0.1.7 Introduction to Grammar

Grammar is a mechanism to describe the languages.

A grammar (G) is defined as a quadruple

$$G = (V, T, S, P)$$

where

V	=	Finite set of objects called VARIABLES
T	=	Finite set of objects called TERMINAL SYMBOLS
$S \in V$	=	Start variables
P	=	Finite set of Productions.

A production rule P is of the form

$$x \rightarrow y$$

Given a string w, of the form $w = uxv$, we can use the production rule $x \rightarrow y$ and obtain a new string $z = uyv$.

The set of all strings obtained by using Production rules is the "Language" generated by the Grammar.

If the grammar $G = (V, T, S, P)$ then

$$L(G) = \{w \in T^* : S \overset{*}{\Rightarrow} w\}$$

If $W \in L(G)$, then the sequence

$$S \Rightarrow w_1 \Rightarrow w_2 \Rightarrow w_3 \;\ldots\ldots\; \Rightarrow w_n \Rightarrow w$$

is a "derivation" of the sentence w.

The string $S, w_1, w_2, \ldots\ldots\; w_n$, which contain variables as well as terminals, are called "SENTENTIAL FORMS" of the derivation.

Example 0.1.52: Given a Grammar $G = (\{S\}, \{a, b\}, S, P)$ with P defined as

$$S \to aSb,$$
$$S \to \lambda$$

(i) Obtain a sentence in language generated by G and the sentential form
(ii) Obtain the language $L(G)$.

Solution

$$\begin{aligned} S &\Rightarrow aSb \\ &\Rightarrow aaSbb \\ &\Rightarrow aabb \end{aligned}$$

Therefore we have $S \overset{*}{\Rightarrow} aabb$.

(i) Sentence in the language generated by $G = aabb$.
Sentential form = $aaSbb$.
(ii) The rule $S \to aSb$ is recursive.

All sentential forms will have the forms

$$w_i = a^i S\, b^i$$

Applying the production rule $S \to aSb$, we get

$$a^i S\, b^i \Rightarrow a^{i+1} S\, b^{i+1}$$

This is true for all i.

In order to get a sentence we apply $S \to \lambda$.

Therefore we get

$$S \overset{*}{\Rightarrow} a^n S\, b^n \Rightarrow a^n b^n$$

Therefore $L(G) = \{a^n b^n;\ n \geq 0\}$.

Example 0.1.53: Obtain a Grammar which generates the language

$$L = \{a^n b^{n+1} : n \geq 0\}$$

Solution

With $L = \{a^n b^n : n \geq 0\}$, the grammar

$$G = (\{S\}, \{a, b\}, S, P)$$

with production rules $S \to aSb, S \to \lambda$.

Therefore $L = \{a^n b^{n+1} : n \geq 0\}$ is obtained by generating an extra b.

This is done with a production rule

$$S \to Ab.$$

Hence the grammar G is given by

$G = (\{S, A\}\{a, b\}, S, P)$ with production rules given by

$$\begin{aligned} &S \to Ab, \\ &A \to aAb \\ &A \to \lambda \end{aligned}$$

Example 0.1.54: Obtain the language L produced by G with production rules

$$\begin{aligned} &S \to SS, \\ &S \to \lambda \\ &S \to aSb \\ &S \to bSa \end{aligned}$$

Solution

It is known from the given production rules that G has equal number of a's and b's.

If w starts with an 'a' and ends with a 'b', then $w \in L$ has the form

$$w = a\, w_1\, b$$

where $w_1 \in L$.

If w starts with a 'b' and ends with an 'a' then $w \in L$ has the form

$$w = b\, w_1\, a$$

where $w_1 \in L$..

As a string in L can begin and end with the same symbol, the string shoud be of the form

$$w = w_1 w_2$$

where w_1 and w_2 are in L, produced by $S \to SS$.

This generates the language

$$L = \{w : n_a(w) = n_b(w)\}$$

where $n_a(w)$ and $n_b(w)$ denotes the number of a's and number of b's in the string w, respectively.

