Teoria da Computação
Conceitos Básicos
Hierarquia de Chomsky

Prof. Jefferson Magalhães de Morais

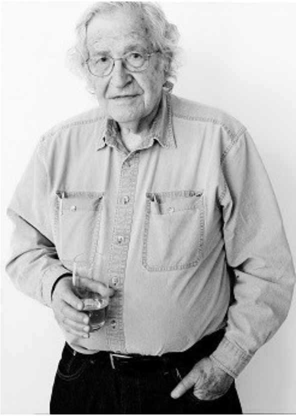
16 de março de 2021

Hierarquia de Chomsky

- Linguista Noam Chomsky impulsionou o estudo sistemático das linguagens formais no final da década de 1950 (**classificação de linguagens**)

- A teoria de autômatos já se apresentada relativamente evoluída à época, porém as linguagens formais ainda não tinha se estabelecido como disciplina

- Chomsky publicou dois artigos que contribuíram para a concentração de pesquisas na área das linguagens formais e teoria de autômatos

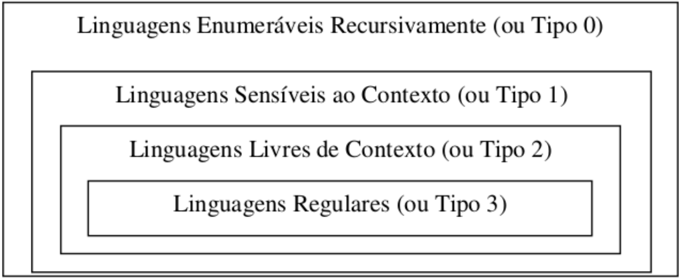


  - A **Hierarquia de Chomsky** é a classificação das linguagens proposta por ele

Hierarquia de Chomsky

- Define quatro classes distintas de linguagens
  - Tipo 0: linguagens recursivamente enumeráveis
  - Tipo 1: linguagens sensíveis ao contexto
  - Tipo 2: linguagens livres de contexto
  - Tipo 3: linguagens regulares
- Essas linguagens são geradas por gramáticas customizadas
- Restrições são aplicadas ao formato das produções $\alpha \to \beta$

Hierarquia de Chomsky



Linguagens regulares: tipo 3

- É a classe de linguagens mais simples dentro da Hierarquia de Chomsky

- Podem ser geradas pelas gramáticas
  - **Linear à direita**
    1. $\alpha \in N$
    2. $\beta \in \Sigma, \beta \in N, \beta \in \Sigma N$ ou $\beta = \varepsilon$

    Exemplo
    $G_1 = (\{0, 1, 2, 3, S, A\}, \{0, 1, 2, 3\}, \{S \to 0A, S \to 1S, S \to A, A \to 2, A \to 3\}, S)$

  - **Linear à esquerda**
    1. $\alpha \in N$
    2. $\beta \in \Sigma, \beta \in N, \beta \in N\Sigma$ ou $\beta = \varepsilon$

    Exemplo
    $G_2 = (\{0, 1, 2, 3, S, A\}, \{0, 1, 2, 3\}, \{S \to S2, S \to S3, S \to A, A \to 1, A \to 0\}, S)$

Linguagens livre de contexto: tipo 2

- Uma gramática é **livre de contexto** se obedecer as regras
  1. $\alpha \in N$
  2. $\beta \in V^*$

  Exemplo: $G_3 = (\{0, 1, S\}, \{0, 1\}, \{S \rightarrow 0S1, S \rightarrow \varepsilon\}, S)$

- Gramáticas desse tipo geram linguagens denominadas **livres de contexto**
- Toda gramática do tipo 3 é também uma gramática do tipo 2
- Mas nem toda gramática do tipo 2 é também gramática do tipo 3
- As gramáticas G_1 e G_2 são lineares e livres de contexto. Já a G_3 é livre de contexto apenas

Linguagens sensíveis ao contexto: tipo 1

- Uma gramática é **sensível ao contexto** se obedecer as regras
  1. $\alpha \in V^*NV^*$
  2. $\beta \in V^*$
  3. $|\beta| >= |\alpha|$
- Esse tipo de gramática não permite, *a priori*, gerar a cadeia vazia
- Entretanto, é comum considerar L sendo sensível ao contexto, mesmo que $\varepsilon \in L$, se $L - \{\varepsilon\}$ puder ser gerada por uma gramática sensível ao contexto **Exemplo**:

  $G_4 = (\{a, b, c, S, X, Y\}, \{a, b, c\}, \{S \rightarrow aXb, S \rightarrow aXa, Xa \rightarrow bc, Xb \rightarrow cb\}, S)$

- As gramática lineares G_1 e G_2 são também sensíveis ao contexto. A gramática livre de contexto G_3 não é sensível ao contexto, devido à presença da produção $S \rightarrow \varepsilon$

Linguagens recursivamente enumeráveis: tipo 0

- Uma gramática é **recursivamente enumerável** (ou **irrestrita**) se obedecer as regras
  1. $\alpha \in V^*NV^*$
  2. $\beta \in V^*$

  Exemplo: $G_5 = (\{a, b, c, S, X, Y\}, \{a, b, c\}, \{S \rightarrow aXb, S \rightarrow aXa, Xa \rightarrow c, Xb \rightarrow c, X \rightarrow \varepsilon\}, S)$ é irrestrita, porém não é sensível ao contexto

- As gramáticas G_1, G_2, G_3 e G_4 são todas irrestritas

Resumindo

Linguagem

- Toda linguagem do tipo $i, 0 \leq i \leq 3$ é gerada por uma gramática do tipo i
- Toda linguagem do tipo $i, 1 \leq i \leq 3$ é também uma linguagem do tipo $i - 1$
- L será do tipo 1 sse $L - \{\varepsilon\}$ for gerada por alguma gramática do tipo 1

Gramática

- Toda gramática do tipo 3 (linear) é também do tipo 2 (livre de contexto)
- Nem toda gramática do tipo 2 também é do tipo 1 (sensível ao contexto)
- São do tipo 1 apenas aquelas que não possuem produções em que $\beta = \varepsilon$
- Toda gramática do tipo 1 é também do tipo 0 (irrestrita)

Conclusão

- Cada linguagem demanda, de acordo com sua complexidade, classes de reconhecedores progressivamente mais poderosos
- É possível reconhecer linguagens do tipo i com reconhecedores para linguagens do tipo $i - 1, 1 \leq i \leq 3$
- O custo computacional de modelos mais complexos só justifica a sua utilização quando a linguagem em questão exige tal poder de reconhecimento, caso contrário construir sempre modelos mais simples

Tipo	**Classe de linguagens**	**Modelo de gramática**	**Modelo de reconhecedor**
0	Recursivamente enumeráveis	Irrestrita	Máquina de Turing
1	Sensíveis ao contexto	Sensível ao contexto	Máquina de Turing com fita limitada
2	Livres de contexto	Livre de contexto	Autômato de pilha
3	Regulares	Linear (direita ou esquerda)	Autômato finito