COMP 335, A Solution to Assignment 2

Definition. In questions below $n_a(w)$ and $n_b(w)$ indicate the number of a's and the number of b's in string w.

Problem 1. *Find a regular expression for the following languages over* $\Sigma = \{a, b\}^*$.

(a) $\{w : w \textit{ does not contain substring } bbb\}$

(b) $\{w : w \textit{ contains exactly three } a\textit{'s and it ends with } abb\}$

(c) $\{w : n_a(w) \mod 3 = 0 \textit{ and } w \textit{ begins with } ab\}$

(d) $\{w : (n_a(w) + n_b(w)) \mod 3 \geq 2\}$

(e) $\{w : |w| \mod 2 = 0\}$

(f) $\{a^m b^n : mn > 4\}$

Answer:

(a) $(a + ba + bba)^*(\lambda + b + bb)$

(b) $b^*ab^*ab^*abb$

(c) $abb^*ab^*ab^*(b^*ab^*ab^*ab^*)^*$

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

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

(f) $aa^*bbbbbb^* + aaa^*bbbb^* + aaaa^*bbb^* + aaaaa^*bbb^* + aaaaaa^*bb^*$

■

Problem 2. *Find an NFA that accepts the language defined by the following regular expression*

$$ab(a + ab)^*(a + aa^*) + (abab)^* + (aaa^* + b)^*$$

Answer:
The answer is depicted in Figure 1. ■

Problem 3. *Find a regular expression for the following regular languages*

(a) $\{w : 2n_a(w) + 3n_b(w) \textit{ is even }\}$

(b) $\{w : (n_a(w) - n_b(w)) \mod 3 = 1\}$

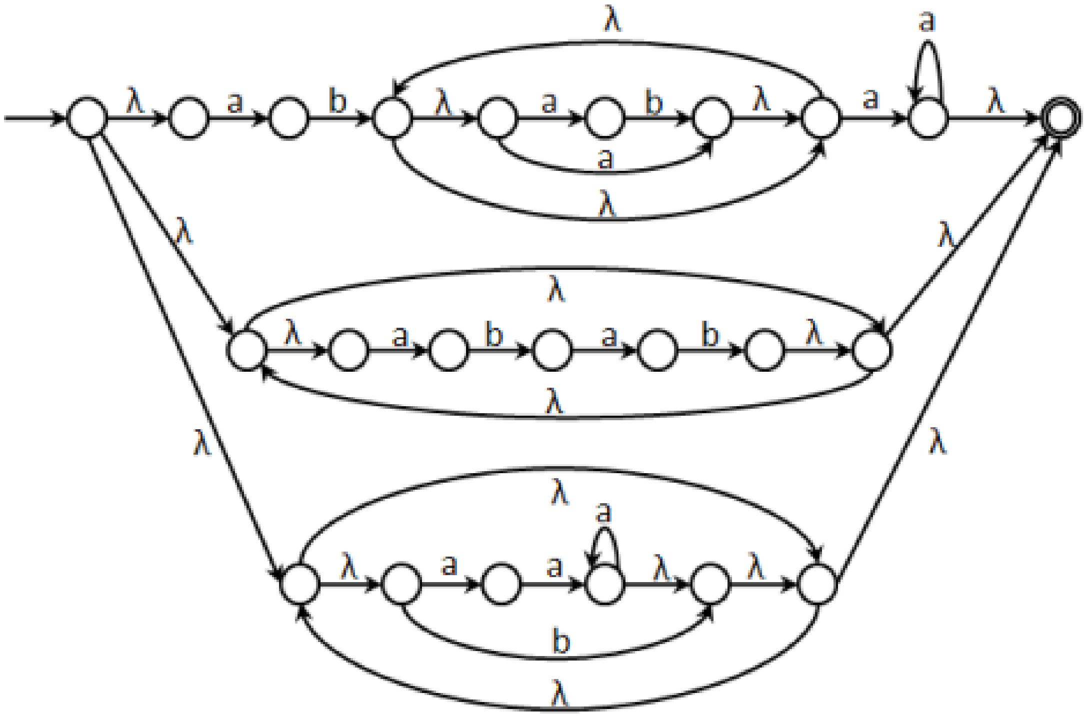

Figure 1: Answer to question 2.

