Parallel and Distributed Parsing: A Practical Viewpoint

Abstract

Parallelism and distribution are two concepts that have become increasingly prominent in computing. In recent years, given the ongoing evolution of computing platforms, the only way to effectively handle the ever-growing amount of data for processing is to exploit exposed parallelism whenever possible entirely or to use distributed systems to achieve multi-machine parallelism. However, parallel and distributed parsing is a field in which much research is still needed. It appears that the limited adoption of distributed parallel analyzers may be attributed to the intrinsic difficulty of adapting sequential algorithms to parallel environments and specific challenges associated with language analysis in this context.

The renewed focus on operator precedence grammars suggests a potential path to overcome these challenges and advance research in this field. In this document, we will discuss the convenience of using parallel parsing to increase the performance of the core when processing a substantial amount of data. We will briefly explore deterministic parsing based on OPGs (Operator Precedence Grammars) and its application to various programming languages. Additionally, we will highlight the advantages of parallel parsing regarding performance and code execution.

We will conclude the presentation by discussing the need for more practical efforts in parallel parsing algorithms and emphasizing the benefits of parallel lexing in code execution and speedup.