Linguagens livres de contexto

Luís Felipe

UFF

15 de Maio de 2023

Luís Felipe

15/05/23

LLC

- Uma linguagem livre de contexto (LLC) é uma linguagem que pode ser gerada por uma gramática livre de contexto (GLC).
  - Uma gramática livre de contexto (GLC) é uma gramática $G = (V, T, S, R)$ onde as regras estão no formato $x \to v$, onde $x \in V$ e $v \in (V \cup T)^*$.
- É fácil ver que uma GR é um caso particular de uma GLC.
- Consequência: Toda linguagem regular é uma LLC.

Luís Felipe

15/05/23

Exemplos:

1.
$$
\begin{aligned}
G &= (T, V, S, R) \\
T &= \{(,)\} \\
V &= \{x\} \\
S &= x \\
R &= \{x \to (x)x \mid \varepsilon\}
\end{aligned}
$$

2.
$$
\begin{aligned}
G &= (T, V, S, R) \\
T &= \{(,)\} \\
V &= \{x\} \\
S &= x \\
R &= \{x \to \ (x) \mid xx \mid \varepsilon\}
\end{aligned}
$$

Luís Felipe
15/05/23

Exemplos:

3. Palíndromos

$$\begin{array}{rcl} G & = & (T, V, S, R) \\ T & = & \{0, 1\} \\ V & = & \{x\} \\ S & = & x \\ R & = & \{x \to 0x0 \mid 1x1 \mid \varepsilon \mid 0 \mid 1\} \end{array}$$

4. $\{0^n 1^n \mid n \geq 0\}$

$$\begin{array}{rcl} G & = & (T, V, S, R) \\ T & = & \{0, 1\} \\ V & = & \{x\} \\ S & = & x \\ R & = & \{x \to 0x1 \mid \ \varepsilon\} \end{array}$$

Luís Felipe

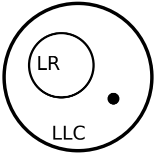
15/05/23

Exemplos:

5. $\{a^i b^i c^k \ : \ i, k \geq 0\}$.

$$
\begin{array}{lcl}
G & = & (T, V, S, R) \\
T & = & \{a, b, c\} \\
V & = & \{S', S_1, S_2\} \\
S & = & S' \\
R & = & \{S' \to S_1 S_2 \\
 & & S_1 \to \ aS_1 b \mid \varepsilon \\
 & & S_2 \to \ cS_2 \mid \varepsilon\}
\end{array}
$$

OBS.: Vimos que a linguagem dos Exemplos 4 e 5 não são regulares (lema do bombeamento). Logo, LR $\subset$ LLC.



Luís Felipe
15/05/23

Derivação em GLC

Exemplo:

$$
\begin{array}{lcl}
G & = & (V, T, S, R) \\
T & = & \{a, b, c\} \\
V & = & \{X, Y, Z\} \\
S & = & X \\
R & = & \{X \to aYb \mid bbY, \\
 & & Y \to ZcX \mid ca, \\
 & & Z \to ZaaZ \mid \varepsilon\}
\end{array}
$$

Luís Felipe

15/05/23

Derivação

- Derivação mais à esquerda: a cada passo da derivação, a variável mais à esquerda é substituída (Notação: $\overset{E+}{\Rightarrow}$)
- Derivação mais à direita: a cada passo da derivação, a variável mais à direita é substituída (Notação: $\overset{D+}{\Rightarrow}$)
- Voltando ao exemplo com derivação mais à esquerda...
- Mais à direita...

