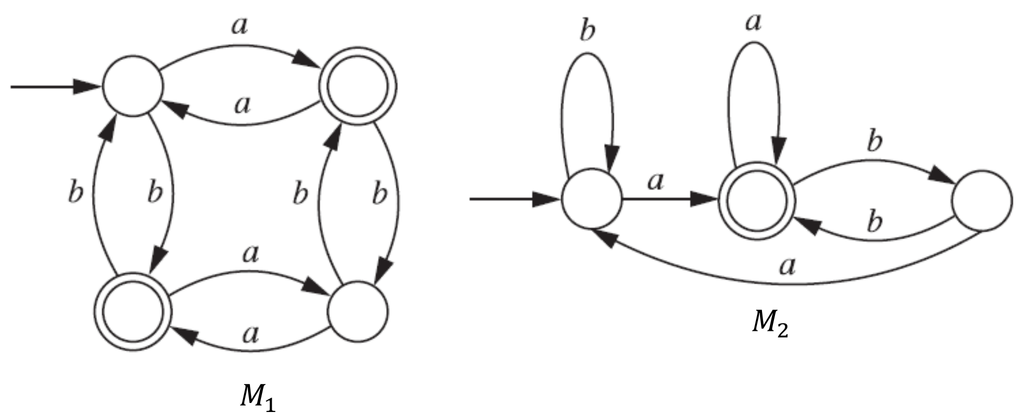
CMPS 130
Spring 2016
Homework Assignment 7 Solutions

Problems are from Martin 4th edition.

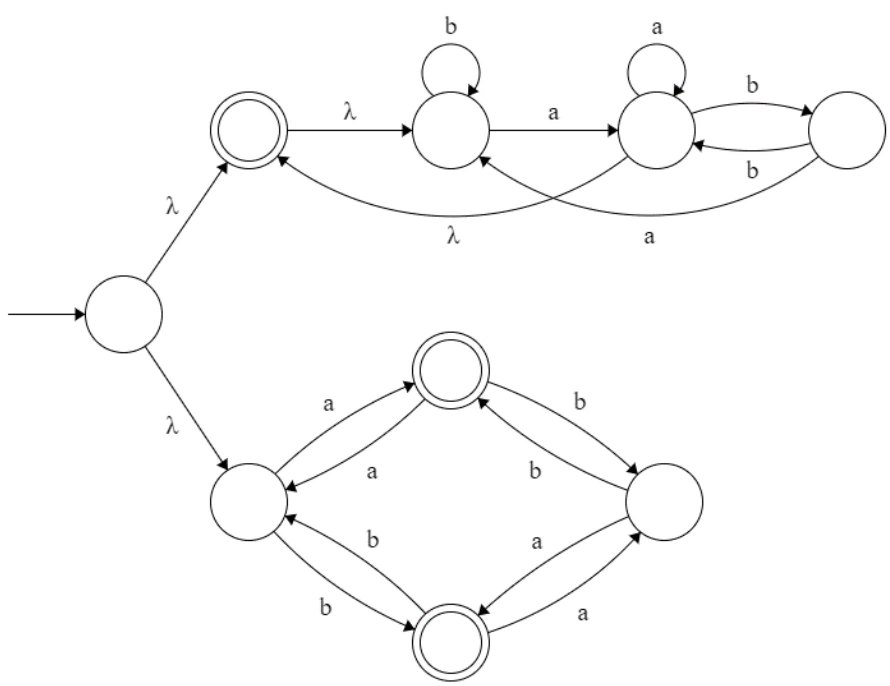
Chapter 3 (p.117): 49ab
Chapter 4 (p.154): 1bcefgh, 3abc, 4ab, 7, 10ac

1. Problem 3.49ab
   Figure 3.39 shows FAs M_1 and M_2 accepting languages L_1 and L_2 respectively. Draw NFAs accepting each of the following languages, using the constructions in the proof of Theorem 3.25.
   a. $L_2^* \cup L_1$
   b. $L_2 L_1^*$

