

Linguagens Formais e Autômatos

Hierarquia de Chomsky

Cristiano Lehrer, M.Sc.

Gramáticas (1/3)

- Formalmente, uma gramática G pode ser definida como sendo uma quádrupla:

$$G = (V, \Sigma, P, S)$$

- V é o **vocabulário** da gramática.
 - Corresponde a um conjunto (finito e não-vazio) de símbolos.
- Σ é o conjunto (finito e não-vazio) dos símbolos **terminais** da gramática.
 - Também denominado **alfabeto**.
- P é o conjunto (finito e não-vazio) de produções ou regras de substituição da gramática.
- S é a **raiz** (ou **símbolo inicial**) da gramática, $S \in V$.

Gramáticas (2/3)

- Adicionalmente, define-se $N = V - \Sigma$ como sendo o conjunto dos símbolos não-terminais da gramática.
 - Σ corresponde ao conjunto dos símbolos que podem ser justapostos para compor as sentenças da linguagem que se está definindo.
 - N corresponde ao conjunto dos símbolos intermediários (classes sintáticas) utilizados na estruturação e na geração de sentenças, sem no entanto fazer parte das mesmas.
 - Σ , N e P são conjuntos finitos e não-vazios.
 - P é o conjunto das produções gramaticais, que obedecem à forma geral:
 - $\alpha \rightarrow \beta$, com $\alpha \in V^*NV^*$ e $\beta \in V^*$

Gramáticas (3/3)

- P é o conjunto das produções gramaticais, que obedecem à forma geral:

$$\alpha \rightarrow \beta, \text{ com } \alpha \in V^*NV^* \text{ e } \beta \in V^*$$

- Em outras palavras, α é uma cadeia qualquer constituída por elementos de V , contendo pelo menos um símbolo não-terminal, e β é uma cadeia qualquer, eventualmente vazia, de elementos de V .

Hierarquia de Chomsky (1/3)

- Classificação de gramáticas formais descrita em 1959 pelo linguista Noam Chomsky.
- Possui 4 níveis, sendo que os dois últimos níveis são amplamente utilizados na descrição de linguagens de programação e na implementação de interpretadores e compiladores.
 - Tipo 0 – Gramáticas com estrutura de frase ou irrestrita.
 - Tipo 1 – Gramáticas sensíveis ao contexto.
 - Tipo 2 – Gramáticas livres de contexto.
 - Tipo 3 – Gramáticas regulares.

Hierarquia de Chomsky (2/3)

Hierarquia de Chomsky	Gramática	Linguagem	Reconhecedor
Tipo 0	Irrestrita	Recursivamente Enumeráveis	Máquina de Turing
Tipo 1	Sensível ao Contexto	Sensíveis ao Contexto	Máquina de Turing com Fita Limitada
Tipo 2	Livre de Contexto	Livres de Contexto	Autômato de Pilha
Tipo 3	Linear	Regulares	Autômato Finito

Hierarquia de Chomsky (3/3)

Linguagens Recursivamente Enumeráveis

Linguagens Sensíveis ao Contexto

Linguagens Livres de Contexto

Linguagens Regulares

Tipo 3 (1/2)

- Dá-se o nome de Gramática Linear à Direita àquela cujas produções obedecem todas às seguintes condições:
 - $\alpha \in N$
 - $\beta \in \Sigma, \beta \in N, \beta \in \Sigma N$ ou $\beta = \varepsilon$, de forma não exclusiva.

$$G1 = (\{S, A\}, \{0, 1, 2, 3\}, P1, S)$$
$$P1 = \{S \rightarrow 0S \mid 1S \mid A$$
$$A \rightarrow 2 \mid 3 \}$$

Tipo 3 (2/2)

- Dá-se o nome de Gramática Linear à Esquerda àquela cujas produções obedecem todas às seguintes condições:
 - $\alpha \in N$
 - $\beta \in \Sigma, \beta \in N, \beta \in N\Sigma$ ou $\beta = \varepsilon$, de forma não exclusiva.

$$G_2 = (\{S, A\}, \{0, 1, 2, 3\}, P_2, S)$$

$$P_2 = \{S \rightarrow S2 \mid S3 \mid A$$

$$A \rightarrow 1 \mid 0 \}$$

Tipo 2

- Uma gramática é dita livre de contexto se as suas produções possuírem apenas um símbolo não-terminal em seu lado esquerdo, e uma combinação qualquer de símbolos terminais e não-terminais no lado direito, formalmente definido como:
 - $\alpha \in N$
 - $\beta \in V^*$

$$G_3 = (\{S\}, \{0, 1\}, P_3, S)$$

$$P_3 = \{S \rightarrow 0S1 \mid \varepsilon\}$$

Tipo 1

- Gramáticas sensíveis ao contexto decorrem da restrição, imposta ao formato das produções, de que o comprimento da cadeia do lado direito de cada produção seja no mínimo igual ao comprimento da cadeia do lado esquerdo, não havendo, portanto, possibilidade de redução do comprimento das formas sentenciais quando da reutilização de derivações em gramáticas deste tipo:
 - $\alpha \in V^*NV^*$
 - $\beta \in V^*$
 - $|\beta| \geq |\alpha|$

$$G_4 = (\{S, X, Y\}, \{a, b, c\}, P_4, S)$$

$$P_4 = \{S \rightarrow aXb \mid aXa \\ Xa \rightarrow bc \\ Xb \rightarrow cb \}$$

Tipo 0

- Gramáticas sobre as quais não é imposta nenhuma restrição quanto ao formato de suas produções, exceto pelo fato de que o lado esquerdo das mesmas deva sempre conter pelo menos um símbolo não-terminal:
 - $\alpha \in V^*NV^*$
 - $\beta \in V^*$
 - Não se exige a validade de qualquer relação restritiva entre $|\beta|$ e $|\alpha|$

$G_5 = (\{S, X, Y\}, \{a, b, c\}, P_5, S)$

$P_5 = \{S \rightarrow aXb \mid aXa$

$Xa \rightarrow c$

$Xb \rightarrow c$

$X \rightarrow \varepsilon \}$