Journal of Automata, Languages and Combinatorics **11** (2006) 2, 147–159

GENERATING ALL CIRCULAR SHIFTS BY CONTEXT-FREE GRAMMARS IN CHOMSKY NORMAL FORM

Peter R. J. Asveld

Department of Computer Science, Twente University of Technology
P. O. Box 217, 7500 AE Enschede, the Netherlands
e-mail: `infprja@cs.utwente.nl`

ABSTRACT

Let $\{a_1, a_2, \ldots, a_n\}$ be an alphabet of n symbols and let C_n be the language of circular or cyclic shifts of the word $a_1a_2\ldots a_n$; so $C_n = \{a_1a_2\ldots a_{n-1}a_n, a_2a_3\ldots a_na_1, \ldots, a_na_1\ldots a_{n-2}a_{n-1}\}$. We discuss a few families of context-free grammars G_n $(n \geq 1)$ in Chomsky normal form such that G_n generates C_n. The grammars in these families are investigated with respect to their descriptional complexity, i. e., we determine the number of nonterminal symbols $\nu(n)$ and the number of rules $\pi(n)$ of G_n as functions of n. These ν and π happen to be functions bounded by low-degree polynomials, particularly when we focus our attention to unambiguous grammars. Finally, we introduce a family of minimal unambiguous grammars for which ν and π are linear.

Keywords: Context-free grammar, Chomsky normal form, permutation, circular shift, cyclic shift, descriptional complexity, unambiguous grammar

1. Introduction

In [2] we considered context-free grammars G_n $(n \geq 1)$ in Chomsky normal form that generate the finite language L_n consisting of all permutations over the alphabet $\Sigma_n = \{a_1, a_2, \ldots, a_n\}$. A few families $\{G_n\}_{n\geq 1}$ of such grammars for $\{L_n\}_{n\geq 1}$ have been studied in [2] with respect to their grammatical complexity; more precisely, the number of nonterminal symbols $\nu(n)$ and the number of production rules $\pi(n)$ of G_n have been determined as functions of n.

In this paper we restrict ourselves to a subset of specific permutations over Σ_n, viz. to the so-called circular or cyclic shifts. Assuming a linear order on Σ_n, e. g., $a_1 < a_2 < \cdots < a_n$, the set C_n of *circular shifts* (also called *cyclic shifts*) over Σ_n is

$$C_n = \{a_1a_2\ldots a_{n-1}a_n, a_2a_3\ldots a_na_1, a_3a_4\ldots a_1a_2, \ldots, a_na_1\ldots a_{n-2}a_{n-1}\}.$$

So C_n is obtained from the word $a_1a_2\ldots a_n$ by repetitively moving the first symbol to the right-hand end of the string. Clearly, C_n contains n strings, which is much less than the $n!$ elements of L_n [7, 12]. Considered as a set of permutations of $a_1a_2\ldots a_n$ the set C_n forms a cyclic subgroup of order n of the symmetric group $\mathfrak{S}_n$ over Σ_n