



HAL
open science

Extension into trees of first order theories

Khalil Djelloul, Thi-Bich-Hanh Dao

► **To cite this version:**

Khalil Djelloul, Thi-Bich-Hanh Dao. Extension into trees of first order theories. The 8th International conference on artificial intelligence and symbolic computation, 2006, China. pp. LNAI, Vol 4120. P 53-67. hal-00202317

HAL Id: hal-00202317

<https://hal.science/hal-00202317>

Submitted on 5 Jan 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Extension of First-Order Theories into Trees

Khalil Djelloul¹ and Thi-Bich-Hanh Dao²

¹ Faculty of Computer Science, University of Ulm, Germany.
DFG research project "Glob-Con"

² Laboratoire d'Informatique Fondamentale d'Orléans.
Bat. 3IA, rue Léonard de Vinci. 45067 Orléans, France.

Abstract. We present in this paper an automatic way to combine any first-order theory T with the theory of finite or infinite trees. First of all, we present a new class of theories that we call *zero-infinite-decomposable* and show that every decomposable theory T accepts a decision procedure in the form of six rewriting rules which for every first order proposition give either true or false in T . We present then the axiomatization T^* of the extension of T into trees and show that if T is flexible then its extension into trees T^* is zero-infinite-decomposable and thus complete. The flexible theories are theories having elegant properties which enable us to eliminate quantifiers in particular cases.

1 Introduction

The theory of finite or infinite trees plays a fundamental role in programming. Recall that Alain Colmerauer has described the execution of Prolog II, III and IV programs in terms of solving equations and disequations in this theory [6, 9, 2]. He has first introduced in Prolog II the unification of infinite terms together with a predicate of non-equality [8]. He has then integrated in Prolog III the domain of rational numbers together with the operations of addition and subtraction and a linear dense order relation without endpoints [5, 7]. He also gave a general algorithm to check the satisfiability of a system of equations, inequations and disequations on a combination of trees and rational numbers. Finally, in Prolog IV, the notions of list, interval and boolean have been added [10, 2].

We present in this paper an idea of a general extension of the model of Prolog IV by allowing the user to incorporate universal and existential quantifiers to Prolog closes and to decide the validity or not validity of any first-order proposition (sentence) in a combination of trees and first-order theories. For that:

- (1) we give an automatic way to generate the axiomatization of the combination of any first order theory T with the theory of finite or infinite trees,
- (2) we present simple conditions on T and only on T so that the combination of T with the theory of finite or infinite trees is complete and accepts a decision algorithm in the form of six rewriting rules which for every proposition give either true or false.

One of major difficulties in this work resides in the fact that the two theories can possibly have non-disjoint signatures. Moreover, the theory of finite or infinite trees does not accept full elimination of quantifiers.

The emergence of general constraint-based paradigms, such as constraint logic programming [19], constrained resolution [3] and what is generally referred to as *theory reasoning* [1], rises the problem of combining decision procedure for solving general first order constraints. Initial combinations results were provided by R. Shostak in [27] and in [28]. Shostak's approach is limited in scope and not very modular. A rather general and completely modular combination method was proposed by G. Nelson and D. Oppen in [21] and then slightly revised in [22]. Given, for $i = 1, \dots, n$ a procedure P_i that decides the satisfiability of quantifier-free formulas in the theory $T_1 \cup \dots \cup T_n$. A declarative and non-deterministic view of the procedure was suggested by Oppen in [24]. In [30], C. Tinelli and H. Harandi followed up on this suggestion describing a non-deterministic version of the Nelson-Oppen approach combination procedure and providing a simpler correctness proof. A similar approach had also been followed by C. Ringeissen in [26] which describes the procedure as a set of a derivation rules applied non-deterministically.

All the works mentioned above share one major restriction on the constraint languages of the component reasoners: they must have disjoint signatures, i.e. no function and relation symbols in common. (The only exception is the equality symbol which is however regarded as a logical constant). This restriction has proven really hard to lift. A testament of this is that, more than two decades after Nelson and Oppen's original work, their combination results are still state of the art.

Results on non-disjoint signatures do exists, but they are quit limited. To start with, some results on the union of non-disjoint equational theories can be obtained as a byproduct of the research on the combination of term rewriting systems. Modular properties of term rewriting systems have been extensively investigated (see the overviews in [23] and [18]). Using some of these properties it is possible to derive combination results for the word problem in the union of equational theories sharing constructors³. Outside the work on modular term rewriting, the first combination result for the word problem in the union of non-disjoint constraint theories were given in [16] as a consequence of some combination techniques based on an adequate notion of (shared) constructors. C. Ringeissen used similar ideas later in [25] to extend the Nelson-Oppen method to theories sharing constructors in a sense closed to that of [16].

Recently, C. Tinelli and C. Ringeissen have provided some sufficient conditions for the Nelson-Oppen combinability by using a concept of stable Σ -freeness [29], a natural extension of Nelson-Oppen's stable-infiniteness requirement for theories with non-disjoint signatures. As for us, we present a natural way to combine the theory of finite or infinite trees with any first order theory T which can possibly have a non-disjoint signature. A such theory is denoted by T^* and does not accept full elimination of quantifiers which makes the decision proce-

³ The word problem in an equational theory T is the problem of determining whether a given equation $s = t$ is valid in T , or equivalently, whether a disequation $\neg(s = t)$ is (un)satisfiable in T . In a term rewriting system, a constructor is a function symbol that does not appear as the top symbol of a rewrite rule's left-hand-side.

ture not evident. To show the completeness of T^* we give simple conditions on T and only on T so that its combination with the theory of finite or infinite trees, i.e. T^* , is complete and accepts a decision procedure which using only six rewriting rules is able to decide the validity or not validity of any first order constraints in T^* .

