Automata and Theory of Computation

(Book: Formal Languages and Automata 5th Edition)

Hyunjun Kim

Exercises for Section 4.1: Closure Properties of Regular Languages

1. Fill in the details of the constructive proof of closure under intersection in Theorem 4.1.

   Proof. Let $L_1 = L(M_1)$ and $L_2 = L(M_2)$ such that $M_1 = (Q, \Sigma, \delta_0, q_0, F_1)$ and $M_2 = (P, \Sigma, \delta_1, p_0, F_2)$ are dfa's. We construct from M_1 and M_2 a combined automaton $\hat{M}$, whose state set $\hat{Q} = Q \times P$ consists of pairs (p_i, q_j), and whose transition function $\hat{\delta}$ is such that $\hat{M}$ is in state (p_i, q_j) whenever M_1 is in state q_i and M_2 is in state p_j. This is achieved by taking $\hat{\delta}((q_i, p_j), a) = (q_k, p_l)$ whenever

   $$\delta_1(q_i, a) = q_k$$

   and

   $$\delta_2(p_j, a) = p_l.$$

   $\hat{F}$ is defined as the set of all (q_i, p_j), such that $q_i \in F_1$, and $p_j \in F_2$. Then it shows $w \in L_1 \cap L_2$ if and only if it is accepted by $\hat{M}$. Consequently $L_1 \cap L_2$ is regular. Q.E.D.

2. Use the construction in Theorem 4.1 to find nfa's that accept

   (a) $L((a + b)a^*) \cap L(baa^*)$,

   (b) $L(ab^*a^*) \cap L(a^*b^*a)$.

   Solution. (a) A dfa for $L((a + b)a^*)$ is given by

   $$\delta(q_0, a) = q_1, \delta(q_0, b) = q_1, \delta(q_1, a) = q_1, \delta(q_1, b) = q_t,$$

   with q_t a trap state and final state q_1. A dfa for $L(baa^*)$ is given by

   $$\delta(p_0, a) = p_t, \delta(p_0, b) = p_1, \delta(p_1, a) = p_2, \delta(p_1, b) = p_t, \delta(p_2, a) = p_2, \delta(p_2, b) = p_t,$$

   with p_t a trap state and final state p_2. From this we find

   $$\delta((q_0, p_0), a) = (q_1, p_t), \delta((q_0, p_0), b) = (q_1, p_1), \delta((q_1, p_1), a) = (q_1, p_2),$$

   etc. When we complete this construction, we see that the only final state is (q_1, p_2) and that $L((a + b)a^*) \cap L(baa^*) = L(baa^*)$.

   (b) A dfa for $L(ab^*a^*)$ is given by

   $$\delta(q_0, a) = q_1, \delta(q_0, b) = q_t,$$
   $$\delta(q_1, a) = q_2, \delta(q_1, b) = q_1,$$
   $$\delta(q_2, a) = q_2, \delta(q_2, b) = q_t,$$

   with q_t a trap state and final states q_1 and q_2. Likewise, a dfa for $L(a^*b^*a)$ is given by

   $$\delta(p_0, a) = p_1, \delta(p_0, b) = p_3,$$
   $$\delta(p_1, a) = p_2, \delta(p_1, b) = p_3,$$
   $$\delta(p_2, a) = p_2, \delta(p_2, b) = p_3,$$
   $$\delta(p_3, a) = p_4, \delta(p_3, b) = p_3,$$
   $$\delta(p_4, a) = p_t, \delta(p_4, b) = p_t,$$

with p_t a trap state and the three final states p_1, p_2, and p_4. From this we construct new nfa below

$$\delta((q_0, p_0), a) = (q_1, p_1), \delta((q_0, p_0), b) = (q_t, p_3),$$
$$\delta((q_1, p_1), a) = (q_2, p_2), \delta((q_1, p_1), b) = (q_1, p_3),$$
$$\delta((q_1, p_3), a) = (q_2, p_4), \delta((q_1, p_3), b) = (q_1, p_3),$$
$$\delta((q_2, p_2), a) = (q_2, p_2), \delta((q_2, p_2), b) = (q_t, p_3),$$
$$\delta((q_2, p_4), a) = (q_2, p_t), \delta((q_2, p_4), b) = (q_t, p_t).$$

From this construction,we see that the final states are (q_1, p_1), (q_2, p_2), and (q_2, p_4). And that $L(ab^*a^*) \cap L(a^*b^*a) = L(a(a^* + b^*a))$.

Example 4.1. Show that the family of regular languages is closed under difference.

In other words, we want to show that if L_1 and L_2 are regular, then $L_1 - L_2$ is necessarily regular also.

The needed set identity is immediately obvious from the definition of a set difference, namely

$$L_1 - L_2 = L_1 \cap \overline{L_2}.$$

The fact that L_2 is regular implies that $\overline{L_2}$ is also regular. Then, because of the closure of regular languages under intersection, we know that $L_1 \cap \overline{L_2}$ is regular, and the argument is complete.

