

The Language hardtyped

BNF-converter

May 10, 2022

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 hardtyped

Identifiers

Identifiers $\langle Ident \rangle$ are unquoted strings beginning with a letter, followed by any combination of letters, digits, and the characters `_` `'`, reserved words excluded.

Literals

Double-precision float literals $\langle Double \rangle$ have the structure indicated by the regular expression $\langle digit \rangle + \cdot \langle digit \rangle + (\text{'e'-'?'}\langle digit \rangle +)?$ i.e. two sequences of digits separated by a decimal point, optionally followed by an unsigned or negative exponent.

Integer literals $\langle Int \rangle$ are nonempty sequences of digits.

String literals $\langle String \rangle$ have the form `"x"`, where *x* is any sequence of any characters except `"` unless preceded by `\`.

BoolValue literals are recognized by the regular expression `{"true"} | {"false"}`

UnitValue literals are recognized by the regular expression `{"unit"}`

Print literals are recognized by the regular expression `{"print"}`

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 hardtyped are the following:

```
Boolean Integer Real
String  Unit   as
in      let
```

The symbols used in hardtyped are the following:

```
;  /\  {
}  (  )
+  -  *
/  =  .
:
```

Comments

Single-line comments begin with `//`.

Multiple-line comments are enclosed with `/*` and `*/`.

The syntactic structure of hardtyped

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.

$$\begin{array}{lcl} \langle ListExpr \rangle & ::= & \epsilon \\ & | & \langle Expr \rangle \\ & | & \langle Expr \rangle ; \langle ListExpr \rangle \end{array}$$

$$\begin{aligned}
\langle Expr \rangle & ::= & /\backslash \langle FuncArg \rangle \{ \langle InExpr \rangle \} \\
& | & \langle Expr \rangle (\langle Expr \rangle) \\
& | & \langle Print \rangle (\langle Expr \rangle) \\
& | & \langle Expr \rangle + \langle Expr \rangle \\
& | & \langle Expr \rangle - \langle Expr \rangle \\
& | & \langle Expr \rangle * \langle Expr \rangle \\
& | & \langle Expr \rangle / \langle Expr \rangle \\
& | & + \langle Expr \rangle \\
& | & - \langle Expr \rangle \\
& | & \text{let } \langle VarDec \rangle = \langle Expr \rangle \\
& | & \text{let } \langle VarDec \rangle = \langle Expr \rangle \text{ in } \langle Expr \rangle \\
& | & \text{let } \langle VarDec \rangle = \langle Expr \rangle \text{ as } \langle Type \rangle \\
& | & \langle Integer \rangle \\
& | & \langle Double \rangle \\
& | & \langle String \rangle \\
& | & \langle Ident \rangle \\
& | & (\langle Expr \rangle) \\
\langle InExpr \rangle & ::= & \langle InExpr \rangle ; \langle InExpr \rangle \\
& | & \langle InExpr \rangle ; \\
& | & \langle Expr \rangle \\
\langle FuncArg \rangle & ::= & \langle VarDec \rangle . \langle FuncArg \rangle \\
& | & \langle VarDec \rangle . \\
\langle VarDec \rangle & ::= & \langle Ident \rangle : \langle Type \rangle \\
& | & \langle Ident \rangle \\
\langle Type \rangle & ::= & \text{Integer} \\
& | & \text{Real} \\
& | & \text{Boolean} \\
& | & \text{String} \\
& | & \text{Unit}
\end{aligned}$$