INFORMATIQUE THÉORIQUE ET APPLICATIONS

M. DAUCHET
E. TIMMERMAN

Continuous monoids and yields of infinite trees

Informatique théorique et applications, tome 20, nº 3 (1986), p. 251-274
<http://www.numdam.org/item?id=ITA_1986__20_3_251_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 20, n° 3, 1986, p. 251 à 274)

CONTINUOUS MONOIDS AND YIELDS OF INFINITE TREES (*)

by M. DAUCHET (1) and E. TIMMERMAN (2)

Communicated by A. ARNOLD

Abstract. – *We define in a canonical algebraic way the structure of (order-) continuous monoid and the notion of the yield of an infinite tree. We prove the decidability of the equality of the yields of two regular infinite trees.*

Résumé. – *Nous spécifions, dans l'esprit des catégories, ce que doivent être un monoïde continu et le feuillage d'un arbre infini (la continuité est prise au sens des C.P.O.). Nous prouvons la décidabilité de l'égalité des feuillages pour les arbres infinis réguliers.*

INTRODUCTION

Infinite words (MacNaughton [8], Nivat [10], Nivat-Perrin [11]) and Infinite trees (Courcelle [4], Nivat [9]) have been studied a lot. Courcelle [2] and Timmerman [13] associate with one infinite tree a frontier [2] or a yield [13] as a generalization of the yield of a finite tree.

There are two goals in this paper:

– to specify, in a categorical way what is a continuous monoid and to construct the free structure,

– to deduce the notion of the (free) yield of an infinite tree and to prove the decidability of the equality of the yields of two infinite regular trees.

A monoid M is continuous if it is provided with a partial order such that M is also a C.P.O. and the concatenation is continuous. We construct and describe the free continuous monoid $W^{\infty}(\Sigma)$ generated by an alphabet Σ.

(*) Received in March 1985, revised in June 1985.
(1) Laboratoire d'Informatique Fondamentale de Lille U.A. 369 C.N.R.S., Université de Lille-I, U.E.R. d'I.E.E.A. 59655 Villeneuve-d'Ascq Cedex, France.
(2) I.U.T. « A » de Lille-I, Département Informatique, B.P. n° 179, 59653 Villeneuve-d'Ascq Cedex, France.

The set $W(\Sigma)$ of finite words is the quotient $(\Sigma \cup \Omega)^*/\Omega \equiv \Omega\Omega$, where Ω is a new symbol (the bottom element). $W^\infty(\Sigma)$ is constructed from $W(\Sigma)$ by completion and $W(\Sigma)$ is the finitary basis of the algebraic C.P.O. $W^\infty(\Sigma)$.

If we consider a congruence on $(\Sigma \cup \{\Omega\})^*$ that can be represented by a confluent and noetherian term rewriting system (Huet [7], Courcelle [3]), and if the normal form mapping N is increasing, then $N(W^\infty(\Sigma))$ is a continuous monoid. For instance, we deduce the usual infinite words (Nivat [10]) from the congruence $\Omega a \equiv \Omega$ (for every $a \in \Sigma$).

Here we consider infinite trees in the usual sense (Courcelle [4], Nivat [10]) with the syntactic order. Then an application φ of infinite trees into a continuous monoid is a yield-application *iff* φ is continuous and, for every tree $f(t_1, \ldots, t_n)$, we have :

$$\varphi(f(t_1, \ldots, t_n)) = \varphi(t_1) \cdot \ldots \cdot \varphi(t_n).$$

Then, we consider initial yields (in the categorical sense). They are words in $W^\infty(\Sigma)$.

Courcelle [2] has introduced frontiers of infinite trees as a generalization of yields of trees to the infinite case. But "frontier" is a different notion than "yield": if the frontiers of two infinite trees are equal, so are their yields, but not conversely. Intuitively, frontier takes into account more information about the structure of the infinite branches (see the examples in part II of this paper).

Unfortunately, the frontier is not continuous. This is the reason why we introduce the yield.

The problem of decidability of the equality of the frontiers of regular trees has been recently solved by Thomas [12]. In part III of this paper, we give a decision algorithm for equality of yields of regular trees. The principle is to construct a rational language from a regular tree, this language being a directed subset (in C.P.O. sense) whose lub is the yield of the tree. This is done in two different ways:

– constructing an automaton from a system that represents a regular tree,

– using Heilbrunner's results [6] and transforming his regular expressions (that represent frontiers) into rational ones that represent the languages we are looking for.

I. CONTINUOUS MONOIDS

I.1. Definitions

Let M be a set provided with an operation $\circ$ called concatenation and with a partial order $<$.

(*a*) *Continuous monoid*

DEFINITION : *M is a continuous monoid* iff *it is both:*

– *a monoid* (*w.r.t. the concatenation*),

– *a C.P.O.* (*Complete Partial Order*)
and satisfies:

– *the order relation is compatible with the concatenation, i. e.* $(x<y \;\&\; z<t)$ *implies* $x \circ z < y \circ t$,

– *the concatenation is continuous, i. e.* $\mathrm{lub}(\mathrm{D}) \circ \mathrm{lub}(\mathrm{D}') = \mathrm{lub}(\mathrm{D} \circ \mathrm{D}')$ *for all directed subsets D and D′ of M.*

Let us remark that the set $D \circ D' = \{x \circ y / x \in D \;\&\; y \in D'\}$ is a directed subset of M whenever D and D' are directed subsets, since the order is compatible with the concatenation.

Let us note, for a continuous monoid M:

$\circ_M$: the concatenation,

λ_M: the neutral element,

$<_M$: the order,

and $\perp_M$: the least element of M
(we sometimes omit the subscript M).

PROPERTY: $\perp_M \circ \perp_M = \perp_M$

Proof: $\perp_M <_M \perp_M \circ \perp_M$

$$\left.\begin{array}{l}\perp_M <_M \lambda_M \\ \perp_M <_M \perp_M\end{array}\right\} \text{thus } \perp_M \circ \perp_M <_M \lambda_M \circ \perp_M = \perp_M$$

by the compatibility of $\leqq_M$ with $\circ$, and by antisymmetry one gets the equality.

(*b*) *Morphisms*

DEFINITION: *Let M, P be continuous monoids, a mapping* $h: M \to P$ *is a morphism* (*of continuous monoids*) *if it is:*

– *a morphism of monoid,*

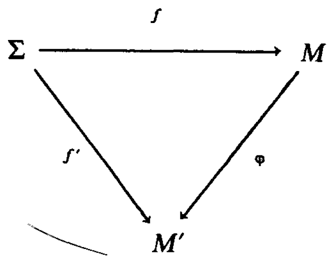
– *a morphism of CPO*

(*i. e. a continuous mapping such that* $h(\perp_M) = \perp_P$)

(*c*) *Free continuous monoid*

Let Σ be an alphabet. We consider, from now on, the continuous monoids generated by Σ; i. e. the triples (Σ, f, M) denoted by $\Sigma \xrightarrow{f} M$ where M is a continuous monoid and f a mapping from Σ into M.

