

Universidade Federal de Alagoas
Instituto de Computação
Bacharelado em Ciência da Computação

Especificação dos Tokens - Isengard++

Márcio Henrique Vieira de Oliveira
Michael Miller Rodrigues Cardoso

Maceió - AL, 2021

Sumário

Sumário	1
1. Linguagem de desenvolvimento	2
2. Enumeração das categorias de tokens	2
3. Expressões regulares auxiliares	3
4. Expressões regulares dos lexemas	4

1. Linguagem de desenvolvimento

Os analisadores da linguagem de programação **Isengard++** foram desenvolvidos na linguagem **Java**, especificamente a versão JDK 11.0.12, utilizando a IntelliJ IDEA 2021.1.2.

2. Enumeração das categorias de tokens

```
public enum TokensEnum{
    EOF,
    ID,

    // palavras reservadas
    PR_BEGIN,
    PR_END,
    PR_FUNCTION,
    PR_MAIN,
    PR_VOID,
    PR_NULL,
    PR_RETURN,
    PR_AND,
    PR_OR,
    PR_NOT,
    PR_IF,
    PR_ELSE,
    PR_WHILE,
    PR_FOR,
    PR_INT,
    PR_FLOAT,
    PR_CHAR,
    PR_STR,
    PR_BOOL,
    PR_INPUT,
    PR_OUTPUT,
    PR_OUTPUTLN,
    PR_TRUE,
    PR_FALSE,

    // constantes
    CT_INT,
    CT_FLOAT,
    CT_CHAR,
    CT_STR,
```

```

//operadores
OP_ATR,
OP_RELEQUAL,
OP_RELDIF,
OP_AD,
OP_SUB,
OP_MULT,
OP_DIV,
OP_RES,
OP_GREATER,
OP_LESS,
OP_GREATERT,
OP_LESST,
OP_NOTUNI,
OP_CONCAT,
OP_SIZE,

// delimitadores
AB_PAR,
FEC_PAR,
AB_COL,
FEC_COL,
TERMINAL,
SEP,

// erros
ERR_ID,
ERR_CHAR,
ERR_NUM,
ERR_PR,
ERR_SYM
}

```

3. Expressões regulares auxiliares

```
small_letter = '[:lower:]'
```

```
digits = '[:digit:]]+'
```

```
alphanumeric = '[:alnum:]'
```

```
symbol = '[:punct:]]{-}[\\"']'
```

ans = '[{alphanumeric}{symbol}[:space:]]'

4. Expressões regulares dos lexemas

Identificadores	
ID	'{small_letter}{alphanumeric}{1,15}'

Palavras reservadas	
PR_BEGIN	'Begin'
PR_END	'End'
PR_FUNCTION	'Funct'
PR_MAIN	'Main'
PR_VOID	'Void'
PR_NULL	'Null'
PR_RETURN	'Return'
PR_AND	'And'
PR_OR	'Or'
PR_NOT	'Not'
PR_IF	'If'
PR_ELSE	'Else'
PR_WHILE	'While'
PR_FOR	'For'
PR_INT	'Int'
PR_FLOAT	'Float'
PR_CHAR	'Char'
PR_STR	'Str'

PR_BOOL	'Bool'
PR_INPUT	'Input'
PR_OUTPUT	'Output'
PR_OUTPUTLN	'Outputln'
PR_TRUE	'True'
PR_FALSE	'False'
CT_INT	'{digits}'
CT_FLOAT	'{digits}\.{digits}'
CT_CHAR	'\{ans}?\'
CT_STR	'\"{ans}*\"'

Operadores	
OP_ATR	'='
OP_RELEQUAL	'=='
OP_RELDIF	'!='
OP_AD	'+'
OP_SUB	'_'
OP_MULT	'*'
OP_DIV	'/'
OP_RES	'%'
OP_GREATER	'>'
OP_LESS	'<'
OP_GREATER_T	'>='
OP_LESS_T	'<='
OP_NOTUNI	'~'
OP_CONCAT	'&'
OP_SIZE	'@'

Delimitadores	
AB_PAR	'(
FEC_PAR)'
AB_COL	'[
FEC_COL]'
TERMINAL	','
SEP	','

Erros Léxicos	
ERR_ID	Erro na formação de identificador
ERR_CHAR	Erro na formação de um Char ou Str
ERR_NUM	Erro na formação de Int ou Float
ERR_PR	Erro na formação de palavra reservada
ERR_SYM	Erro de símbolo desconhecido