

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 ReservedDict:
    reservedWords = {'Initiate': ["RESE_INITIATE", 1],
                     'Halt': ["RESE_HALT", 2],
                     'Central': ["RESE_CENTRAL", 3],
                     'Funcao': ["RESE_FUNCAO", 4],
                     'Retorna': ["RESE_RETORNA", 5],
                     'Int': ["RESE_INT", 6],
                     'Str': ["RESE_STR", 7],
                     'Float': ["RESE_FLOAT", 8],
                     'Char': ["RESE_CHAR", 9],
                     'Bool': ["RESE_BOOL", 10],
                     "Array": ["RESE_ARRAY", 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_ESCREVERPL", 23],
                     "E": ["OPE_CONJUN", 24],
                     "Ou": ["OPE_DISJUN", 25]}

    operators = {"+": ["OPE_ADI", 26],
                 "-": ["OPE_SUB", 27],
                 "*": ["OPE_MULTI", 28],
                 "/": ["OPE_DIV", 29],
                 "%": ["OPE_REST", 30],
                 "=": ["OPE_ATRI", 31],
                 "¬": ["OPE_UNARINEG", 32],
                 "OPE_RELA": ["OPE_RELA", 33],
```

```

"<": ["OPE_MENORQ", 34],
">": ["OPE_MAIORQ", 35],
"OPE_MENORI": ["OPE_MENORI", 36],
"OPE_MAIORI": ["OPE_MAIORI", 37],
"!": ["OPE_NEGA", 38],
"@": ["OPE_CONCAT", 39]}

identifier = {"ID": 40,
              "CNST_INT": 41,
              "CNST_FLOAT": 42,
              "BOOL_VALUE": 43}

delimiters = {"CNST_CHAR": 44,
              "CNST_STR": 45,
              "DELI_INITIATE": 46,
              "DELI_HALT": 47,
              '(': ["DELI_OPAREN", 48],
              ')': ["DELI_CPAREN", 49],
              ']': ["DELI_ENBRA", 50],
              '[': ["DELI_OPBRA", 51],
              ',': ["DELI_COMMA", 52],
              ';': ["DELI_SECOL", 53]}

errors = {"ERR_UNK": 54,
          "ERR_IND": 55,
          "ERR_NUMER": 56,
          "ERR_PR": 57,
          "ERR_CHR": 58}

etc = {"ETC_EOF": 59,
       "ETC_COMMENT": 60}

```

3. Expressões Regulares Auxiliares

letras_maiusculas = '[:upper:]'

letras_minusculas = '[:lower:]'

dígitos = '[:digit:]'

alfanumerico = '[:alnum:]'

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

ans = '{alfanumerico}{simbolo}'

cadeiaChar = '\{ans}+[:space:][:word:]+\'

4. Lexemas

4.1 Palavras Reservadas

VALOR NUMÉRICO	NOME SIMBÓLICO	ER
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	'Array'	RESE_ARRAY
12	'Vazio'	RESE_VAZIO
13	'Verdade'	RESE_VERDADE
14	'Falso'	RESE_FALSO
15	'Se'	RESE_SE
16	'SeNao'	RESE_SENAO
17	'Loop'	RESE_LOOP
18	'Enquanto'	RESE_ENQUANTO
19	'Nulo'	RESE_NULO
20	'Pare'	RESE_PARE
21	'Ler'	RESE_LER
22	'Escrever'	RESE_ESCREVER
23	'Escreverpl'	RESE_ESCREVER
24	'E'	OPE_CONJUN

25	'Ou'	OPE_DISJUN
----	------	------------

4.2 Operadores

VALOR NUMÉRICO	NOME SIMBÓLICO	ER
26	'+'	OPE_ADI
27	'-'	OPE_SUB
28	'*'	OPE_MULTI
29	'/'	OPE_DIV
30	'%'	OPE_REST
31	'='	OPE_ATRI
32	'¬'	OPE_UNARINEG
33	'==' '!='	OPE_RELA
34	'<'	OPE_MENORQ
35	'>'	OPE_MAIORQ
36	'<='	OPE_MENORI
37	'>='	OPE_MAIORI
38	'!'	OPE_NEGA
39	'@'	OPE_CONCAT

4.3 Indentificadores

VALOR NUMÉRICO	NOME SIMBÓLICO	ER
40	'ID'	ID
41	'Int'	CNST_INT
42	'Float'	CNST_FLOAT
42	'Bool'	BOOL_VALUE

4.4 Delimitadores

VALOR NUMÉRICO	NOME SIMBÓLICO	ER
44	'Char'	CNST_CHAR
45	'Str'	CNST_STR
46	'Initiate'	DELI_INITIATE
47	'Halt'	DELI_HALT
48	'('	DELI_OPAREN
49	')'	DELI_CPAREN
50	'['	DELI_ENBRA
51	']'	DELI_OPBRA
52	','	DELI_COMMA
53	','	DELI_SECOL

4.5 Erros

VALOR NUMÉRICO	NOME SIMBÓLICO	ER
54	'Unknown'	ERR_UNK
55	'Identifier'	ERR_IND
56	'Numeric'	ERR_NUMER
57	'Palavra Reservada'	ERR_PR
58	'Character'	ERR_CHR

4.6 Etc

VALOR NUMÉRICO	NOME SIMBÓLICO	ER
59	'EOF'	ETC_EOF
60	'COMMENT'	ETC_COMMENT