Résolution naïve de contraintes non linéaires*

Alain Colmerauer†

Mars 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 $x =$ constant which are entailed by S.

The preliminary phase of the naive algorithm consists of introducing intermediate variables for splitting S in two subsystems, a linear one and a non-linear one containing only constraints of the form $z = x \times y$, where x, y et 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 $x =$ 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 ω which is outside of the domain of the reals. The operations are extended by taking ω as the value as soon as one of the arguments is ω. 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 ω. Rational numbers are kept as constants and variables are not allowed to take the value ω.

*Cette recherche a t́é effectuée dans le cadre du projet *Basic Research Action 3012: Compulog.* Une aide a aussi été obtenue dans le cadre du Greco de programmation du CNRS.

†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

Résumé

Nous étudions un algorithme naïf et incomplet pour résoudre des systèmes non linéaires de contraintes construits à partir de variables réelles, de constantes rationnelles, des opérations $-$, $+$, $\times$ et des relations $\geq$, $>$, $=$, $\neq$. Par *résoudre un système* S nous entendons deux choses : tout d'abord, décider si S admet au moins une solution, ensuite, déterminer l'ensemble des équations de la forme $x =$ constante qui sont impliquées par S.

Dans une phase préliminaire on introduit des variables intermédiaires pour scinder S en deux sous-systèmes, l'un linéaire, l'autre non-linéaire et constitué d'équations de la forme $z = x \times y$ où x, y et z sont des variables. L'algorithme naïf proprement dit se résume à répéter deux actions jusqu'à ce que le système n'évolue plus ou que la partie linéaire n'ait plus de solutions. La première action consiste à résoudre la partie linéaire de S. La deuxième action consiste à considérer les équations de la forme $x =$ constante qui sont impliquées par la partie linéaire de S et à remplacer dans les membres droits des équations non-linéaires de S chacun des x par la constante correspondante.

On montre que l'algorithme naïf devient un algorithme complet si l'on se place dans la structure non-standard suivante des réels : La sémantique de la multiplication est modifiée en considérant que le produit de deux nombres irrationnels est un élément étrange ω distincts de tous les réels. Les opérations sont prolongées en considérant que le résultat est ω dès que l'un des arguments est ω. Une entorse à ce principe est faite pour la multiplication par zéro qui produit toujours zéro. Toutes les relations, y compris la relation $=$, sont considérées comme satisfaites dès que l'un au moins des arguments est ω. Les rationnels sont conservés comme constantes et les variables sont astreintes à ne jamais prendre la valeur ω.

Introduction

Il est possible de résoudre des contraintes non-linéaires portant sur des nombres réels. Alfred Tarski l'a montré dès 1930 et a publié ce résultat, bien plus tard, en se plaçant dans le cadre général des quantifications existentielles et universelles [12]. En 1973 George Collins [1] a proposé un algorithme informatisable pour résoudre le même problème. Cet algorithme dit de « décomposition cylindrique » a été le point de départ de nombreux développements. Malheureusement il a été montré que l'algorithme de décomposition cylindrique avait une complexité doublement exponentielle dans le nombre de changements de type des quantificateur [6]. Pour le cas purement existentiel, qui nous intéresse ici, on obtient une complexité simplement exponentielle dans le nombre de variables. En fait en pratique le nombre limite de variables que l'on peut traiter se situe entre 5 et 10. Signalons enfin que pour le cas purement existentiel il existe d'autres algorithmes. Le lecteur intéressé trouvera une comparaison de leurs différentes complexités en [7].

Les concepteurs des trois principaux langages de programmation logique par contraintes Prolog III [3], CLP(R) [9] et CHIP [5] ont donc été bien inspirés de se contenter de résoudre des contraintes essentiellement linéaires. Ils ont ainsi profité de deux algorithmes efficaces [8] : l'algorithme de Gauss, qui sert à éliminer des variables, et l'algorithme du simplexe de George Dantzig [4], qui sert à optimiser une fonction linéaire portant sur des variables soumises à des inégalités linéaires de type $\geq$.

Cependant beaucoup de problèmes qui ont une formulation non-linéaire peuvent être résolus par un algorithme naïf qui consiste à mêler la résolution de la partie linéaire du problème avec l'intégration de contraintes devenues linéaires du fait qu'il a été possible de déterminer la valeur de certains coefficients.

Ce papier est consacré à clarifier et à justifier cet algorithme naïf. Il est organisé en huit parties suivies d'une conclusion. La partie 1 présente informellement l'algorithme naïf sur un exemple. La partie 2 introduit une terminologie pour parler de contraintes indépendamment du domaine ou plus exactement de la structure dans laquelle on se place. Les parties 3 et 4 sont consacrées à une structure standard pour les réels et à une structure non-standard dans laquelle le domaine, les opérations et les relations sont légèrement modifiés. La partie 5 établit le résultat principal de ce papier : l'algorithme naïf résout tout système de contraintes même non-linéaires, non pas dans la structure standard, mais dans la structure non-standard. Les parties 6, 7 et 8 sont consacrées à la démonstration de théorèmes utilisés dans les parties précédentes.

1 Exemple de résolution naïve

Plaçons nous dans le domaine des nombres réels et considérons le système non-linéaire d'équations S_n de la forme

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

où T_0 est le système

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

et T_{i+1} le système

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

Essayons de résoudre S_2 qui est donc

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

Introduisons des variables intermédiaires u_i, v_i, w_i, z_i, z'_i et faisons apparaître tous les produits non-linéaires sous forme de contraintes $z = xy$ où x, y et z sont des variables.

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

Nous pouvons maintenant partitionner le système en une partie purement linéaire et une partie purement non-linéaire.

$$\left\{\begin{array}{lcl} x_0 & = & 1, \\ y_0 & = & 2, \\ u_1+v_1 & = & 0, \\ x_1+y_1 & = & w_1, \\ z_0 & = & x_0-y_0, \\ z'_0 & = & 2x_0+y_0+1, \\ u_2+v_2 & = & 0, \\ x_2+y_2 & = & w_2, \\ z_1 & = & x_1-y_1, \\ z'_1 & = & 2x_1+y_1+1 \end{array}\right\} \cup \left\{\begin{array}{lcl} u_1 & = & x_0x_1, \\ v_1 & = & y_0y_1, \\ w_1 & = & z_0z'_0, \\ u_2 & = & x_1x_2, \\ v_2 & = & y_1y_2, \\ w_2 & = & z_1z'_1 \end{array}\right\}$$

Résolvons la partie linéaire ou, plus exactement, en ne s'occupant que d'elle, rendons explicites toutes les équations implicites qui sont de la forme $x=k$, avec x variable et k constante. Pour ceci on remarquera que la 1ère, la 2ème, la 5ème et la 6ème équation forment un sous-système de 4 équations à 4 inconnues.

