Journal of Automata, Languages and Combinatorics **13** (2008) 2, 145–156

THRESHOLD PROPERTIES OF SOME PERIODIC FACTORS OF WORDS OVER A FINITE ALPHABET

Ioan Tomescu
Faculty of Mathematics and Computer Science, University of Bucharest
Str. Academiei, 14, 010014 Bucharest, Romania
e-mail: ioan@fmi.unibuc.ro

ABSTRACT

This paper deals with periodic words of the form w_0^k over an alphabet A of cardinality m, where w_0 is fixed and contains $p \leq m$ different letters (p is also fixed). Let k_n be a sequence of positive integers. It is shown that if $\limsup_{n\to\infty} pk_n / \ln n < 1 / \ln m$ then almost all words of length n over A contain the factor $w_0^{k_n}$ as $n \to \infty$, but if $\limsup_{n\to\infty} pk_n / \ln n > 1 / \ln m$ then this property is not longer true. Also, if $\liminf_{n\to\infty} pk_n / \ln n > 1 / \ln m$ then almost all words of length n over A do not contain the factor $w_0^{k_n}$.

Moreover, if there exists $\lim_{n\to\infty}(\ln n - pk_n \ln m) = \delta \in \mathbb{R}$, then the proportion of words of length n containing the factor $w_0^{k_n}$ approaches $1 - \exp(-(1 - 1/m^p)\exp(\delta))$ as $n \to \infty$.

Keywords: Word, factor, recurrence relation, pseudo-Vandermonde system, generating function

1. Notation and Preliminary Results

We consider a finite alphabet A of cardinality $|A| = m$. A factor of a word $v \in A^*$ is a word $u \in A^*$ for which there exist $p, q \in A^*$ such that $v = puq$ [3]. A factor u of a word v can occur in v in different positions, each of those being uniquely determined by the length of the prefix of v preceding u. For example, abc occurs in $abcababc$ in positions 0 and 5. The first occurrence of u in v is that occurrence having minimum position. If $\alpha = a^k \in A^*$ is a fixed word of length $|\alpha| = k \geq 1$ having all letters equal to a, let $L(n)$ denote the number of words $w \in A^*$ such that $|w| = n$ and w does not contain the factor α. We recall the following results from [8, 9]:

Lemma 1 [8] *Numbers $L(n)$ satisfy the following recurrence relation:*

$$L(n+1) = mL(n) - (m-1)L(n-k) \tag{1}$$

for every $n \geq k$ and $L(p) = m^p$ for $0 \leq p \leq k-1$ and $L(k) = m^k - 1$. Moreover, the number of words $w \in A^*$ such that $|w| = n$ and w does not contain a fixed factor $\beta = \beta_1 \dots \beta_k \neq \alpha$ of length k over A is less than or equal to $L(n)$. □