BRAC UNIVERSITY

Kha – 224, Merul Badda, Dhaka, Bangladesh

CSE331 : Automata and Computability

Assignment 03

Summer 2024

Q1. Which of the following statements about Context Free Grammar (CFG) is true? (2 points)

1. Context Free Grammar (CFG) can describe only Context Free Languages.
2. Context Free Grammar (CFG) can describe all the Non Regular Languages and Regular Languages.
3. Context Free Grammar (CFG) can describe both Context Free Languages and Regular Language.

Write an explanation for your answer.

Answer : CFG can describe both context-free languages and regular languages, as regular languages are a subset of context-free languages.

Q2. Consider the following languages.

A) $L = \{ w \in \{a,b,c\}^* : a^n b^n c^n, \text{ where } n \geq 0\}$

B) $L = \{ w \in \{0,1\}^* : ww\}$

C) $L = \{ w \in \{0,1\}^* : ww^R\}$ [w^R means reverse of w. For example, if w = 10110, then w^R = 01101.]

D) $L = \{ w \in \{a,b,c\}^* : a^i b^j c^k, \text{ where } 0 \leq i \leq j \leq k \}$

Which languages are Context Free Languages?

1. Only C
2. B and C
3. C and D
4. A, B and C

Write an explanation for your answer.

Answer: Option C is a palindromes problem.

Q3. Let, L is a context-free language. Then L* will also be a context-free language.

The statement is

1. True
2. False
3. Not enough information to answer

Write an explanation for your answer.

Answer: Kleene star of a context-free language is always context-free.

Q4. Consider the following Context Free Grammar.

S → AB | BC
A → aA | Aa | a | ε
B → bBb | bb
C → c | ε

Choose the correct regular expression for the language covered by the given CFG

1. a*(bb)* | (bb)*c*
2. a^+(bb)* | $(bb)^+$c?
3. $a^+(bb)^+$ | $(bb)^+(c|\varepsilon)$
4. $a^*(bb)^+$ | $(bb)^+(c|\varepsilon)$
5. a^+(bb)* | (bb)*(c|ε)

Write an explanation for your answer.

Answer: A produces a*, B produces sequences of bb and C can either produce c or ε.

Q5. A → 1A | 0A1 | 01

Which of the following string can be generated by the given CFG.

There could be more than one correct answer, choose all of them.

1. 000111
2. 001111
3. 001011
4. 0100111
5. 1000111

Write an explanation for your answer.

1. A → 0A1 → 00A11 → 000111
2. A → 0A1 → 01A1 → 010A11 → 0100111
3. A → 1A → 10A1 → 100A11 → 1000111

Q6. You are given a Context Free Grammar.

S→ AB

A→ aA | ε

B→ Bb | ε

Now consider the following languages.

a) L = { w ∈ {a,b}*: $a^n b^m$, where n,m ≥ 0 }

b) L = { w ∈ {a,b}*: $a^n b^n$, where n ≥ 0 }

c) L = { w ∈ {a,b}*: number of a and number of b is unequal in w}

d) L = { w ∈ {a,b}*: all a in w always precede b}

Which of the following statements are true regarding the language generated by this grammar?

Note, for a language L, the CFG will be correct if and only if it can parse all the strings, w ∈ L, and doesn't parse any string, w ∉ L.

1. Only A
2. Only B
3. A and C
4. A and D
5. B and D

Write an explanation for your answer.

Answer : As S → AB it can generates $a^n b^m$ successfully and all a always precede b.

Q7. L = { w ε {a,b,c}*: $a^m b^n c^k$ | n,m ≥ 0 & k=3m+2}

Let's say, we have the following four CFGs labeled as (A) to (D).

CFG A:

S → aScc | Y

Y → bY | ccc

CFG B:

S → aSccc | Y

Y → bY | cc

CFG C:

$S \rightarrow Xcc$

$X \rightarrow aXccc \mid Z$

$Z \rightarrow bZ \mid \varepsilon$

CFG D:

$S \rightarrow aSccc \mid bS \mid cc$

What will be the correct CFG for the language L?

Note, for a language L, the CFG will be correct if and only if it can parse all the strings, $w \in L$, and doesn't parse any string, $w \notin L$.

1. Only B
2. Only C
3. A and C
4. A and D
5. B and C
6. B and D

Write an explanation for your answer.