$$\left\{\begin{array}{lcl} x_0 & = & 1, \\ y_0 & = & 2, \\ z_0 & = & -1, \\ z'_0 & = & 5, \\ & & \\ u_1+v_1 & = & 0, \\ x_1+y_1 & = & w_1, \\ u_2+v_2 & = & 0, \\ x_2+y_2 & = & w_2, \\ z_1 & = & x_1-y_1, \\ z'_1 & = & 2x_1+y_1+1 \end{array}\right\} \cup \left\{\begin{array}{lcl} u_1 & = & x_0x_1, \\ v_1 & = & y_0y_1, \\ w_1 & = & z_0z'_0, \\ u_2 & = & x_1x_2, \\ v_2 & = & y_1y_2, \\ w_2 & = & z_1z'_1 \end{array}\right\}$$

Remplaçons maintenant dans la partie non-linéaire toute variable x, pour laquelle il existe dans la partie linéaire une équation $x=k$, par k et rapatrions dans la partie linéaire toutes les nouvelles équations linéaires ainsi créées.

$$\left\{\begin{array}{lcl} x_0 & = & 1, \\ y_0 & = & 2, \\ z_0 & = & -1, \\ z'_0 & = & 5, \\ u_1+v_1 & = & 0, \\ x_1+y_1 & = & w_1, \\ u_2+v_2 & = & 0, \\ x_2+y_2 & = & w_2, \\ z_1 & = & x_1-y_1, \\ z'_1 & = & 2x_1+y_1+1, \\ & & \\ u_1 & = & x_1, \\ v_1 & = & 2y_1, \\ w_1 & = & -5 \end{array}\right\} \cup \left\{\begin{array}{lcl} u_2 & = & x_1x_2, \\ v_2 & = & y_1y_2, \\ w_2 & = & z_1z'_1 \end{array}\right\}$$

Résolvons à nouveau la partie linéaire, la 5ème, la 6ème et les 5 dernières équations formant un système de 7 équations à 7 inconnues.

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

Remplaçons maintenant à nouveau dans la partie non-linéaire toute variable x, pour laquelle il existe dans la partie linéaire une équation $x = k$, par k et rapatrions dans la partie linéaire toutes les nouvelles équations linéaires ainsi créées.

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

Résolvons le nouveau système qui est maintenant purement linéaire. Nous obtenons finalement,

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

Cet algorithme naïf étant intégré dans Prolog III [3, 11], il est possible de résoudre le système S_n par le programme :

$$\begin{array}{l} \text{sequence}(\langle\langle 1,2\rangle\rangle) \rightarrow; \\ \text{sequence}(s \cdot \langle\langle x,y\rangle\rangle \cdot \langle\langle x',y'\rangle\rangle) \rightarrow \\ \qquad \text{sequence}(s \cdot \langle\langle x,y\rangle\rangle), \\ \qquad \{x' \times x + y' \times y = 0, \\ \qquad x' + y' = (x-y) \times (2x+y+1)\}; \end{array}$$

Pour caculer la suite de couple

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

on pose la requête[1]

$$\text{sequence}(s), \{|s| = 7\}?$$

qui produit

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

Revenons à notre algorithme naïf. Si nous étions partis du système $\{x^2 - 2x + 1 = 0\}$ nous aurions abouti à quelque chose de la forme $\{z - 2x + 1 = 0,\ z = x \times x\}$ sans découvrir que x valait 2 et, pire, si nous étions partis de $\{x \times x = -4\}$ nous aurions abouti à $\{z = -4,\ z = x \times x\}$ sans découvrir que le système n'avait pas de solutions. On peut donc se poser la question : qu'est-ce que l'algorithme naïf résout au juste dans ces cas là ? Répondre à cette question est le but de l'article.

2 Terminologie

Appellons *structure* un quadruplet

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

constitué d'un domaine D, d'un sous-domaine D' de D, d'un ensemble F d'opérations sur D et d'un ensemble R de relations sur D. A chaque opération et à chaque relation est associé un entier

[1] Dans la version commerciale de Prolog III, la contrainte $|s| = 7$ s'écrit $s :: 7$ et le point d'interrogation devient un point virgule.

non négatif n, le nombre de ses positions. Une opération f à n positions est une application de type $D^n \to D$. Comme d'habitude, les opérations à zéro position sont appelées constantes et sont assimilées à des éléments du domaine. Une relation r à n positions est un sous-ensemble de D^n et plutôt que d'écrire $(a_1, \ldots, a_n) \in r$ on écrit $r(a_1, \ldots, a_n)$.

Pour désigner les éléments du sous-domaine D' on se donne une bonne fois pour toutes un ensemble *universel* infini V de variables et l'on considère deux sortes de formules : les *termes* pour représenter des éléments du domaine et les *contraintes*, pour représenter des propriétés de certains de ses éléments. Plus précisément, un terme est un mot construit sur l'alphabet $V \cup F$ et défini récursivement comme suit : un terme de profondeur 0 est un mot de l'une des deux formes

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

avec $x \in V$ et $a \in D$, et un terme de profondeur $k+1$ est un mot de la forme

$$ft_1 \cdots t_n,$$

où f est une opération à n positions, où l'un au moins des t_i est un terme de profondeur k et où les autres t_i sont des termes de profondeurs au plus égales à k. Les contraintes sont des mots construits sur l'alphabet $V \cup F \cup R$ et qui sont de la forme

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

où r est une relation à n positions et où chaque t_i est un terme.

Une *affectation* d'un sous-ensemble W de variables est une application

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

Une *affectation* tout court est une affectation de l'ensemble universel V de variables. Une telle affectation σ se prolonge naturellement en une application $\sigma^\star$ de l'ensemble des termes dans le domaine en considérant que

$$\begin{array}{rcl} \sigma^\star(x) & = & \sigma(x), \\ \sigma^\star(a) & = & a, \\ \sigma^\star(ft_1 \cdots t_n) & = & f(\sigma^\star(t_1), \ldots, \sigma^\star(t_n)). \end{array}$$

L'affectation σ est *solution* de la contrainte (1) si

$$r(\sigma^\star(t_1), \ldots, \sigma^\star(t_n)).$$

Un *système* de contraintes est un ensemble fini de contraintes. Une *solution* d'un système S de contraintes est une affectation qui est solution de toutes les contraintes de S. Si W est un sous-ensemble de V, alors une *solution sur* W de S est une affectation de W qui coïncide sur W avec une solution de S. Deux systèmes sont *équivalents* s'ils ont le même ensemble de solutions. Deux systèmes sont *équivalents sur* W s'ils ont le même ensemble de solutions sur W.

3 Structure standard pour les réels

Soit $\mathbf{R}$ l'ensemble des nombres réels et $\mathbf{Q}$ l'ensemble des nombres rationnels. Nous appellerons *structure standard* la structure

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

qui a

- pour domaine, l'ensemble des nombres réels,
- pour sous-domaine, l'ensemble de tous les nombres réels, (les variables couvrent tout le domaine),

– pour opérations, les nombres rationnels considérés comme opérations à zéro position (c'est-à-dire comme constantes), l'opération usuelle à une position $-$, et les opérations usuelles à deux positions $+$ et $\times$,

– pour relations, les relations à deux positions usuelles $>$, $\geq$, $=$, $\neq$.

