

# Hierarquia de Chomsky

Linguagens Formais A

Prof. Giovani Rubert Libreotto

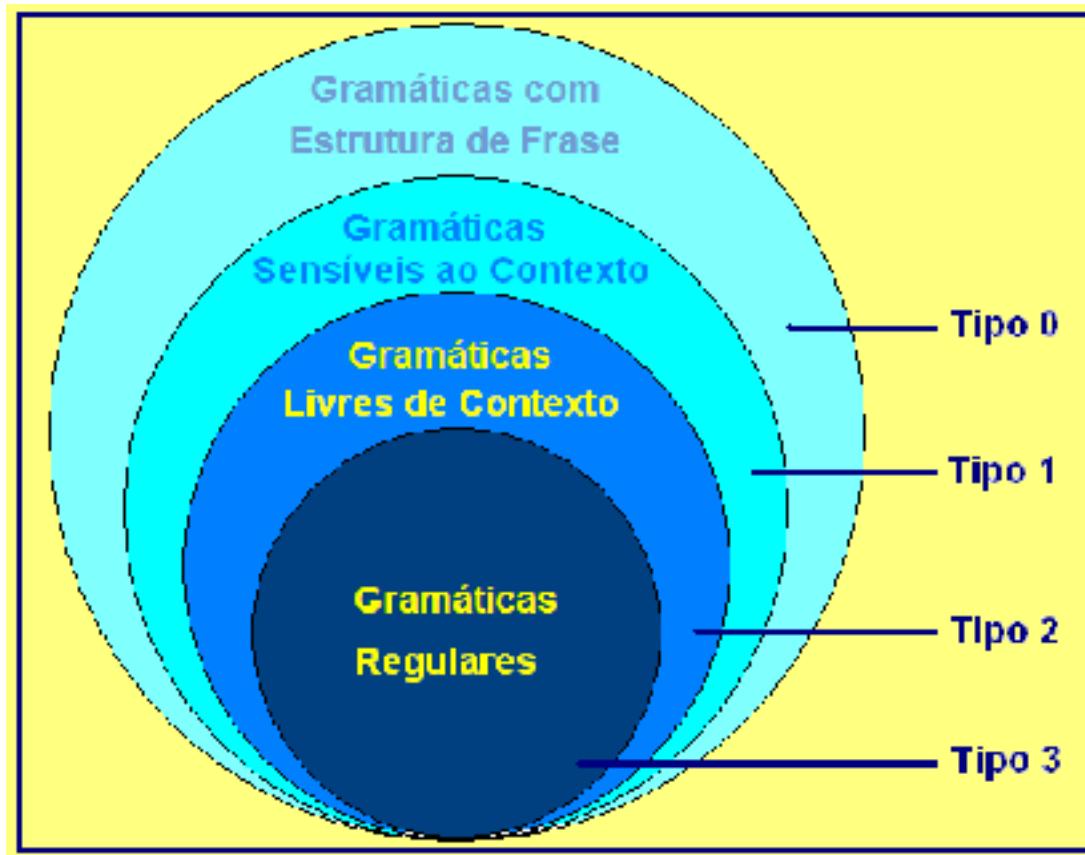
# Hierarquia de Chomsky

- Hierarquia de Chomsky é a classificação de gramáticas formais descrita em 1959 pelo linguista Noam Chomsky.
- Esta classificação possui 4 níveis, sendo que os dois últimos níveis (os níveis 2 e 3) são amplamente utilizados na descrição de linguagem de programação e na implementação de interpretadores e compiladores.
- Mais especificamente, o nível 2 é utilizado em análise sintática (computação) e o nível 3 em análise léxica.

# Hierarquia de Chomsky

- A classificação das gramáticas começa pelo tipo 0, com maior nível de liberdade em suas regras, e aumentam as restrições até o tipo 3.
- Cada nível é um super conjunto do próximo.
- Logo, uma gramática de tipo n é consequentemente uma linguagem de tipo n - 1.

# Hierarquia de Chomsky



# Gramática Regular

- Também conhecida como de **Tipo 3**, é uma restrição sobre a forma das produções.
- Pode-se criar uma nova classe de gramáticas de grande importância no estudo dos compiladores por possuírem propriedades adequadas para a obtenção de reconhecedores simples.
- Que também podem ser denominada de Expressão regular.