Answer : B and C generates all the strings where number of c's = 3(number of a's) + 2 and they can parse all the strings, $w \in L$, and doesn't parse any string, $w \notin L$.

Q8. $L = \{ w \in \{a,b\}^* : w = a^i b^j \text{ where } i = 2+j, j \geq 0\}$

Let's say, we have the following four CFGs labeled as (A) to (C).

CFG A:
$S \rightarrow aaN$

$N \rightarrow aNb \mid \varepsilon$

CFG B:
$S \rightarrow aSb \mid aa$

CFG C:
$S \rightarrow aSb \mid N$

$N \rightarrow aaN \mid \varepsilon$

What will be the correct CFG for the language L?

Note, for a language L, the CFG will be correct if and only if it can parse all the strings, w ∈ L, and doesn't parse any string, w ∉ L.

1. A and C
2. A and B
3. B and C
4. A, B and C

Write an explanation for your answer.

Answer : CFG A & B always generates those strings where i = 2 + j and they can parse all the strings, w ∈ L, and doesn't parse any string, w ∉ L.

Q9. Consider the following two Context Free Grammars. Do these two grammars (G1 and G2) represent the same language?

G1:	G2:
$S \rightarrow AB$	$S \rightarrow ABC$
$A \rightarrow 1A00 \mid C$	$A \rightarrow 1A00 \mid \varepsilon$
$C \rightarrow 0C \mid 0$	$B \rightarrow 0B \mid 0$
$B \rightarrow 1B \mid \varepsilon$	$C \rightarrow 1C \mid \varepsilon$

1. Yes
2. No

Write an explanation for your answer.

Answer : G1 and G2 always generates same strings, so they represent the same language.

Q10. L = { w ∈ {a,b,c}*: $a^m b^n c^k$ | n = m/2 or n = k - 2 ; where n ≥ 0 , m is even & k ≥ 2}

Let's say, we have the following 4 CFGs labeled as (A) to (D).

CFG A:
$S \rightarrow XY \mid PQ$
$X \rightarrow aXbb \mid \varepsilon$
$Y \rightarrow ccY \mid c$

$P \rightarrow aP \mid b$
$Q \rightarrow bQc \mid c$

CFG B:

$S \rightarrow XY \mid PQ$
$X \rightarrow aXb \mid \varepsilon$
$Y \rightarrow cYb \mid c$
$P \rightarrow aP \mid \varepsilon$
$Q \rightarrow bQcc \mid c$

CFG C:

$S \rightarrow XY \mid PQ$

$X \rightarrow aXb \mid c$
$Y \rightarrow cYb \mid \varepsilon$
$P \rightarrow aP \mid \varepsilon$
$Q \rightarrow bQcc \mid c$

CFG D:
$S \rightarrow XY \mid PQ$
$X \rightarrow aaXb \mid \varepsilon$
$Y \rightarrow cY \mid cc$
$P \rightarrow aaP \mid \varepsilon$
$Q \rightarrow bQc \mid cc$

What will be the correct CFG for the language L?

Note, for a language L, the CFG will be correct if and only if it can parse all the strings, $w \in L$, and doesn't parse any string, $w \notin L$.

1. A
2. B
3. C
4. D

Write an explanation for your answer.

Answer : D always generates those strings where $n = m/2$ or $n = k - 2$; where $n \geq 0$, m is even & $k \geq 2$.

Q11. $L = \{w \in \{0,1\}^*: w_1w_2w_3$: where $|w_1| = |w_2|$ or $|w_3| = |w_1|\}$

Let's say, we have the following 4 CFGs labeled as (A) to (D).

CFG A:
$S \rightarrow RQ \mid P$
$P \rightarrow 0P1 \mid 1P0 \mid \varepsilon$
$Q \rightarrow 0Q \mid 1Q \mid \varepsilon$
$R \rightarrow 0R0 \mid 1R1 \mid 0R1 \mid 1R0 \mid Q$

CFG B:
$S \rightarrow PR \mid Q$
$P \rightarrow 0P0 \mid 1P1 \mid \varepsilon$
$Q \rightarrow 0Q \mid 1Q \mid \varepsilon$
$R \rightarrow 0R0 \mid 1R1 \mid 0R1 \mid 1R0 \mid Q$

