Parallel parsing based upon general multigenerative grammar systems

Author: Ing. Zbyněk Sopuch

The topic of my PhD thesis deals with parsing based upon well-know grammar systems and their modifications. My master's thesis is related to this PhD thesis and it is based on parallel parsing upon multigenerative grammar systems with rules in the special forms.

The multigenerative systems were defined and described in the PhD thesis by Ing. Roman Lukáš, Ph.D. This system is composed from several (*n*) context-free grammars which are working parallel and which are synchronized by a special component Q. This special component contains sets of nonterminals or sets of rules. Each grammar generates its own string and the result generated by the whole system is called a multiform[1]. All multiforms, which can be generated by one multigenerative system, are together called the *n*-language[2].

There are defined three types of the multigenerative systems:

- *Canonical multigenerative grammar systems*
- *General multigenerative grammar systems*
- *Hybrid multigenerative grammar systems*

However there is described only parsing of canonical multigenerative grammar systems in Lukáš's PhD thesis. The parsing of the general multigenerative grammars systems was discussed in my master's thesis, but how I mentioned above, I considered only the special forms of rules based on the Cocke-Younger-Kasami form.

Undoubtedly these special forms of rules had effect on the parsing and they were building new borders in the process of the syntax analysis. The process and borders were different even when I changed this special form a little – it's discussed in my master's thesis too.

In this work I want to continue in my and Lukáš's work and explore and define these borders. I want to try to find the ways of parallel parsing based upon general multigenerative grammar systems. The process of parsing should be bottom-up and at the beginning it should be working over *n*-languages.

If the analysis of the problem is going well, I would like to expand my research to the field of parallel parsing based upon special modes of *n*-languages. These special modes, which Lukáš mentioned in his work, are: *the mode of first component*, *the union mode* and the *concatenate mode*. There is first analysis of this problem in my master's thesis too and I consider these modes as really important because there are a lot of situations in common grammar areas where we deal with strings and languages, not with multiforms and *n*-languages.

[1] The term "multiform" means multiset of strings.
[2] The term "*n*-language" means language, which contains multiforms instead of regular strings.