1. Right- or Left-Linear Grammar:

Linear Grammar: A grammar in which each production contains at most one nonterminal in its right-hand side of any production.

Right-linear grammar (Definition): G = (V, T, S, P) is said to be right-linear if all productions are of the form: $A \to xB$, $A \to x$, where A, B∈V and x ∈ T*.

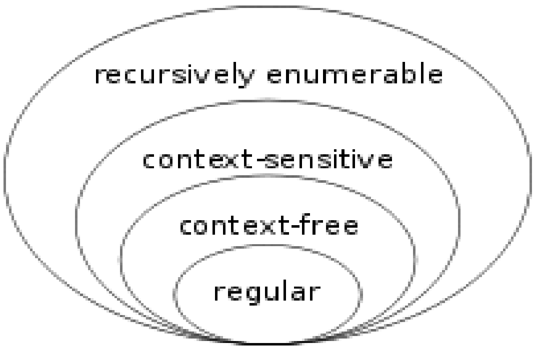
Left-linear grammar (Definition): G = (V, T, S, P) is said to be left-linear if all productions are of the form: $A \to Bx$, $A \to x$, where A, B ∈ V and x ∈ T*.

Example 1:

Find L(G) where G = ({S, S1, S2}, {a, b}, S, P) with
$S \to S1ab$,
$S1 \to S1ab \mid S2$,
$S2 \to a$.
Answer: This is a left-linear grammar.
$S \Rightarrow S1ab \Rightarrow S1abab \Rightarrow S2abab \Rightarrow aabab$.
Then
L(G) = {aabw |w ∈ (ab)*}.

2. Hierarchy of Grammars (Chomsky Hierarchy):

The Chomsky hierarchy classifies grammars according to syntactic restrictions on rules as following. Let G = (∑, V, S, P) be a grammar.
1. G is called a **Type-0** grammar or an **unrestricted** grammar.
2. G is a **Type-1** or **context-sensitive** grammar.
3. G is a **Type-2** or **context-free grammar**.
4. G is a **Type-3** or **regular** grammar.



2.1 An Unrestricted Grammar:

A set of production rules of the form $\alpha \to \beta$ where α and β are arbitrary strings of terminal and non-terminal symbols. The rules of these grammars do not have the restriction above, their left-hand sides may contain any string of terminal and /or non-terminal symbols, provided there is at least one non-terminal symbol.

Example 2:

L = {w ∈ {a, b, c}$^+$: number of a's, b's and c's is the same}

S → ABCS
S → ABC
AB → BA
BC → CB
AC → CA
BA → AB
CA → AC
CB → BC
A → a
B → b
C → c

Exercise 1: what is the language of the following grammar?

S → aBSc
S → aBc
Ba → aB
Bc → bc
Bb → bb

Exercise 2: Let G be the grammar <N, ∑, P, S>, where N = {S}, ∑ = {a, b}, and P are S→ϵ , S →aSbS.

a. Find all the strings that are directly derivable from SaS in G.
b. Find all the derivations in G that start at S and end at ab.
c. Find all the sentential forms (sequences) of G of length 4 at most.

Exercise 3: Find all the derivations of length 3 at most that start at S in the grammar <N, ∑, P, S> whose production rules are:

S →AS
aS →bb
A →aa

2.2 A Context-Sensitive Grammar (CSG):

A production rules of the grammar have the form α → β and |β| ≥|α|, i.e. no production rule is length-decreasing.

A language L is context-sensitive if it is generated by some context-sensitive grammar. Context-Sensitive grammars may have more than one symbol on the left-hand-side of their grammar rules, provided that at least one of them is a non-terminal and the number of symbols on the left-hand-side does not exceed the number of symbols on the right-hand-side.

Example3: The following grammar is context-sensitive (CSG).

