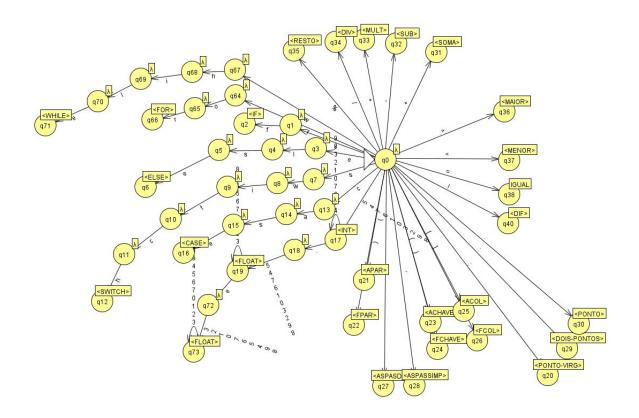
Exemplo de um Analisador Léxico construído através da Análise Preditiva

Expressões regulares para os tokens:

Token	ER	Exemplos de lexemas
<if></if>	(i)(f)	if
<else></else>	(e)(l)(s)(e)	else
<switch></switch>	(s)(w)(i)(t)(c)(h)	switch
<case></case>	(c)(a)(s)(e)	case
<while></while>	(w)(h)(i)(l)(e)	while
<for></for>	(f)(o)(r)	for
<int></int>	(0 9)+	0, 25, 5555, 64
<float></float>	. e	. , e
<apar></apar>	((
<fpar></fpar>))
<aspasdup></aspasdup>	"	н
<aspassimp></aspassimp>	í	í
<achave></achave>	{	{
<fchave></fchave>	}	}
<acol></acol>	[[
<fcol></fcol>]]
<ponto-virg></ponto-virg>	,	,
<dois-pontos></dois-pontos>	:	:
<ponto></ponto>		
<soma></soma>	+	+
	-	-
<mult></mult>	*	*
<div></div>	/	/
<resto></resto>	&	&
<dif></dif>	!	!
<igual></igual>		
<maior></maior>	>	>
<menor></menor>	<	<
<vazio></vazio>	\n \t	
<eof></eof>	Representação do fim	

Máquina de Moore:



Teste 1:

```
if(1+1 == 2)
```

Console de saída:

<IF><APAR><INT><SOMA><INT><IGUAL><INT><FPAR>

Teste 2:

```
while(1*2 != 3)
```

Console de saída:

<WHILE><APAR><INT><MULT><INT><DIF><IGUAL><INT><FPAR>

Teste 3:

```
if(1+1 < 4){
}
while(1+1 >= 2){
}
```

Console de saída:

<IF><APAR><INT><SOMA><INT><MENOR><INT><FPAR><ACHAVE><FCHAVE><WHIL
E><APAR><INT><SOMA><INT><MAIOR><IGUAL><INT><FPAR><ACHAVE><FCHAVE>

Teste 4:

1\$1

Console de saída:

<INT>

Erro léxico: caractere encontrado:\$

era(m) esperados(s):

{'IF': ['i', 'f', '<IF>'], 'ELSE': ['e', 'I', 's', 'e', '<ELSE>'], 'SWITCH': ['s', 'w', 'i', 't', 'c', 'h', '<SWITCH>'], 'CASE': ['c', 'a', 's', 'e', 'CASE'], 'WHILE': ['w', 'h', 'i', 'I', 'e', '<WHILE>'], 'FOR': ['f', 'o', 'r', '<FOR>'], 'INT': ['0', '1', '2', '3', '4', '5', '6', '7', '8', '9', '<INT>'], 'FLOAT': ['.', 'e', '<FLOAT>'], 'APAR': ['(', '<APAR>'], 'FPAR': [')', '<FPAR>'], 'ASPASDUP': ["", '<ASPASDUP>'], 'ASPASSIMP': [""", '<ASPASSIMP>'], 'ACHAVE': ['\f', '<ACHAVE>'], 'FCHAVE': ['\f', '<FCHAVE>'], 'ACOL': ['[', '<ACOL>'], 'FCOL': [']', '<FCOL>'], 'PONTO-VIRG': ['\f', '<PONTO-VIRG': ['\f', '<BONTO-VIRG': ['\f', '<BONTO-VIRG': ['\f', '<BUB>'], 'MULT': ['\f', '<MULT>'], 'DIV': ['\f', '<DIV>'], 'RESTO': ['\f', '<RESTO>'], 'DIF': ['\f', '<DIF>'], 'IGUAL': ['=', '<IGUAL>'], 'MAIOR': '>', 'MENOR': '<', 'VAZIO': [' ', '\n', '\t'], 'EOF': "} <INT>

