SIKSHA O ANUSANDHAN UNIVERSITY

Department Of Comp. Sc. & Engg., ITER, BBSR-30

B.Tech.(5^{th} Sem., CSE &CSIT)

Theory of Computation

Sample Questions -

1. Answer each part for the following context-free grammar G

$$\begin{aligned} R &\rightarrow XRX \mid S \\ S &\rightarrow aTb \mid bTa \\ T &\rightarrow XTX \mid X \mid \epsilon \\ X &\rightarrow a \mid b \end{aligned}$$

   (a) What are the variables of G?
   (b) What are the terminals of G?
   (c) Which is the start variable of G?
   (d) Give three strings in $L(G)$
   (e) Give three strings not in $L(G)$
   (f) True or False: $T \Rightarrow aba$
   (g) True or False: $T \overset{*}{\Rightarrow} aba$
   (h) True or False: $T \Rightarrow T$
   (i) True or False: $T \overset{*}{\Rightarrow} T$
   (j) True or False: $XXX \overset{*}{\Rightarrow} aba$
   (k) True or False: $X \overset{*}{\Rightarrow} aba$
   (l) True or False: $T \overset{*}{\Rightarrow} XX$
   (m) True or False: $T \overset{*}{\Rightarrow} XXX$
   (n) True or False: $S \overset{*}{\Rightarrow} \epsilon$
   (o) Give a description in english of $L(G)$

2. Let G be the context-free grammar

$$\begin{aligned} S &\rightarrow aB \mid bA \\ A &\rightarrow a \mid aS \mid bAA \\ B &\rightarrow b \mid bS \mid aBB \end{aligned}$$

   For the string *aaabbabbba* find a
   (i) leftmost derivation, (ii) rightmost derivation, (iii) parse tree.

3. The following context free grammar generates the grammar of the labguage of all strings of even length:

$$\begin{aligned} S &\rightarrow AS \mid \epsilon \\ A &\rightarrow aa \mid ab \mid ba \mid bb \end{aligned}$$

   Give the leftmost and rightmost derivation for the following strings:
   (i) *aabbba*, (ii) *baabab*, (iii) *aaabbb*

4. The following Context-Free Grammar(CFG) generates the languages consisting of all strings of even length:
   T → BT | BTT | ϵ
   B → aa | ab | ba | bb

   Give leftmost and rightmost derivations for the following strings:

   (a) aabbba
   (b) baabab

5. Here is a context-free grammar:

$$S \rightarrow AS \mid SB \mid 0$$
$$A \rightarrow BA \mid AS \mid 1$$
$$B \rightarrow SB \mid BA \mid 0$$

   Note that each of the right sides AS, SB, and BA occurs twice. Does the string 01100 belonging to the language of the grammar.

6. If the string ***baaba*** is present in the language of the following grammar, Give a leftmost derivation of the string.

$$\begin{aligned} &S \rightarrow AB \mid BC \\ &A \rightarrow BA \mid a \\ &B \rightarrow CC \mid b \\ &C \rightarrow AB \mid a \end{aligned}$$

7. For the given CFG ($G : S \rightarrow AB \mid b,\ A \rightarrow BB \mid a,\ B \rightarrow AB \mid b$). Choose the string that can not be derived from the grammar; *aabbb*, *aabb*, *ababab*,and *abbb*.

8. In the context-free grammar below, S is the start symbol, 'a' and 'b' are terminals and ϵ denotes the empty string:

$$S \rightarrow aSA \mid bSb \mid b \mid a \mid \epsilon$$

Is the string 'babaaabab' a part of the language generated by the given grammar.

9. Consider the CFG G defined by productions:

$$S \rightarrow aS \mid Sb \mid a \mid b$$

   (a) Prove by induction on the string length that no string in $L(G)$ has ba as a substring.

   (b) Describe $L(G)$ informally. Justify your answer using part(a).

10. Consider the CFG G defined by productions:

$$S \rightarrow aSbS \mid bSaS \mid \epsilon$$

Prove that $L(G)$ is the set of all strings with an equal number of a's and b's.

