



UNIVERSITY OF DEBRECEN
FACULTY OF INFORMATICS



PhD Referee's Report

Author: Ondřej Soukup

Title: Context-Freeness Resulting from Tree-restricted Grammars

This PhD thesis deals with one of the fundamental questions of formal language theory. In many cases it is very difficult problem to determine the exact position of a certain language in the Chomsky hierarchy. It is an especially hard problem to decide if a language belongs to the family of context-free languages, for example a 27 years old open problem is if Q - the language of all primitive words over a nontrivial alphabet - is context-free or context-sensitive. For context-free languages a positive lemma has not existed so far to help solving this decision problem, the well known pumping lemmas give only necessary but not sufficient conditions. The main result of this dissertation shows a new tool, which can be used to solve this decision problem by giving necessary and sufficient conditions for languages - defined by different kind of non-context-free grammars - to be context-free. By this tool one can give a positive proof for a language to be context-free.

Details:

The thesis contains 8 main chapters, a reference list, and a list of publications.

- 1.) In the first chapter the author position the problem in the theory of formal languages, and gives an overview of the thesis.
- 2.) In the second chapter the necessary definitions and some theorems are given. This part contains all the basic concepts of formal language theory, which will be used in the following chapters. Although there are a few minor mistakes - e.g. definition of inverse homomorphism - this section is well written and easy to follow. The most important part is the introduction some special grammar types which are important for the thesis: regular-controlled grammars, scattered-context grammars, and cooperating distributed grammar systems. The examples are useful they help understanding these concepts. Normal forms are also introduced and some theorems - with constructive proofs - show the position of the introduced grammar types in normal forms in the Chomsky hierarchy. Here are also some nice examples showing that the introduced grammars generate languages out of the context-free language family.
- 3.) The third chapter deals with the derivation trees of the previously introduced grammars and introduces new concepts which are used in section 6 in the main theorems and the proofs of these theorems.



UNIVERSITY OF DEBRECEN
FACULTY OF INFORMATICS



4.) Chapter 4 shows the recent proof techniques, which can be used to decide the position of a given language in the Chomsky hierarchy. This chapter contains two parts, positive and negative proofs. Unfortunately very few positive proofs are given, and until now there were no positive proofs for context-free languages. Negative proof techniques are based on the well-known Bar-Hillel lemma, and other different versions of pumping lemmas for different language families, however there are some pumping lemmas for regular languages which give necessary and sufficient conditions and there is a well-known workspace theorem which can also be used to reach positive and negative results for context-sensitive languages.

5.) The fifth chapter is a short one extending the Chomsky hierarchy with the previously introduced grammars.

6.) The main chapter of the thesis contains the new results reached by the author. There are 4 sections, each of them deals with one of the grammar types.

The main theorem of the first section gives necessary and sufficient conditions for a language generated by a general grammar to be context-free by limitations of the derivation trees, namely, by limitations of the neighbouring paths of the derivation trees. This is a constructive proof, it shows the exact method to create a context-free grammar which is equivalent to the original grammar - satisfying the conditions - generating the same language. After giving the construction it is proved that the original and the created grammars are equivalent. (This proof uses the binary form of the grammar, it should be mentioned in the proof or in the theorem, not just before the theorem.)

The second section deals with regular-controlled grammars, the theorems give limitation for path-changing steps and show that the languages generated by regular-controlled grammars satisfying the conditions are context-free of index $k+1$. There are also corollaries for context-free languages of an infinite index, propagating regular-controlled grammars and matrix grammars.

In the third section the author investigates the scattered control grammars in binary forms and gives necessary and sufficient conditions for a language generated by these grammars to be context-free by limiting the number of the context-dependent nodes.

The last section of the main chapter shows four theorems for distributed grammar systems. These theorems give necessary and sufficient conditions for languages generated by cooperating distributed grammar systems to be context-free. Here we can also see nice constructive proofs.

7.) In chapter 7 the author shows the applications of the new theorems in practice. This chapter contains three examples. In each example there is a language defined by different - non-context-free - grammar and the author demonstrates that using the theorems of the previous section it can be shown that these languages are context-free.

8.) The chapter on conclusions also contains open problems, which is a nice contribution to the thesis.



UNIVERSITY OF DEBRECEN
FACULTY OF INFORMATICS



9.) References are well selected.

10.) The author's list of publication contains 3 conference papers and 3 journal articles in high quality journals, which is more than enough for a PhD degree.

Summary:

The topic of the thesis is interesting and not marginal at all, the author have been searching answers in one of the main questions of this scientific area. The results are strong and the proofs are clear. The structure of the dissertation is well organized the style is nice and easy to follow. In my opinion this PhD dissertation is above the standard, I recommend the acceptance of the thesis.

18 September 2017
Debrecen, Hungary

Dr. Géza Horváth
associate professor
University of Debrecen
Faculty of Informatics
Department of Computer Science