Teste 5:

else_

Console de saída:

<ELSE>

Erro léxico: caractere encontrado:

era(m) esperados(s):

{'IF': ['i', 'f', '<IF>'], 'ELSE': ['e', 'I', 's', 'e', '<ELSE>'], 'SWITCH': ['s', 'w', 'i', 't', 'c', 'h', '<SWITCH>'], 'CASE': ['c', 'a', 's', 'e', 'CASE'], 'WHILE': ['w', 'h', 'i', 'I', 'e', '<WHILE>'], 'FOR': ['f', 'o', 'r', '<FOR>'], 'INT': ['0', '1', '2', '3', '4', '5', '6', '7', '8', '9', '<INT>'], 'FLOAT': ['.', 'e', '<FLOAT>'], 'APAR': ['(', '<APAR>'], 'FPAR': [')', '<FPAR>'], 'ASPASDUP': [""', '<ASPASDUP': [""', '<ASPASDUP': ['"', '<ACHAVE': ['\], '<FCHAVE>'], 'ACOL': ['[', '<ACOL>'], 'FCOL': [']', '<FCOL>'], 'PONTO-VIRG': ['\], '<PONTO-VIRG>'], 'DOIS-PONTOS': ['\], '<DOIS-PONTOS>'], 'PONTO': ['\], '<PONTO': ['\], '<RESTO': ['\], '<RESTO>'], 'DIF': ['\], '<DIF>'], 'IGUAL': ['=', '<IGUAL>'], 'MAIOR': '>', 'MENOR': '<', 'VAZIO': [', '\n', '\t'], 'EOF': "}