Answer:

(a) $a^*(ba^*ba^*)^*$

(b) $(ba)^*(a+bb)(ab+(b+aa)(ba)^*(a+bb))^*$

Solution 1:

A DFA for this language is depicted in Figure 2, and then reduced to an initial state and a final state by eliminating state 2 by adding transitions that would have otherwise use state 2 as the middle state. The answer then is concluded from this machine.

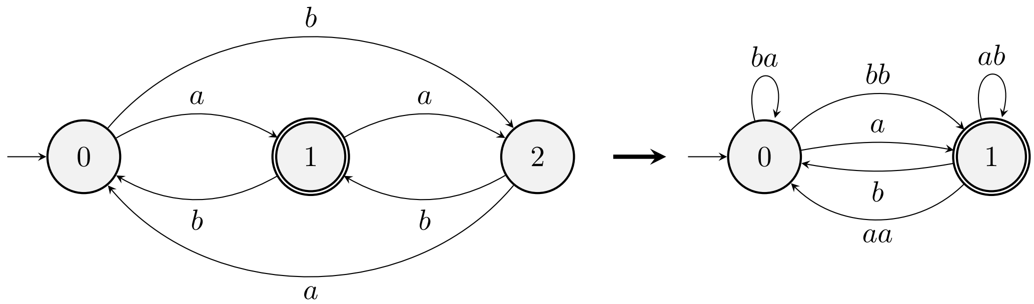

Figure 2: Question 3, part b: To the left, the original DFA. To the right, the reduced version.

Solution 2:

Scan a valid string w from left to right and consider $(n_a(w)-n_b(w)) \mod 3$ after each character, starting with $(n_a(w)-n_b(w)) \mod 3 = 0$. Let us refer to $(n_a(w)-n_b(w)) \mod 3$ as f.

When $f=1$ is first seen (guaranteed to see it at least once by validity of w), we have scanned a string of the form $(ba)^*(a+bb)$. Upon existence, consider the next character: If it is b, $f=0$ is reached, just like the beginning of our scan. Else if it is a, look at the next character (should exist, (why?)) ab will take us back to $f = 1$, just like after the first occurrence of $f = 1$, and aa will take us to $f = 0$, just like the beginning. Therefore, upon existence of next character after first occurrence of $f = 1$ until next time that $f = 1$ happens we have scanned a string of the form $(ab + (b + aa)(ba)^*(a + bb))$. This concludes the answer.

■

Problem 4. *Find a regular grammar for the following languages*

(a) $L(aa^*(ab + a)^*)$

(b) $L((aab^*ab)^*)$

(c) $L = \{w : (n_a(w) - n_b(w)) \mod 3 = 1\}$

Answer:

(a) $S \to aA$
$A \to aA \mid aB \mid \lambda$
$B \to bA$

(b) $S \to aA \mid \lambda$
$A \to aB$
$B \to bB \mid aC$
$C \to bS$

(c) $S \to bA \mid aB$
$A \to aS \mid bB$
$B \to bS \mid aA \mid \lambda$

■

Problem 5. *Prove that the following langauges are not regular*

(a) $L = \{ww : w \in \{a, b\}^*\}$

(b) $L = \{a^n b^k : n \geq k\}$

(c) $L = \{a^n b^m c^k : n + m > k > 0\}$

(d) $L = \{w \in \{a, b\}^* : 2n_a(w) = 3n_b(w)\}$

Answer:

(a) Note that the language is infinite. Let us assume that L is a regular language. Let m be the integer from the pumping lemma. Consider the string $w = a^m b^m a^m b^m$. Clearly, $|w| \geq m$ and w belongs to L. According to the pumping lemma w can be represented as $w = uyv$ where $|uy| \leq m$, and $|y| \geq 1$. From the form of string w and the $|uy| \leq m$ condition follows that y consists of only a's; therefore, $y = a^k$ for some $k \geq 1$. By the pumping lemma, the string $uy^2v = a^{m+k} b^m a^m b^m$ is in L, which is in contradiction to the definition of L. Therefore, L is not a regular language.