Cette structure étant des plus classiques, nous utiliserons les notations infixées et les raccourcis habituels pour noter les termes et les contraintes. En particulier nous écrivons $t_1 - t_2$ au lieu de $t_1 + (-t_2)$ et ferons usage de $\sum$ pour les sommes multiples. Il faut cependant remarquer que si les variables peuvent prendre pour valeurs des nombres irrationnels, les constantes qui figurent dans les termes sont astreintes ici à être des nombres rationnels. Un terme sera considéré comme *linéaire* s'il ne contient pas de sous-terme de la forme $t_1 \times t_2$, où ni t_1 ni t_2 ne sont des constantes. Il s'agit ici d'une définition purement syntaxique et selon cette définition le terme $3 \times x$ est linéaire alors que le terme $(2 + 1) \times x$ ne l'est pas. Un système construit dans la structure standard sera dit standard. Un système standard qui ne fait intervenir que des termes linéaires sera dit linéaire.

Par la suite nous aurons besoin de deux propriétés importantes des systèmes linéaires. Voici la première.

Propriété 3.1 *Si un système standard linéaire admet au moins une solution alors il admet une solution entièrement rationnelle.*

Par solution entièrement rationnelle nous entendons une solution σ telle que pour toute variable x la valeur $\sigma(x)$ soit un nombre rationnel. Cette propriété est l'objet du théorème 6.2 auquel nous consacrons la partie 6. Nous pouvons passer à la deuxième propriété.

Propriété 3.2 *Soient S un système standard linéaire et W un sous-ensemble de l'ensemble universel de variables. Les deux propositions qui suivent sont équivalentes.*

1. *Pour chaque $x \in W$ il existe deux solutions σ et τ de S telles que $\sigma(x)$ et $\tau(x)$ soient des réels distincts.*
2. *Il existe deux solutions σ et τ de S telles que pour chaque $x \in W$ les valeurs $\sigma(x)$ et $\tau(x)$ soient des nombres irrationnels distincts.*

Cette propriété est l'objet du théorème 7.5 auquel nous consacrons la partie 7.

4 Structure non standard pour les réels

On modifie maintenant la structure standard pour aboutir à la structure *non-standard*

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

définie comme suit :

– Le domaine est l'ensemble des nombres réels complété par un élément ω distinct de tous les nombres réels.

– Le sous-domaine est l'ensemble des nombres réels (les variables ne peuvent prendre la valeur ω).

– Les opérations sont les nombres rationnels considérés comme opérations à zéro position (c'est-à-dire comme constantes), l'opération à une position $\dot{-}$, et les opérations à deux positions $\dot{+}$ et $\dot{\times}$. Ces opération $\dot{-}, \dot{+}, \dot{\times}$ coïncident avec les opérations $-, +, \times$, lorsqu'aucun opérande n'est ω, et produisent la valeur ω, dès qu'un opérande est ω. Il y a deux exceptions á cette règle générale :

$$\begin{array}{l} a \dot{\times} b = \omega, \text{ si } a \text{ et } b \text{ sont irrationnels,} \\ 0 \dot{\times} \omega = \omega \dot{\times} 0 = 0. \end{array}$$

– Les relations sont les relations à deux positions $\dot{>}$, $\dot{\geq}$, $\dot{=}$, $\dot{\neq}$. Elles coïncident avec les relations $>, \geq, =, \neq$, lorsqu'aucun opérande n'est ω, et sont satisfaites, dès qu'un opérande est ω.

L'opération $\dot{\times}$ n'est pas distributive par rapport à $\dot{+}$. En effet, si l'on considère le nombre irrationnel π, on a

$$\begin{array}{lcl} \pi \dot{\times} (\pi \dot{+} (\dot{-}\pi)) & = & 0, \\ (\pi \dot{\times} \pi) \dot{+} (\pi \dot{\times} (\dot{-}\pi)) & = & \omega. \end{array}$$

Les opérations $\dot{+}$ et $\dot{\times}$ sont cependant commutatives et associatives. Seule l'associativité de la multiplication non-standard n'est pas évidente. On montre que les produits $(a \dot{\times} b) \dot{\times} c$ et $a \dot{\times} (b \dot{\times} c)$ sont égaux en distinguant quatre formes possibles du triplet (a, b, c). Premièrement, le triplet contient 0 : les produits sont alors égaux à 0. Deuxièmement, le triplet ne contient pas 0 mais contient ω : les produits sont alors égaux à ω. Troisièmement, le triplet ne contient ni 0 ni ω et contient au plus un irrationnel : les produits sont alors égaux à $a \times b \times c$. Quatrièmement, le triplet ne contient ni 0 ni ω et contient au moins deux irrationnels : les produits sont alors égaux à ω, (on se sert du fait que le produit d'un rationnel et d'un irrationnel est un irrationnel).

Il est à noter que si l'on avait décidé que $\omega \times 0 = \omega$, au lieu de $\omega \times 0 = 0$, la multiplication non-standard ne serait plus associative car, si l'on considère toujours le nombre irrationnel π, on aurait eu $\pi \dot{\times} (\pi \dot{\times} 0) = 0$ et $(\pi \dot{\times} \pi) \dot{\times} 0 = \omega$.

A tout terme t, toute contrainte c et tout système S dans la structure standard on peut associer le terme $\dot{t}$, la contrainte $\dot{c}$ et le système $\dot{S}$ dans la structure non-standard obtenue en remplaçant les opérations et les relations $-, +, \times, >, \geq, =, \neq$ par les opérations et les relations correspondantes $\dot{-}, \dot{+}, \dot{\times}, \dot{>}, \dot{\geq}, \doteq, \dot{\neq}$. On peut alors montrer par récurrence sur la profondeur du terme t que, pour toute affectation σ de l'ensemble universel V de variables,

$$\sigma^\star(t) \neq \sigma^\star(\dot{t}) \text{ entraîne } \sigma^\star(\dot{t}) = \omega$$

et donc si σ satisfait la contrainte c alors σ satisfait aussi la contrainte $\dot{c}$. Il s'ensuit que :

Propriété 4.1 *Toute solution du système standard S est solution du système non-standard associé $\dot{S}$.*

Si le terme t est linéaire alors, du fait que ω ne figure pas comme constante, qu'aucune variable ne peut prendre la valeur ω et qu'aucune opération $\dot{\times}$ ne peut produire la valeur ω, on a toujours $\sigma^\star(t) = \sigma^\star(\dot{t})$. L'affectation σ satisfait donc la contrainte linéaire c si et seulement si σ satisfait aussi la contrainte $\dot{c}$. Il s'ensuit que :

Propriété 4.2 *Si S est linéaire alors le système standard S et le système non-standard associé $\dot{S}$ ont le même ensemble de solutions.*

La relation $\doteq$ n'est pas une relation d'équivalence car elle n'est pas transitive. On a $1 \doteq \omega$ et $\omega \doteq 2$, mais on n'a pas $1 \doteq 2$. Cependant la relation $\doteq$, comme la vraie égalité, permet d'introduire des variables intermédiaires pour nommer des termes.

