Example 7.59: Convert the following grammar

$$A \rightarrow BC$$
$$B \rightarrow CA \mid b$$
$$C \rightarrow AB \mid a$$

into GNF

Solution:

Let $A = A_1$, $B = A_2$, $C = A_3$ and the resulting grammar is

$$A_1 \rightarrow A_2A_3$$
$$A_2 \rightarrow A_3A_1 \mid b$$
$$A_3 \rightarrow A_1A_2 \mid a$$

First two productions are of the form: $A_i \rightarrow A_j\alpha$ where $i < j$. So, we consider A_3-production.

Consider A_3-production: Substituting for A_1 in A_3-production we get

$$A_3 \rightarrow A_1A_2 \mid a = (A_2A_3)A_2 \mid a = A_2A_3A_2 \mid a$$

Again replacing the first A_2 in A_3-production we get,

$$A_3 \rightarrow A_2A_3A_2 \mid a = (A_3A_1 \mid b)\, A_3A_2 \mid a$$
$$= A_3A_1A_3A_2 \mid bA_3A_2 \mid a$$

we get the resulting A_3-production as

$$A_3 \rightarrow A_3A_1A_3A_2 \mid bA_3A_2 \mid a$$

which is having left recursion. After eliminating left recursion, we get

$$A_3 \rightarrow bA_3A_2 \mid a \mid bA_3A_2Z \mid aZ$$
$$Z \rightarrow A_1A_3A_2 \mid A_1A_3A_2Z$$

Now, all A_3-productions are in GNF.

Consider A_2-productions: Since all A_3-productions are in GNF, substituting A_3-productions in A_2-production we get

$$A_2 \rightarrow (bA_3A_2 \mid a \mid bA_3A_2Z \mid aZ)\, A_1 \mid b$$
$$= bA_3A_2A_1 \mid aA_1 \mid bA_3A_2ZA_1 \mid aZA_1 \mid b$$

which is in GNF. Now, all A_2-productions are in GNF.

Consider Z-productions: Since A_1-productions are in GNF, substituting A_1-production in Z-production we get Z-productions also in GNF as shown below:

$$Z \rightarrow A_1A_3A_2 \mid A_1A_3A_2Z$$
$$= (bA_3A_2A_1A_3 \mid aA_1A_3 \mid bA_3A_2ZA_1A_3 \mid aZA_1A_3 \mid bA_3)\, A_3A_2$$
$$(bA_3A_2A_1A_3 \mid aA_1A_3 \mid bA_3A_2ZA_1A_3 \mid aZA_1A_3 \mid bA_3)\, A_3A_2Z$$

which can be written as

$$Z \rightarrow bA_3A_2A_1A_3A_3A_2 \mid aA_1A_3A_3A_2 \mid bA_3A_2ZA_1A_3A_3A_2 \mid$$
$$aZA_1A_3\ A_3A_2 \mid bA_3A_3A_2$$
$$Z \rightarrow bA_3A_2A_1A_3A_3A_2\ Z \mid aA_1A_3A_3A_2\ Z \mid bA_3A_2ZA_1A_3A_3A_2Z \mid$$
$$aZA_1A_3\ A_3A_2Z \mid bA_3A_3A_2Z$$

Since all productions are in GNF, the resulting grammar is also in GNF. So, the final grammar obtained in GNF notation is

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

where

$V = \{A_1, A_2, A_3, Z\}$

$T = \{a, b\}$

$P = \{$

$$A_1 \rightarrow bA_3A_2A_1A_3 \mid aA_1A_3 \mid bA_3A_2ZA_1A_3 \mid aZA_1A_3 \mid bA_3$$
$$A_2 \rightarrow bA_3A_2A_1 \mid aA_1 \mid bA_3A_2ZA_1 \mid aZA_1 \mid b$$
$$A_3 \rightarrow bA_3A_2 \mid a \mid bA_3A_2Z \mid aZ$$
$$Z \rightarrow bA_3A_2A_1A_3A_3A_2 \mid aA_1A_3A_3A_2 \mid bA_3A_2ZA_1A_3A_3A_2 \mid$$
$$aZA_1A_3\ A_3A_2 \mid bA_3A_3A_2$$
$$Z \rightarrow bA_3A_2A_1A_3A_3A_2Z \mid aA_1A_3A_3A_2\ Z \mid bA_3A_2ZA_1A_3A_3A_2Z \mid$$
$$aZA_1A_3\ A_3A_2Z \mid bA_3A_3A_2Z$$

$\}$

A_1 is the start symbol

Example 7.60: Convert the following grammar

$$S \rightarrow AA \mid 0$$
$$A \rightarrow SS \mid 1$$

into GNF

Solution:

Let $S = A_1$ and $A = A_2$. After substitution, the resulting grammar obtained is shown below:

$$A_1 \rightarrow A_2\ A_2 \mid 0$$
$$A_2 \rightarrow A_1\ A_1 \mid 0$$

First production is of the form: $A_i \rightarrow A_j\alpha$ where $i < j$. So, we consider A_2-production.

