

# Naive solving of non-linear constraints\*

Alain Colmerauer<sup>†</sup>

March 1992

## Abstract

In this paper we study a naive and incomplete algorithm for solving systems of non-linear constraints. These constraints are expressed with variables ranging over reals, rational constants, the operations  $-$ ,  $+$ ,  $\times$  and the relations  $\geq$ ,  $>$ ,  $=$ ,  $\neq$ . By *solving* a system  $S$  we understand: first, deciding whether  $S$  has at least one solution; second, computing the set of equations of the form  $x = \text{constant}$  which are entailed by  $S$ .

The preliminary phase of the naive algorithm consists of introducing intermediate variables for splitting  $S$  into two subsystems, a linear one and a non-linear one containing only constraints of the form  $z = x \times y$ , where  $x$ ,  $y$  and  $z$  are variables. The naive algorithm itself will repeat two actions until it reaches a stable system or a linear part that has no solution. The first action is to solve the linear part of  $S$ . The second action is to consider the equations of the form  $x = \text{constant}$  that are entailed by the linear part of  $S$  and to replace each variable  $x$  by the corresponding constant in the right-hand sides of the non-linear equations .

We show that the naive algorithm turns out to be complete in the following non-standard structure for reals: multiplication is modified by regarding the product of two irrational numbers as an element  $\omega$  which is outside of the domain of the reals. The operations are extended by taking  $\omega$  as the value as soon as one of the arguments is  $\omega$ . An exception to this principle is made for multiplication by zero, which always produces zero. All the relations, the  $=$  relation included, are considered to be satisfied as soon as one of their arguments is  $\omega$ . Rational numbers are kept as constants and variables are not allowed to take the value  $\omega$ .

## Introduction

It is possible to solve non-linear constraints over real numbers. It was Alfred Tarski who obtained this result in 1930 in the general context of the theory of existential and universal quantification. The result was published much later in [12]. In 1973 George Collins [1] proposed an implementable algorithm for the same problem. This algorithm is known under the name “cylindrical algebraic decomposition” and it has given rise to numerous further developments. Unfortunately it has been shown that the complexity of the cylindrical algebraic decomposition algorithm is doubly exponential in the number of quantifier alternations [6]. For the purely existential case which is of interest here the complexity is singly exponential in the number of variables. It turns out that in practice the largest number of variables that one can treat lies somewhere between 5 and 10. Let us finally mention that for the purely existential case other algorithms have been developed. The interested reader will find a comparison of their different complexities in [7].

---

\*The present research was carried out within the context of the *Compulog* Basic Research Action 3012. Support was also given by the *Greco de programmation* of the CNRS.

<sup>†</sup>Groupe Intelligence Artificielle, Unité de Recherche Associée au CNRS 816, Faculté des Sciences de Luminy, Case 901, 163 Avenue de Luminy, 13288 Marseille cedex 9, France. E-mail: colmer@gia.univ-mrs.fr

The originators of the three principal constraint logic programming languages Prolog III [3], CLP(R) [9] and CHIP [5] were thus well guided in restricting themselves to essentially linear constraints. As a result they have taken advantage of two efficient algorithms [8]: the Gaussian algorithm, which is used for the elimination of variables, and the simplex algorithm of George Dantzig [4], which serves to optimize a linear function on variables constrained by linear  $\geq$  inequalities.

However, many problems formulated in a non-linear fashion can be solved by a naive algorithm which consists in combining the solving of the linear part of the problem with the integration of constraints that have become linear because the values of certain coefficients were determined. In the present paper we intend to clarify and justify this naive algorithm. The paper consists of eight parts followed by a conclusion. Part 1 is an informal presentation of the algorithm with respect to an example. Part 2 introduces the terminology for speaking about constraints independently of the domain, more precisely, of the structure in which one is working. Parts 3 and 4 are devoted to the standard structure of the reals and to a non-standard<sup>1</sup> structure of the reals in which the domain, the operations and the relation are slightly modified. Part 5 contains the main result of the paper: the naive algorithm solves every system of constraints, including non-linear ones, not in the standard but rather in the non-standard structure. Parts 6, 7 and 8 are devoted to the proofs of the theorems used in preceding parts.

## 1 An example of naive solving

In the domain of real numbers let us consider the system of non-linear equations  $S_n$  of the form

$$S_n = T_0 \cup T_1 \cup \dots \cup T_n,$$

where  $T_0$  is the system

$$\left\{ \begin{array}{l} x_0 = 1, \\ y_0 = 2 \end{array} \right\}$$

and  $T_{i+1}$  the system

$$\left\{ \begin{array}{l} x_i \times x_{i+1} + y_i \times y_{i+1} = 0, \\ x_{i+1} + y_{i+1} = (x_i - y_i)(2x_i + y_i + 1) \end{array} \right\}.$$

Let us try to solve the system  $S_2$  which is

$$\left\{ \begin{array}{l} x_0 = 1, \\ y_0 = 2, \\ x_0 x_1 + y_1 y_0 = 0, \\ x_1 + y_1 = (x_0 - y_0)(2x_0 + y_0 + 1), \\ x_1 x_2 + y_1 y_2 = 0, \\ x_2 + y_2 = (x_1 - y_1)(2x_1 + y_1 + 1) \end{array} \right\}.$$

---

<sup>1</sup>The term "non-standard" used in this paper is not to be confused with the one used in non-standard analysis.

We introduce the intermediate variables  $u_i, v_i, w_i, z_i, z'_i$  and express all the non-linear products in constraints of the form  $z = xy$ , where  $x, y$  and  $z$  are variables.

$$\left\{ \begin{array}{l} x_0 = 1, \\ y_0 = 2, \\ u_1 + v_1 = 0, \\ u_1 = x_0 x_1, \\ v_1 = y_0 y_1, \\ \\ x_1 + y_1 = w_1, \\ w_1 = z_0 z'_0, \\ z_0 = x_0 - y_0, \\ z'_0 = 2x_0 + y_0 + 1, \\ \\ u_2 + v_2 = 0, \\ u_2 = x_1 x_2, \\ v_2 = y_1 y_2, \\ \\ x_2 + y_2 = w_2, \\ w_2 = z_1 z'_1, \\ z_1 = x_1 - y_1, \\ z'_1 = 2x_1 + y_1 + 1 \end{array} \right\}$$

We can now partition the system into a purely linear and a purely non-linear part.

$$\left\{ \begin{array}{l} x_0 = 1, \\ y_0 = 2, \\ u_1 + v_1 = 0, \\ x_1 + y_1 = w_1, \\ z_0 = x_0 - y_0, \\ z'_0 = 2x_0 + y_0 + 1, \\ u_2 + v_2 = 0, \\ x_2 + y_2 = w_2, \\ z_1 = x_1 - y_1, \\ z'_1 = 2x_1 + y_1 + 1 \end{array} \right\} \cup \left\{ \begin{array}{l} u_1 = x_0 x_1, \\ v_1 = y_0 y_1, \\ w_1 = z_0 z'_0, \\ u_2 = x_1 x_2, \\ v_2 = y_1 y_2, \\ w_2 = z_1 z'_1 \end{array} \right\}$$

Let us solve the linear part or, more precisely, leaving aside the non-linear part, let us make explicit all the hidden equations of the form  $x = k$ , where  $x$  is a variable and  $k$  a constant. In this connection, observe that the 1st, 2nd, 5th and 6th equation form a subsystem of 4 equations with 4 unknowns.