(b) Note that the language is infinite. Let us assume that L is a regular language. Let m be the integer from the pumping lemma. For getting a contradiction let us consider the string $w = a^m b^m$. Clearly, $|w| \geq m$ and w belongs to L. By the pumping lemma, w can be represented as $w = uyv$, where $|uy| \leq m$, and $|y| \geq 1$. From the form of string w and the $|uy| \leq m$ condition follows that y consists only of a's; therefore, $y = a^k$ for some $k \geq 1$. Now let us look at the string $w_0 = uy^0v = uv = a^{m-k}b^m$. By the pumping lemma w_0 belongs to L, but this is impossible because $m - k < m$. This contradiction proves that our initial assumption about regularity of L was incorrect and L is not a regular language.

(c) Again we prove by contradiction. Assume that L is a regular language. Then from the properties of regular languages we know that if L is regular, then its reversal $L^R = \{c^k b^m a^n : 0 < k < m + n\}$ is also regular. We will prove that L^R is not regular, and so contradiction. Thus, our assumption, that L is regular, was not true, and so L is not regular.

Now we will prove that L^R is not regular. Since L^R is infinite then pumping lemma applies. Let m be the integer from the pumping lemma for language L^R. For getting a contradiction let us pick $w = c^m a^m b^m$, which satisfies both $|w| \geq m$ and $w \in L^R$. By the pumping lemma, w can be represented as $w = uyv$ where $|uy| \leq m$, and $|y| \geq 1$. From the form of string w and the $|uy| \leq m$ condition follows that y consists only of c's; therefore, $y = c^k$ for some $k \geq 1$. Now consider the string $w_{2m+1} = uy^{2m+1}v$. By the pumping lemma, w_{2m+1} belongs to L^R; however, w_{2m+1} does not satisfy the definition of L^R since the number of leading c's is greater than the number of b's and a's that follows. This contradiction proves that our initial assumption about regularity of L^R was incorrect; consequently, L is not a regular language.

(d) Note that the language is infinite. Let us assume that L is a regular language. Let m be the integer from the pumping lemma. Consider the string $w = a^{3m}b^{2m}$, which satisfies both conditions $|w| \geq m$ and $w \in L$. By the pumping lemma, w can be represented as $w = uyv$ where $|uy| \leq m$, and $|y| \geq 1$. From the form of string w and the $|uy| \leq m$ condition follows that y consists only of a's; therefore, $y = a^k$ for some $k \geq 1$. Now let us consider the string $w_0 = uy^0v = uv = a^{3m-k}b^{2m}$. By the pumping lemma, w_0 belongs to L, but this is not possible because $3 \times 2m > 2(3m - k)$. Therefore, L is not a regular language.

■

Problem 6. *Find context free grammars for the following languages*

(a) $L = \{a^n b^m : n, m \geq 0 \textit{ and } n \neq m\}$

(b) $L = \{a^n b^m c^k : n \geq m \geq 0 \textit{ or } m \geq k \geq 0\}$

(c) $L = \{a^n b^m c^k : n + m > k > 0\}$

(d) $L = \{w \in \{a, b, c\}^* : n_b(w) = n_a(w) + n_c(w)\}$

Answer:

(a) $S \to AC \mid CB$
$C \to aCb \mid \lambda$
$A \to aA \mid a$

$B \to Bb \mid b$

(b) $S \to X \mid Y$
$X \to AC$
$A \to aAb \mid aA \mid \lambda$
$C \to Cc \mid \lambda$
$Y \to BD$
$B \to aB \mid \lambda$
$D \to bDc \mid bD \mid \lambda$

(c) $S \to aAc \mid bBc$
$A \to aAc \mid B \mid C$
$B \to bBc \mid D$
$C \to aC \mid D \mid a$
$D \to Db \mid b$

The set of strings each variable represents:

A: $\{a^n b^m c^k : n + m > k \geq 0\}$
B: $\{b^m c^k : m > k \geq 0\}$
C: $\{a^n b^m : n + m \geq 1\}$
D: $\{b^m : m \geq 1\}$

(d) $S \to \lambda \mid B \mid A \mid C$
$B \to bSaS \mid bScS$
$A \to aSbS$
$C \to cSbS$

■