ST. CLOUD STATE
UNIVERSITY

Dept. of Computer Science

CSCI 402/502: Introduction to Theory of Computation
Instructor: Pranava K. Jha

Closure of Regular sets under Reversal

For a string w, let w^{R} denote the reversal of w, and for a set L, let Rev(L) denote the set $\{w^R: w \in L\}$. That regularity of L implies regularity of Rev(L) may be proved in at least three different ways:

1. Start with a *DFA*, say M, which accepts a regular language L, and build an *NFA* (with ε-transitions) to accept Rev(L) by appropriately introducing a new start state, say p, with an ε-transition to each of the final states of M (that are no longer final states of the new machine), reversing the arcs of M, and making the start state of M the sole final state of the new machine.

2. Start with a *right-linear grammar*, say G, which generates a regular language L, and build a grammar by systematically reversing the right side of each production of G. The resulting grammar generates Rev(L) and is necessarily *left-linear*, hence Rev(L) is regular. (Equivalently, start with a left-linear grammar and likewise build a right-linear grammar.)

3. Start with a *regular expression*, say **r**, for a regular language L, and build a regular expression for Rev(L). To that end, proceed by induction on the number of operators in **r**.

The present note amplifies the third scheme outlined above. First a few lemmas.

Lemma 1: $\mathrm{Rev}(L_1 \cup L_2) = \mathrm{Rev}(L_1) \cup \mathrm{Rev}(L_2)$.■

Lemma 2: $\mathrm{Rev}(L_1 \cdot L_2) = \mathrm{Rev}(L_2) \cdot \mathrm{Rev}(L_1)$.

Proof: Let $w \in \mathrm{Rev}(L_1 \cdot L_2)$. Then w may be written as $w = (xy)^{\mathrm{R}}$ where $x \in L_1$ and $y \in L_2$. Now, $(xy)^{\mathrm{R}} = y^{\mathrm{R}}x^{\mathrm{R}}$ that is clearly in $\mathrm{Rev}(L_2) \cdot \mathrm{Rev}(L_1)$. Accordingly, $\mathrm{Rev}(L_1 \cdot L_2) \subseteq \mathrm{Rev}(L_2) \cdot \mathrm{Rev}(L_1)$.

For the reverse inclusion, let $w \in \mathrm{Rev}(L_2) \cdot \mathrm{Rev}(L_1)$. Then w may be written as $w = y^{\mathrm{R}}x^{\mathrm{R}}$ where $y \in L_2$. and $x \in L_1$. Now, $y^{\mathrm{R}}x^{\mathrm{R}} = (xy)^{\mathrm{R}}$ that is clearly in $\mathrm{Rev}(L_1 \cdot L_2)$. Accordingly, $\mathrm{Rev}(L_2) \cdot \mathrm{Rev}(L_1) \subseteq \mathrm{Rev}(L_1 \cdot L_2)$. ■

Lemma 3: $\text{Rev}(L^*) = (\text{Rev}(L))^*$.

Proof: It is clear that the empty string ε is in each of $\text{Rev}(L^*)$ and $(\text{Rev}(L))^*$. In what follows, all strings are of length at least one.

Let w be a typical element of $\text{Rev}(L^*)$. Then $w = x^R$ for some $x \in L^*$. Note that x may be written as $x = x_1 \ldots x_n$, where $n \geq 1$ and $x_i \in L$ for $1 \leq i \leq n$. Now, $w = x^R = (x_1 \ldots x_n)^R = x_n^R \ldots x_1^R$. Since x_i^R is in $\text{Rev}(L)$ for $1 \leq i \leq n$, it is clear that $x_n^R \ldots x_1^R$ (that is equal to w) is in $(\text{Rev}(L))^*$. Thus, $\text{Rev}(L^*) \subseteq (\text{Rev}(L))^*$.

For the reverse inclusion, let w be a typical element of $(\text{Rev}(L))^*$. Then $w = w_1 \ldots w_n$, where $n \geq 1$ and $w_i \in \text{Rev}(L)$ for $1 \leq \text{i} \leq n$. This means that $w_i^R \in L$ for $1 \leq \text{i} \leq n$, i.e., $w_n^R \ldots w_1^R \in L^*$. Since $w_n^R \ldots w_1^R = (w_1 \ldots w_n)^R = w^R$, it is clear that w^R is in L^*, and hence $(w^R)^R$ (that is equal to w) is in $\text{Rev}(L^*)$. Thus, $(\text{Rev}(L))^* \subseteq \text{Rev}(L^*)$. ■

Reversal of the Kleene star of a set is equal to the Kleene star of the reversal.

Theorem: If L is a regular set, then so is $\text{Rev}(L)$.

Proof: Let L be a regular set. Accordingly, L is denoted by a regular expression, say $\mathbf{r}$. It suffices to show that $\text{Rev}(L)$ is denoted by a regular expression. Let k be the number of operators in $\mathbf{r}$. Induct on k to prove the claim.

For $k = 0$, $\mathbf{r}$ is of the form ϕ, ε or $\boldsymbol{a}$ where a is a member of the alphabet. Accordingly, L is equal to one of $\varnothing$, $\{\varepsilon\}$ and $\{a\}$, whence $\text{Rev}(L) = L$, and the claim follows.

For $k \geq 1$, $\mathbf{r}$ is of one of the following forms: $\mathbf{r_1} + \mathbf{r_2}$, $\mathbf{r_1} \cdot \mathbf{r_2}$ and $\mathbf{r_1}^*$ where $\mathbf{r_1}$ and $\mathbf{r_2}$ are themselves regular expressions. Let L_1 and L_2 be the languages denoted by $\mathbf{r_1}$ and $\mathbf{r_2}$, respectively. It is clear that the number of operators in each of $\mathbf{r_1}$ and $\mathbf{r_2}$ is strictly less than k. By induction hypothesis, $\text{Rev}(L_1)$ and $\text{Rev}(L_2)$ are denoted by regular expressions, say $\mathbf{s_1}$ and $\mathbf{s_2}$, respectively.

(i) $\mathbf{r} = \mathbf{r_1} + \mathbf{r_2}$**:** In this case, $L = L_1 \cup L_2$. By Lemma 1, $\text{Rev}(L_1 \cup L_2) = \text{Rev}(L_1) \cup \text{Rev}(L_2)$ that is clearly denoted by the regular expression $\mathbf{s_1} + \mathbf{s_2}$.

(ii) $\mathbf{r} = \mathbf{r_1} \cdot \mathbf{r_2}$**:** In this case, $L = L_1 \cdot L_2$. By Lemma 2, $\text{Rev}(L_1 \cdot L_2) = \text{Rev}(L_2) \cdot \text{Rev}(L_1)$ that is clearly denoted by the regular expression $\mathbf{s_2} \cdot \mathbf{s_1}$.

(iii) $\mathbf{r} = \mathbf{r_1}^*$**:** In this case, $L = L_1^*$. By Lemma 3, $\text{Rev}(L_1^*) = (\text{Rev}(L_1))^*$ that is clearly denoted by the regular expression $\mathbf{s_1}^*$.■

——o——