$S \rightarrow aBCT|aBC$
$T \rightarrow ABCT|ABC$
$BA \rightarrow AB$
$CA \rightarrow AC$
$CB \rightarrow BC$
$aA \rightarrow aa$,
$aB \rightarrow ab$
$bB \rightarrow bb$,
$bC \rightarrow bc$
$cC \rightarrow cc$

Example 4: The following grammar is context-sensitive.

$S \rightarrow aTb \mid ab$
$aT \rightarrow aaTb \mid ac$.

What is the language of the grammar?

$\{ab\} \cup \{a^{n+1}cb^{n+1} \mid n \geq 0\}$. This language is context-free, it has the grammar

$S \rightarrow aTb \mid ab$, and $T \rightarrow aTb \mid c$. Any context-free language is context sensitive.

2.3 A Context-Free Grammar (CFG):

A production rules of the grammar have the form $\alpha \rightarrow \beta$, each production in P satisfies:

$|\alpha|=1$; i.e., α is a single nonterminal.

A language generated from a context-free grammar is called a context-free language. Any context-free language is context sensitive.

The grammars are called context free because – since all rules only have a nonterminal on the left-hand side – one can always replace that nonterminal symbol with what is on the right-hand side of the rule.

Example 5: $\{a^nb^nc^n \mid n \geq 0\}$ is context-sensitive but not context-free.

Here is a **CSG**.

S →Ɛ | abc | aTBc
T → abC | aTBC
CB → BC
B → b.
C → c.

Derive aaabbbccc.

S ⇒ aTBc ⇒ aaTBCBc ⇒ aaabCBCBc ⇒ aaabBCCBc ⇒ aaabbCCBc⇒aaabbCBCc ⇒ aaabbBCCc ⇒ aaabbbCCc ⇒ aaabbbCcc ⇒ aaabbbccc.

Example 6:

Let L(G1) = $\{0^n1^n | n \geq 0\}$ and L(G2) = $\{0^n\#1^n | n \geq 0\}$. Given two CFLs, it is easy to construct a CFG for their **union**, e.g., combining CFGs for L(G1) and L(G2):

$S \rightarrow S_1 | S_2$
$S_1 \rightarrow 0S_11 | \epsilon$
$S_2 \rightarrow 0S_21 | \#$

Example 7:

S→abS
S→a
L(G)=(ab)*a

2.4 Regular Grammar:

G is a *Type-3* or *right-linear* or *regular grammar* if each production has one of the following three forms: A→cB, A → c, A →ϵ; where A, B are non-terminals (with B = A allowed) and c is a terminal.

The **regular languages** are subset of the context-free languages.

Such a grammar restricts its rules to a single nonterminal on the left-hand side. The a right-hand side consisting of a single terminal, possibly followed (or preceded, but not both in the same grammar) by a single nonterminal.

Regular languages can be considered as special types of **context free languages**, i.e. all regular languages are CF languages but not all CF languages are regular.

Example 8:

The following grammar is unrestricted.

S → TbC
Tb → c
cC → Sc | Ɛ

This grammar is not context-sensitive, not context-free, and not regular. But can transform it into S → Sc | Ɛ. So, the language of the grammar is regular.

Regular grammar generates regular languages as in following examples:

Example 9:

S→Aab

A→Aab|B

B→a

L(G)=aab(ab)*

Example 10:

The CFG ({S}, {a, b}, S, P) with P consisting of the following productions:

S → aSb

S →ϵ

The grammar is not regular because of the ***b*** on the right of S.

It generates the language $a^n b^n$ where $n \geq 0$. This is not a regular language but it can be generated by a context free grammar is therefore a context free language.

Exercise 1:

G = ({S}, {0, 1}, {S → 0S1|ϵ}, S)

- Is ϵ in L(G)?
- Is 01 in L(G)?
- Is 0011 in L(G)?
- Is $0^n 1^n$ in L(G)?

What language is defined by the following G?

S →ϵ

S → 0S1

What language is defined by the following G?

S →ϵ

S → 0S0

S → 1S1

Exercise 2:

What is language generated by this grammar G given by the productions

S → 0S0 | 0B0

B → 1B | 1