

Construção de Compiladores 1 - 2015.1 - Prof. Daniel Lucrédio

Lista 02 - Análise Léxica

1. Identifique, em uma tabela, todos os tokens que compõem os programas seguintes. Cada linha da tabela será um token. A tabela terá 3 colunas: na coluna 1, especifique a cadeia (lexema) correspondente; na coluna 2, especifique sua classe (identificador, palavra ou símbolo reservado, número, comentário, etc.); e na coluna 3 especifique o padrão utilizado no reconhecimento (em português mesmo).

a) Pascal

```
function max(i, j: integer): integer;
{ retorna o maior dos inteiros entre i e j }
begin
if i > j then max := i
else max := j
end;
```

b) C

```
int max(i, j) int i, j;
/* retorna o maior dos inteiros entre i e j */
{
return i > j ? i : j
}
```

c) Em qual dos dois programas apresentados nas letras acima (a e b) foram identificados mais tokens?

2. Faça o papel do analisador léxico e “quebre” os seguintes programas, escritos na linguagem ALGUMA, em um fluxo de tokens, no formato <tipo,valor>, onde tipo é um dos tipos da linguagem ALGUMA, e valor aponta para a tabela de símbolos, caso necessário. Monte também a tabela de símbolos.

a)

```
:DECLARACOES
numero1:INT
numero2:INT
numero3:INT
aux:INT

:ALGORITMO
% Coloca 3 números em ordem crescente
LER numero1
LER numero2
LER numero3
SE numero1 > numero2 ENTAO
    INICIO
        ATRIBUIR numero2 A aux
        ATRIBUIR numero1 A numero2
        ATRIBUIR aux A numero1
```

```

    FIM
SE numero1 > numero3 ENTAO
    INICIO
    ATRIBUIR numero3 A aux
    ATRIBUIR numero1 A numero3
    ATRIBUIR aux A numero1
    FIM
SE numero2 > numero3 ENTAO
    INICIO
    ATRIBUIR numero3 A aux
    ATRIBUIR numero2 A numero3
    ATRIBUIR aux A numero2
    FIM
IMPRIMIR numero1
IMPRIMIR numero2
IMPRIMIR numero3

```

b)

```

:DECLARACOES
numero:INT
potencia2:INT
potencia3:INT

:ALGORITMO
% Ler um numero
LER numero
SE numero = 0 ENTAO % zero elevado a qualquer coisa é zero
    INICIO
    ATRIBUIR 0 A potencia2
    ATRIBUIR 0 A potencia3
    FIM
SENAO SE numero = 1 ENTAO % um elevado a qualquer coisa é um
    INICIO
    ATRIBUIR 1 A potencia2
    ATRIBUIR 1 A potencia3
    FIM
SENAO INICIO
    ATRIBUIR numero * numero A potencia2
    ATRIBUIR numero * (numero * numero) A potencia33
FIM
% Mostrar resultados
IMPRIMIR numero
IMPRIMIR ' ao quadrado é igual a '
IMPRIMIR potencia3
IMPRIMIR '\n'
IMPRIMIR numero
IMPRIMIR ' ao cubo é igual a '
IMPRIMIR potencia3
IMPRIMIR '\n'

```

c) Os programas anteriores possuem algum erro léxico? Se sim, qual?

3. Analise o código a seguir, na linguagem ALGUMA, e aponte os erros léxicos, se houverem (Obs: desconsidere os números de linhas, são apenas para sua referência)

```
1.  :ALGORITMO
2.  // Ler um numero
3.  LER 123numero
4.  LER numero345
5.  LER ENTAO SENAO numero == 0
6.  FIM
7.  SE var1 ><=> 22 ENTAO
8.    SE var2 != 33 ENTAO
9.    SENAO
10. INICIOFIM
11. INICIO FIM
12. %%%%%%%%%%%%%%% Mostrar resultados
13.
14. :DECLARACOES
15. 123numero:INT
16. Numero345=INT
17. Algoritmo && FLOAT
18. :ALGORITMO
19. ATRIBUIR 123numero A 123numero
```

4. Quais são os motivos para se separar conceitualmente a análise léxica da sintática? Explique cada um dos motivos.

5. Qual o papel do buffer duplo?

6. Qual a vantagem dos geradores de analisadores léxicos sobre os analisadores construídos à mão?

7. Escreva expressões regulares para os conjuntos de caracteres a seguir ou se não for possível escrever uma expressão regular para um determinado conjunto de caracteres, justifique.

- a) Cadeias de letras maiúsculas começando e terminando com a (minúsculo).
- b) Cadeias de dígitos que representam números pares.
- c) Cadeias de 0s e 1s com um número par de 0s.
- d) Cadeias de 0s e 1s nas quais os 0s ocorrem em pares (um 0 seguido de outro 0).
- e) Cadeias de 0s e 1s compostas por um único 1 rodeado pelo mesmo número de 0s à esquerda e à direita.
- f) Cadeias de dígitos tais que todos os dígitos ímpares, se ocorrerem, ocorrem antes de todos os dígitos pares (se ocorrerem).
- g) Todas as cadeias de letras minúsculas que contêm as cinco vogais em ordem.
- h) Comentários, consistindo em uma cadeia cercada por /* e */, sem um /* intercalado
- i) Todas as cadeias de as e bs que não contêm a subsequência abb.
- j) Endereços IP
- k) Endereços de email
- l) Datas, no formato dd/mm/aaaa
- m) Números reais com notação científica (ex: 10.4E13)

8. Desenhe um diagrama de transições para cada expressão regular do exercício 7. Cuidado com a necessidade de retroceder caracteres (Símbolo *).

9. Escreva um pseudocódigo (algoritmo) para cada diagrama de transição do exercício 8.