$$\left\{ \begin{array}{l} x_0 = 1, \\ y_0 = 2, \\ z_0 = -1, \\ z'_0 = 5, \\ \\ u_1 + v_1 = 0, \\ x_1 + y_1 = w_1, \\ u_2 + v_2 = 0, \\ x_2 + y_2 = w_2, \\ z_1 = x_1 - y_1, \\ z'_1 = 2x_1 + y_1 + 1 \end{array} \right\} \cup \left\{ \begin{array}{l} u_1 = x_0 x_1, \\ v_1 = y_0 y_1, \\ w_1 = z_0 z'_0, \\ u_2 = x_1 x_2, \\ v_2 = y_1 y_2, \\ w_2 = z_1 z'_1 \end{array} \right\}$$

In the non-linear part let us now substitute  $k$  for every variable  $x$ , for which we have the equation  $x = k$  in the linear part. We then bring back the new linear equations in the linear part.

$$\left( \begin{array}{l} x_0 = 1, \\ y_0 = 2, \\ z_0 = -1, \\ z'_0 = 5, \\ u_1 + v_1 = 0, \\ x_1 + y_1 = w_1, \\ u_2 + v_2 = 0, \\ x_2 + y_2 = w_2, \\ z_1 = x_1 - y_1, \\ z'_1 = 2x_1 + y_1 + 1, \\ \\ u_1 = x_1, \\ v_1 = 2y_1, \\ w_1 = -5 \end{array} \right) \cup \left\{ \begin{array}{l} u_2 = x_1 x_2, \\ v_2 = y_1 y_2, \\ w_2 = z_1 z'_1 \end{array} \right\}$$

We now solve again the non-linear part where the 5th, 6th and the 5 last equations form a system of 7 equations with 7 unknowns.

$$\left( \begin{array}{l} x_0 = 1, \\ y_0 = 2, \\ z_0 = -1, \\ z'_0 = 5, \\ x_1 = -10, \\ y_1 = 5, \\ u_1 = -10, \\ v_1 = 10, \\ w_1 = -5, \\ z_1 = -15, \\ z'_1 = -14, \\ \\ u_2 + v_2 = 0, \\ x_2 + y_2 = z_2 \end{array} \right) \cup \left\{ \begin{array}{l} u_2 = x_1 x_2, \\ v_2 = y_1 y_2, \\ w_2 = z_1 z'_1 \end{array} \right\}$$

In the non-linear part let us now again substitute  $k$  for every variable  $x$ , for which we have the equation  $x = k$  in the linear part. We then bring back the new linear equations

in the linear part.

$$\left\{ \begin{array}{l} x_0 = 1, \\ y_0 = 2, \\ z_0 = -1, \\ z'_0 = 5, \\ x_1 = -10, \\ y_1 = 5, \\ u_1 = -10, \\ v_1 = 10, \\ w_1 = -5, \\ z_1 = -15, \\ z'_1 = -14, \\ u_2 + v_2 = 0, \\ x_2 + y_2 = z_2, \\ \\ u_2 = -10x_2, \\ v_2 = 5y_2, \\ w_2 = 210 \end{array} \right\}$$

Let us now solve the new purely linear system. We finally obtain,

$$\left\{ \begin{array}{l} x_0 = 1, \\ y_0 = 2, \\ z_0 = -1, \\ z'_0 = 5, \\ x_1 = -10, \\ y_1 = 5, \\ u_1 = -10, \\ v_1 = 10, \\ w_1 = -5, \\ z_1 = -15, \\ z'_1 = -14, \\ x_2 = 70, \\ y_2 = 140, \\ u_2 = -700, \\ v_2 = 700, \\ w_2 = 210 \end{array} \right\}.$$

Since the naive algorithm is integrated to Prolog III [3, 11], it is possible to solve the system  $S_n$  by the program:

```
sequence(⟨⟨1, 2⟩⟩) →;
sequence(s · ⟨⟨x, y⟩⟩ · ⟨⟨x', y'⟩⟩) →
sequence(s · ⟨⟨x, y⟩⟩),
{x' × x + y' × y = 0,
x' + y' = (x - y) × (2x + y + 1)};
```

To calculate the sequence of pairs

$$s = \langle \langle x_0, y_0 \rangle, \dots, \langle x_6, y_6 \rangle \rangle,$$

we put the query<sup>2</sup>

```
sequence(s), { |s| = 7 }?
```

---

<sup>2</sup>In the commercial version of Prolog III, the constraint  $|s| = 7$  is written  $s :: 7$  and the question mark is replaced by a semicolon.

and we get the result

$$\{s = \langle\langle 1, 2 \rangle, \langle -10, 5 \rangle, \langle 70, 140 \rangle, \langle -39340, 19670 \rangle, \langle 1160707030, 2321414060 \rangle, \langle -10777926478252781260, 5388963239126390630 \rangle, \langle 87122774377966800110603263954929000070, 174245548755933600221206527909858000140 \rangle\rangle\}.$$

Let us return to our naive algorithm. If we had begun with the system  $\{x^2 - 2x + 1 = 0\}$  we would have arrived at something like  $\{z - 2x + 1 = 0, z = x \times x\}$  without having discovered that  $x$  equals 1 and, even worse, if we had started with  $\{x \times x = -4\}$  we would have arrived at  $\{z = -4, z = x \times x\}$  without discovering that the system had no solutions. We can now raise the question: what exactly does the naive algorithm solve in these cases? The answer to this question constitutes the content of this paper.

## 2 Terminology

We call *structure* a 4-tuple

$$(D, D', F, R),$$

where  $D$  is a domain,  $D'$  a subdomain of  $D$ ,  $F$  a set of operations on  $D$  and  $R$  a set of relations over  $D$ . With each operation and each relation is associated its arity, a non-negative integer  $n$ . An  $n$ -place operation  $f$  is a mapping of type  $D^n \rightarrow D$ . As usual, 0-place operations are called constants and are identified with elements of the domain. An  $n$ -place relation  $r$  is a subset of  $D^n$  and instead of writing  $(a_1, \dots, a_n) \in r$ , we write  $r(a_1, \dots, a_n)$ .

To refer to the elements of subdomain  $D'$ , we assume as given an infinite *universal* set  $V$  of variables and we introduce two types of expressions: *terms* to refer to the elements of the domain and *constraints*, to formulate properties of these elements. More precisely, a term is a word constructed from the alphabet  $V \cup F$  and defined recursively as follows: a term of depth 0 is a word of one of the two forms

$$x \text{ or } a,$$

where  $x \in V$  and  $a \in D$ , and a term of depth  $k + 1$  is a word of the form

$$ft_1 \cdots t_n,$$

where  $f$  is an  $n$ -place operation, where at least one  $t_i$  is a term of depth  $k$  and where the others are terms at most of depth  $k$ . The constraints are words constructed from the alphabet  $V \cup F \cup R$  and have the form

$$rt_1 \cdots t_n, \tag{1}$$

where  $r$  is an  $n$ -place relation and where each  $t_i$  is a term.

