INFORMATIQUE THÉORIQUE ET APPLICATIONS

F. BOSSUT
B. WARIN

On a code problem concerning planar acyclic graphs

Informatique théorique et applications, tome 25, nº 3 (1991), p. 205-218.

<http://www.numdam.org/item?id=ITA_1991__25_3_205_0>

© AFCET, 1991, tous droits réservés.

L'accès aux archives de la revue « Informatique théorique et applications » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/legal.php). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 25, n° 3, 1991, p. 205 à 218)

ON A CODE PROBLEM CONCERNING PLANAR ACYCLIC GRAPHS (*)

by F. Bossut (1) and B. Warin (2)

Communicated by A. Arnold

Abstract. – *We prove the unsolvability of a code problem in the case of connected planar dags.*

Résumé. – *Nous prouvons l'indécidabilité du problème du code dans les dags planaires connexes.*

INTRODUCTION

In a free algebraic structure, a finite subset C is a code if and only any objet cannot be obtained by two different compositions (modulo the axioms of the structure) of elements of C. The code problem consists in determining if a finite subset is a code.

So, in the case of words, a subset $C=\{w_1, w_2, \ldots, w_n\}$ is a code if and only if for every i and $j\in[n]$, $w_i . C^* \cap w_j . C^*=\varnothing$. As it is possible to determine if a rational set is empty, this equivalence proves that the code problem is solvable [12].

In the case of trees, let us consider the usual substitution on trees (*see* for example Arnold and Dauchet [1]). A tree t is said non linear if a variable occurs more than once. Then if we substitute a tree u to such a variable of t, we duplicate u. For instance, if $t=b(x, x)$, we obtain $b(u, u)$, also denoted by $t.u$. Then it is easy to formulate the code problem on trees in the same way. In the general (linear or not) case of trees, Dauchet [5] proves that this problem is unsolvable. In the linear case (*i.e.* when every tree of C is linear) the corresponding problem is solvable by using the decidability of the emptyness problem for the rational forests. This problem has been thoroughly studied by Nivat [10, 11].

(*) Received March 1988, revised Novembre 1988.

(1) Laboratoire d'Informatique Fondamentale de Lille, U.A.-C.N.R.S. 369, Université de Lille-Flandres-Artois, 59655 Villeneuve-d'Ascq Cedex, France.

We consider the "smallest generalization" of linear trees, *i.e.* the class of planar directed ordered acyclic graphs (**pdags**). In Bossut and Warin [4] it is shown that the code problem for pdags is reducible to the code problem for pairs of words, which is also unsolvable. Here we consider only connected pdags. We could have thought that the code problem for connected pdags was reducible to the code problem for (non linear) trees by equalities of the form of figure 1.

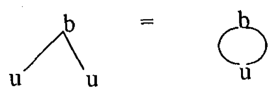


Figure 1

But it is not true for several reasons. The most intuitive one is that, in the left-hand tree of figure 1, each occurrence of u can have a distinct decomposition when, in the right-hand pdag, there is only one occurrence of u therefore only one decomposition of u.

This paper is organized as follows: in section 1, we define the algebraic frame in which we define our pdags. In section 2, we reduce the code problem to the emptiness problem for languages of words by means of derivations graphs of phrase-structured grammars [6, 9, 13].

1. PRELIMINARIES

We extend the notion of d-dags introduced by Kamimura and Slutzki [8].

1.1. DEFINITIONS OF pdags. — A *doubly-ranked alphabet* Σ is a finite set of letters on which are defined two mappings into N called *head-rank* and *tail-rank*.

For n, m integer, we denote by

Σ_m the set of letters of tail-rank m;

${}_n\Sigma$ the set of letters of head-rank n;

$${}_n\Sigma_m = {}_n\Sigma \cap \Sigma_m.$$

We interpret a letter of ${}_n\Sigma_m$ as a labelled node of a graph that has n ordered inputs and m ordered outputs. We define $M(\Sigma)$, the set of pdags over Σ, as $\bigcup_{m,\, n \in N} {}_nM(\Sigma)_m$ where the sets ${}_nM(\Sigma)_m$ of pdags over Σ with n inputs and m outputs are recursively defined by:

(i) if $a \in {}_n\Sigma_m$ then $a \in {}_nM(\Sigma)_m$;

(ii) if $\delta_i \in {}_{pi}M(\Sigma)_{qi}$ for $i=1$, 2 then the *parallel composition* of δ_1 and δ_2, denoted by $\delta_1 \theta \delta_2$, belongs to ${}_{p1+p2}M(\Sigma)_{q1+q2}$;

(iii) if $\delta_i \in {}_{pi}M(\Sigma)_{qi}$ for $i=1,2$ and $q1=p2$ then the *serial composition* of δ_1 and δ_2, denoted by $\delta_1 . \delta_2$, belongs to ${}_{p1}M(\Sigma)_{q2}$. The drawing of figure 2 represents the result of this composition.

