Complete First-Order Axiomatization of Finite or Infinite M-extended Trees

Khalil Djelloul[1] and Thi-Bich-Hanh Dao[2]

[1] Laboratoire d'Informatique Fondamentale de Marseille.
Parc scientifique et technologique de Luminy. 13288 Marseille, France.
[2] Laboratoire d'Informatique Fondamentale d'Orléans.
Bat. 3IA, rue Léonard de Vinci. 45067 Orléans, France.

Abstract. We present in this paper an axiomatization of the structure of finite or infinite M-extended trees. This structure is an intuitive combination of the structure of finite or infinite trees with another structure M and expresses semantically an extension to trees of the model M. Having a structure $M = (D_M, F_M, R_M)$, we define the structure of finite or infinite M-extended trees $Ext_M = (D, F, R)$ whose domain D consists of trees labelled by elements of $D_M \cup F$, where F is an infinite set of function symbols containing F_M and another infinite set of function symbols disjoint from F_M. For each n-ary function symbol $f \in F$, the operation $f(a_1, .., a_n)$ is evaluated in M and produces an element of D_M if $f \in F_M$ and all the a_i are elements of D_M, or is a tree whose root is labelled by f and whose immediate children are $a_1, .., a_n$ otherwise. The set of relations R contains R_M and another relation which distinguishes the elements of D_M from the others. Using a first-order axiomatization T of M, we give a first-order axiomatization $\mathcal{T}$ of the structure Ext_M and show that if T is *flexible* then $\mathcal{T}$ is *complete*. The flexible theories are particular theories whose function and relation symbols have some elegant properties which enable us to handle formulae more easily.

1 Introduction

Recall that a tree built on a set E is essentially a hierarchized set of nodes labelled by the elements of E. To each element e of E corresponds an operation f, called *construction operation*, which, starting from a sequence $a_1, \ldots, a_n$ of trees, builds the tree whose top node is labelled e and whose sequence of immediate children is $a_1, \ldots, a_n$.

The algebra of finite or infinite trees plays a fundamental act in computer science: it is a model for composed data known as *record* in Pascal or *structure* in C. The construction operation corresponds to the creation of a new record, i.e. of a cell containing an elementary information possibly followed by n cells, each one pointing to a record. Circuit of pointers correspond to infinite trees.

As early as 1976, G. Huet proposed an algorithm for unifying infinite terms, that is solving equations in that algebra [11]. B. Courcelle has studied the properties of infinite trees in the scope of recursive program schemes [6]. A. Colmerauer has described the execution of Prolog II, III and IV programs in terms of solving equations and disequations in that algebra [4, 3, 1]. The unification of finite terms, i.e. solving conjunctions of equations in the theory of finite trees has first been studied by A. Robinson [18]. Some better algorithms with better complexities has been proposed after by M.S. Paterson and M.N.Wegman [16] and A. Martelli and U. Montanari [15]. Solving conjunctions of equations in the theory of infinite trees has been studied by G. Huet [11], by A. Colmerauer [4] and by J. Jaffar [12]. Solving conjunctions of equations and disequations in the theory of possibly infinite trees has been studied by A. Colmerauer [4] and H.J. Bürckert [2]. An incremental algorithm for solving conjunctions of equations and disequations on rational trees has been proposed after by V.Ramachandran and P. Van Hentenryck [17]. On the other hand, there exists a quantification elimination algorithm which transforms a first-order formula into a boolean combination of simple ones. In the case of infinite trees with a finite set of function symbols we can refer to the work of M.J. Maher [14] and H. Comon [5]. M.J. Maher has summarized all these cases and proposed a complete axiomatizations for different sets of trees equipped with construction operations [14].

