

Universidade Federal de Alagoas  
Instituto de Computação

Compiladores  
Especificação dos Tokens

Carllos Eduardo Ferreira Lopes e Rafel Emílio Lima Alves  
2021.1

# 1. Linguagem

O analisador léxico da linguagem de programação Pyragua foi implementado utilizando a versão 3.9 do Python.

## 2. Enumeração dos tokens

class Tokens:

```
    tokenDict = { # Palavras Reservadas
        "RESE_INITIATE": 1,
        "RESE_HALT": 2,
        "RESE_CENTRAL": 3,
        "RESE_FUNCAO": 4,
        "RESE_RETORNA": 5,
        "RESE_INT": 6,
        "RESE_STR": 7,
        "RESE_FLOAT": 8,
        "RESE_CHAR": 9,
        "RESE_BOOL": 10,
        "RESE_VAZIO": 11,
        "RESE_VERDADE": 12,
        "RESE_FALSO": 13,
        "RESE_SE": 14,
        "RESE_SENAO": 15,
        "RESE_LOOP": 16,
        "RESE_ENQUANTO": 17,
        "RESE_NULO": 18,
        "RESE_PARE": 19,
        "RESE_LER": 20,
        "RESE_ESCREVER": 21,
        "RESE_ESCREVERPL": 22,
        # Operadores
        "OPE_CONJUN": 23,
        "OPE_DISJUN": 24,
        "OPE_ADI": 25,
        "OPE_SUB": 26,
        "OPE_MULTI": 27,
        "OPE_DIV": 28,
        "OPE_REST": 29,
        "OPE_POTEN": 30,
        "OPE_ATRI": 31,
        "OPE_UNARINEG": 32,
        "OPE_IGUAL": 33,
```

```

"OPE_DIFE": 34,
"OPE_MENORQ": 35,
"OPE_MAIORQ": 36,
"OPE_MENORI": 37,
"OPE_MAIORI": 38,
"OPE_NEGA": 39,
"OPE_CONCAT": 40,
# Identificador
"ID": 41,
# Constantes literais
"CNST_INT": 42,
"CNST_FLOAT": 43,
"CNST_BOOL": 44,
"CNST_CHAR": 45,
"CNST_STR": 46,
# Delimitadores
"DELI_OPAREN": 47,
"DELI_CPAREN": 48,
"DELI_ENBRA": 49,
"DELI_OPBRA": 50,
"DELI_COMMA": 51,
"DELI_SECOL": 52,
"DELI_EOF": 53,
# Erros
"ERR_UNK": 54,
"ERR_IND": 55,
"ERR_NUMER": 56,
"ERR_PR": 57,
"ERR_CHR": 58,
# Comentario
"COMM_TOKEN": 59}

```

### 3. Expressões Regulares Auxiliares

letra	'[a-zA-z]'
digito	'[0-9]'

## 4. Lexemas

### 4.1 Palavras Reservadas

Palavras Reservadas		
1	'Initiate'	RESE_INITIATE
2	'Halt'	RESE_HALT
3	'Central'	RESE_CENTRAL
4	'Funcao'	RESE_FUNCAO
5	'Retorna'	RESE_RETORNA
6	'Int'	RESE_INT
7	'Str'	RESE_STR
8	'Float'	RESE_FLOAT
9	'Char'	RESE_CHAR
10	'Bool'	RESE_BOOL
11	'Vazio'	RESE_VAZIO
12	'Verdade'	RESE_VERDADE
13	'Falso'	RESE_FALSO
14	'Se'	RESE_SE
15	'SeNao'	RESE_SENAO
16	'Loop'	RESE_LOOP
17	'Enquanto'	RESE_ENQUANTO
18	'Nulo'	RESE_NULO
19	'Pare'	RESE_PARE
20	'Ler'	RESE_LER
21	'Escrever'	RESE_ESCREVER
22	'Escreverpl'	RESE_ESCREVER

## 4.2 Operadores

Operadores		
23	'&'	OPE_CONJUN
24	' '	OPE_DISJUN
25	'+'	OPE_ADI
26	'-'	OPE_SUB
27	'*'	OPE_MULTI
28	'/'	OPE_DIV
29	'%'	OPE_REST
30	'^'	OPE_POTEN
31	'='	OPE_ATRI
32	'-'	OPE_UNARINEG
33	'=='	OPE_IGUAL
34	'!='	OPE_DIFE
35	'<'	OPE_MENORQ
36	'>'	OPE_MAIORQ
37	'<='	OPE_MENORI
38	'>='	OPE_MAIORI
39	'!'	OPE_NEGA
40	'@'	OPE_CONCAT

## 4.3 Indentificador

Identificador		
41	'(letra)(letra   digito   '_' )*'	ID

## 4.4 Constantes Literais

Constantes Literais		
42	'(digito)*'	CNST_INT
43	'(digito)+(.) (digito)+'	CNST_FLOAT
44	'("Verdadeiro"   "Falso")'	CNST_BOOL
45	'(') (letra) (' )'	CNST_CHAR
46	'(") (.) * (" )'	CNST_STR

## 4.5 Delimitadores

Delimitadores		
47	'('	DELI_OPAREN
48	')'	DELI_CPAREN
49	'['	DELI_ENBRA
50	']'	DELI_OPBRA
51	','	DELI_COMMA
52	','	DELI_SECOL
53	'EOF'	DELI_EOF

## 4.6 Erros

Erros		
54	Token desconhecido	ERR_UNK
55	Nome mal formado	ERR_IND
56	Número mal formado	ERR_NUMER
57	Palavra reservada mal formada	ERR_PR

58	Caractere mal formado	ERR_CHR
----	-----------------------	---------

## 4.7 Comentário

59	'#'	COMM_TOKEN
----	-----	------------