Closure properties of regular languages

Deepak D'Souza

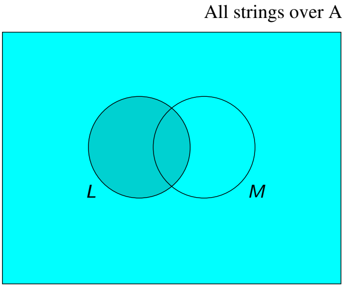
Department of Computer Science and Automation
Indian Institute of Science, Bangalore.

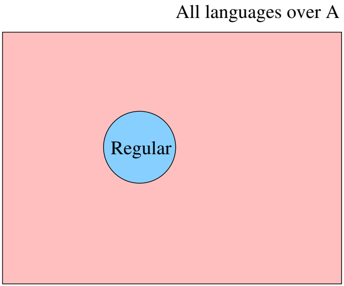
14 August 2012

Outline

Closure properties

- Class of Regular languages is closed under
  - Complement, intersection, and union.
  - Concatenation, Kleene iteration.
- Non-deterministic Finite-state Automata (NFA) = DFA.





Closure under complementation

- Idea: Flip final states.
- Formal construction:
  - Let $\mathcal{A} = (Q, s, \delta, F)$ be a DFA over alpahet A.
  - Define $\mathcal{B} = (Q, s, \delta, Q - F)$.
  - Claim: $L(\mathcal{B}) = A^* - L(A)$.

Proof of claim

- $L(\mathcal{B}) \subseteq A^* - L(\mathcal{A})$.

$$\begin{aligned} w \in L(\mathcal{B}) &\implies \widehat{\delta}(s, w) \in (Q - F). \\ &\implies \widehat{\delta}(s, w) \notin F \\ &\implies w \notin L(\mathcal{A}) \\ &\implies w \in A^* - L(\mathcal{A}). \end{aligned}$$

- $L(\mathcal{B}) \supseteq A^* - L(\mathcal{A})$.

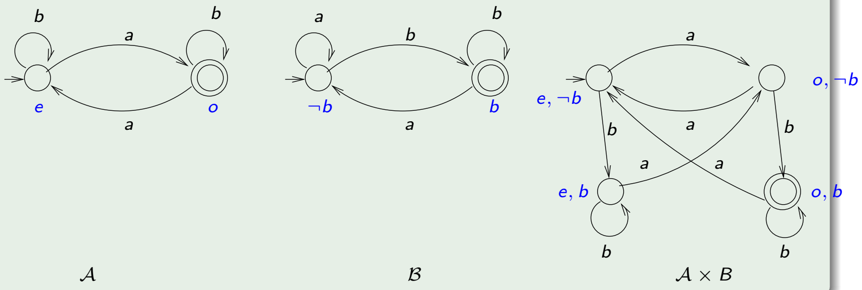
Closure under intersection

Product construction. Given DFA's $\mathcal{A} = (Q, s, \delta, F)$, $\mathcal{B} = (Q', s', \delta', F')$, define product $\mathcal{C}$ of A and B:

$$\mathcal{C} = (Q \times Q', (s, s'), \delta'', F \times F'),$$

where $\delta''((p, p'), a) = (\delta(p, a), \delta'(p', a))$.

Product construction example



Correctness of product construction

Claim: $L(\mathcal{C}) = L(\mathcal{A}) \cap L(\mathcal{B})$.

Proof of claim $L(\mathcal{C}) = L(\mathcal{A}) \cap L(\mathcal{B})$.

- $L(\mathcal{C}) \subseteq L(\mathcal{A}) \cap L(\mathcal{B})$.

$$\begin{aligned} w \in L(\mathcal{C}) &\implies \widehat{\delta}''((s,s'),w) \in F \times F'. \\ &\implies (\widehat{\delta}(s,w), \widehat{\delta}'(s',w)) \in F \times F' \text{ (by subclaim)} \\ &\implies \widehat{\delta}(s,w) \in F \text{ and } \widehat{\delta}'(s',w) \in F' \\ &\implies w \in L(\mathcal{A}) \text{ and } w \in L(\mathcal{B}) \\ &\implies w \in L(\mathcal{A}) \cap L(\mathcal{B}). \end{aligned}$$

- $L(\mathcal{C}) \supseteq L(\mathcal{A}) \cap L(\mathcal{B})$.

Subclaim: $\widehat{\delta}''((s,s'),w) = (\widehat{\delta}(s,w), \widehat{\delta}'(s',w))$.

