CMPS 130
Final Review Problems

These problems are meant as review of the material covered since the last midterm. Homework assignments 6, 7 and 8, and their solutions, should also be considered as review for this material. Consult the previous review sheets and midterm solutions for review of prior material.

1. Write down productions for CFGs that generate the following languages over {a, b}. In each case make up a few strings in the language and give derivations of those strings.

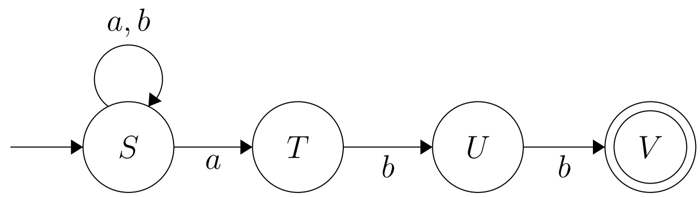
   a. $NonPal = \{x \in \{a,b\}^* | x^r \neq x\}$
   $S \to aSa \mid bSb \mid aUb \mid bUa$
   $U \to \lambda \mid aU \mid bU$
   $S \Rightarrow aUb \Rightarrow a\lambda = ab$
   $S \Rightarrow aSa \Rightarrow aaSaa \Rightarrow aaaUbaa \Rightarrow aaa\lambda baa = aaabaa$
   $S \Rightarrow bUa \Rightarrow baUa \Rightarrow baaUa \Rightarrow baabUa \Rightarrow baabbUa \Rightarrow baabbaUa \Rightarrow baabba\lambda a = baabbaa$

   b. $AeqB = \{x \in \{a,b\}^* | n_a(x) = n_b(x)\}$
   $S \to aA \mid bB \mid \lambda$
   $A \to bS \mid aAA$
   $B \to aS \mid bBB$
   $S \Rightarrow aA \Rightarrow abS \Rightarrow ab\lambda = ab$
   $S \Rightarrow aA \Rightarrow abS \Rightarrow abaA \Rightarrow ababS \Rightarrow abab\lambda = abab$
   $S \Rightarrow aA \Rightarrow abS \Rightarrow abbB \Rightarrow abbaS \Rightarrow abba\lambda = abba$

   c. $AgtB = \{x \in \{a,b\}^* | n_a(x) > n_b(x)\}$
   $S \to B \mid BS$
   $T \to aA \mid bB \mid \lambda$
   $A \to bT \mid aAA$
   $B \to aT \mid bBB$
   $S \Rightarrow B \Rightarrow aT \Rightarrow a\lambda = a$
   $S \Rightarrow BS \Rightarrow BB \Rightarrow aTaT \Rightarrow a\lambda a\lambda = aa$
   $S \Rightarrow B \Rightarrow aT \Rightarrow abB \Rightarrow abaT \Rightarrow aba\lambda = aba$

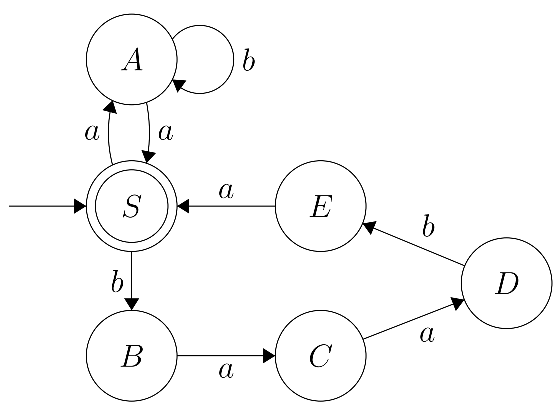
2. Write down productions for regular grammars that generate the following regular languages over $\{a, b\}$. Notation: $L(r)$ denotes the regular language corresponding to the regular expression r. Hint: first draw an NFA for the language, eliminate any λ-transitions, then write down a regular grammar.

   a. $L((a + b)^*abb)$



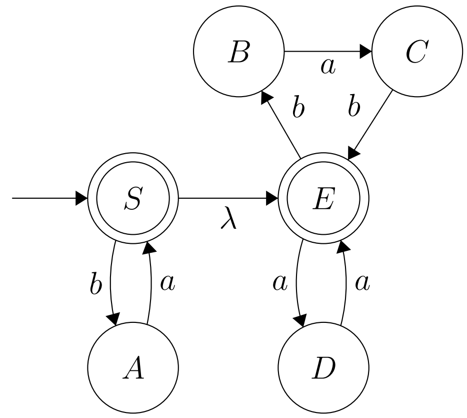
$S \to aS \mid bS \mid aT$
$T \to \ \mid bU$
$U \to bV$
$V \to \lambda$

