Generalized LL Parsing Generalization

Semyon Grigorev
Saint Petersburg State University
7/9 Universitetskaya nab.
St. Petersburg, 199034 Russia
semen.grigorev@jetbrains.com, Anastasiya Ragozins, Artyom Gorokhov

February 3, 2017

Today data for parsing is not only linear string, and context-free grammar is not only programming language specification. Classical example is a graph parsing where input is a graph and grammar is a paths constraints specification. Also you can see such generalizations of parsing like Multi-string parsing presented at Parsing@SLE-2016, Abstract parsing [3], ETC. All of them are special cases of the Bar-Hillel theorem and can be generalized, but todey many of them are separated solutions.

So, goal of our work is an abstract framework for parsing based on geteralization of GLL parsing algorithm [2] which proposed by Scott and J. Also we want to investigate practical areas of application. Current the next tasks.

- Context-free path querying for graph data bases. Our GLL-basd graph-aprsing algorithm is faster then presented at WWW [1].
- Context-free pattern search in metagenomical assemblies. Not only regular but also CF-compressed input processing which is actual for metagenomic assembly precessing. Sequitur compression algorithm.
- Multiple input parsing (Scott ...).
- Error recovery as a graph parsing.

References

[1] Ch, "Contex-free rdf querying", *Theoretical Computer Science*, V. 516, p. 101–120, January 2014

[2] Scott, "GLL parsing", *Theoretical Computer Science*, V. 516, p. 101–120, January 2014

[3] ccc, "Abstract LR", *Theoretical Computer Science*, V. 516, p. 101–120, January 2014