

Atividade – Data entrega 23/09/2024.

Aula 3 – Análise Léxica (Implementação de um Autômato)

Observações:

- A atividade pode ser realizada individual ou em duplas.
- A linguagem escolhida para a implementação fica a critério do(s) desenvolvedor(es).
- A Atividade é parte integrante da GA (02 pontos).
- O trabalho será apresentado no dia 23 de setembro de 2024.
- Deverá ser entregue os códigos-fontes devidamente comentados, na mesma data da apresentação no moodle, Bom trabalho!

Etapa 1 - Implementação

Implemente um autômato para reconhecer linguagens cujos tokens podem ser:

- Identificadores iniciados por uma letra, podendo possuir na sequência números e/ou letras.
- Constantes numéricas formadas por um ou mais números inteiros na casa da dezena, ou seja, até o valor 99.
- Constantes numéricas formadas por números reais, também na casa da dezena, ou seja, valor máximo de 99.99 (identificadas pelo ponto). Sempre com duas casas decimais.
- Identificadores no formato de comentários de linha, padrão da Linguagem C (//).
- As palavras reservadas da linguagem C: `int`, `double`, `float`, `real`, `break`, `case`, `char`, `const`, `continue`.

Etapa 2 – Carregamento

O analisador deverá carregar um arquivo-texto contendo um padrão por linha e reconhecer o token especificado. Sendo que o arquivo, deverá ser especificado pelo usuário no momento da execução do programa.

Etapa 3 - Apresentação

Ao final da análise, o autômato deverá exibir:

- Os tokens de entrada (e a linha onde eles aparecem – tabela léxica).
- A tabela de símbolos.
- A lista das linhas onde os erros aparecem (caso tenham erros no arquivo), bem como a palavra, não sendo especificado qual erro.

Por exemplo, se o usuário entrar com o arquivo contendo os dados abaixo:

```
a
a
int
asd
as123
99
99.
99.99
```

```
float
real a
real
double _c
double
  as_asd
asd
// Joao da Silva
-- Teste
Joao da Silva --
char
```

O programa deverá emitir o seguinte arquivo de saída e apresentar na tela, os itens:

Tokens de Entrada

- [1] IDENTIFICADOR 1
- [2] IDENTIFICADOR 1
- [3] INT
- [4] IDENTIFICADOR 2
- [5] IDENTIFICADOR 3
- [6] NÚMERO INTEIRO 4
- [8] NÚMERO REAL 5
- [9] FLOAT
- [11] REAL
- [13] DOUBLE
- [15] IDENTIFICADOR 2
- [16] COMENTÁRIO
- [19] CHAR

Tabela de Símbolos

- 1 - a
- 2 - asd
- 3 - as123
- 4 - 99
- 5 - 99.99

Erros nas linhas:

- 7 (99.)
- 10 (real a)
- 12 (double _c)
- 14 (as_asd)
- 17 (-- Teste)
- 18 (Joao da Silva)