This paper is organized in five sections followed by a conclusion. This introduction is the first section. In Section 2, we recall the basic definitions of signature, model, theory and vectorial quantifier. In section 3, after having presented a new quantifier called *zero-infinite*, we present a new class of theories that we call *zero-infinite-decomposable*. The main idea behind this class of theories consists in decomposing each quantified conjunction of atomic formulas into three embedded sequences of quantifications having very particular properties, which can be expressed with the help of three special quantifiers denoted by $\exists?$, $\exists!$, $\exists_{o\infty}^{\psi(u)}$ and called *at-most-one*, *exactly-one*, *zero-infinite*. We end this section by six rewriting rules which for every zero-infinite-decomposable theory T and for every proposition φ give either true or false in T . The correctness of our algorithm shows the completeness of the zero-infinite-decomposable theories. In Section 4, we give a general way to generate the axioms of T^* using those of T and show that if T is flexible then T^* is zero-infinite-decomposable and thus complete. The flexible theories are theories having elegant properties which enable us to eliminate quantifiers in particular cases. We end this section by some fundamental flexible theories.

The zero-infinite-decomposable theories, the decision procedure in zero-infinite-decomposable theories, the axiomatization of T^* and the flexible theories are our main contribution in this paper. A full version of this paper with detailed proofs can be found in [13] and in the Ph.D thesis of K. Djelloul [14] (chapters 3 and 4).

2 Preliminaries

Let V be an infinite set of *variables*. Let S be a set of symbols, called a *signature* and partitioned into two disjoint sub-sets: the set F of *function symbols* and the set R of *relation symbols*. To each function symbol and relation is linked a non-negative integer n called *its arity*. An n -ary symbol is a symbol of arity n . A 0-ary function symbol is called a *constant*.

An *S-formula* is an expression of the one of the eleven following forms:

$$\begin{aligned} & s = t, \quad r t_1 \dots t_n, \quad \text{true}, \quad \text{false}, \\ & \neg\varphi, (\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \rightarrow \psi), (\varphi \leftrightarrow \psi), \\ & (\forall x \varphi), (\exists x \varphi), \end{aligned} \tag{1}$$

with $x \in V$, r an n -ary relation symbol taken from F , φ and ψ shorter *S*-formulas, s , t and the t_i 's *S*-terms, that are expressions of the one of the two following forms

$$x, \quad f t_1 \dots t_n,$$

with x taken from V , f an n -ary function symbol taken from F and the t_i shorter S -terms

The S -formulas of the first line of (1) are called *atomic*, and *flat* if they are of the one of the five following forms:

$$\text{true, false, } x_0 = f x_1 \dots x_n, x_0 = x_1, r x_1 \dots x_n,$$

with the x_i 's possibly non-distinct variables taken from V , $f \in F$ and $r \in R$.

If φ is an S -formula then we denote by $\text{var}(\varphi)$ the set of the free variables of φ . An S -proposition is an S -formula without free variables. The set of the S -terms and the S -formulas represent a *first-order language with equality*.

An S -structure is a couple $M = (D, F)$, where D is a non-empty set of *individuals* of M and F a set of functions and relations in D . We call *instantiation* or *valuation* of an S -formula φ by individuals of M , the $(S \cup D)$ -formula obtained from φ by replacing each free occurrence of a free variable x in φ by the same individual i of D and by considering each element of D as 0-ary function symbol.

An S -theory T is a set of S -propositions. We say that the S -structure M is a model of T if for each element φ of T , $M \models \varphi$. If φ is an S -formula, we write $T \models \varphi$ if for each S -model M of T , $M \models \varphi$. A theory T is called *complete* if for every proposition φ , one and only one of the following properties holds: $T \models \varphi$, $T \models \neg\varphi$.

Let M be a model. Let $\bar{x} = x_1 \dots x_n$ and $\bar{y} = y_1 \dots y_n$ be two words on \mathbf{v} of the same length. Let φ and $\varphi(\bar{x})$ be M -formulas. We write

$$\begin{aligned} \exists \bar{x} \varphi & \quad \text{for } \exists x_1 \dots \exists x_n \varphi, \\ \forall \bar{x} \varphi & \quad \text{for } \forall x_1 \dots \forall x_n \varphi, \\ \exists ? \bar{x} \varphi(\bar{x}) & \quad \text{for } \forall \bar{x} \forall \bar{y} \varphi(\bar{x}) \wedge \varphi(\bar{y}) \rightarrow \bigwedge_{i \in \{1, \dots, n\}} x_i = y_i, \\ \exists ! \bar{x} \varphi & \quad \text{for } (\exists \bar{x} \varphi) \wedge (\exists ? \bar{x} \varphi). \end{aligned}$$

The word \bar{x} , which can be the empty word ε , is called *vector of variables*. Note that the formulas $\exists ? \varepsilon \varphi$ and $\exists ! \varepsilon \varphi$ are respectively equivalent to *true* and to φ in any model M .

3 Zero-infinite-decomposable theories

In this section, let us fix a signature $S^* = F^* \cup R^*$. Thus, we can allow ourself to remove the prefix S^* from the following words: formulas, equations, theories and models. We will also use the abbreviation *wfnv* for “*without new free variables*”. We say that an S -formula φ is equivalent to a wfnv S -formula ψ in T if $T \models \varphi \leftrightarrow \psi$ and ψ does not contain other free variables than those of φ .

3.1 Zero-infinite quantifier [15]

Let M be a model and T a theory. Let $\Psi(u)$ be a set of formulas having at most one free variable u . Let φ and φ_j be M -formulas. We write

$$M \models \exists_o^{\Psi(u)} x \varphi(x), \tag{2}$$

if for each instantiation $\exists x \varphi'(x)$ of $\exists x \varphi(x)$ by individuals of M one of the following properties holds:

- the set of the individuals i of M such that $M \models \varphi'(i)$, is infinite,
- for every finite sub-set $\{\psi_1(u), \dots, \psi_n(u)\}$ of elements of $\Psi(u)$, the set of the individuals i of M such that $M \models \varphi'(i) \wedge \bigwedge_{j \in \{1, \dots, n\}} \neg \psi_j(i)$ is infinite.

