CS 181 PRACTICE MIDTERM 2A

You may state without proof any fact taught in lecture.

1 Describe the languages generated by the following context-free grammars with alphabet $\{a, b\}$. You may provide a verbal description or a regular expression, as appropriate.

a. $S \to XY$
$X \to Xa \mid Xb \mid a$
$Y \to aY \mid bY \mid b$

b. $S \to aT \mid bT \mid \varepsilon$
$T \to aS$

c. $S \to \Sigma S \mid aS_1$
$S_1 \to \Sigma S_1 \mid aS_2$
$S_2 \to \Sigma S_2 \mid aS_3$
$S_3 \to \Sigma S_3 \mid \varepsilon$
$\Sigma \to a \mid b$

2 Give context-free grammars for the following languages:

a. strings of the form $a^n \# a^m$, where n and m are nonnegative integers with $n \neq m$;

b. strings over the alphabet $\{a, b\}$ in which some prefix contains more b's than a's.

3 Draw a pushdown automaton for the complement of the language $\{a^n b^n : n \geqslant 0\}$ over the alphabet $\{a, b\}$.

4 Consider the context-free grammar

$$\begin{aligned} S &\to \Sigma S \Sigma \mid \Sigma T a \\ T &\to \Sigma\Sigma T \mid \varepsilon \\ \Sigma &\to a \mid b. \end{aligned}$$

a. Describe the language generated by this grammar.

b. Prove that this grammar is ambiguous.

c. Give an equivalent unambiguous grammar.

5 For languages A and B, define $A \diamond B = \{uv : u \in A, v \in B, |u| = |v|\}$. Explain how to construct a PDA for $A \diamond B$ from DFAs for A and B.

6 Prove or disprove the following claim: if the grammar (V, Σ, R, S) generates L, then the grammar $(V, \Sigma, R \cup \{S \to SS \mid \varepsilon\}, S)$ generates L^*.

7 For each of the following languages L over the alphabet $\{a, b\}$, determine whether it is context-free and prove your answer:

a. $\{a^n b a^m b a^k : m > n + k\}$;

b. strings that do not contain $aaba$;

c. strings that begin with $a^n b^n a^n$ for some $n \geqslant 1$.

SOLUTIONS

CS 181 PRACTICE MIDTERM 2A

You may state without proof any fact taught in lecture.

1 Describe the languages generated by the following context-free grammars with alphabet $\{a, b\}$. You may provide a verbal description or a regular expression, as appropriate.

a. $S \to XY$
$X \to Xa \mid Xb \mid a$
$Y \to aY \mid bY \mid b$

b. $S \to aT \mid bT \mid \varepsilon$
$T \to aS$

c. $S \to \Sigma S \mid aS_1$
$S_1 \to \Sigma S_1 \mid aS_2$
$S_2 \to \Sigma S_2 \mid aS_3$
$S_3 \to \Sigma S_3 \mid \varepsilon$
$\Sigma \to a \mid b$

Solution:

a. strings that begin with a and end with b;

b. $(aa \cup ba)^*$;

c. strings with at least three a's.

2 Give context-free grammars for the following languages:

a. strings of the form $a^n\#a^m$, where n and m are nonnegative integers with $n \neq m$;

b. strings over the alphabet $\{a, b\}$ in which some prefix contains more b's than a's.

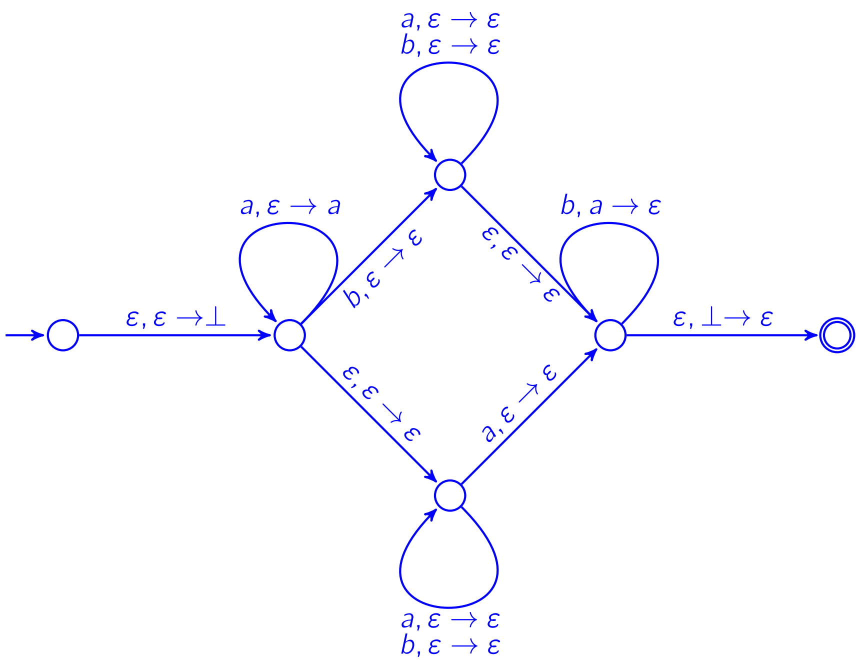
Solution.

