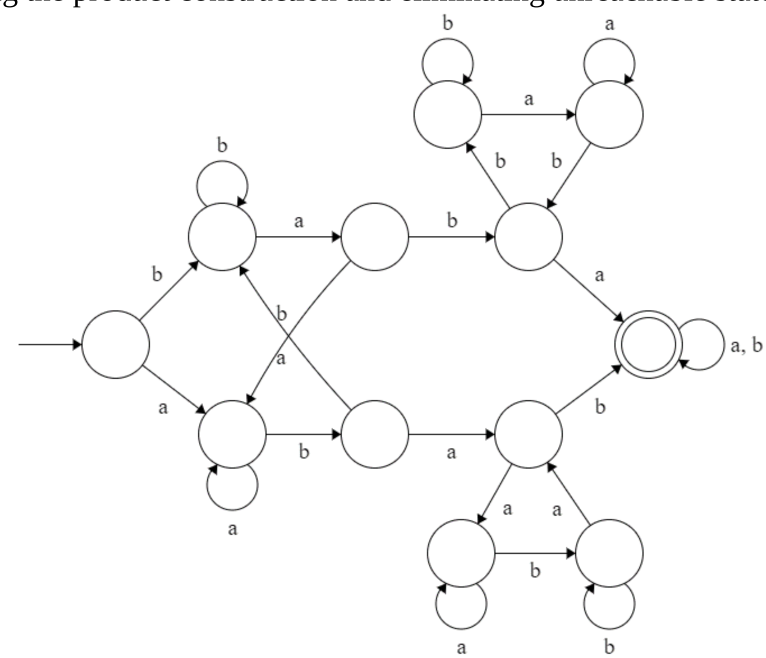
CMPS 130
Spring 2016
Homework Assignment 4 **Solutions**

Problems are from Martin 4th edition.

Chapter 2 (p.77): 1k, 14, 17ab, 21a, 26, 33, 35, 40abc

1. Problem 2.1k
   In each part below, draw an FA accepting the indicated language over $\{a, b\}$.
   k. The language of all strings containing both *aba* and *bab* as substrings.

   Solution:
   After using the product construction and eliminating unreachable states, we arrive at



2. Problem 2.14
   Let *z* be a fixed string of length *n* over the alphabet {a, b}. Using the argument in Example 2.5, we can find an FA with $n + 1$ states accepting the language of all strings in $\{a, b\}^*$ that end in *z*. The states correspond to the $n + 1$ distinct prefixes of *z*. Show that there can be no FA with fewer than $n + 1$ states accepting this language.

   Proof:
   We will display a set of $n + 1$ pairwise L-distinguishable strings in $\{a, b\}$. The result then follows from Theorem 2.21 (page 59). Let $S \subseteq \{a, b\}^*$ consist of the $n + 1$ distinct prefixes of *z*. We claim that this is the required set. Let $x, y \in S$ be two different prefixes of *z*. Then necessarily $|x| \neq |y|$, say for definiteness that $|x| < |y|$. We must show that *x* and *y* are *L*-distinguishable. To do this we must find a string *u* such that one of *xu* and *yu* is in *L*, and the other is not. Choose the string *u* so that $yu = z$. In other words, pick *u* to be the corresponding suffix to the prefix *y* of *z*. Then clearly $yu \in L$ since being equal to *z*, it certainly ends in *z*. On the other hand $|xu| = |x| + |u| < |y| + |u| = |yu| = |z|$, and therefore *xu*, being shorter than *z*, does not end in *z*. Thus $xu \notin L$, showing that *x* and *y* are *L*-distinguishable. Since *x* and *y* were arbitrary, any pair of distinct members of *S* are *L*-distinguishable, as required.

3. Problem 2.17ab
   Let L be the language $AnBn = \{ a^n b^n \mid n \geq 0 \}$.
   a. Find two distinct strings x and y in $\{a, b\}^*$ that are not L-distinguishable.
   b. Find an infinite set of pairwise L-distinguishable strings.

