Linguagens Formais e Autómatos / Compiladores

Gramáticas independentes do contexto (GIC)

Artur Pereira / Miguel Oliveira e Silva <{artur,mos}@ua.pt>

DETI, Universidade de Aveiro

Sumário

Gramáticas

Definição de gramática

Uma gramática é um quádruplo $G = (T, N, P, S)$, onde

- T é um conjunto finito não vazio de símbolos terminais;
- N, sendo $N \cap T = \emptyset$, é um conjunto finito não vazio de símbolos não terminais;
- P é um conjunto de produções (ou regras de rescrita), cada uma da forma $\alpha \to \beta$;
- $S \in N$ é o símbolo inicial.

α e β são designados por **cabeça da produção** e **corpo da produção**, respetivamente.

No caso geral

$$\begin{array}{lcl} \alpha & \in & (T \cup N)^+ \\ \beta & \in & (T \cup N)^* \end{array}$$

GRAMÁTICAS INDEPENDENTES DO CONTEXTO – GIC

$\mathcal{D}$ Uma gramática $G = (T, N, P, S)$ diz-se **independente do contexto** se, para qualquer produção $(\alpha \to \beta) \in P$, as duas condições seguintes são satisfeitas

$$\begin{array}{lcl} \alpha & \in & N \\ \beta & \in & (T \cup N)^* \end{array}$$

- A linguagem gerada por uma gramática independente do contexto diz-se independente do contexto
  - as gramáticas regulares são independentes do contexto
- As gramáticas independentes do contexto são fechadas sob as operações de reunião, concatenação e fecho
  - mas não o são sob as operações de intersecção e complementação.

Derivação

$\mathcal{D}$ Dada uma palavra $\alpha\, A\, \beta$ e uma produção $A \rightarrow v$, chama-se **derivação direta** à rescrita de $\alpha\, A\, \beta$ em $\alpha\, v\, \beta$, denotando-se

$$\alpha\, A\, \beta \Rightarrow \alpha\, v\, \beta$$

$\mathcal{D}$ Dada uma palavra $\alpha\, A\, \beta$, com $\beta \in T^*$, e uma produção $A \rightarrow v$, chama-se **derivação direta à direita** à rescrita de $\alpha\, A\, \beta$ em $\alpha\, v\, \beta$, denotando-se

$$\alpha\, A\, \beta \stackrel{D}{\Rightarrow} \alpha\, v\, \beta$$

$\mathcal{D}$ Dada uma palavra $\alpha\, A\, \beta$, com $\alpha \in T^*$, e uma produção $A \rightarrow v$, chama-se **derivação direta à esquerda** à rescrita de $\alpha\, A\, \beta$ em $\alpha\, v\, \beta$, denotando-se

$$\alpha\, A\, \beta \stackrel{E}{\Rightarrow} \alpha\, v\, \beta$$

Derivação

$\mathcal{D}$ Chama-se **derivação à direita** a uma sucessão de zero ou mais derivações diretas à direita, denotando-se

$$\alpha \stackrel{D}{\Rightarrow}{}^* \beta$$

ou, equivalentemente,

$$\alpha = \alpha_0 \stackrel{D}{\Rightarrow} \alpha_1 \stackrel{D}{\Rightarrow} \cdots \stackrel{D}{\Rightarrow} \alpha_n = \beta$$

onde n é o comprimento da derivação.

$\mathcal{D}$ Chama-se **derivação à esquerda** a uma sucessão de zero ou mais derivações diretas à esquerda, denotando-se

$$\alpha \stackrel{E}{\Rightarrow}{}^* \beta$$

ou, equivalentemente,

$$\alpha = \alpha_0 \stackrel{E}{\Rightarrow} \alpha_1 \stackrel{E}{\Rightarrow} \cdots \stackrel{E}{\Rightarrow} \alpha_n = \beta$$

onde n é o comprimento da derivação.