We write $T \models \exists_o^\Psi x \varphi(x)$, if for every model M of T we have $M \models \exists_o^\Psi x \varphi(x)$.

This infinite quantifier holds only for infinite models, i.e. models whose set of elements are infinite. Note that if $\Psi(u) = \{\text{false}\}$ then (2) simply means that if $M \models \exists x \varphi(x)$ then M contains an infinity of individuals i such that $M \models \varphi(i)$. The intuitions behind this definition come from an aim to eliminate a conjunction of the form $\bigwedge_{i \in I} \neg \psi_i(x)$ in complex formulas of the form $\exists \bar{x} \varphi(x) \wedge \bigwedge_{i \in I} \neg \psi_i(x)$ where I is a finite (possibly empty) set and the $\psi_i(x)$ are formulas which do not accept full elimination of quantifiers.

3.2 Zero-infinite-decomposable theory [15]

A theory T is called *zero-infinite-decomposable* if there exists a set $\Psi(u)$ of formulas, having at least one free variable u , a set A of formulas closed under conjunction, a set A' of formulas of the form $\exists \bar{x} \alpha$ with $\alpha \in A$, and a sub-set A'' of A such that:

1. every formula of the form $\exists \bar{x} \alpha \wedge \psi$, with $\alpha \in A$ and ψ a formula, is equivalent in T to a wnfv formula of the form:

$$\exists \bar{x}' \alpha' \wedge (\exists \bar{x}'' \alpha'' \wedge (\exists \bar{x}''' \alpha''' \wedge \psi)),$$

with $\exists \bar{x}' \alpha' \in A'$, $\alpha'' \in A''$, $\alpha''' \in A$ and $T \models \forall \bar{x}'' \alpha'' \rightarrow \exists! \bar{x}''' \alpha'''$,

2. if $\exists \bar{x}' \alpha' \in A'$ then $T \models \exists ? \bar{x}' \alpha'$ and for every free variable y in $\exists \bar{x}' \alpha'$, one at least of the following properties holds:
 - $T \models \exists ? y \bar{x}' \alpha'$,
 - there exists $\psi(u) \in \Psi(u)$ such that $T \models \forall y (\exists \bar{x}' \alpha') \rightarrow \psi(y)$,
3. if $\alpha'' \in A''$ then
 - the formula $\neg \alpha''$ is equivalent in T to a wnfv formula of the form $\bigvee_{i \in I} \alpha_i$ with $\alpha_i \in A$,
 - for every x'' , the formula $\exists x'' \alpha''$ is equivalent in T to a wnfv formula which belongs to A'' ,
 - for every variable x'' , $T \models \exists_o^\Psi x'' \alpha''$,
4. every conjunction of flat formulas is equivalent in T to a wnfv disjunction of elements of A ,
5. if the formula $\exists \bar{x}' \alpha' \wedge \alpha''$ with $\exists \bar{x}' \alpha' \in A'$ and $\alpha'' \in A''$ has no free variables then \bar{x} is the empty vector, α' is the formula *true* and α'' is either the formula *true* or *false*.

3.3 A decision procedure for zero-infinite-decomposable theories [14]

Let T be a zero-infinite-decomposable theory. The sets $\Psi(u)$, A , A' and A'' are known and fixed.

Definition 3.3.1 A normalized formula φ of depth $d \geq 1$ is a formula of the form $\neg(\exists \bar{x} \alpha \wedge \bigwedge_{i \in I} \varphi_i)$, where I is a finite possibly empty set, $\alpha \in A$, the φ_i are normalized formulas of depth d_i with $d = 1 + \max\{0, d_1, \dots, d_n\}$, and all the quantified variables have distinct names and different from those of the free variables.

Property 3.3.2 Every formula is equivalent in T to a normalized formula.

Definition 3.3.3 A final formula is a normalized formula of the form

$$\neg(\exists \bar{x}' \alpha' \wedge \alpha'' \wedge \bigwedge_{i \in I} \neg(\exists \bar{y}'_i \beta'_i)), \quad (3)$$

with I a finite possibly empty set, $\exists \bar{x}' \alpha' \in A'$, $\alpha'' \in A''$, $\exists \bar{y}'_i \beta'_i \in A'$, α'' is different from the formula false, all the β'_i 's are different from the formulas true and false.

Property 3.3.4 Let φ be a conjunction of final formulas without free variables. The conjunction φ is either the formula true or the formula \neg true.

Property 3.3.5 Every normalized formula is equivalent in T to a conjunction of final formulas.

Proof. We give bellow six rewriting rules which transform a normalized formula of any depth d into a conjunction of final formulas equivalent in T . To apply the rule $p_1 \implies p_2$ on a normalized formula p means to replace in p , the subformula p_1 by the formula p_2 , by considering the connector \wedge associative and commutative.

