

# Gramáticas e Hierarquia de Chomsky

Guilherme Dantas de Andrade Rodrigues e João Pedro Araujo

28/05/2023

## Resumo

Neste estudo dirigido, exploramos os conceitos de gramáticas formais e a Hierarquia de Chomsky. Iniciamos com a definição das gramáticas formais e sua relação tanto com a linguagem natural quanto com as linguagens de programação. Discutimos as regras de produção e os símbolos terminais e não terminais presentes nas gramáticas. Em seguida, examinamos cada classe da Hierarquia de Chomsky. Começamos com as gramáticas irrestritas da Classe Tipo 0, que possuem a maior capacidade de expressão. Em seguida, avançamos para as gramáticas sensíveis ao contexto da Classe Tipo 1, as gramáticas livres de contexto da Classe Tipo 2 e as gramáticas regulares da Classe Tipo 3, cada uma com suas propriedades e exemplos ilustrativos. Concluímos ressaltando a importância das gramáticas e da Hierarquia de Chomsky no estudo das estruturas de dados e na ciência da computação. Esses conceitos fornecem uma base teórica sólida para a descrição e análise das linguagens formais, permitindo a criação de linguagens de programação, desenvolvimento de compiladores, análise sintática e processamento de linguagem natural, entre outras aplicações.

## Introdução

Neste estudo dirigido, pretendemos compreender os conceitos de gramáticas formais e explorar a hierarquia de Chomsky como modelo teórico para classificar essas gramáticas. Gramáticas formais são sistemas de regras que definem a estrutura de uma linguagem e são usadas para descrever tanto linguagens naturais como inglês e português quanto linguagens artificiais como a linguagem de programação C. A importância das gramáticas formais e da hierarquia de Chomsky é evidente no estudo das estruturas de dados e sua aplicação na ciência da computação. Essas ferramentas fornecem uma base sólida para entender a estrutura das linguagens e são essenciais na hora de projetar linguagens de programação onde as regras de sintaxe e semântica são definidas. Além disso, as gramáticas formais são usadas na análise e geração de linguagens formais, essenciais para a construção de compiladores e intérpretes. Para a ciência da computação utilizamos de suas aplicações e no estudo de dados em algumas situações tendo elas como: Análise e geração de linguagens formais, Design de linguagens de programação, compiladores e interpretadores, autômatos e expressão regulares e teoria da complexidade. Com isso podemos concluir que as gramáticas formais e a hierarquia de Chomsky desempenham um papel central no estudo das estruturas

de dados e na ciência da computação fornecendo ferramentas e conceitos fundamentais tanto para estrutura de dados, quanto para design de linguagens de programação e também para análise de programas.

## 1 Conceitos Básicos

Como conceitos básicos temos que as gramáticas formais são sistemas de regras que definem a estrutura de uma linguagem. Elas são usadas para descrever tanto as linguagens naturais, como o inglês e o português, quanto às linguagens de programação. As gramáticas formais consistem em um conjunto de regras de produção que especificam como sequências de símbolos podem ser combinadas para formar sentenças válidas em uma linguagem. Na linguagem natural, as gramáticas formais ajudam a descrever a estrutura e as regras que governam a formação de frases e sentenças corretas. Elas definem a sintaxe da linguagem, como a ordem das palavras, a concordância gramatical e a estrutura das sentenças. Nas linguagens de programação, as gramáticas formais são utilizadas para definir a sintaxe da linguagem, ou seja, as regras que determinam como os programas devem ser escritos. Essas regras incluem a forma correta de declarar variáveis, escrever comandos, estruturar blocos de código, entre outros elementos da linguagem.

## 2 Hierarquia de Chomsky:

Na Hierarquia de Chomsky temos 4 tipos de classes, iniciando da classe Tipo 0 até a classe Tipo 3. 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.

### 2.1 Tipo 0

Gramáticas com estrutura de frase ou irrestrita: 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 deve sempre conter pelo menos um símbolo não-terminal. Não há restrições nas regras de produção.

### 2.2 Tipo 1

Gramáticas sensíveis ao contexto: 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. As regras de produção possuem restrições contextuais, o lado esquerdo da regra deve ter um número de símbolos não menor que o lado direito.

### 2.3 Tipo 2

Gramáticas livres de contexto: Uma gramática é dita livre de contexto se as suas produções possuem apenas um símbolo não-terminal em seu lado esquerdo, e uma combinação qualquer de símbolos terminais e não-terminais no lado direito. As regras de produção têm a forma  $A \rightarrow \beta$ , onde  $A$  é um único símbolo não terminal e  $\beta$  é uma sequência de símbolos (terminais ou não terminais).

## 2.4 Tipo 3

Tipo 3 - Gramáticas regulares: As regras de produção são da forma  $A \rightarrow aB$  ou  $A \rightarrow a$ , onde  $A$  e  $B$  são símbolos não terminais, " $a$ " é um símbolo terminal e a produção é restrita a ter apenas um símbolo não terminal no lado esquerdo

## 3 Gramáticas

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  $\beta = \epsilon$ , de forma não exclusiva. 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  $\beta = \epsilon$ , de forma não exclusiva.