Implementação do analisador léxico por intermédio de análise preditiva:

```
import sys
              from abc import ABC, abstractmethod
          v class Analisador(ABC):
                     def init (self):
                               self.NOME DEFAULT ARQUIVO ENTRADA = 'entrada.txt'
                               self.tokens = {
                                      'IF': ['i','f', '<IF>'],

'ELSE': ['e','l','s','e', '<ELSE>|'],

'SMITCH': ['s','w','i','t','c','h', '<SWITCH>'],

'CASE': ['c','a','s','e', 'CASE'],

'WHILE': ['w','h','i','l','e', '<WHILE>'],

'FOR': ['f','o','r', '<FOR>'],

'INT': ['o','1','2','3','4','5','6','7','8','9','<INT>'],

'FLOAT': ['.','e', '<FLOAT>'],

'APAR': ['(', '<APAR>'],

'ASPASDUP': ['"', '<ASPASDUP>'],

'ASPASSIMP': [""', '<ASPASSIMP>'],

'ACHAVE': ['f', '<ACHAVE>'],

'FCHAVE': ['f', '<ACHAVE>'],

'FCOL': [']', '<FCOL>'],

'PONTO-VIRG': [';','<PONTO-VIRG>'],

'DOIS-PONTOS': [':', '<PONTO-VIRG>'],

'SOMA': ['+','<SOMA>'],

'SUB': ['-', '<SUB>'],

'MULT': ['*', '<OIV>'],

'RESTO': ['%', '<RESTO>'],

'DIF': ['f', '<DIF>'],

'IGUAL': ['=', '<IGUAL>'],

'MAIOR': '<',

'WAZIO': [''.' \n', '\t'],
19
                                        'MENOR': '<'
                                        'VAZIO': [' ', '\n', '\t'],
                                        'EOF':
                               }
                      @abstractmethod
                      def Analisador(self, _nomeArquivoEntrada):
                               self.nomeArquivoEntrada = self._nomeArquivoEntrada
                      def Analisador(self):
                               self.nomeArquivoEntrada = self.NOME DEFAULT ARQUIVO ENTRADA
         v class AnalisadorLexico(Analisador):
                      def __init__(self,_nomeArquivoEntrada):
                               super().__init__()
                               self.proxCaractere = ''
                               self.linha = 1
                               self.posicao = -1
                               self.tokenReconhecido = ''
                               self.s = []
```

```
try:
              self.file = open(_nomeArquivoEntrada, "r")
              self.entrada = self.file.read()
              self.file.close()
              self.leProxCaractere()
          except:
              print("Erro na leitura do arquivo" + _nomeArquivoEntrada)
      def leProxCaractere(self):
          try:
              self.posicao += 1
              self.proxCaractere = self.entrada[self.posicao]
          except IndexError:
              self.proxCaractere ='<EXIT>'
      def proxCaractereIs(self, s):
          if self.proxCaractere in s:
              return True
              return False
v class MyAnalisadorLexico(AnalisadorLexico):
      def __init__(self,_nomeArquivoEntrada):
           super().__init__(_nomeArquivoEntrada)
      def s0(self):
          if(self.proxCaractereIs(self.tokens['INT'])):
              self.leProxCaractere()
              self.s1()
          elif(self.proxCaractereIs(self.tokens['IF'][0])):
              self.leProxCaractere()
              self.s2()
          elif(self.proxCaractereIs(self.tokens['ELSE'][0])):
              self.leProxCaractere()
              self.s3()
          elif(self.proxCaractereIs(self.tokens['WHILE'][0])):
              self.leProxCaractere()
              self.s4()
          elif(self.proxCaractereIs(self.tokens['SWITCH'][0])):
              self.leProxCaractere()
```

```
self.s5()
elif(self.proxCaractereIs(self.tokens['CASE'][0])):
    self.leProxCaractere()
    self.s6()
elif(self.proxCaractereIs(self.tokens['FLOAT'])):
    self.leProxCaractere()
    self.s7()
elif(self.proxCaractereIs(self.tokens['FOR'][0])):
    self.leProxCaractere()
    self.s8()
elif(self.proxCaractereIs(self.tokens['APAR'])):
    self.leProxCaractere()
    self.s9()
elif(self.proxCaractereIs(self.tokens['FPAR'])):
    self.leProxCaractere()
    self.s10()
elif(self.proxCaractereIs(self.tokens['ASPASDUP'])):
    self.leProxCaractere()
    self.s11()
elif(self.proxCaractereIs(self.tokens['ASPASSIMP'])):
    self.leProxCaractere()
    self.s12()
elif(self.proxCaractereIs(self.tokens['ACHAVE'])):
    self.leProxCaractere()
    self.s13()
elif(self.proxCaractereIs(self.tokens['FCHAVE'])):
    self.leProxCaractere()
    self.s14()
elif(self.proxCaractereIs(self.tokens['ACOL'])):
    self.leProxCaractere()
    self.s15()
elif(self.proxCaractereIs(self.tokens['FCOL'])):
    self.leProxCaractere()
    self.s16()
elif(self.proxCaractereIs(self.tokens['PONTO-VIRG'])):
    self.leProxCaractere()
    self.s17()
elif(self.proxCaractereIs(self.tokens['DOIS-PONTOS'])):
    self.leProxCaractere()
    self.s18()
elif(self.proxCaractereIs(self.tokens['PONTO'])):
    self.leProxCaractere()
    self.s19()
```

```
elif(self.proxCaractereIs(self.tokens['50MA'])):
          self.leProxCaractere()
          self.s20()
     elif(self.proxCaractereIs(self.tokens['SUB'])):
          self.leProxCaractere()
          self.s21()
     elif(self.proxCaractereIs(self.tokens['MULT'])):
          self.leProxCaractere()
          self.s22()
     elif(self.proxCaractereIs(self.tokens['DIV'])):
    self.leProxCaractere()