$$\begin{aligned}
 (1) \quad & \neg \left[\begin{array}{l} \exists \bar{x} \alpha \wedge \varphi \wedge \\ \neg(\exists \bar{y} \text{ true}) \end{array} \right] \implies \text{true} \\
 (2) \quad & \neg \left[\exists \bar{x} \alpha \wedge \text{false} \wedge \varphi \right] \implies \text{true} \\
 (3) \quad & \neg \left[\begin{array}{l} \exists \bar{x} \alpha \wedge \\ \bigwedge_{i \in I} \neg(\exists \bar{y}_i \beta_i) \end{array} \right] \implies \neg \left[\begin{array}{l} \exists \bar{x}' \bar{x}'' \alpha' \wedge \alpha'' \wedge \\ \bigwedge_{i \in I} \neg(\exists \bar{x}''' \bar{y}_i \alpha''' \wedge \beta_i) \end{array} \right] \\
 (4) \quad & \neg \left[\begin{array}{l} \exists \bar{x} \alpha \wedge \varphi \wedge \\ \neg(\exists \bar{y}' \beta' \wedge \beta'') \end{array} \right] \implies \left[\begin{array}{l} \neg(\exists \bar{x} \alpha \wedge \varphi \wedge \neg(\exists \bar{y}' \beta')) \wedge \\ \bigwedge_{i \in I} \neg(\exists \bar{x} \bar{y}' \alpha \wedge \beta' \wedge \beta'' \wedge \varphi) \end{array} \right] \\
 (5) \quad & \neg \left[\begin{array}{l} \exists \bar{x} \alpha \wedge \\ \bigwedge_{i \in I} \neg(\exists \bar{y}'_i \beta'_i) \end{array} \right] \implies \neg \left[\begin{array}{l} \exists \bar{x}' \alpha' \wedge \alpha''_* \\ \bigwedge_{i \in I'} \neg(\exists \bar{y}'_i \beta'_i) \end{array} \right] \\
 (6) \quad & \neg \left[\begin{array}{l} \exists \bar{x} \alpha \wedge \varphi \wedge \\ \neg \left[\begin{array}{l} \exists \bar{y}' \beta' \wedge \beta'' \wedge \\ \bigwedge_{i \in I} \neg(\exists \bar{z}'_i \delta'_i) \end{array} \right] \end{array} \right] \implies \left[\begin{array}{l} \neg(\exists \bar{x} \alpha \wedge \varphi \wedge \neg(\exists \bar{y}' \beta' \wedge \beta'')) \wedge \\ \bigwedge_{i \in I} \neg(\exists \bar{x} \bar{y}' \bar{z}_i \alpha \wedge \beta' \wedge \beta'' \wedge \delta'_i \wedge \varphi) \end{array} \right]
 \end{aligned}$$

with α an element of A , φ a conjunction of normalized formulas and I a finite possibly empty set. In the rule (3), the formula $\exists \bar{x} \alpha$ is equivalent in T to a decomposed formula of the form $\exists \bar{x}' \alpha' \wedge (\exists \bar{x}'' \alpha'' \wedge (\exists \bar{x}''' \alpha'''))$ with $\exists \bar{x}' \alpha' \in A'$, $\alpha'' \in A''$, $\alpha''' \in A$, $T \models \forall \bar{x}'' \alpha'' \rightarrow \exists! \bar{x}''' \alpha'''$ and $\exists \bar{x}''' \alpha'''$ different from $\exists \varepsilon \text{ true}$. All the β_i 's belong to A . In the rule (4), the formula $\exists \bar{x} \alpha$ is equivalent in T to a decomposed formula of the form $\exists \bar{x}' \alpha' \wedge (\exists \bar{x}'' \alpha'' \wedge (\exists \varepsilon \text{ true}))$ with $\exists \bar{x}' \alpha' \in A'$ and $\alpha'' \in A''$. The formula $\exists \bar{y}' \beta'$ belongs to A' . The formula β'' belongs to A'' and is different from the formula true . Moreover, $T \models (\neg \beta'') \leftrightarrow \bigvee_{i \in I} \beta''_i$ with $\beta''_i \in A$. In the rule (5), the formula $\exists \bar{x} \alpha$ is not of the form $\exists \bar{x} \alpha_1 \wedge \alpha_2$ with $\exists \bar{x} \alpha_1 \in A'$ and $\alpha_2 \in A''$, and is equivalent in T to a decomposed formula of the form $\exists \bar{x}' \alpha' \wedge (\exists \bar{x}'' \alpha'' \wedge (\exists \varepsilon \text{ true}))$ with $\exists \bar{x}' \alpha' \in A'$ and $\alpha'' \in A''$. Each formula $\exists \bar{y}'_i \beta'_i$ belongs to A' . The set I' is the set of the $i \in I$ such that $\exists \bar{y}'_i \beta'_i$ does not contain free occurrences of a variable of \bar{x}'' . Moreover, $T \models (\exists \bar{x}'' \alpha'') \leftrightarrow \alpha''_*$ with $\alpha''_* \in A''$. In the rule (6), $I \neq \emptyset$, $\exists \bar{y}' \beta' \in A'$, $\exists \bar{z}'_i \delta'_i \in A'$ and $\beta'' \in A''$.

Let ψ be a formula without free variables, the *decision* of ψ proceeds as follows:

1. Transform the formula ψ into a normalized formula φ which is equivalent to ψ in T .
2. While it is possible, apply the rewriting rules on φ . At the end, we obtain a conjunction ϕ of final formulas.

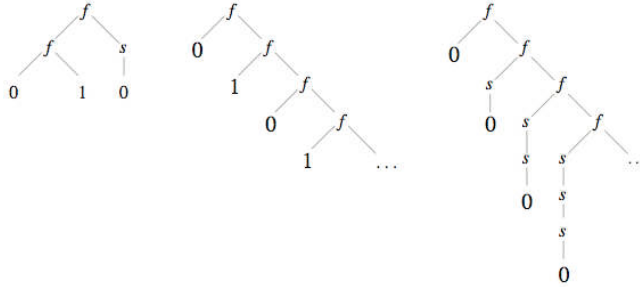
According to Property 3.3.5, the application of the rules on a formula ψ without free variables produces a wnfv conjunction ϕ of final formulas, i.e. a conjunction ϕ of final formulas without free variables. According to Property 3.3.4, ϕ is either the formula true , or the formula $\neg \text{true}$, i.e. the formula false .

Corollary 3.3.6 *If T is zero-infinite-decomposable then T is complete and accepts a decision procedure in the form of six rewriting rules which for every proposition give either true or false in T .*

4 Extension of first-order theories into trees

4.1 The structure of finite or infinite trees

Trees are well known objects in the computer science world. Here are some of them:



Their nodes are labeled by the symbols 0,1,s,f, of respective arities 0,0,1,2, taken

from a set F of functional symbols which we assume to be infinite. While the first tree is a *finite tree* (it has a finite set of nodes), the two others are *infinite trees* and have an infinite set of nodes. We denote by A the set of all trees⁴ constructed on F .