Given a subset  $W$  of the set  $V$  of variables, an *assignment to  $W$*  is a mapping

$$\sigma : W \rightarrow D'.$$

When  $W$  is not explicitly mentioned,  $W$  is assumed to be the universal set  $V$  of variables. Such an assignment  $\sigma$  of  $V$  extends naturally to a mapping  $\sigma^*$  from the set of terms into the domain by taking

$$\begin{aligned} \sigma^*(x) &= \sigma(x), \\ \sigma^*(a) &= a, \\ \sigma^*(ft_1 \cdots t_n) &= f(\sigma^*(t_1), \dots, \sigma^*(t_n)). \end{aligned}$$

An assignment  $\sigma$  is a *solution* of the constraint (1) if

$$r(\sigma^*(t_1), \dots, \sigma^*(t_n)).$$

A *system* of constraints is a finite set of constraints. A *solution* of a system  $S$  of constraints is an assignment which is a solution of all the constraints of  $S$ . If  $W$  is a subset of  $V$ , then a *solution over  $W$*  of  $S$  is an assignment to  $W$  which agrees with a solution of  $S$  over  $W$ . Two systems are *equivalent* if they have the same set of solutions. Two systems are *equivalent over  $W$*  if they have the same set of solutions over  $W$ .

### 3 Standard Structure for the reals

Let  $\mathbf{R}$  be the set of real numbers and  $\mathbf{Q}$  the set of rational numbers.

By *standard structure* we mean the structure

$$(\mathbf{R}, \mathbf{R}, \mathbf{Q} \cup \{-, +, \times\}, \{>, \geq, =, \neq\})$$

which has

- as its domain, the set of real numbers,
- as its subdomain, the set of real numbers (the variables cover the entire domain),
- as its operations, the rational numbers considered as zero-place operations (that is, as constants), the usual one-place operation  $-$ , and the usual two-place operations  $+$  and  $\times$ ,
- as its relations, the usual two-place relations  $>$ ,  $\geq$ ,  $=$ ,  $\neq$ .

As this structure is classic, we will use the infix notations and the standard abbreviations to refer to terms and constraints. In particular, we will write  $t_1 - t_2$  instead of  $t_1 + (-t_2)$  and we will use  $\sum$  for multiple sums. We should observe, however, that while variables can have as their value irrational numbers, the constants occurring in terms are restricted to rational numbers. A term is *linear* if it does not contain a subterm of the form  $t_1 \times t_2$ , where neither  $t_1$  nor  $t_2$  are constants. This definition is purely syntactic and according to this definition the term  $3 \times x$  is linear, whereas the term  $(2 + 1) \times x$  is not. A system constructed on the standard structure is called *standard*. A standard system containing only linear terms is called *linear*.

In what follows we will need two important properties of linear systems. Here is the first one:

**Property 3.1** *If a standard linear system admits at least one solution then it admits at least one completely rational solution.*

By completely rational solution we mean a solution  $\sigma$  such that for every variable  $x$  the value  $\sigma(x)$  is a rational number. This property forms the content of theorem 6.2 which is dealt with in part 6. We now come to the second property.

**Property 3.2** *Let  $S$  be a standard linear system and  $W$  a subset of the universal set of variables. The two following propositions are equivalent.*

1. *For all  $x \in W$  there exist two solutions  $\sigma$  and  $\tau$  of  $S$  such that  $\sigma(x)$  and  $\tau(x)$  are distinct reals.*
2. *There exist two solutions  $\sigma$  and  $\tau$  of  $S$  such that for all  $x \in W$  the values  $\sigma(x)$  and  $\tau(x)$  are distinct irrational numbers.*

This property forms the content of theorem 7.5 which is dealt with in part 7.

## 4 Non-standard structure for the reals

We now modify the standard structure to obtain a *non-standard structure*

$$(\mathbf{R} \cup \{\omega\}, \mathbf{R}, \mathbf{Q} \cup \{\dot{-}, \dot{+}, \dot{\times}\}, \{\dot{>}, \dot{\geq}, \dot{=}, \dot{\neq}\})$$

defined as follows:

- The domain is the set of real numbers extended by an element  $\omega$  distinct from all the real numbers.
- The subdomain is the set of real numbers (variables cannot take the value  $\omega$ ).
- The operations are the rational numbers considered as zero-place relations (that is, as constants), the one-place operation  $\dot{-}$ , and the two-place operations  $\dot{+}$  and  $\dot{\times}$ . These operations  $\dot{-}, \dot{+}, \dot{\times}$  coincide with the operations  $-, +, \times$ , when all operands are distinct from  $\omega$ , and yield the value  $\omega$ , wherever an operand is  $\omega$ . There are two exceptions to this general rule:

$$\begin{aligned} a \dot{\times} b &= \omega, \text{ if } a \text{ and } b \text{ are irrational,} \\ 0 \dot{\times} \omega &= \omega \dot{\times} 0 = 0. \end{aligned}$$

- The relations are the two-place relations  $\dot{>}, \dot{\geq}, \dot{=}, \dot{\neq}$ . They coincide with the relations  $>, \geq, =, \neq$ , if no operand is  $\omega$ , and are satisfied, whenever an operand is  $\omega$ .

The operation  $\dot{\times}$  does not distribute over  $\dot{+}$ . Indeed, if one considers the irrational number  $\pi$ , we have  $\pi \dot{\times} (\pi \dot{+} (\dot{-}\pi)) = 0$  but  $(\pi \dot{\times} \pi) \dot{+} (\pi \dot{\times} (\dot{-}\pi)) = \omega$ .

The operations  $\dot{+}$  and  $\dot{\times}$  are however commutative and associative. Only the associativity of non-standard multiplication is not obvious. To show that the products  $(a \dot{\times} b) \dot{\times} c$  and  $a \dot{\times} (b \dot{\times} c)$  are equal we consider four possible forms of the triplet  $(a, b, c)$ . First, the triplet contains 0: the products are then equal to 0. Second, the triplet does not contain 0, but does contain  $\omega$ : the products are then equal to  $\omega$ . Third, the triplet contains neither 0 nor  $\omega$  and it contains at most one irrational number: the products are then equal to  $a \times b \times c$ . Fourth, the triplet does not contain 0 or  $\omega$  but contains at least two irrational numbers: the products are then equal to  $\omega$ , (this follows from the fact that the product of a rational and an irrational is an irrational).

It should be observed that if we had decided that  $\omega \times 0 = \omega$ , instead of  $\omega \times 0 = 0$ , non-standard multiplication would have been no longer associative, because if we consider again the irrational number  $\pi$ , one would have had  $\pi \dot{\times} (\pi \dot{\times} 0) = 0$  and  $(\pi \dot{\times} \pi) \dot{\times} 0 = \omega$ .

With each term  $t$ , each constraint  $c$  and each system  $S$  in the standard structure, one can associate the term  $\dot{t}$ , the constraint  $\dot{c}$  and the system  $\dot{S}$  in the non-standard structure obtained by replacing the operations and relations  $-, +, \times, >, \geq, =, \neq$  by the corresponding operations and relations  $\dot{-}, \dot{+}, \dot{\times}, \dot{>}, \dot{\geq}, \dot{=}, \dot{\neq}$ . We can then show by induction on the depth of the term  $t$  that for every assignment  $\sigma$  to the universal set of variables  $V$ ,

$$\sigma^*(t) \neq \sigma^*(\dot{t}) \text{ implies } \sigma^*(\dot{t}) = \omega$$

and therefore that if  $\sigma$  satisfies the constraint  $c$ , then  $\sigma$  also satisfies the constraint  $\dot{c}$ . It follows that:

**Property 4.1** *Every solution of the standard system  $S$  is a solution of the associated non-standard system  $\dot{S}$ .*

If we now take the term  $t$  to be linear then, since  $\omega$  is not among the constants, since no variable can take the value  $\omega$  and since no operation can yield the value  $\omega$ , we always have  $\sigma^*(t) = \sigma^*(\dot{t})$  and therefore  $\sigma$  satisfies the constraint  $c$  if and only if  $\sigma$  also satisfies the constraint  $\dot{c}$ . It follows that:

**Property 4.2** *If  $S$  is linear, then the standard system  $S$  and the associated non-standard system  $\dot{S}$  have the same set of solutions.*

The relation  $\dot{=}$  is not an equivalence relation, since it is not transitive. We have  $1 \dot{=} \omega$  and  $\omega \dot{=} 2$ , but not  $1 \dot{=} 2$ . However, the relation  $\dot{=}$ , like true equality, allows the introduction of intermediate variables to name terms.