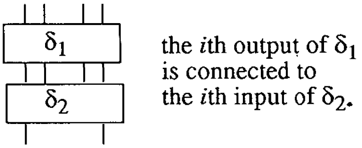


Figure 2

And these operators satisfy the following axiomatic equalities:

(iv) A particular pdag of ${}_1M(\Sigma)_1$, denoted by ε, is the unit element of the serial composition in the following sense:

For any integer p, let ε_p be $\varepsilon\theta\varepsilon\theta \dots \theta\varepsilon$ (p times), then for $\delta \in {}_pM(\Sigma)_q$, $\varepsilon_p . \delta = \delta . \varepsilon_q = \delta . \varepsilon_0$ denotes the absence of ε.

(v) The serial and parallel compositions are associative operators, and if $a.a'$ and $b.b'$ are defined then $(a.a')\theta(b.b')=(a\theta b).(a'\theta b')$.

Example. – For $a \in {}_1\Sigma_2$ and $b \in {}_4\Sigma_2$, the pdag obtained by the serial composition of the parallel composition of ε, a and ε with b is denoted by $(\varepsilon\theta a \theta \varepsilon).b$ and can be represented by the graph of figure 3.

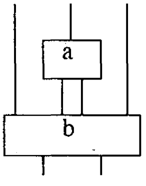


Figure 3

The underlying algebraic structure for $M(\Sigma)$ is the *free magmoid generated by* Σ. More details about this structure can be found in Arnold and Dauchet [1, 2], Schnorr [13], Hotz [7], Bossut and Warin [3].

Σ^* denotes the free monoid generated by the alphabet Σ and the operation θ.

We said that δ' is a *subgraph* of a pdag δ if for some pdags δ_1, δ_2, δ_3, δ_4 we have

$$\delta=\delta_1 . (\delta_2 \,\theta\, \delta' \,\theta\, \delta_3) . \delta_4$$

For δ' *subgraph* of a pdag δ, we shall say that:

δ' is a *proper subgraph* of δ if δ' is not equal to δ.

δ' is *an initial subgraph* of δ if δ_1 belongs to ε^*.

From the properties of the operators . and θ, it is easy to state that

$$\forall \delta \in M(\Sigma), \quad \exists \delta_1, \delta_2, \ldots, \delta_n \in (\Sigma \cup \varepsilon)^* \quad \text{such that } \delta=\delta_1 . \delta_2 \ldots \delta_n$$

and

$$\forall i,\ 1 \leqq i < n, \qquad \text{if} \quad \delta_i . \delta_{i+1}=\gamma' \,\theta\, a \,\theta \gamma'' \quad \text{with } a \in {}_p\Sigma_q \quad \text{then } \delta_i=\delta_i' \,\theta \varepsilon_p \,\theta\, \delta_i''$$

and

$$\delta_{i+1}=\delta_{i+1}' \,\theta\, a \,\theta\, \delta_{i+1}'' \qquad \text{where} \quad \gamma'=\delta_i' . \delta_{i+1}' \quad \text{and} \quad \gamma''=\delta_i'' . \delta_{i+1}''.$$

In such a decomposition of a pdag δ, δ_n is called the *yield* of δ. Intuitively, the yield of a pdag is, from left to right, the sequence of the labels of the nodes which have no successor.

1.2. DEFINITIONS. — Let A and B be two doubly-ranked alphabets, $M(A)$ and $M(B)$ the free magmoids generated by A and B. A mapping μ from A into $M(B)$ *respects the double rank* if and only if

$$\text{for any integers } p \text{ and } q, \quad \delta \in {}_pA_q \quad \Rightarrow \quad \mu(\delta) \in {}_pM(B)_q$$

An injective mapping μ from A into $M(B)$ that respects the double rank is a *coding mapping* if and only if its homomorphic extension to $M(A)$ is still injective. If μ is a coding mapping, we say that $\mu(A)$ is a *code*.

We say that $k \in M(A)$ is a *decomposition of* δ *over* A if $\mu(k)=\delta$.

A pair (k_1, k_2) of distinct decompositions of a pdag δ is said to be *irreducible* if there exists no k, k_1', k_2' elements of $M(A)$ such that

$$k_1=k . k_1' \qquad \text{and} \qquad k_2=k . k_2'$$

or

$$k_1=k_1' . k \qquad \text{and} \qquad k_2=k_2' . k.$$

From the above mentioned definitions, $\mu(A)$ is not a code if there exist k_1 and k_2, elements of $M(A)$, such that $\mu(k_1)=\mu(k_2)$. If such a pair exists, we show that an irreducible pair exists too.