Derivação

Exemplo de derivação

Considere, sobre o alfabeto $T = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, a gramática seguinte

$$
\begin{aligned}
S &\to \varepsilon \mid \mathtt{a}\ B \mid \mathtt{b}\ A \mid \mathtt{c}\ S \\
A &\to \mathtt{a}\ S \mid \mathtt{b}\ A\ A \mid \mathtt{c}\ A \\
B &\to \mathtt{a}\ B\ B \mid \mathtt{b}\ S \mid \mathtt{c}\ B
\end{aligned}
$$

$\mathcal{Q}$ Determine as derivações à direita e à esquerda da palavra `aabcbc`

$\mathcal{R}$

à direita

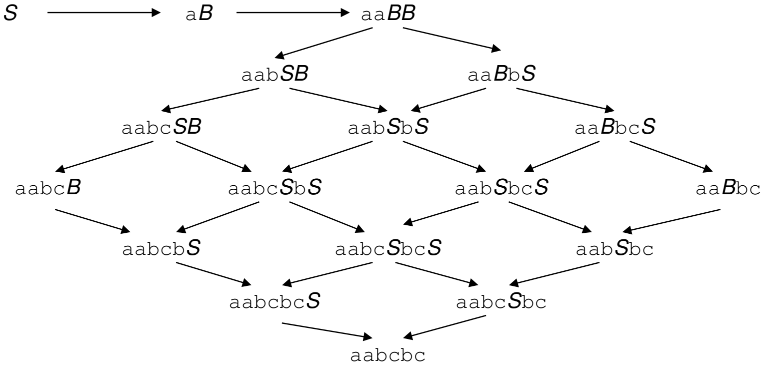
$$
\begin{aligned}
& S \Rightarrow \mathtt{a}B \Rightarrow \mathtt{aa}BB \Rightarrow \mathtt{aa}B\mathtt{b}S \Rightarrow \mathtt{aa}B\mathtt{bc}S \\
& \Rightarrow \mathtt{aa}B\mathtt{bc} \Rightarrow \mathtt{aab}S\mathtt{bc} \Rightarrow \mathtt{aabc}S\mathtt{bc} \Rightarrow \mathtt{aabcbc}
\end{aligned}
$$

à esquerda

$$
\begin{aligned}
& S \Rightarrow \mathtt{a}B \Rightarrow \mathtt{aa}BB \Rightarrow \mathtt{aab}SB \Rightarrow \mathtt{aabc}SB \\
& \Rightarrow \mathtt{aabc}B \Rightarrow \mathtt{aabcb}S \Rightarrow \mathtt{aabcbc}S \Rightarrow \mathtt{aabcbc}
\end{aligned}
$$

Derivação

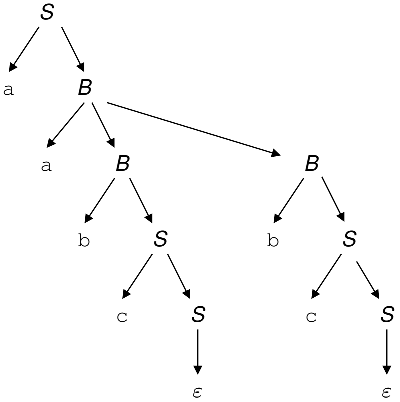
O grafo seguinte capta as alternativas de derivação. Considera-se novamente a palavra `aabcbc` e a gramática anterior



ÁRVORE DE DERIVAÇÃO

Uma **árvore de derivação** (*parse tree*) é uma representação de uma derivação onde os nós-ramos são elementos de N e os nós-folhas são elementos de T. Considera-se a gramática anterior

A árvore de derivação da palavra `aabcbc` na gramática anterior é



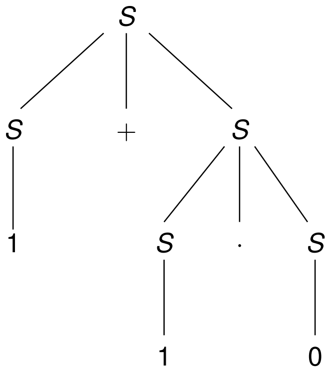
AMBIGUIDADE

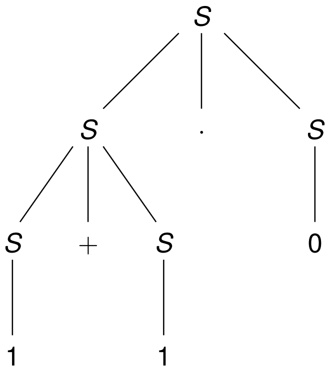
Considere a gramática

$$S \rightarrow S \texttt{ + } S \mid S \texttt{ . } S \mid \neg\ S \mid \texttt{(} S \texttt{)} \mid \texttt{0} \mid \texttt{1}$$

e desenhe as árvores de derivação das palavras `1+1` e `1+1` · `0`.