   Solution:
   a. The strings b and bb are L-indistinguishable. Indeed if z is any string over $\{a, b\}$ then neither bz nor bbz belong to $AnBn$, and hence $bI_L bb$ as claimed.
   b. We claim any pair of distinct strings in $\{ a^n \mid n \geq 0 \}$ are L-distinguishable. To see this suppose $n, m \in \mathbb{N}$ with $n \neq m$. Then b^n distinguishes a^n from a^m since $a^n b^n \in L$ while $a^m b^n \notin L$. The set $\{ a^n \mid n \geq 0 \}$ is clearly infinite, so we are done.

4. Problem 2.21a
   For each of the following languages L, show that the elements of the infinite set $\{a^n \mid n \geq 0\}$ are pairwise L-distinguishable.
   a. $L = \{ a^n b a^{2n} \mid n \geq 0 \}$

   Solution:
   Pick any pair of elements in the set $\{a^n \mid n \geq 0\}$, say a^{n_1} and a^{n_2} where $n_1 \neq n_2$. Then $z = ba^{2n_1}$ distinguishes between them since $a^{n_1}z = a^{n_1}ba^{2n_1} \in L$ while $a^{n_2}ba^{n_1} \notin L$.

5. Problem 2.26
   The pumping lemma says that if M accepts a language L, and if n is the number of states of M, then for every $x \in L$ satisfying $|x| \geq n$, there exist u, v and w with $x = uvw$ and such that (1) $|uv| \leq n$, (2) $|v| > 0$ (i.e. $v \neq \lambda$), and (3) $uv^i w \in L$ for all $i \geq 0$. Show that the statement provides no information if L is finite: If M accepts a finite language L, and n is the number of states of M, then L can contain no strings of length n or greater.

   Proof:
   Suppose M is an FA with n states accepting a finite language L. Assume, to get a contradiction, that there exists $x \in L$ with $|x| \geq n$. Then by the pumping lemma there exist u, v and w with $x = uvw$, and satisfying (1)-(3) above. In particular by (3) $uv^i w \in L$ for all $i \geq 0$. But the strings $uv^i w$ are all distinct. This follows since, by (2), they are all of different lengths $\left|uv^i w\right| = |uvw| + |v|^{i-1} = n + |v|^{i-1}$. Therefore L contains infinitely many strings, contrary to hypothesis. This contradiction shows that no such $x \in L$ exists.

6. Problem 2.33
   Let x be a string of length n in $\{a, b\}^*$, and let $L = \{x\}$. How many equivalence classes does I_L have? Describe them.

   Solution:
   L has exactly $n + 2$ equivalences classes. Let $u_0, u_1, u_2, \dots, u_n$ be the $n + 1$ prefixes of x. We claim that the equivalence classes of I_L are the singleton sets $\{u_i\}$ for $0 \leq i \leq n$, together with the set of non-prefixes of x: $\{a, b\}^* - \{u_0, u_1, u_2, \dots, u_n\}$. Let $v_0, v_1, v_2, \dots, v_n$ be the suffixes corresponding to the above prefixes, i.e. $u_i v_i = x$ for $0 \leq i \leq n$.

   - Each prefix u_i ($0 \leq i \leq n$) is L-distinguishable from every other string in $\{a, b\}^*$: Let y be any string other than u_i. Then $u_i v_i = x \in L$ but $yv_i \neq x$, and hence $yv_i \notin L$. Thus v_i distinguishes u_i from y with respect to L, and therefore $[u_i] = \{u_i\}$ for $0 \leq i \leq n$.

- Any pair y, z of non-prefixes of x are L-indistinguishable: Since y is a non-prefix, there is no $w \in \{a, b\}^*$ for which $yw = x$. Likewise there is no $w \in \{a, b\}^*$ for which $zw = x$. Therefore for all $w \in \{a, b\}^*$, both $yw \notin L$ and $zw \notin L$, hence yI_Lw. Thus $[\text{any non-prefix of } x] = \{\text{all non-prefixes of } x\}$.

7. Problem 2.35
Let $L \subseteq \Sigma^*$ be any language. Show that if $[\lambda]$ (the equivalence class of I_L containing λ) is not $\{\lambda\}$, then it is infinite.

Proof:
If $[\lambda] \neq \{\lambda\}$, then $[\lambda]$ contains some string other than λ. Say $x \in [\lambda]$ and $x \neq \lambda$. Therefore $xI_L\lambda$, which says that $\forall z$: $xz \in L \leftrightarrow \lambda z \in L$, i.e. $\forall z$: $xz \in L \leftrightarrow z \in L$. We use (weak) induction on n to show that for all $n \geq 0$: $x^n \in [\lambda]$. The result then follows since the strings $\{x^n \mid n \geq 0\} \subseteq [\lambda]$, being different lengths, are all distinct, and hence $[\lambda]$ is infinite.