3. In Example 4.1 we showed closure under difference for regular languages, but the proof was nonconstructive. Provide a constructive argument for this result, following the approach used in the argument for intersection in Theorem 4.1.

Proof. Let $L_1 = L(M_1) and L_2 = L(M_2)$ such that $M_1 = (Q, \Sigma, \delta_0, q_0, F_1)$ and $M_2 = (P, \Sigma, \delta_1, p_0, F_2)$ are dfa's. We construct from M_1 and M_2 a combined automaton $\hat{M}$, whose state set $\hat{Q} = Q \times P$ consists of pairs (p_i, q_j) and whose transition function $\hat{\delta}$ is such that $\hat{M}$ is in state (p_i, q_j) whenever M_1 is in state q_i and M_2 is in state p_j. This is achieved by taking $\hat{\delta}((q_i, p_j), a) = (q_k, p_l)$ whenever

$$\delta_1(q_i, a) = q_k$$

and

$$\delta_2(p_j, a) = p_l.$$

$\hat{F}$ is defined as set of all (q_i, p_j), such that $q_i \in F_1$, and $p_j \notin F_2$. Then it shows $w \in L_1 - L_2$ if and only if it is accepted by $\hat{M}$. Consequently $L1 - L2$ is regular. Q.E.D.

Theorem 4.3. Let h be a homomorphism. If L is a regular language, then its homomorphic image $h(L)$ is also regular. The family of regular languages is therefore closed under arbitrary homomorphisms.

Proof. Let L be a regular language denoted by some regular expression r. We find $h(r)$ by substituting $h(a)$ for each symbol $a \in \Sigma$ of r. It can be shown directly by an appeal to the definition of a regular expression that the result is a regular expression. It is equally easy to see that the resulting expression denotes $h(L)$. All we need to do is to show that for every $w \in L(r)$, the corresponding $h(w)$ is in $L(h(r))$ and conversely that for every v in $L(h(r))$ there is a w in L, such that $v = h(w)$. Leaving details as an exercise, we claim that $h(L)$ is regular. Q.E.D.

4. In the proof of Theorem 4.3, show that $h(r)$ is a regular expression. Then show that $h(r)$ denotes $h(L)$.

Proof. A regular expression for $h(r)$ can be obtained by simply applying $h(a)$ for each symbol $a \in \Sigma$ of r. Now, we have to prove that the set of strings accepted by $h(r)$ is a subset of those accepted by $h(L)$, and the set of strings accepted by $h(L)$ is a subset of those accepted by $h(r)$, thus proving they are equivalent.

$h(r) \subseteq h(L)$: All strings accepted by $h(r)$ must first follow the original rules of r, but each symbol a in Σ of r has been substituted with $h(a)$ and thus each string accepted by $h(r)$ must be accepted by $h(L)$.

$h(L) \subseteq h(r)$: All strings accepted by $h(L)$ must first follow the regular expression r, but rather than use the old symbols, will have them put through h first. Since this is the definition of $h(r)$, it is easy to see. Q.E.D.

5. Show that the family of regular languages is closed under finite union and intersection, that is, if L_1, L_2, ..., L_n are regular, then

$$L_U = \bigcup_{i=\{1,2,...,n\}} L_i$$

and

$$L_I = \bigcap_{i=\{1,2,...,n\}} L_i$$

are also regular.

Proof. Induction on n to show that if L_1, L_2, ..., L_n are regular, $L_U = \cup_{i=\{1,2,...n\}} L_i$ is also regular.

<u>Base case</u>: If $n = 1$, the expanded right side is L_1 and it is regular. So the argument holds when $n = 1$, now we have a base case.

<u>Inductive step</u>: Let $k \geq 1$ be an arbitrary natural number.
Let assume the induction hypothesis: the argument holds for all values of n up to some k.
Then, with $n = n + 1$, $L_U = \bigcup_{i=\{1,2,..,n+1\}} L_i = \bigcup_{i=\{1,2,..,n\}} L_i \cup L_{n+1}$. From the latest formula, the left operand of $\cup$ is regular by inductive hypothesis, and the right operand is also regular immediately by the argument. Since the family of regular languages are closed under union, now we see that the argument holds for $n = n + 1$, consequently holds for any arbitrary n, $n \geq 1$. Q.E.D.

Showing that the family of regular languages is closed under finite intersection will taken easy as above.

6. The *symmetric difference* of two sets S_1 and S_2 is defined as

$$S_1 \ominus S_2 = \{x : x \in S_1 \text{ or } x \in S_2, \text{ but } x \text{ is not in both } S_1 \text{ and } S_2\}.$$

Show that the family of regular languages is closed under symmetric difference.

Proof.

$$\begin{aligned} S_1 \ominus S_2 &= (S_1 \cup S_2) - (S_1 \cap S_2) \\ &= (S_1 \cup S_2) \cap \overline{(S_1 \cap S_2)} \end{aligned}$$

We know the family of regular languages is closed under union, intersection, and complementation. Now we see that it is also closed under symmetric difference. Q.E.D.