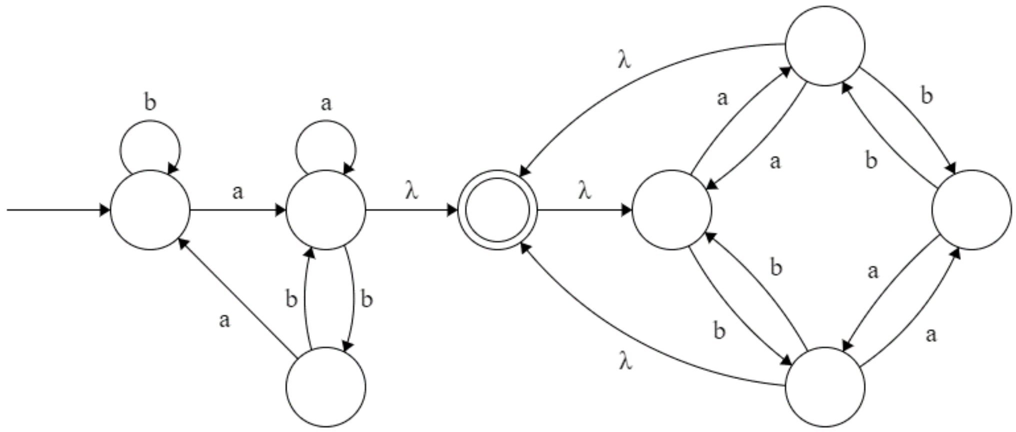


Solution:

a.



b.



2. Problem 4.1bcefgh
In each case below, say what language (a subset of $\{a, b\}^*$) is generated by the context-free grammar with the indicated productions.
b. $S \to SS \mid bS \mid a$
c. $S \to SaS \mid b$
e. $S \to TT$
$T \to aT \mid Ta \mid b$
f. $S \to aSa \mid bSb \mid aAb \mid bAa$
$A \to aAa \mid bAb \mid a \mid b \mid \lambda$
g. $S \to aT \mid bT \mid \lambda$
$T \to aS \mid bS$
h. $S \to aT \mid bT$
$T \to aS \mid bS \mid \lambda$

Solution: (Note to grader: answers only are sufficient for full credit on this problem.)

b. $L(G) = L((a + b)^*a)$, strings that end in a.
c. $L(G) = L(b(ab)^*) = L((ba)^*b)$, strings of alternating a's and b's that start and end with b.
e. $L(G)$ consists of all strings that contain exactly two b's.
f. $L(G)$ consists of all non-palindromes that can be changed into palindromes by changing a single character from a to b or b to a, i.e. strings that are one character away from being a palindrome.
g. $L(G)$ consists of all even length strings.
h. $L(G)$ consists of all odd length strings.

3. Problem 4.3abc
   In each case below, find a CFG generating the given language.
   a. The set of odd-length strings in $\{a, b\}^*$ with middle symbol a.
   b. The set of even-length strings in $\{a, b\}^*$ with the two middle symbols equal.
   c. The set of odd-length strings in $\{a, b\}^*$ whose first, middle, and last symbols are all the same.

   Solution: (Note to grader: answers only are sufficient for full credit on this problem.)

   a. $S \to aSa \mid aSb \mid bSa \mid bSb \mid a$.

   Another way: $\begin{cases} S \to XSX \mid a & \text{odd length strings with middle character } a \\ X \to a \mid b & \text{strings of length 1} \end{cases}$

   b. $S \to aSa \mid aSb \mid bSa \mid bSb \mid aa \mid bb$.