PROPOSITION 1.2.1: *If a pdag δ admits a pair of distinct decompositions, then there exists a pdag δ' that admits an irreducible pair of decompositions.*

Proof. — Let (k_1, k_2) be a pair of distinct decompositions of δ, either (k_1, k_2) is irreducible or there exists a pair (k'_1, k'_2) of decompositions of a proper subgraph of δ and so on. As δ is a finite graph, δ has not an infinite number of proper subgraphs, then one of them admits an irreducible pair of decompositions.

So, we have

PROPOSITION 1.2.2: *$\mu(A)$ is not a code if and only if there exists a pdag of $M(B)$ that admits an irreducible pair of decompositions over A.*

2. CODE AND DERIVATION PDAGS

2.1. DEFINITION: A *phrase-structure grammar* is a system $G=\langle \Gamma, T, P\rangle$ where:

- Γ is a finite set of letters.
- $T \subset \Gamma$, is a set of terminal letters.
- The set P consists of expressions of the form $\alpha \to \beta$ with $\alpha, \beta \in \Gamma^+$, P is called the set of *production rules*. We do not consider here the productions of the form $\alpha \to \lambda$ where λ is the empty sentence.

We define, in a classical way, for $A \in \Gamma - T$, *the language generated by G from axiom A* (*see* Hopcroft and Ullman [6]) and we denote it by $L(G, A)$.

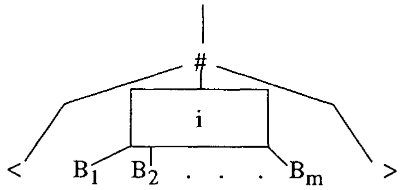
Let $G=\langle \Gamma, T, P\rangle$ be a phrase-structured grammar, $A \in \Gamma - T$ be an axiom and k be the number of rules of P. We associate with G and A two sets of pdags $CG1$ and $CG2$ of $M(\Sigma)$, where Σ is the doubly-ranked alphabet defined by:

- if $a \in \Gamma$ then $a \in {}_1\Sigma_1$; the set of such letters will be still denoted by Γ.
- if $a \in T$ then $a' \in {}_1\Sigma_1$; the set of such letters will be denoted T'.
- for i, n and $m \in N$, if $a_1 \ldots a_n \to b_1 \ldots b_m$ is the rule number i of P then $i \in {}_n\Sigma_m$;
- $\langle\ ,\ \rangle \in {}_1\Sigma_1$ and $\# \in {}_1\Sigma_3$ are three new symbols.

2.1.1. *Construction of $CG1$*

(*a*) If the rule number i of P is of the form $A \to b_1 \ldots b_m$, we define the set:

$CG\,1\,(i)' = \Bigg\{$

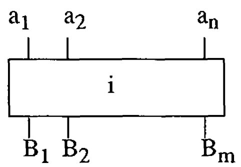


such that for j=1,...m
Bj ∈ { bj,b'j} if bj ∈ T, and
Bj= bj otherwise $\Bigg\}$

Otherwise $CG\,1\,(i)' = \varnothing$.

(*b*) If the rule number i of P is of the form $a_1 \ldots a_n \to b_1 \ldots b_m$, we define the set:

$CG\,1\,(i)'' = \Bigg\{$



such that for j=1,...m
Bj ∈ { bj,b'j} if bj ∈ T, and
Bj= bj otherwise $\Bigg\}$

Finally, let us set:

$$CG\,1' = \bigcup_{i=1}^{k} CG\,1\,(i)',$$

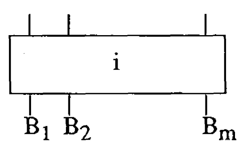
$$CG\,1'' = \bigcup_{i=1}^{k} CG\,1\,(i)'' \qquad \text{and} \qquad CG\,1 = CG\,1' \cup CG\,1''$$

where k is the number of rules of P.

2.1.2. *Construction of CG* 2

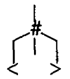
If the rule number i of P is of the form $a_1 \ldots a_n \to b_1 \ldots b_m$, we define the set:

$CG\,2\,(i) = \Bigg\{$



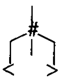
such that for j=1,...m
Bj ∈ { bj.bj, b'j} if bj ∈ T, and
Bj= bj.bj otherwise $\Bigg\}$

And

$$CG\,2=\left[\bigcup_{i=1}^{k} \mathrm{CG}\,2\,(\mathrm{i})\right]\cup\{$$  $$\}$$

The elements of $CG\,1$ and $CG\,2$ are pdags of $M(\Sigma)$ and can be considered as images under a mapping μ of a doubly-ranked-alphabet $\overline{CG}=\overline{CG\,1}\cup\overline{CG\,2}$.