CFG C:
$S \rightarrow PQ \mid R$
$P \rightarrow 0P0 \mid 1P1 \mid 0P1 \mid 1P0 \mid \varepsilon$
$Q \rightarrow 0Q \mid 1Q \mid \varepsilon$
$R \rightarrow 0R0 \mid 1R1 \mid 0R1 \mid 1R0 \mid Q$

CFG D:

$S \rightarrow PQ \mid R$
$P \rightarrow 0P0 \mid 1P1 \mid 0P1 \mid 1P0 \mid \varepsilon$
$Q \rightarrow 0Q \mid 1Q$
$R \rightarrow 0R0 \mid 1R1 \mid 0R1 \mid 1R0 \mid P$

What will be the correct CFG for the language L?
Note, for a language L, the CFG will be correct if and only if it can parse all the strings, $w \in L$, and doesn't parse any string, $w \notin L$.

1. A and C
2. B and C
3. C and D
4. A, B and C
5. A, B, C and D

Write an explanation for your answer.

Answer : A, B and C always generates those strings where $|w_1| = |w_2|$ or $|w_3| = |w_1|$

Q12. $L = \{w \varepsilon \{0,1\}^*: w_1\#w_2\#w_3 : \text{where } w_3 = w_1^R \}$

Given a string w over some alphabet Σ, let w^R be its reverse. For example, if w = 10110, then $w^R = 01101$.

Let's say, we have the following 4 CFGs labeled as (A) to (C).

CFG A:

$S \rightarrow 1S0 \mid 0S1 \mid \#P\#$

$P \rightarrow 0P \mid 1P \mid \varepsilon$

CFG B:

$S \rightarrow 0S \mid 1S \mid \#P\#$

$P \rightarrow 0P0 \mid 1P1 \mid \varepsilon$

CFG C:

$S \rightarrow 0S0 \mid 1S1 \mid \#P\#$

$P \rightarrow 0P \mid 1P \mid \varepsilon$

What will be the correct CFG for the language L?

Note, for a language L, the CFG will be correct if and only if it can parse all the strings, $w \in L$, and doesn't parse any string, $w \notin L$.

1. A
2. B
3. C

Write an explanation for your answer.

Answer : C always generates those strings ($w_1\#w_2\#w_3$) where $w_3 = w_1^R$

Q13. $L = \{ w \varepsilon \{0,1\}^*: w \in \Sigma^* \mid$ every prefix of w has at least as many 0's as 1's $\}$

Clarification: L states "Every prefix of w contains equal or more 0s than 1s."

Select the correct CFG for L:

A) $S \rightarrow 0S1S \mid 0S \mid \varepsilon$

B) $S \rightarrow SS \mid 0S1 \mid 1 \mid \varepsilon$

C) $S \rightarrow SS \mid 0S1 \mid 0 \mid \varepsilon$

D) $S \rightarrow 0S1S \mid 1S \mid \varepsilon$

A prefix is a string consisting of several first letters of a given string, without any reorders. An empty prefix is also a valid prefix. For example, the string "abcd" has 5 prefixes: empty string, "a", "ab", "abc" and "abcd".

1. a and d
2. a and b
3. a and c
4. a, b and d
5. a, b and c
6. a, b, c and d

Write an explanation for your answer.

Answer : a and c always generate those strings where every prefix of w has at least as many 0's as 1's.

Q14. Let $\Sigma = \{a, b\}$ and let $L = \{ a^n b^m \mid n, m \in \mathbb{N} \text{ and } n \leq m \leq 5n \}$. The CFG for L is given below.

Here is a CFG of the language

$S \rightarrow aSb? \mid \varepsilon$
$A \rightarrow b \mid \varepsilon$

What will be the '?' such that the grammar can describe the language, L correctly.

Fill out the question mark with one or more variables/terminals. Write down the missing string only.

Write an explanation for your answer.

Answer: Replace the ? with AAAA and it will correctly describe the language L

Q15.

$E \rightarrow E + P \mid P$
$P \rightarrow P * Q \mid Q$
$Q \rightarrow id$

Is the grammar ambiguous?

1. Yes
2. No

Write an explanation for your answer.

Answer : This grammar has only 1 Left Most Derivation (LMD), 1 Right Most Derivation (RMD) and 1 Parse Tree (PT).