# Gramática Regular

- $S ::= a \ S \mid u \ S \mid g \ S \mid c \ S \mid a \mid u \mid g \mid c$
- Ex: uaggcuacgccagcau

# Gramática Livre do Contexto

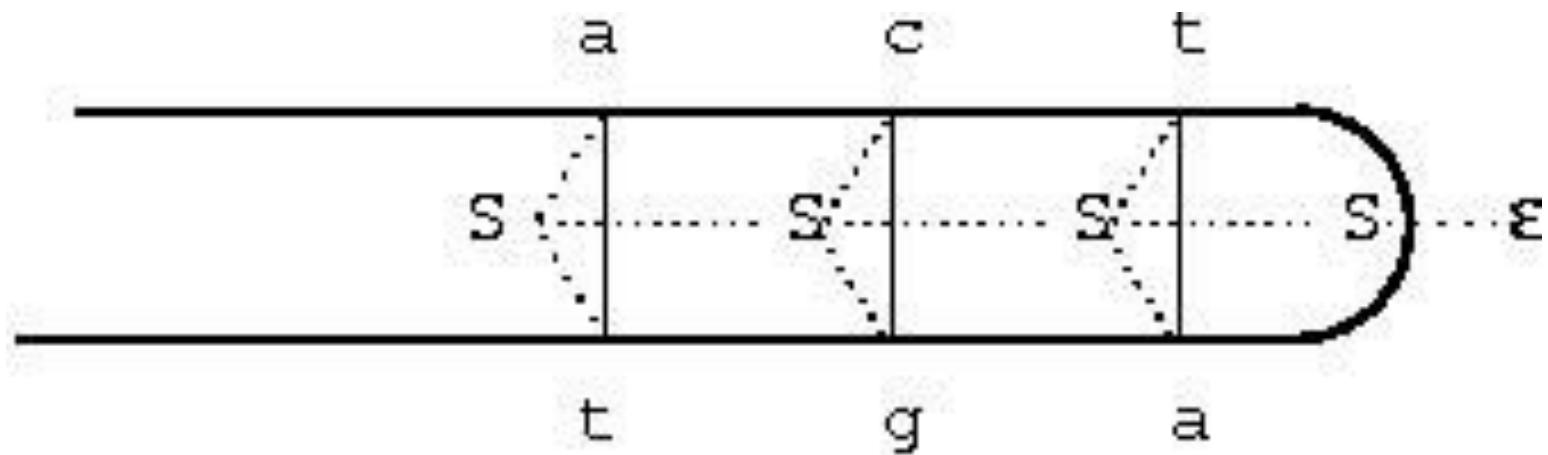
- Também conhecida como de **Tipo 2**, são aquelas em que é levantado o condicionamento das substituições impostas pelas regras definidas pelas produções.
- Este condicionamento é eliminado impondo às produções uma restrição adicional, que restringe as produções à forma geral:

$$A \rightarrow \beta$$

onde  $A \in V_n$  e  $\beta \in (V_n \cup V_t)^*$

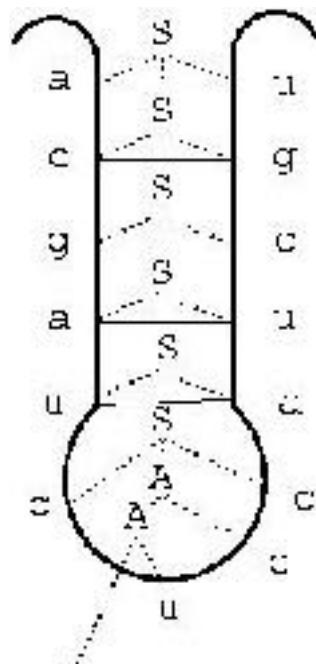
# Gramática Livre do Contexto (1)

- $S ::= a S u \mid u S a \mid c S g \mid g S c \mid \epsilon$



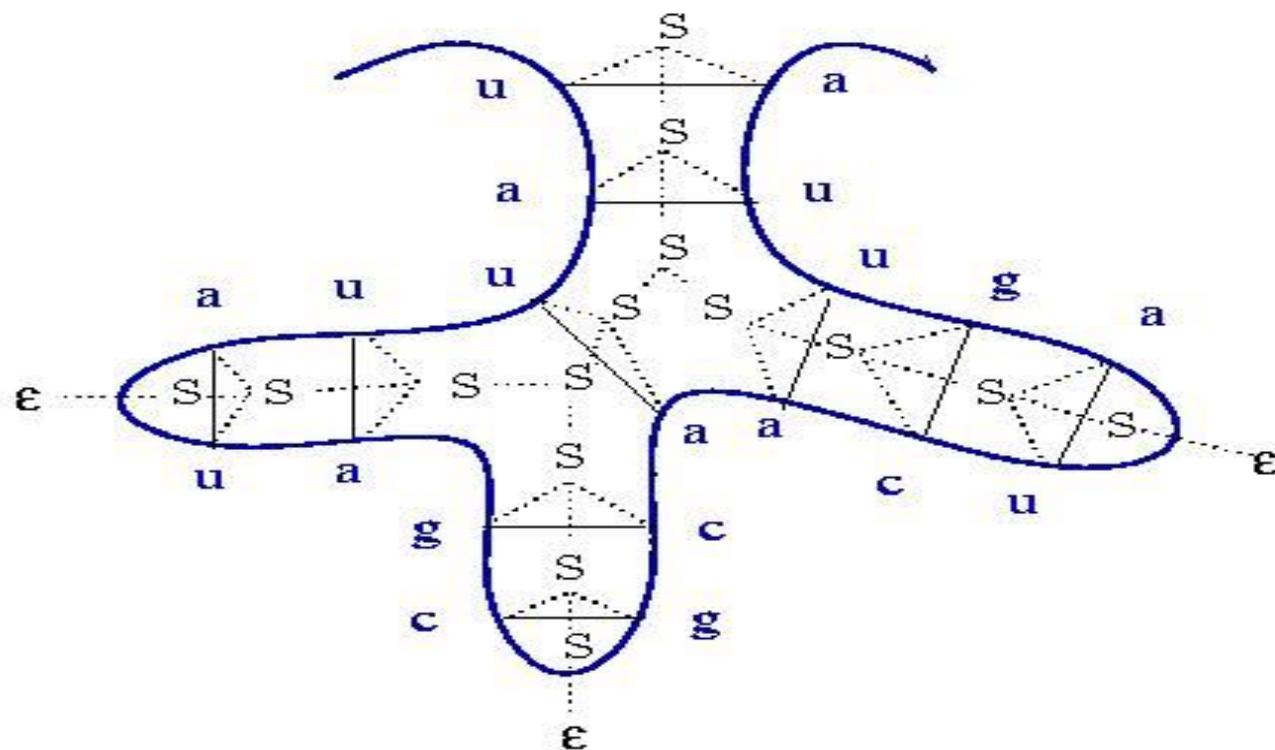
# Gramática Livre do Contexto (2)