A morphism of continuous monoids generated by Σ is a morphism (of continuous monoids) $\varphi : M \to M'$ such that the following diagram commutes:



i. e. $\forall a \in \Sigma, f'(a) = \varphi(f(a))$.

This forms a category.

The initial object of the category is called a free continuous monoid (generated by Σ), whenever it exists and it is then unique up to isomorphism.

I.2. Construction of the free continuous monoid

Let Ω be a symbol not in Σ (Ω is the symbol for "undefined").

We consider the free monoid $(\Sigma \cup \{\Omega\})^*$, i. e. the set of finite words over $\Sigma \cup \{\Omega\}$ provided with the usual concatenation.

Let $\leqq_\Omega$ denote the relation on $(\Sigma \cup \{\Omega\})^*$ defined by:

for all x, y

$$x \leqq_\Omega y \ \overline{\text{iff}} \begin{cases} - \ x \in \Sigma^* \ \& \ x = y \\ \text{or} \\ - \ x = x_0 \Omega x_1 \Omega \ldots x_{n-1} \Omega x_n, \ x_i \in \Sigma^* \\ y = x_0 y_1 x_1 y_2 \ldots x_{n-1} y_n x_n, \ y_j \in (\Sigma \cup \{\Omega\})^* \\ n \geqq 1 \end{cases}$$

that is: y is obtained from x by substituting arbitrary words for occurrences of Ω.

It is easy to check that $\leqq_\Omega$ is the least preorder on $(\Sigma \cup \{\Omega\})^*$ compatible with the concatenation such that Ω is less than every word.

Let $\simeq$ denote the congruence on $(\Sigma \cup \{\Omega\})^*$ generated by: $\Omega\Omega \simeq \Omega$.

Then, $v \simeq w$ iff $v \leqq_\Omega w$ and $w \leqq_\Omega v$ for all v, w, and thus the quotient set $(\Sigma \cup \{\Omega\})^*/\simeq$ is (partially) ordered by $\leqq_\Omega$ (same notation for the preorder on words and the order on congruence's classes).

Let us denote by $W(\Sigma)$ the monoid $(\Sigma \cup \{\Omega\})^*/\simeq$.

The canonical representative of an element of $W(\Sigma)$ is the shortest word of the class, i. e. the word which has no two successive occurences of the symbol Ω; it is obtained as a normal form by the confluent and Noetherian rewriting system: $\Omega\Omega \to \Omega$.

We always identify an element (also called word) of $W(\Sigma)$ with its canonical representative.

The concatenation on $W(\Sigma)$ is the corresponding quotient operation.

The empty word ε (more precisely the class $\{\varepsilon\}$ of ε) is the neutral element for this operation.

The order $(\leqq_\Omega)$ on $W(\Sigma)$ is compatible with the concatenation, and $\Omega\Omega^*$, (the class of Ω) is the least element of $W(\Sigma)$.

$W(\Sigma)$ is not a C.P.O.: it can be completed by ideal completion (standard construction), which gives an ω-algebraic C.P.O. denoted by $W^\infty(\Sigma)$ whose least element is Ω. $W(\Sigma)$ is the finitary basis of the C.P.O. $W^\infty(\Sigma)$, hence:

– $\forall w \in W^\infty(\Sigma)$, $\exists D \subseteqq W(\Sigma)$ such that $w = \mathrm{lub}(D)$

and

– $\forall D \subseteqq W(\Sigma)$, D being a directed subset $\mathrm{lub}(D) \in W(\Sigma)$ iff D is finite.

The concatenation on $W(\Sigma)$ extends by continuity to $W^\infty(\Sigma)$, and thus, by construction, $W^\infty(\Sigma)$ is a continuous monoid generated by Σ:

$$\Sigma \xrightarrow{\text{id}} W^\infty(\Sigma)$$

with id the identity mapping

PROPERTY: *$W^\infty(\Sigma)$ is the free continuous monoid generated by Σ.*

See [5] for the proof, which is straightforward.

We will call 'infinite term" a non-finite element of $W^\infty(\Sigma)$.

I.3. (usual) infinite words as forming a continuous monoid

(a) Preliminary

PROPERTY: *If $\equiv$ is a congruence (of monoids) on $W(\Sigma)$ that can be oriented into a confluent and Noetherian term rewriting system (t.r.s.), then there is an isomorphism of monoids N between $W(\Sigma)/\equiv$ and $N(W(\Sigma))$ in the following* *sense:*

— *for all m, N(m) denotes the normal form of m associated with the t.r.s.*

— *$N(W(\Sigma))$ is the set of normal forms and the concatenation, denoted by .,* *is defined by:*

$$N(m).N(m')=N(m.m')$$

Example: The congruence defined by $\forall a \in \Sigma$, $\Omega a \equiv \Omega$ is oriented into the Noetherian and confluent t.r.s.

$$\forall a \in \Sigma, \qquad \Omega a \to \Omega.$$

The normal form N is such that: for any $m_1, \ldots, m_p \in \Sigma^*$

$$N(m_1 \Omega \ldots \Omega m_p)=m_1 \Omega$$

N is an isomorphism of $W(\Sigma)/\equiv$ on $N(W(\Sigma))=\Sigma^* \Omega \cup \Sigma^*$ and, in $N(W(\Sigma))$, for all m, $m' \in \Sigma^*$, one has:

$$m\Omega.m'\Omega=N(m\Omega m'\Omega)=m\Omega,$$
$$m.m'\Omega=mm'\Omega,$$

and

$$m\Omega.m'=m\Omega,$$
$$m.m'=mm'.$$

PROPERTY: *If N is monotone, then*

(*a*) *N extends* (*by continuity*) *to* $W^\infty(\Sigma) \to W^\infty(\Sigma)$,

(*b*) $N(W^\infty(\Sigma))$ *provided with the induced concatenation and with the order* $\leq_\Omega$ (*limited to this subset of* $W^\infty(\Sigma)$) *is a continuous monoid. It is isomorphic to* $W^\infty(\Sigma)/\equiv$.

See [5] for the proof.

(*b*) *Infinite words*

An infinite word is a mapping $w: N_+ \to \Sigma$ with Σ the alphabet and N_+ the set of positive integers.

Σ^ω denotes the set of infinite words over Σ and $\Sigma^\infty=\Sigma^* \cup \Sigma^\omega$ the set of words (over Σ). It is well known that Σ^∞ provided with the prefix order is a (ω-algebraic) C.P.O. whose finitary basis is Σ^*. It is also a monoid (with the usual concatenation).

Since the prefix order is not compatible with concatenation, Σ^∞ is not a continuous monoid.

Let $\equiv$ be the congruence on $W(\Sigma)$ previously defined ($\Omega a \equiv \Omega$).