**Property 4.3** *Let  $c[t]$  be a non-standard constraint in which an occurrence of a term  $t$  has been chosen and let  $c[x]$  be the same constraint in which the chosen occurrence of  $t$  has been replaced by a variable  $x$  which does not occur in  $c[t]$ . The systems*

$$\{c[t]\} \text{ and } \{c[x], x \dot{=} t\}$$

*are equivalent over the subset of variables  $V - \{x\}$ .*

This property forms the content of theorem 8.3 which is dealt with in section 8.

## 5 Solving non-standard systems

We now have all the necessary elements to show that the naive algorithm figuring in our introduction is a complete algorithm for solving non-standard systems of constraints. We must however agree upon the meaning of the word ‘solving’. Intuitively, the problem is to determine the set of solutions of a system of constraints  $S$  in a given mathematical structure. Since this set can be infinite, as for instance in the system  $\{0 \leq x+y, x+y \leq 4\}$ , it is not always possible to explicitly enumerate its elements. We must therefore be a little less ambitious and for our part, we will stipulate that *solving* a system  $S$  consists of two things: to determine if  $S$  has at least one solution, and if so, to produce a *solved* system  $T$ , which is equivalent to  $S$  over the set of variables of  $S$ . The notion of “solved system” is defined as following:

**Definition 5.1** *A system  $T$  is solved if it has at least one solution and if it is of the form*

$$\{x_1 \approx a_1, \dots, x_n \approx a_n\} \cup T', \quad (2)$$

- *where  $\{x_1, \dots, x_n\}$  is the set of variables  $x_i$  of  $V$  whose value  $\sigma(x_i)$  is the same in every solution  $\sigma$  of  $T$ ,*
- *where the  $a_i$ ’s are constants of the structure under consideration,*
- *where  $\approx$  is a binary relation of the structure under consideration which coincides with the equality relation each time its operands are constants or elements of the sub-domain of the variables.*

According to this definition the relation  $\approx$  will be the relation  $=$  in the standard structure and the relation  $\dot{=}$  in the non-standard structure.

It must be noted that the existence of a solved system, equivalent to a given one, can be a problem because of the lack of some constants. For example, in our standard structure, it is not possible to solve (in terms of the definition we have just given) the system  $\{x \times x = 2, x \geq 0\}$ , given the fact that the number  $\sqrt{2}$  is not rational and thus not one of the constants. It would have been necessary to accept all the algebraic numbers as constants. We will return to this point in our conclusion.

If we restrict ourselves to standard linear systems, the problem just mentioned no longer exists due to the following property.

**Property 5.1** *For every standard linear system which has at least one solution there exists an equivalent solved standard linear system.*

Indeed, let  $S$  be a linear standard system which has at least one solution and let  $\{x_1, \dots, x_n\}$  be the (necessarily finite) set of variables  $x_i$  of  $V$  whose value  $\sigma(x_i)$  is the same real number  $a_i$  in every solution  $\sigma$  of  $S$ . If one of the  $a_i$ 's is an irrational number, there is a contradiction with the property 3.1 which states that there must exist at least one completely rational solution. All the  $a_i$ 's are thus rational numbers and the system  $\{x_1 = a_1, \dots, x_n = a_n\} \cup S$  is in solved form and equivalent to  $S$ .

We can thus assume that we have an algorithm which allows us to solve any standard linear system in the sense specified above. For a detailed account of such an algorithm we recommend [8] and for a more general view [10]. Given property 4.2 concerning the equivalence of standard linear systems and the associated non-standard ones, the same algorithm can be used to solve non-standard linear systems. We can therefore assume that we have an algorithm for solving linear non-standard system and propose the following algorithm for solving non-linear non-standard systems.

**Algorithm 5.1** *Let  $S$  be a non-standard system that is to be solved. We consider the pair  $(S_1, S_2)$  of non-standard systems which is initialized at the beginning to  $(S, \emptyset)$  and we carry out action 1.*

1. *As long as  $S_1$  contains an occurrence of a term or subterm of the form  $s \dot{\times} t$ , where neither  $s$  nor  $t$  are constants, introduce three new variables  $x, y, z$ , replace this occurrence of  $s \dot{\times} t$  by  $z$ , add to  $S_1$  the constraints  $x \doteq s, y \doteq t$  and add to  $S_2$  the constraint  $z \doteq x \dot{\times} y$ . Proceed then to action 2.*
2. *Solve the linear system  $S_1$ . If  $S_1$  has no solution, stop and conclude that the original system  $S$  has no solution. Else replace  $S_1$  with its solved form and proceed to action 3.*
3. *As long as the system  $S_2$  contains a constraint of the form  $z \doteq x \dot{\times} y$ , and the system  $S_1$  contains a constraint of the form  $x \doteq a$  or  $y \doteq a$ , where  $a$  is a constant, remove the constraint  $z \doteq x \dot{\times} y$  from  $S_2$  and add the constraint  $z \doteq a \dot{\times} y$  or  $z \doteq x \dot{\times} a$  to  $S_1$ , as the case may be. Proceed then to action 4.*
4. *If action 3 modified the pair  $(S_1, S_2)$ , repeat action 2. Else stop here and exhibit  $S_1 \cup S_2$  as the solved form of  $S$ .*

There is no difficulty in showing that the algorithm always terminates. Indeed, every time one executes action 2 after action 4 the number of constraints of  $S_2$  diminishes and the number of times that one performs these actions is therefore finite. Let us now show that the answers yielded by the algorithm are correct.

We establish first that after every transformation of the pair  $(S_1, S_2)$  the system  $S_1 \cup S_2$  remains equivalent to itself over the set of variables which it contains. This is true for the transformation 1 given property 4.3. This is equally true in the case of transformation 2. This is also true for transformation 3 given that in the constraint  $x \doteq a$  or  $y \doteq a$  the relation  $\doteq$  behaves like a true equality. We conclude that the system  $S_1 \cup S_2$  is always equivalent to the system  $S$  over the set of variables of  $S$ .

When one detects in action 2 that  $S_1$  is not solvable, it is therefore correct to conclude that  $S$  is not solvable. The only thing left to show is that the final system  $S_1 \cup S_2$  that was produced as an answer at the end of action 2 is solved. This final system is of the form

$$T \cup \{u_1 \doteq v_1 \dot{\times} w_1, \dots, u_m \doteq v_m \dot{\times} w_m\}, \quad (3)$$

where  $T$  is a solved linear system, i.e. of the form (2), and where the  $u_i, v_i, w_i$  are variables, the  $v_i, w_i$  being distinct from the variables  $x_i$  of  $T$ . As the system  $T$  is solved, for every variable  $y$  other than  $x_1, \dots, x_n$  there exist at least two solutions  $\sigma_y$  and  $\tau_y$  of  $T$  such that  $\sigma_y(y) \neq \tau_y(y)$ . Due to properties 4.2 and 3.2 there thus exist two solutions  $\sigma$  and  $\tau$  of  $T$  such that for every variable  $y$  other than  $x_1, \dots, x_n$  the numbers  $\sigma(y)$  and  $\tau(y)$

are irrational and distinct. The assignments  $\sigma$  and  $\tau$  satisfy every equation  $u_i \doteq v_i \times w_i$  because  $\sigma(v_i) \times \sigma(w_i) = \omega$  and  $\tau(v_i) \times \tau(w_i) = \omega$ . The assignments  $\sigma$  and  $\tau$  are thus solutions of the system (3) and since for every variable  $y$  other than  $x_1, \dots, x_n$  we have  $\sigma(y) \neq \tau(y)$ , the system (3) is solved.

