By $(A, \overline{a}_k) \models F(\overline{x}_k)$ we mean that the formula $F(\overline{x}_k)$ is satisfied in A when the free occurrences of variables $\overline{x}_k$ in $F(\overline{x}_k)$ are replaced by the elements $\overline{a}_k$ of A. This is equivalent to $A \models F(\overline{a}_k)$.

3 Theory of Finite Trees

Global Proviso. Throughout the paper Σ denotes a *finite* functional signature containing *at least one* constant symbol. Hence $T(\Sigma)$, the Herbrand universe over Σ, is non-empty. □

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

The good well-known news, due to Mal'cev and Kunen, is that the theory is decidable.

Theorem 2 ([Mal71, Kun87b, Mah88, Hod93]) *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 [Mal71, Kun87b, Mah88, Hod93]. 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* $\exp_n(0)$, *i.e, a tower of 2's of height* n. *A decision problem is* elementary recursive 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 *not elementary recursive*. This disproves K. Kunen's claim [Kun87a] that the theory of finite trees is *PSPACE*-complete:

Theorem 4 ([Vor96a]) *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 $\exp_\infty(\lfloor cn \rfloor)$ *for some* $c > 0$ *and infinitely many* $n \in \omega$, *where* n *is the length of input.* □

The same applies to different variations of the theory, like the theories of *rational*, *feature*, and *rational feature* trees (for the definitions of these theories see, e.g., [Mah88, AKPS94, Smo92]).

4 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* $s_j =^d t_j$ *for* $j = 1, \ldots, m$. □

Thus, in contrast to the usual equality, comparing two random terms for the approximate $=^d$ equality takes time *at most exponential* in d.

Now we are ready to define the main subject of discourse in this paper. We give two definitionally equivalent [Hod93] formalizations of the bounded theory of finite trees: first, in a signature *with function symbols* and, second, *in a purely relational* signature.

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 universe* $T(\Sigma)$. □

We use the epithet "functional" and the superscripts ^{f} to stress the presence of function symbols and to distinguish the above formulation of the theory from the "relational" one we consider in the sequel.

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.* $\square$

A typical reduction step is $x =^{d+1} y \;\rightsquigarrow \bigvee_{f\in\Sigma} \exists x_1, y_1, \ldots, x_{ar(f)}, y_{ar(f)}$

$$\Big(x = f(x_1, \ldots, x_{ar(f)}) \wedge y = f(y_1, \ldots, y_{ar(f)}) \wedge \bigwedge_{i=1}^{ar(f)} x_i =^d y_i\Big).$$

By iteratively applying such reductions to all occurrences of the approximate equality predicates $=^d$ one can transform any $\mathcal{F}^f_{bnd}(\Sigma)$-sentence into an equivalent sentence of the usual theory of finite trees, and then use a decision procedure for that theory to settle the validity of the initial sentence.

We would like to stress, however, that this 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 operates 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, by replacing them with predicate symbols. We first define a relational signature corresponding to a functional one, then introduce a canonical model of this relational signature, and finally define the bounded theory of trees as the first-order theory of this model.

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 written in binary and called ranks. $\square$

The intended semantics of the relational language is captured by the following standard model.

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 universe* $T(\Sigma)$ *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* $\mathcal{M} \models f_p^d(s,t)$ *iff*
  $T(\Sigma) \models \exists x_1 \dots x_{p-1} x_{p+1} \dots x_{ar(f)} \big(s =^d f(x_1 \dots x_{p-1}, t, x_{p+1} \dots x_{ar(f)})\big)$.

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

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.* □

Remark 11 By definition, both $Th^f_{bnd}(T(\Sigma))$ and $Th^p_{bnd}(T(\Sigma))$ are *complete* theories, i.e., for every sentence ϕ either ϕ or or $\neg\phi$ belongs to a theory.

4.1 Relational vs. Functional Formalization

There is no essential difference between functional and relational theories.

Proposition 12 *The functional and the relational bounded theories of trees are* definitionally equivalent*, see [Hod93].* □

Informally this means that both theories may be interpreted in each other. It follows that the relational bounded theory of trees is also decidable. We describe in Section 7 a decision procedure for the *relational* bounded theory of trees. This does not lead to the loss of generality, since any formula in $\mathcal{F}^f_{bnd}(\Sigma)$ can be effectively transformed into an equivalent formula of $\mathcal{F}^p_{bnd}(\Sigma)$ (see Proposition 13). As is shown in Section 7 the main parameter influencing the decision complexity of the relational bounded theory of trees $Th^p_{bnd}(T(\Sigma))$

is the number of quantifiers in the prenex form of a formula. The following proposition shows that the transformation of an arbitrary formula of $\mathcal{F}^f_{bnd}(\Sigma)$ into a prenex formula of $\mathcal{F}^p_{bnd}(\Sigma)$ gives *at most a linear increase* of the number of quantifiers.

