

LINGUAGEM LIVRE DE CONTEXTO

As Linguagens Livres de Contexto é um reconhecedor de linguagens, capaz de aceitar palavras corretas (cadeia, sentenças) da linguagem. Por exemplo, os autômatos. Um gerador de linguagens é um dispositivo capaz de sintetizar todas as palavras de uma linguagem, é o caso das gramáticas.

Autômatos Finito e Expressões Regulares definem Linguagens Regulares. Linguagem Regulares não permitem construções contendo alinhamentos sintáticos. Assim foi adicionado a possibilidade de alinhamentos às Linguagens Regulares (LRs), obtendo outra classe de linguagens, as Linguagens Livres de Contexto.

A classe das LLCs é de fundamental importância para o estudo da informática, pois:

- a) Compreende um universo mais amplo que as LR, permitindo tratar questões como:
 - Parênteses Balanceados,
 - Construções Bloco-Estruturadas
 - Outras estruturas próprias de linguagens como C, Pascal, etc.
- b) Os algoritmos que as implementam são simples e possuem uma boa eficiência.
- c) Aplicações: analisadores sintáticos, tradutores de linguagens e processadores de texto etc.

O estudo das LLC é desenvolvido a partir de um formalismo gerador (gramática) e um formalismo operacional ou reconhecedor (autômato) como segue:

- a) **Gramáticas Livres de Contexto:** São gramáticas onde as regras de produção são definidas de forma mais livre do que nas gramáticas regulares,
- b) **Autômato de Pilha:** Possui a estrutura básica de um AFD ao qual é associado uma memória auxiliar na forma de pilha e a facilidade de não-determinismo.

GRAMÁTICA LIVRE DE CONTEXTO

As Linguagens Livres Contexto são desenvolvidas a partir das Gramáticas Livres de Contexto (GLC). A GLC é um tipo mais complexo de geradores de linguagem, as quais materializam um completo entendimento do procedimento de construção das palavras pertencentes à linguagem. A GLC é aplicada em compiladores e conversores de documentos.

Exemplo: GLC para a linguagem $a^n b^n$

Regras

$$1- S \rightarrow \lambda$$

$$2- S \rightarrow a S b$$

Exemplo de funcionamento para a palavra aabb

$$S \rightarrow a S b \quad (\text{aplicamos a regra 2})$$

$$aa S bb \quad (\text{aplicamos em S a regra 2})$$

$$aa \lambda bb \quad (\text{aplicamos em S a regra 1})$$

Resultado: aabb

Definição:

Há quatro conceitos importantes em uma descrição de uma gramática:

- Um conjunto finito de símbolos, chamados **terminais** ou **símbolos terminais**;
- Conjunto finito de variáveis, chamados de **categorias sintáticas** ou **não terminais**.
- Uma das variáveis representa a linguagem que está sendo definida, chamada **símbolo de início**.

Existe um conjunto finito de produções ou regras que representam a definição recursiva de uma linguagem. Cada produção consiste em:

- Uma variável que está sendo (parcialmente) definida pela produção. Essa variável é chamada de cabeça (head) da produção
- O símbolo de produção, \rightarrow
- Uma palavra de zero ou mais terminais e variáveis. Esta parte é chamada de corpo.
- O lado esquerdo das produções possui exatamente uma variável.

Definição Formal

Uma Gramática Livre de Contexto (GLC) G é uma quádrupla (V, T, P, S) , onde:

V conjunto finito de variáveis ou símbolo não-terminais

T (conjunto de **terminais**) é um subconjunto de V

P (conjunto de **regras** (ou produções)) é um subconjunto finito de $(V - T) \times V^*$

S (**símbolo inicial** (ou variável de início)) é um elemento de V

Exemplo

Definir a gramática que gera palíndromos. “Um Palíndromo é uma cadeia lido da mesma forma, da direita para a esquerda e da esquerda para a direita”.

Exemplo1: 0110, 11011, 000, ϵ são palíndromos

Exemplo2: 011, 0101, 1010 não são palíndromos.

Gramática

$G = (\{P, \{0, 1\}, A, S)$

Onde A é dada por:

1. $S \rightarrow \lambda$

2. $S \rightarrow 0$

3. $S \rightarrow 1$

4. $S \rightarrow 0 S 0$

5. $S \rightarrow 1 S 1$

Redução do espaço utilizando a seguinte notação :

ou $S \rightarrow \lambda | 01110 S 011 S 1$

Gerar a palavra $w = 00100100$

$S \rightarrow$ 0S0
00S00
001S100
0010S0100
0010 λ 0100

Derivações

Aplicando as regras de produção em ordens diferentes, uma gramática pode gerar muitas cadeias. O conjunto de todas tais cadeias é a linguagem definida ou gerada pela gramática. Existe duas formas de derivação : derivação Mais à Esquerda e Mais à Direita

Exemplo: Linguagem Livre de Contexto

A linguagem gerada pela GLC abaixo é composta por expressões aritméticas contendo colchetes balanceados, dois operandos e um operador:

$G = (\{E\}, \{+, *, [,], x\}, P, E)$, onde

$P = \{ E \rightarrow E+E \mid E*E \mid [E] \mid x \}$.