The mapping $N: W(\Sigma) \to W(\Sigma)$ associated with the congruence is monotone.

The set of normal forms $\mathrm{N}(W(\Sigma))$ is equal to $\Sigma^* \cup \Sigma^* \Omega$.

The order $\leqq_\Omega$, limited to this set, can be defined by:

$$x \leqq_\Omega y \text{ iff } \begin{cases} - \ x \in \Sigma^* \text{ and } x = y \\ \text{or} \\ - \ x = x_1 \Omega, \ x_1 \in \Sigma^* \\ y = x_1 y_1, \ y_1 \in \Sigma^* \cup \Sigma^* \Omega \end{cases}$$

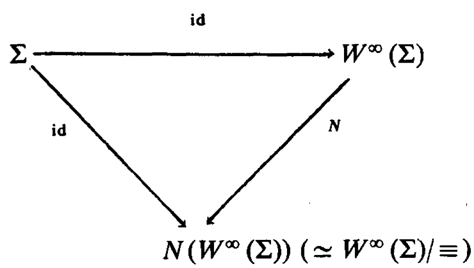
It is very similar to the prefix order.

N extends to $W^\infty(\Sigma)$ and

$$N(W^\infty(\Sigma)) = \Sigma^* \cup \Sigma^* \Omega \cup \{\text{lub}(D)/D \text{ directed subset } \subseteqq \Sigma^* \Omega\}$$

Let us remark that a directed subset D of $\Sigma^* \Omega$ is an increasing sequence (w.r.t. $\leqq_\Omega$).

The set $\{\text{lub}(D)/D$ directed subset of $\Sigma^* \Omega\}$ corresponds exactly to Σ^ω, and thus $N(W^\infty(\Sigma)) = \Sigma^* \Omega \cup \Sigma^\infty$ which is a continuous monoid generated by Σ:



II. THE YIELD OF AN INFINITE TREE

II.1. Trees

Let Σ be a ranked alphabet and Ω be the symbol of arity 0 such that $\Omega \notin \Sigma$ (it means "undefined").

Let us denote by Σ_i, $i \in N$, the subset of Σ of symbols of arity i; Σ_0 is the set of constant symbols.

We consider $T_\Omega^\infty(\Sigma)=T^\infty(\Sigma \cup \{\Omega\})$ the set of finite and infinite trees over $\Sigma \cup \{\Omega\}$.

For each tree t in $T_\Omega^\infty(\Sigma)$, $\mathrm{dom}(t)$ denotes its tree-domain and $fr(t)$ the string of terminal nodes ordered by lexicographic order in the tree-domain t being considered as a partial mapping $t: N_+^* \to \Sigma \cup \{\Omega\}$ with the usual properties.

The syntactic order on trees is defined by:

$$t<t' \overline{\text{iff}}\ \mathrm{dom}(t) \subseteq \mathrm{dom}(t') \text{ and for all } w \text{ in } \mathrm{dom}(t),$$
$$\text{if } t(w) \neq \Omega \text{ then } t'(w)=t(w)$$

$T_\Omega^\infty(\Sigma)$, provided with the syntactic order, is an ω-algebraic C.P.O.; Ω is the least element, and $T_\Omega(\Sigma)$, the set of finite trees over $\Sigma \cup \{\Omega\}$, its finitary basis.

The set of maximal trees, w.r.t. the syntactic order, is $T^\infty(\Sigma)$ the set of trees over Σ.

The yield operation on finite trees is the mapping:

$$fg: \quad T_\Omega(\Sigma) \to (\Sigma \cup \{\Omega\})^*$$

defined by:

for all t in $T_\Omega(\Sigma)$:

- $fg(t)=t$ if $t \in \Sigma_0 \cup \{\Omega\}$,
- $fg(t)=fg(t_1)\ldots fg(t_p)$ whenever

$$t=f(t_1, \ldots, t_p),\ f \in \Sigma_p.$$

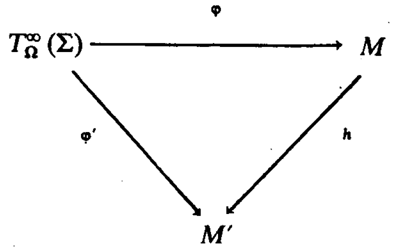
II.2. Yield operation and initial yield

DEFINITION: *We call yield a mapping* φ: $T_\Omega^\infty(\Sigma) \to M$, *where* M *is a continuous monoid, such that:*

- φ *is strict:* $\varphi(\Omega)=\perp_M$,
- φ *is monotonous:* $t<t'$ *implies* $\varphi(t)<_M \varphi(t')$,
- φ *is continuous:* $\varphi(\mathrm{lub}(D))=\mathrm{lub}(\varphi(D))$ *for all directed subset* D *of* $T_\Omega^\infty(\Sigma)$ *and*
- $\varphi(t)=\varphi(t_1)\circ \ldots \circ \varphi(t_p)$ *whenever*

$$t=f(t_1, \ldots, t_p), \qquad f \in \Sigma_p$$

DEFINITION: *A yield* $\varphi : T_\Omega^\infty(\Sigma) \to M$, *is initial* iff *for any yield,* $\varphi' : T_\Omega^\infty(\Sigma) \to M'$, *there exists one and only one morphism (of continuous monoids)* $h : M \to M'$ *such that* $\varphi' = h \circ \varphi$ *i.e. such that the diagram commutes:*



Immediate property: The yield $\varphi : T_\Omega^\infty(\Sigma) \to W^\infty(\Sigma_0)$ *defined by:*

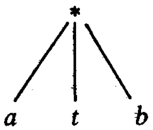
$$\varphi(a) = a \quad \text{for all } a \in \Sigma_0$$

is the initial yield.

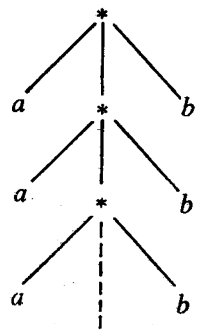
Example: Let M be $\Sigma_0^* \Omega \cup \Sigma^\infty$, the continuous monoid previously defined. Let $\varphi_M : T_\Omega^\infty(\Sigma) \to M$ be the yield defined by $\varphi_M(a) = a$ for all a in Σ_0. Then the morphism $h : W^\infty(\Sigma_0) \to M$ is defined by $h(a) = a$ for all a in Σ_0.

