

GHEORGHE PĂUN

**Some consequences of a result of Ehrenfeucht
and Rozenberg**

RAIRO. Informatique théorique, tome 14, n° 1 (1980), p. 119-122

http://www.numdam.org/item?id=ITA_1980__14_1_119_0

© AFCET, 1980, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Informatique théorique » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). 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/>

SOME CONSEQUENCES OF A RESULT OF EHRENFEUCHT AND ROZENBERG (*)

by Gheorghe PĂUN

Communiqué par J. BERSTEL

Abstract. — Recently, Ehrenfeucht and Rozenberg have proved that there are context-free languages that are not EDTOL languages. In the present note, some consequences of this result are pointed out with reference to certain open problems related to matrix languages.

Résumé. — Récemment, Ehrenfeucht et Rozenberg ont prouvé qu'il existe des langages indépendants du contexte qui ne sont pas des EDTOL-langages. Dans la présente note, nous indiquons quelques conséquences de ce résultat, en résolvant certains problèmes ouverts sur les langages matriciels.

1. NOTATIONS

In order to save space we shall omit the definitions and we shall specify only the notations we use:

— \mathcal{L} (EDTOL) is the family of languages generated by EDTOL systems (see [6]);

— \mathcal{M}_f is the family of languages generated by matrix grammars of finite index [1];

— \mathcal{SM} is the family of simple matrix languages [7] and \mathcal{SM}_f is the family of finite index languages in \mathcal{SM} ;

— \mathcal{L}_2 is the family of context-free languages;

— $\text{Ind}_m(L)$, $\text{Ind}_{cf}(L)$ denote the index of a language L according to matrix grammars, respectively, to context-free grammars;

— D_i is the Dyck language over the vocabulary $\{a_1, \dots, a_i, a'_1, \dots, a'_i\}$, that is the language generated by the context-free grammar with the rules $S \rightarrow SS$, $S \rightarrow a_j S a'_j$, $j = 1, 2, \dots, i$, $S \rightarrow \lambda$ [16].

(*) Reçu en mai 1979, et dans sa version définitive en juillet 1979.

(¹) University of Bucharest, Division of Systems Studies, Bucuresti, Romania.

2. OPEN PROBLEMS RELATED TO THE FAMILIES $\mathcal{M}_f, \mathcal{SM}$

The index of a matrix grammar was defined in [1] where it is proved that for any infinite language L in \mathcal{M}_f , the set of lengths of strings in L contains an infinite arithmetical progression (a similar result holds for languages in \mathcal{L}_2). In [1] it is asked (problem A) whether or not any context-free language has a finite index (clearly, according to matrix grammars).

The problem was partially solved by Salomaa [15] by proving that $\text{Ind}_{cf}(D_1)$ is infinite. The existence of a context-free language of infinite index according to matrix grammars remained open.

In [3], Cremers, Mayer, Weiss claim that $\text{Ind}_m(D_1)=2$, but, as in [8] it was pointed out, the proof in [3] is wrong; the finiteness of $\text{Ind}_m(D_1)$ remained open. Note that if any Dyck language would be in \mathcal{M}_f , then $\mathcal{L}_2 \subset \mathcal{M}_f$ since the family \mathcal{M}_f is a full-AFL [11] and any language L in \mathcal{L}_2 can be written as $h(D_i \cap R)$, where h is a homomorphism and R is a regular language [16].

In [7] Ibarra introduces the so-called simple matrix grammars, a class of grammars generating languages with many context-free-like properties. Similar properties (closure properties, decision results, pumping lemmas) were proved for the family \mathcal{M}_f too (see [11, 12, 14]). The relation between the families \mathcal{M}_f and \mathcal{SM} was formulated as an open problem in various contexts [9, 11, 12]; in fact, we obviously have $\mathcal{M}_f - \mathcal{SM} = \emptyset$ since the language $\{a^n b^n c^n \mid n \geq 1\}^*$ is in \mathcal{M}_f but not in \mathcal{SM} [7]. Therefore, the problem asks for a language in \mathcal{SM} which has an infinite matrix index. A stronger formulation [12] asks for context-free languages which are not in \mathcal{M}_f (problem related to the paper [3]).

In [9] we proved that $\mathcal{SM}_f \subset \mathcal{M}_f$ and one asks whether or not $\mathcal{SM}_f \subseteq \mathcal{SM}$ is a proper inclusion. The problem was solved in [10] by proving that D_1 does not belong to \mathcal{SM}_f . In a similar way it follows that any $D_i, i \geq 1$, is not in \mathcal{SM}_f .

3. THE SETTLEMENT OF THE ABOVE OPEN PROBLEMS

The result in [4] can be easily used to solve all the above open problems. Indeed, we have:

THEOREM: $\text{Ind}_m(D_i)$ is infinite for any $i \geq 2$.