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 7). It differs only by a constant factor for both theories:

Proposition 13 *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.*

Proof. First convert a formula into a *flat form* containing at most one function symbol per atom. A typical conversion is $x =^{d+1} f(\ldots, g(\ldots), \ldots) \rightsquigarrow \forall u\Big(u =^d g(\ldots) \Rightarrow x =^{d+1} f(\ldots, u, \ldots)\Big)$. This gives at most linear increase of the number of quantifiers. Second, transform the resulting formula into prenex form by using a standard procedure. This does not increase the number of quantifiers. Finally replace atoms $y =^d c$ and $y =^d f(x_1 \ldots x_k)$ with $Is_c(y)$ and $\wedge_{i=1}^k f_i^d(y, x_i)$ respectively. $\square$

Main Theorem.

1. *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.*
2. *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 $c > 0$.*
3. *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.* $\square$

By Propositions 12 and 13, it suffices to prove the claim for the bounded theory of trees in the companion relational signature $\widehat{\Sigma}_=$. We present the elementary decision procedure in Section 7.

5 Ferrante-Rackoff Games for Decidability

In this section we briefly survey a complexity-tailored refinement of the Ehrenfeucht-Fraïssé games due to Ferrante and Rackoff, following Chapter 2 of [FR79]. We discuss only a small fragment of their general techniques, which is needed to decide a theory of a *single* structure.

The classical Ehrenfeucht-Fraïssé method, see [Ehr61, Hod93], gives criteria, in terms of partial isomorphisms or back-and-forth games, of *indistinguishability* of two structures by first-order formulas. Consequently, if any couple of structures of a theory are indistinguishable, the theory is *complete*, and hence decidable. The *drawback* is that to prove decidability of a theory one has to have its *explicit axiomatization.* An explicit axiomatization may be problematic, as in the case of semantically defined theories, e.g., a first-order theory of a single structure. This is exactly the case we deal with.

The advantage of the Ferrante-Rackoff game techniques is that it works without explicit axiomatizations for the theories of classes of models, in particular, for a theory of a single structure. *The whole game decision method due to Ferrante-Rackoff consists in proving, by means of an Ehrenfeucht-Fraïssé-like game, that quantifiers can be replaced by bounded quantifiers,* running over **finite subsets** of a structure. Therefore, testing the validity of a quantified formula amounts to checking its matrix on a finite set of elements of the domain. Moreover, by analyzing the sizes of the finite sets in consideration, one gets upper complexity bounds on the decision problem.

Further we consider first-order languages *with relation symbols only.* The modification of the Ehrenfeucht-Fraïssé games for languages with unlimited use of function symbols is not difficult, but more elaborate, and could be found in [Hod93]. This was our aim in trading function symbols for predicates in Section 4.

5.1 Boundedness and Reduction to Bounded Quantification

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 settle whether $\exists x\, \Phi(x)$ is true or false, one just needs to check *finitely many small candidates for witnesses.* This forms the basis of the decision method. In this section we formally define boundedness and show how it leads to decidability. In Section 5.3 we explain how games are used to establish boundedness.

For decidability and complexity considerations, we associate with the elements of a structure a *norm*, i.e., a function from the domain of a structure to ω. This is particularly simple for structures built of syntactical material, like terms, trees, as in the case we deal with.

Definition 14 *Given a constant term t of a finite functional signature Σ with constants, the* norm *of t denoted by $|t|$ is defined as the maximal nesting of function symbols in t. In other words, $|c| = 0$ for $c \in Const(\Sigma)$, and $|f(t_1 \dots t_n)| = 1 + \max_{i=1}^{n}(|t_i|)$ for $f \in Fun(\Sigma)$.* □

Notation For an element a of a structure 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 element a_i of the k-tuple $\overline{a}_k$ one has $a_i \leq m$. □

Proviso. Throughout this section we suppose that all models we consider are the models of *finite purely relational signatures.* □

Remark 15 We discuss the needed modifications for *infinite* signatures in Section 7. This is necessary because even though Σ is finite, the companion relational signature $\widehat{\Sigma}_=$ is *always infinite.*

We are ready to introduce the main technical definition of boundedness underlying the decision method.

Definition 16 (Boundedness, [FR79]) *Suppose A is a model and $H : \omega^3 \to \omega$ is a function.*

Let for every $n, k, m \in \omega$, every $\overline{a}_k \in A^k$ such that $\overline{a}_k \leq m$, and every formula $\Phi(\overline{x}_{k+1})$ of quantifier depth $\leq n$ the following be true:

$$\begin{array}{ll} \textit{if} & A \models \exists x_{k+1}\, \Phi(\overline{a}_k, x_{k+1}), \\ \textit{then} & A \models \Phi(\overline{a}_k, a_{k+1}) \textit{ for some } a_{k+1} \leq H(n,k,m). \end{array}$$

*In this case we write $A \models (\exists x_{k+1} \leq H(n,k,m))\, \Phi(\overline{a}_k, x_{k+1})$ and say that A is H-*bounded. □

If a model is known to be H-bounded, then the following simple theorem suggests a straightforward decision procedure for its full first-order theory.