Let us denote by α the element of

$\overline{CG\,2}$ such that $\mu(\alpha)=$ 

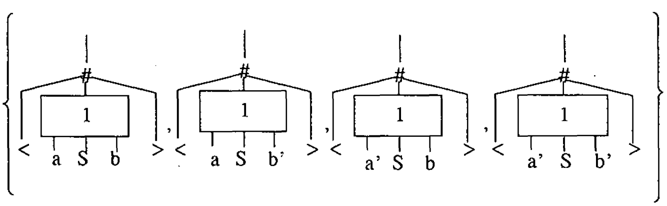
2.2. *Example.* – Let $G=\langle \Gamma, T, P\rangle$ be a grammar where

$$\Gamma=\{A, B, S, a, b, c\},\qquad T=\{a, b, c\}$$

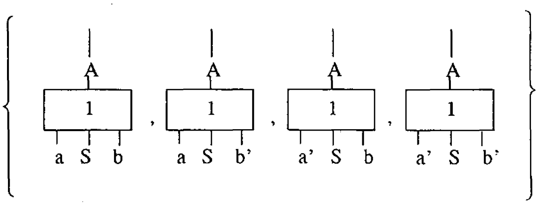
and P contains the following productions:

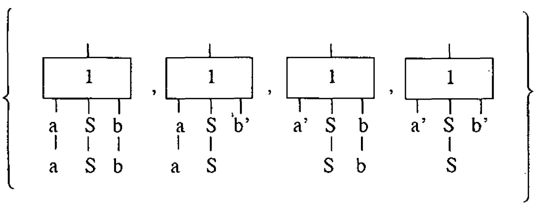
rule 1 : $A\to a\,S\,b$, *rule* 2 : $\to AB$, *rule* 3 : $a\,A\to c$, *rule* 4 : $B\,b\to c$.

Let A be the axiom then we have:

$CG\,1\,(1)'=$ 

$CG\,1\,(2)'=CG\,1\,(3)'=CG\,1\,(4)'=\varnothing$

$CG\,1\,(1)''=$ 

$CG\,2\,(1)=$ 

We let to the reader the construction of the other sets.

LEMMA 2.3: *$L(G, A) \neq \varnothing \Rightarrow CG$ is not a code.*

Sketch of proof: If $L(G, A) \neq \varnothing$ then there exists a word $m \in T^*$ such that m derives from A. Let δ be the pdag associated with this derivation. From δ, we construct a pdag γ of $M(\Sigma)$ as follows:

- we replace the root A by α;
- we replace the leaves by the corresponding primed letters;
- we replace each letter a of Γ which is neither root nor leave by $\begin{matrix} a \\ | \\ a \end{matrix}$;

Such a pdag will be said to be associated with this derivation. It can be decomposed over $\overline{CG1}$ and also over $\overline{CG2}$. Indeed, the pdag γ is a *juxtaposition* of rules of P so it can be decomposed over $\overline{CG1}$ and as the rules are correctly *linked*, it can be also decomposed over $\overline{CG2}$. Moreover γ is of the form of figure 4 because the first applied rule derives the axiom.

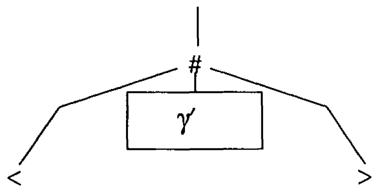


Figure 4

2.3.1. *Example* :Let us consider the grammar of example 2.2, and the following derivation

$$A \to a\,S\,b \to a\,A\,B\,b \to c\,B\,b \to cc$$

Figure 5 shows the derivation graph corresponding to this derivation, its associated pdag of $M(\Sigma)$ and its decompositions over $\overline{CG2}$ and $\overline{CG1}$.

Conversely, we prove that if CG is not a code then a pdag admits a decomposition over $\overline{CG1}$ and another over $\overline{CG2}$ and therefore $L(G, A)$ is not empty. This proof requires technical preliminary lemmas.

LEMMA 2.4: *Let δ be a pdag of $M(\Sigma)$ that admits an irreducible pair (k_1, k_2) of decomposition over $\overline{CG}$, then*

$$\exists p, q \in N, k'_1, k'_2 \in M(\overline{CG}), \qquad x \in \overline{CG1'}$$

such that

$$k_1 = (\varepsilon_p \theta x \theta \varepsilon_q) . k'_1$$
$$k_2 = (\varepsilon_p \theta \alpha \theta \varepsilon_q) . k'_2$$

Sketch of proof : Let us choose an arbitrary letter ρ in the first level of then there exist x, $y \in \overline{CG}$ such that ρ appears in the first level of $\mu(x)$ and $\mu(y)$. So, roughly speaking, $\mu(x)$ and $\mu(y)$ must be "superposable" in such a manner that they have at least this letter ρ in common. Now, if we examine the elements of CG, either $x = y$ or $x = \alpha$ and $y \in \overline{CG 1'}$. But x, y are different because (k_1, k_2) is an irreducible pair of decompositions.