In this paper, we give and justify an axiomatization of the structure of finite or infinite M-extended trees. This structure is an intuitive combination of the structure of trees with another structure M and can be seen semantically as an extension to trees of the model M. Having a structure $M = (D_M, F_M, R_M)$ together with its domain D_M, its set of operations F_M and its set of relations R_M, we define the M-extended trees structure $Ext_M = (D, F, R)$ whose domain D consists of trees labelled by elements of $D_M \cup F$, where F is an infinite set of function symbols containing F_M and another infinite set of function symbols disjoint from F_M. For each n-ary function $f \in F$, the operation $f(a_1, ..., a_n)$ is evaluated in M and produces an element of D_M if $f \in F_M$ and all the a_i are elements of D_M, or is a tree whose root is labelled by f and whose immediate children are $a_1, ..., a_n$ otherwise. The set of relations R is built essentialy from R_M. In the case where M is the set of rational numbers together with the operations of addition and substraction and a linear dense order relation we can refer to Prolog III and IV whose execution has been modelized by A. Colmerauer [4, 1] using this M-extended trees.

The paper is organized in four sections followed by a conclusion. This introduction is the first section. In the second section we recall the Maher's structure of finite or infinite trees and introduce the M-extended trees structure for any model M. In the third section, we present our general sufficient conditions for the completeness of any first-order theory. Then, having a first-order axiomatization T of M, we give a first-order axiomatization $\mathcal{T}$ of finite or infinite M-extended trees. Finally we present in the fourth section a new class of theories that we call *flexible* and show that if T is flexible then $\mathcal{T}$ is complete. To show the completeness of $\mathcal{T}$ for any flexible theory T we use the general sufficient conditions presented in the third section. The definition of the M-extended trees, the axiomatization of $\mathcal{T}$, the definition of flexible theories and the proof of the completeness of $\mathcal{T}$ for every flexible theory T are our main contribution in this paper.

2 Extension to trees of a model M

2.1 Finite or infinite trees

Let F be an infinite set of *function symbols* and R be a set of *relation symbols*. To each element of $F \cup R$ is associated an integer, its *arity*. The arities are non-negative for elements of F and are positive for elements of R. An n-ary symbol is a symbol with arity n. A *constant* is a 0-ary symbol.

Let N be a set of words of positive integers, including the empty word ϵ. Let "." denote concatenation of word. A *tree built on* F is a mapping $a : E \to F$, for some non-empty subset E of N such that each element $i_1 \dots i_k$ (with $k \geq 0$) satisfies two conditions: (1) if $k > 0$ then $i_1 \dots i_{k-1} \in E$ and (2) if $a(i_1 \dots i_k) = f$ and f has arity n, then $i_1 \dots i_k i_{k+1} \in E$ if and only if $1 \leq i_{k+1} \leq n$.

The subtree of the tree a at $n \in E$ is the mapping $a' : E' \to F$ where $D' = \{d | n.d \in E\}$ and $a'(d) = a(n.d)$.

The set of all trees built on F is denoted A. To each n-ary function symbol f we associate a function from A^n to A also denoted f such that $f(a_1, \dots, a_n) = a$ where $a(\epsilon) = f$ and $a(i.d) = a_i(d)$ for $1 \leq i \leq n$ and d a node. These functions are called construction operations. The set of trees A with these construction operations forms the trees structure or trees algebra.

2.2 Finite or infinite M-extended trees structure

We are given now once for all a structure $M = (D_M, F_M, R_M)$ with its domain D_M, its set of functions F_M and its set of relations R_M. Let F be an infinite set of function symbols containing the set F_M and another infinite set of function symbols disjoint from F_M. Let R be the set of relation symbols $R_M \cup \{p\}$, with p a unary relation symbols which does not belong to R_M. The extension to trees of the model M, quite simply called M-extended trees model is the model $Ext_M = (D, F, R)$ defined as follows:

- the domain D is the set of the trees built on $F \cup D_M$ where each element $f \in F$ of arity n is considered as a label of arity n and each element of D_M is considered as a label of arity 0,
- to each n-ary element f of F is associated a function $f : D^n \to D$ such that $f(a_1, .., a_n)$ is the result of f on $(a_1, .., a_n)$ in D_M, if $f \in F_M$ and $a_i \in D_M$ for all i, and is the result of the construction operation f on $(a_1, .., a_n)$ otherwise,
- to each n-ary relation symbols r of $R - \{p\}$ is associated the set $r^{Ext_M} = r^M$. To the unary relation symbols p is associated the set $p^{Ext_M} = D_M$.