Para a palavra `1+1` · `0` tem-se





AMBIGUIDADE

- $\mathcal{D}$ Diz-se que uma palavra é derivada **ambiguamente** se possuir duas ou mais árvores de derivação distintas
- $\mathcal{D}$ Diz-se que uma gramática é **ambígua** se possuir pelo menos uma palavra gerada ambiguamente

- Frequentemente é possível definir-se uma gramática não ambígua que gere a mesma linguagem que uma ambígua
- No entanto, há gramáticas **inerentemente ambíguas**

  Ex:

$$L = \{\mathtt{a}^i\mathtt{b}^j\mathtt{c}^k \mid i = j \vee j = k\}$$

Ambiguidade

Considere novamente a gramática seguinte:

$$S \rightarrow S \texttt{ + } S \mid S \texttt{ . } S \mid \neg\ S \mid \texttt{(} S \texttt{)} \mid \texttt{0} \mid \texttt{1}$$

$\mathcal{Q}$ Obtenha uma gramática não ambígua equivalente

$\mathcal{R}$

$$S \rightarrow K \mid S \texttt{ + } K \mid S \texttt{ . } K \mid \neg\ S$$

$$K \rightarrow \texttt{0} \mid \texttt{1} \mid \texttt{(} S \texttt{)}$$

$\mathcal{Q}$ Desenhe a árvode de derivação da palavra `1+1.0`

Projeto de gramáticas: exemplo #1

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\mathtt{a}, \mathtt{b}\}$, determine uma gramática independente do contexto que represente a linguagem

$$L_1 = \{\omega \in T^* : \#(\mathtt{a}, \omega) = \#(\mathtt{b}, \omega)\}$$

$\mathcal{R}_1$

$$S \to \varepsilon \mid \mathtt{a}\ S\ \mathtt{b}\ S \mid \mathtt{b}\ S\ \mathtt{a}\ S$$

$\mathcal{Q}$ A gramática é ambígua? (Analise a palavra `aabbab`.)

Projeto de gramáticas: exemplo #1: solução #2

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\texttt{a}, \texttt{b}\}$, determine uma gramática independente do contexto que represente a linguagem

$$L_1 = \{\omega \in T^* : \#(\texttt{a}, \omega) = \#(\texttt{b}, \omega)\}$$

$\mathcal{R}_2$

$$S \rightarrow \varepsilon \mid \texttt{a}\ B \mid \texttt{b}\ A$$
$$A \rightarrow \texttt{a}\ S \mid \texttt{b}\ A\ A$$
$$B \rightarrow \texttt{a}\ B\ B \mid \texttt{b}\ S$$

$\mathcal{Q}$ A gramática é ambígua? (Analise a palavra `aababb`.)

PROJETO DE GRAMÁTICAS: EXEMPLO #1: SOLUÇÃO #3

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\texttt{a}, \texttt{b}\}$, determine uma gramática independente do contexto que represente a linguagem

$$L_1 = \{\omega \in T^* : \#(\texttt{a}, \omega) = \#(\texttt{b}, \omega)\}$$

$\mathcal{R}_3$

$$S \to \varepsilon \mid \texttt{a}\ B\ S \mid \texttt{b}\ A\ S$$

$$A \to \texttt{a} \mid \texttt{b}\ A\ A$$

$$B \to \texttt{a}\ B\ B \mid \texttt{b}$$

$\mathcal{Q}$ A gramática é ambígua?

Projeto de gramáticas: exemplo #2

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, determine uma gramática independente do contexto que represente a linguagem

$$L_2 = \{\omega \in T^* \, : \, \#(\mathtt{a}, \omega) = \#(\mathtt{b}, \omega)\}$$

$\mathcal{R}$

$$S \rightarrow \varepsilon \mid \mathtt{a}\ B\ S \mid \mathtt{b}\ A\ S \mid \mathtt{c}\ S$$
$$A \rightarrow \mathtt{a} \mid \mathtt{b}\ A\ A \mid \mathtt{c}\ A$$
$$B \rightarrow \mathtt{a}\ B\ B \mid \mathtt{b} \mid \mathtt{c}\ B$$

$\mathcal{Q}$ A gramática é ambígua?

