

Parsing techniques – Earley Parser

Jakub Jeleň (xjelen07), Vladimír Ščeňák (xscesn00)@stud.fit.vutbr.cz

During parsing, the compiler tries to determine, whether the source text given from lexical analysis in the form of lexical symbols form a sentence that corresponds to the grammar of the translated language. During the implementation generally two basic approaches were used, translation from the top-down or bottom-up. Names of these methods correspond to procedure of a derivation tree creation. This approach also corresponds to two basic classes of grammars - LL and LR grammars, which describe some bigger subsets of context-free languages. They are used in constructing a compiler, but they can only handle with restricted classes of languages.

On the other hand, the Earley's parser algorithm can parse all context-free languages. Earley's parser can also be described as a breadth-first top-down parser with bottom-up recognition, still, we prefer to treat it as a bottom-up method, for it can handle left recursion directly but needs special measures to handle ϵ -rules. Furthermore, the basic principle of Earley's parser algorithm will be described as well as its methods, efficiency, and benefits.