I. The claim holds for $n = 0$ since $\lambda \in [\lambda]$. We can take as base case $n = 1$ since it was given that $x \in [\lambda]$. Note this is the same as $\forall z$: $xz \in L \leftrightarrow z \in L$.

II. Let $n \geq 1$ and assume $x^n \in [\lambda]$, which is equivalent to $\forall z$: $x^n z \in L \leftrightarrow z \in L$. We must show that $x^{n+1} \in [\lambda]$. This is equivalent to $\forall w$: $x^{n+1}w \in L \leftrightarrow w \in L$. Let $w \in \Sigma^*$ be chosen arbitrarily. Then

$$\begin{aligned} w \in L &\leftrightarrow x^n w \in L && \text{by the induction hypothesis with } z = w \\ &\leftrightarrow x(x^n w) \in L && \text{by the base case with } z = x^n w \\ &\leftrightarrow x^{n+1} w \in L && \text{by the associative law} \end{aligned}$$

Since w was arbitrary, $\forall w$: $x^{n+1}w \in L \leftrightarrow w \in L$, whence $x^{n+1} \in [\lambda]$, and the induction is complete.

8. Problem 2.40abc
Consider the language $L = AEqB = \{\, x \in \{a, b\}^* \mid n_a(x) = n_b(x) \,\}$.
a. Show that if $n_a(x) - n_b(x) = n_a(y) - n_b(y)$, then xI_Ly.
b. Show that if $n_a(x) - n_b(x) \neq n_a(y) - n_b(y)$, then x and y are L-distinguishable.
c. Describe all the equivalence classes of I_L.

Solution:
a. **Proof:** Suppose $n_a(x) - n_b(x) = n_a(y) - n_b(y)$. Let $z \in \{a, b\}^*$ be chosen arbitrarily. Then

$$\begin{aligned} xz \in L &\leftrightarrow n_a(xz) = n_b(xz) && \text{by definition of } L \\ &\leftrightarrow n_a(x) + n_a(z) = n_b(x) + n_b(z) \\ &\leftrightarrow n_a(x) - n_b(x) = n_b(z) - n_a(z) \\ &\leftrightarrow n_a(y) - n_b(y) = n_b(z) - n_a(z) && \text{by the given assumption} \\ &\leftrightarrow n_a(y) + n_a(z) = n_b(y) + n_b(z) \\ &\leftrightarrow n_a(yz) = n_b(yz) \\ &\leftrightarrow yz \in L && \text{by definition of } L \end{aligned}$$

Since z was arbitrary we have $\forall z$: $xz \in L \leftrightarrow yz \in L$, and hence xI_Ly.

b. **Proof:** Suppose $n_a(x) - n_b(x) \neq n_a(y) - n_b(y)$. We have two cases to consider.

Case 1: $n_a(x) \geq n_b(x)$. Let $k = n_a(x) - n_b(x)$ and $z = b^k$. Then $n_b(xz) = n_b(x) + k = n_a(x) = n_a(xz)$, so that $xz \in L$. On the other hand $n_b(yz) = n_b(y) + k \neq n_a(y) = n_a(yz)$, whence $yz \notin L$. Thus b^k distinguishes *x* from *y* in this case.

Case 2: $n_a(x) < n_b(x)$. Let $k = n_b(x) - n_a(x)$ and $z = a^k$. Then $n_a(xz) = n_a(x) + k = n_b(x) = n_b(xz)$, and again $xz \in L$. But $n_a(yz) = n_a(y) + k \neq n_b(y) = n_b(yz)$, so again $yz \notin L$. Therefore a^k distinguishes *x* from *y* in this case.

In both cases *x* and *y* are L-distinguishable, as claimed.

c. There is one I_L equivalence class for each natural number k, namely $\{y \mid n_a(y) - n_b(y) = k\}$. In other words, $[x] = \{y \mid n_a(y) - n_b(y) = n_a(x) - n_b(x)\}$. This follows directly from (a) and (b) above.