Por exemplo, a expressão $[x+x]*x$ pode ser gerada pela seguinte seqüência de derivações:

$E \rightarrow E*E$
 $[E]*E$
 $[E+E]*E$
 $[x+E]*E$
 $[x+x]*E$
 $[x+x]*x$

ÁRVORE DE DERIVAÇÃO

Para uma determinada GLC, a representação da derivação de palavras na forma de árvore, denominada *Árvore de Derivação* é como se segue:

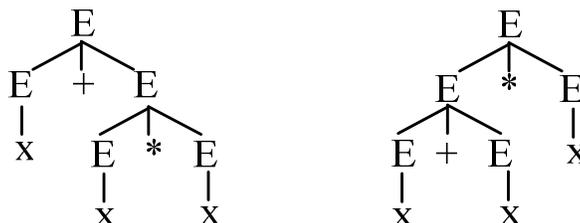
- A raiz é o símbolo inicial da GLC;
- Os vértices interiores obrigatoriamente são variáveis. Se A é um vértice interior e X_1, X_2, \dots, X_n são os filhos de A , então $A \rightarrow X_1X_2\dots X_n$ é uma produção da gramática e os vértices X_1, X_2, \dots, X_n estão ordenados da esquerda para a direita.
- Um vértice folha é um símbolo terminal ou um símbolo vazio. Neste caso o vazio é o único filho de seu pai ($A \rightarrow \lambda$).

Exemplo: Árvores de Derivação

- $G = (\{S\}, \{a, b\}, P, S)$, onde $P = \{S \rightarrow aSb \mid \lambda\}$.
- $G = (\{E\}, \{+, *, [,], x\}, P, E)$, onde $P = \{E \rightarrow E+E \mid E*E \mid [E] \mid x\}$.

Árvore de Derivação x Derivações

Uma única árvore de derivação pode representar derivações distintas de uma mesma palavra ou expressão. Na árvore representada na figura, a palavra $x + x * x$ pode por exemplo ser gerada pelas seguintes derivações:



AMBIGUIDADE

Uma GLC é dita uma gramática *ambígua* se existe uma palavra que possui duas ou mais árvores de derivação. Ex: $x + x * x$.

Exercício – Gramáticas Livre de Contexto

1) Verifique se as seguintes palavras são reconhecidas pela gramática

OBS: USE A CRIATIVIDADE PARA GERAR MAIS TRES PALAVRAS

1) $S \rightarrow SS \mid a S a \mid b S b \mid \lambda$

- a) abba
- b) aabbaaaa

2) $S \rightarrow a b A$
 $A \rightarrow a A \mid b A \mid B$
 $B \rightarrow ba$

- a) abbbaba
- b) abba

3) $S \rightarrow a A \mid b A$
 $A \rightarrow a B \mid b B$
 $B \rightarrow bC$
 $C \rightarrow a C \mid b C \mid \epsilon$

- a) abbab
- b) bbababba

4) $S \rightarrow a S \mid b S \mid B$
 $B \rightarrow aa \mid bb$

- a) ababaa
- b) babbbb

2) Desenvolver as GLCs capazes de produzir as seguintes linguagens:

a) $L2 = \{ \& \}$

b) $L3 = \{ a, b \}^*$

c) $L4 = \{ w \mid w \text{ e' palíndromo em } \Sigma = \{ a, b \}^* \}$

d) $L5 = \{ ww^r \mid w \text{ é palavra em } \Sigma = \{ a, b \}^* \}$

e) $L6 = \{ w \mid w \text{ é palavra de } \Sigma = \{ (,) \}^* \text{ com parênteses balanceados} \}$

f) $L7 = \{b^m a b^n a b^p \mid \text{para qualquer } m, n \text{ e } p > 0\}$

g) $L8 = \{a^n b^m a^n \mid \text{para qualquer } n > 0 \text{ e } m \geq 0\}$

3) Seja G a gramática $G = (V, T, P, S)$, onde

$W = \{O, A, N, V, S\}$

$T = \{\text{Jim, grande, verde, queijo, comeu}\}$

$P = \left\{ \begin{array}{l} O \rightarrow NVN, \\ N \rightarrow S, \\ N \rightarrow NA, \\ A \rightarrow \text{grande}, \\ A \rightarrow \text{verde}, \\ S \rightarrow \text{queijo}, \\ S \rightarrow \text{Jim}, \\ V \rightarrow \text{comeu} \end{array} \right. \quad S = \{O\}$

Gramática representando uma parte do idioma português. O significa oração, A – adjetivo, S – substantivo, V – verbo e N – complemento nominal.

Derivar as seguintes das seguintes palavras

Jim comeu queijo

Jim grande comeu queijo verde

Queijo grande comeu Jim

Queijo grande comeu queijo grande verde grande verde verde

Jim verde comeu Jim grande verde

4) Encontre a Expressão Regular para seguinte Gramática $G = (V, T, P, S)$, onde $V = \{a, b, \}$

5) Verifica se a gramática abaixo é ambígua ou não e justifica sua resposta:

$S \rightarrow SS \mid a S a \mid b S b \mid \epsilon$

(i) *abba*

(ii) *aabbaaaa*