## 6 The rational solution theorem

What remains to be shown is the correctness of properties 3.1, 3.2 and 4.3. We begin with the first property. We will be concerned with the standard structure  $\Sigma$  as defined in part 3.

**Lemma 6.1 (Block of solutions)** *Let  $\sigma$  be a solution of a linear system  $S$  containing only constraints of type  $>$  and  $\neq$ . There always exists a strictly positive real number  $h$  such that every assignment  $\tau$  that satisfies the condition*

$$|\tau(x) - \sigma(x)| \geq h, \quad \text{for all } x \in V, \quad (4)$$

*is a solution of  $S$ .*

*Proof.* Let us first consider the case where the system  $S$  contains only constraints of type  $>$ . The system  $S$  can always be put in the form

$$\{t_1 > 0, \dots, t_m > 0\},$$

where every  $t_i$  is a term of the form

$$b_i + \sum_{j=1}^n a_{ij} y_j,$$

and where the  $b_i$ 's and the  $a_{ij}$ 's are real numbers and  $y_1, \dots, y_n$  the variables occurring in  $S$ . Let  $\sigma$  be a solution of  $S$ , that is, an assignment such that for every  $i$  we have

$$\sigma^*(t_i) > 0.$$

If  $m = 0$  or  $n = 0$  or all the  $a_{ij}$ 's are zero, every assignment  $\tau$  of  $V$  is a solution of  $S$  and the property is proved. We can therefore assume that  $m \geq 1$ , that  $n \geq 1$ , that at least one  $a_{ij}$  is not zero and introduce three strictly positive reals  $k, a, h$  such that

$$k < \min\{\sigma^*(t_i)\}, \quad a = \max\{|a_{ij}|\}, \quad h = \frac{k}{na}.$$

Let  $\tau$  be an assignment to  $V$  that respects condition (4). For every  $j$  we have

$$a \times |\sigma(y_j) - \tau(y_j)| \leq k/n$$

and thus for all  $i$  and  $j$  we have successively

$$\begin{aligned} |a_{ij}| \times |\sigma(y_j) - \tau(y_j)| &\geq k/n, \\ a_{ij} \times (\sigma(y_j) - \tau(y_j)) &\geq k/n, \\ a_{ij} \tau(y_j) &\geq a_{ij} \sigma(y_j) - k/n. \end{aligned}$$

By taking the sum with respect to  $j$  and by adding to each member the quantity  $b_i$  we obtain successively and for all  $i$

$$\begin{aligned} b_i + \sum_{j=1}^n a_{ij} \tau(y_j) &\geq b_i + \sum_{j=1}^n a_{ij} \sigma(y_j) - k, \\ \tau^*(t_i) &\geq \sigma^*(t_i) - k, \\ \tau^*(t_i) &> 0. \end{aligned}$$

The assignment  $\tau$  is therefore a solution of  $S$ .

It remains to be shown that the property also holds when in  $S$  the number  $n$  of constraints of the form  $s \neq t$  is not zero. For this it suffices to consider the  $2^n$  systems  $S_i$  obtained by replacing in  $S$  every constraint of the form  $s \neq t$  either by the constraint  $s > t$ , or by the constraint  $t > s$ . We observe that for every system  $S_i$ , each solution of  $S_i$  is a solution of  $S$ . If  $\sigma$  is a solution of  $S$  then there exists necessarily a  $S_i$  admitting  $\sigma$  as solution. It follows from the property that we have just proved, that there exists a strictly positive real number  $h$  such that every assignment  $\tau$  satisfying the condition 4 is a solution of  $S_i$  and thus of  $S$ .

**Theorem 6.2 (Rational Solution)** *If a standard linear system admits at least one solution then it admits at least one completely rational solution.*

Proof. Let us suppose first of all that the system  $S$  under consideration contains no constraint of type  $\geq$ . We now use induction on the number  $n$  of constraints of type  $=$  of  $S$ . Let  $\sigma$  be a solution of  $S$ . If  $n = 0$  then it follows from the property 6.1 concerning the existence of a block of solutions, that there exists a strictly positive real number  $h$  such that every assignment  $\tau$  which satisfies the condition

$$\sigma(x) - h \leq \tau(x) \leq \sigma(x) + h, \text{ for all } x \in V$$

is a solution of  $S$ . Since one can always insert a rational number between two reals, we can choose for every  $x$  a rational number  $a_x$  such that

$$\sigma(x) - h < a_x < \sigma(x) + h.$$

The assignment  $\tau$  defined by  $\tau(x) = a_x$ , for every  $x \in V$ , is thus a completely rational solution of  $S$ . Let us suppose that the property holds for  $n$  and let us show that it holds for  $n + 1$ . Let  $S$  be a solvable linear system containing  $n + 1$  constraints of type  $=$  and let  $c$  be one of these constraints. Since the set of rationals (together with operations  $+$  and  $\times$ ) is a subfield of the reals, the constraint  $c$  can always be put in one of two forms

$$0 = 0 \text{ or } x = t,$$

where  $t$  is a term of the standard structure in which  $x$  does not occur. If  $c$  is of the form  $0 = 0$ , the system  $S - \{c\}$  admits, by induction assumption, a completely rational solution, which is also a completely rational solution of  $S$ . If  $c$  is of the form  $x = t$ , let  $T$  be the system obtained by replacing in  $S - \{c\}$  every occurrence of  $x$  by  $t$ . By induction assumption,  $T$  admits a completely rational solution  $\rho$ , and since the systems  $S$  and  $T \cup \{x = t\}$  are equivalent and since  $T$  does not contain any occurrence of  $x$ , the assignment  $\rho'$  defined by  $\rho'(x) = \rho^*(t)$  and  $\rho'(y) = \rho(y)$ , for every variable  $y$  other than  $x$ , is a completely rational solution of  $S$ .

It remains to be shown that the property holds when in  $S$  the number  $n$  of constraints of the form  $s \geq t$  is not zero. We consider the  $2^n$  systems  $S_i$  obtained by replacing in each  $S$  each constraint of the form  $s \geq t$  first by the constraint  $s = t$ , then by the constraint  $s > t$ . We observe that for every system  $S_i$ , any solution of  $S_i$  is a solution of  $S$ . If  $\sigma$  is a solution of  $S$  then there necessarily exists a  $S_i$  admitting  $\sigma$  as solution. From what we have shown it follows that this  $S_i$  admits a completely rational solution  $\rho$  and from the preceding remark it follows that  $\rho$  is also a solution of  $S$ . The system  $S$  thus admits a completely rational solution.

