COL352 Lecture 7

Contents

1 Recap

Completion of proof of the second part of the proof of the class of regular languages being closed under concatenation.

2 Definitions

Definition 1

Let $L \in \Sigma^*$ be any language. We define the language L^* as follows:

$$L^* = L_0 \cup L_1 \cup \cdots = \bigcup_{n=0}^{\infty} L^n$$

where $L^0 = \{\epsilon\}, L_1 = L, L^n = L \cdot L \cdots L$ where there are n instances of L.

Definition 2

Let Σ be a finite alphabet. A regular expression over Σ is any expression that is in one of the following forms:

1. $\emptyset$.
2. ϵ.
3. a, where $a \in \Sigma$.
4. $(R_1 \cup R_2)$ where R_1, R_2 are regular expressions over Σ.
5. $(R)^*$ where R is a regular expression over Σ.

Definition 3

The language of a regular expression R, denoted by $\mathcal{L}(R)$ is defined as follows:

1. $\mathcal{L}(\emptyset) = \emptyset$.
2. $\mathcal{L}(\epsilon) = \{\epsilon\}$.
3. $\mathcal{L}(a) = \{a\}$.
4. $(\mathcal{R}_\infty \cup \mathcal{R}_\in) = \mathcal{L}(R_1) \cup \mathcal{L}(R_2)$.
5. $\mathcal{L}\left((R)^*\right) = \left(\mathcal{L}(R)\right)^*$

3 Content

3.1 Closure under *

Theorem 1

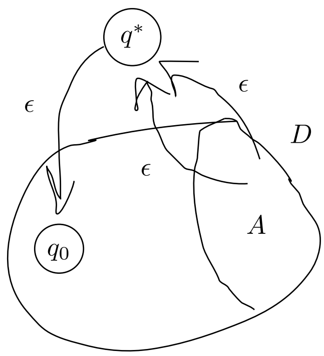
The class of regular languages is closed under $*$.

Proof. Let L be regular. Let D be a DFA recognizing L. Construct an NFA recognizing L^*.

For the DFA $D = (Q, \Sigma, \delta, q_0, A)$, consider the NFA $N = (Q \uplus q^*, \Sigma, \Delta, \{q^*\}, \{q^*\})$.

where

$$\Delta = \{(q, a, q') \mid \delta(q, a) = q' \wedge q, q' \in Q \wedge a \in A\} \cup \{(q_a, \epsilon, q^*) \mid q_a \in A\} \cup (q^*, \epsilon, q_0)$$



Claim 1.1

If N accepts x, then $x \in L^*$.

Proof. Consider an accepting run of N on x Suppose it visits q^* $n+1$ times. Then we claim $x \in L^n$.

Proof. Induction on n. $n = 0 \implies$ empty string.
Now suppose $n > 0$. Consider the second-last time the run visits q^*. After that state, it transitions to q_0, and that is a run of D on the substring corresponding to whatever is left in x after that state, and the part of the run before that is an accepting run of N on the prefix of x that was deleted, and it visits q^* exactly n times. Hence $x = x'x''$ where $x' \in L^{n-1}$ by the inductive hypothesis and $x'' \in L$ as D accepts x''. Hence we have $x \in L^n$, as required. ■

Hence we are done. ■

Claim 1.2

If $x \in L^*$, then N accepts x.

Proof. $x \in L^* \implies \exists n$ such that $x \in L^n$. Let n be the least integer such that $x \in L^n$. Now we proceed with a proof by induction on n.
Base case: $n = 0$ means empty string, so we are done.
Now consider $n > 0$. Then $x \in L^n \implies x = x_1 \cdot x_2$ such that $x_1 \in L_{n-1}$ and $x_2 \in L$. By induction hypothesis, N accepts x_1. Construct an accepting run of N on x as follows:
[run of N on x_1 starting and ending with q^*] ϵ [accepting run of D on x_2 starting at q_0 and ending at $q \in A$] ϵ q^*. ■

■

Note 1

In general, to prove closure of the class of regular languages under some binary operation:

1. Assume L_1, L_2 are some arbitrary regular languages. Let D_1, D_2 be the DFAs recognizing L_1, L_2 respectively.
2. Construct an NFA N recognizing $L_1 \text{op} L_2$.
3. Argue that $x \in L_1 \text{op} L_2$ iff N accepts x.

3.2 Regular expressions

Theorem 2

Suppose $L \in \Sigma^*$ is a language generated by some regular expression over Σ. Then L is a regular language.