naive-ebnf: \LaTeX\ Package for EBNF in Plain Text*

Yegor Bugayenko
yegor256@gmail.com
2023-07-14, 0.0.13

NB! Large ENBF snippets may take too long to render!

1 Introduction

This package helps render an Extended Backus-Naur Form using plain text notation:

\begin{ebnf}
\langle \lambda\text{-Expr} \rangle \rightarrow \langle \text{Var} \rangle \ |
\langle \lambda \text{ Var} \ . \ . \ \text{Expr} \rangle \ |
\langle ( \text{Expr} \ \text{Expr} ) \rangle
\end{ebnf}

The \texttt{ebnf} environment doesn’t add any formatting to the paragraph, but only replaces the plain text symbols, such as “:=” and “<Var>” with proper \LaTeX\ commands. The following syntax is understood inside the \texttt{ebnf} environment:

- \texttt{:=} separates the left-hand side from the right-hand side of the production rule;
- \texttt{<...>} denotes a non-terminal (variable);
- “...” denotes a terminal symbol;
- ‘...’ denotes a special non-printable terminal symbol, like ‘\texttt{EOL}’;
- (...|...) denotes a series of options to choose from;
- /.../ denotes a regular expression, like /\texttt{[a-z]}+/;
- [...] denotes an optional substitution;
- {...} denotes a zero or more times repetition;
- || denotes an indented vertical bar at the beginning of the string.

*The sources are in GitHub at \texttt{yegor256/naive-ebnf}
Attention: The usage of some symbols is prohibited inside terminals. Instead, the following substitutions are recommended:

- \lparen and \rparen instead of "( and ") (from the mathtools package);
- \langle and \rangle instead of "< and >";
- \lbrace and \rbrace instead of "{ and }" (also mathtools);
- \lbrack and \rbrack instead of "[ and ]" (also mathtools);
- \vert instead of "|".

They would look even better, if the following notation is used:

- \char\( ( and \char\) instead of "( and ");
- \char\< and \char\> instead of "< and >";
- \char\{ and \char\} instead of "{ and }";
- \char\[ and \char\] instead of "[ and ]".

width There is an optional argument of ebnf environment, which sets the width of the left-hand side of each rule (the default width is 6em):

\begin{ebnf}[1.5in]
\langle VeryLongVariable \rangle \rightarrow \langle X \rangle \mid \langle Y \rangle
\langle X \rangle \rightarrow "X" \text{ EOL}
\langle Y \rangle \rightarrow "Y"
\end{ebnf}

Inside the text, terminals, non-terminals, and special terminals may be formatted using three supplementary commands:

\EbnfTerminal \EbnfNonTerminal \EbnfSpecial

The non-terminal \EbnfNonTerminal{Var} in $\lambda$-calculus may be equal to $v_1, v_2, \ldots$. Application starts with "( and ends with ").

It’s possible to use them in math-mode too, for example:

If "( $f_1$ (\EbnfNonTerminal{$\lambda$-Var}) ") is always true, then $f_1$ is a tautology.

A regular expression is possible too:

If \EbnfTerminal{} $f_1$ \EbnfNonTerminal{$\lambda$-Var} \EbnfTerminal{} is always true, then \EbnfTerminal{} is a tautology.
2 Package Options

It’s possible to configure the behavior of the package with the help of a few package options:

bw   By default, some colors are used in the rendered grammar. However, the bw package option disables any colors and makes sure the grammar is black-and-white:
\usepackage[bw]{naive-ebnf}

trail The ebnf environment is doing pre-processing of the LATEX commands provided and then let LATEX render them. It may be useful to see the output generated by the pre-processing. The trail option (with a file name) asks the package to save the content of the environment after the pre-processing into the file:
\usepackage[trail=log.tex]{naive-ebnf}