## 7 Theorem of irrational solutions

To prove property 3.2 we again consider the standard structure defined in 3. Along the way we establish the well-known theorem of independence of  $\neq$  constraints. Another proof

of this theorem which applies to numerous infinite structures [2] can be found in [10]. We need two new notations. If  $a$  is a real and  $\sigma$  and  $\tau$  are assignments to the same subset  $W$  of variables then  $a\sigma$  and  $\sigma + \tau$  are assignments to  $W$  defined by

$$[a\sigma](x) = a\sigma(x),$$

$$[\sigma + \tau](x) = \sigma(x) + \tau(x).$$

**Lemma 7.1 (Convexity)** *Let  $S$  be a standard system without  $\neq$  constraints. Any linear combination  $\sigma = k_1\sigma_1 + \dots + k_n\sigma_n$  of solutions  $\sigma_i$  of  $S$ , where the  $k_i$ 's are non-negative reals such that  $k_1 + \dots + k_n = 1$ , is also a solution of  $S$ .*

*Proof.* We can always assume that all the  $k_i$ 's are not zero, since if  $p$  of them were zero it would be sufficient to prove the lemma for  $n - p$  instead of  $n$ . Let us consider any constraint  $c$  of  $S$ . This constraint can be put into the form

$$a_1x_1 + \dots + a_nx_m \succ b,$$

where the  $x_i$ 's are variables which occur in  $S$  and where  $\succ$  denotes one of the relations  $\geq, >, =$ . By taking into account successively that  $\sigma_i$  is a solution of  $S$ , that  $k_i$  is strictly positive, that  $\sigma = \sum k_i\sigma_i$  and that  $\sum k_i = 1$  we get

$$\begin{aligned} a_1\sigma_i(x_1) + \dots + a_n\sigma_i(x_m) &\succ b, \\ a_1k_i\sigma_i(x_1) + \dots + a_nk_i\sigma_i(x_m) &\succ bk_i, \\ a_1\sum k_i\sigma_i(x_1) + \dots + a_n\sum k_i\sigma_i(x_m) &\succ b\sum k_i, \\ a_1\sigma(x_1) + \dots + a_n\sigma(x_m) &\succ b. \end{aligned}$$

The assignment  $\sigma$  satisfies therefore the constraint  $c$ . This constraint  $c$  being any constraint of  $S$ , it follows that  $\sigma$  is a solution of  $S$ .

**Lemma 7.2 (Pseudo-convexity)** *Let  $c_1, \dots, c_n$  be  $n$  constraints of type  $\neq$  and let  $S$  be a linear standard system without  $\neq$  constraints. Let  $\sigma$  be a solution of  $S$  and  $\tau$  a solution of  $S \cup \{c_1, \dots, c_n\}$ . There exist at most  $n$  reals  $k$  such that  $0 \geq k \geq 1$  and such that the assignment  $k\sigma + (1 - k)\tau$  is not a solution of  $S \cup \{c_1, \dots, c_n\}$ .*

*Proof.* Let  $k$  be a real and let  $\rho$  be the assignment  $\rho = k\sigma + (1 - k)\tau$ . Let us consider a constraint  $c_i$ . This constraint can always be put in the form  $t_i \neq 0$ , where  $t_i$  is a linear term. We have

$$\rho^*(t_i) = \tau^*(t_i) + k(\sigma^*(t_i) - \tau^*(t_i)).$$

Due to the fact that  $\tau$  satisfies the constraint  $c_i$ , we have  $\tau^*(t_i) \neq 0$ . There exists therefore at most one real  $k$  such that  $\rho^*(t_i) = 0$ , that is to say such that  $\rho$  does not satisfy  $c_i$ . It follows that there exist at most  $n$  reals  $k$  such that  $\rho$  does not satisfy  $\{c_1, \dots, c_n\}$ . According to the previous lemma, if  $0 \geq k \geq 1$  the assignment  $\rho$  is a solution of  $S$ . There exist therefore at most  $n$  reals  $k$  such that  $0 \geq k \geq 1$  and such that  $\rho$  is not a solution of  $S \cup \{c_1, \dots, c_n\}$ .

**Theorem 7.3 (Independence of the  $\neq$  constraints)** *Given a linear standard system  $S$  and  $n$  constraints  $c_1, \dots, c_n$  of type  $\neq$ , the two following propositions are equivalent.*

1. *Each of the  $n$  systems  $S \cup \{c_i\}$  has at least one solution.*
2. *The system  $S \cup \{c_1, \dots, c_n\}$  has at least one solution.*

Proof. Obviously proposition 2 entails proposition 1. To prove that proposition 1 entails proposition 2 let us assume that proposition 1 is true and let us show by induction on  $n$  that proposition 2 is true. If  $n = 1$  proposition 2 coincides with proposition 1. If  $n \geq 2$  we can assume that proposition 2 is true for  $n - 1$ , that is to say, that the system

$$S \cup \{c_1, \dots, c_{n-1}\} \quad (5)$$

admits a solution  $\sigma$ . Due to the fact that proposition 1 is assumed to be true, the system

$$S \cup \{c_n\} \quad (6)$$

also admits a solution  $\tau$ . Let  $k$  be a real and let  $\rho$  be the assignment  $\rho = k\sigma + (1 - k)\tau$ . According to the pseudo-convexity property 7.2, the number of reals  $k \in [0, 1]$  such that  $\rho$  is not a solution of system (5) is at most equal to  $n - 1$ . The number of reals  $k \in [0, 1]$  such that  $\rho$  is not a solution of system (6) is equal to the number of reals  $(1 - k) \in [0, 1]$  such that  $\rho$  is not a solution of system (6). According to the pseudo-convexity property 7.2, this last number is at most equal to 1. Therefore there is an infinite number of reals  $k \in [0, 1]$  such that  $\rho$  is simultaneously a solution of both systems. Hence the system  $S \cup \{c_1, \dots, c_n\}$  is solvable.

**Corollary 7.4 (Multiple solutions)** *Let  $S$  be a standard linear system and let  $x_1, \dots, x_n$  be  $n$  variables taken from the universal set of variables. The two following propositions are equivalent.*

1. *For each  $x_i$  there exist two solutions  $\sigma_i$  and  $\tau_i$  of  $S$  such that  $\sigma_i(x_i) \neq \tau_i(x_i)$ .*
2. *There exist two solutions  $\sigma$  and  $\tau$  of  $S$  such that for each  $x_i$  we have  $\sigma(x_i) \neq \tau(x_i)$ .*

Proof. Let  $X$  be the finite set of variables which occur in  $S$  or in  $\{x_1, \dots, x_n\}$ . To each variable  $x$  of  $X$  let us associate a distinct variable  $x'$  which is not in  $X$ . Let  $S'$  denote the system  $S$  in which every variable  $x$  has been replaced by the corresponding variable  $x'$ . Propositions 1 and 2 are then respectively equivalent to the two propositions

1. each system  $S \cup S' \cup \{x_i \neq x'_i\}$  is solvable,
2. the system  $S \cup S' \cup \{x_1 \neq x'_1, \dots, x_n \neq x'_n\}$  is solvable,

which, according to theorem 7.3, are equivalent.

