

The Language PCFGNLTK

BNF-converter

October 25, 2016

This document was automatically generated by the *BNF-Converter*. It was generated together with the lexer, the parser, and the abstract syntax module, which guarantees that the document matches with the implementation of the language (provided no hand-hacking has taken place).

The lexical structure of PCFGNLTK

Literals

Terminal literals are recognized by the regular expression $(\langle letter \rangle \mid \langle digit \rangle \mid \text{'_'})(\langle letter \rangle \mid \langle digit \rangle \mid \text{'_'} \mid \text{'-'})^*$

Reserved words and symbols

The set of reserved words is the set of terminals appearing in the grammar. Those reserved words that consist of non-letter characters are called symbols, and they are treated in a different way from those that are similar to identifiers. The lexer follows rules familiar from languages like Haskell, C, and Java, including longest match and spacing conventions.

The reserved words used in PCFGNLTK are the following:

There are no reserved words in PCFGNLTK.

The symbols used in PCFGNLTK are the following:

$\begin{array}{ll} \rightarrow & \backslash \mathbf{n} \quad [\\] & \text{' } \quad | \end{array}$

Comments

Single-line comments begin with #.

There are no multiple-line comments in the grammar.

The syntactic structure of PCFGNLTK

Non-terminals are enclosed between \langle and \rangle . The symbols $::=$ (production), $|$ (union) and ϵ (empty rule) belong to the BNF notation. All other symbols are terminals.

$$\langle Grammar \rangle ::= \langle ListRule \rangle$$
$$\begin{aligned} \langle Rule \rangle & ::= \langle LHS \rangle \rightarrow \langle ListRHS \rangle \\ & | \epsilon \end{aligned}$$
$$\begin{aligned} \langle ListRule \rangle & ::= \epsilon \\ & | \langle Rule \rangle \backslash \mathbf{n} \langle ListRule \rangle \end{aligned}$$
$$\langle LHS \rangle ::= \langle Terminal \rangle$$
$$\langle Prob \rangle ::= [\langle Double \rangle]$$
$$\begin{aligned} \langle RHS \rangle & ::= \langle ListSYM \rangle \langle Prob \rangle \\ & | ' \langle Terminal \rangle ' \langle Prob \rangle \end{aligned}$$
$$\langle SYM \rangle ::= \langle Terminal \rangle$$
$$\begin{aligned} \langle ListRHS \rangle & ::= \epsilon \\ & | \langle RHS \rangle \\ & | \langle RHS \rangle | \langle ListRHS \rangle \end{aligned}$$
$$\begin{aligned} \langle ListSYM \rangle & ::= \epsilon \\ & | \langle SYM \rangle \langle ListSYM \rangle \end{aligned}$$