We introduce in A a set of construction operations⁵, one for each element $f \in F$, which is the mapping $(a_1, \dots, a_n) \rightarrow b$, where n is the arity of f and b the tree whose initial node is labeled by f and the sequence of sons is (a_1, \dots, a_n) , and which be schematized as:



We thus obtain *the structure of finite or infinite trees* constructed on F , which we denote by (A, F) .

4.2 Theory of finite or infinite trees

Let S be a signature containing only an infinite set of function symbols F . Michael Maher has introduced the S -theory of finite or infinite trees [20]. The axiomatization of this S -theory is the set of the S -propositions of the one of the following forms:

- 1 $\forall \bar{x} \forall \bar{y} f \bar{x} = f \bar{y} \rightarrow \bigwedge_i x_i = y_i$,
- 2 $\forall \bar{x} \forall \bar{y} \neg f \bar{x} = g \bar{y}$,
- 3 $\forall \bar{x} \exists ! \bar{z} \bigwedge_i z_i = f_i(\bar{z}, \bar{x})$,

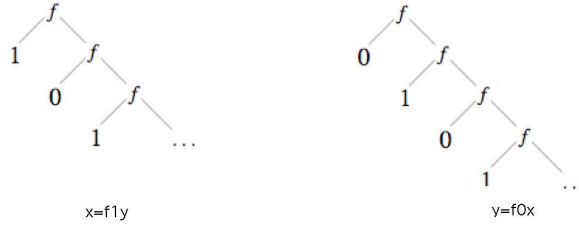
with f and g two distinct function symbols taken from F , \bar{x} a vector of variables x_i , \bar{y} a vector of variables y_i , \bar{z} a vector of distinct variables z_i and $f_i(\bar{x}, \bar{z})$ an S -term which begins with an element of F followed by variables taken from $\bar{x}\bar{z}$.

The first axiom is called *axiom of explosion*, the second one is called *axiom of conflict of symbols* and the last one is called *axiom of unique solution*.

We show that this theory has as model the structure of finite or infinite trees [12]. For example, using axiom 3, we have $T \models \exists ! x y x = f1y \wedge y = f0x$. The individuals x and y represents the two following trees in the structure of finite or infinite trees:

⁴ More precisely, we define first a node to be a word constructed on the set of strictly positive integers. A *tree* a on F , is then a mapping of type $a : E \rightarrow F$, where E is a non-empty set of nodes, each one $i_1 \dots i_k$ (with $k \geq 0$) satisfies two conditions: (1) if $k > 0$ then $i_1 \dots i_{k-1} \in E$ and (2) if the arity of $a(i_1 \dots i_k)$ is n , then the set of the nodes E of the form $i_1 \dots i_k i_{k+1}$ is obtained by giving to i_{k+1} the values $1, \dots, n$.

⁵ In fact, the construction operation linked to the n -ary symbol f of F is the mapping $(a_1, \dots, a_n) \rightarrow b$, where the a_i 's are any trees and b is the tree defined as follows from the a_i 's and their set of nodes E_i 's: the set E of nodes of a is $\{\varepsilon\} \cup \{ix | x \in E_i \text{ and } i \in \{1, \dots, n\}\}$ and, for each $x \in E$, if $x = \varepsilon$, then $a(x) = f$ and if x is of the form iy , with i being an integer, $a(x) = a_i(y)$.



4.3 Axiomatization of the theory $T + Tree$ or T^*

Let us fix now a signature S containing a set F of function symbols and a set R of relation symbols, as well as a signature S^* containing:

- an infinite set $F^* = F \cup F_A$ where F_A is an infinite set of function symbols disjoint from F .
- a set $R^* = R \cup \{p\}$ of relation symbols, containing R , and an 1-ary relation symbol p .

Let T be an S -theory. The extension of the S -theory T into trees is the S^* -theory denoted by T^* and whose set of axioms is the infinite set of the following S^* -propositions, with \bar{x} a vector of variables x_i and \bar{y} a vector of variables y_i :

1. Explosion: for all $f \in F^*$:

$$\forall \bar{x} \forall \bar{y} \neg p f \bar{x} \wedge \neg p f \bar{y} \wedge f \bar{x} = f \bar{y} \rightarrow \bigwedge_i x_i = y_i$$

2. Conflict of symbols: Let f and g be two distinct function symbols taken from F^* :

$$\forall \bar{x} \forall \bar{y} f \bar{x} = g \bar{y} \rightarrow p f \bar{x} \wedge p g \bar{y}$$

3. Unique solution

$$\forall \bar{x} \forall \bar{y} \left(\bigwedge_i p x_i \right) \wedge \left(\bigwedge_j \neg p y_j \right) \rightarrow \exists! \bar{z} \bigwedge_k (\neg p z_k \wedge z_k = t_k(\bar{x}, \bar{y}, \bar{z}))$$

where \bar{z} is a vector of distinct variables z_i , $t_k(\bar{x}, \bar{y}, \bar{z})$ an S^* -term which begins by a function symbol $f_k \in F^*$ followed by variables taken from $\bar{x}, \bar{y}, \bar{z}$, moreover, if $f_k \in F$, then the S^* -term $t_k(\bar{x}, \bar{y}, \bar{z})$ contains at least a variable from \bar{y} or \bar{z}

4. Relations of R : for all $r \in R$,

$$\forall \bar{x} r \bar{x} \rightarrow \bigwedge_i p x_i$$

5. Operations of F : for all $f \in F$,

$$\forall \bar{x} p f \bar{x} \leftrightarrow \bigwedge_i p x_i$$

(if f is 0-ary then this axiom is written $p f$)

6. Elements not in T : for all $f \in F^* - F$,

$$\forall \bar{x} \neg pf\bar{x}$$

7. Existence of an element satisfying p (only if F does not contain 0-ary function symbols):

$$\exists x px,$$

8. Extension into trees of the axioms of T : all axioms obtained by the following transformations of each axiom φ of T : While it is possible to replace all sub-formula of φ which is of the form $\exists \bar{x} \psi$, but not of the form $\exists \bar{x} (\bigwedge px_i) \wedge \psi'$, by $\exists \bar{x} (\bigwedge px_i) \wedge \psi$ and every sub-formula of φ which is of the form $\forall \bar{x} \psi$, but not of the form $\forall \bar{x} (\bigwedge px_i) \rightarrow \psi'$, by $\forall \bar{x} (\bigwedge px_i) \rightarrow \psi$.