11. For the given two CFGs ($G_1 : S_1 \rightarrow aE \mid b,\ E \rightarrow a$) , ($G_2 : S_2 \rightarrow AB,\ A \rightarrow BC,\ B \rightarrow CC \mid b,\ C \rightarrow AB \mid a$. Compute (i) G_1G_2 (ii) $G_1 \cup G_2$ (iii) Kleen closure($*$ closure) of G_1

12. Give context-free grammars that generate the following languages.

   (a) $L_1 = \left\{ a^n b^m c^m d^{2n} \mid n \geqslant 0,\ m > 0 \right\}, \Sigma = \{a, b, c, d\}$.

   (b) $L_2 = \{ a^n b^m \mid 0 \leq n \leq m \leq 2n \}, \Sigma = \{a, b\}$.

   (c) $L_3 = \{ a^n b^m a^n \mid n > 0,\ m > 0 \}, \Sigma = \{a, b\}$.

   (d) $L_4 = \left\{ a^n b^m c^k \mid k = n + m \right\}, \Sigma = \{a, b, c\}$.

   (e) $L_5 = \left\{ a^n b^m c^k \mid k \neq n + m \right\}, \Sigma = \{a, b, c\}$.

   (f) $L_6 = \{ 10^n 1^n \mid n > 0 \} \bigcup \left\{ 110^n 1^{2n} \mid n > 0 \right\}, \Sigma = \{0, 1\}$.

   (g) $L_7 = \{ w \mid w \textit{ starts and ends with the same symbol} \}, \Sigma = \{0, 1\}$.

   (h) $L_8 = \{ w \mid |w| \textit{ is odd} \}, \Sigma = \{0, 1\}$.

   (i) $L_9 = \{ w \mid |w| \textit{ is odd and its middle symbol is } 0\}, \Sigma = \{0, 1\}$.

   (j) $L_{10} = \left\{ w\#x \mid w^R \textit{ is a substring of } x,\ \textit{where } w, x \in \{ a,\ b \}^* \right\}, \Sigma = \{a, b, \#\}$.

   (k) $L_{11} = \{0^i 1^j 2^k \mid i + j \geq 2k\}, \Sigma = \{0, 1, 2\}$.

   (l) $L_{12} = \{a^i b^j c^k \mid i, k \geq 0, j > 0 \textit{ and } j > i + k\}, \Sigma = \{a, b, c\}$.

   (m) $L_{13} = \{a^i b^j \mid i \leq j \leq 2i\}, \Sigma = \{a, b\}$.

13. Find CFG for $\Sigma = \{a, b\}$ that generate the sets of

   (a) all strings with exactly one a.

   (a) all strings with at least one a.

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

   (a) all strings with at ;east three a's.

14. Find CF for the following languages on $\Sigma = \{a\}$

(a) $L = \{w \mid |w| \bmod 3 = 0\}$

(a) $L = \{w \mid |w| \bmod 3 > 0\}$

(a) $L = \{w \mid |w| \bmod 3 \neq |w| \bmod 2\}$

(a) $L = \{w \mid |w| \bmod 3 \geq |w| \bmod 2\}$

15. Let $L = \{a^n b^n \ : \ n \geq 0\}$.

    (a) Show that L^2 is context-free.

    (b) Show that L^k is context-free for any given $k \geq 1$.

    (c) Show that $\bar{L}$ and L^* are context-free.

16. Design CFG for $L = \{a^i b^j c^k d^l \ : \ i + k = j + l, \ i, k, j, l \geq 0\}$, where $\Sigma = \{a, b, c, d\}$.

    (a) From your CFG determine the start variable and the set of variables.

    (b) Give derivations of the following strings: $a^4bc^3d^6, \ b^5c^6d, \ ad, \ acd^2$.

    (c) For the string $w = a^2b^2c^2d^2$, give different yields that one can get while deriving w from S.