3 Theory of finite or infinite M-extended trees

Let V an infinite countable set of *variables.* A *term* is an expression of the form x or $ft_1 \dots t_n$ where $n \geq 0$, f an n-ary symbol in F and the t_i's are shorter terms. A *M-term* is either a variable or a term whose function symbols are elements of F_M. A *formula* is an expression of the forms

$$s = t,\ rt_1..t_n,\ \mathit{true},\ \mathit{false},\ \neg(\varphi),\ (\varphi \wedge \psi),\ (\varphi \vee \psi),\ (\varphi \rightarrow \psi),\ (\varphi \leftrightarrow \psi),\ \exists x\varphi,\ \forall x\varphi,$$

where $x \in V$, s, t and the t_i's are terms, r is an n-ary relation symbol in R and φ and ψ are shorter formulae. Formulae of the first form are called *equations* and of the second form *relations.* A *M-equation* is an equation of M-terms and a *M-relation* is a relation $rt_1...t_n$ with $r \in R_M$ and the t_i's M-terms.

An occurrence of a variable x in a formula is *bound* if it occurs in a sub-formula of the form $(\exists x\varphi)$ or $(\forall x\varphi)$. It is *free* otherwise. The *free variables* of a formula are those which have at least a free occurrence in the formula. For each formula φ, we denote by $\mathit{var}(\varphi)$ the set of all free variables of φ.

We call *instantiation* of a formula φ by individuals of D_M the obtained formula from φ in which for each free variable x in φ, we replace each free occurrence of x by the same individual i of D_M.

3.1 Theory and complete theory

Let $\bar{x} = x_1 \dots x_n$ and $\bar{y} = y_1 \dots y_n$ be two *vectors of variables* of the same length. Let ψ, ϕ, φ and $\varphi(\bar{x})$ be formulae. We write

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

Note that the formulae $\exists?\varepsilon\varphi$ and $\exists!\varepsilon\varphi$ are respectively equivalent to *true* and to φ in any model M. Theses quantifiers are just convenient notations and can be expressed in the first-order level.

Definition 3.1.1 *Let $\Psi(u)$ be a set of formulas having at most u as a free variable. We write $M \models \exists_{o\ \infty}^{\Psi(u)} x\,\varphi(x)$, iff for any instantiation $\exists x\,\varphi'(x)$ of $\exists x\,\varphi(x)$ by individuals of D_M one of the following properties holds:*

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

A *theory* is a set of propositions. We say that the model M is a *model of* T iff for each element φ of T, $M \models \varphi$. If φ is a formula, we write $T \models \varphi$ iff for each model M of T, $M \models \varphi$. A theory T is *complete* if for each proposition φ, either $T \models \varphi$ or $T \models \neg\varphi$. A *complete axiomatization* of a structure M is a recursive set T of propositions such that for each proposition φ, $T \models \varphi$ iff $M \models \varphi$.

In what follows we use the abbreviation wnfv for *"without new free variables"*. By saying a formula φ is equivalent to a wnfv formula ψ in T we mean $T \models \varphi \leftrightarrow \psi$ and ψ does not contain other free variables than those of φ. The following theorem states general sufficient conditions for the completeness of a theory T.

Theorem 3.1.2 *[9, 10] A theory T is complete if there exists a set $\Psi(u)$ of formulas, having at most u as free variable, a set A of formulas, closed under conjunction and renaming, 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 flat atomic formula is equivalent in T to a wnfv Boolean combination of basic formulas of the form $\exists \bar{x}\alpha$ with $\alpha \in A$,*
2. *every formula without free variables of the form $\exists \bar{x}'\alpha' \wedge \alpha''$ with $\exists \bar{x}'\alpha' \in A'$ and $\alpha'' \in A''$ is equivalent either to false or to true in T,*
3. *every formula of the form $\exists \bar{x}\, \alpha \wedge \psi$, with $\alpha \in A$ and ψ any 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 $T \models \forall \bar{x}''\alpha'' \rightarrow \exists! \bar{x}'''\alpha'''$,
4. *if $\exists \bar{x}'\alpha' \in A'$ then $T \models \exists ? \bar{x}'\, \alpha'$ and for each free variable y in $\exists \bar{x}'\alpha'$, at least one of the following properties holds:*
   - *$T \models \exists ? y\bar{x}'\, \alpha'$,*
   - *there exists $\psi(u) \in \Psi(u)$ such that $T \models \forall y\, (\exists \bar{x}'\, \alpha') \rightarrow \psi(y)$,*
5. *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 each x'', the formula $\exists x''\alpha''$ is equivalent in T to a wnfv formula which belongs to A'',*
   - *for each x'', $T \models \exists_{o\,\infty}^{\Psi(u)} x''\, \alpha''$.*