**Theorem 7.5 (Irrational solutions)** *Let  $S$  be a linear standard system and  $W$  a subset of the universal set of variables. The two following propositions are equivalent.*

1. *For all  $x \in W$  there exist two solutions  $\sigma$  and  $\tau$  of  $S$  such that  $\sigma(x)$  and  $\tau(x)$  are distinct reals.*
2. *There exist two solutions  $\sigma$  and  $\tau$  of  $S$  such that for all  $x \in W$  the values  $\sigma(x)$  and  $\tau(x)$  are distinct irrational numbers.*

Proof. As proposition 2 is a particular case of proposition 1, it is sufficient to prove that 1 entails 2. Moreover, due to the fact that we only need to consider the variables of  $W$  which occur in  $S$ , we can assume that  $W$  is a finite set  $\{x_1, \dots, x_n\}$ . According to the previous corollary, if proposition 1 is true there exist two solutions  $\sigma$  and  $\tau$  of  $S$  such that for each  $x_i$  we have  $\sigma(x_i) \neq \tau(x_i)$ . According to the pseudo-convexity property 7.2 there exists a non-countable subset  $A \subset \mathbf{R}$  such that any assignment  $\rho$  of the form

$$\rho = \sigma + a(\tau - \sigma), \quad \text{with } a \in A,$$

is a solution of  $S$ . Due to the fact that for each  $x_i$  we have  $[\tau - \sigma](x_i) \neq 0$ , the mappings

$$\varphi_i : a \mapsto \rho(x_i), \quad \varphi : a \mapsto \rho$$

are injective, that is to say that  $a \neq b$  entails  $\varphi_i(a) \neq \varphi_i(b)$  and  $\varphi(a) \neq \varphi(b)$ . The set of reals

$$B = \bigcup_{i=1}^n \varphi_i^{-1}(\mathbf{Q})$$

is countable since the set  $\mathbf{Q}$  of rationals is countable, since the  $\varphi_i$ 's are injective mappings and since a finite union of countable sets produces a countable set. It follows that the set  $A - B$  is infinite. Let  $a$  and  $b$  be two distinct elements of  $A - B$ . The assignments  $\sigma' = \varphi(a)$  and  $\tau' = \varphi(b)$  are solutions of  $S$  and, due to the fact that the mapping  $\varphi_i$  is injective, are such that for each  $x_i$  the values  $\sigma'(x_i)$  and  $\tau'(x_i)$  are distinct irrational numbers.

## 8 Intermediate variables theorem

To prove the property 4.3 of the relation  $\dot{=}$ , we will consider a more general structure than the non-standard. Instead of taking  $\mathbf{Q}$  as the set of constants we will take the whole domain  $\mathbf{R} \cup \{\omega\}$ .

**Lemma 8.1 (Value of a term containing  $\omega$ )** *Let  $\sigma$  be an assignment, let  $t[\omega]$  be a term in which an occurrence of  $\omega$  has been chosen and let  $t[a]$  be the same term in which the chosen occurrence of  $\omega$  has been replaced by the real  $a$ . Exactly one of the three following propositions is true.*

1.  $\sigma^*(t[\omega]) = \omega$  and there exists a real  $a$  such that  $\sigma^*(t[a]) = \omega$ .
2.  $\sigma^*(t[\omega]) = \omega$ , there exists no real  $a$  such that  $\sigma^*(t[a]) = \omega$  and for all reals  $b$  there exists a real  $a$  such that  $\sigma^*(t[a]) = b$ .
3.  $\sigma^*(t[\omega])$  is a real  $b$  and for all reals  $a$  we have  $\sigma^*(t[a]) = b$ .

*Proof.* Let us proceed by induction on the depth  $i$  of the term  $t[\omega]$ . If  $i = 0$  then the term  $t[\omega]$  can only be  $\omega$  and proposition 2 is true for  $a = b$ . Let us assume that the lemma is true for  $i$  and let us show that the lemma is also true for  $i = n + 1$ . Let us consider the term  $t[\omega]$  of depth  $n + 1$ . There are three possible cases:

The term  $t[\omega]$  is of the form  $\dot{-}s[\omega]$ , where  $s[\omega]$  is a term of depth  $n$  containing the chosen occurrence of  $\omega$ . By assumption one of the three propositions of the lemma is true for  $s[\omega]$ . If it is proposition 1, then proposition 1 is true for  $t[\omega]$ . If it is proposition 2, then proposition 2 is true for  $t[\omega]$  by taking  $-b$  instead of  $b$ . If it is proposition 3, then proposition 3 is true for  $t[\omega]$  by taking  $-b$  instead of  $b$ .

The term  $t[\omega]$  is of the form  $r \dot{+} s[\omega]$  or  $s[\omega] \dot{+} r$ , where  $r$  and  $s[\omega]$  are terms of depth  $i \geq n$ , and where the term  $s[\omega]$  contains the chosen occurrence of  $\omega$ . Either  $\sigma^*(r) = \omega$  and proposition 1 is true for  $t[\omega]$ . Or  $\sigma^*(r)$  is a real. By assumption one of the three propositions of the lemma is then true for  $s[\omega]$ . If it is proposition 1, then proposition 1 is true for  $t[\omega]$ . If it is proposition 2, then proposition 2 is true for  $t[\omega]$  by taking  $b + \sigma^*(r)$  instead of  $b$ . If it is proposition 3 then proposition 3 is true for  $t[\omega]$  by taking  $b + \sigma^*(r)$  instead of  $b$ .

The term  $t[\omega]$  is of the form  $r \dot{\times} s[\omega]$  or  $s[\omega] \dot{\times} r$ , where  $r$  and  $s[\omega]$  are terms of depth  $i \geq n$ , and where the term  $s[\omega]$  contains the chosen occurrence of  $\omega$ . By assumption one of the three propositions of the lemma is true for  $s[\omega]$ . If it is proposition 1 then, according as the value of  $\sigma^*(r)$  is or is not 0, proposition 3 or proposition 1 is true for  $t[\omega]$ . If it is proposition 2 then, according to whether the value  $\sigma^*(r)$  is 0, a non zero rational or an

irrational, proposition 3, proposition 2 or proposition 1 is true for  $t[\omega]$ . If it is proposition 3 then, according to whether the non-standard product  $\sigma^*(s[\omega]) \times \sigma^*(r)$  is  $\omega$  or a real, proposition 1 or proposition 3 is true for  $t[\omega]$ .

**Lemma 8.2 (Implicit existential quantification)** *Let  $\sigma$  be an assignment, let  $c[\omega]$  be a constraint in which an occurrence of  $\omega$  has been chosen and let  $c[a]$  be the same constraint in which the chosen occurrence of  $\omega$  has been replaced by the real  $a$ . The following propositions are equivalent:*

1. Assignment  $\sigma$  satisfies constraint  $c[\omega]$ .
2. There exists a real  $a$  such that  $\sigma$  satisfies  $c[a]$ .

Proof. Let  $\succ$  denote one of the four relations  $>$ ,  $\geq$ ,  $\doteq$ ,  $\neq$ . We will assume that the constraint  $c[\omega]$  is of the form  $s \succ t[\omega]$ , the positions of the terms  $s$  and  $t[\omega]$  being irrelevant in the following. Propositions 1 and 2 can then be stated as

1. we have  $\sigma^*(s) \succ \sigma^*(t[\omega])$ ,
2. there exists a real  $a$  such that  $\sigma^*(s) \succ \sigma^*(t[a])$ .

Assume property 1 is true. There are three cases. First,  $\sigma^*(t[\omega])$  is a real. Let  $a$  be any real. According to the previous lemma  $\sigma^*(t[a]) = \sigma^*(t[\omega])$  and therefore proposition 2 is true for this  $a$ . Second,  $\sigma^*(t[\omega]) = \omega$  and there exists a real  $a$  such that  $\sigma^*(t[a]) = \omega$ . Proposition 2 is then true for this  $a$ . Thirdly,  $\sigma^*(t[\omega]) = \omega$  and there exists no real  $a$  such that  $\sigma^*(t[a]) = \omega$ . If one refers to the definitions of relations  $\succ$  there always exists a real  $b$  such that  $\sigma^*(s) \succ b$ . According to the previous lemma there always exists a real  $a$  such that  $b = \sigma^*(t[a])$  and proposition 2 is true for this  $a$ .

