

## Context-Free Languages: Answers

1.

- a. The following strings can be produced: aa, aab, aba, baa
- b. Four distinct derivations are as follows:

$S \Rightarrow AA \Rightarrow bAA \Rightarrow baA \Rightarrow babA \Rightarrow babbA \Rightarrow babbAb \Rightarrow babbab$

$S \Rightarrow AA \Rightarrow AAb \Rightarrow Aab \Rightarrow Abab \Rightarrow Abbab \Rightarrow bAbbab \Rightarrow babbab$

$S \Rightarrow AA \Rightarrow bAA \Rightarrow bAAb \Rightarrow bAbAb \Rightarrow bAbbAb \Rightarrow babbAb \Rightarrow babbab$

$S \Rightarrow AA \Rightarrow AAb \Rightarrow bAAb \Rightarrow bAbAb \Rightarrow babAb \Rightarrow babbAb \Rightarrow babbab$

2.

- a. The context-free grammar with start symbol S is as follows:

$S \rightarrow aSa$

$S \rightarrow bSb$

$S \rightarrow c$

- b. The context-free grammar with start symbol S is as follows:

$S \rightarrow aSa$

$S \rightarrow bSb$

$S \rightarrow \epsilon$

- c. The context-free grammar with start symbol S is as follows:

$S \rightarrow aSa$

$S \rightarrow bSb$

$S \rightarrow a$

$S \rightarrow b$

$S \rightarrow \epsilon$

3.

- a. The derivation of the string ababba is as follows:

$S \Rightarrow aB \Rightarrow abS \Rightarrow abaB \Rightarrow ababS \Rightarrow ababbA \Rightarrow ababba$

- b. In order to prove this, we need to jointly prove the following:

All strings derived from S have the same number of a's and b's

All strings derived from A have one more a than b

All strings derived from B have one more b than a

**Base cases:** the base cases are those production rules which have no non-terminals on the right-hand side. These are as follows:

$A \rightarrow a$

$B \rightarrow b$

The required property follows immediately from these rules.

**Inductive hypothesis:** we assume that the required property is true for S, A and B.

**Inductive cases:** we show that the required property follows from the inductive hypothesis for those production rules which have non-terminals on the right-hand side

$S \rightarrow aB$

B generates strings which have one more b than a (by the inductive hypothesis). Therefore, strings generated from the right-hand side will contain the same number of a's and b's

$S \rightarrow bA$

A generates strings which have one more b than a (by the inductive hypothesis). Therefore, strings generated from the right-hand side will contain the same number of a's and b's

$A \rightarrow aS$

S generates strings which have the same number of a's and b's (by the inductive hypothesis). Therefore, strings generated from the right-hand side will contain one more a than b

$A \rightarrow BAA$

A generates strings which have one more a than b (by the inductive hypothesis). B generates strings which have one more b than a (by the inductive hypothesis). Therefore, strings generated from the right-hand side will contain one more a than b

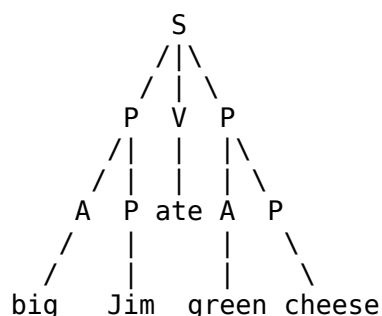
$B \rightarrow bS$

S generates strings which have the same number of a's and b's (by the inductive hypothesis). Therefore, strings generated from the right-hand side will contain one more b than a

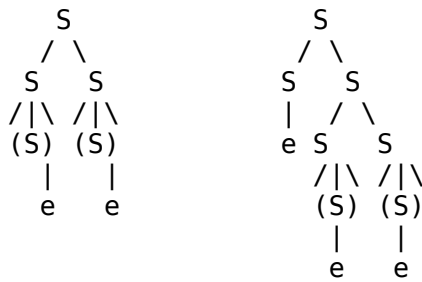
$B \rightarrow ABB$

A generates strings which have one more a than b (by the inductive hypothesis). B generates strings which have one more b than a (by the inductive hypothesis). Therefore, strings generated from the right-hand side will contain one more b than a

4. The parse tree is as follows:



5. Two different parse trees are as follows:



An equivalent unambiguous grammar is as follows:

$S \rightarrow ST$   
 $S \rightarrow \epsilon$   
 $T \rightarrow (S)$

6.