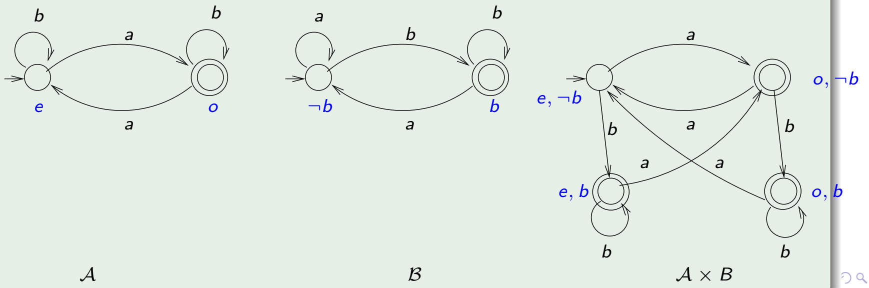
Closure under union

- Follows from closure under complement and intersection since $L_1 \cup L_2 = \overline{\overline{L_1} \cap \overline{L_2}}$.

Closure under union

- Follows from closure under complement and intersection since $L_1 \cup L_2 = \overline{\overline{L_1} \cap \overline{L_2}}$.
- Can also do directly by product construction: Given DFA's $\mathcal{A} = (Q, s, \delta, F)$, $\mathcal{B} = (Q', s', \delta', F')$, define $\mathcal{C}$: $\mathcal{C} = (Q \times Q', (s, s'), \delta'', (F \times Q') \cup (Q \times F'))$, where $\delta''((p, p'), a) = (\delta(p, a), \delta(p', a))$.

Union construction



Principle of Mathematical Induction

- $\mathbb{N} = \{0, 1, 2 \ldots\}$
- $P(n)$: A statement P about a natural number n.
- Example:
  - $P(n) =$ "n is even."
  - $P_1(n) =$ "Sum of the numbers $1 \ldots n$ equals $n(n+1)/2$."
  - $P_2(n) =$ "For all $w \in A^*$, if length of w is n then $\widehat{\delta}''((s, s'), w) = (\widehat{\delta}(s, w), \widehat{\delta}'(s', w))$."

Principle of Induction

If a statement P about natural numbers

- is true for 0 (i.e $P(0)$ is true), and,
- is true for $n+1$ whenever it is true for n (i.e. $P(n) \implies P(n+1)$)

then P is true of all natural numbers (i.e. "For all n, $P(n)$" is true).

Proof of subclaim

Exercise: Prove the Subclaim:

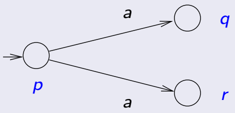
$$\widehat{\delta}''((s, s'), w) = (\widehat{\delta}(s, w), \widehat{\delta}'(s', w)).$$

using induction.

Nondeterministic Finite-state Automata (NFA)

- Allows multiple start states.
- Allows more than one transition from a state on a given letter.

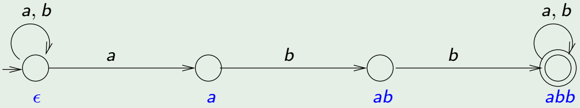
Non-deterministic transitions



- A word is accepted if there is some path on it from a start to a final state.

Example NFA's

NFA for "contains *abb* as a subword"



NFA definition

- Mathematical representation of NFA
  - $\mathcal{A} = (Q, S, \Delta, F)$, where $S \subseteq Q$, and $\Delta : Q \times A \to 2^Q$.
  - Define relation $p \xrightarrow{w} q$ which says there is a path from state p to state q labelled w.
    - $p \xrightarrow{\epsilon} p$
    - $p \xrightarrow{ua} q$ iff there exists $r \in Q$ such that $p \xrightarrow{u} r$ and $q \in \Delta(r, a)$.
  - Define $L(\mathcal{A}) = \{w \in A^* \mid \exists s \in S, f \in F : \; s \xrightarrow{w} f\}$.
- NFA $\to$ DFA: Subset construction
  - Example: determinize NFA for "contains *abb*."
  - Formal construction
  - Correctness

Closure under concatenation and Kleene iteration

- Concatenation of languages:

$$L \cdot M = \{u \cdot v \mid u \in L,\ v \in M\}.$$

- Kleene iteration of a language:

$$L^* = \{\epsilon\} \cup L \cup L^2 \cup L^3 \cup \cdots ,$$

where

$$\begin{aligned} L^n &= L \cdot L \cdots L \ (n \text{ times}). \\ &= \{w_1 \cdots w_n \mid \text{each } w_i \in L\}. \end{aligned}$$