Theorem 17 (p. 30 [FR79]) *Let a model A be H-bounded and $Q_1x_1Q_2x_2 \ldots Q_kx_k\, \Phi(\overline{x}_k)$ be a sentence with $Q_i \in \{\forall, \exists\}$ and a quantifier-free matrix $\Phi(\overline{x}_k)$. Suppose that $m_0 \leq m_1 \leq m_2 \leq \ldots \leq m_k$ is a sequence of natural numbers such that $H(k-i, i-1, m_{i-1}) \leq m_i$ for $1 \leq i \leq k$.*

$$\begin{array}{ll} \textit{Then} & A \models\ Q_1x_1Q_2x_2 \ldots Q_kx_k\, \Phi(\overline{x}_k) \\ \textit{if and only if} & A \models\ (Q_1x_1 \leq m_1) \ldots (Q_kx_k \leq m_k)\, \Phi(\overline{x}_k). \end{array}$$

Proof. By induction [FR79]. We generalize and prove it for infinite signatures in Section 6. □

We have the following simple

Corollary 18 *Suppose all the premises of Theorem 17 are true and for every $m \in \omega$ there exist **only finitely many** elements in A with norm $\leq m$ then $Th(A)$, the first-order theory of A, is decidable.* □

This is immediate, since deciding a theory reduces to the routine verification of unquantified formulas over finite domains.

Remark 19 (Very Important) This explains why the finiteness restriction on the functional signature Σ is *crucial.* For if Σ is infinite, for every

$m \in \omega$ there are *infinitely many* ground terms of signature Σ of norm $\leq m$. Thus, the replacement of unbounded quantification by the bounded one in Theorem 17 *does not yield decidability.* Notice that even though we assume the functional signature Σ finite, the companion relational signature $\widehat{\Sigma}_=$ introduced in Definition 8 is *always infinite.* So we must be extremely careful in applying Theorem 17 to the *relational* bounded theory of finite trees, which is formalized in the infinite signature $\widehat{\Sigma}$. In Section 7 we spend additional effort to reduce everything to the case of finite signatures. □

5.2 Calculating Upper Space Complexity Bounds

As a by-product, Theorem 17 gives us the following simple way to obtain upper complexity bounds. Suppose, in conditions of Theorem 17 one needs space at most $S(m_i)$ to write down a representation of an arbitrary element x with norm $|x| \leq m_i$ for $1 \leq i \leq k$. Then to check $A \models Q_1 x_1 Q_2 x_2 \ldots Q_k x_k \, \Phi(\overline{x}_k)$, it suffices, by Theorem 17, to generate all k-tuples of elements $\overline{x}_k$ such that $x_1 \leq m_1, \ldots, x_k \leq m_k$ and to check the validity of the quantifier-free matrix $\Phi(\overline{x}_k)$ for each such k-tuple. The latter check does not usually use much additional space. Thus, the space $\sum_{i=1}^{k} S(m_i)$ to cycle through the representations of all k-tuples of elements $\overline{x}_k$ satisfying $x_1 \leq m_1, \ldots, x_k \leq m_k$ is enough. This gives the upper space complexity bound for the decision complexity of a theory in question. We use these considerations in Section 7.5.

5.3 Proving Boundedness by Games

Thus the essence of the above method consists in demonstrating that a structure is H-bounded for an appropriate function H. In the rest of this section we present, following [FR79], Ch. 2, the Ehrenfeucht-Fraïssé game technique for proving H-boundedness.

First we need to define the indistinguishability equivalence relations $\equiv_{n,k}$

Definition 20 *Suppose A, B are two structures of the same signature.*

Let $n, k \in \omega$, and $\overline{a}_k \in A^k$, $\overline{b}_k \in B^k$. Write $(A, \overline{a}_k) \equiv_{n,k} (B, \overline{b}_k)$ if and only if for every formula $\Phi(\overline{x}_k)$ of quantifier depth $\leq n$ one has

$$A \models \Phi(\overline{a}_k) \text{ if and only if } B \models \Phi(\overline{b}_k).$$ □

So, $(A, \overline{a}_k) \equiv_{n,k} (B, \overline{b}_k)$ means that $(A, \overline{a}_k)$ and $(B, \overline{b}_k)$ are *indistinguishable* by formulas of quantifier depth $\leq n$. In particular, $(A, \overline{a}_k) \equiv_{0,k} (B, \overline{b}_k)$ means that $(A, \overline{a}_k)$ and $(B, \overline{b}_k)$ satisfy the same quantifier-free formulas or, equivalently, the same atomic formulas. The usefulness of the $\equiv_{n,k}$ relations is as follows. Let $(A, \overline{a}_k) \equiv_{n,k} (B, \overline{b}_k)$ and we have to verify the validity of formulas with quantifier depth at most n in $(A, \overline{a}_k)$. Suppose that for some reason we prefer $(B, \overline{b}_k)$, as smaller, more convenient, intuitive. In this case we may safely switch to $(B, \overline{b}_k)$ and use it in verification instead of $(A, \overline{a}_k)$.

