RAIRO. INFORMATIQUE THÉORIQUE

L. BOASSON
S. HORVÁTH

On languages satisfying Ogden's lemma

RAIRO. Informatique théorique, tome 12, nº 3 (1978), p. 201-202
<http://www.numdam.org/item?id=ITA_1978__12_3_201_0>

R.A.I.R.O. Informatique théorique/Theoretical Computer Science
(vol. 12, n° 3, 1978, p. 201 à 202)

ON LANGUAGES SATISFYING OGDEN'S LEMMA (*)

by L. Boasson ([1]) and S. Horváth ([2])

Communicated by M. Nivat

Abstract. — *We show that various types of non-context-free languages satisfying Ogden's lemma can very easily be constructed. A simple answer is then given to a question of* [6] *and* [7].

Define an "Ogden-like" language as a language L satisfying Ogden's lemma [3, 4]:

Definition: A language L is *Ogden-like* if there exists an integer k_0 such that if in any word f of L any k_0 or more positions (= occurrences of letters) are marked, f has a factorization $f = aubvc$ satisfying:

(1) $au^n bv^n c \in L, \quad \forall n \geqq 0$;

(2) either each of a, u and b, or each of b, v and c contains a marked position;

(3) ubv contains at most k_0 such positions.

Ogden's lemma, which is stronger than the classical Bar-Hillel's lemma [2, 4, 5], ensures that any context-free language is Ogden-like. The aim of this short note is to prove, as an answer to the third question of [6], the following:

Proposition: *There exists properly context-sensitive, properly recursive, properly recursively enumerable and non-recursively enumerable languages which are Ogden-like.*

By properly context-sensitive, we mean non context-free context sensitive, and analogously for properly recursive and properly recursively enumerable.

Proof: The proof is very simple. Consider first a subset P of $\mathbf{N}$, and define:

$$A_P = \{ (ab)^n \mid n \in P \},$$

$$B_P = A_P \cup X^* \{ aa, bb \} X^*,$$

(*) Received March 1978.

([1]) Université de Picardie et L. A. « Informatique théorique et Programmation ».
([2]) Eötvös Loránd University, Dpt. of Computer Mathematics, Budapest, Hongrie.

both being languages over $X=\{a, b\}$.

Obviously, B_P is context-free if A_P is. Moreover, B_P is properly context-sensitive (resp. recursive, recursively enumerable) if A_P is and B_P is not recursively enumerable if A_P is not.

It is obvious too that B_P is Ogden-like with $k_0=4$.

The proposition is then proved by chosing A_P properly context-sensitive (resp. recursive, resp. recursively enumerable) or not recursively enumerable. (It is well known that such languages A_P do exist, *see* [4] or [5] for instance.) $\equiv$

REMARK 1: The family of Ogden-like languages is closed under union, product, star and homomorphism. It is not closed under inverse homomorphism, and under intersection with regular sets.

REMARK 2: The family of Ogden-like languages (belonging as usual to some countably infinite alphabet) has cardinality C (=continuum). The proof is exactly the same as the proof of the corollary of theorem 3 in [6].

REMARK 3: The languages B_P are built exactly on the same principle as the non-regular language of [1] satisfying the pumping lemma of regular sets. Moreover, B_P does satisfy this lemma. It can even be observed that B_P is never properly context-free. However, it is easily shown that the proof of the proposition could be done with

$$A'_P=\{a^n b^n \mid n \in P\},$$
$$B'_P=\{a^n b^m \mid n \neq m\} \cup A'_P,$$

which gives other examples of Ogden-like languages in the various classical families.

REFERENCES

1. J. M. AUTEBERT, L. BOASSON and G. COUSINEAU, *A Note on 1-Locally Linear Languages*, Information and Control, Vol. 37, No. 1, 1978, p. 1-4.
2. Y. BAR-HILLEL, M. PERLES and E. SHAMIR, *On Formal Properties of Simple Phrase Structure Grammars*, Zeitschr. Phonetik., Sprachwiss., Vol. 14, 1961, p. 143-172.
3. W. OGDEN, *A Helpful Result for Proving Inherent Ambiguity*, Math. System Theory, Vol. 2, No. 3, 1968, p. 191-194.
4. A. V. AHO and J. D. ULLMAN, *The Theory of Parsing, Translation and Compiling*, Vol. I, Parsing, Prentice-Hall, 1971.
5. A. SALOMAA, *Formal Languages*, Academic Press, New York, London, 1973.
6. S. HORVÁTH, *The Family of Languages Satisfying Bar-Hillel's Lemma*, this issue of the R.A.I.R.O., Informatique théorique.
7. D. WISE, *A Strong Pumping Lemma for Context-Free Languages*, Theoretical Computer Science, Vol. 3, 1976, p. 359-369.