PROJETO DE GRAMÁTICAS: EXEMPLO #3: SOLUÇÃO #1

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, determine uma gramática independente do contexto que represente a linguagem

$$L_3 = \{\omega \in T^* \;:\; \#(\mathtt{a},\omega) = \#(\mathtt{b},\omega) \wedge \forall_{i \le |\omega|}\; \#(\mathtt{a}, \mathrm{prefix}(i,\omega)) \ge \#(\mathtt{b}, \mathrm{prefix}(i,\omega))\}$$

$\mathcal{R}_1$

$$S \to \varepsilon \mid \mathtt{a}\ S\ \mathtt{b}\ S \mid \mathtt{c}\ S$$

$\mathcal{Q}$ A gramática é ambígua?

PROJETO DE GRAMÁTICAS: EXEMPLO #3: SOLUÇÃO #2

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, determine uma gramática independente do contexto que represente a linguagem

$$\begin{aligned} L_3 \quad = \quad & \{\omega \in T^* \;:\; \#(\mathtt{a}, \omega) = \#(\mathtt{b}, \omega) \wedge \\ & \quad \forall_{i \leq |\omega|} \;\; \#(\mathtt{a}, \text{prefix}(i, \omega)) \geq \#(\mathtt{b}, \text{prefix}(i, \omega))\} \end{aligned}$$

$\mathcal{R}_2$

$$S \to \varepsilon \mid \mathtt{a}\;B \mid \mathtt{c}\;S$$

$$B \to \mathtt{a}\;B\;B \mid \mathtt{b}\;S \mid \mathtt{c}\;B$$

$\mathcal{Q}$ A gramática é ambígua? (Analise a palavra $\mathtt{aababb}$.)

Reunião de GIC

$\mathcal{D}$ Sejam $G_1 = (T_1, N_1, P_1, S_1)$ e $G_2 = (T_2, N_2, P_2, S_2)$ duas gramáticas independentes do contexto quaisquer, com $N_1 \cap N_2 = \emptyset$. A gramática $G = (T, N, P, S)$ onde

$$\begin{array}{rcl} T & = & T_1 \ \cup \ T_2 \\ N & = & N_1 \ \cup \ N_2 \ \cup \ \{S\} \quad \text{com} \quad S \notin (N_1 \cup N_2) \\ P & = & \{S \to S_1, S \to S_2\} \ \cup \ P_1 \ \cup \ P_2 \end{array}$$

é independente do contexto e gera a linguagem $L = L(G_1) \cup L(G_2)$.

- As novas produções $S \to S_i$, com $i = 1, 2$, permitem que G gere a linguagem $L(G_i)$

Reunião de GIC: exemplo

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, determine uma gramática independente do contexto que represente a linguagem

$$L_4 \quad = \quad \{\, \omega \in T^* \ :\ \#(\mathtt{a},\omega) = \#(\mathtt{b},\omega) \ \vee\ \#(\mathtt{a},\omega) = \#(\mathtt{c},\omega) \,\}$$

$\mathcal{R}$

- Para $L_{4a} = \{\, \omega \in T^* \ :\ \#(\mathtt{a},\omega) = \#(\mathtt{b},\omega) \,\}$ tem-se

$$S_1 \rightarrow \varepsilon \mid \mathtt{a}\ S_1\ \mathtt{b}\ S_1 \mid \mathtt{b}\ S_1\ \mathtt{a}\ S_1 \mid \mathtt{c}\ S_1$$

- Para $L_{4b} = \{\, \omega \in T^* \ :\ \#(\mathtt{a},\omega) = \#(\mathtt{c},\omega) \,\}$ tem-se

$$S_2 \rightarrow \varepsilon \mid \mathtt{a}\ S_2\ \mathtt{c}\ S_2 \mid \mathtt{c}\ S_2\ \mathtt{a}\ S_2 \mid \mathtt{b}\ S_2$$

