

## Disciplina: Construção de Compiladores

### Projeto: Desenvolvimento de um Compilador

#### Parte I: Analisador Léxico (SCANNER)

**Sílvia Bandeira**

Professor Adjunto

Escrever, em C/C++ ou Java, a função SCAN para a linguagem abaixo:

**Observação 1:** o arquivo a ser compilado será passado ao seu compilador **via argumento da linha de comando**

**Observação 2:** Imprimir **apenas mensagens de erro**.

**Observação 3:** A mensagem deve ser **clara e específica** de erro, sempre que for o caso, **e em qualquer fase do compilador**. Formato: "ERRO na linha *n*, coluna *m*, ultimo token lido *t*: *mensagem específica do erro*"

Símbolos

letra ::= [a-z]

dígito ::= [0-9]

id ::= (letra | "\_" ) (letra | "\_" | dígito)\*

oprelacional ::= < | > | <= | >= | == | !=

oparitmético ::= "+" | "-" | "\*" | "/" | "="

especial ::= ")" | "(" | "{" | "}" | "," | ";"

palreservada ::= main | if | else | while | do | for | int | float | char

inteiro ::= dígito+

float ::= dígito\*.dígito+

char ::= 'letra' | 'dígito' // Uma constante do tipo **char** (entre aspas simples)

## Aspectos Gerais da Linguagem

- Linguagem de formato livre.
- Linguagem é *case sensitive*, ou seja “WHILE” != “while”
- As palavras reservadas são delimitadas, no programa fonte, por brancos, operadores aritméticos ou símbolos especiais;
- Os comentários são delimitados por
  - “//” - indicando comentário até o final da linha
  - “/\*” e “\*/”

## Atribuições do Scanner

- Retornar dois resultados: classificação e *lexema*
- No caso de o *token* ser um identificador ou uma constante devem ser retornados classificação e *lexema*
- Se for uma palavra reservada, operador, caracter especial ou delimitador: classificação apenas
- Emitir mensagem **clara e específica** de erro, sempre que for o caso, **e em qualquer fase do compilador**. Formato: "ERRO na linha *n*, coluna *m*, ultimo token lido *t*: *mensagem*"
  - Considere o TAB como equivalente a 4 colunas

## Procedimento:

- Adotar uma representação interna para cada tipo de *token*
  - Lembre-se que, por exemplo, "+" é diferente de "-", portanto não podem ter mesma representação
- Organizar a tabela de palavras reservadas
- Construir um diagrama de estados, com ações semânticas, para a identificação dos símbolos
- Escrever a função SCAN.