Propriété 4.3 *Soit $c[t]$ une contrainte non-standard dans laquelle on a privilégié une occurrence du terme t et soit $c[x]$ la même contrainte dans laquelle on a remplacé l'occurrence privilégiée de t par une variable x qui ne figure pas dans $c[t]$. Les systèmes*

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

sont équivalents sur le sous-ensemble de variables $V - \{x\}$.

Cette propriété est l'objet du théorème 8.3 auquel nous consacrons la partie 8.

5 Résolution de systèmes non standards

Nous avons maintenant tous les éléments nécessaires pour montrer que l'algorithme naïf de notre introduction est un algorithme complet pour résoudre des systèmes non-standards de contraintes. Encore faut-il se mettre d'accord sur la signification du mot « résoudre ». Intuitivement il s'agit de déterminer l'ensemble des solutions d'un système S de contraintes dans une structure mathématique donnée. Cet ensemble pouvant être infini comme, par exemple, pour le système $\{0 \leq$

$x + y,\ x + y \leq 4\}$, il n'est pas possible d'énumérer explicitement ses éléments. Il faut donc se montrer moins ambitieux et, pour notre part, nous conviendrons que *résoudre* un système S consiste en deux choses : à déterminer si S a au moins une solution et, dans l'affirmative, à produire un système *résolu* T, équivalent à S sur l'ensemble des variables de S. La notion de « système résolu » est définie comme suit :

Définition 5.1 *Un système T est* résolu *s'il admet au moins une solution et s'il est de la forme*

$$\{x_1 \approx a_1, \ldots, x_n \approx a_n\} \cup T', \tag{2}$$

- *où $\{x_1, \ldots, x_n\}$ est l'ensemble des variables x_i de V dont la valeur $\sigma(x_i)$ est la même dans toute les solution σ de T,*
- *où les a_i sont des constantes de la structure considérée,*
- *où $\approx$ est une relation binaire de la structure considéré qui coïncide avec l'égalité chaque fois que ses arguments sont des constantes ou des éléments du sous-domaine des variables.*

En accord avec cette définition la relation $\approx$ sera la relation $=$ dans la structure standard et la relation $\doteq$ dans la structure non-standard.

Il faut remarquer que l'existence d'un système résolu, équivalent à un système donné, peut être problématique du fait de l'absence de certaines constantes. Par exemple dans notre structure standard, il n'est pas possible de résoudre (dans le sens que nous venons de définir) le système $\{x \times x = 2,\ x \geq 0\}$ du fait que le réel $\sqrt{2}$ ne figure pas parmi les constantes. Il aurait fallu accepter comme constantes tous les nombres algébriques. Nous reviendrons sur ce point dans notre conclusion. Cependant si l'on se limite à des systèmes standards linéaires le problème évoqué ne se pose pas du fait de la propriété qui suit.

Propriété 5.2 *Pour tout système standard linéaire qui a au moins une solution il existe un système standard linéaire résolu qui lui est équivalent.*

En effet soit S un système standard linéaire ayant au moins une solution et soit $\{x_1, \ldots, x_n\}$ l'ensemble (nécessairement fini) des variables x_i de V dont la valeur $\sigma(x_i)$ est le même réel a_i dans toute solution σ de S. Si un des a_i est un nombre irrationnel il y a contradiction avec la propriété 3.1 qui veut qu'il existe au moins une solution entièrement rationnelle. Tous les a_i sont donc des nombres rationnels et le système $\{x_1 = a_1, \ldots, x_n = a_n\} \cup S$ est sous forme résolue et est équivalent à S.

Nous pouvons donc supposer que nous disposons d'un algorithme permettant de résoudre tout système standard linéaire au sens que nous avons précisé. Pour un aperçu détaillé d'un tel algorithme nous renvoyons le lecteur à [8] et pour un aperçu plus global à [10]. Du fait de la propriété 4.2 d'équivalence des systèmes linéaires standard et de leurs systèmes non standard associés, le même algorithme peut être utilisé pour résoudre des système linéaires non standard. Nous pouvons donc aussi supposer que nous disposons d'un algorithme permettant de résoudre tout système linéaire non-standard et proposer l'algorithme suivant pour résoudre les systèmes non-standards.

Algorithme 5.3 *Soit à résoudre un système non-standard S. On considère le couple (S_1, S_2) de systèmes non standards qui au départ est initialisé à $(S, \emptyset)$ et on effectue l'action 1.*

1. *Tant que dans S_1 figure une occurrence d'un terme ou d'un sous-terme de la forme $s \dot{\times} t$, où ni s ni t ne sont des constantes, introduire trois nouvelles variables x, y, z, remplacer cette occurrence de $s \dot{\times} t$ par z, ajouter à S_1 les contraintes $x \doteq s$, $y \doteq t$ et ajouter à S_2 la contrainte $z \doteq x \dot{\times} y$. Exécuter ensuite l'action 2.*
2. *Résoudre le système linéaire S_1. Si S_1 n'a pas de solutions, s'arrêter là et conclure que le système S d'origine n'a pas de solutions. Sinon remplacer S_1 par sa forme résolue et effectuer l'action 3.*
3. *Tant que dans le système S_2 figure une contrainte de la forme $z \doteq x \dot{\times} y$, et que dans le système S_1 figure une contrainte de la forme $x \doteq a$ ou $y \doteq a$, où a est une constante, enlever la contrainte $z \doteq x \dot{\times} y$ de S_2 et ajouter, suivant le cas, la contrainte $z \doteq a \dot{\times} y$ ou $z \doteq x \dot{\times} a$ à S_1. Effectuer ensuite l'action 4.*

4. Si l'action 3 a modifié le couple (S_1, S_2), effectuer à nouveau l'action 2. Sinon s'arrêter là et fournir $S_1 \cup S_2$ comme forme résolue de S.

Il n'y a pas de problème pour montrer que l'algorithme se termine toujours. En effet chaque fois que l'on exécute l'action 2 après l'action 4 le nombre de contraintes de S_2 diminue et donc le nombre de fois que l'on exécute ces actions est fini. Montrons maintenant que les réponses fournies par l'algorithme sont correctes.

Etablissons tout d'abord qu'après chaque transformation du couple (S_1, S_2) le système $S_1 \cup S_2$ reste équivalent à lui même sur l'ensemble des variables qui y figuraient. C'est vrai dans la transformation 1 du fait de la propriété 4.3. C'est bien entendu vrai dans la transformation 2. C'est vrai dans la transformation 3 du fait que dans la contrainte $x \doteq a$ ou $y \doteq a$ la relation $\doteq$ se comporte comme une vraie égalité. On en conclut que le système $S_1 \cup S_2$ est constamment équivalent au système S sur l'ensemble des variables de S.