Q16. You are given two Context Free Grammar G1 and G2.

G1:
$S \rightarrow (S)S \mid \varepsilon$

G2:
S→(S) | SS | ε

Now, consider the string w = (( ) (( ))) ((( ) ( ))) ( )

Which of the following statements is true?

1. w can be generated using only G1
2. w can be generated using only G2
3. w can be generated using G1 and G2
4. w cannot be generated using either G1 or G2

Write an explanation for your answer.

Answer : G1 and G2 both can generate the string w because both can handle balanced parentheses.

Q17. Consider the two Context free grammars G1 and G2. Which of the following is L(G1) ∩ L(G2)?

G1:	G2:
S → ASA	S → ASA
S → 0	S → ε
A → 0 | 1	A → 0 | 1

1. L = {w ∈ Σ* | w contains 0 in the middle.}
2. L = {w ∈ Σ* | length of w is even.}
3. L = {w ∈ Σ* | w is a even length string with 0 in the middle.}
4. L = {w ∈ Σ* | w is a odd length string with 0 in the middle.}
5. None of the above.

Write an explanation for your answer.

Answer : Length of G1 is odd and length of G2 is even so their intersection is not possible.

Q18. Consider the following Context Free Grammar, G.

G:
P → XP1 | XQ0
Q → XQ1 | XR0
R → XRX | #
X → 0 | 1

Now, answer the following questions.

What is the shortest string can be parsed from the grammar, G? If there are multiple correct answer write only one of those.

Answer : 01#00

Write an explanation for your answer.

Answer : P → XQ0 → 0Q0 → 0XR00 → 01R00 → 01#00

The length of a string, w can be expressed by |w|.

Find out how many distinct strings can be generated using the given Context Free Grammar, G, such that 9 <= |w| <= 11. Write the numeric value only.

Answer : 1008

Write an explanation for your answer.

Answer : Using the grammar production rules, by counting and summing all the possible combinations, we obtain 1008 distinct strings with lengths between 9 and 11.

Q19. Consider the following Context Free Grammar, G.

G:

S → aaS | abS | baS | bbS | X
X → aaY | baY
Y → aY | bY | ε

Write a eight length string that starts with "bb" and has only one parse tree. Write the string only.

Answer : bbababba

Write an explanation for your answer.

Answer : S → bbS → bbabS → bbababS → bbababX → bbababbaY → bbababba

Q20. L = {w ε {a,b,c}*: $a^i b^j c^k$| i, j, k ≥ 0 and if i > j then k = i - j, else k = 0}

Which of the following CFGs generates L?

CFG A:
S → AB | AC
A → aA | ε
B → bB | b
C → aCc | ε

CFG B:
S → AB | A
A → aA | a | ε
B → bBc | b

CFG C:
S → A | B
A → aAb | Ab | ε
B → aBc | aCc
C → aCb | ε

CFG D:
S → AB | AC
A → aAb | ε
B → bB | ε
C → aCc | ε

Note, for a language L, the CFG will be correct if and only if it can parse all the strings, w ∈ L, and doesn't parse any string, w ∉ L.

1. A
2. B
3. C
4. D

Write an explanation for your answer.

Answer : C always generates those strings ($a^i b^j c^k$) where i, j, k ≥ 0 and if i > j then k = i - j, else k = 0

Q21. L = {w ε {a, b}*: each b in w is followed by at least two a}

Note, for a language L, the CFG will be correct if and only if it can parse all the strings, w ∈ L, and doesn't parse any string, w ∉ L.

Which of the following Context Free Grammar can generate L?

S → XY
X → aX | ε

$Y \rightarrow ZY \mid \varepsilon$
$Z \rightarrow PQ$
$P \rightarrow ba$
$Q \rightarrow aQ \mid a$

1. Correct
2. Incorrect

$P \rightarrow aP \mid bQ \mid \varepsilon$
$Q \rightarrow aR \mid bT$
$R \rightarrow aS \mid bT$
$S \rightarrow aS \mid bQ \mid \varepsilon$
$T \rightarrow aT \mid bT$

1. Correct
2. Incorrect

$S \rightarrow AbR \mid A$
$A \rightarrow aA \mid \varepsilon$
$R \rightarrow aAabR \mid A$

1. Correct
2. Incorrect

ALL ANSWERS ARE NOT CORRECT