Let t be the tree defined by the equation:

$t =$



so that $t =$



Considering finite approximations of t, we get:

$$\varphi(t) = \text{lub}\{a^n \Omega b^n / n \in N\} = \text{lub}(a^* \Omega b^*)$$

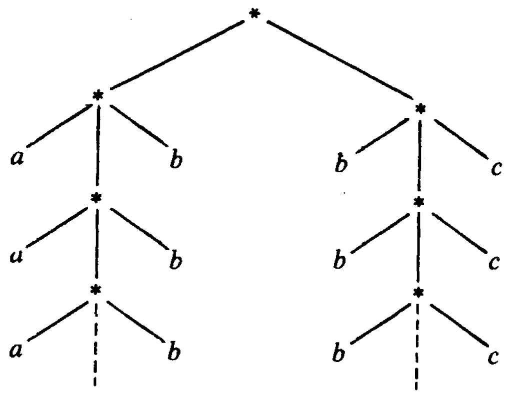
and

$$\varphi_M(t) = \text{lub}\{a^n . \Omega . b^n / n \in N\} = \text{lub}\{a^n \Omega / n \in N\}$$

thus

$$\varphi_M(t) = a^\omega = h(\varphi(t))$$

Let t' be the tree:



$$\varphi(t') = \text{lub}\{a^n \Omega b^{n+p} \Omega c^p / n, p \in N\} = \text{lub}(a^* \Omega b^* \Omega c^*)$$

$$\varphi_M(t') = h(\varphi(t')) = \text{lub}\{h(a^n \Omega b^{n+p} \Omega c^p)/n, p \in N\}$$

$$= \text{lub}\{a^n . \Omega . b^{n+p} . \Omega . c^p / n, p \in N\}$$

$$= \text{lub}\{a^n . \Omega / n \in N\} = a^\omega.$$

II.3. Yield and frontier

The notion of frontier of an infinite tree, has been introduced by Courcelle [2]. It is based on the definition of arrangements or generalized infinite words. We just recall the main definitions and properties and then compare yield and frontier, infinite terms and arrangements.

Arrangements

Let X be an alphabet. An arrangement over X is a triple $u = \langle D, \pi, h \rangle$ consisting of a set D, a total order π on D, and a mapping $h: D \to X$. u is said countable whenever D is.

$A_\omega(X)$ denotes the set of countable arrangements. The words of X^* are identified with the finite arrangements $\langle [n], \leqq, h \rangle$ where $[n] = \{1, 2, \ldots, n\}$ and $\leqq$ the natural order on integers. The arrangements $\langle N_+, \leqq, h \rangle$ correspond to the infinite words of X^ω.

An equivalence relation on arrangements is defined as follows: if $u = \langle D, \pi, h \rangle$ and $u' = \langle D', \pi', h' \rangle$, u and v are said equivalent ($u \equiv v$) iff there exists a bijective order preserving mapping $f: D \to D'$ such that $h = h' \circ f$. The concatenation of arrangement can also be defined (*see* Courcelle [2] and Heilbrunner [6]).

We consider $T^{\infty}(\Sigma)$ the set of trees over Σ [i. e. the set of maximal trees in $T_{\Omega}^{\infty}(\Sigma)$].

Let us recall that a tree is locally finite iff every branch leads at least to a leaf.

$T^{loc}(\Sigma)$ denotes the set of locally finite trees.

DEFINITION: *For each tree* $t: \mathrm{dom}(t) \to \Sigma$, *the "frontier" of t is the countable arrangement* $\psi(t) = \langle fr(t), \leq_l, t \rangle$ *(there* $\leq_l$ *is the lexicographic order).*

PROPERTIES:

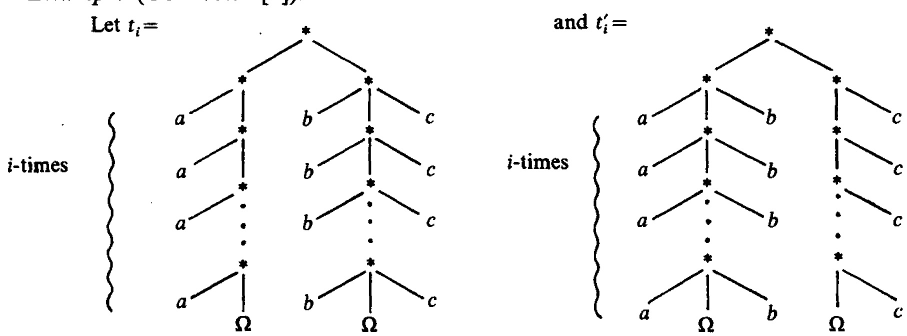
- *If* $t = f(t_1, \ldots, t_p)$, $f \in \Sigma_p$ *then:* $\psi(t) = \psi(t_1) \ldots \psi(t_p)$,
- *for each finite tree* $t: \psi(t) \equiv fg(t)$.

THEOREM (Courcelle):

$$\mathscr{A}_{\omega}(\Sigma_0) = \{\psi(t)/t \in T^{\infty}(\Sigma)\} = \{\varepsilon\} \cup \{\psi(t)/t \in T^{loc}(\Sigma)\}.$$

Let us recall that ψ is not "continuous" in $T_{\Omega}^{\infty}(\Sigma)$, i. e.: if $(t_i)_{i \in N}$ and $(t'_i)_{i \in N}$ are increasing sequences of trees in $T_{\Omega}(\Sigma)$ such that $\psi(t_i) \equiv \psi(t'_i)$ for all i in N, this does not imply $\psi(\mathrm{lub}\{t_i/i \in N\}) \equiv \psi(\mathrm{lub}\{t'_i/i \in N\})$.

Example (Courcelle [2]*):*



For all $i \in N$ we have:

$$\psi(t_i) \equiv \psi(t'_i) \equiv a^i \Omega b^i \Omega c^i$$

but

$$\psi(\mathrm{lub}\{t_i/i \in N\}) \not\equiv \psi(\mathrm{lub}\{t'_i/i \in N\}).$$

Consequence: For all $t, t' \in T^{loc}(\Sigma)$

$$\varphi(t) = \varphi(t') \text{ does not imply } \psi(t) \equiv \psi(t').$$

On this example, if:

$$t = \operatorname{lub}\{t_i/i \in N\} \text{ and } t' = \operatorname{lub}\{t'_i/i \in N\}$$

$$\varphi(t) = \varphi(t') = \operatorname{lub}\{a^i \Omega b^i \Omega c^i/i \in N\} = \operatorname{lub}(a^* \Omega b^* \Omega c^*).$$

Conversely, we have the following:

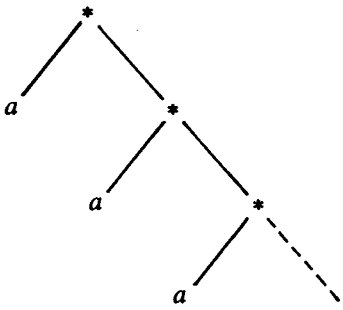
PROPERTY: *For all* $t, t' \in T^{\text{loc}}(\Sigma)$

$$\psi(t) \equiv \psi(t') \text{ implies } \varphi(t) = \varphi(t').$$

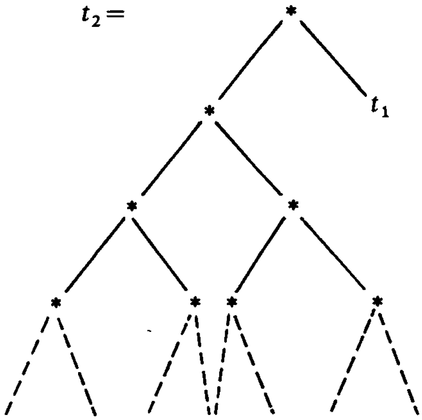
Note that this property does not generally hold in $T^\infty(\Sigma)$.