Luís Felipe

15/05/23

Exemplos:

- Gramática para gerar expressões aritméticas bem formadas

  - Primeira Tentativa:

$$
\begin{aligned}
G &= (V, T, S, R)\\
T &= \{+, -, *, /, (,), \mathrm{num}\}\\
V &= \{E\}\\
S &= E\\
R &= \{E \to E + E \mid E - E \mid E * E \mid E/E \mid (E) \mid \mathrm{num}\}
\end{aligned}
$$

  - Derivação mais à esquerda:

$$E \Rightarrow E + E \Rightarrow \mathrm{num} + E \Rightarrow \mathrm{num} + E * E \Rightarrow \mathrm{num} + \mathrm{num} * E \Rightarrow \mathrm{num} + \mathrm{num} * \mathrm{num}$$

  - Outra derivação mais à esquerda para a mesma palavra:

$$E \Rightarrow E * E \Rightarrow E + E * E \Rightarrow \mathrm{num} + E * E \Rightarrow \mathrm{num} + \mathrm{num} * E \Rightarrow \mathrm{num} + \mathrm{num} * \mathrm{num}$$

- Tá rolando uma ambiguidade

Luís Felipe

15/05/23

Exemplos:

- Segunda Tentativa:

$$
\begin{aligned}
G &= (V, T, S, R) \\
T &= \{+, -, *, /, (,), \text{num}\} \\
V &= \{E\} \\
S &= E \\
R &= \{E \to (E + E) \mid (E - E) \mid (E * E) \mid (E/E) \mid (E) \mid \text{num}\}
\end{aligned}
$$

- Derivação mais à esquerda para ((num+num) $*$ num):

$$
E \Rightarrow (E * E) \Rightarrow ((E + E) * E) \Rightarrow ((\text{num}) + E) * E \Rightarrow
$$
$$
((\text{num} + \text{num}) * E) \Rightarrow ((\text{num} + \text{num}) * \text{num})
$$

- Existe outra derivação mais à esquerda para a mesma palavra? Teríamos que começar pela regra $E \to (E + E)$.

Luís Felipe

15/05/23

Pra que tantos parênteses?

- Será que conseguimos uma que funcione com menos parênteses?
- Dica: vamos aumentar o número de variáveis (3 regras de precedência $\rightarrow$ pelo menos 3 variáveis)

  - Terceira Tentativa:
  $$\begin{aligned} G &= (V, T, S, R) \\ T &= \{+, -, *, /, (,), \text{num}\} \\ V &= \{E, T, F\} \\ S &= E \\ R &= \{E \rightarrow E + T \mid E - T \mid T, \\ & \quad T \rightarrow T * F \mid T/F \mid F, \\ & \quad F \rightarrow (E) \mid \text{num}\} \end{aligned}$$
  - Derivação mais à esquerda para num+num $*$ num:
  $E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow \text{num} + T \Rightarrow \text{num} + T * F \Rightarrow \text{num} + F * F \Rightarrow \text{num} + \text{num} * F \Rightarrow \text{num} + \text{num} * \text{num}$
  - Haveria ambiguidade? Teríamos que começar pela regra $E \rightarrow T$.

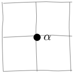
Luís Felipe

15/05/23

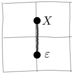
Árvores de Análise Sintática

Uma árvore de análise sintática (AAS) é definida indutivamente:

Casos Básicos:

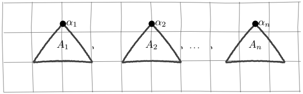
1. 

, onde $\alpha \in T$ é uma AAS

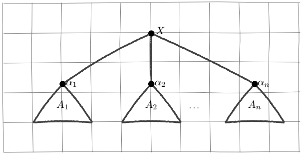
2. 

, onde $x \to \varepsilon \in R$ é uma AAS.

Regra de Composição:

Se 

são AAS's e $x \to \alpha_1\alpha_2\ldots\alpha_n$, então



Luís Felipe

15/05/23

OBS

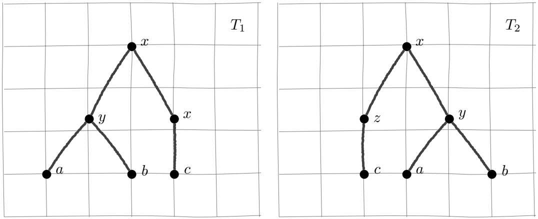
1. Cada nó de uma AAS é rotulado por um símbolo de $T \cup V \cup \{\varepsilon\}$
2. Folhas são rotuladas por símbolos de $T \cup \{\varepsilon\}$.
3. Nós interiores são rotulados por variáveis.
4. Uma AAS que possui raiz rotulada por uma variável x é chamada x-árvore. Estamos especialmente interessados em S-árvores, onde S é o símbolo inicial da gramática.

