Closure properties and regular expressions

Arno Pauly

February 12, 2021

Notation

- $\wedge$, $\vee$ denote boolean *and* and *or*
- $\cup$ is the union of sets, i.e. $A \cup B := \{x \mid x \in A \vee x \in B\}$
- $\cap$ is the intersection of sets,
  i.e. $A \cap B := \{x \mid x \in A \wedge x \in B\}$

A closure property

Theorem

If L_1,L_2 are regular languages, then so are $L_1 \cap L_2$ and $L_1 \cup L_2$.

The technical tool

Definition

Given two finite automata $\mathcal{A}_i = (V_i, s_i, \delta_i, F_i)_{i \in \{1,2\}}$, let their product be $(V_1 \times V_2, (s_1, s_2), \delta_\times, F)$ where

1. $((q_1, q_2), a, (q_1', q_2')) \in \delta_\times$ iff $(q_i, a, q_i') \in \delta_i$ for both $i \in \{1, 2\}$ (non-deterministic case)
2. $\delta_\times((q_1, q_2), a) = (\delta_1(q_1), \delta_2(q_2))$ (deterministic case)
3. and typically $F = \{(q_1, q_2) \mid q_1 \in F_1 \vee q_2 \in F_2\}$ or $F = F_1 \times F_2 = \{(q_1, q_2) \mid q_1 \in F_1 \wedge q_2 \in F_2\}$

Concatenation

Definition

Given languages L_1, L_2, let their concatenation be

$$L_1 \circ L_2 := \{uw \mid u \in L_1 \wedge w \in L_2\}$$

(sometimes written at $L_1 L_2$).

Theorem

If L_1, L_2 are regular, then so is $L_1 \circ L_2$.

Kleene-star

Definition

Given a language L, let $L^0 = \{\varepsilon\}$, $L^{n+1} = LL^n$ and $L^* = \bigcup_{n\in\mathbb{N}} L^n$.

Theorem

If L is regular, so is L^*.

Regular expressions

Definition

Regular expressions are defined as follows:

1. $\emptyset$ is a regular expression.
2. ε is a regular expression.
3. a is a regular expression for each $a \in \Sigma$.
4. $R|Q$ is a regular expression whenever R and Q are.
5. RQ is a regular expression whenever R and Q are.
6. R^* is a regular expression whenever R is.

Warning!!

Regex in Perl are not regular expressions.

Meaning

1. $\emptyset$ denotes the empty language.
2. ε denotes the language $\{\varepsilon\}$
3. a denotes the language $\{a\}$
4. $R|Q$ denotes the language given by the union of the languages denoted by R,Q
5. RQ denotes the language given by the concatenation of the languages denoted by R,Q
6. R^* denotes the language given by the Kleene star of the language denoted by R

Connection

Theorem

A language is regular if and only if there is a regular expression denoted by it.

- That regular expressions denote regular languages follows from the closure properties we saw today.
- We'll leave the other direction for later.