   b. $L((ab^*a + baaba)^*)$

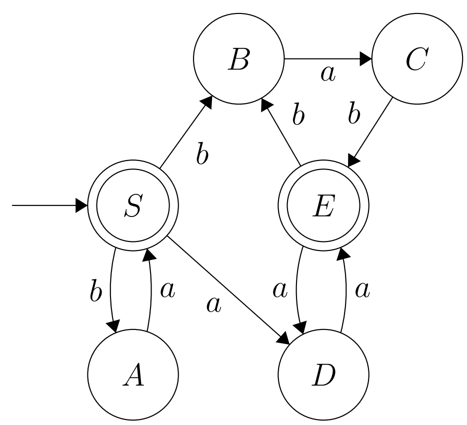


$S \to aA \mid bB \mid \lambda$
$A \to aS \mid bA$
$B \to bC$
$C \to aD$
$D \to bE$
$E \to aS$

c. $L((ba)^*(bab + aa)^*)$



	a	b	λ		a	b
S	—	A	E		D	A,B
A	S	—	—		S	—
B	C	—	—		C	—
C	—	E	—		—	E
D	E	—	—		E	—
E	D	B	—		D	B



$S \to bA \mid bB \mid aD \mid \lambda$
$A \to aS$
$B \to aC$
$C \to bE$
$D \to aE$
$E \to aD \mid bB \mid \lambda$

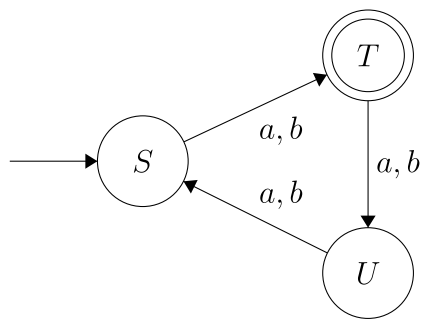
3. Write productions for a CFG that generates the language $L = \{a^i b^j |\ 2i = 3j\}$. Hint: look at hw7 solutions problem 4.10c.

   Let $i = 3n$, $j = 2n$, $\forall n \in \mathbb{N}$, then $2i = 2(3n) = 6n = 3(2n) = 3j$
   Then $S \Rightarrow^n a^{3n}b^{2n}$

   $S \rightarrow aaaSbb \mid \lambda$

4. Describe the language (a subset of $\{a, b\}^*$) generated by the CFG with the following productions. Hint: consider hw7 solutions problems 4.1gh.

$$\begin{cases} S \rightarrow aT \mid bT \\ T \rightarrow aU \mid bU \mid \lambda \\ U \rightarrow aS \mid bS \end{cases}$$



   $L(G)$ consists of all strings of length $3n + 1$ for $n \geq 0$.

5. Prove that the union of two CFLs is a CFL. (This is part (1) of Theorem 4.9 whose proof is on p.136 of the text. The proof is also in the lecture notes from 5-24-16 and 5-26-16.)

   Proof. We construct a CFG $G_u = (V_u, \Sigma, S_u, P_u)$ generating $L_1 \cup L_2$, as follows. S_u is a new variable not in either V_1 or V_2,

$$V_u = V_1 \cup V_2 \cup \{S_u\}$$

   and

$$P_u = P_1 \cup P_2 \cup \{S_u \rightarrow S_1 | S_2\}$$

   For every $x \in L_1 \cup L_2$, we can derive x in the grammar G_u by starting with either $S_u \rightarrow S_1$ or $S_u \rightarrow S_2$ and continuing with the derivation in either G_1 or G_2. On the other hand, if $S_u \Rightarrow^*_{G_u} x$, the first step in any derivation must be either $S_u \Rightarrow S_1$ or $S_u \Rightarrow S_2$, because those are the only productions with left side S_u. In the first case, the remaining steps must involve productions in G_1, because no variables in V_2 can appear, and so $x \in L_1$; similarly, in the second case $x \in L_2$. Therefore, $L(G_u) = L_1 \cup L_2$. $\square$

6. Let $\mathcal{F}$ denote the set of finite languages and $\mathcal{R}$ the set of regular languages over an alphabet Σ.

   a. Prove that every $L \in \mathcal{F}$ is a CFL. (Use the recursive definition of $\mathcal{F}$, structural induction and Theorem 4.9.)