Example 0.1.55: Given $G_1 = (\{A, S\}, \{a, b\}, S, P_1)$ with P_1 defined by the production rules

$$S \to aAb \mid \lambda$$
$$A \to aAb \mid \lambda$$

show that $L(G_1) = \{a^n b^n : n \geq 0\}$.

Also show that G_1 is equivalent to G where $G = (\{S\}, \{a, b\}, S, P)$ where P is given by

$$S \to aSb$$
$$S \to \lambda.$$

Solution

Given P_1 as

$$S \to aAb$$
$$S \to \lambda$$
$$S \to aAb$$
$$A \to \lambda.$$

$S \to \lambda$ produces a string with zero length. ($n = 0$)

$$\begin{aligned} S &\Rightarrow aAb \\ &\Rightarrow a\lambda b \\ &\Rightarrow ab \end{aligned} \qquad \begin{aligned} S &\Rightarrow aAb \\ &\Rightarrow aaAbb \\ &\Rightarrow aabb \\ &\Rightarrow a^2 b^2 \quad \text{and so on} \end{aligned}$$

Therefore $L(G_1) = \{a^n b^n : n \geq 0\}$.

Given $G = (\{S\}, \{a, b\}, S, P)$ where P is $S \to aSb$, $S \to \lambda$.

The rule $S \to aSb$ is recursive.

All sentential forms will have the forms

$$w_i = a^i S b^i$$

Applying the production rule $S \to aSb$, we get

$$a^i S b^i \Rightarrow a^{i+1} S b^{i+1}$$

This is true for all i.

In order to get a sentence we apply $S \to \lambda$.

Therefore we get

$$S \overset{*}{\Rightarrow} a^n S b^n \Rightarrow a^n b^n$$

Hence $$L(G) = \left\{a^n b^n \ : n \geq 0\right\}$$

Hence G_1 is equivalent to G as both the grammars are given by $\left\{a^n b^n \ : n \geq 0\right\}$.

Example 0.1.56: Given a grammar G defined by the production rules

$$\begin{aligned} &S \to AB \\ &A \to Aa \\ &B \to Bb \\ &A \to a \\ &B \to b. \end{aligned}$$

Show that the word $$w = a^2 b^4 \in L(G),$$

where L is a language determined by G.

Solution

$$\begin{aligned} S &\Rightarrow AB \\ &\Rightarrow AaB \\ &\Rightarrow aaB \\ &\Rightarrow aaBb \\ &\Rightarrow aaBbb \\ &\Rightarrow aaBbbb \\ &\Rightarrow aabbbb \\ &\Rightarrow a^2 b^4 \end{aligned}$$

Hence the word $w = a^2 b^4 \in L(G)$.

Example 0.1.57: Find grammars for $\Sigma = \{a, b\}$ that generate the sets of

(a) all strings with exactly one 'a'

(b) all strings with at least one 'a'

(c) all strings with no more than three a's.

Solution

(a) Given $\Sigma = \{a, b\}$
We are able to write the grammar G which produces all strings with exactly one 'a' whose production rules are

$$A \to aSb$$
$$S \to Sb$$
$$S \to \in$$

(b) For all strings with at least one 'a': Production rules of Grammar C are

$$A \to aSb$$
$$S \to bSa$$
$$S \to \in$$

(c) For all strings with no more than three a's

$$L = \left\{ a^n b^m \middle| n \le 3,\ m \ge 0 \right\}$$

with production rules

$$A \to aSb$$
$$S \to aBb$$
$$B \to aCb$$
$$C \to bC$$
$$C \to b \mid \in.$$

Example 0.1.58: Give a simple description of the language generated by the grammar with productions

$$(a)\quad S \to aA,\quad A \to bS,\quad S \to \lambda \qquad\qquad (b)\quad S \to Aa,\quad A \to B,\quad B \to Aa.$$

Soution

(a) For the given production rules

$$S \to aA$$
$$A \to bS$$
$$S \to \lambda$$

we have the language L given by

$$L = \left\{ a^n b^n \,\middle|\, n \ge 1 \right\}$$

(b) For the given production rules

$$S \to Aa$$

$$A \rightarrow B$$
$$B \rightarrow Aa$$

There is no language L produced as there is no proper termination.