**Roteiro para iniciação do analisador léxico.**

**1º Passo: Pesquisa sobre Geradores de Analisadores Léxicos/Sintáticos em Python**

Python tem várias ferramentas para análise léxica e sintática. Uma das mais conhecidas é o **PLY** (Python Lex-Yacc), que é uma implementação pura do Python das ferramentas populares **Lex** e **Yacc** encontradas em linguagens como **C**. **PLY** usa o mesmo princípio, mas é escrito em Python e se integra bem com a linguagem. Outras opções incluem **SLY** (*Successor of PLY*), que é mais moderna e usa as funcionalidades mais recentes do *Python*.

**2º Passo: Executar e Analisar um Exemplo**

Uma vez que escolha a ferramenta (como o PLY), você pode baixar um exemplo de código fonte, como uma simples análise de expressões matemáticas (evitando calculadoras), e estudar como o código está estruturado. Entenda as regras léxicas que identificam os tokens e as regras sintáticas que definem como esses tokens são combinados para formar expressões válidas.

**3º Passo: Modificar o Exemplo**

Para criar um sistema de matrícula de alunos, você precisará definir tokens para representar nomes, identificadores de cursos, números de matrícula, etc. Você também precisará definir uma gramática que possa interpretar comandos como matricular um aluno, desmatricular, listar alunos em um curso, etc.

Definições e Regras Léxicas:

Tokens podem incluir palavras-chave como **MATRICULAR**, **DESMATRICULAR**, **CURSO**, **ALUNO**, números e strings para nomes.

Expressões regulares seriam criadas para identificar esses tokens. Por exemplo, um nome pode ser uma sequência de caracteres alfabéticos com espaço, um identificador de curso pode ser uma combinação de letras e números, etc.

Regras Gramaticais:

* A gramática para os comandos pode parecer com algo assim:

comando: **MATRICULAR ALUNO em CURSO**

**|DESMATRICULAR ALUNO de CURSO**

**|LISTAR ALUNOS em CURSO**

* Precedência e associação podem ser definidas se necessário para resolver ambiguidades.

Ações Semânticas:

* Nas ações, você definiria o que acontece quando um comando é reconhecido. Por exemplo, ao matricular um aluno, você adicionaria o aluno a uma lista de alunos matriculados no curso correspondente.

Exemplo de Árvore de Derivação:

Para o comando MATRICULAR "João Silva" em "COMP101", você mostraria uma árvore com a raiz sendo o comando, se ramificando para MATRICULAR, ALUNO, em, CURSO.

Exemplo Executando o Código Modificado:

Você executaria o código para mostrar a matrícula e desmatrícula de alunos, etc. Aqui você mostraria o código em execução através de testes unitários ou scripts que demonstram a funcionalidade.

Principais Dificuldades Encontradas:

Durante o desenvolvimento, você pode encontrar dificuldades como resolver ambiguidades na gramática, garantir que as ações semânticas estejam corretas e integrar o analisador com um banco de dados ou uma estrutura de dados para armazenar as informações dos alunos.

Para começar a codificar, você primeiro decidiria sobre a ferramenta (por exemplo, PLY) e começaria a definir os tokens e a gramática. A partir daí, você desenvolveria as ações semânticas correspondentes. Você gostaria de começar com um exemplo de código para definir tokens usando PLY ou outra ferramenta de sua preferência?