3 Implementation

First, we process package options:
\RequirePackage{pgfkeys}
\pgfkeys{/ebnf/.cd,
  bw/.store in=ebnf@bw,
  trail/.store in=ebnf@trail,
Then, we include a few packages, mostly to deal with MakeX3 expressions:
\RequirePackage{expl3}
\color{gray}{\texttt{Finally, we include \texttt{xcolor} to colorize the output a bit:}}
\begin{verbatim}
\makeatletter\ifdefined\ebnf@bw\else
\RequirePackage{xcolor}\fi
\newcommand\ebnf@color[2]{\ifdefined\ebnf@bw#2\else\textcolor{#1}{#2}\fi}
\makeatother
\EbnfTerminal Then, we define a command to render a single terminal:
\makeatletter
\newcommand\EbnfTerminal[1]{{\relax\ifmmode\ttfamily\else\sffamily\fi#1}}
\makeatother
\EbnfTerminal Then, we define a command to render a single non-terminal:
\makeatletter
\newcommand\EbnfNonTerminal[1]{{\relax\ifmmode\langle\else\texttt{\langle}\fi#1\relax\ifmmode\rangle\else\texttt{\rangle}\fi}}
\makeatother
\EbnfSpecial Then, we define a command to render a single non-terminal:
\makeatletter
\newcommand\EbnfSpecial[1]{{#1}}
\makeatother
\EbnfRegex Then, we define a command to render a regular expression:
\makeatletter
\newcommand\EbnfRegex[1]{{\relax\ifmmode\ttfamily/#1/\else#1\fi}}
\makeatother
\EbnfOptional Then, we define supplementary commands:
\makeatletter
\newcommand\ebnf@optional[1]{\ebnf@color{gray}{[#1]}}
\newcommand\ebnf@repetition[1]{\ebnf@color{gray}{\{}#1\ebnf@color{gray}{\}}}\ExplSyntaxOn
\newcommand\ebnf@grouping[1]{\ebnf@color{gray}{(}#1\ebnf@color{gray}{)}}
\ExplSyntaxOff
Then, we define the ebnf environment:

```latex
\ExplSyntaxOff
\newcommand\ebnf@special[1]{
  \tl_set:Nn \l_ebnf_tl {}\tl_set_rescan:Nnn \l_ebnf_tl {} { #1 }
  \ebnfSpecial{\l_ebnf_tl}
}\ExplSyntaxOn
\newcommand\ebnf@nonterminal[1]{
  \tl_set:Nn \l_ebnf_tl {}
  \tl_set_rescan:Nnn \l_ebnf_tl {} { #1 }
  \ebnfNonTerminal{\l_ebnf_tl}
}\newcommand\ebnf@regexp[1]{
  \tl_set:Nn \l_ebnf_tl {}
  \tl_set_rescan:Nnn \l_ebnf_tl {} { #1 }
  \ebnfRegex{\l_ebnf_tl}
}\newcommand\ebnf@to {
  \ebnf@color{gray}{\( \to \)}
}\newcommand\ebnf@alternation {
  \ebnf@color{gray}{\( \vert \)}
}\makeatother
ebnf Then, we define the ebnf environment:
\begin{itemize}
\item \textbf{ebnf} Then, we define the ebnf environment:
\item \textbf{ebnf} Then, we define the ebnf environment:
\item \textbf{ebnf} Then, we define the ebnf environment:
\end{itemize}
Change History

0.0.1
General: First draft. .......................... 3

0.0.11
\texttt{ebnf}: Many bugs fixed in the area of regular expression matching. .... 5

0.0.2
General: Proper parsing of grouping. . 3
Substitutions suggested for special symbols. ............... 3
\texttt{\textbackslash EbfnNonTerminal}: New command \texttt{\textbackslash EbfnNonTerminal} added, to enable rendering non-terminal symbols outside of the \texttt{ebnf} environment. ...................... 4
New command \texttt{\textbackslash EbfnTerminal} added, to enable rendering terminal symbols outside of the \texttt{ebnf} environment. ...................... 4

0.0.3
\texttt{\textbackslash EbfnTerminal}: Quotes fixed in both text and math modes. .............. 4

0.0.4
\texttt{ebnf}: Any symbols are allowed inside \texttt{\textbackslash EbfnNonTerminal} commands and inside the \texttt{ebnf} environment, where non-terminals are mentioned. ....................... 5

0.0.5
General: New package option \texttt{trail} added, to enable saving of the generated \TeX{} content to a file, for debugging purposes. ...................... 3

0.0.6
\texttt{\textbackslash EbfnSpecial}: New command \texttt{\textbackslash EbfnSpecial} added, to enable rendering of special non-printable terminal symbols outside of the \texttt{ebnf} environment. ...................... 4

0.0.8
\texttt{\textbackslash EbfnRegex}: New command \texttt{\textbackslash EbfnRegex} added, to enable rendering of regular expressions outside of the \texttt{ebnf} environment. . 4
Index

Numbers written in italic refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; numbers in roman refer to the code lines where the entry is used.

### Symbols
- \( \) (25, 27, 65, 67, 96)
- \( \) (25, 27, 65, 67, 96)
- \[ \) (25, 27, 65, 67, 96)
- \{ \) (25, 27, 65, 67, 96)
- \} (25, 27, 65, 67, 96)
- \| (25, 27, 65, 67, 96)

### Numbers
- \2 (76, 82, 84)
- \c (88, 92, 97, 102, 105, 107, 109, 111, 113, 115, 117)
- \closeout (124)
- \cs (70, 89, 94, 99)

### E
- \ebnf (69, 73, 76, 78, 80, 82, 84, 88, 89, 92, 93, 94, 97, 98, 99, 102, 103, 105, 107, 109, 111, 113, 115, 117, 118, 119, 123, 128)
- \ebnf@alternation (66)
- \ebnf@bw (4, 10, 14)
- \ebnf@color (10, 19, 21, 25, 27, 37, 39, 41, 65, 67)
- \ebnf@grouping (40)
- \ebnf@nonterminal (53)
- \ebnf@optional (36)
- \ebnf@regexp (58)
- \ebnf@repetition (38)
- \ebnf@special (48)
- \ebnf@terminal (43)
- \ebnf@to (64)

### O
- \ebnf@trail (5, 120, 122, 126)
- \ebnf@write (121, 122, 123, 124)

### P
- \openout (122)

### R
- \regexp (75, 77, 79, 81, 83, 86, 90, 95, 100, 104, 106, 108, 110, 112, 114, 116)
- \relax (18, 19, 21, 25, 26, 27, 30, 33, 35, 71)
- \RequirePackage (1, 9, 11)

### S
- \sffamily (19, 21, 26)
- \space (125, 126)

### T
- \textcolor (14)
- \textsf (18, 30, 33)

### U
- \unexpanded (123)

### V
- \vert (67)

### W
- \write (123)