Journal of Automata, Languages and Combinatorics **21** (2016) 4, 339 – 349

DELETING POWERS IN WORDS

John Machacek

Department of Mathematics, Michigan State University
619 Red Cedar Road, East Lansing, MI 48824
machace5@math.msu.edu

ABSTRACT

We consider the language consisting of all words such that it is possible to obtain the empty word by iteratively deleting powers. It turns out that in the case of deleting squares in binary words this language is regular, and in the case of deleting squares in words over a larger alphabet the language is not regular. However, for deleting squares over any alphabet we find that this language can be generated by a linear indexed grammar which is a mildly context-sensitive grammar formalism. In the general case we show that this language is generated by an indexed grammar.

Keywords: formal languages, indexed grammars, combinatorics on words

1. Introduction

Let Σ be a finite alphabet. For a word $w \in \Sigma^*$ and $a \in \Sigma$ we let $|w|$ denote the length of w and $|w|_a$ denote the number of occurrences of the letter a. For an integer $p > 0$, a *pth-power* is a p-fold repetition $u^p = uu \cdots u$ of non-empty word $u \in \Sigma^*$. As an example, $(ab)^3 = ababab$ is a 3rd-power. Given a word $w = a_1 a_2 \cdots a_n \in \Sigma^*$, we say w contains the word $u \in \Sigma^*$ if $u = a_i a_{i+1} \cdots a_j$ for some $1 \leq i \leq j \leq n$. A word is called *pth-power-free* if it contains no pth-powers. For $p = 2$ and $p = 3$, we will refer to pth-powers as *squares* and *cubes* respectively. We let Σ_k denote an alphabet of size k, and we typically consider $\Sigma_k \subsetneq \Sigma_{k+1}$. Also, for $k = 2$ and $k = 3$ we call elements of Σ_k^* *binary* and *ternary* words respectively where $\Sigma_2 = \{a, b\}$ and $\Sigma_3 = \{a, b, c\}$. We denote the class of regular languages, context-free languages, and context-sensitive languages by **REG**, **CFL**, and **CSL** respectively. These classes of languages are standard, and we will assume the reader has familiarity with them. We denote the class of indexed languages and linear indexed languages by **IL** and **LIL** respectively. Definitions of the grammars which generate indexed languages and linear indexed languages will be given in Section 3.

Given a word $w \in \Sigma^*$ and an integer $p > 0$ we consider the possible outcomes of iteratively deleting pth-powers from w until we have a pth-power-free word. In particular, we are interested in when we can obtain the empty word ϵ. Let us now consider an example.