Convention. When structures A and B are clear from the context, we will write $\overline{a}_k \equiv_{n,k} \overline{b}_k$ instead of $(A, \overline{a}_k) \equiv_{n,k} (B, \overline{b}_k)$. □

The following theorem allows us to establish the H-boundedness of a structure by means of a "back-and-forth" game. Ferrante and Rackoff [FR79] prove its variant for a class of models. For our purposes we need to formulate it for *just one* model.

For the ease of application, the theorem is stated for the relations $E_{n,k}$ refinining the $\equiv_{n,k}$ relations defined above.

Theorem 21 (pp. 34–35, [FR79]) *Let A be a structure, $H : \omega^3 \to \omega$ be a function, and binary relations $E_{n,k}$ satisfy the following properties for all $n, k, m \in \omega$, $\overline{a}_k, \overline{b}_k \in A^k$:*

- $\overline{a}_k \; E_{0,k} \; \overline{b}_k \; \Rightarrow \; \overline{a}_k \equiv_{0,k} \overline{b}_k$; (1)

- *if* $\overline{a}_k \; E_{n+1,k} \; \overline{b}_k$ *and* $\overline{b}_k \leq m$, *then for every* $a_{k+1} \in A$ *there exists* $b_{k+1} \in A$ *such that* $b_{k+1} \leq H(n, k, m)$ *and* $\overline{a}_{k+1} \; E_{n,k+1} \; \overline{b}_{k+1}$. (2)

THEN:

- $\overline{a}_k \; E_{n,k} \; \overline{b}_k \Rightarrow \overline{a}_k \equiv_{n,k} \overline{b}_k$ *for all* $n, k \in \omega$. (3)
- *The structure A is H-bounded.* (4)

Proof. See [FR79], pp. 35–36. We generalize and prove it for infinite signatures in Section 6. □

Condition (2) (without boundedness) of Theorem 21 is well known in model theory as the *"back-and-forth condition"* [CK73, Hod93] (by symmetry we need only the "forth" part), and has a natural interpretation in terms of games and winning strategies [Ehr61, Hod93]: whenever $\overline{a}_k$ and $\overline{b}_k$ are in a "good" relation, whichever element a_{k+1} the "spoiler" (usually $\forall$belard) chooses, the "duplicator" (usually ∃loise) can always respond by choosing a correct b_{k+1} to maintain the resulting $\overline{a}_{k+1}$ and $\overline{b}_{k+1}$ in a "good" relation.

We say that the element b_{k+1} in (2) is **small**, because its size depends only on n, k and the maximum of sizes of $\overline{b}_k$, and **does not depend on the sizes of** $\overline{a}_k$ and a_{k+1}. Therefore, the condition (2) stipulates that whenever $\overline{a}_k\ E_{n+1,k}\ \overline{b}_k$, then for any a_{k+1} there exists a **small** b_{k+1} satisfying $\overline{a}_{k+1}\ E_{n,k+1}\ \overline{b}_{k+1}$. This boundedness condition is absent from the classical formulation of games [Ehr61, Hod93], but is very useful to establish decidability, guaranteed by the condition (4) and Theorem 17. The condition (3) guarantees that the relations $E_{n,k}$ *refine* the relations $\equiv_{n,k}$. Thus the k-tuples equivalent modulo $E_{n,k}$ are $\equiv_{n,k}$ -equivalent.

Remark 22 . Theorem 21 allows, in particular, to use directly the relations $\equiv_{n,k}$ instead of $E_{n,k}$ for establishing H-boundedness of structures (in which case there is no need to stipulate (1) and demonstrate (3)). However, working with refinements $E_{n,k}$ of $\equiv_{n,k}$ is easier in practice, because $\equiv_{n,k}$ are formulated in terms of quantified formulas and are usually difficult to deal with, whereas $E_{n,k}$ may be formulated in a different language, tailored to the problem domain, and even in a richer and more expressive language. We will see an example in Section 7. □

6 Ferrante-Rackoff's Games in Infinite Signatures

In the next section we prove our Main Theorem by applying Ferrante-Rackoff's game techniques described in Section 5 by Theorems 17 and 21. 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, see Remark 19. 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.

6.1 Local Boundedness

Definition 23 *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 $D \in \omega$.

We need to modify correspondingly the notion of boundedness, cf., Definition 16. Recall that $\mathcal{M}$ is the canonical relational model of the bounded theory of trees (see Definition 9).

Definition 24 (Local Boundedness) *Let $H : \omega^4 \to \omega$ be a function.*

*We say that $\mathcal{M}$, the canonical relational model of the bounded theory of trees, is H-*locally bounded *if and only if 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 $\Phi(\overline{x}_k, x_{k+1})$ of quantifier depth $\leq n$ with $k+1$ free variables of signature $\widehat{\Sigma}^D_=$ the following is true:*

$$\begin{array}{ll} \textit{if} & \mathcal{M} \models \exists x_{k+1} \Phi(\overline{a}_k, x_{k+1}) \\ \textit{then} & \mathcal{M} \models \Phi(\overline{a}_k, a_{k+1}) \textit{ for some } a_{k+1} \leq H(n, k, m, D). \end{array}$$

Remark 25 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.* $\square$

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

6.2 Local Boundedness Implies Decidability

Local boundedness yields decidability and provides means to settle upper complexity bounds, quite similar to Theorem 17.

