

RELATÓRIO 2 – Estrutura das Linguagens

PARTE 1:

Um gerador de gramática de linguagem pode gerar todas as cadeias de texto possíveis que correspondem a um padrão determinado. Em algumas linguagens, expressões regulares são usadas para ver se uma string que você possui pode combinar um padrão. Uma expressão regular é uma maneira concisa de definir um padrão de texto. Na ciência da computação, um padrão de expressão regular define uma gramática. Todas as palavras, que correspondem ao padrão, são ditas no "idioma" definido pela gramática. O analisador léxico, conhecido como lexer, converte uma sequência de caracteres em uma sequência de tokens. O token é uma string que possui significado. Já o parser é o analisador sintático que converte uma lista de tokens em uma estrutura de dados.

YACC (Yet Another Compiler Compiler) é um analisador sintático desenvolvido por Stephen C. Johnson da AT&T para o sistema operacional Unix. Ele gera um analisador sintático, parte do compilador responsável por fornecer sentido sintático a um determinado código fonte, baseado numa gramática formal. O resultado é um código para o analisador sintático em C. O yacc costumava ser o gerador de analisador sintático padrão na maioria dos sistemas Unix, mas acabou sendo suplantado por versões mais modernas ainda que compatíveis.

GNU BISON é um software livre para compilador escrito para o projeto GNU, e disponível para quase todos os sistemas operacionais. É compatível com o yacc, e oferece muitas melhorias se comparado com o antigo software. Ele é utilizado em conjunto com analisadores flex.

LEX é um programa que gera analisadores léxicos. Ele é geralmente usado com o yacc. Escrito originalmente por Eric Schmitd e Mike Lesk, ele é o gerador de analisador léxico padrão em diversos sistemas Unix. O lex lê um fluxo de entrada especificando um analisador que mapeia expressões regulares em blocos de código, e retorna um código fonte implementando o analisador. Apesar do gerador ser genérico e poder se adequar a diferentes linguagens de programação, atualmente, somente a geração de código C é suportada.

FLEX (the fast lexical analyzer generator) software desenvolvido na linguagem C como uma alternativa para o lex. Tem a mesma finalidade do lex. Apesar de possuir código aberto, o flex não faz parte do projeto GNU.

PARTE 2:

O arquivo scan.l define a estrutura léxica do programa e está dividido em três partes:

Definição: ocorre a definição do alfabeto, tendo incluso a standard do C e o arquivo y.tab.h, gerado por um analisador léxico.

Regras: liga o alfabeto com os tipos.

Código em C: comportamento do programa.

O arquivo gram.y define sintaticamente o programa e é dividido em três partes:

Tokens: os tipos da linguagem.

Sintaxe: como devem ser as expressões presentes na linguagem.

Código em C: comportamento do programa.