Formal Languages and Automata Theory

Test 1

Maximum marks: 40 | Time: 22-02-2022, 10:15am | Duration: 1 hour 15 minutes

---

1. Consider the following language over the alphabet $\Sigma = \{a, b\}$.

$$L = \left\{ uvwu^r \;\middle|\; u, v, w \in \Sigma^+,\ |v| = |w| + 1 \right\},$$

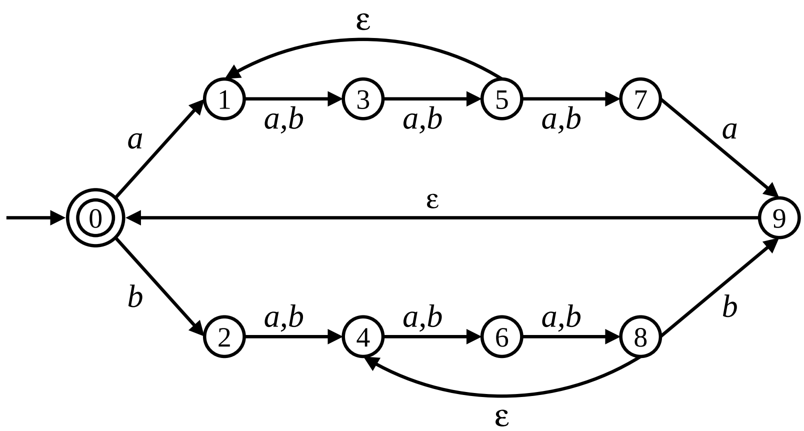
where u^r stands for the reverse of the string u, and $|x|$ is the length of the string x.

(a) Write a regular expression for L. **(4)**

Solution $a\Big((a+b)(a+b)(a+b)\big((a+b)(a+b)\big)^*\Big)a \,+\, b\Big((a+b)(a+b)(a+b)\big((a+b)(a+b)\big)^*\Big)b.$

(b) Argue that the language of the ε-NFA of Figure 1 is L^*. **(4)**

Figure 1: ε-NFA for Q1



Solution Let us call this ε-NFA N. If N can reach State 9, it accepts by making an ε-transition to State 0. Let us remove this ε-transition, and make 9 as the final state and 0 a non-final state. Easy inspection shows that both the sets $\{x \mid 7 \in \hat{\Delta}(1, x)\}$ and $\{x \mid 8 \in \hat{\Delta}(2, x)\}$ are equal to the set of all odd-length strings over $\{a, b\}$ of length at least 3. Therefore, the modified NFA accepts all the odd-length strings of length $\geqslant 5$ and starting and ending with the same symbol. These are precisely all the strings in L. The given NFA accepts ε (the only member of L^0). Now, let $w = w_1 w_2 \ldots w_n \in L^*$ for some $n \geqslant 1$ and with each $w_i \in L$. Under the inductive hypothesis that N accepts $w_1 w_2 \ldots w_{n-1}$, we have $0 \in \hat{\Delta}(0, w_1 w_2 \ldots w_{n-1})$. But then, $\hat{\Delta}(0, w_1 w_2 \ldots w_n)$ contains 9, and so 0 by the ε-transition from 9 to 0.

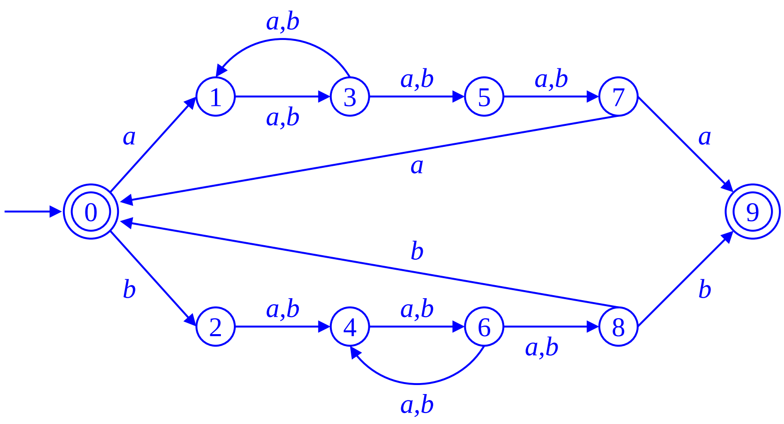
(c) If Δ denotes the transition function of the ε-NFA of Figure 1, find $\hat{\Delta}(0, aabbaba)$. Show all the states that the ε-NFA can be in, after reading each symbol from the given input string. **(4)**

