



### Regras para Entrega:

1. A data da entrega é exclusivamente através do sistema de atividades práticas, a partir de 10/06/2023;
2. A nota do trabalho considera:
  1. **Legibilidade do código:** código indentado, com nomes de variáveis simples e significativos e devidamente comentado. Nos comentários iniciais deve constar o nome dos autores.
  2. **Correção do código:** o código deve ser em python, ou notebook python e funcionar sem alterações e sem erros.
  3. **Originalidade do código:** cópia de código implica nota zero para todas as cópias.
  4. **Pontualidade da entrega:** atrasos na entrega acarretam descontos cumulativos na nota, assim, entrega em 10/06/2023, após 23:55 acarreta desconto de 1 ponto na nota do trabalho, nos dias seguintes, mais 3 pontos de desconto por dia de atraso.
3. Os grupos para execução do trabalho devem ter no máximo 06 (seis) alunos.

### Cenário:

Em um compilador, a análise do código fonte é separada em duas partes: análise léxica e análise sintática.

A análise léxica tem a função de isolar as palavras chaves, símbolos especiais, símbolos especiais compostos, identificadores e constantes, transformando-os em códigos convenientes para a análise sintática.

Os tokens são sequências de caracteres que representam algum símbolo terminal da linguagem. No caso do nosso Pascal simplificado<sup>1</sup> (mini-pascal), são exemplos de tokens:

- begin, end, +, -, \*, /, :=, valorTotal, 1234

O analisador léxico, ou scanner, processa o programa fonte, caracter por caracter, retornando os códigos dos tokens encontrados. Para os tokens que são identificadores ou números, além do código, o scanner devolve a própria sequência de caracteres, ou o valor do número na representação interna, produzindo uma tabela de símbolos.

A tabela de símbolos é uma estrutura em forma de lista encadeada, hash, ou árvore balanceada que armazena as informações dos tokens de identificadores e números para uso na análise sintática.

A estrutura genérica da tabela de símbolos é mostrada na Tabela 1.

Tabela 1: Estrutura Genérica da Tabela de Símbolos

Nome do Token	Tipo do Token	Parâmetros do Token
Nome de variável, função ou procedure	Ver especificação da gramática	Lista de parâmetros necessários definida pelo projetista

As palavras reservadas da linguagem como nomes de comandos e de especificação de tipos também são armazenadas em uma estrutura tipo lista encadeada, hash, ou árvore balanceada para facilitar a pesquisa e identificação.

### Problema:

Implemente um analisador léxico que analise códigos fonte em Pascal simplificado e produza a tabela de símbolos para o analisador sintático.

1. A documentação do projeto deve incluir as hipóteses, o desenvolvimento e explicações necessárias ao entendimento do analisador léxico desenvolvido.
2. A forma da implementação da tabela sintática e da lista de palavras reservadas fica a critério da equipe e a documentação do projeto deve conter as justificativas da escolha.
3. As saídas do analisador léxico devem indicar os erros, sua posição no código no formato (linha, coluna) e, sempre que possível, uma sugestão de solução.

---

<sup>1</sup> A especificação do Pascal simplificado está disponível na página da disciplina, no moodle.