Assume property 2 is true. If  $\sigma^*(t[\omega])$  is a real, according to the previous lemma,  $\sigma^*(t[a]) = \sigma^*(t[\omega])$  and hence proposition 1 is true. If  $\sigma^*(t[\omega]) = \omega$ , due to the fact that one of the operands of  $\succ$  is equal to  $\omega$ , proposition 1 is true.

**Theorem 8.3 (Intermediate variables)** *Let  $c[t]$  be a non-standard constraint in which an occurrence of a term  $t$  has been chosen and let  $c[x]$  be the same constraint in which the chosen occurrence of  $t$  has been replaced by a variable  $x$  which does not occur in  $c[t]$ . The systems*

$$\{c[t]\} \quad \text{and} \quad \{c[x], x \doteq t\}$$

*are equivalent over the subset of variables  $V - \{x\}$ .*

Proof. Let us assume that  $\sigma$  is a solution of system  $\{c[t]\}$  and let us show that there exists a real  $a$  such that  $\sigma$  is a solution of the system  $\{c[a], a \doteq t\}$ . If  $\sigma^*(t)$  is a real, then we take  $a = \sigma^*(t)$ . If  $\sigma^*(t) = \omega$  then, according to the previous lemma, there exists a real  $a$  such that  $\sigma$  is a solution of  $\{c[a]\}$  and thus also of  $\{c[a], a \doteq t\}$ .

Let us assume that  $\sigma$  is a solution of the system  $\{c[a], a \doteq t\}$ , where  $a$  is a real, and let us show that  $\sigma$  is a solution of system  $\{c[t]\}$ . If  $\sigma^*(t)$  is a real, then  $a = \sigma^*(t)$  and therefore  $\sigma$  is a solution of  $\{c[t]\}$ . If  $\sigma^*(t) = \omega$  then, according to the previous lemma,  $\sigma$  is a solution of  $\{c[t]\}$ .

## Conclusion

In this paper we have used three properties of the set of rational numbers and this only in the proofs of theorems 6.2 and 7.5. Here are these properties:

1. Between two reals one can always find a rational.

2. The rational numbers form a subfield of the field of reals, that is, the set of rational numbers is closed under the operations  $-$ ,  $+$ ,  $\times$  and the inverse for multiplication of a non-zero rational number is always a rational number.
3. The set of rational numbers is countable.

We could therefore have replaced  $\mathbf{Q}$  by any other subset  $\mathbf{Q}'$  of the reals with these three properties. The set of constants of standard and non-standard constants would then be  $\mathbf{Q}'$  and  $\omega$  would be the non-standard product of two reals not belonging to  $\mathbf{Q}'$ . A good candidate for  $\mathbf{Q}'$  would be the set of algebraic numbers. It would suffice to replace the word “rational” by “algebraic” and the word “irrational” by “transcendant”.<sup>3</sup>

Other changes are possible. In order to define the numeric part of Prolog III in [3] we used a variant of the non-standard defined here. This variant does not involve the use of the element  $\omega$ , but rather the replacement of the multiplication operation by a three-place relation  $\mu$  defined as follows:  $\mu(a, b, c)$  means that  $a, b$  and  $c$  are reals such that  $c = a \times b$  or that  $a$  and  $b$  are irrationals and that  $c$  is an arbitrary real.

We conclude this paper by giving an indication of the expressive power of non-standard multiplication. In the standard structure one can do without the relations  $>$ ,  $\neq$  and  $\geq$ . Indeed, if we replace

$$\begin{aligned} t_1 > t_2 & \text{ by } t_1 \geq t_2, t_1 \neq t_2, \\ t_1 \neq t_2 & \text{ by } x \times (t_1 - t_2) = 1, \\ t_1 \geq t_2 & \text{ by } t_1 - t_2 = x \times x, \end{aligned}$$

where  $x$  designates a new variable, one transforms every standard system  $S$  in a system of equations  $S'$  that is equivalent to  $S$  on the set of variables of  $S$ . If one applies the same procedure to a non-standard system  $\dot{S}$ , the definitions of the operations and the relations (page 8) permit only the first two replacements:

$$\begin{aligned} t_1 \dot{>} t_2 & \text{ by } t_1 \dot{\geq} t_2, t_1 \dot{\neq} t_2, \\ t_1 \dot{\neq} t_2 & \text{ by } x \dot{\times} (t_1 - t_2) \dot{=} 1. \end{aligned}$$

Non-standard multiplication thus allows doing without the relations  $\dot{>}$  and  $\dot{\neq}$ , but not without the relation  $\dot{\geq}$ . The simplex algorithm is thus as useful as ever!

## Acknowledgements

I thank Gérard Rauzy for having sketched the proof of theorem 7.5 and Frédéric Benhamou for many interesting discussions. I also thank Franz Guenther for his help with the English version of this paper.

## References

- [1] Collins George E., Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. *Lectures Notes in Computer Science*, 134-183. Springer-Verlag, Berlin, 1975, Vol. 33.
- [2] Colmerauer Alain, Equations and inequations on finite and infinite trees, invited lecture, in *Proceedings of the International Conference on Fifth Generation Computer Systems*, p 85-99, Tokyo, November 1984.

---

<sup>3</sup>It is perhaps surprising that the algebraic numbers are countable. Indeed, they are the real numbers which are the roots of polynomials with integer coefficients. Let us denote by  $P_n$  the set of polynomials with integer coefficients of the form  $\sum_{i=0}^n a_i x^i$ , with  $0 \leq |a_i| \leq n$ , for  $i = 1, \dots, n$ . The set  $P_n$  is finite and since a polynomial of degree  $n$  cannot have more than  $n$  roots, the set  $A_n$  of algebraic numbers definable by  $P_n$  is also finite. Since the set of algebraic numbers is equal to  $\bigcup_{n=1}^{\infty} A_n$ , it is countable.

- [3] Colmerauer Alain, An introduction to Prolog III. *Communications of the ACM*, 33(7), 69-90, July 1990.
- [4] Dantzig George B., *Linear Programming and Extension*. Princeton University Press, 1963.
- [5] Dincbas Mehmet, Pascal Van Hentenryck, Helmut Simonis, Abderrahmane Aggoun, T. Graf, and F. Berthier. The Constraint Logic Programming Language CHIP. In *Proceedings on the International Conference on Fifth Generation Computer Systems FGCS-88*, Tokyo, Japan, December 1988.
- [6] Grigor'ev D. Yu. The complexity of deciding Tarski algebra. *Journal of symbolic Computation*, 5(1,2):65-108, 1988.
- [7] Hong Hoon, Comparison of several decision algorithms for the existential theory of the reals. *Technical Report 91-41.0*, Research Institute for Symbolic Computation, Johannes Kepler University, Linz, Austria, September 1991.
- [8] Imbert Jean Louis et Pascal van Hentenryck, On the handling of disequations in CLP over linear rational arithmetic. In this book.
- [9] Jaffar Joxan, Spiro Michaylov, Peter Stuckey and Roland Yap, The CLP( $\mathcal{R}$ ) language and system. *Technical Report RC 16292 (#72336) 11/15/90*, IBM Research Division, November 1990.
- [10] Lassez Jean Louis and Ken McAloon, A canonical form for generalized linear constraints. *IBM technical report RC 15004*. (To appear in *Journal of Symbolic Computation*).
- [11] Prolog III, Version 1.1, Reference and User's Manual, PrologIA, Marseille, March 1990.
- [12] Tarski Alfred, *A Decision Method for Elementary Algebra and Geometry*. Second edition, University of California Press, Berkeley, May 1951.