Proof: In [2] one introduces the “parallel matrix grammars” as usual context-free matrix grammars with parallel derivation (each rule rewrites all occurrences of its left-hand side in the string to be rewritten). Let us denote by \mathcal{PM} the family of languages generated by these grammars. In [2] it is proved that $\mathcal{M}_f \subset \mathcal{PM}$, strict inclusion.

In [13] it is proved that, in fact, we have $\mathcal{PM} = \mathcal{L}$ (EDTOL), therefore $\mathcal{M}_f \subset \mathcal{L}$ (EDTOL). Following [4], there are context-free languages which are not in \mathcal{L} (EDTOL). Thus there are context-free languages which do not belong to \mathcal{M}_f as well. Consequently, any context-free generator (any context-free language whose smallest AFL containing it equals the family of context-free languages) is not a matrix language of finite index. Any language D_i , $i \geq 2$, is a context-free generator (see example 5.1.1, p. 139 [5]), hence $\text{Ind}_m(D_i)$ is infinite for any $i \geq 2$.

Consequently, (1) \mathcal{L}_2 and \mathcal{M}_f are incomparable (clearly, \mathcal{M}_f contains non-context-free languages), (2) \mathcal{PM} and \mathcal{M}_f are incomparable, (3) $\text{Ind}_m(D_i) = \text{Ind}_{cf}(D_i) = \infty$, $i \geq 2$, and thus all the problems listed in section 2 are settled.

It remains open the problem to find context-free languages L such that $\text{Ind}_m(L) < \text{Ind}_{cf}(L)$ [or even $\text{Ind}_m(L) < \text{Ind}_{cf}(L) = \infty$]. Particularly, we do not know whether $\text{Ind}_m(D_1)$ is finite or not.

Some problems settled by the result in [4] were considered in [4] too. Perhaps there are other problems which can be solved using the same result.

Note: We are very indebted to the referee for pointing out a mistake in an earlier version of this paper.

REFERENCES

1. B. BRAINERD, *An Analog of a Theorem About Context-Free Languages*, Inform. Control, Vol. 11, 1968, pp. 561-568.
2. S. CIOBOTARU, *The Parallel Matrix Grammars and the Thue Languages*, Bull. Math. Soc. Sc. Math. R.S.R., Vol. 27, (70), 1978, pp. 269-278.
3. A. B. CREMERS, O. MAYÈR and K. WEISS, *On the Complexity of Regulated Context-Free Rewriting*, Intern. Bericht, No. 4, Univ. Karlsruhe, Fakultät für Informatik, 1973.
4. A. EHRENFEUCHT and G. ROZENBERG, *On Some Context-Free Languages that Are Not Deterministic ETOL Languages*, R.A.I.R.O./Theoretical Computer Sc., Vol. 11, No. 4, 1977, pp. 273-291.
5. S. GINSBURG, *Algebraic and Automata-Theoretic Properties of Formal Languages*, North-Holl. Publ. Comp., Amsterdam, Oxford, 1975.
6. G. HERMAN and G. ROZENBERG, *Developmental Systems and Languages*, North-Holl. Publ. Comp., Amsterdam, 1975.
7. O. IBARRA, *Simple Matrix Languages*, Inform. Control, Vol. 17, 1970, pp. 359-394.
8. Gh. PĂUN, *On the Index of Grammars and Languages*, Inform. Control, Vol. 35, 1977, pp. 259-266.
9. Gh. PĂUN, *On the Generative Capacity of Simple Matrix Grammars of Finite Index*, Inform. Processing Letters, Vol. 7, 1978, pp. 100-102.

10. Gh. PĂUN, *On the Generative Capacity of Some Classes of Grammars with Regulated Rewriting*, Math. Found. of Computer Sc. Symp., Olomouc, Czechoslovakia, 1979.
11. Gh. PĂUN, *On the Family of Finite Index Matrix Languages*, J. Computer and Syst. Sc., Vol. 18, 1979, pp. 267-280.
12. Gh. PĂUN, *Matrix Grammars*, Editura Stiintifică si Enciclopedică, Bucuresti, 1980 (in Romanian).
13. Gh. PĂUN, *On the Generative Capacity of Parallel Matrix Grammars*, Bull. Math. Soc. Sc. Math. R.S.R., Vol. 23, (71), 1979, pp. 367-371.
14. Gh. PĂUN, *Some Context-Free-Like Properties of Matrix Languages of Finite Index*, Bull. Math. Soc. Sc. Math. R.S.R. (in press).
15. A. SALOMAA, *On the Index of Context-Free Languages*, Inform. Control, Vol. 14, 1969, pp. 474-477.
16. A. SALOMAA, *Formal Languages*, Academic Press, New York and London, 1973.