THE ALGEBRA OF REGULAR EXPRESSIONS

Reminder of Basic Definitions and Some Basic Proofs

(1) For languages $L, M \subseteq \Sigma^*$; $L+M$, $L.M$ and L^* are interpreted as follows :

$L+M = L \cup M$; $L.M = \{w \mid w = u.v ;\ u \in L ; v \in M\}$; $L^* = \cup_{i=0,+\infty} L^i$ where $L^i := L.L. \ldots .L$ *(i times)*

(2) $(L+M)^* = (L^*. M^*)^*$

Proof of (2):

Let $u \in (L+M)^*$ then by definition $u = u_1 .u_2 . \ldots . u_k$ for some integer $k \geq 0$ where for each j $u_j \in L+M$. But $L \subseteq L^* \subseteq L^*. e \subseteq L^* . M^*$ and $M \subseteq M^* \subseteq e . M^* \subseteq L^* . M^*$ and thus $u_j \in L^*. M^* + L^*. M^* = L^*. M^*$ and therefore $(L+M)^* \subseteq (L^* . M^*)^*$

Conversely let $u \in (L^*.M^*)^*$ then by definition $u = u_1 .u_2 . \ldots . u_k$ where $u_j \in L^*.M^*$ hence $u_j = v_j^1 . v_j^2 \ldots .v_j^{l(j)} . w_j^1 . w_j^2 \ldots .w_j^{p(j)}$ where $v_j^m \in L$ and $w_j^m \in M$. Hence $u = z_1 . z_2 . \ldots z_q$ where $q = \sum_{j=1,k} l(j)+p(j)$ and each $z_j \in L+M$. This proves that $(L^*.M^*)^* \subseteq (L +M)^*$ which proves that $(L +M)^* = (L^*. M^*)^*$

(3) $(L+M)^* = (L^*+M^*)^*$

Proof of (3):

Since $L \subseteq L^*$ and $M \subseteq M^*$ it follows that $(L+M)^* \subseteq (L^*+M^*)^*$.

Conversely let $u \in (L^*+M^*)^*$ then $u = (v_1+w_1). \ \ldots \ . (v_k + w_k)$ where for each j $v_j \in L^*$ and $w_j \in M^*$. We show that $u \in (L^*. M^*)^*$ by using induction on k. For $k=1$ $v_1 \in L^* \subseteq L^*. e \subseteq L^*.M^* \subseteq (L^*. M^*)^*$ similarly $w_1 \in M^* \subseteq e. M^* \subseteq L^*.M^* \subseteq (L^*. M^*)^*$ hence $v_1+w_1 \subseteq (L^*. M^*)^*$. Now assume statement holds for $k-1$, hence $z := (v_1+w_1). \ \ldots \ . (v_{k-1} + w_{k-1}) \in (L^*. M^*)^*$. But using the above reasoning for v_1+w_1 it follows that $v_k+w_k \in (L^*. M^*)^*$ and therefore $u = z . (v_k+w_k) \in (L^*. M^*)^*. (L^*. M^*)^* = (L^*. M^*)^*$ using the obvious identity $K^* . K^* = K^*$ for any language K. This proves that $(L^*+M^*)^* \subseteq (L^*. M^*)^*$, but by (2) $(L+M)^* = (L^*. M^*)^*$ hence $(L^*+M^*)^* \subseteq (L + M)^*$ and (3) is proved.

(4) $(L.M)^* \subseteq (L^*M^*)^*$ and $(L.M)^* = (L^*M^*)^*$ iff $e \in L$ and $e \in M$

Proof of (4):

First statement is obvious using $L \subseteq L^*$ and $M \subseteq M^*$.

To prove the second one assume $e \in L$ and $e \in M$

and let $u \in (L^* . M^*)^*$ then $u = v_1 . w_1 \ldots . v_k . w_k$ where $v_j \in L^*$ and $w_j \in M^*$ therefore

$vj = y_j^1 . \ldots . y_j^{l(j)}$ and $wj = z_j^1 . \ldots . z_j^{p(j)}$ with $y_j^m \in L$ and $z_j^m \in M$. Hence

$u = q_1 . \ldots . q_r$ where $r = \sum_{j=1,k} (l(j)+p(j))$ where each $q_i \in L$ or $q_i \in M$. Using the assumption we can write $u = q'_1 . \ldots . q'_{r'}$ by adding an empty string in between the q_j strings ,if necessary, so that we have for each $j=1, \ldots , r'$, $q'_j \in L$ and $q'_{j+1} \in M$. This proves that $u \in (L.M)^*$ To prove the converse result we present counter-examples that violate the assumption $e \in L$ and $e \in M$.

Suppose $e \notin L$ choose $L = 0.0^*$ and $M = 1^*$ then $1 \in (L^*.M^*)^*$ whereas $1 \notin (L.M)^*$; alternatively if $e \notin M$ choose $L = 0^*$ and $M = 1.1^*$ then $0 \in (L^*.M^*)^*$ whereas $0 \notin (L.M)^*$.

Homework #2 *due March 25, Thursday 2021, before recitation*

(1) Using either the results or the techniques used above try to simplify the following expressions and prove your simplification.

(i) $(0+1)^*.1.(0+1) +(0+1)^*.1.(0+1)$

(ii) $(((0^*.1^*)+1)^*(0+1)^*)^*$

(iii) $(L+M^*)^*$

(iv) $(L.M^*)^*$

(2) Convert the *regular expression* $((0.0^*.(1.1))+ 0.1)^*$ into an ***ε-NFA***

(3) Problems from the textbook

3.1.1 (b) and (c)

3.1.4 (b) and (c)

3.2.1 (c) and (d)

3.2.3