          self.s23()
     elif(self.proxCaractereIs(self.tokens['RESTO'])):
          self.leProxCaractere()
          self.s24()
     elif(self.proxCaractereIs(self.tokens['DIF'])):
          self.leProxCaractere()
          self.s25()
     elif(self.proxCaractereIs(self.tokens['IGUAL'])):
    self.leProxCaractere()
          self.s26()
     elif(self.proxCaractereIs(self.tokens['MAIOR'])):
          self.leProxCaractere()
          self.s27()
     elif(self.proxCaractereIs(self.tokens['MENOR'])):
         self.leProxCaractere()
          self.s28()
     elif(self.proxCaractereIs(self.tokens['VAZIO'])):
    self.leProxCaractere()
          self.s0()
          if(self.proxCaractere == '<EXIT>'):
             sys.exit()
          print('\nErro léxico: caractere encontrado:'+self.proxCaractere)
         print('era(m) esperados(s):')
          print(self.tokens)
         self.leProxCaractere()
         self.s0()
def s1(self):
     self.tokenReconhecido = '<INT>'
     if(self.proxCaractereIs(self.tokens['INT'])):
```

```
self.leProxCaractere()
        self.s1()
def s2(self):
    self.tokenReconhecido = '<IF>'
    if(self.proxCaractereIs(self.tokens['IF'])):
        self.leProxCaractere()
        self.s2()
def s3(self):
    self.tokenReconhecido = '<ELSE>'
    if(self.proxCaractereIs(self.tokens['ELSE'])):
        self.leProxCaractere()
        self.s3()
def s4(self):
    self.tokenReconhecido = '<WHILE>'
    if(self.proxCaractereIs(self.tokens['WHILE'])):
        self.leProxCaractere()
        self.s4()
def s5(self):
    self.tokenReconhecido = '<SWITCH>'
    if(self.proxCaractereIs(self.tokens['SWITCH'])):
        self.leProxCaractere()
        self.s5()
def s6(self):
    self.tokenReconhecido = '<CASE>'
    if(self.proxCaractereIs(self.tokens['CASE'])):
        self.leProxCaractere()
        self.s6()
def s7(self):
    self.tokenReconhecido = '<FLOAT>'
    if(self.proxCaractereIs(self.tokens['FLOAT'])):
        self.leProxCaractere()
        self.s7()
def s8(self):
```

```
self.tokenReconhecido = '<FOR>'
    if(self.proxCaractereIs(self.tokens['FOR'])):
        self.leProxCaractere()
        self.s8()
def s9(self):
    self.tokenReconhecido = '<APAR>'
    if(self.proxCaractereIs(self.tokens['APAR'])):
        self.leProxCaractere()
        self.s9()
def s10(self):
    self.tokenReconhecido = '<FPAR>'
    if(self.proxCaractereIs(self.tokens['FPAR'])):
        self.leProxCaractere()
        self.s10()
def s11(self):
    self.tokenReconhecido = '<ASPASDUP>'
    if(self.proxCaractereIs(self.tokens['ASPASDUP'])):
        self.leProxCaractere()
        self.s11()
def s12(self):
    self.tokenReconhecido = '<ASPASSIMP>'
    if(self.proxCaractereIs(self.tokens['ASPASSIMP'])):
        self.leProxCaractere()
        self.s12()
def s13(self):
    self.tokenReconhecido = '<ACHAVE>'
    if(self.proxCaractereIs(self.tokens['ACHAVE'])):
        self.leProxCaractere()
        self.s13()
def s14(self):
    self.tokenReconhecido = '<FCHAVE>'
    if(self.proxCaractereIs(self.tokens['FCHAVE'])):
```

```
self.leProxCaractere()
        self. s14()
def s15(self):
    self.tokenReconhecido = '<ACOL>'
    if(self.proxCaractereIs(self.tokens['ACOL'])):
        self.leProxCaractere()
        self.s15()
def s16(self):
    self.tokenReconhecido = '<FCOL>'
    if(self.proxCaractereIs(self.tokens['FCOL'])):
        self.leProxCaractere()
        self.s16()
def s17(self):
    self.tokenReconhecido = '<PONTO-VIRG>'
    if(self.proxCaractereIs(self.tokens['PONTO-VIRG'])):
        self.leProxCaractere()
        self.s17()
def s18(self):
    self.tokenReconhecido = '<DOIS-PONTOS>'
    if(self.proxCaractereIs(self.tokens['DOIS-PONTOS'])):
        self.leProxCaractere()
        self.s18()
def s19(self):
    self.tokenReconhecido = '<PONTO>'
    if(self.proxCaractereIs(self.tokens['PONTO'])):
        self.leProxCaractere()
        self.s19()
def s20(self):
    self.tokenReconhecido = '<SOMA>'
    if(self.proxCaractereIs(self.tokens['50MA'])):
        self.leProxCaractere()
        self.s20()
def s21(self):
```

```
self.tokenReconhecido = '<SUB>'
    if(self.proxCaractereIs(self.tokens['SUB'])):
        self.leProxCaractere()
        self.s21()
def s22(self):
    self.tokenReconhecido = '<MULT>'
    if(self.proxCaractereIs(self.tokens['MULT'])):
        self.leProxCaractere()
        self.s22()
def s23(self):
    self.tokenReconhecido = '<DIV>'
    if(self.proxCaractereIs(self.tokens['DIV'])):
        self.leProxCaractere()
        self.s23()
def s24(self):
    self.tokenReconhecido = '<RESTO>'
    if(self.proxCaractereIs(self.tokens['RESTO'])):
        self.leProxCaractere()
        self. s24()
def s25(self):
    self.tokenReconhecido = '<DIF>'
    if(self.proxCaractereIs(self.tokens['DIF'])):
        self.leProxCaractere()
        self. s25()
def s26(self):
    self.tokenReconhecido = '<IGUAL>'
    if(self.proxCaractereIs(self.tokens['IGUAL'])):
        self.leProxCaractere()
        self. s26()
def s27(self):
    self.tokenReconhecido = '<MAIOR>'
    if(self.proxCaractereIs(self.tokens['MAIOR'])):
        self.leProxCaractere()
        self.s27()
def s28(self):
    self.tokenReconhecido = '<MENOR>'
```

