

ANTLR - Another Tool For Language Recognition

Discentes:

Fernando Henrique Callata, Murillo Gordo, Marcos Rafael Lapa de Sousa,
Saulo Calixto.



Objetivos:

- 1 - Introduzir o conceito de parser e a ferramenta
- 2 - Noções de como criar uma gramática formal
- 3 - Uso prático do ANTLR

Introdução



Parsing

- análise sintática
- transforma um texto na entrada em uma estrutura de dados
- mais fácil e menos limitado em relação ao uso de expressões regulares



Ferramenta ANTLR

- Gerador de parser.
- Produtividade maior do que criar seu próprio parser na mão, pode focar em atingir seus objetivos ao invés de detalhes de implementação do parser.
- Utilizado em muitos softwares como Twitter em seu motor de busca, o Hibernate, Netbeans, etc.

Gramática



Gramática formal

- uma gramática formal é um conjunto de regras de formação de cadeias, estas regras descrevem como formar cadeias, que são válidas de acordo com a sintaxe da linguagem,
- importante lembrar que ao formar uma gramática, esta não atribui sentido (ou semântica) a linguagem, apenas valida a sintaxe.



Abordagens para criar uma gramática

Top Down (de cima para baixo):

- Define regras de seções e organização do arquivo primeiro
- regras mais baixo nível depois
- definição dos tokens(a menor parte) por último
- necessário conhecimento teórico sobre o arquivo que está sendo trabalhado



Abordagens para criar uma gramática

Bottom up (de baixo para cima):

- Define primeiro os tokens, como vão ser capturados, como são agrupados
- aos poucos definindo expressões básicas sobre eles
- por último nos preocupamos com organização de arquivo
- mais intuitivo
- pode necessitar refatoração para poder cumprir os objetivos

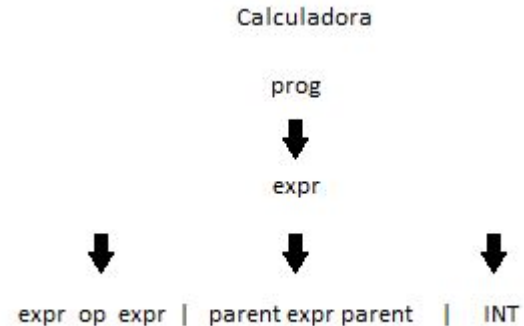


Regras Léxicas

- Primeiro são definidas as regras léxicas para uma calculadora
 - os tokens serão: número (INT), operando (op) e parênteses (parent)
-
- INT : ('0' ... '9')+ ;
 - op : (' + ' , ' - ' , ' * ' , ' / ') ;
 - parent : (' (' , ') ') ;

Regras do parser

- Para representar a calculadora, declara-se uma seção prog (de programa).
- Prog será composto de uma expressão expr, aqui é colocada a lógica de uma expressão matemática.
- Uma expr pode ser composta de:
 - Duas expr separadas por um operador (op)
 - Uma expr entre parênteses (parent)
 - E para finalizar ela também pode ser um número (INT)



Exemplo prático