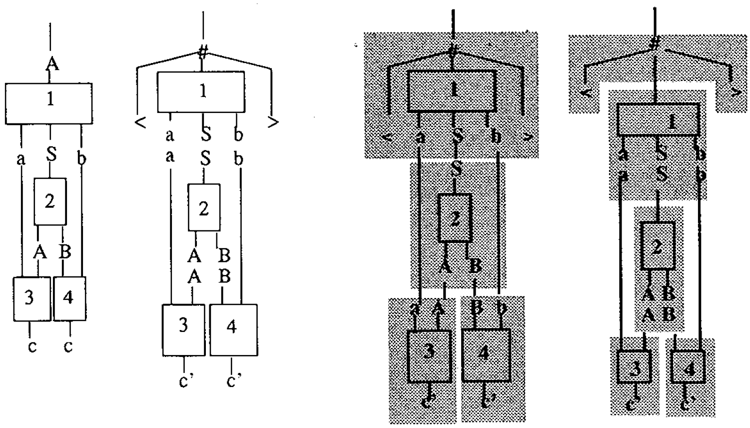
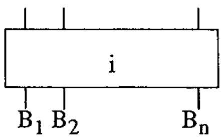


Figure 5



where i ∈ {1,...,k}
and for j ∈ {1,...,n}
$B_j \in \Gamma \cup T'$

Figure 6

NOTATION : Each element d of $CG 1$ can be written in an unique way as $\mu(\alpha) . (\varepsilon \theta d'' \theta \varepsilon)$ or as $d' . d''$ where $d' \in \Gamma^*$ and d'' is of the form of figure 6. Let $CG 0$ be the set of the such d'.

Then, with each d'' is associated an unique $\bar{d}''$ of $(\Gamma \cup \varepsilon)^*$ such that $d'' . \bar{d}'' \in CG 2$.

The next lemma proves that for any pdag that admits an irreducible pair of decompositions over $\overline{CG}$, its decompositions can be constructed by induction.

LEMMA 2.5: *For each irreducible pair* (k_1, k_2) *of decompositions over* $\overline{CG}$ *of a pdag* δ, *we can exhibit three sequences:* $(\delta_i)_{i\in N}$ *of elements of* $M(\overline{CG1})$; $(\delta'_i)_{i\in N}$ *of elements of* $M(\overline{CG2})$, $(\lambda_i)_{i\in N}$ *of elements of* $(\Gamma\cup\varepsilon)^*$ *such that:*

(1) $\delta_0\in\overline{CG1'}$, $\delta'_0\in\alpha.\overline{CG2}$ and $\mu(\delta_0).\lambda_0=\mu(\delta'_0)$; *and* $\forall i\in N$

(2) $\exists q_1, q_2, k'_1, k'_2\in M(\overline{CG})$: $k_1=(q_1\,\theta\,\delta_i\,\theta\,q_2).k'_1$ *and* $k_2=(q_1\,\theta\delta'_i\,\theta\,q_2).k'_2$

(3) $\mu(\delta_i).\lambda_i=\mu(\delta'_i)$

(4 *a*) *either* $\lambda_i\in\varepsilon^*$

(4 *b*) *or* $\exists\, d'\in CG0$; $d''\in(\Gamma\cup\varepsilon)^*$; $a\in\overline{CG1}$ *and* $b\in\overline{CG2}$; $n, m\in N$ *such that*

$$\mu(a)=d'.d'' \quad \textit{and} \quad \mu(b)=d''.\bar{d}''$$

$$\delta_{i+1}=\delta_i.(\varepsilon_n\,\theta\,a\,\theta\,\varepsilon_m)$$

$$\delta'_{i+1}=\delta'_i.(\varepsilon_n\,\theta\,b\,\theta\,\varepsilon_m)$$

$$\lambda_i=\lambda'_i\,\theta\,d'\,\theta\,\lambda''_i \qquad \textit{where} \quad \lambda'_i\in(\Gamma\cup\varepsilon)^n \quad \textit{and} \quad \lambda''_i\in(\Gamma\cup\varepsilon)^m$$

$$\lambda_{i+1}=\lambda'_i\,\theta\,\bar{d}''\,\theta\,\lambda''_i.$$

Proof.

Step 1 :(1), (2) *and* (3) *are true for* $i=0$.

From lemma 2.4, $\exists p, q\in N$, $k'_1, k'_2\in M(\overline{CG})$, $x\in\overline{CG1'}$ such that

$$k_1=(\varepsilon_p\,\theta\,x\,\theta\varepsilon_q).k'_1, \qquad k_2=(\varepsilon_p\,\theta\alpha\theta\varepsilon_q).k'_2 \qquad \text{and} \qquad \mu(x)=\alpha.(\varepsilon\theta\,d''\,\theta\varepsilon).$$

As $\mu(k1)=\mu(k2)$, d'' must be the beginning of an element of CG, now there exists an unique $b\in\overline{CG2}$ such $\mu(b)=d''.\bar{d}''$.