3.2 Axiomatization of the structure of M-extended trees

M. Maher has introduced a complete axiomatization of the structure of finite or infinite trees built on an infinite set F [14]. The axiomatization is the set of propositions of the following forms:

$$\begin{array}{ll} 1 & \forall \bar{x} \forall \bar{y}\, f\bar{x} = f\bar{y} \rightarrow \bigwedge_i x_i = y_i, \\ 2 & \forall \bar{x} \forall \bar{y}\, \neg f\bar{x} = g\bar{y}, \\ 3 & \forall \bar{x} \exists! \bar{z}\, \bigwedge_i z_i = t_i(\bar{z}, \bar{x}), \end{array}$$

where $f, g \in F$, x, y, z are variables, $\bar{x}$ is vector of variables x_i, $\bar{y}$ is vector of variables y_i, $\bar{z}$ is vector of distinct variables z_i and where $t_i(\bar{x}, \bar{z})$ is a term which begins by an element of F followed by variables taken from $\bar{x}$ or $\bar{z}$.

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

Let T be an axiomatization of the structure $M = (D_M, F_M, R_M)$. Using this axiomatization, let us now define an axiomatization $\mathcal{T}$ of the structure of finite or infinite M-extended trees together with the sets F and R (defined in section 2.2) as function and relation symbols.

Definition 3.2.1 An axiomatization $\mathcal{T}$ of the structure of finite or infinite M-extended trees is the set of propositions of the following forms where $\bar{x}$, $\bar{y}$ are vectors of variables x_i, y_i.

1. explosion: for each $f \in F$
$$\forall \bar{x} \forall \bar{y}\ \neg pf\bar{x} \wedge \neg pf\bar{y} \wedge f\bar{x} = f\bar{y} \rightarrow \bigwedge_i x_i = y_i$$
2. conflict of symbols: for f and g distinct symbols in F
$$\forall \bar{x} \forall \bar{y}\ f\bar{x} = g\bar{y} \rightarrow pf\bar{x} \wedge pg\bar{y}$$

3. unique solution

$$\forall \bar{x} \forall \bar{y} \; (\bigwedge_i px_i) \wedge (\bigwedge_j \neg py_j) \rightarrow \exists! \bar{z} \bigwedge_k (pz_i \wedge z_k = t_k(\bar{x}, \bar{y}, \bar{z}))$$

where $\bar{z}$ is a vector of distinct variables z_i, $t_k(\bar{x}, \bar{y}, \bar{z})$ is a term expressed by a function symbol f_k followed by variables taken from $\bar{x}, \bar{y}, \bar{z}$, moreover, if $f_k \in F_M$, the term $t_k(\bar{x}, \bar{y}, \bar{z})$ contains at least one variable from $\bar{y}$ or $\bar{z}$

4. relations of R_M: for each $r \in R_M$,

$$\forall \bar{x} \; r\bar{x} \rightarrow \bigwedge_i px_i$$

5. operations of F_M: for each $f \in F_M$,

$$\forall \bar{x} \; pf\bar{x} \leftrightarrow \bigwedge_i px_i$$

(this axiom, in the case of f a constant in F_M, becomes pf)

6. elements not in M: for each $f \in F - F_M$,

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

7. existence of at least one element satisfying p (only if F_M does not contains 0-arity function symbols):

$$\exists x \, px,$$

8. the extension of axioms of T: all axioms obtained by the following transformation of an axiom φ of T : While it is possible 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 all 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 $\forall \bar{x} \, (\bigwedge px_i) \rightarrow \psi$.