- $S ::= a \ S \ u \ | \ u \ S \ a \ | \ c \ S \ g \ | \ g \ S \ c \ | \ A$
- $A ::= A \ a \ | \ A \ u \ | \ A \ c \ | \ g \ A \ | \ a \ | \ u \ | \ c \ | \ g$



# Gramática Livre do Contexto (3)

- $S ::= a \: S \: u \mid u \: S \: a \mid c \: S \: g \mid g \: S \: c \mid ss$



# Gramática com Estrutura de Frase (Irrestrita)

- Também conhecida como **Tipo 0**, são aquelas às quais nenhuma limitação é imposta.
- O universo das linguagens que se podem definir através dos mecanismos gerativos definidos pela gramática corresponde exatamente ao conjunto das linguagens que esta classe de gramática é capaz de gerar.

# Gramática Sensível ao Contexto

- Também conhecida como **Tipo 1** ou Gramáticas Irrestritas.
- Se as regras de substituição forem sujeitas à restrição de que nenhuma substituição possa reduzir o comprimento da forma sentencial à qual a substituição é aplicada, cria-se uma classe chamada sensíveis ao contexto ou tipo 1.

# Gramática Sensível ao Contexto

- A gramática formal  $G=(V, T, P, S)$  é sensível ao contexto se todas as regras em  $P$  são da forma  $\alpha A \beta \rightarrow \alpha \gamma \beta$
- em que:
  - $A \in V$  (isto é,  $A$  é um único não-terminal),
  - $\alpha, \beta \in (V \cup P)^*$  (ou seja,  $\alpha$  e  $\beta$  são sequências de não-terminais e símbolo terminal) e
  - $\gamma \in (V \cup P)^+$  (isto é,  $\gamma$  é uma sequência não vazia de não-terminais e terminais).

# Gramática Sensível ao Contexto

- Essa gramática gera a linguagem não-livre do contexto canônica  $\{a^n b^n c^n \mid n \geq 1\}$ :

1.  $S \rightarrow aSBC$
2.  $S \rightarrow aBC$
3.  $CB \rightarrow HB$
4.  $HB \rightarrow HC$
5.  $HC \rightarrow BC$
6.  $aB \rightarrow ab$
7.  $bB \rightarrow bb$
8.  $bC \rightarrow bc$
9.  $cC \rightarrow cc$

# Gramática Sensível ao Contexto

- A gramática seguinte gera a linguagem de cópia não-livre do contexto,  $C=\{xx|x\in\{a,b\}^*\}$ :
  1.  $S \rightarrow aAS|bBS|T$
  2.  $Aa \rightarrow aA$
  3.  $Ba \rightarrow aB$
  4.  $Ab \rightarrow bA$
  5.  $Bb \rightarrow bB$
  6.  $BT \rightarrow Tb$
  7.  $AT \rightarrow Ta$
  8.  $T \rightarrow \epsilon$

# Resumindo: Classes de Gramáticas

- Tipo 3: Regulares:
  - $A \rightarrow a \mid aB$
- Tipo 2: Livres de Contexto:
  - $A \rightarrow aBCfd$
- Tipo 1: Sensíveis ao Contexto:
  - $aAD \rightarrow aBCfd$

# Hierarquia de Chomsky

Linguagens Formais A

Prof. Giovani Rubert Libreotto