- a. Yes
- b. No
- c. No
- d. No
- e. No
- f. Yes
- g. Yes
- h. Yes
- i. Yes
- j. No

7. The pushdown automaton with start state  $s$  and accepting state  $f$  has the following transitions:

$((s, (\epsilon), (s, ()))$   
 $((s, (, (s, \epsilon))$   
 $((s, \epsilon, \epsilon), (f, \epsilon))$

8. The context-free grammar with start symbol  $S$  is as follows:

$S \rightarrow aSa$   
 $S \rightarrow aSb$   
 $S \rightarrow bSb$   
 $S \rightarrow bSa$   
 $S \rightarrow a$

9. The pushdown automaton with start state  $s$  and accepting state  $f$  has the following transitions:

$((s, \epsilon, \epsilon), (q1, \epsilon))$   
 $((s, \epsilon, \epsilon), (q4, \epsilon))$   
 $((q1, \epsilon, \epsilon), (q3, \epsilon))$   
 $((q1, a, \epsilon), (q1, a))$   
 $((q1, b, a), (q2, \epsilon))$

((q2,b,a),(q2, $\epsilon$ ))  
 ((q2, $\epsilon$ , $\epsilon$ ),(q3, $\epsilon$ ))  
 ((q3,c, $\epsilon$ ),(q3, $\epsilon$ ))  
 ((q3, $\epsilon$ , $\epsilon$ ),(f, $\epsilon$ ))  
 ((q4, $\epsilon$ , $\epsilon$ ),(f, $\epsilon$ ))  
 ((q4,a, $\epsilon$ ),(q4, $\epsilon$ ))  
 ((q4,b, $\epsilon$ ),(q5,b))  
 ((q5,b, $\epsilon$ ),(q5,b))  
 ((q5,c,b),(q6, $\epsilon$ ))  
 ((q6,c,b),(q6, $\epsilon$ ))  
 ((q6, $\epsilon$ , $\epsilon$ ),(f, $\epsilon$ ))

10. In order to show that a language is not context-free using the pumping lemma, we must show that there are strings  $v$ ,  $w$ ,  $x$ ,  $y$  and  $z$  such that  $|wy| > 0$  and  $vw^kxy^kz$  is an element of the language for all  $k \geq 0$ . We assume that the language is context-free and obtain a contradiction.

- a. Let  $s = a^n b^{n+1} c^{n+2}$ .  $s$  must equal  $vwxyz$ , where  $v$ ,  $w$ ,  $x$ ,  $y$  and  $z$  satisfy the conditions above.

If  $wy$  contains at least one  $a$ , then it can contain no  $c$ 's. Therefore,  $vw^2xy^2z$  must contain at least  $n+1$   $a$ 's and exactly  $n+2$   $c$ 's, which is impossible if the string is in the language.

If  $wy$  contains no  $a$ 's, then it must contain either  $b$  or  $c$ . Therefore,  $vw^0xy^0z = vxz$  has either fewer than  $n+1$   $b$ 's or fewer than  $n+2$   $c$ 's, but in either case exactly  $n$   $a$ 's, which is impossible if the string is in the language.

- b. Let  $s = a^n b^n c^n$ .  $s$  must equal  $vwxyz$ , where  $v$ ,  $w$ ,  $x$ ,  $y$  and  $z$  satisfy the conditions above.

If  $wy$  contains at least one  $a$ , then it can contain no  $c$ 's. Therefore,  $vw^0xy^0z = vxz$  must contain less than  $n$   $a$ 's, but exactly  $n$   $c$ 's, which is impossible if the string is in the language.

If  $wy$  contains no  $a$ 's, then  $vw^2xy^2z$  contains either more than  $n$   $b$ 's or more than  $n$   $c$ 's, but exactly  $n$   $a$ 's, which is impossible if the string is in the language.