### 3.1 Definição

A classe Tipo 0: Consiste em gramáticas irrestritas, também conhecidas como gramáticas sem restrições ou gramáticas recursivamente enumeráveis. Nesse tipo de gramática, não há restrições nas regras de produção, permitindo a geração de qualquer sequência de símbolos. Propriedades: As gramáticas irrestritas são as mais poderosas em termos de expressividade, pois podem gerar linguagens complexas e arbitrariamente grandes. Elas podem representar qualquer linguagem formal, incluindo linguagens naturais e linguagens de programação.

Exemplos:  $S \rightarrow aSb \mid \epsilon$

A classe Tipo 1: Consiste em gramáticas sensíveis ao contexto, onde as regras de produção são restritas de acordo com o contexto em que são aplicadas. Nesse tipo de gramática, o lado esquerdo da regra deve ter um número igual ou maior de símbolos do que o lado direito

Exemplos:  $S \rightarrow aSb \mid ab$

As gramáticas sensíveis ao contexto são um modelo formal que permite descrever e reconhecer linguagens sensíveis ao contexto, ou seja, linguagens cuja estrutura depende do contexto em que são utilizadas. Elas são usadas para representar estruturas complexas em linguagens naturais e em algumas linguagens de programação.

A classe Tipo 2: Composta por gramáticas livres de contexto, onde as regras de produção possuem uma forma fixa, em que um símbolo não terminal é substituído por uma sequência de símbolos terminais e não terminais.

Exemplos:  $S \rightarrow aSb \mid ab$

As gramáticas livres de contexto são um modelo formal amplamente utilizado para descrever e reconhecer linguagens livres de contexto. Elas são utilizadas na teoria da linguagem, no design de linguagens de programação e em análises sintáticas de compiladores. A notação de Backus-Naur (BNF) é comumente usada para representar gramáticas livres de contexto.

A classe Tipo 3: Consiste em gramáticas regulares, onde as regras de produção são mais restritas do que nas classes anteriores. As regras de produção são da forma  $A \rightarrow aB$  ou  $A \rightarrow a$ , em que  $A$  e  $B$  são símbolos não terminais, ' $a$ ' é um símbolo terminal e a produção é restrita a ter apenas um símbolo não terminal à esquerda.

Exemplos:  $S \rightarrow aS \mid b$

As gramáticas regulares são um modelo formal utilizado para descrever e reconhecer linguagens regulares, que são linguagens simples e bem estruturadas. Elas são amplamente utilizadas na teoria da linguagem, em processamento de linguagem natural e em expressões regulares, que são uma ferramenta poderosa para manipular e descrever padrões em linguagens regulares.

### 3.2 Relações entre as classes

A Hierarquia de Chomsky estabelece uma classificação hierárquica das gramáticas formais em quatro níveis ou classes, conforme mencionado anteriormente. Essas classes são organizadas de forma que cada classe inclui a classe abaixo dela, resultando em um conjunto de relações de inclusão e exclusão entre as classes. Cada classe de gramática da Hierarquia de Chomsky possui diferentes capacidades de geração e reconhecimento de linguagens formais. As diferentes classes de gramáticas da Hierarquia de Chomsky fornecem modelos formais com diferentes capacidades de expressão e aplicação, sendo utilizadas em diversos campos da ciência da computação, linguística e teoria da linguagem.

## Referências

1. **OpenAI. GPT-3.5. OpenAI, 2021. Disponível em:**  
*[https : //openai.com/research/gpt – 3 – 5](https://openai.com/research/gpt-3-5).*
2. **Hopcroft, J. E., Motwani, R., Ullman, J. D. (2006). Introduction to Automata Theory, Languages, and Computation. Addison-Wesley.**
3. **Sipser, M. (2012). Introduction to the Theory of Computation. Cengage Learning.**
4. **Monteiro, S. Conceitos Elementares da Teoria da Computação. Disponível em:**  
*[https : //www.cin.ufpe.br/ tbl2/pesq/trabalho1nccword.pdf](https://www.cin.ufpe.br/tbl2/pesq/trabalho1nccword.pdf) Pinto, F.*
5. **Gramáticas de formas como modelo computacional teórico. Disponível em:**  
*[https : //lume.ufrgs.br/bitstream/handle/10183/26341/000757792.pdf?sequence = 1](https://lume.ufrgs.br/bitstream/handle/10183/26341/000757792.pdf?sequence=1)*
6. **USP. Linguagens Formais Hierarquia de Chomsky. Disponível em:**  
*[http : //wiki.icmc.usp.br/images/a/a5/Aula14outTiposGramaticas.pdf](http://wiki.icmc.usp.br/images/a/a5/Aula14outTiposGramaticas.pdf)*
7. **Cristiano Lehrer, M.Sc. Hierarquia de Chomsky. Disponível em:**  
*[http : //www.ybadoo.com.br/tutoriais/lfa/09/LFA09.pdf](http://www.ybadoo.com.br/tutoriais/lfa/09/LFA09.pdf)*