

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL  
FACOM - FACULDADE DE COMPUTAÇÃO

COMPILADORES I - 2023/2  
PROFA. BIANCA DE ALMEIDA DANTAS

**Trabalho Prático**  
**Descrição da 1<sup>a</sup> Etapa – Análise Léxica**

## 1 DESCRIÇÃO

A primeira parte do trabalho prático de nossa disciplina consiste na implementação, usando a linguagem C++, do analisador léxico para a linguagem hipotética X++ descrita no livro “Como Construir um Compilador Utilizando Ferramentas Java” de Márcio Eduardo Delamaro. O analisador léxico deve ser capaz de reconhecer os *tokens* da linguagem e **ignorar espaços em branco e comentários**.

**Tokens:**

1. Identificadores (ID): Uma sequência de letras, dígitos e *underscores*, começando sempre com uma letra ou com *underscore*. A linguagem distingue letras maiúsculas e minúsculas;
2. Números inteiros (INTEGER\_LITERAL): uma sequência de dígitos decimais, que denotam o número inteiro correspondente;
3. Operadores (OP): <, >, <=, >=, +, -, \*, /, %, =, == e !=;
4. Separadores (SEP): (, ), [, ], {, }, ;, . e ,;
5. Literais do tipo *string* são delimitados por “e”.
6. Comentários: os comentários devem ser ignorados. Existem dois tipos: comentário de linha (iniciado com // e indo até o final da linha) e comentário de bloco (iniciado com /\* e encerrado em \*/ sem aninhamentos);
7. Espaços em branco: devem ser ignorados.

Observe que as palavras reservadas da linguagem, nesta etapa, serão reconhecidas como **identificadores**. Assim que estudarmos a estrutura da tabela de símbolos, iremos usá-la como auxiliar para diferenciar as palavras reservadas dos identificadores. As palavras reservadas são as seguintes: **class, extends, int, string, break, print, read, return, super, if, else, for, new, constructor**.

## 2 ESPECIFICAÇÕES

- O trabalho prático poderá ser realizado em grupos de, no máximo, 3 alunos **sem exceções**.
- A linguagem C++ deverá ser utilizada na implementação do trabalho.
- A entrega de todas as etapas deve ser realizada até o dia: **17/11/2023**.