Example 3.2.2 Let M be the structure of the rational numbers together with the operations of addition, substraction and a linear dense order relation without endpoints. In this case D_M is the set of the rational numbers, $F_M = \{+, -, 0, 1\}$ and $R_M = \{<\}$. Let a be a positive integer and let $t_1, ..., t_n$ be terms. Let us denote by:

– $t_1 < t_2$, the term $< t_1 t_2$,
– $t_1 + t_2$, the term $+t_1 t_2$,
– $t_1 + t_2 + t_3$, the term $+t_1(+t_2 t_3)$,
– $-at_1$, the term $\underbrace{(-t_1) + \cdots + (-t_1)}_{a}$,
– $0t_1$, the term 0,
– at_1, the term $\underbrace{t_1 + \cdots + t_1}_{a}$,
– a the term $\underbrace{1 + \cdots + 1}_{a}$.

The axiomatization T of the structure M is of the 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 $\forall x \, nx = 0 \rightarrow x = 0$,
6_n $\forall x \, \exists! y \, ny = x$, $\quad (n \neq 0)$
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 $0 < 1$.

Using the transformations of Definition 3.2.1, the axiomatization $\mathcal{T}$ of the M-extended trees theory is of the form:

1 $\forall\bar{x}\forall\bar{y}\,((\neg p\, f\bar{x}) \wedge (\neg p\, f\bar{y}) \wedge f\bar{x} = f\bar{y}) \rightarrow \bigwedge_i x_i = y_i,$
2 $\forall\bar{x}\forall\bar{y}\, f\bar{x} = g\bar{y} \rightarrow p\, f\bar{x} \wedge p\, g\bar{y},$
3 $\forall\bar{x}\forall\bar{y}\,(\bigwedge_{i\in I} p\, x_i) \wedge (\bigwedge_{j\in J} \neg p\, y_j) \rightarrow (\exists!\bar{z}\, \bigwedge_{k\in K} (\neg p\, z_k \wedge z_k = t_k(\bar{x},\bar{y},\bar{z}))),$
4 $p0,$
5 $p1,$
6 $\forall x\forall y\, x < y \rightarrow (p\, x \wedge p\, y),$
7 $\forall x\forall y\, p\, x + y \leftrightarrow p\, x \wedge p\, y,$
8 $\forall x\, p\, - x \leftrightarrow p\, x,$
9 $\forall\bar{x}\, \neg p\, h\bar{x},$
10 $\forall x\forall y\,(p\, x \wedge p\, y) \rightarrow x + y = y + x,$
11 $\forall x\forall y\forall z\,(p\, x \wedge p\, y \wedge p\, z) \rightarrow x + (y + z) = (x + y) + z,$
12 $\forall x\, p\, x \rightarrow x + 0 = x,$
13 $\forall x\, p\, x \rightarrow x + (-x) = 0,$
14_n $\forall x\, p\, x \rightarrow (nx = 0 \rightarrow x = 0),$
15_n $\forall x\, p\, x \rightarrow \exists! y\, p\, y \wedge ny = x, \quad (n \neq 0)$
16 $\forall x\, p\, x \rightarrow \neg x < x,$
17 $\forall x\forall y\forall z\, p\, x \wedge p\, y \wedge p\, z \rightarrow ((x < y \wedge y < z) \rightarrow x < z),$
18 $\forall x\forall y\,(p\, x \wedge p\, y) \rightarrow (x < y \vee x = y \vee y < x),$
19 $\forall x\forall y\,(p\, x \wedge p\, y) \rightarrow (x < y \rightarrow (\exists z\, p\, z \wedge x < z \wedge z < y)),$
20 $\forall x\, p\, x \rightarrow (\exists y\, p\, y \wedge x < y),$
21 $\forall x\, p\, x \rightarrow (\exists y\, p\, y \wedge y < x),$
22 $\forall x\,\forall y\,\forall z\,(p\, x \wedge p\, y \wedge\, p\, z) \rightarrow (x < y \rightarrow (x + z < y + z)),$
23 $0 < 1,$

where f and g are two distinct function symbols taken from F, $h \in F - F_M$, x, y, z are variables, $\bar{x}$ is a vector of variables x_i, $\bar{y}$ is a vector of variables y_i, $\bar{z}$ is vector of distinct variables z_i and where $t_k(\bar{x},\bar{y},\bar{z})$ is a term which begins by a function symbol f_k element of F followed by variables taken from $\bar{x}$ or $\bar{y}$ or $\bar{z}$, moreover, if $f_k \in F_M$ then $t_k(\bar{x},\bar{y},\bar{z})$ contains at least a variable taken from $\bar{y}$ or $\bar{z}$. This theory has been used by A. Colmerauer to modelize the execution of Prolog III and IV [4, 1].