Theorem 26 *Suppose that $\mathcal{M}$ is H-locally bounded.*

Let $Q_1x_1Q_2x_2\ldots Q_kx_k\,\Phi(\overline{x}_k)$ be 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 that

$$m_0 \leq m_1 \leq m_2 \leq \ldots \leq m_k$$

is a sequence of natural numbers such that

$$H(k-i, i-1, m_{i-1}, D) \leq m_i \text{ for } 1 \leq i \leq k.$$

Then

$$\begin{aligned} \mathcal{M} \models\ & Q_1x_1Q_2x_2\ldots Q_kx_k\,\Phi(\overline{x}_k) \text{ if and only if} \\ \mathcal{M} \models\ & (Q_1x_1 \leq m_1)\ldots(Q_kx_k \leq m_k)\,\Phi(\overline{x}_k). \end{aligned}$$

Proof. We prove by induction in $i \in \{1, \ldots, k+1\}$ that

$$\begin{aligned} \mathcal{M} \models\ & Q_1x_1\ldots Q_kx_k\ \Phi(\overline{x}_k) \text{ iff} \\ \mathcal{M} \models\ & (Q_1x_1 \leq m_1)\ldots(Q_{i-1}x_{i-1} \leq m_{i-1})Q_ix_i\ldots Q_kx_k\ \Phi(\overline{x}_k). \end{aligned}$$

The base case, $i = 1$, is obviously true, since it is just

$$\mathcal{M} \models Q_1x_1\ldots Q_kx_k\ \Phi(\overline{x}_k) \ \Leftrightarrow\ \mathcal{M} \models Q_1x_1\ldots Q_kx_k\ \Phi(\overline{x}_k).$$

Assume the induction hypothesis is true for some $i \in \{1, \dots, k\}$; we must prove it for $i + 1$. Consider any $\overline{a}_{i-1} \in T(\Sigma)^{i-1}$ such that $a_j \leq m_j$ for $1 \leq j \leq i - 1$. Then, since $\mathcal{M}$ is H-locally bounded,

$$\begin{aligned}\mathcal{M} \models\ & Q_i x_i [Q_{i+1} x_{i+1} \dots Q_k x_k \Phi(\overline{a}_{i-1}, x_i, \dots, x_k)] \ \Leftrightarrow \\ \mathcal{M} \models\ & (Q_i x_i \leq H(k-i, i-1, m_{i-1}, D)[Q_{i+1} x_{i+1} \dots Q_k x_k \Phi(\overline{a}_{i-1}, x_i, \dots, x_k)].\end{aligned}$$

(Note: This follows directly from the definition of local boundedness only if Q_i is $\exists$, but it is easy to see (since $\forall = \neg\exists\neg$) that it should also hold for Q_i equal $\forall$).

Since $H(k-i, i-1, m_{i-1}, D) \leq m_i$ by assumption, for all $\overline{a}_{i-1} \in T(\Sigma)^{i-1}$ such that $a_j \leq m_j$ for $1 \leq j \leq i - 1$, we have:

$$\begin{aligned}\mathcal{M} \models\ & Q_i x_i [Q_{i+1} x_{i+1} \dots Q_k x_k \Phi(\overline{a}_{i-1}, x_i, x_{i+1}, \dots, x_k)] \ \Leftrightarrow \\ \mathcal{M} \models\ & (Q_i x_i \leq m_i)[Q_{i+1} x_{i+1} \dots Q_k x_k \Phi(\overline{a}_{i-1}, x_i, x_{i+1}, \dots, x_k)]. \qquad (5)\end{aligned}$$

Therefore, by inductive hypothesis and (5) respectively,

$$\begin{aligned}\mathcal{M} \models\ & Q_1 x_1 \dots Q_k x_k \ \Phi(\overline{x}_k) \ \Leftrightarrow \\ \mathcal{M} \models\ & (Q_1 x_1 \leq m_1) \dots (Q_{i-1} x_{i-1} \leq m_{i-1}) \\ & \qquad \mathbf{Q}_i x_i Q_{i+1} x_{i+1} \dots Q_k x_k \Phi(\overline{a}_{i-1}, x_i, x_{i+1}, \dots, x_k) \ \Leftrightarrow \\ \mathcal{M} \models\ & (Q_1 x_1 \leq m_1) \dots (Q_{i-1} x_{i-1} \leq m_{i-1}) \\ & \qquad (Q_i x_i \leq m_i) Q_{i+1} x_{i+1} \dots Q_k x_k \Phi(\overline{a}_{i-1}, x_i, x_{i+1}, \dots, x_k).\end{aligned}$$

Thus, the inductive hypothesis is true for $i + 1$ and we are done. □

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, as described in Section 5.2. We return back to these calculations in Sections 7.3–7.5.

6.3 Proving Local Boundedness by Games

To prove local boundedness, necessary to apply Theorem 26, we need an auxiliary notion of *indistinguishability* of tuples by formulas of bounded quantifier depth and bounded rank of predicate symbols. We first define the local analog of the $\equiv_{n,k}$ relations, cf., Definition 20.