   We denote by $\mathcal{F}$ the subset of $2^{\{a,b\}^*}$ (the set of languages over $\{a, b\}$) defined as follows:

   1. $\emptyset$, $\{\lambda\}$, $\{a\}$, and $\{b\}$ are elements of $\mathcal{F}$.
   2. For every L_1 and every L_2 in $\mathcal{F}$, $L_1 \cup L_2 \in \mathcal{F}$.
   3. For every L_1 and every L_2 in $\mathcal{F}$, $L_1L_2 \in \mathcal{F}$.

   Theorem 4.9: If L_1 and L_2 are context-free languages over an alphabet Σ, then $L_1 \cup L_2$, L_1L_2, and L_1^* are also CFLs.

   Proof. By structural induction

   I. $\emptyset$ is a CFL : $P = \emptyset$
   $\{\lambda\}$ is a CFL : $P = \{S \to \lambda\}$
   $\{\sigma\}$ is a CFL : $P = \{S \to \sigma\}$ (for any $\sigma \in \Sigma$)

   II. Let $L_1, L_2 \subseteq \Sigma^*$ and assume L_1, L_2 both CFLs.
   must show (i) $L_1 \cup L_2$ is a CFL and (ii) L_1L_2 is a CFL.
   (i) and (ii) follow directly from theorem 4.9.

   Therefore every $L \in \mathcal{F}$ is a CFL. □

   b. Prove that every $L \in \mathcal{R}$ is a CFL. (Use the recursive definition of $\mathcal{R}$, structural induction and Theorem 4.9. This is also problem 4.21 in the text.)

   If Σ is an alphabet, the set $\mathcal{R}$ of regular languages over Σ is defined as follows.

   1. The language $\emptyset$ is an element of $\mathcal{R}$, and for every $a \in \Sigma$, the language $\{a\}$ is in $\mathcal{R}$.
   2. For any two languages L_1 and L_2 in $\mathcal{R}$, the three languages $L_1 \cup L_2$, L_1L_2, and L_1^* are elements of $\mathcal{R}$.

   Proof. By structural induction

   I. $\emptyset$ is a CFL : $P = \emptyset$
   $\{\sigma\}$ is a CFL : $P = \{S \to \sigma\}$ (for any $\sigma \in \Sigma$)

   II. Let $L_1, L_2 \subseteq \Sigma^*$ and assume L_1, L_2 both CFLs.
   must show (i) $L_1 \cup L_2$ is a CFL, (ii) L_1L_2 is a CFL and (iii) L_1^* is a CFL.
   (i), (ii) and (iii) follow directly from theorem 4.9.

   Therefore every $L \in \mathcal{R}$ is a CFL. □

7. For each of the DFAs pictured in figure 2.44 (p.79 of the text), write down a regular grammar that generates the language accepted by the DFA.