Lorsque dans l'action 2 on détecte que S_1 n'est pas soluble il est donc correct de conclure que S n'est pas soluble. Il ne reste plus qu'à montrer que le système final $S_1 \cup S_2$ que l'on fournit comme réponse à la fin de l'action 2 est résolu. Ce système final est de la forme

$$T \cup \{u_1 \doteq v_1 \dot{\times} w_1, \ \ldots \ , u_m \doteq v_m \dot{\times} w_m\}, \tag{3}$$

où T est un système linéaire résolu, donc de la forme (2), et où les u_i,v_i,w_i sont des variables, les v_i,w_i étant distincts des variables x_i de T. Le système T étant résolu, pour toute variable y autre que $x_1, \ldots, x_n$ il existe au moins deux solutions σ_y et τ_y de T tels que $\sigma_y(y) \neq \tau_y(y)$. D'après les propriété 4.2 et 3.2 il existe donc deux solution σ et τ de T telles que pour toute variable y autre que $x_1, \ldots, x_n$ les nombres $\sigma(y)$ et $\tau(y)$ soient irrationnels et distincts. Les affectations σ et τ satisfont chaque équation $u_i \doteq v_i \dot{\times} w_i$ car $\sigma(v_i) \dot{\times} \sigma(w_i) = \omega$ et $\tau(v_i) \dot{\times} \tau(w_i) = \omega$. Les affectations σ et τ sont donc solutions du système (3) et, du fait que pour toute variable y autre que $x_1, \ldots, x_n$ on ait $\sigma(y) \neq \tau(y)$, le système (3) est résolu.

6 Théorème de la solution rationnelle

Il ne reste plus qu'à démontrer l'exactitude des propriétés 3.1, 3.2 et 4.3. Commençons par la première. On se place donc dans la structure standard Σ définie dans la partie 3.

Lemme 6.1 (Pavé de solutions) *Soit σ une solution d'un système linéaire S comportant uniquement des contraintes de type $>$ et $\neq$. Il existe toujours un réel h strictement positif tel que toute affectation τ qui satisfait à la condition*

$$|\tau(x) - \sigma(x)| \leq h, \quad \textit{pour tout } x \in V, \tag{4}$$

soit solution de S.

Preuve. Considérons d'abord le cas où le système S ne comporte que des contraintes de type $>$. Le système S peut toujours être mis sous la forme

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

où chaque t_i est un terme de la forme

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

les b_i et les a_{ij} étant des nombres réels et $y_1, \ldots, y_n$ étant les variables apparaissant dans S. Soit σ une solution de S, c'est-à-dire une affectation telle que pour tout i on ait

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

Si $m = 0$ ou $n = 0$ ou si tous les a_{ij} sont nuls, toute affectation τ de V est solution de S et la propriété est démontrée. On peut donc supposer que $m \geq 1$, que $n \geq 1$, qu'au moins un a_{ij} n'est pas nul et introduire trois réels strictement positifs k, a, h tels que

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

Soit τ une affectation de V qui respecte la condition (4). Pour tout j on a

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

et donc pour tout i et j on a successivement

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

En sommant par rapport à j et en ajoutant à chaque membre la quantité b_i on obtient successivement et pour tout i

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

L'affectation τ est donc solution de S.

Il ne reste plus qu'à démontrer que la propriété est vraie lorsque dans S le nombre n de contraintes de la forme $s \neq t$ n'est pas nul. Il suffit alors de considérer les 2^n systèmes S_i obtenus en remplaçant dans S chaque contrainte de la forme $s \neq t$ soit par la contrainte $s > t$, soit par la contrainte $t > s$. On remarque que quelque soit le système S_i, toute solution de S_i est solution de S. Si σ est une solution de S alors il existe forcément un S_i admettant σ pour solution. D'après la propriété que nous venons de démontrer, il existe alors un réel strictement positif h tel que toute affectation τ qui satisfait à la condition 4 soit solution de S_i et donc de S.

Théorème 6.2 (Solution rationnelle) *Si un système standard linéaire admet au moins une solution alors il admet une solution entièrement rationnelle.*

Preuve. Supposons tout d'abord que le système considéré S ne contienne aucune contrainte de type $\geq$. Raisonnons par récurrence sur le nombre n de contraintes de type $=$ de S. Soit σ une solution de S. Si $n = 0$ alors d'après la propriété 6.1 d'existence d'un pavé de solutions, il existe un réel h strictement positif tel que toute affectation τ qui satisfait à la condition

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

soit solution de S. Du fait qu'entre deux réels on peut toujours glisser un nombre rationnel, pour chaque variable x on peut choisir un nombre rationnel a_x tel que

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

L'affectation τ définie par $\tau(x) = a_x$, pour tout $x \in V$, est donc une solution entièrement rationnelle de S. Supposons la propriété vraie pour n et démontrons qu'elle est vraie pour $n+1$. Soit S un système linéaire soluble contenant $n+1$ contrainte de type $=$ et soit c une de ces contraintes. Du fait que l'ensemble des rationnels forment un sous-corps des réels, la contrainte c peut toujours être mise sous l'une des deux formes

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

où t est un terme de la structure standard dans lequel x ne figure pas. Si c est de la forme $0 = 0$, le système $S - \{c\}$ admet, par supposition, une solution entièrement rationnelle, qui est aussi une solution entièrement rationnelle de S. Si c est de la forme $x = t$, soit T le système obtenu en remplaçant dans $S - \{c\}$ toute occurence de x par t. Par supposition T admet une solution ρ, entièrement rationnelle, et du fait que les systèmes S et $T \cup \{x = t\}$ sont équivalents et que T ne contient aucune occurrence de x l'affectation ρ' définie par $\rho'(x) = \rho^\star(t)$ et $\rho'(y) = \rho(y)$, pour toute variable y autre que x, est une solution entièrement rationnelle de S.

Il ne reste plus qu'à démontrer que la propriété est vraie lorsque dans S le nombre n de contraintes de la forme $s \geq t$ n'est pas nul. On considére les 2^n systèmes S_i obtenus en remplaçant dans S chaque contrainte de la forme $s \geq t$ d'abord par la contrainte $s = t$, puis par la contrainte

$s > t$. On remarque que quelque soit le système S_i, toute solution de S_i est solution de S. Si σ est une solution de S alors il existe forcément un S_i admettant σ pour solution. D'après ce que nous venons de démontrer ce système S_i admet une solution entèrement rationnelle ρ et d'après la remarque précédente ρ est aussi solution de S. Le système S admet donc une solution entièrement rationnelle.

7 Théorème des solutions irrationnelles

Pour montrer la propriété 3.2 on se place toujours dans la structure standard définie dans la partie 3. Au passage on établit le théorème bien connu de l'indépendance des contraintes de type $\neq$, dont on trouve une autre démonstration en [10] et qui s'applique à beaucoup de structures infinies [2]. Nous utilisons deux notations nouvelles. Si a est un réel, et σ et τ des affectations d'un même sous-ensemble W de variables alors $a\sigma$ et $\sigma + \tau$ sont des affectations de W définies par

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

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

Lemme 7.1 (Convexité) *Soit S un système standard ne contenant pas de contraintes de type $\neq$. Toute combinaison linéaire $\sigma = k_1\sigma_1 + \cdots + k_n\sigma_n$ de solutions σ_i de S, où les k_i sont des réels non négatifs et tels que $k_1 + \cdots + k_n = 1$, est aussi une solution de S.*

Preuve. On peut toujours supposer que tous les k_i sont distincts de 0 car s'il y en a p de nuls il suffit de démontrer le lemme pour $n - p$ au lieu de n. Considérons une contrainte c quelconque de S. Elle peut être mise sous la forme

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

où les x_i sont les variables qui apparaissent dans S et où $\succ$ désigne l'une des relations $\geq, >, =$. En tenant compte successivement de ce que σ_i est solution de S, de ce que k_i est strictement positif, de ce que $\sigma = \sum k_i\sigma_i$ et de ce que $\sum k_i = 1$ on obtient

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

L'affectation σ satisfait donc la contrainte c. Cette contrainte c étant une contrainte quelconque de S, il s'ensuit que σ est solution de S.