Therefore $\delta_0=x$, $\delta'_0=\alpha.(\varepsilon\theta\,b\,\theta\varepsilon)$ and $\lambda_0=\varepsilon\theta\,\bar{d}''\,\theta\varepsilon$.

Step 2 : $\forall i\in N$, (2) *and* (3) $\Rightarrow$ (4).

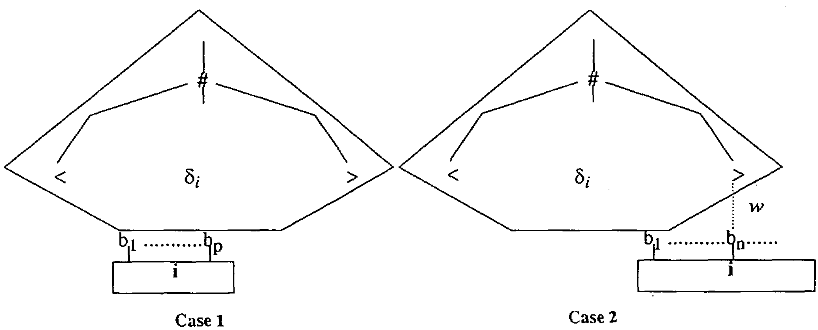


Figure 7

If $\lambda_i \notin \varepsilon^*$ then some letters that appear in λ_i belong to the first level of an element of $\mu(\overline{CG1''})$. On Figure 7, we have represented the two cases that might occur.

Assume that case 2 arises:

So, there exist a path from the node labelled $>$ to the node labelled **i**. Let w be the word composed of the labels of the nodes along the rightmost path from $>$ to **i**.

Let us denote by *PCG* the set of all the paths that go from the top to the bottom of element of *CG*. So *PCG* can be defined as follows:

$$PCG = \{ aib,\ ibb/a_1 \ldots a \ldots a_n \to b_1 \ldots b \ldots b_m \text{ rule number } i \text{ of } P \}$$
$$\cup \{ aib',\ ib'/a_1 \ldots a \ldots a_n \to b_1 \ldots b \ldots b_m \text{ rule number } i \text{ of } P \text{ and } b \in T \}$$
$$\cup \{ \#\, ib/A \to b_1 \ldots b \ldots b_m \text{ rule number } i \text{ of } P \}$$
$$\cup \{ \#\, ib'/A \to b_1 \ldots b \ldots b_m \text{ rule number } i \text{ of } P \text{ and } b \in T \} \cup \{ \#\langle,\ \#\rangle,\ \# \}.$$

As $\mu(\delta_i).\lambda_i = \mu(\delta'_i)$, then $w = w' b_n i$ where w' and $w' b_n$ can be decomposed over *PCG*. This implies that there exist $b_p\, jb_n$ or $\#\, jb_n$ member of *PCG* such that

(i) $w' = w''.(b_p j b_n)$ and $w' b_n = w'' b_p.(j b_n b_n)$ with w'' and $w''.b_p \in PCG^*$

or

(ii) $w' = w''.(\#\, jb_n)$ and $w' b_n = w''\,\#.(jb_n b_n)$ with w'' and w'' and $w''.\# \in PCG^*$.

In the first case (i), we will never reach the label $>$. In the second one (ii), it means that δ should have the form of figure 8.

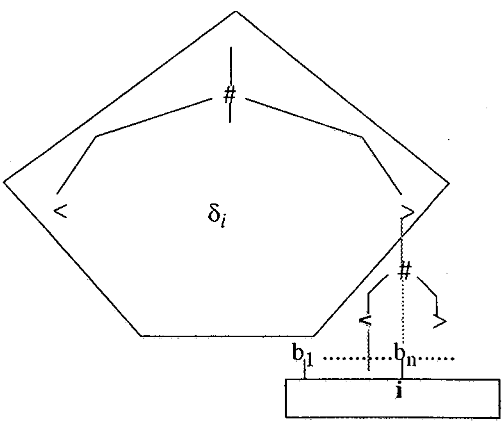


Figure 8

Thus we come back to an analogous situation in which appears a new path from a node labelled by $<$ to the node labelled by **i**. Therefore, this assumption leads to a contradiction.

So only case 1 can occur, and all the letters of the first level of this element of $\mu(\overline{CG\,1''})$ appear in λ_i, then

$$\exists\, d' \in CG\,0, \quad n, m \in N \quad \text{such that } \lambda_i = \lambda_i' \,\theta\, d' \,\theta\, \lambda_i''$$

where

$$\lambda_i' \in (\Gamma \cup \varepsilon)^n \qquad \text{and} \qquad \lambda_i'' \in (\Gamma \cup \varepsilon)^m.$$

So we can construct δ_{i+1}, δ'_{i+1} and λ_{i+1} as exposed in $(4\,b)$.

Step 3 : $\forall\, i \in N$, *if property* $(4\,b)$ *is true for i then properties* (2) *and* (3) *are true for* $i+1$.