Example:

$t_1 =$



$t_2 =$



$$\varphi(t_1) = \operatorname{lub}\{a^n \Omega/n \in N\},$$

$$\varphi(t_2) = \operatorname{lub}\{\Omega a^n \Omega/n \in N\} \neq \varphi(t_1)$$

and

$$\psi(t_1) \equiv \psi(t_2) \equiv a^\omega.$$

These properties show that the frontier is a finer operation than the yield, and that the elements of $W^{\text{loc}}(\Sigma_0)$ (i. e. the set of yields of locally finite trees) are classes of countable arrangements.

More precisely, there exists an onto mappling $s:\mathscr{A}_{\omega}(\Sigma_0) \to W^{loc}(\Sigma_0)$ such that: for all $t \in T^{loc}(\Sigma)$

$$\varphi(t) = s(\psi(t))$$

and

$$u \equiv v \text{ implies } s(u) = s(v).$$

II.3. Yields of regular trees

From now on, we only consider the initial yield $\varphi: T_{\Omega}^{\infty}(\Sigma) \to W^{\infty}(\Sigma_0)$ or its restriction to $T^{\infty}(\Sigma)$.

II.3.1. Systems of equations

A tree is regular

– iff it has a finite number of distinct subtrees,

– iff it is component of the unique solution of a regular system of equations (*see* Courcelle [4]).

Let V be a set of syntactic variables (arity 0); a regular system (of equations) is a system of the form:

$$S = \langle x_1 = u_1, \ldots, x_n = u_n \rangle$$

with $x_1, \ldots, x_n$ in V, and with the u_i' s in $T(\Sigma \cup V)$ and not in V.

S can also be viewed as a deterministic regular tree grammar (one production rule for each syntactic variable). The solution of S is the n-uple of trees of $T^{\infty}(\Sigma):(t_1, \ldots, t_n)$ defined by:

$$t_i = \text{lub}\,\{h(t)/x_i \overset{*}{\underset{S}{\to}} t\}$$

where h substitutes the symbol Ω for the variables.

Let $\bar{S}$ be the derived system, associated with S, and defined by:

$$\bar{S} = \langle x_1 = \varphi(u_1), \ldots, x_n = \varphi(u_n) \rangle$$

$\bar{S}$ is an algebraic system of equations (of words) in which for each x_i, there is only one equation with left-hand side x_i.

Conversely, each algebraic system of words:

$$T = \langle x_1 = w_1, \ldots, x_n = w_n \rangle, \qquad x_i \in V,$$

$$w_i \in (\Sigma_0 \cup V)^*$$

is the derived system, $\bar{S}$, of some regular system of trees, S.

Let h be the morphism, $h : (\Sigma_0 \cup V)^* \to W(\Sigma_0)$ defined by:

$h(x_i) = \Omega$ for each x_i in V,

$h(\alpha) = \alpha$ for each letter α in Σ_0.

It can be noted that $h(u) \leqq_\Omega h(v)$ whenever $u \overset{*}{\underset{\bar{S}}{\to}} v$; so the set $L_i = \{h(w) / x_i \overset{*}{\underset{\bar{S}}{\to}} w\}$ is a directed subset of $W(\Sigma_0)$, and thus has a lub in $W^\infty(\Sigma_0)$).

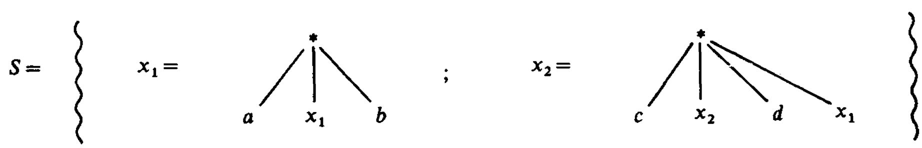
PROPERTY: *The n-uple* $(\mathrm{lub}(L_1), \ldots, \mathrm{lub}(L_n))$ *is the least solution of the system* $\bar{S}$ *in* $(W^\infty(\Sigma_0))^n$ *and is equal to* $(\varphi(t_1), \ldots, \varphi(t_n))$, *where* $(t_1, \ldots, t_n)$ *is the solution of* S.

Proof: A solution of $\bar{S}$ is a n-uple $(w_1, \ldots, w_n)$ of $(W^\infty(\Sigma_0))^n$ satisfying the equations, that is, such that for $i = 1, \ldots, n : w_i = \theta(u_i)$ with θ the morphism $\theta : W^\infty(\Sigma_0 \cup V) \to W^\infty(\Sigma_0)$ defined by:

- $\theta(\Omega) = \Omega$,
- $\theta(\alpha) = \alpha$ for all α in Σ_0,
- $\theta(x_i) = w_i$ for all x_i in V.

It is quite immediate that $\mathrm{lub}(L_i) = \varphi(t_i)$ and that $(\varphi(t_1), \ldots, \varphi(t_n))$ is a solution of $\bar{S}$. A simple induction on the length of derivation shows that $z \leqq_\Omega w_i$ whenever z is in L_i and $(w_1, \ldots, w_n)$ is a solution of $\bar{S}$, and thus $\mathrm{lub}(L_i) \leqq_\Omega w_i$ for $i = 1, \ldots, n$.

Example: Consider the following regular system:



Then

$$\bar{S} = \{x_1 = a\, x_1\, b;\ x_2 = c\, x_2\, d\, x_1\},$$

$$L_1 = \{a^n\, \Omega\, b^n / n \in N\},$$

$$L_2 = \bigcup_{p \geqq 0} c^p\, \Omega\, (d\, L_1)^p.$$

It is easy to see that L_1 is equivalent (same lub) to the rational language $a^* \Omega b^*$, and L_2 to $c^* \Omega (da^* \Omega b^*)^*$; thus the (least) solution of $\bar{S}$ is

$$(\text{lub}(a^* \Omega b^*), \text{lub}(c^* \Omega (da^* \Omega b^*)^*)).$$

In general, the solution of certain particular classes of algebraic systems (the linear and quasi-rational ones) can easily be expressed with rational languages; that is not so evident for arbitrary system.

Remark: Let us recall that an element of $W(\Sigma_0)$ is a class of congruence $(\Omega \simeq \Omega\Omega)$ of the free monoid $(\Sigma_0 \cup \{\Omega\})^*$. Given a language L of $(\Sigma_0 \cup \{\Omega\})^*$, one consider L, in $W(\Sigma_0)$, as the set of classes which are represented i. e. $\{w \in W(\Sigma_0)/\exists u \in L \ \& \ u \in w\}$.