4.4 Example: Extension of linear dense order relations without endpoints into trees

Let F be an empty set of function symbols and let R be a set of relation symbols containing only the relation symbol $<$ of arity 2. If t_1 and t_2 are terms, then we write $t_1 < t_2$ for $<(t_1, t_2)$. Let T_{ord} be the theory of linear dense order relation without endpoints, whose signature is $S = F \cup R$ and whose axioms are the following propositions:

- 1 $\forall x \neg x < x$,
- 2 $\forall x \forall y \forall z (x < y \wedge y < z) \rightarrow x < z$,
- 3 $\forall x \forall y x < y \vee x = y \vee y < x$,
- 4 $\forall x \forall y x < y \rightarrow (\exists z x < z \wedge z < y)$,
- 5 $\forall x \exists y x < y$,
- 6 $\forall x \exists y y < x$.

Let now F^* be an infinite set of function symbols and $R^* = \{<, p\}$ a set of relation symbols containing the 2-ary relation symbol $<$ and the 1-ary relation symbol p . Let S^* be the signature $F^* \cup R^*$. According to the transformations of axioms given in Section 4.3, the axiomatization of the extension of the theory T_{ord} into trees is the S^* -theory T_{ord}^* whose axioms are the following propositions:

- 1 $\forall \bar{x} \forall \bar{y} \neg pf\bar{x} \wedge \neg pf\bar{y} \wedge f\bar{x} = f\bar{y} \rightarrow \bigwedge_i x_i = y_i$
- 2 $\forall \bar{x} \forall \bar{y} f\bar{x} = g\bar{y} \rightarrow pf\bar{x} \wedge pg\bar{y}$
- 3 $\forall \bar{x} \forall \bar{y} (\bigwedge_i px_i) \wedge (\bigwedge_j \neg py_j) \rightarrow \exists! \bar{z} \bigwedge_k (\neg pz_k \wedge z_k = f_k(\bar{x}, \bar{y}, \bar{z}))$
- 4 $\forall x \forall y x < y \rightarrow (px \wedge py)$,
- 5 $\forall \bar{x} \neg pf\bar{x}$,
- 6 $\exists x px$,
- 7 $\forall x px \rightarrow \neg x < x$,
- 8 $\forall x \forall y \forall z px \wedge py \wedge pz \rightarrow ((x < y \wedge y < z) \rightarrow x < z)$,
- 9 $\forall x \forall y (px \wedge py) \rightarrow (x < y \vee x = y \vee y < x)$,
- 10 $\forall x \forall y (px \wedge py) \rightarrow (x < y \rightarrow (\exists z pz \wedge x < z \wedge z < y))$,
- 11 $\forall x px \rightarrow (\exists y py \wedge x < y)$,
- 12 $\forall x px \rightarrow (\exists y py \wedge y < x)$,

where f and g are distinct function symbols taken from F^* , x, y, z variables, \bar{x} a vector of variables x_i , \bar{y} a vector of variables y_i , \bar{z} a vector of distinct variables z_i and $f_k(\bar{x}, \bar{y}, \bar{z})$ a term which begins by an element f_k of F^* followed by variables taken from $\bar{x}\bar{y}\bar{z}$.

5 Completeness of T^*

We have given a general axiomatization of T^* using the axioms of T , what about the completeness of T^* ? Are all the extensions into trees complete theories? While in [15] we have shown the completeness of a combination of trees and rational numbers, in this paper the challenge is to use general properties that hold not only for rational numbers but for a large set of different theories T_i and that make T_i^* zero-infinite-decomposable and thus complete.

Let $S = F \cup R$ be a signature and T an S -theory. Let $S^* = F^* \cup R^*$ be another signature with F^* an infinite set of function symbols containing F and $R^* = R \cup \{p\}$. Let T^* be the S^* -theory of the extension of T into trees. Suppose that the variables of V are ordered by a linear dense order relation without endpoints denoted \succ .

5.1 Flexible theory

Definition 5.1.1 We call leader of an S -equation α the greatest variable x in α , according to the order \succ , such that $T \models \exists!x\alpha$.

Definition 5.1.2 A conjunction of S -atomic formulas α is called formatted in T if

- α does not contain sub-formulas of the form $f_1 = f_2$ or $rf_1\dots f_n$ or $y = x$, where all the f_i 's are 0-ary function symbols taken from F , $r \in R$ and $x \succ y$,
- each S -equation of α has a distinct leader which has no occurrences in other S -equations or S -relations of α ,
- if α' is the conjunction of all the S -equations of α then for all $x \in \text{var}(\alpha')$ we have $T \models \exists?x\alpha'$.

Definition 5.1.3 The theory T is called flexible if for each conjunction α of S -equations and for each conjunction β of S -relations:

1. $\alpha \wedge \beta$ is equivalent in T to a formatted conjunction of atomic formulas wnfv,
2. the S -formula $\neg\beta$ is equivalent in T to a disjunction wnfv of S -equations and S -relations,
3. for all $x \in V$
 - the S -formula $\exists x\beta$ is equivalent in T to false, or to a wnfv conjunction of S -relations,
 - for all $x \in V$, we have $T \models \exists_o^\infty\{f^{aux}\}x\beta$.