Let a, b be such that $\mu(a) = d'.d''$ and $\mu(b) = d''.\bar{d}''$.

As $a \in \overline{CG\,1''}$ and appears in k_1, $\delta_{i+1} \in M(\overline{CG\,1})$ and $k_1 = (q_1 \,\theta\, \delta_{i+1} \,\theta\, q_2).k_1''$ for $k_1'' \in M(\overline{CG})$.

As $b \in \overline{CG\,2}$ and appears in k_2, $\delta'_{i+1} \in M(\overline{CG\,2})$ and $k_2 = (q_1 \,\theta \delta'_{i+1} \,\theta\, q_2).k_2''$ for $k_2'' \in M(\overline{CG})$.

As $\bar{d}'' \in \Gamma^*$, $\lambda_{i+1} \in (\Gamma \cup \varepsilon)^*$ and $\mu(\delta_{i+1}).\lambda_{i+1} = \mu(\delta'_{i+1})$.

LEMMA 2.6: *For each irreducible pair* (k_1, k_2) *of decompositions of a pdag* δ *over* $\overline{CG}$, *one belongs to* $M(\overline{CG\,1})$, *the other to* $M(\overline{CG\,2})$ *and* $\delta = \#.(\langle \theta\,\delta'\,\theta \rangle)$ *for* δ' *pdag.*

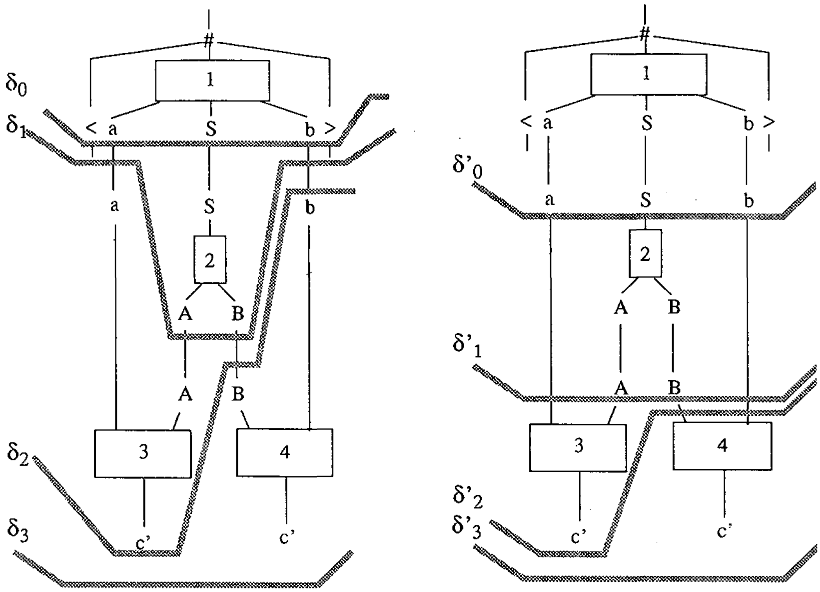
Proof : As k_1, k_2 end, we deduce from lemma 2.5 that there exists j such that $\lambda_j \in \varepsilon^*$. So $\mu(\delta_j) = \mu(\delta_j')$ and (δ_j, δ_j') is a pair of decompositions of a subgraph of δ. From (2) of lemma 2.5, δ_j, δ_j' are respectively initial subgraphs of k_1 and k_2. As (k_1, k_2) is irreducible, we conclude that $(\delta_j, \delta_j') = (k_1, k_2)$ and $\mu(\delta_j) = \mu(\delta_j') = \delta$.

From $(4\,b)$ of lemma 2.5, for $i \geqq 0$, δ_i' is an initial subgraph of δ'_{i+1}.

So δ_0' is an initial subgraph of $\delta_j'(\mu(\delta_j') = \delta)$. From (1) of lemma 2.5, we can conclude that $\delta = \#.(\langle \theta\,\delta'\,\theta \rangle)$ for δ' pdag.

Figure 9 presents the sequences $(\delta_i)_{i \in N}$, $(\delta_i')_{i \in N}$, $(\lambda_i)_{i \in N}$ associated with the pdag of example 2.3.1.

LEMMA 2.6 bis: *If* $\delta = \#.(\langle \theta\,\delta'\,\theta \rangle)$ *has two decompositions, the first over* $\overline{CG\,1}$ *and the second over* $\overline{CG\,2}$, δ *is associated with a terminal derivation in G from A.*



$\delta'_0 = \delta_0 . \lambda_0$ where $\lambda_0 = (\varepsilon\ \theta\ a\ \theta\ S\ \theta\ b\ \theta\ \varepsilon)$

$\delta'_1 = \delta_1 . \lambda_1$ where $\lambda_1 = (\varepsilon\ \theta\ a\ \theta\ A\ \theta\ B\ \theta\ b\ \theta\ \varepsilon)$

$\delta'_2 = \delta_2 . \lambda_2$ where $\lambda_2 = (\varepsilon\ \theta\ \varepsilon\ \theta\ B\ \theta\ b\ \theta\ \varepsilon)$

$\delta'_3 = \delta_3 . \lambda_3$ where $\lambda_3 = (\varepsilon\ \theta\ \varepsilon\ \theta\ \varepsilon\ \theta\ \varepsilon)$

Figure 9

Proof : Let us denote by $\varphi(\delta)$ the yield of δ. Let j be the integer such that $\lambda_j \in \varepsilon^*$.

