

ON THE COMPLEXITY OF A MILDLY CONTEXT-SENSITIVE LANGUAGE CLASS

DÁVID ANGYAL^(A) BENEDEK NAGY^(B) GYÖRGY VASZIL^(A)

^(A)*Department of Computer Science, Faculty of Informatics
University of Debrecen, Kassai út 26, 4028 Debrecen, Hungary
{angyal.david,vaszil.gyorgy}@inf.unideb.hu*

^(B)*Department of Mathematics, Faculty of Arts and Sciences
Eastern Mediterranean University, Famagusta, North-Cyprus, Mersin-10, Turkey
nbenedek.inf@gmail.com*

ABSTRACT

We study the mildly context-sensitive class of languages generated by controlled linear grammars with context-free (exact) control, or equivalently, accepted by a family of two-head pushdown automata and a family of input reversal pushdown automata. (A controlled grammar is with exact control, if there exists a valid derivation in the grammar for each word of the control language. A pushdown automaton is input reversal, if it is able to reverse the yet unread part of its input word.) We present a simple normal form for linear grammars with context-free exact control, and study the problem of parsing the generated languages. We give a new parsing algorithm based on Earley's method having a time complexity of $\mathcal{O}(n^5)$.

Keywords: mildly context-sensitive languages, controlled grammars, parsing

1. Introduction

To support the idea that the “world is not context-free”, the authors of the monograph [4] list seven circumstances where context-free languages are not enough to characterize the phenomena in question, one of the most important being the description of natural languages. Mildly context-sensitive language families are investigated by a strong linguistic motivation [11], because the superclasses of the class of context-free languages in the Chomsky hierarchy, i. e., the classes of context-sensitive and recursively enumerable languages seem to be too complex and too impractical to handle from the computational point of view [9, 16]. Thus, other formalisms have been created to reduce the size of the analyzed languages to only semi-linear ones (or languages with the constant grow property, as required by Joshi in [11]); but still including all the context-free languages, together with some non-context-free context sensitive ones, such as the three linguistically important languages: the copy language,