

**UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE COMPUTAÇÃO
CIÊNCIA DA COMPUTAÇÃO**



**UNIVERSIDADE FEDERAL
DE ALAGOAS**

Especificações - Tokens (BFS)

**Bruna Leal Torres Silva
Eirene de Oliveira Fireman
Samuel Lucas Vieira Lins Barbosa**

Sumário

Sumário	1
A linguagem de desenvolvimento	2
Enumeração de categoria de Tokens	2
Expressões regulares dos lexemas	3

1. A linguagem de desenvolvimento

A linguagem de programação BFS tem seu desenvolvimento de analisador léxico feito na linguagem Java, mais especificamente o OpenJDK 11.

2. Enumeração de categoria de Tokens

```
public enum Tokens {
```

```
    EOF,
```

```
    ID,
```

```
    CT_INT,
```

```
    CT_FLOAT,
```

```
    CT_STRING,
```

```
    CT_CHAR,
```

```
    PR_MAIN,
```

```
    PR_FN,
```

```
    PR_RETURN,
```

```
    PR_VOID,
```

```
    PR_INT,
```

```
    PR_FLOAT,
```

```
    PR_CHAR,
```

```
    PR_STRING,
```

```
    PR_BOOL,
```

```
    PR_SYSIN,
```

```
    PR_SYSOUT,
```

```
    PR_IF,
```

```
    PR_ELIF,
```

```
    PR_ELSE,
```

```
    PR_WHILE,
```

```
    PR_BREAK,
```

```
    PR_FOR,
```

```
    PR_TRUE,
```

```
    PR_FALSE,
```

```
    PR_ARRAY,
```

```
    BOOL_VALUE,
```

```
    UNDER,  
    OPR_ADD,  
    OPR_SUB,  
    OPR_MULT,  
    OPR_DIV,  
    OPR_MOD,  
    OPR_IGUAL,  
    OPR_DIF,  
    OPR_DIGUAL,  
    OPR_MAIOR,  
    OPR_MENOR,  
    OPR_MAIORIG,  
    OPR_CONC,  
    OPR_MENORIG,  
    OPR_INVERS,  
  
    OPR_AND,  
    OPR_OR,  
    OPR_NOT,  
  
    OP_CHAVES,  
    CL_CHAVES,  
    OP_PAR,  
    CL_PAR,  
    OP_COLC,  
    CL_COLC,  
  
    S_VIRG,  
    S_PVIRG,  
  
    ERR_ID,  
    ERR_CHAR,  
    ERR_NUM,  
    ERR_PR,  
    ERR_DESC,  
}
```

3. Expressões regulares auxiliares

digitos = '[:digit:]'

alfanumerico = '[:alnum:]'

simbolos = '[:punct:]]{-}[\\']'

ans = '{[alfanumerico]{simbolos}[:space:]'

Obs.: Padrão Flex.

4. Expressões regulares dos lexemas

IDENTIFICADORES	
ID	'(_[:letter:]]{alfanumerico}*){1,16}'

PALAVRAS RESERVADAS	
PR_MAIN	'main'
PR_FN	'function'
PR_RETURN	'return'
PR_VOID	'void'
PR_INT	'int'
PR_FLOAT	'float'
PR_CHAR	'char'
PR_STRING	'string'
PR_BOOL	'bool'
PR_SYSIN	'SysIn'
PR_SYSOUT	'SysOut'
PR_IF	'if'
PR_ELIF	'elif'
PR_ELSE	'else'

PR_WHILE	'while'
PR_BREAK	'break'
PR_FOR	'for'
PR_TRUE	'true'
PR_FALSE	'false'
PR_ARRAY	'array'
BOOL_VALUE	'true false'

CONSTANTES	
CT_INT	'{digitos}+'
CT_FLOAT	'{digitos}+\.{digitos}+'
CT_CHAR	'\{ans}?\'
CT_STRING	'\"{ans}*\"'

OPERADORES	
OPR_ADD	'+'
OPR_SUB	'_'
OPR_MULT	'*'
OPR_DIV	'/'
OPR_MOD	'%'
OPR_INVERS	'~'
OPR_IGUAL	'='
OPR_DIGUAL	'=='
OPR_DIF	'!='
OPR_MAIOR	'>'
OPR_MENOR	'<'

OPR_MAIORIG	'>='
OPR_MENORIG	'<='
OPR_AND	'and'
OPR_OR	'or'
OPR_NOT	'not'
OPR_CONC	'&'

DELIMITADORES	
OP_CHAVES	{'
CL_CHAVES	'}
OP_PAR	('
CL_PAR)'
OP_COLC	['
CL_COLC	']
S_VIRG	','
S_PVIRG	','
UNDER	'_'

ERROS LÉXICOS	
ERR_ID	Identificador mal formado
ERR_CHAR	Caracter ou cadeia mal formado
ERR_NUM	Número mal formado
ERR_PR	Palavra reservada mal formada
ERR_DESC	Símbolo desconhecido