a. $S \to AT \mid TA$
$T \to aTa \mid \#$
$A \to aA \mid a$

b. $S \to EbX$
$E \to aEbE \mid bEaE \mid \varepsilon$ (generates all strings with equally many a's and b's)
$X \to aX \mid bX \mid \varepsilon$ (generates all strings)

3 Draw a pushdown automaton for the complement of the language $\{a^n b^n : n \geqslant 0\}$ over the alphabet $\{a, b\}$.

Solution. The complement of $\{a^n b^n : n \geqslant 0\}$ is given by the grammar $S \to aSb \mid bX \mid Xa$, where X generates Σ^*. This gives the following PDA.



4 Consider the context-free grammar

$$\begin{aligned} &S \to \Sigma S \Sigma \mid \Sigma T a \\ &T \to \Sigma\Sigma T \mid \varepsilon \\ &\Sigma \to a \mid b. \end{aligned}$$

a. Describe the language generated by this grammar.

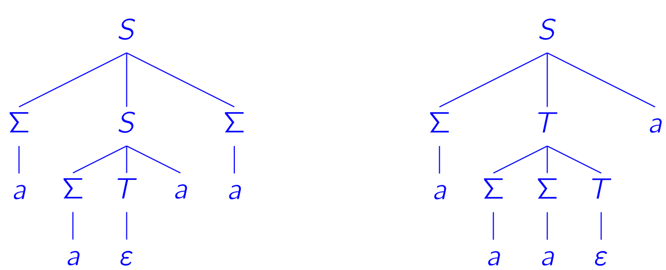
b. Prove that this grammar is ambiguous.

c. Give an equivalent unambiguous grammar.

Solution.

a. Even-length strings that contain an a in their second half.

b. The string $aaaa$ has at least two parse trees:

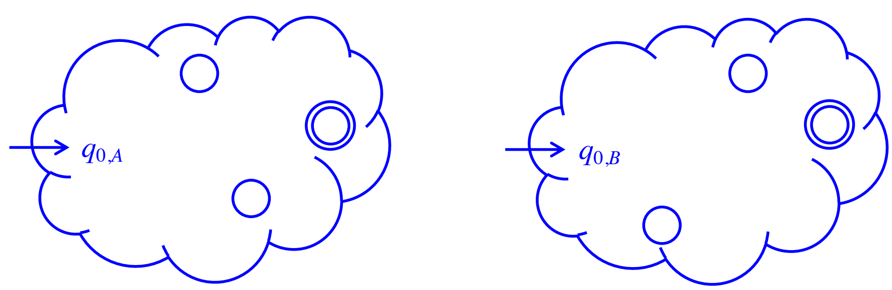


c. Change the first rule to $S \to \Sigma S b \mid \Sigma T a$.

