

Generalized LL Parsing Generalization

Semyon Grigorev, Anastasiya Ragozina, Artyom Gorokhov

Saint Petersburg State University

7/9 Universitetskaya nab.

St. Petersburg, 199034 Russia

semen.grigorev@jetbrains.com, ragozina.anastasiya@gmail.com,

gorohov.art@gmail.com

Today data for parsing is not only linear string, and context-free grammar is not only programming language specification. Classical example is a context-free path querying for graph data bases where input is a graph and grammar is a paths constraints specification. Also you can find such generalizations of parsing as multiple input GLL parsing presented at Parsing@SLE-2016 by Elizabeth Scott and Adrian Johnstone, Abstract LR parsing [4] and other thechniques for dynamically generated strings parsing, etc. We have some experience in these areas 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. We have some exerieence in graph parsing and GLL. Our GLL-basd graph-aprsing algorithm is faster then presented at ISWC-2016 [2]. End some ideas of application: Error recovery as a graph parsing.

All of them are special cases of the Bar-Hillel [1] theorem on CF and regular language inpersection, and can be generalized, but today many of them are separated solutions. Thus, the goal of our work is an abstract framework for parsing based on geteralization of GLL parsing algorithm [3] which proposed by Elizabeth Scott and Adrian Johnstone. Also we want to investigate practical areas of application and create solutions based on our framework to demonstare its practical value.

References

- [1] Bar-Hillel, Yehoshua, Micha Perles, and Eliahu Shamir. “On formal properties of simple phrase structure grammars.” *Sprachtypologie und Universalienforschung* 14 (1961): 143–172.
- [2] Zhang, Xiaowang, et al. “Context-free path queries on RDF graphs.” *International Semantic Web Conference*. Springer International Publishing, 2016. 632–648.
- [3] Scott, Elizabeth, and Adrian Johnstone. “GLL parsing.”, *Electronic Notes in Theoretical Computer Science*, 253.7 (2010): 177–189.
- [4] Doh, Kyung-Goo, Hyunha Kim, and David A. Schmidt. “Abstract LR-parsing.”, *Formal Modeling: Actors, Open Systems, Biological Systems.*, Springer, 2011. 90–109.