4 Completeness of $\mathcal{T}$

We suppose that the variables of V are ordered by a strict linear dense order relation denoted by $\succ$. We call *leader* of an M-equation α the greatest variable x of all variables in α, according to the order $\succ$, such that $M \models \exists! x\alpha$.

4.1 Flexible structure

The model M is called *flexible* if for each conjunction α of M-equations and each conjunction β of M-relations:

1. α is equivalent in M either to *false* or to a wnfv conjunction α' of M-equations whose each element has a distinct leader which has one and only occurrence in α', and for all variable $x \in \mathit{var}(\alpha')$ we have $M \models \exists! x\alpha'$,
2. if β does not contain variables then $M \models \beta$ or $M \models \neg\beta$,
3. the formula $\neg\beta$ is equivalent in M to a wnfv disjunction of M-equations and M-relations,
4. for all $x \in V$
   - the formula $\exists x\, \beta$ is equivalent in M either to *false* or to a quantifier free conjunction of M-relations,

- for all $x \in V$ and for all instantiation $\exists x\, \beta'(x)$ of $\exists x\, \beta(x)$ by individuals of D_M, either $M \models \neg\exists x \beta'(x)$ or there exists an infinite set of individuals i of D_M such that $M \models \beta'(i)$.

A theory T is called flexible iff all its models are flexible.

Property 4.1.1 *If T is flexible then it is complete.*

4.2 Blocks and solved blocks in $\mathcal{T}$

Definition 4.2.1 *A* block *is a conjunction α of formulae of the following forms:*

- *true, false, px, $\neg px$,*
- *$x = y$, $x = fx_1 \dots x_n$, with $f \in F$,*
- *$t_1 = t_2 \wedge \bigwedge_{i=1}^n px_i$, where $\{x_1, \dots, x_n\}$ is the set of variables which occur in the M-equation $t_1 = t_2$,*
- *$rt_1 \dots t_n$, where $r \in R_M$ and the t_i's are M-terms,*

and such that α contains px or $\neg px$ for each variable $x \in \mathrm{var}(\alpha)$. A relation block *is a block without equations. An* equation block *is a block without M-relations and where each variable has an occurrence in at least one equation.*

Definition 4.2.2 *If a block α has a sub-formula of the form*

$$x_0 = t_0(x_1) \wedge x_1 = t_1(x_2) \wedge \cdots \wedge x_{n-1} = t_{n-1}(x_n) \wedge \bigwedge_{i=0}^{n-1} \neg px_i,$$

where x_{i+1} has an occurrence in the term $t_i(x_{i+1})$, then the variable x_n and the equation $x_{n-1} = t_{n-1}(x_n)$ are called reachable *from x_0 in α.*

Property 4.2.3 *Let α be a block. If all the variables of $\bar{x}$ are reachable in α from free variables of $\exists \bar{x}\alpha$, then $\mathcal{T} \models \exists ? \bar{x}\alpha$.*

Definition 4.2.4 A block α is called *well-typed* iff α does not contain sub-formulae of one of the following forms:

- $p\,x \wedge \neg p\,x$,
- $x = h\bar{y} \wedge p\,x$, with $h \in F - F_M$,
- $x = f_0 \wedge \neg p\,x$, with f_0 a constant of F_M,
- $x_0 = fx_1...x_n \wedge \neg p\,x_0 \wedge \bigwedge_{i=1}^n p\,x_i$, with $f \in F_M$,
- $x_0 = fx_1...x_n \wedge p\,x_0 \wedge \neg p\,x_i$, with $f \in F$
- $x_0 = x_1 \wedge p\,x_0 \wedge \neg p\,x_1$,
- $x_0 = x_1 \wedge \neg p\,x_0 \wedge p\,x_1$,
- $rt_1...t_n \wedge \neg p\,x_i$ with $r \in R_M$ and x_i a variable which occurs in the M-relation $rt_1...t_n$.