Definition 27 ($\equiv_{n,k}^D$ 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. $\square$

Now we state and prove the following *local variant* of Ferrante-Rackoff's Theorem 21 for the game-style proof of local boundedness by means of refinements $E_{n,k}^D$ of the $\equiv_{n,k}^D$ relations.

The theorem simplifies the proof of local boundedness, by reducing it to the proof of two conditions (6) and (7), familiar as the *back-and-forth* conditions in Ehrenfeucht-Fraïssé games [Ehr61, Hod93], but with additional boundedness constraints. It also takes into consideration the modified notion of boundedness, adapted for infinite signatures.

Theorem 28 *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 $E_{n,k}^D$ satisfying properties (6), (7) for all $n, k, m, D \in \omega$, and $\overline{a}_k, \overline{b}_k \in T(\Sigma)^k$:*

- $\overline{a}_k \; E_{0,k}^D \; \overline{b}_k \;\Rightarrow\; \overline{a}_k \equiv_{0,k}^D \overline{b}_k.$ (6)
- *If $\overline{a}_k \; E_{n+1,k}^D \; \overline{b}_k$ 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 $\overline{a}_{k+1} \; E_{n,k+1}^D \; \overline{b}_{k+1}$.* (7)

THEN:

- $\overline{a}_k \; E_{n,k}^D \; \overline{b}_k \Rightarrow \overline{a}_k \equiv_{n,k}^D \overline{b}_k$ *for all $n, k, D \in \omega$.* (8)
- *The model $\mathcal{M}$ is H-locally bounded.* (9)

Proof. We first prove (8) by induction on n. It follows from (6) that (8) is true for $n = 0$ and all $k, D \in \omega$. So the base case is true.

Assume that (8) is true for some n and all $k, D \in \omega$. We must prove it for $n+1$. Suppose that $(\mathcal{M}, \overline{a}_k) \; E_{n+1,k}^D \; (\mathcal{M}, \overline{b}_k)$ and $\Phi(\overline{x}_k)$ is a formula of quantifier depth $n+1$ of signature $\widehat{\Sigma}_{=}^D$. We must demonstrate that $\mathcal{M} \models \Phi(\overline{a}_k) \Leftrightarrow$

$\mathcal{M} \models \Phi(\overline{b}_k)$. Since every formula of a given quantifier depth is equivalent to a boolean combination of formulas beginning with existential quantifiers of the same signature and the same (or less) quantifier depth, it suffices to demonstrate that $\mathcal{M} \models \exists x_{k+1}\Psi(\overline{a}_k, x_{k+1}) \Leftrightarrow \mathcal{M} \models \exists x_{k+1}\Psi(\overline{b}_k, x_{k+1})$, where Ψ is of quantifier depth n. By symmetry it suffices to show only $\Rightarrow$.

Assume that $\mathcal{M} \models \exists x_{k+1}\Psi(\overline{a}_k, x_{k+1})$ and let $a_{k+1} \in T(\Sigma)$ be such that $\mathcal{M} \models \Psi(\overline{a}_k, a_{k+1})$. Since $(\mathcal{M}, \overline{a}_k)\ E^D_{n+1,k}\ (\mathcal{M}, \overline{b}_k)$, we have, by (7), that for some $b_{k+1} \in T(\Sigma)$, $(\mathcal{M}, \overline{a}_{k+1})\ E^D_{n,k+1}\ (\mathcal{M}, \overline{b}_{k+1})$. Hence, by induction hypothesis, $(\mathcal{M}, \overline{a}_{k+1}) \equiv^D_{n,k+1} (\mathcal{M}, \overline{b}_{k+1})$, and, since $\mathcal{M} \models \Psi(\overline{a}_k, a_{k+1})$, we also have $\mathcal{M} \models \Psi(\overline{b}_k, b_{k+1})$. Thus, $\mathcal{M} \models \exists x_{k+1}\Psi(\overline{b}_k, x_{k+1})$, and (8) is proved.

We now prove (9). Let $n, k, D \in \omega$ be arbitrary, and $\Phi(\overline{x}_k)$ be of quantifier depth $\leq n$ and of signature $\widehat{\Sigma}^D_{=}$. Suppose that $\mathcal{M} \models \exists x_{k+1}\Phi(\overline{a}_k, x_{k+1})$. Then for some $a_{k+1} \in T(\Sigma)$ one has $\mathcal{M} \models \Phi(\overline{a}_k, a_{k+1})$.

Since, obviously, $(\mathcal{M}, \overline{a}_k)\ E^D_{n+1,k}\ (\mathcal{M}, \overline{a}_k)$, for some $a'_{k+1} \leq H(n, k, m, D)$ one has $(\mathcal{M}, \overline{a}_k, a_{k+1})\ E^D_{n,k+1}\ (\mathcal{M}, \overline{a}_k, a'_{k+1})$, by assumption (7). By the already proved property (8), $(\mathcal{M}, \overline{a}_k, a_{k+1}) \equiv^D_{n,k+1} (\mathcal{M}, \overline{a}_k, a'_{k+1})$. Since $\mathcal{M} \models \Phi(\overline{a}_k, a_{k+1})$, it should also be $\mathcal{M} \models \Phi(\overline{a}_k, a'_{k+1})$. Thus $\mathcal{M}$ is H-locally bounded. $\square$

7 Upper Bounds for the Bounded Theories of Trees

In this section we apply Theorem 28 to prove the local boundedness of the bounded theory of finite trees, and then use Theorem 26 to conclude its decidability and to settle the upper complexity bounds.