17. Consider the following Context-Free Grammar(CFG) as :

    $$\begin{aligned} &\mathrm{E} \to \mathrm{E} + \mathrm{T} \mid \mathrm{T} \\ &\mathrm{T} \to \mathrm{T} * \mathrm{F} \mid \mathrm{F} \\ &\mathrm{F} \to (\mathrm{E}) \mid \mathrm{id} \end{aligned}$$

    (a) Derive the string; id + id ∗ id .

    (b) Show that this grammar is ambiguous.

18. Convert following CFG to CNF form.

    $$\mathrm{S} \longrightarrow \mathrm{S} \cap \mathrm{S} \mid \mathrm{S} \cup \mathrm{S} \mid \sim\mathrm{S} \mid (\mathrm{S}) \mid \mathrm{p} \mid \mathrm{q}$$

19. Let G be the context-free grammar with productions;

    $$\mathrm{S} \to \mathrm{aS} \mid \mathrm{aSbS} \mid \mathrm{c}$$

    and let $\mathrm{G_1}$ be the other context-free grammar with productions;

    $$\mathrm{S_1} \to \mathrm{T} \mid \mathrm{U}, \ \mathrm{T} \to \mathrm{aTbT} \mid \mathrm{c}, \ \mathrm{U} \to \mathrm{aS_1} \mid \mathrm{aTbU}$$

    (i) Show that G is ambiguous.

    (ii) Show that G and $\mathrm{G_1}$ generate the same language.

    (iii) Convert G into Chomsky Normal Form(CNF).

20. Construct two parse trees corresponding to the given string **2 + 3** ∗ **9** for the CFG G:

    $$\begin{aligned} &\mathrm{S} \to \mathrm{S} + \mathrm{S} \mid \mathrm{S} * \mathrm{S} \mid \mathrm{N} \\ &\mathrm{N} \to 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \end{aligned}$$

21. Suppose G is a CFG in CNF and w is a string in $L(G)$ of length n, then how long is a derivation of w in G. Justify your answer.

22. (a) Write a context-free grammar for the language $L_2 = \{ \alpha \in \{ a, b, c\}^* \mid \#a(\alpha) + \#b(\alpha) = \#c(\alpha)\}$. Here, $\#d(\alpha)$ means the number of occurrences of the symbol d in the string α, where $d \in \{a, b, c\}$. Write only the productions in your grammar, and mention which is the start symbol.

    (b) Convert the grammar of Part (a) to Chomsky normal form. Show all the relevant steps briefly.

23. Prove that $L = \{w \ : \ n_a(w) = n_b(w)$ and w does not contain the substring $aab\}$ is a CFL. Where $n_a(w)$ is the number of a's in w and $n_b(w)$ is the number of b's in w.

24. What do you mean by a regular language and a context free language? For a given grammar(G) $S \to 0S0/00$ find the language of the grammar($L(G)$). Is $L(G)$ a CFL or Regular or both? Justify your answer. What is the set relation between set of CFL and set of RL over an arbitrary Σ.

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

    (a) Is L a CFL? If yes then give context-free grammar that generates L, if no then prove.

    (b) w be a string in L such that $w = aabbccdd$, give a leftmost derivation of w.

(c) If there are any other leftmost derivation of w then derive it. What can you say about ambiguity of the grammar of L.

26. Consider the grammar:

$$S \rightarrow SbS|a$$

(a) Show in particular that the string *abababab a*, has two leftmost derivations and two parse trees.

(b) Can you give any other string which has the above properties.

27. Consider the grammar:

$$\begin{aligned} E &\rightarrow I \mid E+E \mid E*E \mid (E) \\ I &\rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \end{aligned}$$

Give two leftmost derivations and parse trees for each of the sentential form $5 + (E + E * E)$, $(5 + E) * E + E$.

28. Show that the following grammar is ambiguous.

$$\begin{aligned} S &\rightarrow AB \mid aaB \\ A &\rightarrow a \mid Aa \\ B &\rightarrow b \end{aligned}$$

29. Consider the grammar G:

$$\begin{aligned} S &\rightarrow A1B \\ A &\rightarrow 0A \mid \epsilon \\ B &\rightarrow 0B \mid 1B \mid \epsilon \end{aligned}$$

(a) What is $L(G)$?

