

Huygens College Reflection

Assignment 1, Tuesday, Nov. 10, 2015

1. In the following table, give for every language a word that is in it and a word that is not in it. The alphabet is $\Sigma = \{a, b\}$.

e	$\in \mathcal{L}(e)$	$\notin \mathcal{L}(e)$
aa^*		
$(a \cup b)(a \cup b)^*$		
$(ab^*)^*$		
$a^*(ba^*)^*$		
$ab^*(a \cup \epsilon \cup aa)$		

2. Consider the languages $L_1 = \mathcal{L}((abba)^*)$, $L_2 = \mathcal{L}(a(bba)^*)$, $L_3 = \mathcal{L}((a(bba)^*)^*)$.
- Show that each of these languages is different.
 - For which pairs L_i, L_j (with $i \neq j$) do we have $L_i \subseteq L_j$? Prove your answer.
 - For which of these languages do we have $L_k L_k \neq L_k$? Prove your answer.

Answer Info:

- λ and $abbaabba$ are in L_1 and L_3 , but not in L_2 . So $L_2 \neq L_1$, $L_2 \neq L_3$.
(That they are in L_1 and L_3 is straightforward. That $\lambda \notin L_2$ as well. That $abbaabba \notin L_2$ follows from the fact that the words of L_2 in length order are a , $abba$, $abbabba$, $abbabbabba$, \dots , so $abbaabba \notin L_2$.)
 - a is in L_2 and L_3 but not in L_1 . So $L_1 \neq L_3$.
- $L_2 \subseteq L_3$, because L_3 is the Kleene-* of L_2 , and therefore contains it by definition. $L_1 \subseteq L_3$. Proof: The words in L_1 are exactly the words $(abba)^n$ for $n \geq 0$. The words in L_3 are exactly the words

$$a(bba)^{m_1} a(bba)^{m_2} \dots a(bba)^{m_n}$$

for $n \geq 0$ and $m_i \geq 0$ for all i ($1 \leq i \leq n$). So, a word in L_1 is also in L_3 , because we can take $m_1 = m_2 = \dots = m_n = 1$ in the pattern for the words of L_3 .

- $L_2 L_2 \neq L_2$, because $aa \in L_2 L_2$, but $aa \notin L_2$. For the others, $L_1 L_1 = L_1$ and $L_3 L_3 = L_3$, because their regular expressions are both of the shape $(-)^*$.

End Answer Info

3. Give a regular expression for the following languages and explain your answer.

-

$$\{w \in \{a, b, c\}^* \mid |w| \geq 3\}.$$

Answer Info:

$(a+b+c)(a+b+c)(a+b+c)(a+b+c)^*$. Explanation: A word in the language is of the shape $x_1 x_2 x_3 w$ with x_1, x_2 and x_3 in $\{a, b, c\}$ and w an arbitrary word over $\{a, b, c\}$. That is x_1, x_2 and x_3 are in $\mathcal{L}(a+b+c)$ and $w \in \mathcal{L}((a+b+c)^*)$.

Obvious alternatives $(a+b+c)^*(a+b+c)(a+b+c)(a+b+c)$, $(a+b+c)(a+b+c)(a+b+c)(a+b+c)(a+b+c)(a+b+c)(a+b+c)^*$

End Answer Info

(b)

$$\{w \in \{a, b\}^* \mid w \text{ begins with } b \text{ and } |w|_b \text{ is even}\}.$$

Answer Info:
 $ba^*ba^*(a^*ba^*ba^*)^*$. *Explanation:* A word in the language should start with a b and contain at least 2 b 's, so it is of the shape $ba^{n_1}ba^{n_2}w$ with $n_1, n_2 \geq 0$ and w having an even number of b 's. A word with an even number of b 's is of the shape $a^{m_1}ba^{m_2}ba^{m_3}b \dots a^{m_{2p}}ba^{m_{2p+1}}$ for $p \geq 0, M_i \geq 0$. This is captured by the regular expression $(a^*ba^*ba^*)^*$.
Obvious alternatives: $ba^*ba^*(ba^*ba^*)^*$, $(ba^*ba^*)^*ba^*ba^*$, $b(a^*ba^*b)^*a^*ba^*$.
End Answer Info

(c)

$$\{w \in \{a, b\}^* \mid bb \text{ doesn't occur in } w\}.$$

Answer Info:
 $a^*(baa^*)^*(b + \lambda)$. *Explanation:* If w is a word in the language, then every b in w should be directly followed by an a except for a possible last b !. So w starts with an arbitrary number of a 's, and then, possibly a sub-word ba , followed again by an arbitrary number of a 's, and so forth, ending with a possible single b . In conclusion, w is of the shape $a^pbaa^{m_1} \dots baa^{m_n}x$ with $x \in \{b, \lambda\}, p \geq 0, n \geq 0$ and each $m_i \geq 0$.
End Answer Info

4. [This exercise is hard now. They show that the subject is non-trivial. Later we will learn methods to solve this more easily]

(a) Show that the language

$$\{w \in \{a, b\}^* \mid aa \text{ occurs exactly twice in } w\}.$$

is regular.

[Hint. Beware of the string aaa !]

Answer Info:
A possible answer is:

$$(b^*(abb^*)^*aaab^*(abb^*)^*(a + \lambda)) + (b^*(abb^*)^*aabb^*(abb^*)^*aab^*(abb^*)^*(a + \lambda)).$$

This solution borrows from 3(c) the regular expression that does not contain aa , either at the end (expression $b^*(abb^*)^*(a + \lambda)$) or at the beginning (expression $b^*(abb^*)^*$). The expression $(b^*(abb^*)^*aaab^*(abb^*)^*(a + \lambda))$ covers the words that contain aaa , the expression $(b^*(abb^*)^*aabb^*(abb^*)^*aab^*(abb^*)^*(a + \lambda))$ covers the words that contain two separate sub-words aa .

Another possible solution (from Twan): $b^*(abb^*)^*aa(a + bb^*(abb^*)aa)(bb^*a)^*b^*$

End Answer Info