Let us now present our main result

Theorem 5.1.4 If T is flexible then T^* is zero-infinite-decomposable.

5.2 Some fundamental flexible theories:

We present in this section the axiomatization of some fundamental flexible theories. Full proofs can be found in [14].

Infinite Clark equational theory: Let Cl be a theory together with an empty set of function and relation symbols and whose axioms is the infinite set of propositions of the following form:

$$(1_n) \quad \forall x_1 \dots \forall x_n \exists y \neg(x_1 = y) \wedge \dots \wedge \neg(x_n = y), \quad (4)$$

where all the variables $x_1 \dots x_n$ are distinct and ($n \neq 0$). The form (4) is called *diagram of axiom* and for each value of n there exists *an axiom* of Cl . This theory Cl has been introduced by Clark [4] and has an infinite set of models each one containing an infinite set of distinct individuals.

Additive rational or real numbers theory with addition and subtraction: Let $F = \{+, -, 0, 1\}$ be a set of function symbols of respective arities 2, 1, 0, 0. Let $R = \emptyset$ be an empty set of relation symbols. The theory Ra of additive rational or real numbers together with addition and subtraction consists in the infinite set of propositions of the following form:

$$\begin{aligned} 1 \quad & \forall x \forall y \ x + y = y + x, \\ 2 \quad & \forall x \forall y \forall z \ x + (y + z) = (x + y) + z, \\ 3 \quad & \forall x \ x + 0 = x, \\ 4 \quad & \forall x \ x + (-x) = 0, \\ 5_n \quad & \forall x \ n \cdot x = 0 \rightarrow x = 0, \\ 6_n \quad & \forall x \ \exists! y \ n \cdot y = x, \\ 7 \quad & \forall x \forall y \forall z \ (x = y) \leftrightarrow (x + z = y + z), \\ 8 \quad & \neg(0 = 1). \end{aligned}$$

with n an non-null integer. This theory has two usual models: rational numbers Q with addition and subtraction in Q and real numbers R with addition and subtraction in R .

Linear dense order theory without endpoints: Let F be an empty set of function symbols and R a set of relation symbols containing only the binary relation symbol $<$. The theory T_{ord} be the theory of the linear dense order without endpoints consists in the set of propositions of the following form:

$$\begin{aligned} 1 \quad & \forall x \ \neg x < x, \\ 2 \quad & \forall x \forall y \forall z \ (x < y \wedge y < z) \rightarrow x < z, \\ 3 \quad & \forall x \forall y \ x < y \vee x = y \vee y < x, \\ 4 \quad & \forall x \forall y \ x < y \rightarrow (\exists z \ x < z \wedge z < y), \\ 5 \quad & \forall x \ \exists y \ x < y, \\ 6 \quad & \forall x \ \exists y \ y < x. \end{aligned}$$

Ordered additive rational or real numbers theory with addition and subtraction: Let $F = \{+, -, 0, 1\}$ be a set of function symbols of respective arities 2, 1, 0, 0. Let $R = \{<\}$ be a set of relation symbols containing only the binary relation symbol $<$. The theory T_{ad} of ordered additive rational or real numbers theory with addition and subtraction consists in the infinite set of propositions of the following form:

- 1 $\forall x \forall y x + y = y + x,$
- 2 $\forall x \forall y \forall z x + (y + z) = (x + y) + z,$
- 3 $\forall x x + 0 = x,$
- 4 $\forall x x + (-x) = 0,$
- 5 _{n} $\forall x n.x = 0 \rightarrow x = 0, \quad (n \neq 0)$
- 6 _{n} $\forall x \exists! y n.y = x, \quad (n \neq 0)$
- 7 $\forall x \neg x < x,$
- 8 $\forall x \forall y \forall z (x < y \wedge y < z) \rightarrow x < z,$
- 9 $\forall x \forall y (x < y \vee x = y \vee y < x),$
- 10 $\forall x \forall y x < y \rightarrow (\exists z x < z \wedge z < y),$
- 11 $\forall x \exists y x < y,$
- 12 $\forall x \exists y y < x,$
- 13 $\forall x \forall y \forall z x < y \rightarrow (x + z < y + z),$
- 14 $0 < 1.$

with n a non-null integer.

6 Conclusion

We have defined in this paper a general idea for the extension of the models of Prolog by giving an automatic way to combine any first order theory T with the theory of finite or infinite trees. To show the completeness of T^* we have introduced the flexible theories and have shown that if T is flexible then T^* zero-infinite-decomposable. The zero-infinite-decomposable theories are first order theories having elegant properties which enable us to decide the validity of any proposition using only six rewriting rules. The main idea behind this rules consists in a local decomposition of quantified conjunctions of hybrid atomic formulas, a partial elimination of quantifiers using the properties of the vectorial quantifiers, and a special distribution to decrease the depth of the formulas.

There exists many practical applications of the extensions into trees of first order theories. First-order constraints on trees can be expressed in a simpler way when they are in the extension into trees of another structure. For example, the constraints representing the moves in two players games introduced by Alain Colmerauer and Thi-Bich-Hanh Dao [11, 12] can be represented by a simpler constraint in the extension into trees of the integers together with the operations of addition and subtraction and a linear dense order relation.

On the other hand, our decision algorithm can decide the validity or not validity of big and complex propositions and can also be applied on formulas having free variables and produces in this case a boolean combination of basic

formulas which does not accept full elimination of quantifiers. Unfortunately, this algorithm is not able to detect formulas having free variables and being always equivalent to *false* or *true* in T^* . It does not warrant that a final formula having at least one free variable is neither true nor false in T^* and can not present the solutions of the free variables in a clear and explicit way. This is why our algorithm is called *decision procedure* and not a general algorithm for solving first order constraints. It would be interesting to transform our decision procedure into a general algorithm for solving any first order constraint in T^* and which presents the solutions of the free variables in a clear and explicit way, as it has been done in [11, 12] for the theory of finite trees and finite or infinite trees. This kind of algorithm needs another work completely different from this one, by introducing syntactic and semantic definitions much more complex than the definition of flexible theories given in this paper. The implementation of a such algorithm will enable us to extend the Prolog language by allowing the user to solve any complex first order constraint, with or without free variables, in many combinations of theories around trees.