(b) Construct an *ambiguous* grammar for this $L(G)$ and demonstrate its ambiguity.

30. Recognize the CFL for the given CFG.

$$\begin{aligned} S &\rightarrow aB|bA \\ A &\rightarrow a|aS|bAA \\ B &\rightarrow b|bS|aBB \end{aligned}$$

Reduce it to Chomsky normal form.

31. Evaluate the following statements as *TRUE* or *FALSE* with proper justification:

(a) There exists context free languages(CFL) such that all context free grammars generating them are ambiguous.

(b) An unambiguous context free grammar(CFG) always has a unique parse tree for each string of the language generated by it.

(c) Both deterministic and non-deterministic pushdown automata(DPDA and NDPDA) always accept the same set of languages.

(d) A finite set of strings from some alphabet is always a regular language.

32. Consider the following grammar G

$$\begin{aligned} S &\rightarrow bS \mid aA \mid b \\ A &\rightarrow bA \mid aB \\ B &\rightarrow bB \mid aS \mid a \end{aligned}$$

Say, $N_a(w)$ and $N_b(w)$ denote the number of a's and b's in a string w. The language $L(G) \subseteq \{a, b\}^+$ generated by the grammar G. Describe mathematically $L(G)$ in terms of $N_a(w)$ and $N_b(w)$.

33. Say, L_{nfa} and L_{ndpda} be the classes of languages accepted by non-deterministic finite automata and non-deterministic pushdown automata respectively. Say, L_{dfa} and L_{dpda} be the classes of languages accepted by deterministic finite automata and deterministic pushdown automata respectively. Comment on the following relation and justify it properly:
$L_{nfa} = L_{dfa}$ and $L_{dpda} \subset L_d pda$

34. Convert the following grammar into its equivalent Chomsky Normal Form(CNF):

$$
\begin{aligned}
S &\rightarrow bS \mid aA \mid b \\
A &\rightarrow bA \mid aB \\
B &\rightarrow bB \mid aS \mid a \mid \epsilon
\end{aligned}
$$

35. Given the following context-free grammar

$$
\begin{array}{lcl}
S & \rightarrow & aAa \mid bBb \mid \epsilon \\
A & \rightarrow & C \mid a \\
B & \rightarrow & C \mid b \\
C & \rightarrow & CDE \mid \epsilon \\
D & \rightarrow & A \mid B \mid ab
\end{array}
$$

   (a) Eliminate ϵ rules.
   (b) Eliminate any unit production in the resulting Grammar.
   (c) Eliminate any useless symbol in the resulting Grammar.
   (d) Put the resulting grammar into Chomsky Normal Form.

36. Repeat the above question for the following grammar

$$
\begin{array}{lcl}
S & \rightarrow & AAA \mid B \\
A & \rightarrow & aA \mid B \\
B & \rightarrow & \epsilon
\end{array}
$$

37. Convert the following CFGs into equivalent CFGs in Chomsky normal form (CNF). Also remove all useless variables if exist.

   (a) $A \rightarrow BAB \mid B \mid \epsilon$
   $B \rightarrow 00 \mid \epsilon$
   (b) $S \rightarrow aSa \mid A \mid C$
   $A \rightarrow bBb \mid bCb \mid E$
   $B \rightarrow bBb \mid \epsilon$
   $C \rightarrow aC \mid bC$
   $D \rightarrow aD \mid \epsilon$
   $E \rightarrow bb \mid bEb$
   (c) $S \rightarrow aXbX$
   $X \rightarrow aY \mid bY \mid \epsilon$
   $Y \rightarrow X \mid c$
   (d) $S \rightarrow AbA$
   $A \rightarrow Aa \mid \epsilon$
   (e) $S \rightarrow (S) \mid SS \mid \epsilon$
   (f) $S \rightarrow aTXb$
   $T \rightarrow XTX \mid \epsilon$
   $X \rightarrow a \mid b$

38. Suppose G be the CFG and w, of length l, is in $L(G)$. How long is a derivation of w in G if G is in CNF.