

Universidade Federal de Alagoas
Instituto de Computação
Compiladores

Onicla - Especificação dos tokens

Ramon Basto Callado Lima Lopes
José Rubens da Silva Brito

Maceió
2020.1

Sumário

Linguagem para desenvolvimento	3
Enumeração com as categorias de Tokens	3
Expressões Regulares Auxiliares	4
Expressões Regulares dos Lexemas	4

1. Linguagem para desenvolvimento

A linguagem de programação Onicla possui os seus analisadores léxico e sintático desenvolvidos na linguagem de programação Java. Onde o Java Development Kit na versão 11.0.10.

2. Enumeração com as categorias de Tokens

```
public enum TokenCategory {  
    EOF,  
    ID,  
    CTE_FLOAT,  
    CTE_INT,  
    CTE_CHAR,  
    OP_ATR,  
    OP_REL,  
    OP_AD,  
    OP_SUB,  
    OP_MULT,  
    OP_DIV,  
    OP_RES,  
    OP_GREATER,  
    OP_LESS,  
    OP_GREATEREQ,  
    OP_LESSEQ,  
    OP_AND,  
    OP_OR,  
    OP_NOT,  
    OP_NOTUN,  
    COMMENT,  
    PR_AND,  
    PR_OR,  
    PR_FUNCTION,  
    PR_REFOUND,  
    PR_IF,  
    PR_ELSE,  
    PR_WHILE,  
    PR_FORREPEAT,  
    PR_INTEGER,  
    PR_FLOAT,
```

```

PR_CHARACTERARRAY,
PR_CHARACTER,
PR_BOOL,
PR_INPUT,
PR_PRINT,
PR_TRUE,
PR_FALSE,
PR_NULL,
PR_BEGIN,
PR_END,
BOOL_VALUE,
AB_BEGIN,
FEC_END,
AB_PAR,
FEC_PAR,
AB_COL,
FEC_COL,
PR_VOID,
PR_MAIN,
TERMINAL,
SEP;
}

```

3. Expressões Regulares Auxiliares

letras_minusculas = [a-z]

letras_maiusculas = [A-Z]

dígitos = [0-9]

alfanumerico = letras_minusculas | letras_maiusculas | digito;

simbolo = ' ' | '.' | ',' | ':' | ';' | '?' | '!' | '+' | '-' | '*' | '\' | '/' | ' ' | '%' | '@' | '&' | '#' | '\$' | 'i' |
 '¿' | '=' | '(' | ')' | '[' | ']' | '{' | '}' | '_' | '"' | "'";

4. Expressões Regulares dos Lexemas

Identificadores	
ID	{' ' letra}, {alfanumérico ' ' }

Palavras reservadas	
PR_VOID	'Void'
PR_MAIN	'Main'
PR_IF	'If'
PR_ELSE	'Else'
PR_WHILE	'While'
PR_FORREPEAT	'Repeat'
BOOL_VALUE	'True' 'False'
CTE_FLOAT	'Float'
CTE_INT	'Integer'
CTE_CHARACTERARRAY	'CharacterArray'
CTE_CHARACTER	'Character'
COMMENT	'#'
PR_BEGIN	'BEGIN'
PR_END	'END'
PR_INPUT	'Input'
PR_PRINT	'Print'
PR_BOOL	'Bool'

PR_TRUE	'True'
PR_FALSE	'False'
PR_NULL	'NULL'

Operadores	
OP_ATR	'='
OP_REL	'==', '!='
OP_AD	'+'
OP_SUB	'-'
OP_MULT	'*'
OP_DIV	'/'
OP_RES	'%'
OP_GREATER	'>'
OP_LESS	'<'
OP_GREATEREQ	'<='
OP_LESSEQ	'>='
OP_AND	'AND'
OP_OR	'OR'
OP_NOT	'!'
OP_NOTUN	'~'

Delimitadores	
AB_BEGIN	'BEGIN'
FEC_END	'END'

AB_PAR	‘(
FEC_PAR)’
AB_COL	[‘
FEC_COL]’
TERMINAL	‘,’
SEP	‘ ’