Exemplo de Analisador Sintático construído através de Análise Preditiva

Gramática Livre de contexto

```
1 S \rightarrow inicio.
 2 inicio -> corpo .
 3 corpo -> cmdIf | cmdWhile | cmdFor | cmdSwitch | cmdSwitchCase .
4 \exp -> < NUM> \exp B .
5 \exp B \rightarrow < SOMA > < NUM > | < SUB > < NUM > | < MULT > < NUM > | < DIV > < NUM > .
6 cmdIf -> <IF> <APAR> exp <FPAR> <ACHA> exp <FCHA> .
 7 cmdWhile -> <WHILE> <APAR> exp <FPAR> <ACHA> exp <FCHA> .
8 cmdFor -> <FOR> <APAR> exp <FPAR> <ACHA> exp <FCHA> .
9 cmdSwitch -> <SWITCH> <APAR> exp <FPAR> <ACHA> listCmdSwitchCase <FCHA> .
10 listCmdSwitchCase -> cmdSwitchCase listCmdSwitchCase .
11 cmdSwitchCase -> <CASE> <DOIS-PONTOS> exp .
12 <SOMA> -> + .
13 <SUB> -> - .
14 <MULT> -> * .
15 <DIV> -> / .
16 <NUM> -> 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 .
```

Teste do analisador sintático:

Teste 1:

```
if(1+1 == 2)
{
}
```

Console de saída:

Análise realizada com sucesso no arquivo file.txt

Teste 2:

```
while(1*2 != 3)
{
}
```

Console de saída:

Análise realizada com sucesso no arquivo file.txt

Teste 3:

```
if*1+1 < 4)
{
}
while(1+1 >= 2)
{
}
```

Console de saída:

Erro sintático: token encontrado: <MULT> era(m) esperados(s): <APAR>

Teste 4:

```
else=
```

Console de saída:

Erro sintático: token encontrado: <IGUAL> era(m) esperados(s): <ACHAR>

Implementação por intermédio da análise preditiva:

