Journal of Automata, Languages and Combinatorics **3** (1998) 2, 85–103

MULTI-BRACKETED CONTEXTUAL GRAMMARS

MARTIN KAPPES
Fachbereich Informatik, Johann Wolfgang Goethe-Universität
D-60054 Frankfurt am Main, Germany
e-mail: kappes@psc.informatik.uni-frankfurt.de

ABSTRACT

Bracketed contextual grammars are contextual grammars with an induced Dyck-structure to control the derivation process and to provide derivation trees. In this paper, we introduce a generalization of bracketed contextual grammars called multi-bracketed contextual grammars. The generative capacity of this class is investigated. It will be shown that context-free languages and the class of languages generated by tree adjoining grammars without local constraints are also generated by natural subclasses of multi-bracketed contextual grammars.

Keywords: contextual grammars, tree adjoining grammars.

1. Motivation and Introduction

Contextual grammars are a formalization of the linguistic idea that more complex well formed strings are obtained by inserting contexts into already well formed strings. They were first introduced by MARCUS in 1969 and have been extensively investigated. For a detailed overview the reader is referred to the monograph [11]; more compressed sources of information are [1] or [4].

Although the model is well motivated from a linguistic point of view, it does not assign any structure as derivation trees for context-free grammars do to the derived words. To overcome this shortcoming, MARTIN-VIDE and PĂUN introduced bracketed contextual grammars in the paper [10]. The generative capacity of bracketed contextual grammars was studied in the paper [8]. The idea to induce a bracket-structure into contextual grammars in order to control the derivation process is also linguistically well-motivated. However it turned out that bracketed contextual grammars seem too restricted since even simple regular languages cannot be generated by arbitrary grammars of this type. Therefore in this paper we introduce a generalization of this model and study some properties from a formal language theory point of view.

2. From Internal Contextual Grammars to Multi-Bracketed Contextual Grammars

Let Σ^* denote the free monoid generated by the finite alphabet Σ and $\Sigma^+ = \Sigma^* - \{\lambda\}$, where λ denotes the empty word. By $|x|$ we denote the length of a string, $|x|_a$ is the