7.1 $E_{n,k}^D$ Relations

The crucial point in application of Theorem 28 is the invention of appropriate refinement relations $E_{n,k}^D$. We first need a simple auxiliary definition.

Definition 29 (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. □

We have the following simple

Proposition 30 *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. $f_p^d(a_i, a_j) \Leftrightarrow f_p^d(b_i, b_j)$;
3. $Is_c(a_i) \Leftrightarrow Is_c(b_i)$. □

The proof is immediate from definitions. Here comes the principal

Definition 31 ($E_{n,k}^D$ Relations) *For $D, n, k \in \omega$ define the binary relation $E_{n,k}^D$ on the set of k-tuples of constant terms of signature Σ as follows:*

$$\overline{a}_k \ E_{n,k}^D \ \overline{b}_k \text{ iff the } 2^n + D\text{-truncations of } \overline{a}_k \text{ and } \overline{b}_k \text{ coincide.} \quad \square$$

We now prove that $E_{n,k}^D$ satisfy conditions (6) and (7) of Theorem 28.

7.2 Basis: Condition (6) of Theorem 28

We must prove

$$\overline{a}_k\ E^D_{0,k}\ \overline{b}_k\ \Rightarrow\ \overline{a}_k\ \equiv^D_{0,k}\ \overline{b}_k. \tag{10}$$

By Definition 31, $\overline{a}_k\ E^D_{0,k}\ \overline{b}_k$ means that $1+D$-truncations of $\overline{a}_k$ and $\overline{b}_k$ coincide. By Definition 27, $\overline{a}_k\ \equiv^D_{0,k}\ \overline{b}_k$ 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 $d \leq D$, $f \in Fun(\Sigma)$, $p \in \{1, \ldots, ar(f)\}$, and $c \in Const(\Sigma)$. Therefore, (10) follows by Proposition 30.

7.3 Inductive Step: Condition (7) of Theorem 28

Suppose $\overline{a}_k\ E^D_{n+1,k}\ \overline{b}_k$, $\overline{b}_k \leq m$, and a_{k+1} is *an arbitrary ground term* of signature Σ. We must prove that for an *appropriate* bounding function H one can *always* choose $b_{k+1} \leq H(n,k,m,D)$ 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}. Indeed, with this choice of b_{k+1} we obviously have $\overline{a}_{k+1}\ E^D_{n,k+1}\ \overline{b}_{k+1}$, because:

- $\overline{a}_k\ E^D_{n+1,k}\ \overline{b}_k$ implies $\overline{a}_k\ E^D_{n,k}\ \overline{b}_k$ (cf., Definition 31),
- the $2^n + D$-truncation of a_{k+1} and b_{k+1} coincide (by choice of b_{k+1}).

It follows that the appropriate bounding function we need is

$$H(n,k,m,D)\ =\ 2^n + D, \tag{11}$$

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 (11). This finishes the proof of the Theorem 28. □

7.4 Decidability

We now apply Theorem 26 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 $H(k-i, i-1, m_{i-1}, D) \le m_i$ for $1 \le i \le k$, where H is the bounding function defined by (11). As our function does not depend on its third argument, we simply let $m_0 = 0$, and $m_i = H(k-i, i-1, *, D) = 2^{k-i} + D$ for $i \in \{1, \ldots, k\}$. Therefore, to decide $Q_1x_1Q_2x_2 \ldots Q_kx_k\, \Phi(\overline{x}_k)$ (with Φ quantifier-free) or, equivalently, $(Q_1x_1 \le m_1) \ldots (Q_kx_k \le m_k)\, \Phi(\overline{x}_k)$ (by Theorem 26), 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.

7.5 Complexity

We now turn to the upper complexity bound of the bounded theory of finite trees. It follows from Theorem 26 that the principal measure of complexity is the number of quantifiers in the prenex form of a formula. From our considerations in Sections 7.3 and 7.4 it follows that this complexity also depends on the maximal rank D of predicate symbols occurring in 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 predicate symbols $=^d$, 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 13) 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 greater than 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, and no function symbols at all. 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 [SM73, Sto77].

8 Conclusions and Further 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 [Vor96a]. 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 [Ven87] 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

[AKPS94] 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.

[Ber80] L. Berman. The complexity of logical theories. *Theor. Comput. Sci.*, 11:71–77, 1980.

[BM80] A. R. Bruss and A. R. Meyer. On time-space classes and their relation to the theory of real addition. *Theor. Comput. Sci.*, 11:59–69, 1980.