Lemme 7.2 (Pseudo-convexité) *Soient n contraintes $c_1, \ldots, c_n$ de type $\neq$ et soit S un système standard linéaire ne contenant pas de contraintes de type $\neq$. Soient σ une solution de S et τ une solution de $S \cup \{c_1, \ldots, c_n\}$. Dans l'interval $[0, 1]$ il existe au plus n réels k tels que l'affectation $k\sigma + (1 - k)\tau$, ne soit pas solution de $S \cup \{c_1, \ldots, c_n\}$.*

Preuve. Soit k un réel et soit l'affectation $\rho = k\sigma + (1 - k)\tau$. Considérons une contrainte quelconque c_i. On peut toujours la mettre sous la forme $t_i \neq 0$, où t_i est un terme linéaire. On a

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

Du fait que τ satisfait la contrainte c_i on a $\tau^\star(t_i) \neq 0$. Il existe donc au plus un réel k tel que $\rho^\star(t_i) = 0$, c'est-à-dire tel que ρ ne satifasse pas c_i. Il s'ensuit qu'il existe au plus n réels k tels que ρ ne satisfasse pas $\{c_1, \ldots, c_n\}$. D'après le lemme précédent, si $k \in [0, 1]$ l'affectation ρ est solution de S. Il existe donc au plus n réeIs $k \in [0, 1]$ tels que ρ ne soit pas solution de $S \cup \{c_1, \ldots, c_n\}$.

Théorème 7.3 (Indépendance des $\neq$) *Etant données un système linéaire standard S et n contraintes $c_1, \ldots, c_n$ de type $\neq$, les deux propositions qui suivent sont équivalentes.*

1. *Chacun des n systèmes $S \cup \{c_i\}$ a au moins une solution.*
2. *Le système $S \cup \{c_1, \ldots, c_n\}$ a au moins une solution.*

Preuve. La proposition 2 entraîne évidemment la proposition 2. Pour montrer que la proposition 1 entraîne la proposition 2 supposons que la proposition 1 soit vraie et montrons par récurrence sur n que la proposition 2 est vraie. Si $n = 1$ la proposition 2 est vraie car sa formulation est la même que celle de la proposition 1. Si $n \geq 2$ nous pouvons supposer que que la proposition 2 est vraie pour $n - 1$, c'est-à-dire que le système

$$S \cup \{c_1, \ldots, c_{n-1}\} \tag{5}$$

admet une solution σ. Du fait que la proposition 1 est supposée vraie, le système

$$S \cup \{c_n\} \tag{6}$$

admet aussi une solution τ. Soit k un réel et soit l'affectation $\rho = k\sigma + (1 - k)\tau$. D'après la propriété 7.2 de pseudo-convexité, le nombre de réels $k \in [0, 1]$ tels que ρ ne soit pas solution du système (5) est au plus égal à $n - 1$. Le nombre de réels $k \in [0, 1]$ tels que ρ ne soit pas solution du système (6) est égal au nombre de réels $(1 - k) \in [0, 1]$ tels que ρ ne soit pas solutions du système (6). D'après la propriété de pseudo-convexité, ce dernier nombre est au plus égal à 1. Il reste donc une infinité de réels $k \in [0, 1]$ tel que ρ soit solution des systèmes 5 et 6 à la fois. Le système $S \cup \{c_1, \ldots, c_n\}$ est donc soluble.

Corollaire 7.4 (Solutions multiples) *Soit S un système standard linéaire et soient n variables $x_1, \ldots, x_n$ prises dans l'ensemble universel de variables. Les deux propositions qui suivent sont équivalentes.*

1. *Pour chaque x_i il existe deux solutions σ_i et τ_i de S telles que $\sigma_i(x_i) \neq \tau_i(x_i)$.*
2. *Il existe deux solutions σ et τ de S telles que pour chaque x_i on ait $\sigma(x_i) \neq \tau(x_i)$.*

Preuve. Soit X l'ensemble fini des variables qui figure dans S ou dans $\{x_1, \ldots, x_n\}$. A chaque variable x de X associons un variable distincte x' qui ne figure pas dans X. Désignons par S' le système S dans lequel on a remplaçé chaque variable x par la variable correspondante x'. Les propositions 1 et 2 sont alors respectivement équivalentes aux deux propositions

1. chaque système $S \cup S' \cup \{x_i \neq x'_i\}$ est soluble,
2. le système $S \cup S' \cup \{x_1 \neq x'_1, \ldots, x_n \neq x'_n\}$ est soluble,

qui, d'après le théorème 7.3, sont équivalentes.

Théorème 7.5 (Solutions irrationnelles) *Soient S un système standard linéaire et W un sous-ensemble de l'ensemble universel de variables. Les deux propositions qui suivent sont équivalentes.*

1. *Pour chaque $x \in W$ il existe deux solutions σ et τ de S telles que $\sigma(x)$ et $\tau(x)$ soient des réels distincts.*
2. *Il existe deux solutions σ et τ de S telles que pour chaque $x \in W$ les valeurs $\sigma(x)$ et $\tau(x)$ soient des nombres irrationnels distincts.*

Preuve. La proposition 2 étant un cas particulier de la proposition 1, il suffit de montrer que 1 entraîne 2. De plus, du fait qu'il suffit de considérer les variables de W qui figurent dans S, on peut supposer que W est un ensemble fini $\{x_1, \ldots, x_n\}$. D'après le corollaire précédent, si la propostion 1 est vraie il existe deux solutions σ et τ de S telle que pour chaque x_i on ait $\sigma(x_i) \neq \tau(x_i)$. D'après la propriété 7.2 de pseudo-convexité il existe un sous-ensemble non dénombrable $A \subset \mathbf{R}$ tel que toute affectation ρ de la forme

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

soit solution de S. Du fait que pour chaque x_i on a $[\tau - \sigma](x_i) \neq 0$, les applications

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

sont injectives, c'est-à-dire que $a \neq b$ entraîne $\varphi_i(a) \neq \varphi_i(b)$ et $\varphi(a) \neq \varphi(b)$. L'ensemble des réels

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

est dénombrable du fait que l'ensemble $\mathbf{Q}$ des rationnels l'est, que φ_i est une application injective et qu'une union finie d'ensembles dénombrables est dénombrable. Il s'ensuit que l'ensemble $A - B$ est infini. Soient a et b deux éléments distincts de $A - B$. Les affectations $\sigma' = \varphi(a)$ et $\tau' = \varphi(b)$ sont solutions de S et, du fait que l'application φ_i est injective, sont telles que pour chaque x_i les valeurs $\sigma'(x_i)$ et $\tau'(x_i)$ sont des nombres irrationnels distincts.