- Finalmente, para $L_4 = L_{4a} \cup L_{4b}$ tem-se

$$\begin{aligned} &S \rightarrow S_1 \mid S_2 \\ &S_1 \rightarrow \varepsilon \mid \mathtt{a}\ S_1\ \mathtt{b}\ S_1 \mid \mathtt{b}\ S_1\ \mathtt{a}\ S_1 \mid \mathtt{c}\ S_1 \\ &S_2 \rightarrow \varepsilon \mid \mathtt{a}\ S_2\ \mathtt{c}\ S_2 \mid \mathtt{c}\ S_2\ \mathtt{a}\ S_2 \mid \mathtt{b}\ S_2 \end{aligned}$$

Concatenação de GIC

$\mathcal{D}$ Sejam $G_1 = (T_1, N_1, P_1, S_1)$ e $G_2 = (T_2, N_2, P_2, S_2)$ duas gramáticas independentes do contexto quaisquer, com $N_1 \cap N_2 = \emptyset$. A gramática $G = (T, N, P, S)$ onde

$$\begin{array}{lcl} T & = & T_1 \ \cup \ T_2 \\ N & = & N_1 \ \cup \ N_2 \ \cup \ \{S\} \quad \text{com} \quad S \notin (N_1 \cup N_2) \\ P & = & \{S \rightarrow S_1 S_2\} \ \cup \ P_1 \ \cup \ P_2 \end{array}$$

é independente do contexto e gera a linguagem $L = L(G_1) \cup L(G_2)$.

- A nova produção $S \rightarrow S_1 S_2$ justapõe palavras de $L(G_2)$ às de $L(G_1)$

Concatenação de GIC: exemplo

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, determine uma gramática independente do contexto que represente a linguagem

$$\begin{aligned} L_5 \quad = \quad & \{\omega_1\omega_2 \ : \ \omega_1, \omega_2 \in T^* \\ & \qquad \wedge \#(\mathtt{a}, \omega_1) = \#(\mathtt{b}, \omega_1) \ \wedge \ \#(\mathtt{a}, \omega_2) = \#(\mathtt{c}, \omega_2)\} \end{aligned}$$

$\mathcal{R}$

- Atendendo a que $L_5 = L_{4a} \cdot L_{4b}$ tem-se

$$\begin{aligned} & S \rightarrow S_1 \ \ S_2 \\ & S_1 \rightarrow \varepsilon \mid \mathtt{a} \ \ S_1 \ \ \mathtt{b} \ \ S_1 \mid \mathtt{b} \ \ S_1 \ \ \mathtt{a} \ \ S_1 \mid \mathtt{c} \ \ S_1 \\ & S_2 \rightarrow \varepsilon \mid \mathtt{a} \ \ S_2 \ \ \mathtt{c} \ \ S_2 \mid \mathtt{c} \ \ S_2 \ \ \mathtt{a} \ \ S_2 \mid \mathtt{b} \ \ S_2 \end{aligned}$$

Fecho de Kleene de GIC

Seja $G_1 = (T_1, N_1, P_1, S_1)$ uma gramática independente do contexto qualquer. A gramática $G = (T, N, P, S)$ onde

$$\begin{array}{rcl} T & = & T_1 \\ N & = & N_1 \cup \{S\} \quad \text{com} \quad S \notin N_1 \\ P & = & \{S \rightarrow \varepsilon, S \rightarrow S_1 S\} \cup P_1 \end{array}$$

é independente do contexto e gera a linguagem $L = (L(G_1))^*$.

- A produção $S \rightarrow \varepsilon$, per si, garante que $L^0(G_1) \subseteq L(G)$
- As produções $S \rightarrow S_1 S$ e $S \rightarrow \varepsilon$ garantem que $L^i(G_1) \subseteq L(G)$, para qualquer $i > 0$

Fecho de Kleene de GIC: exemplo

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, determine uma gramática independente do contexto que represente a linguagem

$$L_6 \quad = \quad \{\omega \in T^* \ : \ \#(\mathtt{a}, \omega) > \#(\mathtt{b}, \omega)\}$$

$\mathcal{R}$

- Para $L_{6a} = \{\omega \in T^* \ : \ \#(\mathtt{a}, \omega) = \#(\mathtt{b}, \omega) + 1\}$ tem-se

$$\begin{aligned}
&S \rightarrow A \\
&X \rightarrow \varepsilon \mid \mathtt{a}\ B \mid \mathtt{b}\ A \mid \mathtt{c}\ X \\
&A \rightarrow \mathtt{a}\ X \mid \mathtt{b}\ A\ A \mid \mathtt{c}\ A \\
&B \rightarrow \mathtt{a}\ B\ B \mid \mathtt{b}\ X \mid \mathtt{c}\ B
\end{aligned}$$

