TOC PROBLEM SET-3

SIDDHANT CHAUDHARY
BMC201953

1. Write rational expressions for the following languages.

(a) The set of strings over $\Sigma = \{a, b, c\}$ containing atleast one a and atleast one b. We just need to ensure that there our string has ab or ba as a subsequence, which can be achieved as follows.

$$[(a \cup b \cup c)^* \cdot a \cdot (a \cup b \cup c)^* \cdot b \cdot (a \cup b \cup c)^*] \cup [(a \cup b \cup c)^* \cdot b \cdot (a \cup b \cup c)^* \cdot a \cdot (a \cup b \cup c)^*]$$

(b) The set of strings of 0's and 1's with at most one pair of consecutive 1's. First, we find a rational expression for the language consisting of all words not containing consecutive $1'$s and not starting with a 1. This language has the expression

$$E_0 = (0^* \cdot (\epsilon \cup 01) \cdot 0^*)^*$$

Then, the expression for the language consisting of all words not containing consecutive 1's and starting with 1 is

$$E_1 = 1 \cdot E_0$$

So we see that

$$E = E_0 \cup E_1$$

is the expression for the language consisting of all words which don't have consecutive 1's. So, the expression for the required language is

$$E \cup [((E \cdot 0) \cup \epsilon) \cdot 1 \cdot 1 \cdot E_0]$$

(c) Strings of 0s and 1s which do not contain 101 as a substring. This expression is easy to construct. Note that a 1 can only be followed by a 1 or 00 or a 0 in the case where the string ends with a 10. So, the rational expression is given below.

$$0^* \cdot (1^* \cdot 0 \cdot 0 \cdot 0^*)^* \cdot 1^* \cdot 0^*$$

The part $1^* \cdot 0^*$ of this expression handles the boundary case where a string ends with a 10.

2. For each of the rational expressions obtained in the previous problem, construct an NFA accepting the corresponding rational language.

Solution: To be completed.

3. Let the rotational closure of a language L be given as

$$\text{ROC}(L) = \{yx | xy \in L\}$$

Show that the set of recognizable languages is closed under rotational closure, and $\text{ROC}(L) = \text{ROC}(\text{ROC}(L))$.

Solution: I was a bit lazy to write this solution. This link has a good explanation.

Date: September 2020.

4. Suppose $K, L \subset \{a, b\}^*$ are languages such that K does not contain the empty word. Show that $X = K^*L$ is the unique solution of the equation

$$X = KX + L$$

where X is an unknown language.

Bonus: What can you say about X if K contains the empty word?

Solution: First, we show that for any such X, $K^*L \subset X$. So let $y \cdot l \in K^*L$, where $y \in K^*$ and $l \in L$. If $y = \epsilon$, then clearly $y \cdot l = l \in X$, because L is a subset of X. If $y \neq \epsilon$, then $y = y_1y_2...y_p$ where $y_i \in K$ for each $1 \leq i \leq p$. Then, observe that $l \in X$, which implies that $y_p \cdot l \in X$, which implies that $y_{p-1}y_p \cdot l \in X$, and continuing this way p times we see that $y_1y_2...y_p \cdot l = y \cdot l \in X$. Hence, $K^*L \subset X$.

Next, we show that $X \subset K^*L$, by induction on the length of the word taken in X. For the base case, let $x \in X$ such that the length of x is 0, i.e $x = \epsilon$. Since K does not contain the empty word, we have that $x \in L$. There are two cases: if $\epsilon \in L$, then in this case, $\epsilon \in K^*L$ is true. If $\epsilon \notin L$, then there is no such x, and hence the base case is vacuously true.

For the inductive case, suppose $x \in X$ and $|x| \leq m \implies x \in K^*L$. Let $x \in X$ be such that $|x| = m + 1$. If $x \in L$, then $x \in K^*L$, and hence we are done. So, suppose $x \notin L$. Hence, $x \in KX$, so that

$$x = k \cdot x_1$$

where $k \in K$ and $x_1 \in X$ is such that $|x| \leq m$. By our hypothesis, $x_1 \in K^*L$, and hence $x = k \cdot x_1 \in K^*L$, completing the proof that $X \subset K^*L$. So, we have shown that $X = K^*L$ is the unique solution to the given equation.

Now, if K contains ϵ, then we have the following claim.

Any Xsuch that $K^*L \subset X$ is a solution to the given equation.

I didn't have time to write a proof for this, but the arguments are very similar to the ones given above.

5. Let L be any language. Define DROPOUT(L) as

$$\text{DROPOUT}(L) := \{xz | xyz \in L, y \in \Sigma, x, z \in \Sigma^*\}$$

Show that the set of regular languages is closed under the DROPOUT operation.

Solution: Let M be a DFA accepting L. We make an ϵ-NFA N as follows. Put two copies of M, say M_1 and M_2 next to each other, and let the initial state be the initial state of M_1, and let the final states be the final states of M_2. The transitions in M_1 and M_2 remain the same, and we add more ϵ-transitions as follows. If $q_1 \xrightarrow{s} q_2$ is a transition in M, add $q_1 \xrightarrow{\epsilon} q_2$ where q_1 is the copy in M_1 and q_2 is the copy in M_2. Then, we claim that exactly those words are accepted which have exactly one letter dropped. If a word is accepted, it must have gone through an ϵ-transition (because the final states reside in M_2), and hence it has exactly one letter dropped. Conversely, if there is a word which is of the form xz with $xyz \in L$ and $y \in \Sigma$, then we can drop the letter y by going through a suitable ϵ-transition. Hence, the class of regular languages is closed under the DROPOUT operation.