8 Théorème des variables intermédiaires

Pour démontrer la propriété 4.3 de la relation $\doteq$, nous nous placerons dans une structure un peu plus générale que la structure non-standard en considérant que l'ensemble des constantes n'est pas $\mathbf{Q}$ mais tout le domaine $\mathbf{R} \cup \{\omega\}$.

Lemme 8.1 (Valeur d'un terme contenant ω) *Soit σ une affectation, soit $t[\omega]$ un terme dans lequel on a privilégié une occurrence de ω et soit $t[a]$ le même terme dans lequel on a remplacé l'occurrence privilégiée de ω par le réel a. Une et une seule des trois propositions qui suivent est vraie.*

1. *$\sigma^\star(t[\omega]) = \omega$ et il existe un réel a tel que $\sigma^\star(t[a]) = \omega$.*
2. *$\sigma^\star(t[\omega]) = \omega$, il n'existe aucun réel a tel que $\sigma^\star(t[a]) = \omega$ et, quel que soit le réel b, il existe un réel a tel que $\sigma^\star(t[a]) = b$.*
3. *$\sigma^\star(t[\omega])$ est un réel b et pour tout réel a on a $\sigma^\star(t[a]) = b$.*

Preuve. Raisonnons par récurrence sur la profondeur i du terme $t[\omega]$. Si $i = 0$ alors le terme $t[\omega]$ ne peut être que ω et la proposition 2 est vraie pour $a = b$. Supposons le lemme vrai pour $i \leq n$ et montrons que le lemme est vrai pour $i = n + 1$. Considérons le terme $t[\omega]$ de profondeur $n + 1$. Trois cas se présentent :

Le terme $t[\omega]$ est de la forme $\dot{-}s[\omega]$, où $s[\omega]$ est un terme de profondeur n contenant l'occurrence privilégiée de ω. Par supposition l'une des trois propositions du lemme est vraie pour $s[\omega]$. Si c'est la proposition 1 alors la proposition 1 est vraie pour $t[\omega]$. Si c'est la proposition 2 alors la proposition 2 est vraie pour $t[\omega]$ en prenant $-b$ au lieu de b. Si c'est la proposition 3 alors la proposition 3 est vraie pour $t[\omega]$ en prenant $-b$ au lieu de b.

Le terme $t[\omega]$ est de la forme $r \dot{+} s[\omega]$ ou $s[\omega] \dot{+} r$, où r et $s[\omega]$ sont des termes de profondeur $i \leq n$, le terme $s[\omega]$ contenant l'occurrence privilégiée de ω. Soit, $\sigma^\star(r) = \omega$ et la proposition 1 est vraie pour $t[\omega]$. Soit, $\sigma^\star(r)$ est un réel. Par supposition l'une des trois propositions du lemme est alors vraie pour $s[\omega]$. Si c'est la proposition 1 alors la proposition 1 est vraie pour $t[\omega]$. Si c'est la proposition 2 alors la proposition 2 est vraie pour $t[\omega]$ en prenant $b + \sigma^\star(r)$ au lieu de b. Si c'est la proposition 3 alors la proposition 3 est vraie pour $t[\omega]$ en prenant $b + \sigma^\star(r)$ au lieu de b.

Le terme $t[\omega]$ est de la forme $r \dot{\times} s[\omega]$ ou $s[\omega] \dot{\times} r$, où r et $s[\omega]$ sont des termes de profondeur $i \leq n$, le terme $s[\omega]$ contenant l'occurrence privilégiée de ω. Par supposition l'une des trois propositions du lemme est vraie pour $s[\omega]$. Si c'est la proposition 1 alors suivant que la valeur de $\sigma^\star(r)$ est ou n'est pas 0 la proposition 3 ou la proposition 1 est vraie pour $t[\omega]$. Si c'est la proposition 2 alors suivant que la valeur de $\sigma^\star(r)$ est 0, un rationnel différent de 0 ou un irrationnel, la proposition 3, la proposition 2 ou la proposition 1 est vraie pour $t[\omega]$. Si c'est la proposition 3 alors, suivant que le produit non standard $\sigma^\star(s[\omega]) \dot{\times} \sigma^\star(r)$ est ω ou un réel, la proposition 1 ou la proposition 3 est vraie pour $t[\omega]$.

Lemme 8.2 (Quantification existentielle implicite) *Soit σ une affectation, soit $c[\omega]$ une contrainte dans laquelle on a privilégié une occurrence de ω et soit $c[a]$ la même contrainte dans lequel on a remplacé l'occurrence privilégiée de ω par le réel a. Les propositions qui suivent sont équivalentes :*

1. *L'affectation σ satisfait la contrainte $c[\omega]$.*

2. Il existe un réel a tel que σ satisfasse $c[a]$.

Preuve. Désignons par $\dot{\succ}$ l'une des quatre relations $\dot{>}$, $\dot{\geq}$, $\doteq$, $\dot{\neq}$. On supposera que la contrainte $c[\omega]$ est de la forme $s \mathrel{\dot{\succ}} t[\omega]$, le même raisonnement s'appliquant si les positions des termes s et $t[\omega]$ étaient inversés. Les propositions 1 et 2 deviennent alors

1. on a $\sigma^\star(s) \mathrel{\dot{\succ}} \sigma^\star(t[\omega])$,
2. il existe un réel a tel que $\sigma^\star(s) \mathrel{\dot{\succ}} \sigma^\star(t[a])$.

Supposons que la proposition 1 soit vraie. Trois cas se présentent. Premièrement, $\sigma^\star(t[\omega])$ est un réel. Soit a un réel quelconque. D'après le lemme précédent $\sigma^\star(t[a]) = \sigma^\star(t[\omega])$ et donc la proposition 2 est vraie pour cet a. Deuxièmement, $\sigma^\star(t[\omega]) = \omega$ et il existe un réel a tel que $\sigma^\star(t[a]) = \omega$. La proposition 2 est alors vraie pour cet a. Troisièmement, $\sigma^\star(t[\omega]) = \omega$ et il n'existe aucun réel a tel que $\sigma^\star(t[a]) = \omega$. Si l'on se reporte aux définitions des relations $\dot{\succ}$ il existe toujours un réel b tel que $\sigma^\star(s) \mathrel{\dot{\succ}} b$. D'après le lemme précédent il existe alors un réel a tel que $b = \sigma^\star(t[a])$ et la proposition 2 est vraie pour cet a.

Supposons que la proposition 2 soit vraie. Si $\sigma^\star(t[\omega])$ est un réel, d'après le lemme précédent, $\sigma^\star(t[a]) = \sigma^\star(t[\omega])$ et donc la proposition 1 est vraie. Si $\sigma^\star(t[\omega]) = \omega$, du fait qu'un des arguments de $\dot{\succ}$ est égal à ω, la proposition 1 est vraie.