- Sabendo que $L_6 = {L_{6a}}^+$ tem-se

$$\begin{aligned}
&S \rightarrow A\ S \mid A \\
&X \rightarrow \varepsilon \mid \mathtt{a}\ B \mid \mathtt{b}\ A \mid \mathtt{c}\ X \\
&A \rightarrow \mathtt{a}\ X \mid \mathtt{b}\ A\ A \mid \mathtt{c}\ A \\
&B \rightarrow \mathtt{a}\ B\ B \mid \mathtt{b}\ X \mid \mathtt{c}\ B
\end{aligned}$$

Símbolos produtivos

- Seja $G = (T, N, P, S)$ uma gramática qualquer
- Um símbolo não terminal A diz-se **produtivo** se for possível transformá-lo num expressão contendo apenas símbolos terminais
- Ou seja, A é produtivo se

$$A \Rightarrow^{+} u \quad \wedge \quad u \in T^{*}$$

- Caso contrário, diz-se que A é **improdutivo**
- Uma gramática é improdutiva se o seu símbolo inicial for improdutivo
- Na gramática

$$\begin{aligned} &S \rightarrow \mathtt{a}\ \mathtt{b} \mid \mathtt{a}\ S\ \mathtt{b} \mid X \\ &X \rightarrow \mathtt{c}\ X \end{aligned}$$

  - S é produtivo, porque $S \Rightarrow \mathtt{ab} \quad \wedge \quad \mathtt{ab} \in T^{*}$
  - X é improdutivo. $X \Rightarrow \mathtt{c}X \Rightarrow \mathtt{cc}X \Rightarrow^{*} \mathtt{c}\cdots\mathtt{c}X$

Símbolos produtivos

- O conjunto símbolos produtivos, N_p, pode ser obtido por aplicação das seguintes regras construtivas

$$\texttt{if}\ (A \to \alpha) \in P\ \texttt{and}\ \alpha \in T^*\ \texttt{then}\ A \in N_p$$
$$\texttt{if}\ (A \to \alpha) \in P\ \texttt{and}\ \alpha \in (T \cup N_p)^*\ \texttt{then}\ A \in N_P$$

- A 1ª regra é um caso particular da 2ª, pelo que pode ser retirada

Algoritmo de cálculo dos símbolos produtivos

```
let Np = ∅,  Pp = P        # Np – símbolos produtivos
repeat
    nothingAdded = TRUE
    foreach (A → α) ∈ Pp do
        if α ∈ (T ∪ Np)* then
            if A ∉ Np then
                Np = Np ∪ {A}
                nothingAdded = false
            Pp = Pp − {A → α}
until nothingAdded or Np = N
```

SÍMBOLOS ACESSÍVEIS

- Seja $G = (T, N, P, S)$ uma gramática qualquer
- Um símbolo terminal ou não terminal x diz-se **acessível** se for possível transformar S (o símbolo inicial) numa expressão que contenha x
- Ou seja, x é acessível se

$$S \Rightarrow^* \alpha x \beta$$

- Caso contrário, diz-se que x é **inacessível**
- Na gramática

$$\begin{array}{l} S \rightarrow \varepsilon \mid \mathtt{a}\ S\ \mathtt{b} \mid \mathtt{c}\ C\ \mathtt{c} \\ C \rightarrow \mathtt{c}\ S\ \mathtt{c} \\ D \rightarrow \mathtt{d}\ X\ \mathtt{d} \\ X \rightarrow C\ C \end{array}$$

  - D, $\mathtt{d}$, e X são inacessíveis
  - Os restantes são acessíveis

Símbolos acessíveis

- O conjunto dos seus símbolos acessíveis, V_A, pode ser obtido por aplicação das seguintes regras construtivas

$$S \in V_A$$
$$\texttt{if } A \rightarrow \alpha B \beta \in P \texttt{ and } A \in V_A \texttt{ then } B \in V_A$$