Solution We have the following transitions:

$$\begin{aligned}
\hat{\Delta}(0, \varepsilon) &= \varepsilon\text{-closure}(\{0\}) &&= \{0\},\\
\hat{\Delta}(0, a) &= \varepsilon\text{-closure}(\{1\}) &&= \{1\},\\
\hat{\Delta}(0, aa) &= \varepsilon\text{-closure}(\{3\}) &&= \{3\},\\
\hat{\Delta}(0, aab) &= \varepsilon\text{-closure}(\{5\}) &&= \{1, 5\},\\
\hat{\Delta}(0, aabb) &= \varepsilon\text{-closure}(\{3, 7\}) &&= \{3, 7\},\\
\hat{\Delta}(0, aabba) &= \varepsilon\text{-closure}(\{5, 9\}) &&= \{0, 1, 5, 9\},\\
\hat{\Delta}(0, aabbab) &= \varepsilon\text{-closure}(\{2, 3, 7\}) &&= \{2, 3, 7\},\\
\hat{\Delta}(0, aabbaba) &= \varepsilon\text{-closure}(\{4, 5, 9\}) &&= \{0, 1, 4, 5, 9\}.
\end{aligned}$$

(d) Convert the ε-NFA of Figure 1 to an equivalent NFA without ε-transitions but with the same set of states. **(4)**

Figure 2: NFA for Q1(d)



Solution The equivalent NFA is given in Figure 2.

2. Let L be a language over an alphabet Σ. Recall that a string x is called a prefix of a string y if $y = xz$ for some string z. For example, all the prefixes of $abbab$ are $\varepsilon, a, ab, abb, abba, abbab$. From L, we generate the language $\text{dupPrefix}(L)$ by duplicating prefixes of strings in L. More precisely, we define

$$\text{dupPrefix}(L) = \left\{ xy \;\middle|\; y \in L, \text{ and } x \text{ is a prefix of } y \right\}.$$

(a) Prove/Disprove: If L is regular, then $\text{dupPrefix}(L)$ must also be regular. **(5)**

Solution *False.* Take $\Sigma = \{a, b\}$, and $L = \mathscr{L}(a^*b) = \{a^n b \mid n \geqslant 0\}$. Suppose that $\text{dupPrefix}(L)$ is regular. Let k be a pumping-lemma constant for $\text{dupPrefix}(L)$. Supply the string $a^k b a^k b \in \text{dupPrefix}(L)$ to the pumping lemma with $u = a^k b$, $v = a^k$, and $w = b$. The lemma returns a decomposition $v = xyz$ with $y = a^l$ for some $l > 0$. Pumping out y, we get the string $uxzw = a^k b a^{k-l} b \in L$. But since $l > 0$, $a^k b$ cannot be a prefix of $a^{k-l} b$, a contradiction.

(b) Prove/Disprove: If L is not regular, then $\text{dupPrefix}(L)$ must also be non-regular. **(5)**

Solution *False:* Take $\Sigma = \{a\}$, and $L = \{a^{n^2} \mid n \geqslant 0\}$ (a language already proved as not regular). We have

$$\text{dupPrefix}(L) = \{a^m \mid m \geqslant 0,\ \ m \neq 3\}$$

which is regular (because its complement is a finite set).

3. In this exercise, we use the Myhill–Nerode theorem to prove that the intersection L of two regular languages L_1 and L_2 (over the same alphabet Σ) is again regular. Let $\equiv_1$ and $\equiv_2$ be Myhill–Nerode (MN) relations for L_1 and L_2, respectively. The equivalence classes $C_1, C_2, \ldots, C_k$ of $\equiv_1$ partition Σ^*. Likewise, the equivalence classes $D_1, D_2, \ldots, D_l$ of $\equiv_2$ partition Σ^*. Define the subsets $E_{ij} = C_i \cap D_j$ of Σ^* for $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, l$. We consider only those E_{ij} that are non-empty. For any fixed i, the non-empty subsets E_{ij} partition C_i. Therefore all the non-empty subsets E_{ij} partition Σ^*.