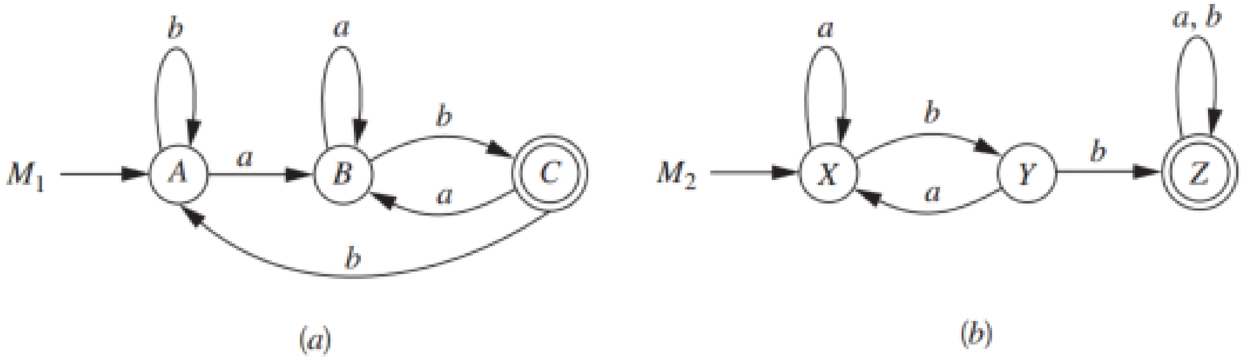


Figure 2.44 |

$(a+b)^*ab$

$A \to bA \mid aB$
$B \to aB \mid bC$
$C \to aB \mid bA \mid \lambda$

$(a+b)^*bb(a+b)^*$

$X \to aX \mid bY$
$Y \to aX \mid bZ$
$Z \to aZ \mid bZ \mid \lambda$

8. Draw a PDA that accepts the language of even length palindromes over $\{a, b\}$.

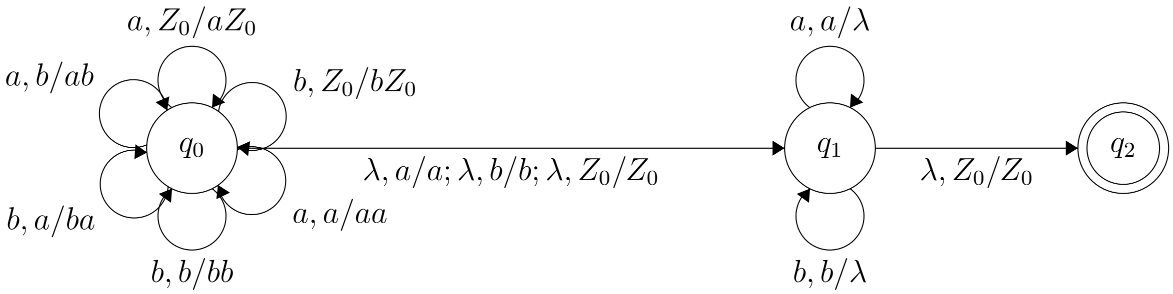


Table 1: Transition Table for a PDA Accepting *EvenPal*

Move Number	**State**	**Input**	**Stack Symbol**	**Move(s)**
1	q_0	a	Z_0	(q_0, aZ_0)
2	q_0	a	a	(q_0, aa)
3	q_0	a	b	(q_0, ab)
4	q_0	b	Z_0	(q_0, bZ_0)
5	q_0	b	a	(q_0, ba)
6	q_0	b	b	(q_0, bb)
7	q_0	λ	Z_0	(q_1, Z_0)
8	q_0	λ	a	(q_1, a)
9	q_0	λ	b	(q_1, b)
10	q_1	a	a	(q_1, λ)
11	q_1	b	b	(q_1, λ)
12	q_1	λ	Z_0	(q_2, Z_0)
	(all other combinations)			none

9. State the Pumping Lemma for CFLs.

   Suppose L is a context-free language. Then there is an integer n so that for every $u \in L$ with $|u| \geq n$, u can be written as $u = vwxyz$, for some strings v, w, x, y, and z satisfying

   1. $|wy| > 0$
   2. $|wxy| \leq n$
   3. for every $m \geq 0$, $vw^m xy^m z \in L$

10. Prove that the language $L = \{x \in \{a, b, c\}^* |\ n_a(x) = n_b(x) = n_c(x)\}$ is not context free.

    Proof. Assume $L = L(G)$. Let $u = a^n b^n c^n \in L$

    Condition 1 implies that the string wxy contains at least one symbol, and condition 2 implies that it can contain no more than two distinct symbols. If σ_1 is one of the three symbols that occurs in wy, and σ_2 is one of the three that doesn't, then the string vw^0xy^0z obtained from u by deleting w and y contains fewer than n occurrences of σ_1 and exactly n occurrences of σ_2. This is a contradiction, because condition 3 implies that vw^0xy^0z must have equal numbers of all three symbols. (We could also have obtained a contradiction by considering $m > 1$ in condition 3.)

    Therefore, L is not context free. $\square$

11. Problem 6.2a on p. 220 of the text. (This is the pumping for CFLs lemma again.)

    show using the pumping lemma that the given language is not a CFL.

    a. $L = \{a^i b^j c^k | i < j < k\}$

    Proof. Assume $L = L(G)$. Let $i = n, j = n + 1, k = n + 2$, then $u = a^n b^{n+1} c^{n+2} \in L$

    Condition 1 implies that the string wxy contains at least one symbol, and condition 2 implies that it can contain no more than two distinct symbols. In particular, wxy cannot contain both an a and a c.

Case 1. If a is not one of the symbols that occurs in wy, then the string vw^0xy^0z obtained from u by deleting w and y contains either $n_b \geq n_c$ so that $j \not< k$ or $n_a \geq n_b$ so that $i \not< j$.

Case 2. If c is not one of the symbols that occurs in wy, then the string vw^2xy^2z obtained from u by adding another w and y contains either $n_b \geq n_c$ so that $j \not< k$ or $n_a \geq n_b$ so that $i \not< j$.

This is a contradiction, because condition 3 implies that vw^mxy^mz for all $m \geq 0$ must have $i < j < k$ for all $a^i b^j c^k$. Therefore, L is not a CFL. $\square$