Definition 4.2.5 *Let t_1 be a term. Let t_2 and t_3 be two M-terms. Let α be a well-typed equation block. Either $x = t_1 \wedge \neg px$ is a sub-formula of α. In this case, x is called α-leader of the equation $x = t_1$. Else $t_2 = t_3 \wedge \bigwedge_{i \in \mathrm{var}(t_2 = t_3)} pi$ is a sub-formula of α. In this case, the greatest variable in $var(t_2 = t_3)$ according to the order $\succ$ such that $\mathcal{T} \models \exists! x\, t_2 = t_3 \wedge \bigwedge_{i \in var(t_2 = t_3)} pi$ is called α-leader of the equation $t_2 = t_3$.*

Definition 4.2.6 *A block α is called* solved block, *iff:*

1. *α is well-typed and does not contain formulae of the form $t_1 = t_2$ or $rt_1...t_n$ with $r \in R_M$ and the t_i's terms which does not contain variables,*
2. *for each equation $x = y$ in α, $x \succ y$,*
3. *each equation in α has a distinct α-leader which does not occur in M-relations of α,*
4. *if px and py are sub-formulas of α with x and y two α-leaders of two equations α_1, α_2 of α then $x \notin \mathrm{var}(\alpha_2)$,*
5. *for all variable x which occurs in an equation of α we have $\mathcal{T} \models \exists ? x\alpha$.*

Property 4.2.7 *Let α be a solved equation block different from the formula* false *and let $\bar{x}$ be the set of the α-leaders of the equations of α. We have $\mathcal{T} \models \exists! \bar{x}\alpha$.*

Property 4.2.8 *If T is flexible then each block is equivalent in $\mathcal{T}$ to a solved block.*

4.3 Completeness of $\mathcal{T}$

Theorem 4.3.1 *If T is a flexible theory then $\mathcal{T}$ is complete.*

We show this theorem using Theorem 3.1.2. The sets $\Psi(u)$, A, A' and A'' are chosen as follows:

- $\Psi(u)$ is the set of the formulae of the form $\exists \bar{y}\, u = f\bar{y} \wedge \neg p\, u$, with f a non 0-ary function symbol taken from F.
- A is the set of blocks.
- A' is the set of the formulae of the form $\exists \bar{x}'\alpha'$, where:
  - all the variables of $\bar{x}'$ are reachable in α' from free variables of $\exists \bar{x}'\alpha'$,
  - α' is a solved equation block, different from the formula *false*, and where the order $\succ$ is such that all the variables of $\bar{x}'$ are greater than the free variables of $\exists \bar{x}'\alpha'$,
  - all the equations of α' of the form $x_0 = f x_1 ... x_n$ with $f \in F - F_M$ are reachable in α' from free variables of $\exists \bar{x}'\alpha'$,
  - if the M-equation $t_1 = t_2$ is a sub-formula of α' then, each variable x_i which occurs in it is either a free variable of $\exists \bar{x}'\alpha'$ or reachable in α' from free variables of $\exists \bar{x}'\alpha'$,
- A'' is the set of solved relation blocks.

5 Conclusion

We have defined in this paper the structure of the M-extended trees for any model M. This structure can be considered as a combination of the structure of finite or infinite trees with the structure M. Having an axiomatization T of M we have given a first-order axiomatization $\mathcal{T}$ of the M-extended trees structure and have shown that if T is flexible then $\mathcal{T}$ is complete. To prove the completeness in this case, we have used our general sufficient condition. From this condition we can extract a general algorithm for solving first-order constraints in $\mathcal{T}$. Due to lack of space we cannot present this algorithm in this paper. Just note that this algorithm uses the block defined in our paper and transforms any formula φ in a particular formula ψ called *solved formula* equivalent to φ in $\mathcal{T}$. In particular if φ has no free variables then ψ is either the formula *true* or the formula *false*. The correctness of our algorithm is another proof of the completeness of $\mathcal{T}$ of each flexible theory T.