(a) Let $L = L_1 \cap L_2$. The partition of Σ^* by non-empty E_{ij} induces an equivalence relation $\equiv$ on Σ^*. Prove that $\equiv$ is an MN relation for L. By the Myhill–Nerode theorem, L is therefore regular. **(6)**

Solution [Right congruence of $\equiv$] Let $x \equiv y$, and $a \in \Sigma$. Then, x and y belong to the same part E_{ij} for some i, j. We have $E_{ij} = C_i \cap D_j$, that is, x and y belong to both C_i and D_j, that is, $x \equiv_1 y$ and $x \equiv_2 y$. Since $\equiv_1$ and $\equiv_2$ are MN relations, $xa \equiv_1 ya$ and $xa \equiv_2 ya$, that is, there exist i', j' such that both $xa, ya \in C_{i'}$, and both $xa, ya \in D_{j'}$. But then, both xa and ya are in $E_{i'j'}$, that is, $xa \equiv ya$.

[$\equiv$ refines L] Let $x \equiv y$. Then, both of them belong to some E_{ij} and therefore to both C_i and D_j, that is, $x \equiv_1 y$ and $x \equiv_2 y$. Suppose that $x \in L$, that is, $x \in L_1$ and $x \in L_2$. Since $\equiv_1$ and $\equiv_2$ are MN relations, we have $y \in L_1$ and $y \in L_2$, that is, $y \in L$. Analogously, we can prove that if $y \in L$, then $x \in L$ too.

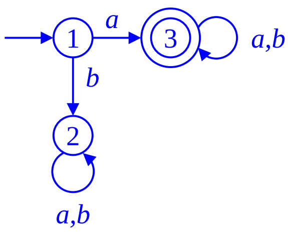
[$\equiv$ has finite index] The maximum possible index of $\equiv$ is kl.

(b) Take $\Sigma = \{a,b\}$. Let $L_1 = \mathscr{L}\big(a(a+b)^*\big)$ and $L_2 = \mathscr{L}\big((a+b)b(a+b)^*\big)$. Then, $L = L_1 \cap L_2 = \mathscr{L}\big(ab(a+b)^*\big)$. Construct the minimal DFA for L_1 and L_2, and deduce the partitions induced by the corresponding coarsest MN relations $\equiv_1$ and $\equiv_2$. Construct the MN relation $\equiv$ for L as described above, and generate an equivalent DFA M from that relation. Prove/Disprove: M is the minimal DFA for L. **(4 + 3 + 1)**

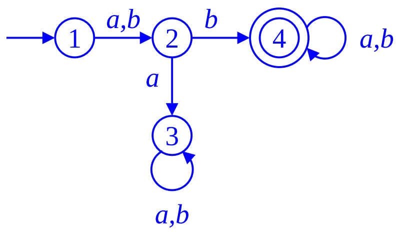
Solution The minimal DFA for L_1 and L_2 are given in Parts (a) and (b) of Figure 3. The partitions induced by these DFA are as follows.

$$C_1 = \{\varepsilon\} \qquad D_1 = \{\varepsilon\}$$
$$C_2 = \mathscr{L}(b(a+b)^*) \qquad D_2 = \{a,b\}$$
$$C_3 = \mathscr{L}(a(a+b)^*) \qquad D_3 = \mathscr{L}((a+b)a(a+b)^*)$$
$$D_4 = \mathscr{L}((a+b)b(a+b)^*)$$

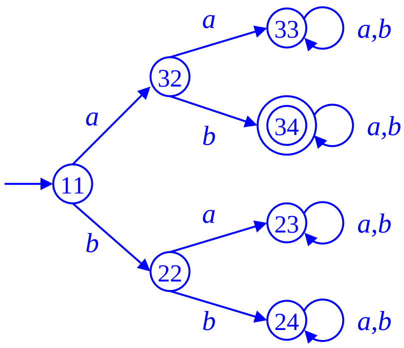
Figure 3: The constructions for Q3(b)



(a) Minimal DFA for L1



(b) Minimal DFA for L2



(c) DFA obtained by the given construction

The non-empty intersections E_{ij} are given below.

$$E_{11} = \{\varepsilon\} \qquad E_{22} = \{b\} \qquad E_{32} = \{a\}$$
$$E_{23} = \mathscr{L}(ba(a+b)^*) \qquad E_{33} = \mathscr{L}(aa(a+b)^*)$$
$$E_{24} = \mathscr{L}(bb(a+b)^*) \qquad E_{34} = \mathscr{L}(ab(a+b)^*)$$

The DFA equivalent to this partition is given in Part (c) of Figure 3. This DFA is not minimal, because the states 33, 22, 23, and 24 are equivalent.