Conversely, given a subset L of $W(\Sigma_0)$ (also called language) one considers if necessary, the corresponding language of $(\Sigma_0 \cup \{\Omega\})^*$ of the canonical representatives of the element of L.

THEOREM: *For every regular tree t, one can find a rational language $R \subseteq (\Sigma_0 \cup \{\Omega\})^*$ such that*

- *R (considered as a subset of $W(\Sigma_0)$) is directed and*
- *$\varphi(t) = \text{lub}(R)$.*

The proof of this theorem consists in solving the algebraic systems of equations and this can be done in two different ways:

- constructing an automaton associated with the system,
- using Heilbrunner's results [6].

These are described later.

Consequence: The equality of the yield of two regular trees is decidable.

Proof: Let t and t' be regular trees and R, R' be rational languages satisfying the previous theorem.

Then:

$$\varphi(t) = \varphi(t'),$$

$$\underline{\text{iff}}\ \text{lub}(R) = \text{lub}(R'),$$

$$\underline{\text{iff}}\ \text{Ideal}(R) = \text{Ideal}(R'),$$

where

$$\text{Ideal}(L) = \{w \in W(\Sigma_0)/\exists v \in L \ \& \ w \leqq_\Omega v\}$$

for all

$$L \subseteqq W^{\infty}(\Sigma_0).$$

Moreover, it is easy to check that the Ideal (more precisely, the set of canonical representatives of the Ideal) of a rational language is a rational language, an expression of which can be constructed.

This establishes the result since the equality of rational languages is decidable.

III.2. Solving algebraic systems: first way

Let S be an algebraic system:

$$S = \begin{cases} x_i = u_i; \text{ with } u_i \in (\Sigma_0 \cup V)^*; \ V = \{x_1, \ldots, x_n\}, \\ \qquad i = 1, \ldots, n; \ n \in N_+. \end{cases}$$

Let $\equiv$ denote the equivalence relation on the set of variables V, defined by:

$$x_i \equiv x_j \ \underline{\text{iff}} \begin{cases} - \ x_i \overset{*}{\underset{S}{\rightarrow}} w x_j w', & w, w' \in (\Sigma_0 \cup V)^* \\ \text{and} & \\ - \ x_j \overset{*}{\underset{S}{\rightarrow}} v x_i v', & v, v' \in (\Sigma_0 \cup V)^*. \end{cases}$$

This relation allows to make a partition of the system S into disjoint "closed" subsystems (that is: all variables appearing at the left part of the equations are equivalent).

The principle is then to solve, regardless of the orthers, each subsystem by considering it as a system. The general solution of the system (i. e. the rational languages) is then obtained by straightforward substitution of the partial ones.

Example:

$$V = \{x_1, x_2, x_3\}$$

$$S = \begin{cases} x_1 = a x_2 b x_1 c x_3 d \\ x_2 = u x_2 v x_3 w \\ x_3 = \alpha x_2 \beta x_3 \gamma \end{cases}$$

we have $x_2 \equiv x_3$ and $x_1 \not\equiv x_2$ and thus the subsystems:

$$S_1 = \{x_1 = a x_2 b x_1 c x_3 d\}$$

$$S_2 = \begin{cases} x_2 = u x_2 v x_3 w \\ x_3 = \alpha x_2 \beta x_3 \gamma \end{cases}$$

S_1 is solved by the rational

$$R_1 = (a x_2 b)^* \Omega (c x_3 d)^*$$

If R_2 and R_3 solve S_2, we just have to substitute R_2 for x_2 and R_3 for x_3 in the expression of R_1 to obtain the rational language corresponding to the first variable of the system S.

The solution of a "closed" subsystem can be obtained by constructing an automaton associated with it.

Let S be a system:

$$S = \begin{cases} x_i = u_i, \\ i = 1, \ldots, p, \end{cases}$$

with $x_i \equiv x_j$ for all i, j.

Let A be the automaton defined in the following way:

– *Alphabet:* $X \cup \{\Omega\}$,

– *Set of states:* $Q = \{q_1, \ldots, q_p\} \cup \{\bar{q}_1, \ldots, \bar{q}_p\}$
i. e. two states per variable.

– *Transitions:* of four different types:

$\forall i = 1, \ldots, p$ (a) $q_i \bullet \xrightarrow{\Omega} \bullet \bar{q}_i$

(b) $q_i \bullet \xrightarrow{\alpha_i} \bullet q_j$ if $u_i = \alpha_i x_j w$
$\alpha_i \in X^*$

(c) $\bar{q}_k \bullet \xrightarrow{\beta_i} \bullet \bar{q}_i$ if $u_i = v x_k \beta_i$
$\beta_i \in X^*$

(d) $\bar{q}_j \bullet \xrightarrow{\gamma_i} \bullet q_k$ if $u_i = w x_j \gamma_i x_k w'$
$\gamma_i \in X^*$

Let us note by R_i the language that is recognized by the automaton with q_i as initial state and $\bar{q}_i$ as unique terminal state, for $i=1, \ldots, p$.

Let h be the morphism of monoids:

$$h: \quad (X \cup V)^* \rightarrow (X \cup \{\Omega\})^*$$

defined by:

$$h(x_i)=\Omega, \qquad \forall i=1, \ldots, p,$$
$$h(\alpha)=\alpha, \qquad \forall \alpha \text{ in } X.$$

LEMME 1: *For any word w such that $x_i \underset{S}{\overset{*}{\rightarrow}} w$, the word $h(w)$ is in R_i.*

Proof: Immediate by definition of the automaton.

LEMME 2: *For any word w in R_i, there exists a word w' such that $x_i \underset{S}{\overset{*}{\rightarrow}} w'$ and $w \leqq_\Omega h(w')$.*

Sketch of the proof:

(i) for all sequences of transitions of the form:

$$q_{j_1} \bullet \xrightarrow{\alpha_1} \bullet\, q_{j_2} \bullet \xrightarrow{\alpha_2} \bullet \ldots \bullet\, q_{j_n} \bullet \xrightarrow{\alpha_n} \bullet\, q_k$$

one has with S:

$$x_{j_1} \underset{S}{\overset{n}{\rightarrow}} \alpha_1 \alpha_2 \ldots \alpha_n x_k w'$$

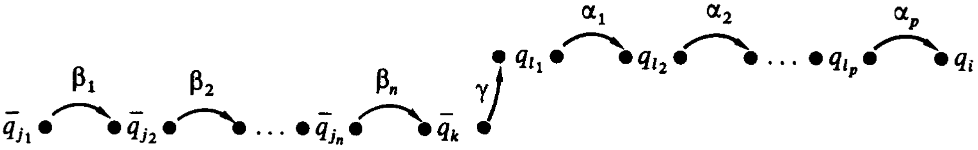
for some $w' \in (X \cup V)^*$.

