

SCATTERED SUBWORD COMPLEXITY OF NON-PRIMITIVE WORDS

SZILÁRD ZSOLT FAZEKAS

*Research Group in Mathematical Linguistics, Universidad Rovira i Virgili
Pl. Imperial Tàrraco 1, Tarragona, Spain
e-mail: szilard.zsolt@estudiants.urv.cat*

and

BENEDEK NAGY

*Research Group in Mathematical Linguistics, Universidad Rovira i Virgili
Pl. Imperial Tàrraco 1, Tarragona, Spain, and
Faculty of Informatics, University of Debrecen
Egyetem tér 1, Debrecen, Hungary
e-mail: nbenedek@inf.unideb.hu*

ABSTRACT

In this paper we analyze primitive words from the point of view of their scattered subwords. The language of primitive words has been the subject of numerous studies. It is the language of the words that are not proper powers of another word. First we take a look at the Parikh-vectors of these words, that is, we consider the commutative closure of languages formed by primitive words. After first looking at the shortest (one letter) subwords we move on towards giving necessary conditions for a word to be non-primitive in terms of scattered subword multiplicity. Furthermore, we prove that knowing the multiplicity of every word in a fixed set of words as scattered subwords in some word w is not enough to decide whether w is primitive or not.

Keywords: Scattered, subword, complexity, primitive, inference

1. Introduction

A widely investigated topic in combinatorics on words and formal language theory is that of the primitive words. From a formal language point of view the most studied question about them is whether the language of all primitive words over a given alphabet is a context-free language. In [4] the authors conjectured that it is not context-free but this problem has been open for more than ten years. In several papers and books about combinatorial properties of words (e.g. [8]) there is a great deal of (combinatorial) results concerning primitive words. We focus on this direction. In particular we give some criteria based on the number of scattered subword occurrences which are sufficient but not necessary for a word to be primitive. In this framework