

ANTLR

ANother Tool for Language Recognition

Allgemeines

ANTLR ist der Parsergenerator von PCCTS (Purdue Compiler Construction Tool Set). Er parst nach dem LL(k) Verfahren, versteht EBNF und besitzt die Möglichkeit Syntaxbäume zu erzeugen.

Lexer und Parser

Der Lexer ist nicht wie gewohnt als DFA gestaltet, sondern verwendet ebenfalls die LL(k)-Analyse. Das ist ein wenig gewöhnungsbedürftig und auch nicht einfacher, bietet aber diverse Vorteile. Man umgeht die Beschränkung auf reguläre Sprachen und semantische Aktionen sind leichter zu implementieren. Der erzeugte Quelltext ist leichter zu lesen und somit lassen sich Fehler besser identifizieren als beim DFA.

Der Parser versteht sich auf EBNF, semantische Aktionen und alles Nötige zur Generierung von Syntaxbäumen.

Tree Parsing

ANTLR kann alle Bäume parsen, die das AST Interface implementieren. Das Tree Parsing nutzt dabei ein "lookahead" von 1. Jede Regel hat implizit einen (automatisch definierten) Ergebnisbaum. Außerdem ist es möglich den Baum im XML Format zu exportieren.

Bewertung

ANTLR ist durch den Lexer ein wenig gewöhnungsbedürftig. Aber nachdem die ersten Hürden genommen sind, bietet es alles, was das Compilerbauerherz begehrt. Der erzeugte Java Code ist gut verständlich. Die möglichen semantischen Aktionen sind vielfältig und eine Möglichkeit der Baumgenerierung sowie das "Tree walking" runden das Komplettpaket ab.

Weitere Informationen findet man unter:

<http://www.antlr.org>