Théorème 8.3 (Variables intermédiaires) *Soit $c[t]$ une contrainte non-standard dans laquelle on a privilégié une occurrence du terme t et soit $c[x]$ la même contrainte dans laquelle on a remplacé l'occurrence privilégiée de t par une variable x qui ne figure pas dans $c[t]$. Les systèmes*

$$\{c[t]\} \quad \textit{et} \quad \{c[x],\ x \doteq t\}$$

sont équivalents sur le sous-ensemble de variables $V - \{x\}$.

Preuve. Supposons que σ soit solution du système $\{c[t]\}$ et montrons qu'il existe un réel a tel que σ soit solution du système $\{c[a],\ a \doteq t\}$. Si $\sigma^\star(t)$ est un réel alors il suffit de prendre $a = \sigma^\star(t)$. Si $\sigma^\star(t) = \omega$ alors, d'après le lemme précédent, il existe un réel a tel que σ soit solution de $\{c[a]\}$ et donc aussi de $\{c[a],\ a \doteq t\}$.

Supposons que σ soit solution du système $\{c[a],\ a \doteq t\}$, où a est un réel, et montrons que σ est solution du système $\{c[t]\}$. Si $\sigma^\star(t)$ est un réel alors $a = \sigma^\star(t)$ et donc σ est solution de $\{c[t]\}$. Si $\sigma^\star(t) = \omega$ alors, d'après le lemme précédent, σ est solution de $\{c[t]\}$.

Conclusion

Dans tout l'exposé nous avons utilisé trois propriétés de l'ensemble $\mathbf{Q}$ des nombres rationnels et ceci uniquement dans les démonstrations des théorèmes 6.2 et 7.5. Voici ces propriétés :

1. Entre deux réels on peut toujours glisser un rationnel.
2. Les rationnels forment un sous-corps du corps des réels, c'est-à-dire que l'ensemble des rationnels est fermé par rapport aux opérations $-, +, \times$ et que l'inverse multiplicatif d'un rationnel non nul est toujours un rationnel.
3. L'ensemble des rationnels est dénombrable.

On aurait donc pu remplacer $\mathbf{Q}$ par tout autre sous-ensemble $\mathbf{Q}'$ des réels qui aurait ces trois propriétés. L'ensemble des constantes des structures standards et non-standards deviendrait alors $\mathbf{Q}'$ et ce serait le produit non-standard de deux réels n'appartenant pas à $\mathbf{Q}'$ qui serait ω. Un bon candidat pour $\mathbf{Q}'$ serait l'ensemble des nombres algébriques. Il suffirait alors de remplacer le mot « rationnel » par « algébrique » et le mot « irrationnel » par « transcendant ».[2]

[2] On peut être surpris que les nombres algébriques soient dénombrables. En effet, ce sont les nombres réels qui sont racines de polynôme à coefficient entiers. Désignons par P_n l'ensemble des polynômes à coefficients entiers de la forme $\sum_{i=0}^{n} a_i x^i$, avec $0 \leq |a_i| \leq n$ pour $i = 1, \ldots, n$. L'ensemble P_n est fini et, comme un polynôme de degré n ne peut avoir plus de n racines, l'ensemble A_n des nombres algébriques définissables par P_n l'est aussi. L'ensemble des nombres algébriques étant égal à $\bigcup_{n=1}^{\infty} A_n$, il est dénombrable.

D'autres changements sont possibles. Pour définir la partie numérique de Prolog III en [3] nous avons utilisé une variante de la structure non-standard présentée ici. Cette variante consiste à ne pas introduire d'élément ω mais à remplacer l'opération de multiplication par une relation μ à trois positions définie comme suit : $\mu(a,b,c)$ signifie que a, b et c sont des réels tels que $c = a \times b$ ou que a et b sont des irrationnels et que c un réel quelconque.

Terminons cet article en donnant une idée du pouvoir expressif de la multiplication non-standard. Dans la structure standard on peut se passer des relations $>$, $\neq$ et $\geq$. En effet en remplaçant

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

où x désigne une nouvelle variable, on transforme tout système standard S en un système d'équation S', équivalent à S sur l'ensemble des variables de S. Si l'on applique la même procédure sur un système non-standard $\dot{S}$, les définitions des opérations et des relations (page 8) n'autorisent que les deux premiers remplacements :

$$\begin{array}{lll} t_1 \mathrel{\dot{>}} t_2 & \text{par} & t_1 \mathrel{\dot{\geq}} t_2,\ t_1 \mathrel{\dot{\neq}} t_2, \\ t_1 \mathrel{\dot{\neq}} t_2 & \text{par} & x \mathbin{\dot{\times}} (t_1 \mathbin{\dot{-}} t_2) \doteq 1. \end{array}$$

La multiplication non-standard permet donc de se passer des relations $\dot{>}$ et $\dot{\neq}$ mais pas de la relation $\dot{\geq}$. L'algorithme du simplexe a encore de beaux jours !

Remerciements

Je remercie Gérard Rauzy pour avoir esquissé la démonstration du théorème 7.5 et Frédéric Benhamou pour nos nombreuses et intéressantes discussions.

Références

[1] Collins George E., Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. *Lectures Notes in Computer Science*, pages 134-183. Springer-Verlag, Berlin, 1975, Vol. 33.

[2] Colmerauer Alain, Equations and Inequations on Finite and Infinite Trees, invited lecture, dans *Proceeddings of the International Conference on Fifth Generation computer Systems*, p 85-99, Tokyo, Novembre 1984.

[3] Colmerauer Alain, An introduction to Prolog III. *Communications of the ACM*, 33(7) : 69-90, Juillet 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, et F. Berthier. The Constraint Logic Programming Language CHIP. Dans *Proceedings on the International Conference on Fifth Generation Computer Systems FGCS-88*, Tokyo, Japon, Decembre 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, Comparaison of several decision algorithms for the existential theory of the reals. *Tecnical Report 91-41.0*, Research Institute for Symbolic Computation, Johannes Kepler University, Linz, Autriche, Septembre 1991.

[8] Imbert Jean Louis et Pascal van Hentenryck, Efficient Handling of Disequations in CLP over Linear Rational Arithmetic, à paraître dans un livre sur la programmation logique par contrainte, probablement publié en 1992 par Mit Press.

[9] Jaffar Joxan, Spiro Michaylov, Peter Stuckey et Roland Yap, The CLP(**R**) language and system. *Technical Report RC 16292 (#72336) 11/15/90*, IBM Research Division, Novembre 1990.

[10] Lassez Jean Louis and Ken McAloon, A canonical form for generalized linear constraints. *IBM technical report RC 15004.* A paraître en 1991 dans *Journal of Symbolic Computation.*

[11] Prolog III, Version 1.1, Manuel de Référence et d'Utilisation, PrologIA, Marseille, Mars 1990.

[12] Tarski Alfred, *A Decision Method for Elementary algebra and Geometry.* Second edition, University of California Press, Berkeley, Mai 1951.