Chomsky Classification of Grammar Type 0 Grammar : Unrestricted Grammar

- In this grammar G=(N,T,P,S) production rules are of the form
- $\alpha \rightarrow \beta$
- LHS of the rule $\alpha \in V^*NV^*$ (V= N U T, α is a string of symbols of the alphabet N U T and α must contain at least one non terminal)
- RHS of the rule $\beta \in V^*$ (β is a string over alphabet V)
- No restriction on the rules of the grammar

Chomsky Classification of Grammar Type 0 Grammar : Unrestricted Grammar

- **AB → CD**
- **AaBD → defBD**
- **ASB → aS**

Chomsky Classification of Grammar Type 1 Grammar : Context Sensitive Grammar

- In this grammar G=(N,T,P,S) production rules are of the form
- $\alpha \rightarrow \beta$
- $\alpha \in V^*NV^*$ (V= N U T, α is a string of symbols of the alphabet N U T and it must contain at least one non terminal)
- $\beta \in V^+$ (β is a string over alphabet V and it cannot by empty string)
- Restriction on rules - $|\alpha| <= |\beta|$ (length of LHS of the rule is less than equal to RHS of the rule)
- Length Increasing Grammar

Chomsky Classification of Grammar Type 1 Grammar : Context Sensitive Grammar

- **ASB → AcB**
- **aABCc → aAbdCc**
- **A → ab**

Chomsky Classification of Grammar Type 1 Grammar : Context Sensitive Grammar

- In this grammar G=(N,T,P,S) production rules are of the form
- $\alpha A \beta \rightarrow \alpha \gamma \beta$
- $\alpha, \beta \in V^*$ (V=N U T, α and β are strings of symbols of the alphabet N U T)
- $A \in N$ (A is a symbol over alphabet N)
- $\gamma \in V^+$ (V=N U T, γ is a string over alphabet N U T, and it can not be empty string)
- We read rule $\alpha A \beta \rightarrow \alpha \gamma \beta$ as A is rewritten as γ in the context of α and β.
- α is called left context and β is called right context of the production rule.
- Indian languages are CSL

Chomsky Classification of Grammar Type 2 Grammar : Context Free Grammar

- In this grammar G=(N,T,P,S) production rules are of the form
- $A \rightarrow \alpha$
- $A \in N$ (A is a symbol of alphabet N)
- $\alpha \in V^*$ (V=N U T, α is a string over alphabet N U T)
- There is no concept of left or right context in the grammar rules.
- LHS of the rule is context free
- Note – The rule $S \rightarrow \varepsilon$ is allowed only if S does not appear on the RHS of any rule in the grammar.
- Most European languages are CFL

Chomsky Classification of Grammar Type 2 Grammar : Context Free Grammar

- **A → aAa**
- **S → ab**
- **A → abB**

Chomsky Classification of Grammar
Type 3 Grammar : Regular Grammar

- In this grammar G=(N,T,P,S) production rules are of the form
- $A \rightarrow aB$ or $A \rightarrow b$
- $A,B \in N$ (A is a symbol of alphabet N)
- $a \in T$ and $b \in T \cup \varepsilon$ (a,b are symbols over alphabet T and b can be empty string also)
- RHS of the rule must contain a terminal followed by non terminal or only terminal symbol
- Note – The rule $S \rightarrow \varepsilon$ is allowed only if S does not appear on the RHS of any rule in the grammar.

Chomsky Classification of Grammar
Type 3 Grammar : Regular Grammar

- **A → a**
- **A → bC**
- **B → c**
- **S → aS**

Chomsky Classification of Grammar Summary

Type	Rules of the form	Restriction	Examples
Type 0 Unrestricted Grammar	**$\alpha \rightarrow \beta$** **$\alpha \in V^*NV^*$ ($V=N \cup T$)** **$\beta \in V^*$**	No restriction	**$AB \rightarrow CD$** **$AaBD \rightarrow defBD$** **$ASB \rightarrow aS$**
Type 1 Context Sensitive Grammar	**$\alpha A \beta \rightarrow \alpha \gamma \beta$** **$\alpha, \beta \in V^*$** **$A \in N$ and $\gamma \in V^+$**	Length of LHS of the rule <= length of RHS of the rule	**$ASB \rightarrow AcB$** **$aABCc \rightarrow aAbdCc$** **$A \rightarrow ab$**
Type 2 Context Free Grammar	**$A \rightarrow \alpha$** **$A \in N$** **$\alpha \in V^*$**	LHS of the rule is context free	**$A \rightarrow aAa$** **$S \rightarrow ab$** **$A \rightarrow abB$**
Type 3 Regular Grammar	**$A \rightarrow aB$ Or $A \rightarrow b$** **$A,B \in N$** **$a \in T$ and $b \in T \cup \varepsilon$**	RHS of the rule has either terminal or terminal followed by non terminal	**$A \rightarrow a$** **$A \rightarrow bC$** **$B \rightarrow c$** **$S \rightarrow aS$**

Grammar
Notation Convention

- **Non terminal symbols** are denoted by upper case alphabets A,B,C,D,...
- **Terminal symbols** are denoted by by lower case alphabets a,b,c,d...
- A **sentential form α** is a string of non terminal and terminal symbols that can be derived from start symbol S in the grammar G [**$\alpha \in (N \cup T)^*$ and $S \Rightarrow^* \alpha$**]
- It is usually denoted by symbols **α, β, γ.**
- A **sentence w** is a string consisting of terminal symbols only that can be derived from start symbol S in the grammar G [**$w \in T^*$ and $S \Rightarrow^* w$]**
- It is usually denoted by symbols **w,x,y,z**
- **Notation ⇒ directly derives or derives in one step**
- **Notation $\Rightarrow^*$ derives in zero or more steps (Reflexive transitive closure of ⇒)**