Currently, we are trying to proof that every extension of a complete theory into trees is complete and may be zero-infinite-decomposable. For that, we expect to add new vectorial quantifiers in the decomposition such as \exists^n which means *there exists n* and $\exists_{n,\infty}^{\psi(u)}$ which means *there exists n or infinite*, in order to increase the size of the set of the zero-infinite-decomposable theories and may be get a much more simple definition than the one defined in this paper. We plan also with Thom Fruehwirth [17] to add to CHR a general mechanism to treat our normalized formulas. This will enable us to implement quickly and easily our algorithms and get a general idea on the expressiveness of first order constraints in combinations of trees and first order theories.

Acknowledgments We thank Alain Colmerauer for our many discussions and his help in this work. We thank him too for the quality of her remarks and advice on how to improve the organization of this paper. We dedicate to him this paper with our best wishes for a speedy recovery.

References

1. Baumgartner, T., Furbach, U. and Petermann, U. A unified approach to theory reasoning. Research report. pages 15-92. Univesitat Koblenz, Germany. 1992.
2. Benhamou, F., Colmerauer, A. and Van caneghem, M. the manuel of Prolog IV, PrologIA, Marseille, France, 1996.
3. Burckert, H.J. A resolution principle for constraint logics. Artificial intelligence. 66, pages 235-271. 1994.
4. Clark, K.L. Negation as failure. Logic and Data bases. Ed Gallaire, H. and Minker, J. Plenum Pub. 1978.
5. Colmerauer, A. Total precedence relations. JACM. Vol 17, pages 14-30. 1970
6. Colmerauer, A. Equations and inequations on finite and infinite trees. Proc. of the Int. Conf. on the Fifth Generation of Computer Systems, pages 85-99. Tokyo, 1984.
7. Colmerauer, A. Prolog in 10 figures. Communications of the ACM. Vol 28(12), pages 1296-1310. 1985.
8. Colmerauer, A. A view of the origins and development of Prolog. Communications of the ACM. Vol 31, pages 26-36. 1988.

9. Colmerauer, A. An introduction to Prolog III. *Communication of the ACM*, 33(7):68-90,1990.
10. Colmerauer, A. Specification of Prolog IV. LIM Technical Report. Marseille university, France. 1996.
11. Colmerauer A. and Dao, TBH., Expressiveness of full first-order constraints in the structure of finite or infinite trees, *Constraints*, Vol. 8, No. 3, 2003, pages 283-302.
12. Dao, TBH. Résolution de contraintes du premier ordre dans la théorie des arbres finis or infinis. Thèse d'informatique, Université de la Méditerranée, décembre 2000.
13. Djelloul, K. and Dao, TBH. Extension of first order theories into trees. Rapport de recherche LIFO RR-2006-06, <http://www.univ-orleans.fr/lifo/rapports.php>.
14. Djelloul, K. Complete theories around trees. Ph.D. thesis. Université de la Méditerranée, June 2006. (<http://khalil.djelloul.free.fr>)
15. Djelloul, K. About the combination of trees and rational numbers in a complete first-order theory. Proceeding of the 5th Int. Conf. on Frontiers of Combining Systems FroCoS 2005, Vienna Austria. LNAI, vol 3717, pages. 106–122.
16. Domenjoud, E., Klay, F. and Ringeissen, C. Combination techniques for non-disjoint equational theories. In proceedings of the 12th international conference on automated deduction. LNAI 814: 267-281. 1994.
17. Frühwirth, T. and Abdelnadhher, S. Essentials of constraints programming. Springer Cognitive technologies. 2002.
18. Gramlich, B. On termination and confluence properties of disjoint and constructor sharing conditional rewrite systems. *Theoretical computer science* 165(1): 97-131. 1996.
19. Jaffar, J. and Maher, M. Constraint logic programming: a survey. *Journal of logic programming*. 19/20, pages 503-581. 1994.
20. Maher, M. Complete axiomatization of the structure of finite, rational and infinite trees. *Technical report, IBM*, 1988.
21. Nelson, G. and Oppen, D. Simplification by co-operating decision procedures. *ACM Transaction on programming languages* 1(2):245-251. 1979.
22. Nelson, G. Combining satisfiability procedures by equality sharing. *Automated theorem proving: After 25 years* vol 29: 201-211. 1984.
23. Ohelbush, E. Modular properties of composable term rewriting systems. *Journal of symbolic computation*, 20(1): 1-41. 1995.
24. Oppen, D. Complexity, convexity and combinations of theories. *Theoretical computer science*. Vol 12. 1980.
25. Ringeissen, C. co-operation of decision procedure for the satisfiability problem. In proceedings of the 1st international workshop on frontiers of combining systems. Kluwer pages 121-140. 1996.
26. Ringeissen, C. Combinaison de resolution de contraintes. These de doctorat, universite de Nancy. 1993.
27. Shostak, R. A practical decision procedure for arithmetic with function symbols. *Journal of the ACM* 26(2): 351-360. 1979.
28. Shostak, R. Deciding combinations of theories. *Journal of the ACM* 31: 1-12. 1984.
29. Tinelli, C. and Ringeissen, C. Unions of non-disjoint theories and combinations of satisfiability procedures. *Theor. Comput. Sci.* 290 (1): 291-353. 2003.
30. Tinelli, C. and Harandi, M. A new correctness proof for the Nelson-Oppen combination procedure. In proceedings of the 1st international workshop on frontiers of combining systems. Kluwer pages 103-120. 1996.