From lemma 2.5, it is easy to state that, for $i \geqq 0$ and $i < j$:

$$\varphi(\delta_i) = \varphi(\delta'_i) = \langle m_i \rangle \qquad \text{with} \quad m_i = m'_i . d' . m''_i.$$

$$\varphi(\delta_{i+1}) = \varphi(\delta'_{i+1}) = \langle m_{i+1} \rangle \qquad \text{with} \quad m_{i+1} = m'_i . d'' . m''_i$$

such that $\mu(a) = d' . d''$ and $\mu(b) = d'' . \overline{d}'$ for $a \in \overline{CG\,1}$ and $b \in \overline{CG\,2}$.

Let h be the morphism from $(\Gamma \cup T')^*$ into Γ^* such that, if $a' \in T'$ then $h(a') = a$ and otherwise, $h(x) = x$.

So, for $i \geqq 0$ and $i < j$, $h(m_i) \to h(m_{i+1})$ in G. And as $\delta_0 \in CG\,1'$, the rule $A \to h(m_0)$ belongs to P. Moreover, $h(m_j) \in T^*$ because if a letter, that do not belong to T', appears in m_j, it will be in λ_j too.

Finally, δ is associated with the following terminal derivation in G

$$A \to h(m_0) \to \ldots \to h(m_i) \to h(m_{i+1}) \to \ldots \to h(m_j).$$

So we can state the last lemma.

LEMMA 2.7: *CG is not a code* $\Rightarrow L(G, A) \neq \emptyset$.

Proof: From proposition 1.2.2 and lemmas 2.6, 2.6*bis*, if CG is not a code, there exists a pdag associated with a terminal derivation in $L(G, A)$ hence $L(G, A) \neq \emptyset$.

Lemma 2.3 and 2.7 reduce our code problem to the emptiness problem for languages of words generated by phrase-structured grammars which is unsolvable. So we have:

THEOREM 2.8: *The code problem for finite sets of connected pdags is unsolvable.*

ACKNOWLEDGMENTS

We would like to thank A. Arnold for his very helpful comments and suggestions.

REFERENCES

1. A. ARNOLD and M. DAUCHET, Théorie des magmoïdes, *RAIRO Inform. Théor. Appl.*, 1978, *12*, 3, pp. 235-257.
2. A. ARNOLD and M. DAUCHET, Théorie des magmoïdes, *RAIRO Inform. Théor. Appl.*, 1979, *13*, 2, pp. 135-154.
3. F. BOSSUT and B. WARIN, Rationalité and reconnaissabilité dans des graphes acycliques, *Ph.D.*, University of Lille-I, 1986.
4. F. BOSSUT and B. WARIN, Problème de décision sur les dags, publication interne IT 107 du L.I.F.L., Université de Lille-I, Villeneuve-d'Ascq, France, 1987.
5. M. DAUCHET, It is undecidable whether a finite set of trees is a code, publication interne IT 109 du L.I.F.L., Université de Lille-I, Villeneuve-d'Ascq, France, 1987.
6. J. E. HOPCROFT and J. D. ULLMAN, Formal languages and their relation to automata, *Addison Wesley*, series in computer science an information processing, 1969.
7. G. HOTZ, Eine Algegraisierung des Syntheseproblems von Schaltkreisen, 1965, EIK 1, pp. 185-205, 209-231.
8. T. KAMIMURA and G. SLUTZKI, Parallel and Two-Way Automata on Directed Ordered Acyclic Graphs, *Inform. and Control*, 1981, *49*, pp. 10-51.
9. LOECK, The Parsing for General Phrase-Structure Grammars, *Inform. and Control*, 1970, *16*. pp. 443-464.
10. M. NIVAT, Code d'arbres, presented to Journées Hispano-Françaises d'Informatique Théorique et Programmation coordonnées avec LANFOR 86, Barcelone, nov. 1986.
11. M. NIVAT, Tree codes, Tapsoft 87, Pise, 1987.
12. C. W. PATTERSON and A. A. SARDINAS, A Necessary and sufficient condition for the unique decomposition of coded messages, *IRE intern. conv. Record*, 1953, *8*, pp. 104-108.
13. C. P. SCHNORR, Transformational classes of grammar, *Inform. and Control*, 1969, *14*, pp. 252-277.