



Análise Léxica

Análise Léxica: 1ª fase de compilação que tem o papel de ler o arquivo fonte (texto-fonte de entrada) em busca de unidades significativas (os *tokens*) instanciadas por lexemas ou átomos.

Também denominado de scanner [AHO, 86], porque varre o arquivo de entrada eliminando comentários e caracteres indesejáveis ao agrupar caracteres com um papel bem definido.



"token", padrão e lexemas"

Em geral, existe um conjunto de cadeias de entrada para os quais o mesmo *token* é produzido como saída.

Esse conjunto de cadeias é descrito por uma regra chamada de um *padrão* associado ao *token* de entrada. O padrão é dito *reconhecer* cada cadeia do conjunto.

Um *lexema* é um conjunto de caracteres no texto-fonte que é reconhecido pelo padrão de algum token [AHO, 86].

Por ex em Pascal: `const pi = 3.1416`, a subcadeia `pi` é um lexema para o token "identificador"



Tabela de Símbolos

A **tabela de símbolos** é utilizada para armazenar as cadeias de caracteres que formam o identificador, e um índice passa a ser utilizado como código de comprimento fixo que represente tal identificador para efeito de análise por parte dos outros módulos do compilador [JOSÉ NETO, 87].



Átomos

Com a tabela de símbolos torna-se possível o mapeamento da cadeia de entrada, formada pelo texto-fonte, em uma cadeia de pares de códigos, de comprimento fixo, cada qual responsável por simbolizar uma das partículas em que o texto-fonte foi decomposto.

A tais pares de códigos, formados pelas informações de classe e de valor, associados à partícula que representam, dá-se o nome de *átomos* [JOSÉ NETO, 87]. Estes representam os símbolos terminais da gramática.



Funções do Analisador Léxico

São funções do Analisador Léxico [JOSÉ NETO, 87]:

- extração e classificação de átomos;
- eliminação de delimitadores e comentários;
- conversão numérica;
- tratamento de identificadores;
- identificação de palavras reservadas;
- iteração com o sistema de arquivos.



Tarefas Adicionais do A. Léxico

São tarefas adicionais do Analisador Léxico:

- atividades ligadas à geração de listagens do texto-fonte;
- recuperação de erros;
- eventual manipulação das tabelas de símbolos.