   Another way: $\begin{cases} S \to XSX \mid aa \mid bb & \\ X \to a \mid b & \text{strings of length 1} \end{cases}$

   c. $\begin{cases} S \to aAa \mid bBb & \\ A \to XAX \mid a & \text{odd length strings with middle character } a \\ B \to XBX \mid b & \text{odd length strings with middle character } b \\ X \to a \mid b & \text{strings of length 1} \end{cases}$

4. Problem 4.4ab
   In both parts below, the productions in a CFG G are given. In each part, show first that for every string $x \in L(G)$, $n_a(x) = n_b(x)$; then find a string $x \in \{a, b\}^*$ with $n_a(x) = n_b(x)$ that is not in $L(G)$.
   a. $S \to SabS \mid SbaS \mid \lambda$
   b. $S \to aSb \mid bSa \mid abS \mid baS \mid Sab \mid Sba \mid \lambda$

   Solution:
   As in earlier assignments, we use the book's notation for the language of strings having the same number of a's and b's: AeqB $= \{\, x \in \{a, b\}^* \mid n_a(x) = n_b(x)\}$. Also we write $S \Rightarrow_G^n x$ to mean that there is an n-step derivation of the string x in the grammar G, or just $S \Rightarrow^n x$ if the grammar is understood.

   a. <u>Claim</u>: if $x \in L(G)$, then $n_a(x) = n_b(x)$.
   Proof: If $x \in L(G)$, then $S \Rightarrow_G^* x$. We show by induction $\forall n \geq 1$: if $S \Rightarrow_G^n x$, then $x \in$ AeqB.
   I. If $n = 1$, then the derivation can only be $S \Rightarrow \lambda$ and indeed $\lambda \in$ AeqB.
   II. Let $n > 1$. Assume for any k in the range $1 \leq k < n$ that if $S \Rightarrow^k y$, then $y \in$ AeqB. We must show that if $S \Rightarrow^n x$, then $x \in$ AeqB. Suppose $x \in \{a, b\}^*$ can be derived in n steps. Since $n > 1$ the first step in the derivation must be either $S \Rightarrow SabS$ or $S \Rightarrow SbaS$. Let us consider the first case only, the other being entirely similar. Let y be the substring of x generated by the leftmost S in $SabS$, and let z be the substring of x generated by the rightmost S. Thus $x = yabz$, and there exist derivations $S \Rightarrow^{k_1} y$ and $S \Rightarrow^{k_2} z$ where both $k_1 < n$ and $k_2 < n$. By the induction hypothesis we have both $y \in$ AeqB and $z \in$ AeqB. It follows that

$$
\begin{aligned}
n_a(x) &= n_a(yabz) \\
&= n_a(y) + 1 + n_a(z) \\
&= n_b(y) + 1 + n_b(z) \\
&= n_b(yabz) \\
&= n_b(x)
\end{aligned}
$$

so $x \in$ AeqB as required.

Claim: The string $aabb \notin$ L(G), although $aabb \in$ AeqB.
Proof: Assume, to get a contradiction, that there exists a derivation $S \Rightarrow^* aabb$. Then the first step must have been $S \Rightarrow SabS$. (Obviously it did not start $S \Rightarrow \lambda$. If it started $S \Rightarrow SbaS$, then the final string would contain the substring *ba*, which $aabb$ does not.) The substring of $aabb$ generated by the left S in $SabS$ is therefore a, and the substring generated by the right S is b. Hence there are derivations $S \Rightarrow^* a$ and $S \Rightarrow^* b$, and so $a, b \in L(G)$. But this is impossible since $a, b \notin$ AeqB, and we showed above that $L(G) \subseteq$ AeqB. We have reached a contradiction, and therefore no such derivation $S \Rightarrow^* aabb$ exists. It follows that $aabb \notin L(G)$.

Note: the string *bbaa* would also serve as a counterexample.

b. Claim: if $x \in L(G)$, then $x \in$ AeqB.
Proof: We again show by induction that $\forall n \geq 1$: if $S \Rightarrow_G^n x$ then $x \in$ AeqB.