(ii) symmetrically, if:

$$\bar{q}_{j_1} \bullet \xrightarrow{\beta_1} \bullet\, \bar{q}_{j_2} \bullet \xrightarrow{\beta_2} \bullet \ldots \bullet\, \bar{q}_{j_n} \bullet \xrightarrow{\beta_n} \bullet\, \bar{q}_k$$

then $k \underset{S}{\overset{n}{\rightarrow}} w' x_{k_1} \beta_1 \beta_2 \ldots \beta_n$, some $w' \in (X \cup V)$.

(iii) for all sequences of transitions of the form:



with S one has:

$$x \overset{n+p+1}{\underset{S}{-\,-\rightarrow}} w' x_{j_1} \beta_1 \beta_2 \ldots \beta_n \gamma \alpha_1 \alpha_2 \ldots \alpha_p x_i w''$$

for some x in V and w', w' in $(X \cup V)^*$

(i), (ii), (iii) are direct consequences of the definition of the automaton.

Let w be in R_i, then $w = w_0 \Omega w_1 \Omega \ldots w_{n-1} \Omega w_n$, $n \in N^+$.

With (i) and (ii) one easily gets:

$$x_i \overset{*}{\underset{S}{\rightarrow}} w_0 v w_n$$

with as least one variable in v. Since all variables are equivalent and by (iii) one successively obtains:

$$v = v_1 x_{j_1} v_2 \overset{*}{\underset{S}{\rightarrow}} v_1 v_3 w_1 v_4 v_2$$

$$v_4 = v_5 x_{j_2} v_6 \overset{*}{\underset{S}{\rightarrow}} v_5 v_7 w_2 v_8 v_6$$

$\vdots$

and so on,

and thus:

$$x_i \overset{*}{\underset{S}{\rightarrow}} w_0 z_1 w_1 z_2 w_2 \ldots z_{n-1} w_{n-1} z_n w_n = w'$$

with the $z_j's$ in $(X \cup V)^*$ and by definition of $\leqq_\Omega$ one has $w \leqq_\Omega h(w')$.

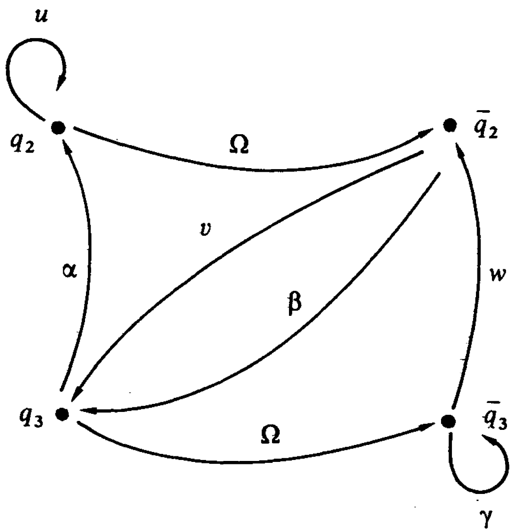
PROPERTY: *The p-uple* $(\mathrm{lub}(R_1), \ldots, \mathrm{lub}(R_p))$ *is the least solution of the system S [with R_i considered in $W(\Sigma)$].*

Proof: Direct consequence of Lemma 1 and Lemma 2.

Example: (continued)

$$S_2 = \begin{cases} x_2 = u x_2 v x_3 w \\ x_3 = \alpha x_2 \beta x_3 \gamma \end{cases}$$

the associated automaton is:



which gives the rational languages:

$$R_2 = u^* \Omega ((v + \beta)(\alpha u^* \Omega + \Omega \gamma^* w))^*$$

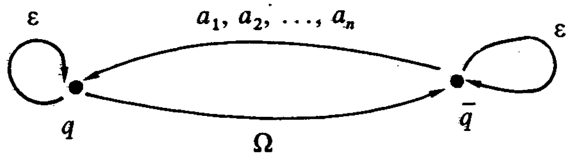
and

$$R_3 = ((\alpha u^* \Omega + \Omega \gamma^* w)(v + \beta))^* \Omega \gamma^*.$$

Other example:

$$S = \{x = x a_1 x a_2 x \ldots x a_n x\}$$

$A =$



$R = (\Omega (a_1 + a_2 + \ldots + a_n))^* \Omega.$

III.3. Solving algebraic systems: second way

Algebraic systems have been studied by Courcelle [2] and Heilbrunner [6] within the framework of countable arrangements and frontiers of infinite trees.

The considered solution of a system in a n-uple of arrangements, each component of which being equivalent (w.r.t. the equivalence of arrangements) to the frontier of the corresponding maximal derivation tree.

The solution is given by "regular expressions" that represent these arrangements.

A regular expression consists in an expression using concatenation, exponentiation to ω and $-\omega$, and shuffle. Without recalling the precise definitions of these operations, we just give characteristic properties.

– *Exponentiation:* If u is an arrangement:

- u^{ω} is the solution of the equation $x=ux$,
- $u^{-\omega}$ is the solution of the equation $x=xu$;

– *Shuffle:* Let $u_1, \ldots, u_n$ be arrangements, $n \in N_+$, and u be the set $u=\{u_1, \ldots, u_n\}$.

The shuffle of u, denoted be u^{η} is the solution of the equation

$$x=xu_1\,xu_2\,x \ldots xu_n\,x.$$

If $u^{\eta}=\langle D, \pi, f\rangle$ with $f: D \rightarrow u$ then one has:

for all x, y in D such that $x \neq y$ and $x\,\pi\,y$;

for all $u \in U$, there exists $z \in D$ such that $x\,\pi\,z$ and $z\,\pi\,y$ and $f(z)=u$.

With any regular expression E, in Heilbrunner's sense, one can associate a rational language of $(\Sigma_0 \cup \{\Omega\})^*$ defined by the rational expression $g(E)$, g being recursively defined as follows:

– $g(\varepsilon)=\varepsilon$,

– $g(a)=a$ for each letter a in Σ_0,

– $g(E.E')=g(E).g(E')$ whenever E and E' are regular expressions,

– $g(E^{\omega})=(g(E))^*\,\Omega$,

– $g(E^{-\omega})=\Omega(g(E))^*$,

– $g(E^{\eta})=(\Omega(g(E_1)+g(E_2)+\ldots+g(E_n)))^*\,\Omega$ whenever $E=\{E_1, \ldots, E_n\}$.

Let s be the mapping $s: \mathscr{A}_{\omega}(\Sigma_0) \rightarrow W^{\mathrm{loc}}(\Sigma_0)$ such that: for all locally finite trees t, $\varphi(t)=s(\psi(t))$.

PROPERTY: *Let E be a regular expression that denotes an arrangement u in* $\mathscr{A}_\omega(\Sigma_0)$; *then* $s(u)=\text{lub}(g(E))$.

Sketch of the proof:

– $u \equiv w$, $w \in X^*$ whenever u is finite and then $g(E)=x=s(u)$.