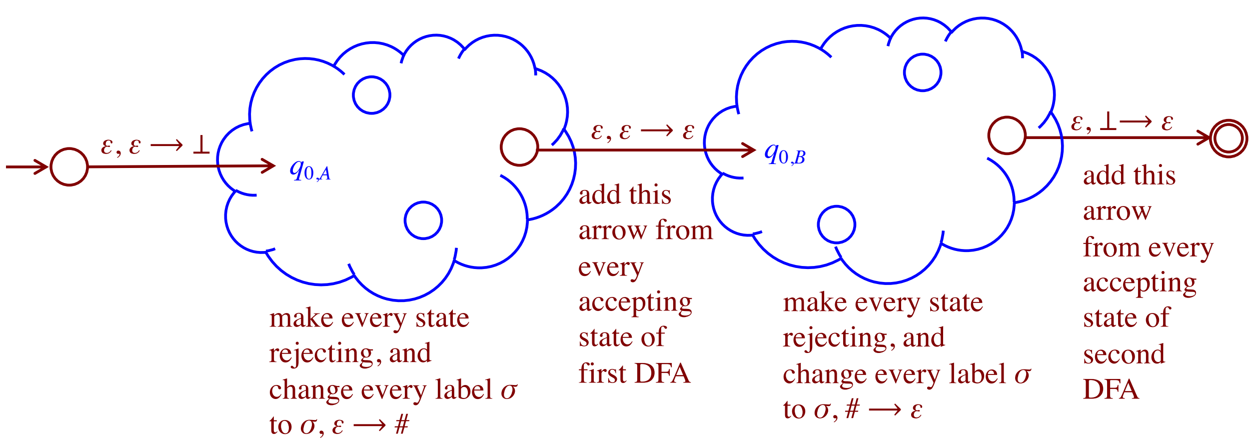
5 For languages A and B, define $A \diamond B = \{uv : u \in A, v \in B, |u| = |v|\}$. Explain how to construct a PDA for $A \diamond B$ from DFAs for A and B.

Solution. Let D_A and D_B be DFAs for A and B, respectively. A PDA for $A \diamond B$ starts by simulating D_A on the first half of the input, using the stack to keep track of the number of symbols read so far. When at an accepting state of D_A, the PDA has the option of proceeding to the next stage. In this second stage, the PDA simulates D_B on the second half of the input, using the stack to ensure that this second half has as many symbols as the first half. When at an accepting state of D_B, the PDA accepts provided that the stack is empty.

The formal construction is as follows. Start with the DFAs for A and B, shown schematically below with their start states $q_{0,A}$ and $q_{0,B}$:



The PDA for $A \diamond B$ is obtained by applying the changes in red:



6 Prove or disprove the following claim: if the grammar (V, Σ, R, S) generates L, then the grammar $(V, \Sigma, R \cup \{S \to SS \mid \varepsilon\}, S)$ generates L^*.

Solution. The claim is false. As a counterexample, consider the grammar $S \to aSb \mid \#$ that generates the language $L = \{a^n \# b^n : n \geqslant 0\}$. The grammar $S \to aSb \mid \# \mid SS \mid \varepsilon$ generates not L^* but a strictly larger language, containing in particular the string ab.

7 For each of the following languages L over the alphabet $\{a, b\}$, determine whether it is context-free and prove your answer:

a. $\{a^n ba^m ba^k : m > n + k\}$;

b. strings that do not contain $aaba$;

c. strings that begin with $a^n b^n a^n$ for some $n \geqslant 1$.

Solution.

a. Context-free. Rewriting, $L = \{(a^n ba^n)a^+(a^k ba^k) : n, k \geqslant 0\}$. Hence the grammar

$$S \to TAT$$
$$T \to aTa \mid b$$
$$A \to aA \mid a.$$

b. Context-free. The set of strings that contain $aaba$ has regular expression $\Sigma^* aaba\Sigma^*$ and is therefore regular. Since regular languages are closed under complement, the language $\overline{\Sigma^* aaba\Sigma^*} = L$ is also regular. This makes L context-free, since every regular language is context-free.

c. Not context-free. Take an arbitrary integer $p \geqslant 1$ and consider the string $w = a^p b^p a^p \in L$. Fix any decomposition $w = uvxyz$ for some strings u, v, x, y, z with $|v| + |y| \neq 0$ and $|vxy| \leqslant p$.

Case 1. If $|u| \geqslant p$, then pumping down produces a string of the form $a^p b^* a^*$ and length less than $3p$. Thus, $uxz \notin L$.

Case 2. If $|u| \leqslant p - 1$, there are three possibilities to consider.

Case 2a. If $v \in a^+ b^+$, then pumping up produces a string that starts with $a^p b^q a^+$ for some $q < p$. Thus, $uv^2xy^2z \notin L$.

Case 2b. If $y \in a^+ b^+$, then pumping up produces a string that starts with $a^p a^* b^q a^+$ for some $q < p$. Thus, $uv^2xy^2z \notin L$.

Case 2c. If $v \notin a^+ b^+$ and $y \notin a^+ b^+$, then pumping up produces a string of the form $a^+ b^+ a^p$ of length greater than $3p$. Thus, $uv^2xy^2z \notin L$.

By the contrapositive form of the pumping lemma, L is not context-free. Note that it is a mistake to pump up in Case 1, or to pump down in Case 2.