There exists a lot real and practical problems which can be represented by full first-order formulae on M-extended trees. We can site for example the works of A. Colmerauer [4, 1] who has realized the execution of Prolog III and IV using the M-extended trees where M is the structure of the rational numbers together with the operations of addition and substraction and linear dense order relation.

On the other hand S. Vorobyov [19] have shown that the problem of deciding if a proposition without free variables is true or not in the trees theory is non-elementary, i.e. the complexity of all algorithm which solve it is not bounded by a tower of powers of $2's$ (with a top down evaluation) with a fixed height. A. Colmerauer and B. Dao [8, 7] have also given a proof of non-elementary complexity of solving constraints in the trees theory. Thus, it is normal that our sufficient condition is complex and the properties of our blocks uses some nonclassical quantifiers. Nevertheless we hope find some interesting class of complexities in the implementation of our algorithm as it has been done in [8] in the theory of finite or infinite trees.

Actually we try to show the completeness of $\mathcal{T}$ where M is the structure of the real numbers together with addition, substraction, multiplication and a linear dense order relation. We also study the complexity and the expressiveness of the first-order constraints in $\mathcal{T}$ as it has done in [7, 8].

Acknowledgements We thank Alain Colmerauer for our many discussions and his help in the organization and the drafting of this paper. We thank him too, for the definitions and proof of his course of DEA. We dedicate to him this paper with our best wishes for a speedy recovery.

References

1. Benhamou F, Colmerauer , Van caneghem M. Le manuel de Prolog IV , PrologIA, Marseille, France, 1996.
2. Bürckert H. Solving disequations in equational theories. In Proc. 9th Conf. on Automated Deduction, LNCS 310, pages 517-526. Springer-Verlag, 1988.
3. Colmerauer A. An introduction to Prolog III. *Communication of the ACM*, 33(7):68–90,1990.
4. Colmerauer A. Equations and inequations on finite and infinite trees. Proc. of the Int. Conf. on the Fifth Generation of Computer Systems, Tokyo, 1984. P. 85–99.
5. Comon H. Unification et disunification : Théorie et applications. PhD thesis, Institut National Polytechnique de Grenoble, 1988.
6. Courcelle B. Equivalences and Transformations of Regular Systems applications to Program Schemes and Grammars, TCS, vol. 42, 1986, p. 1–122.
7. Colmerauer A., Dao. TBH., Expressiveness of full first-order constraints in the algebra of finite or infinite trees, Constraints, Vol. 8, No. 3, 2003, pages 283-302.
8. Dao TBH. Résolution de contraintes du premier ordre dans la théorie des arbres finis ou infinis. Thèse d'informatique, Université de la Méditerranée, décembre 2000.
9. Djelloul K. Complete first-order axiomatisation of the construction of trees on an ordered set. Proceedings of the 2005 Int. Conf. on Foundations of Computer Science (FCS'05) Las Vegas, CSREA Press.
10. 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, P. 106–122.
11. Huet G. Resolution d'equations dans les langages d'ordre 1, 2,$\ldots\omega$. These d'Etat, Universite Paris 7. France,1976.
12. Jaffar J. Efficient unification over infinite terms. New Generation Computing, 2(3): P. 207–219, 1984.
13. Kunen K. Negation in logic programming. Journal of Logic Programming, 4: P. 289–308. 1987.
14. Maher M. Complete axiomatization of the algebra of finite, rational and infinite trees. *Technical report, IBM*, 1988.
15. Matelli A. and Montanari U. An efficient unification algorithm. ACM Trans. on Languages and Systems, 4(2):258-282, 1982.
16. Paterson M and Wegman N. Linear unification. Journal of Computer and Systems Science, 16:158-167, 1978.
17. Ramachandran V. and Van Hentenryck P. Incremental algorithms for constraint solving and entailment over rational trees. In Proc. 13th Conf. Foundations of Software Technology and Theoretical Computer Science, LNCS 761, pp 205-217, 1993.
18. Robinson J.A. A machine-oriented logic based on the resolution principle. JACM, 12(1):23-41, 1965.
19. Vorobyov S. An Improved Lower Bound for Elementary Theories of Trees, Proceeding of the 13th Conference on Automated Deduction. LNAI, vol 1104, pp. 275-287.