

TRABALHO 2 - COMPILADOR PARCIAL USANDO A BIBLIOTECA PLY

O código desenvolvido começa pela definição de tokens e palavras reservadas que o analisador léxico deve validar no código fornecido como input. Alguns dos tokens foram definidos usando-se expressões regulares, como para identificar nomes de variáveis ou números inteiros e de ponto flutuante. O analisador léxico “*lex*” então roda em cima do código de entrada identificando se todas sequências de caracteres correspondem a algum dos tokens ou a uma palavra reservada.

Depois disso, começa a fase de análise sintática e semântica usando-se o analisador sintático “*yacc*”. A análise sintática usa sequências de tokens e palavras reservadas, identificadas na análise léxica, para saber o que cada parte do código faz, as regras de gramática. Se analisa, também, se estas partes do código estão escritas de forma que possa ser entendível, de acordo com a linguagem implementada no compilador, o que deve acontecer nestas partes do código. A primeira sequência a ser identificada é a da função principal (*main*), que deve possuir um bloco de código. Este bloco de código tem expressões, que podem ser declarações e atribuições de valores à variáveis, estruturas condicionais (*if-e/se*) e laços de repetição (*for*, *while* e *do-while*). Cada uma destas estruturas contém, por si só, outros blocos de código. Além disso, implementei “grupos” de tokens para facilitar a escrita das regras de gramática, sendo eles: “tipos” (*char*, *float*, *double*...); “atribuicoes” (=, +=, -= e <<); “operacoes” (+, -, *...); “comparador” (==, <, >, >=...); e, “literal” (números e booleanos).

Na verificação da sintaxe das declarações de variáveis e atribuições de valores a elas, é realizada a ação semântica de preencher a tabela de símbolos do compilador com o valor atribuído a cada variável, o seu tipo e a qual contexto ela pertence. Quanto ao contexto, ele começa com o valor 0, para a função principal, e a cada vez um bloco de código é feito, o valor do contexto sobe em 1, e a cada vez que o bloco é totalmente reconhecido, o valor do contexto diminui em 1.