É importante salientar, que o exemplo abaixo é um complemento da implementação acima, ou seja, para executar o código abaixo, você só precisa juntá-lo ao código acima.

```
v class AnalisadorSintatico(Analisador):
       def __init__(self, _nomeArquivoEntrada):
           self.t = []
           self.A = MyAnalisadorLexico(_nomeArquivoEntrada)
           super().__init__()
           self.leProxToken()
       def leProxToken(self):
           self.A.s0()
      def reconhece(self, t):
           if self.A.tokenReconhecido in t:
                self.leProxToken()
                print('\nErro Sintático: token encontrado:' + self.A.tokenReconhecido)
                print('era(m) esperados(s):' , end = '')
                print(t)
       def proxTokenIs(self, t):
           if self.A.tokenReconhecido in t:
               return True
                return False
v class MyAnalisadorSintatico(AnalisadorSintatico):
      def __init__(self,_nomeArquivoEntrada):
            self.A = AnalisadorSintatico(_nomeArquivoEntrada)
            super().__init__(_nomeArquivoEntrada)
      def inicio(self):
           self.corpo()
      def corpo(self):
           if (self.proxTokenIs(self.tokens['IF'])):
               self.leProxToken()
                self.cmdIf()
                self.corpo()
           elif (self.proxTokenIs(self.tokens['WHILE'])):
    self.leProxToken()
                self.cmdWhile()
                self.corpo()
           elif (self.proxTokenIs(self.tokens['FOR'])):
    self.leProxToken()
                self.cmdFor()
                self.corpo()
           elif (self.proxTokenIs(self.tokens['SWITCH'])):
    self.leProxToken()
                self.cmdSwitch()
self.corpo()
           elif (self.proxTokenIs(self.A.tokens['EOF'])):
    print('\nAnalise Sintática: Concluída')
               print('\nErro Sintático: token encontrado:'+ self.A.tokenReconhecido)
```

```
def exp(self):
    if (self.proxTokenIs(self.tokens['INT'])):
         self.leProxToken()
         if (self.proxTokenIs(self.tokens['SOMA'])):
    self.leProxToken()
    self.exp()
         elif (self.proxTokenIs(self.tokens['SUB'])):
    self.leProxToken()
             self.exp()
         elif (self.proxTokenIs(self.tokens['MULT'])):
             self.leProxToken()
              self.exp()
         elif (self.proxTokenIs(self.tokens['DIV'])):
    self.leProxToken()
             self.exp()
         elif (self.proxTokenIs(self.tokens['RESTO'])):
             self.leProxToken()
self.exp()
         elif (self.proxTokenIs(self.tokens['DIF'])):
             self.leProxToken()
             self.exp()
         elif (self.proxTokenIs(self.tokens['IGUAL'])):
             self.leProxToken()
             self.exp()
         elif (self.proxTokenIs(self.tokens['MAIOR'])):
    self.leProxToken()
              self.exp()
         elif (self.proxTokenIs(self.tokens['MENOR'])):
    self.leProxToken()
             self.exp()
    elif (self.proxTokenIs(self.tokens['INT'])):
         self.leProxToken()
    elif (self.proxTokenIs(self.tokens['VAZIO'])):
         self.leProxToken()
         print('\nErro Sintático: token encontrado:'+ self.A.tokenReconhecido)
         print('era(m) esperados(s):', end = '')
         print( self.t)
def bloco(self):
    self.exp()
def cmdIf(self):
     self.reconhece(self.tokens['APAR'])
     self.exp()
    self.reconhece(self.tokens['FPAR'])
    self.reconhece(self.tokens['ACHAVE'])
    self.bloco()
     self.reconhece(self.tokens['FCHAVE'])
def cmdWhile(self):
    self.reconhece(self.tokens['APAR'])
    self.exp()
```

```
self.reconhece(self.tokens['APAR'])
                self.exp()
                self.reconhece(self.tokens['FPAR'])
                self.reconhece(self.tokens['ACHAVE'])
                self.bloco()
                self.reconhece(self.tokens['FCHAVE'])
            def cmdWhile(self):
                self.reconhece(self.tokens['APAR'])
                self.exp()
                self.reconhece(self.tokens['FPAR'])
                self.reconhece(self.tokens['ACHAVE'])
                self.bloco()
                self.reconhece(self.tokens['FCHAVE'])
            def cmdFor(self):
                self.reconhece(self.tokens['APAR'])
                self.exp()
                self.reconhece(self.tokens['FPAR'])
                self.reconhece(self.tokens['ACHAVE'])
                self.bloco()
                self.reconhece(self.tokens['FCHAVE'])
            def cmdSwitch(self):
                self.reconhece(self.tokens['APAR'])
                self.exp()
                self.reconhece(self.tokens['FPAR'])
                self.reconhece(self.tokens['ACHAVE'])
                self.listCmdSwitchCase()
                self.reconhece(self.tokens['FCHAVE'])
            def listCmdSwitchCase(self):
                self.cmdSwitchCase()
                self.listCmdSwitchCase()
            def cmdSwitchCase(self):
629
                self.reconhece(self.tokens['CASE'])
self.reconhece(self.tokens['INT'])
                self.reconhece(self.tokens['DOIS-PONTOS'])
                self.bloco()
        #A = MyAnalisadorLexico('file.txt')
        #while True:
              print(A.tokenReconhecido, end='')
              if A.proxCaractere == '<EXIT>':
        #
                  break
        tst = MyAnalisadorSintatico('file.txt')
        tst.inicio()
        print('Analise conclida com sucesso')
```