I. This grammar also has only one derivation of length $n = 1$, namely $S \Rightarrow \lambda$, and $\lambda \in$ AeqB.

II. Let $n > 1$ and assume for all k in the range $1 \leq k < n$ that if $S \Rightarrow^k y$, then $y \in$ AeqB. We must show that if $S \Rightarrow^n x$, then $x \in$ AeqB. Suppose $x \in \{a, b\}^*$ has an *n*-step derivation in G. There are six possible first steps in this derivation: $S \Rightarrow aSb$, $S \Rightarrow bSa$, $S \Rightarrow abS$, $S \Rightarrow baS$, $S \Rightarrow Sab$ and $S \Rightarrow Sba$. Let us suppose that the first step was $S \Rightarrow aSb$ (the other cases are entirely similar and are omitted.) Let the substring of *x* generated by *S* in *aSb* be denoted by y. Therefore $x = ayb$ and there exists a derivation $S \Rightarrow^k y$ where $k < n$. By the induction hypothesis we have $y \in$ AeqB, and hence $n_a(x) = 1 + n_a(y) = n_b(y) + 1 = n_b(x)$. Thus $x \in$ AeqB as required.

Claim: The string $aabbbbaa \notin$ L(G), although $aabbbbaa \in$ AeqB.
Proof: Consider all the possible first steps in a derivation $S \Rightarrow^* aabbbbaa$ in *G*.

$S \Rightarrow aSb$: excluded since such a string cannot begin and end with a.
$S \Rightarrow bSa$: excluded since such a string cannot begin and end with a.
$S \Rightarrow abS$: excluded since such a string cannot begin with *aa*.
$S \Rightarrow baS$: excluded since such a string cannot begin with *aa*.
$S \Rightarrow Sab$: excluded since such a string cannot end with *aa*.
$S \Rightarrow Sba$: excluded since such a string cannot end with *aa*.

Since no possible initial step can lead to the string *aabbbbaa*, no such derivation exists and hence $aabbbbaa \notin$ L(G).

5. Problem 4.7
   Describe the language generated by the CFG with productions

$$S \to ST \mid \lambda$$
$$T \to aS \mid bT \mid b$$

   Give an induction proof that your answer is correct.

   Solution: (Note to grader: answers only are sufficient for full credit on this problem.)

   $L(G) = \{a, b\}^*$, the set of all strings over $\{a, b\}$.
   Proof: Obviously $L(G) \subseteq \{a, b\}^*$. It remains to prove the opposite inclusion: $\supseteq$. We show by structural induction that if $x \in \{a, b\}^*$, then $x \in L(G)$.

   I. $\lambda \in L(G)$ because of the derivation $S \Rightarrow \lambda$.
   II. Let $x \in \{a, b\}^*$ and assume $x \in L(G)$, so there exists a derivation $S \Rightarrow^* x$ in G. We must show that there exist derivations $S \Rightarrow^* xa$ and $S \Rightarrow^* xb$ in G, and hence $xa, xb \in L(G)$. The last step in $S \Rightarrow^* x$ must be either to apply the production $S \to \lambda$ or the production $T \to b$. Note that all intermediate strings in a derivation for this grammar have all variables to the right of all terminals since the above productions possess that property.

   Case 1: The last step was to apply $S \to \lambda$. In this case we have $S \Rightarrow^* xS \Rightarrow x\lambda = x$. Alter the last step in the derivation in two ways to obtain:

$$S \Rightarrow^* xS \Rightarrow xST \Rightarrow xT \Rightarrow xaS \Rightarrow xa$$
and
$$S \Rightarrow^* xS \Rightarrow xST \Rightarrow xT \Rightarrow xb$$

   showing that $xa, xb \in L(G)$.

   Case 2: The last step was to apply $T \to b$. Here we have $S \Rightarrow^* yT \Rightarrow yb = x$ for some string y. Again alter the last step in two ways:

$$S \Rightarrow^* yT \Rightarrow ybT \Rightarrow ybaS \Rightarrow yba\lambda = xa$$
and
$$S \Rightarrow^* yT \Rightarrow ybT \Rightarrow ybb = xb$$

   showing that $xa, xb \in L(G)$.

   In either case $xa, xb \in L(G)$, and the proof is complete.

6. Problem 4.10ac
   Find context-free grammars generating each of the languages below.
   a. $\{\, a^i b^j \mid i \le j \,\}$
   c. $\{a^i b^j \mid j = 2i \}$

   Solution: (Note to grader: answers only are sufficient for full credit on this problem.)

   a. $S \to aSb \mid Sb \mid \lambda$
   c. $S \to aSbb \mid \lambda$