Algoritmo de cálculo dos símbolos acessíveis

```
let V_A = {S}, N_X = {S}        # V_A – símbolos acessíveis
repeat
    let A = um elemento de N_X
    N_X = N_X \ {A}
    foreach (A → α) ∈ P do
        foreach x in α do
            if x ∉ V_A then
                V_A = V_A ∪ {A}
                if x ∈ N then
                    N_X = N_X ∪ {A}
until N_X = ∅
```

Gramáticas limpas

- Numa gramática os símbolos inacessíveis e os símbolos improdutivos são **símbolos inúteis**
- Se tais símbolos forem removidos obtem-se uma gramática equivalente
- Diz-se que uma gramática é **limpa** se não possuir símbolos inúteis
- Para limpar uma gramática deve-se:
  - começar por a expurgar dos símbolos improdutivos
  - só depois remover os inacessíveis

Gramáticas limpas: exemplo

$\mathcal{Q}$ Sobre o conjunto de terminais $T = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$, determine uma gramática limpa equivalente à seguinte

$$
\begin{aligned}
S &\to \mathtt{a}\ A\ \mathtt{b} \mid \mathtt{b}\ B \\
A &\to \mathtt{c}\ C \mid \mathtt{b}\ B \mid \mathtt{d} \\
B &\to \mathtt{d}\ D \mid \mathtt{b} \\
C &\to A\ C \mid B\ D \mid E\ D \\
D &\to A\ D \mid B\ C \mid C\ E \\
E &\to \mathtt{a}\ A \mid \mathtt{b}\ B \mid \varepsilon
\end{aligned}
$$

$\mathcal{R}$

1 Sendo inicialmente $N_p = \{\}$, identificação das produções com corpos apenas com elementos de $T^\star$

$$
\begin{aligned}
A &\to \mathtt{d} \\
B &\to \mathtt{b} \\
E &\to \varepsilon
\end{aligned}
$$

que permitem acrescentar A, B e E a N_p, donde $N_p = \{A, B, E\}$

GRAMÁTICAS LIMPAS: EXEMPLO

- Falta analisar

$$S \rightarrow \mathtt{a}\ A\ \mathtt{b} \mid \mathtt{b}\ B$$
$$C \rightarrow A\ C \mid B\ D \mid E\ D$$
$$D \rightarrow A\ D \mid B\ C \mid C\ E$$

2. Identificação de novas produções com corpos apenas com elementos de $T \cup N_p$

$$S \rightarrow \mathtt{a}\ A\ \mathtt{b}$$

que permitem acrescentar S a N_p, donde $N_p = \{S, A, B, E\}$

- Falta analisar

$$C \rightarrow A\ C \mid B\ D \mid E\ D$$
$$D \rightarrow A\ D \mid B\ C \mid C\ E$$

3. Nenhum corpo das produções começadas por C e D tem todos os seus elementos no conjunto $N_p = \{S, A, B, E\}$, pelo que se pode concluir que C e D são improdutivos

GRAMÁTICAS LIMPAS: EXEMPLO

- Retirando os símbolos improdutivos da gramática

$$
\begin{aligned}
S &\rightarrow \texttt{a}\ A\ \texttt{b} \mid \texttt{b}\ B \\
A &\rightarrow \texttt{b}\ B \mid \texttt{d} \\
B &\rightarrow \texttt{b} \\
E &\rightarrow \texttt{a}\ A \mid \texttt{b}\ B \mid \varepsilon
\end{aligned}
$$

  Pode-se passar à fase de remoção dos inacessíveis

  4 S é acessível, porque é o inicial
  5 sendo S acessível, de $S \rightarrow \texttt{a}\ A\ \texttt{b}$, tem-se que A é acessível
  6 sendo S acessível, de $S \rightarrow \texttt{b}\ B$, tem-se que B é acessível
  7 de A só se chega a B, que já foi marcado como acessível
  8 de B não se chega a nenhum não terminal

- Pode-se então concluir que E não é acessível, pelo que a gramática limpa é

$$
\begin{aligned}
S &\rightarrow \texttt{a}\ A\ \texttt{b} \mid \texttt{b}\ B \\
A &\rightarrow \texttt{b}\ B \mid \texttt{d} \\
B &\rightarrow \texttt{b}
\end{aligned}
$$