

Relatório 1ª entrega - Lexer + Parser

Linguagem C

C é uma linguagem de programação compilada de propósito geral, estruturada, imperativa, procedural, padronizada pela Organização Internacional para Padronização (ISO), criada em 1972 por Dennis Ritchie na empresa AT&T Bell Labs para desenvolvimento do sistema operacional Unix (originalmente escrito em Assembly).

C é uma das linguagens de programação mais populares e existem poucas arquiteturas para as quais não existem compiladores para C. C tem influenciado muitas outras linguagens de programação (por exemplo, a linguagem Java), mais notavelmente C++, que originalmente começou como uma extensão para C.

Usamos a ferramenta [ANTLR](#) que é implementada em JAVA e cria automaticamente o lexer e o parser do compilador diante da gramática representada em um arquivo `.g4`. Como já existem muitas referências na internet sobre gramáticas de linguagens de programação, em particular de C que é bastante famoso, fizemos o uso deste [repositório](#).

Desenvolvimento

Ao utilizar o ANTLR nos deparamos com muitos erros ao tentar reconhecer as entradas pois quando pegamos a gramática do C no gitHub, o autor não indicou qual era a regra inicial da gramática para que o ANTLR iniciasse a varredura dos tokens. Esse ponto específico nos gastou boas horas.

Outra dificuldade foi conseguir identificar quando havia um erro no parser, pois o antlr faz toda a leitura e no final devolve os tokens sem uma mensagem muito clara de erro. Nós conseguimos alterar essa detecção e o programa para com uma mensagem ao se deparar com algum erro na entrada.

Testes

Executamos testes sobre if else aninhados, alguns if's sem else e if com apenas uma linha de comando. Sobre parenteses encaixados (exceto o erro relatado acima). Nomes de variáveis, falta de ';', operações aritméticas sem um dos operandos, ou operando incorreto. Falta de chaves. Reconhecimento de estruturas e palavras reservadas da linguagem.

Até onde o grupo conseguiu cobrir, o analisador léxico/sintático está correto, conseguindo indicar os erros e reconhecer entradas corretas.