Luís Felipe

15/05/23

Resultado

- Resultado ou colheita de uma AAS é uma palavra $w \in T^*$ obtida pela concatenação dos símbolos presentes nas folhas, percorrendo as folhas da esquerda para a direita.
- Notação: $r(A)$ ou $c(A)$

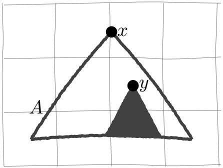


Luís Felipe

15/05/23

OBS

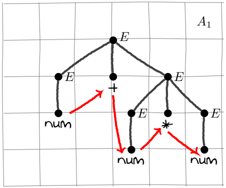
- Se A é uma AAS com raiz rotulada por x e possui um nó interior rotulado por y, a sub-árvore enraizada em y é uma AAS bem formada, chamada y-árvore.

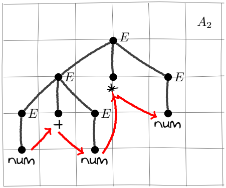


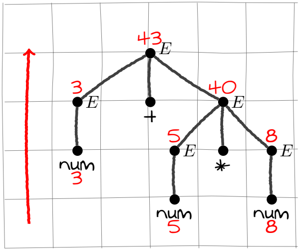
Luís Felipe
15/05/23

Voltando ao exemplo...

Vamos construir as AAS's para *num* $+$ *num* $*$ *num*, de acordo com a primeira gramática.







Luís Felipe
15/05/23

Ambiguidade

- Uma gramática é ambígua se existe ao menos uma palavra $w \in L(G)$ tal que:
  1. Existem ao menos duas AAS's distintas A_1, A_2 tais que $r(A_1) = r(A_2) = w$.
     ou de maneira equivalente,
  2. Existem ao menos duas derivações mais à esquerda (mais à direita) distintas tais que $S \overset{E+}{\Rightarrow}_1 w$ e $S \overset{E+}{\Rightarrow}_2 w$ ($S \overset{D+}{\Rightarrow}_1 w$ e $S \overset{D+}{\Rightarrow}_2 w$)
- Uma linguagem L é inerentemente ambígua se, para toda gramática G tal que $L = L(G)$, temos que G é ambígua.
  OBS.:
  1. A linguagem das expressões aritméticas não é inerentemente ambígua.
  2. Ambiguidade é propriedade de gramáticas.
  3. Ambiguidade inerente é propriedade de linguagens.

Luís Felipe
15/05/23

Ambiguidade inerente

Exemplo: $L = \{a^n b^n c^m d^m : n, m \geq 0\} \cup \{a^n b^m c^m d^n : n, m \geq 0\}$

1. Obtenha uma gramática para L;
2. Verifique que essa gramática é ambígua;
3. Mostre que não existe gramática que gera L e não seja ambígua.

1.

$$
\begin{aligned}
G &= (V, T, S, R)\\
T &= \{a, b, c, d\}\\
V &= \{X, A, B, C, D\}\\
S &= X\\
R &= \{X \to AB \mid C,\\
&\quad A \to aAb \mid ab \mid \varepsilon,\\
&\quad B \to cBd \mid cd \mid \varepsilon,\\
&\quad C \to aCd \mid aDd \mid D \mid \varepsilon,\\
&\quad D \to bDc \mid bc \mid \varepsilon\}
\end{aligned}
$$

2. Verifique possíveis derivações para $aabbccdd$

3. Argumento baseado nas palavras em que $n = m$.