– Let us suppose $E=E_1 . E_2$ with E_1 and E_2 satisfying the property. Then $u=u_1 u_2$ and $s(u)=s(u_1).s(u_2)$ (by property of s); thus $s(u)=\text{lub}(g(E_1)).\text{lub}(g(E_2))$ and, by property of the C.P.O. $W^\infty(\Sigma_0)$, $s(u)=\text{lub}(g(E_1 E_2))=\text{lub}(g(E))$.

– Let α be a letter in Σ_0, and $E=\alpha^\omega$. E is the solution of the equation $x=\alpha x$. Considering the derivation tree t of that equation, one immediatly gets:

$$\varphi(t)=s(u)=\text{lub}(\alpha^* \Omega)=\text{lub}(g(E)).$$

In the same way, if $E=\alpha^{-\omega}$ the result is obtained by using the equation$x=x\alpha$.

– Let $\alpha_1, \ldots, \alpha_n$ be letters in Σ_0

$$\alpha=\{\alpha_1, \ldots, \alpha_n\} \qquad \text{and} \qquad E=\alpha^\eta.$$

E is the solution of the equation $x=x\alpha_1 x \ldots x\alpha_n x$. As before, t is the maximal derivation tree (t is locally finite); $s(u)=s(\psi(t))=\varphi(t)$ and, using the previous method for solving this equation, one gets $\varphi(t)=\text{lub}((\Omega(\alpha_1+\ldots+\alpha_n))^* \Omega)=\text{lub}(g(E))$.

– It remains to be proved that $s(w)=\text{lub}(g(\Theta(E)))$ where w is an arrangment denoted by $\Theta(E)$, whenever an arrangement u is denoted by E and satisfies $s(u)=\text{lub}(g(E))$, and Θ is a substitution of the expression E_α for the letter α such that $s(u_\alpha)=\text{lub}(g(E_\alpha))$. We only need to ramark that $g(\Theta(E))$ is equal to $\tau(g(E))$ with τ the substitution of the expression $g(E_\alpha)$ for the letter α.

Consequence: Let S be an algebraic system

$$S=\begin{cases} x_i=u_i, \\ i=1, \ldots, p \end{cases}$$

and $E_1, \ldots, E_p$ be the regular expressions given by Heilbrunner's algorithm to solve S.

Then the n-uple $(\text{lub}(g(E_1)), \ldots, \text{lub}(g(E_n)))$ is the least solution of the system S, whenever none of the expressions E_i is reduced to the empty word ε (i. e. the maximal derivation trees, from a variable, are locally finite).

Proof: Let $t_1, \ldots, t_p$ be the maximal derivation trees associated with S (obtained from the variables) then E_i represents $\psi(t_i)$ and $\varphi(t_i) = s(\psi(t_i)) = \text{lub}(g(E_i))$ whenever t_i is locally finite and $(\varphi(t_1), \ldots, \varphi(t_p))$ is the least solution of S.

Example: (The same as previously).

Let S be the system:

$$S = \begin{cases} x_1 = a x_2 b x_1 c x_3 d \\ x_2 = u x_2 v x_3 w \\ x_3 = \alpha x_2 \beta x_3 \gamma \end{cases}$$

Heilbrunner's algorithm gives the expressions:

$$E_2 = u^\omega . \{\gamma^{-\omega} wv \alpha u^\omega + \gamma^{-\omega} w \beta\alpha u^\omega\}^\eta . \gamma^{-\omega} w,$$

$$E_3 = \alpha u^\omega \{\gamma^{-\omega} wv \alpha u^\omega + \gamma^{-\omega} w \beta\alpha u^\omega\}^\eta . \gamma^{-\omega}$$

and then:

$$E_1 = (a E_2 b)^{-\omega} . (c E_3 d)^\omega.$$

With the transformation g, one gets:

$$-g(E_2) = u^* \Omega (\Omega (\Omega\gamma^* wv \alpha u^* \Omega + \Omega\gamma^* w \beta\alpha u^* \Omega))^* \Omega\gamma^* w$$

$$\simeq u^* \Omega (\gamma^* wv \alpha u^* \Omega + \gamma^* w \beta\alpha u^* \Omega)^* \Omega\gamma^* w,$$

$$-g(E_3) = \alpha u^* \Omega (\gamma^* wv \alpha u^* \Omega + \gamma^* \Omega + \gamma^* w \beta\alpha u^* \Omega)^* . \Omega\gamma^* w$$

and:

$$-g(E_1) = (ag(E_2) b)^* \Omega (cg(E_3) d^*.$$

It is easy to see that $g(E_1)$, $g(E_2)$ and $g(E_3)$ are, respectively, equivalent (same lub) to the rational languages R_1, R_2 and R_3 obtained by the first method.

REFERENCES

1. G. Comyn, *Objets infinis calculables*, Thèse d'État, Lille-I, mars 1982.
2. B. Courcelle, *Frontiers of Infinite Trees*, in R.A.I.R.O., Informatique Théorique, Vol. 12, 1978, pp. 319-337.
3. B. Courcelle, *Infinite Trees in Normal Form and Recursive Equations Having a Unique Sollution*, Math. Systems Theory, Vol. 13, 1978, pp. 131-180.
4. B. Courcelle, *Fundamental Properties of Infinite Trees*, in Theoretical Computer Science, Vol. 25, 1983, pp. 95-169.

5. M. DAUCHET and E. TIMMERMAN, *Continuous Monoids and Yields of Infinite Trees*, Technical report No. IT-74-85, Lille-I, juin 1985.
6. S. HEILBRUNNER, *An Algorithm for the Solution of Fixed Point Equations for Infinite Words*, in R.A.I.R.O., Informatique Théorique, Vol. 14, 1980, pp. 131-141.
7. G. HUET, *Confluent Reductions: Abstract Properties and Applications to Term Rewriting Systems*, J. A.C.M., Vol. 27, 1980, pp. 797-821.
8. MACNAUGHTON, *Testing and Generating Infinite Sequences by a Finite Automaton*, Information and Control, Vol. 9, 1966, pp. 521-530.
9. M. NIVAT, *Langages algébriques sur le magma libre et sémantique des schémas de programmes*, in Automata, Languages and Programming, 1st Colloquium, Springer-Verlag, 1973, pp. 293-307.
10. M. NIVAT, *Infinite Words, Infinite Trees, Infinite Computations*, in Mathematical Centre Tracts, Vol. 109, 1979, pp. 1-52.
11. M. NIVAT and D. PERRIN, "*Ensembles reconnaissables de mots bi-infinis*". Rapport LITP, 52, 1981.
12. W. THOMAS, *On Frontier of Regular Trees*, (to appear this journal).
13. E. TIMMERMAN, *Yields of Infinite Trees*, Proc. 9th CAAP, Bordeaux, 1984, in Cambridge University Press, pp. 199-311.
14. E. TIMMERMAN, *Feuillages d'arbres infinis*, Thèse de 3e Cycle, Lille-I, juin 1984.