[CK73] C. C. Chang and H.J. Keisler. *Model theory*, volume 73 of *Studies in Logic and the Foundations of Mathematics*. Elsevier, 1973. (3rd edition, 1990).

[Ehr61] A. Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fund. Mathematicæ*, 49:129–141, 1961.

[FR79] J. Ferrante and C. W. Rackoff. *The computational complexity of logical theories*, volume 718 of *Lect. Notes Math.* Springer-Verlag, 1979.

[Hod93] W. Hodges. *Model Theory*, volume 42 of *Encyclopedia of Mathematics and its Applications*. Cambridge Univ. Press, 1993.

[JM94] J. Jaffar and M. J. Maher. Constraint logic programming: A survey. *J. Logic Programming*, 19 & 20:503–581, 1994.

[Kun87a] K. Kunen. Answer sets and negation as failure. In *4th International Conference on Logic Programming*, volume 1, pages 219–228. MIT Press, 1987.

[Kun87b] K. Kunen. Negation in logic programming. *J. Logic Programming*, 4:289–308, 1987.

[Mah88] 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.

[Mal71] 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.

[Qui46] W. V. Quine. Concatenation as a basis for arithmetic. *J. Symb. Logic*, 11(4):105–114, 1946.

[SM73] 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.

[Smo92] G. Smolka. Feature constraint logics for unification grammars. *J. Logic Programming*, 12:51–87, 1992.

[Smu61] R. M. Smullyan. *Theory of Formal Systems*. Princeton University Press, revised edition, 1961.

[Sto77] L. J. Stockmeyer. The polynomial-time hierarchy. *Theor. Comput. Sci.*, 3:1–22, 1977.

[Ven87] K. N. Venkataraman. Decidability of the purely existential fragment of the theory of term algebras. *J. ACM*, 34(2):492–510, 1987.

[Vor96a] 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.

[Vor96b] S. Vorobyov. On the bounded theories of finite trees. In *Asian'96 Computing Science Conference*, volume (to appear) of *Lect. Notes Comput. Sci.* Springer-Verlag, December 1996.

20pt
20pt

I N F O R M A T I K

Below you find a list of the most recent technical reports of the research group *Logic of Programming* at the Max-Planck-Institut für Informatik. They are available by anonymous ftp from our ftp server `ftp.mpi-sb.mpg.de` under the directory `pub/papers/reports`. Most of the reports are also accessible via WWW using the URL `http://www.mpi-sb.mpg.de`. If you have any questions concerning ftp or WWW access, please contact `reports@mpi-sb.mpg.de`. Paper copies (which are not necessarily free of charge) can be ordered either by regular mail or by e-mail at the address below.

Max-Planck-Institut für Informatik
Library
attn. Regina Kraemer
Im Stadtwald
D-66123 Saarbrücken
GERMANY
e-mail: `kraemer@mpi-sb.mpg.de`

MPI-I-96-2-009	D. Basin, N. Klarlund	Beyond the Finite in Automatic Hardware Verification
MPI-I-96-2-007	A. Herzig	SCAN and Systems of Conditional Logic
MPI-I-96-2-006	D. Basin, S. Matthews, L. Vigano	Natural Deduction for Non–Classical Logics
MPI-I-96-2-005	A. Nonnengart	Auxiliary Modal Operators and the Characterization of Modal Frames
MPI-I-96-2-004	G. Struth	Non-Symmetric Rewriting
MPI-I-96-2-003	H. Baumeister	Using Algebraic Specification Languages for Model-Oriented Specifications
MPI-I-96-2-002	D. Basin, S. Matthews, L. Vigano	Labelled Propositional Modal Logics: Theory and Practice
MPI-I-96-2-001	H. Ganzinger, U. Waldmann	Theorem Proving in Cancellative Abelian Monoids
MPI-I-95-011	P. Mutzel	Automatisiertes Zeichnen von Diagrammen
MPI-I-95-010	C. Rüb	On the Average Running Time of Odd-Even Merge Sort
MPI-I-95-009	J. Könemann, C. Schmitz, C. Schwarz	Radix heaps, an efficient implementation for priority queues
MPI-I-95-008	H. J. Ohlbach, R. A.Schmidt, U. Hustadt	Translating Graded Modalities into Predicate Logic
MPI-I-95-007	J. Radhakrishnan, S. Saluja	Lecture Notes Interactive Proof Systems
MPI-I-95-006	95, P. G. Bradford, R. Fleischer, M. Smid	95 A Polylog-Time and $O(n\sqrt{\lg n})$-Work Parallel Algorithm for finding the Row Minima in Totally Monotone Matrices
MPI-I-95-005	J.-H. Hoepmann, M. Papatriantafilou, P. Tsigas	Towards Self-Stabilizing Wait-Free Shared Memory Objects
MPI-I-95-003	P. G. Bradford, R. Fleischer	Matching nuts and bolts faster
MPI-I-95-002	S. Näher, C. Uhrig	LEDA user manual (Version R 3.2)
MPI-I-95-001	M. Smid, C. Thiel, F. Follert, E. Schömer, J. Sellen	Computing a largest empty anchored cylinder, and related problems