Consider A_2-production: Substituting for A_1 in A_2-production we get

$$A_2 \rightarrow A_2A_2\ A_1 \mid 0A_1 \mid 1$$

The above grammar is having left recursion. After eliminating left recursion, we get

$$A_2 \to 0A_1 \mid 1 \mid 0A_1Z \mid 1Z$$

$$Z \to A_2A_1 \mid A_2A_1Z$$

Now, all A_2-productions are in GNF.

Consider A_1-productions: Since all A_2-productions are in GNF, substituting A_2-productions in A_1-production we get

$$A_1 \to 0A_1A_2 \mid 1A_2 \mid 0A_1Z\,A_2 \mid 1Z\,A_2 \mid 0$$

which is in GNF. Now, all A_1-productions are in GNF.

Consider Z-productions: Since A_2-productions are in GNF, substituting A_2-production in Z-production we get Z-productions also in GNF as shown below:

$$Z \to 0A_1A_1 \mid 1A_1 \mid 0A_1Z\,A_1 \mid 1ZA_1$$

$$Z \to 0A_1A_1Z \mid 1A_1Z \mid 0A_1ZA_1Z \mid 1ZA_1Z$$

So, the final grammar obtained which is in GNF is shown below:

$$A_1 \to 0A_1A_2 \mid 1A_2 \mid 0A_1Z\,A_2 \mid 1Z\,A_2 \mid 0$$

$$A_2 \to 0A_1 \mid 1 \mid 0A_1Z \mid 1Z$$

$$Z \to 0A_1A_1 \mid 1A_1 \mid 0A_1Z\,A_1 \mid 1ZA_1$$

$$Z \to 0A_1A_1Z \mid 1A_1Z \mid 0A_1ZA_1Z \mid 1ZA_1Z$$

Exercises

1) Define the following: i) Context-free grammar ii) Derivation
2) What is a sentence or sentential form? Explain with example.
3) Consider the following grammar

$$S \to aCa$$

$$C \to aCa \mid b$$

What is the language generated by this grammar?

4) Obtain a grammar to generate the following languages:
   a. $L = \{a^nb^n : n \geq 0\}$
   b. $L = \{a^{n+1}b^n : n \geq 0\}$
   c. $L = \{a^nb^{n+1} : n \geq 0\}$
   d. $L = \{a^nb^{n+2} : n \geq 0\}$
   e. $L = \{a^nb^{2n} : n \geq 0\}$
5) Obtain a grammar to generate set of all palindromes over $\Sigma = \{a, b\}$.
6) Obtain a grammar to generate $L = \{ww^R \text{ where } w \in \{a, b\}^*\}$.
7) Obtain a grammar to generate strings of all non-palindromes over $\{a, b\}$.

8) Obtain the grammar to generate following languages:

a) $L = \{ 0^m 1^m 2^n \mid m \geq 1 \text{ and } n \geq 0 \}$

b) $L = \{w \mid n_a(w) = n_b(w)\}$

9) What is the language generated by the following grammar

$$S \rightarrow 0A \mid \varepsilon]$$

$$A \rightarrow 1S$$

10) Obtain a CFG to generate a string of balanced parentheses.

11) Obtain a grammar to generate the following languages:

a) $L = \{0^i 1^j \mid i \neq j, i \geq 0 \text{ and } j \geq 0\}$

b) $L = \{a^{n+2} b^m \mid n \geq 0 \text{ and } m > n\}$

c) $L = \{a^n b^m \mid n \geq 0, m > n\}$

d) $L = \{a^n b^{n-3} \mid n \geq 3\}$

e) $L = L_1 L_2$ where $L_1 = \{a^n b^m \mid n \geq 0, m > n\}$ $L_2 = \{0^n 1^{2n} \mid n \geq 0 \}$

f) $L = L_1 \cup L_2$ where $L_1 = \{a^n b^m \mid n \geq 0, m > n\}$ $L_2 = \{0^n 1^{2n} \mid n \geq 0 \}$

g) $L = \{w : |w| \bmod 3 \neq |w| \bmod 2\}$ on $\Sigma = \{a\}$

h) $L = \{w : |w| \bmod 3 \geq |w| \bmod 2\}$ on $\Sigma = \{a\}$

12) Obtain a grammar to generate set of all strings with exactly one *a* when $\Sigma = \{a, b\}$.

13) Obtain a grammar generating all strings with at least one *a* if $\Sigma = \{a, b\}$.

14) Obtain a grammar to generate the set of all strings with no more than three *a*'s when $\Sigma = \{a, b\}$.

