

Compiladores

Léxico

José Luis Seixas Junior

Índice

- Funções do Léxico.
- Entrada.
- Saída:
- Vantagens:

Função do Léxico

- Extração e classificações de átomos;
- Eliminação de delimitadores e comentários;
- Conversão numérica;
- Tratamento de Identificadores;
- Identificação de Palavras Reservadas;
- Interação com o sistema de arquivos;

Função do Léxico

- Eliminação de dependências de identificação:
 - Palavras reservadas serão sempre comandos e, portanto, não poderão ser usadas com identificadores;
 - Espaços, comentários, quebras de linhas;

Análise Léxica

- 1º fase de compilação:
 - Ler o Arquivo de Entrada;
 - Buscar Unidades significativas;
 - Agrupar caracteres com um papel bem definido;
 - Classificar as unidades.

Análise Léxica

- Nomenclaturas:
 - Token: Unidade significativa:
 - Lexema: Conjunto de caracteres no texto que é reconhecido pelo padrão de algum token.
 - Exemplo:
 - Const pi = 3.1415;
→ A subcadeia pi é um Lexema para o token “identificador”;

Análise Léxica

- Agrupamento:
- Linha Exemplo: `const pi = 3.1415;`
 - Token: `3` → Numérico
 - Token: `.` → Símbolo simples
 - Token: `1415` → Numérico
 - Lexema: `3.1415` → Numérico “float”
 - Não necessariamente parte do léxico, mas facilita possíveis análises;

Análise Léxica

- Padrões:
 - Sequências de caracteres que resolvem o mesmo token;
 - Toda entrada do AFD que aceitaria um “identificador”.
 - Podem ser geradas divisões claras do AFD para expressões diferentes;

Entrada

- Arquivo .txt ou .pas;
- Código Pascal simplificado;
- Funções descritas no slide 01!
- Blocos:
 - begin
 - end
 - Exceto comentários.

Saída

- Tabela de Simbolos:
 - Elementos da gramática;
 - Passam as ser usados uma constante de expressão;
 - Numérico; Palavra Reservada; Identificador;
 - Normalmente tipo é usado fator numérico pela facilidade de leitura;
 - Identificador: 0;
 - Palavra Reservada: 1;
 - Numérico: 2;

Saída

- Átomos:
 - Mapeamento da cadeia de entrada em forma de códigos de comprimento fixo;
 - Simbolização de partículas de como o texto é composto;
 - Símbolos TERMINAIS da gramática;

Vantagens

- Elimina a necessidade dos analisadores sintático ou semântico de tratarem tabelas de símbolos;
- Toda função relacionada a terminais será concentrada em uma única parte do compilador;
- Formação de átomos são mais simples que as restantes, permitindo o uso de técnicas de análises mais simples e eficientes;

Vantagens

- Manutenção da clareza conceitual da análise sintática;
- Facilitação da manutenibilidade, pois erros por problemas léxicos serão tratados separadamente;
- Facilitação de otimizações;
- Permite a utilização de linguagens de montagem;

Referências

- AHO, A. V.; SETHI, R.; ULLMAN, J. D.; *Compiladores: Princípios, técnicas e ferramentas*, LTC - Livros Técnicos e Científicos Editora, 1995.
- KOWALTOSKI, T.; *Implementação de Linguagens de Programação*, Editora Guanabara Dois, 1983.