15) Obtain a grammar to generate the following languages:

a) $L = \{w \mid n_a(w) = n_b(w) + 1\}$

b) $L = \{w \mid n_a(w) > n_b(w)\}$

c) $L = \{a^n b^m c^k \mid n + 2m = k \text{ for } n \geq 0, m \geq 0\}$

d) Strings of 0's and 1's having a substring 000

e) $L = \{a^{n1} b^{n1} a^{n2} b^{n2} \ldots\ldots a^{nk} b^{nk} : n, k \geq 0\}$

16) Obtain a grammar to generate integers. Show the derivation for the unsigned number 1965 and signed number +1965.

17) Is the following grammar ambiguous?

$$S \rightarrow aSb \mid SS \mid \varepsilon$$

18) Show that the following grammar is ambiguous.

$$S \rightarrow SbS \mid a$$

19) Given the following grammar:

$$E \rightarrow I \mid E+E \mid E^*E \mid (E)$$

$$I \rightarrow a \mid b \mid c \mid d$$

Obtain the derivations for the strings (a+b)*c*d and a+b*c and the parse tree for each derivation.

20. Obtain a reduced grammar for the grammar shown below:

S → aAa

A → Sb | bCC | aDA

C → ab | aD |

E → aC

D → aAD

21. Find an equivalent grammar without ε-productions for the grammar shown below.

S → aSa | bSb | A

A → aBb | bBa

B → aB | bB | ε

22. Remove the unit productions from the grammar

S → A | B | Cc

A → aBb | B

B → aB | bb

C → Cc | B

23. Eliminate useless productions from the grammar

S → aA | a | B | C

A → aB | ε

B → aA

C → CcD

D → abd

Note: First remove ε and unit productions, then simplify CFG.

24. Obtain the following grammar in CNF

E → E + T | T

T → T * F | F

F → (E) | I

I → a | b | c | Ia | Ib | Ic

25. Obtain the following grammar in CNF and GNF

S → aA | a | B | C

A → aB | ε

B → aA|

C → cCD

D → abd

26. Simplify the following CFG and convert it into CNF

$S \rightarrow AaB \mid aaB$

$A \rightarrow \varepsilon$

$B \rightarrow bbA \mid \varepsilon$

27. Convert the following grammar into GNF

$S \rightarrow abAB$

$A \rightarrow bAB \mid \varepsilon$

$B \rightarrow Aa \mid \varepsilon$

28. What is left recursion? How it can be eliminated?

29. Eliminate left recursion from the following grammar

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

30. Eliminate the useless symbols in the grammar

$S \rightarrow aA \mid bB$

$A \rightarrow aA \mid a$

$B \rightarrow bB$

$D \rightarrow ab \mid Ea$

$E \rightarrow aC \mid d$

31. Eliminate left recursion from the following grammar

$S \rightarrow Ab \mid a$

$A \rightarrow Ab \mid Sa$

32. What is the need for simplifying a grammar?

33. What is a useless symbol? Explain with example.

34. Simplify the following grammar

$S \rightarrow aA \mid a \mid Bb \mid cC$

$A \rightarrow aB$

$B \rightarrow a \mid Aa$

$C \rightarrow cCD$

$D \rightarrow ddd$

35. What is an ε - production? What is a Nullable variable? Explain with example.

36. Eliminate all ε- productions from the grammar

$S \rightarrow ABCa \mid bD$

$A \rightarrow BC \mid b$

$B \rightarrow b \mid \varepsilon$

$C \rightarrow c \mid \varepsilon$

$D \rightarrow d$

37. Eliminate all ε- productions from the grammar

$S \rightarrow BAAB$

$A \rightarrow 0A2 \mid 2A0 \mid \varepsilon$

$B \rightarrow AB \mid 1B \mid \varepsilon$

38. Obtain a grammar to generate an arithmetic expression using the operators +, -, *, / and ^ (indicating power). An identifier can start with any of the letters from {a, b, c} and can be followed by zero or more symbols from {a, b, c}

39. Define the following terms:

a) Derivation tree   b) yield of a tree   c) ambiguous grammar

40. Is the following grammar ambiguous?

$S \rightarrow iCtS \mid iCtSeS \mid a$

$C \rightarrow b$

41. What is dangling else problem? How dangling else problem can be solved?"

42. Eliminate ambiguity from the following ambiguous grammar:

$S \rightarrow iCtS \mid iCtSeS \mid a$

$C \rightarrow b$

43. Convert the following ambiguous grammar into unambiguous grammar

$E \rightarrow E * E \mid E - E$

$E \rightarrow E \wedge E \mid E / E$

$E \rightarrow E + E$

$E \rightarrow (E) \mid \mathbf{id}$

44. What is inherently ambiguous grammar?"

45. Obtain the inherently ambiguous grammar for the following inherently ambiguous languages:

$L = \{a^n b^n c^m d^m \mid m \geq 1, n \geq 1\} \cup \{a^n b^m c^m d^n \mid m \geq 1, n \geq 1\}$

$L = \{a^i b^j c^k \mid i, j, k, \geq 0\}$ and $i = j$ or $j = k\}$

46. What is left recursion? How to eliminate left recursion?

47. What is useless variable? Explain with example.