Formal Languages and Automata

Chapter 4
Properties of Regular Languages

Chuan-Ming Liu

cmliu@csie.ntut.edu.tw

Department of Computer Science and Information Engineering
National Taipei University of Technology
Taipei, TAIWAN
Objectives

Problem: how general regular languages are?

Questions:
1. what will happen when we perform operations on regular languages?
2. what are the properties of different kinds of languages?
3. how can we tell if a given language is regular or not?
Contents

- Closure Properties of Regular Languages
- Elementary Questions about Regular Languages
- Identifying Nonregular Languages
Closure Properties

Theorem:
If $L_1$ and $L_2$ are regular languages, then so are $L_1 \cup L_2$, $L_1 \cap L_2$, $L_1 L_2$, $\overline{L_1}$, and $L_1^*$. In other words, the family of regular languages is closed under union, intersection, concatenation, complement, and star-closure.

Proof:
Example 1

Show that the family of regular languages is closed under difference. (i.e., If $L_1$ and $L_2$ are regular, then $L_1 - L_2$ is also regular.)

Solution: Consider $L_1 - L_2 = L_1 \cap \overline{L_2}$. Since the closure of regular languages under intersection and complement, so $L_1 - L_2$ is regular.
Reversal

Theorem:
The family of regular languages is closed under reversal.
Homomorphism

Definition:
Suppose $\Sigma$ and $\Gamma$ are alphabets. Then a function

$$h : \Sigma \rightarrow \Gamma^*$$

is called a *homomorphism*.

Note: A homomorphism is a substitution in which a single letter is replaced with a string.

Definition:
If $L$ is a language on $\Sigma$, then its *homomorphic image* is defined as

$$h(L) = \{ h(w) : w \in L \}.$$
Example 2

Let $\Sigma = \{a, b\}$ and $\Gamma = \{a, b, c\}$ and define $h$ by

\[
\begin{align*}
  h(a) &= ab \\
  h(b) &= bbc.
\end{align*}
\]

Then, $h(aba) = abbbcab$. The homomorphic image of $L = \{aa, aba\}$ is the language $h(L) = \{abab, abbbcab\}$. 
Example 3

Let $\Sigma = \{a, b\}$ and $\Gamma = \{b, c, d\}$ and define $h$ by

\[
\begin{align*}
    h(a) &= dbcc \\
    h(b) &= bdc.
\end{align*}
\]

If $L$ is a regular language denoted by

\[
r = (a + b^*)(aa)^*,
\]

Then,

\[
r_1 = (dbcc + (bdc)^*)(dbccdbcc)^*
\]

denotes $h(L)$. 
Closure under homomorphism

Theorem:
Let $h$ be a homomorphism. If $L$ is a regular language, then its homomorphic image $h(L)$ is also regular. The family of regular languages is therefore closed under arbitrary homomorphisms.
Right Quotient

Definition:
Let \( L_1 \) and \( L_2 \) be languages on the same alphabet. Then, the right quotient of \( L_1 \) with \( L_2 \) is defined as

\[
L_1 / L_2 = \{ x : xy \in L_1 \text{ for some } y \in L_2 \}.
\]

Example 4:
If

\[
L_1 = \{ a^n b^m : n \geq 1, m \geq 0 \} \cup \{ ba \}
\]

and

\[
L_2 = \{ b^m : m \geq 1 \},
\]

then,

\[
L_1 / L_2 = \{ a^n b^m : n \geq 1, m \geq 0 \}.
\]
Dfa’s for $L_1$ and $L_1/L_2$
Closure under Right Quotient

**Theorem:**
If $L_1$ and $L_2$ are regular languages, then $L_1 / L_2$ is also regular. We say that the family of regular languages is closed under right quotient with a regular language.
Example 5

Find $L_1 / L_2$ for

\[
L_1 = L(a^* baa^*) \\
L_2 = L(ab^*).
\]

\[
L_1 = L(a^* baa^*) \\
L_1 / L_2 = L(a^* ba^*)
\]
Contents

- Closure Properties of Regular Languages
- Elementary Questions about Regular Languages
- Identifying Nonregular Languages
Elementary Questions

Consider the following decision problem: Given a language $L$ and a string $w$, decide if $w \in L$.

**Standard Representation**
A regular language is given in a standard representation if and only if it is described by a finite automaton, a regular expression, or a regular grammar.
Theorems

Theorem:
Given a standard representation of any regular language $L$ on $\Sigma$ and any $w \in \Sigma^*$, there exists an algorithm for determining whether or not $w \in L$.

Theorem:
There exists an algorithm for determining whether a regular language, given in a standard representation, is empty, finite, and infinite.
Equality of Two Languages

Theorem:
Given standard representations of two regular languages $L_1$ and $L_2$, there exists an algorithm to determine whether or not $L_1 = L_2$. 
Contents

- Closure Properties of Regular Languages
- Elementary Questions about Regular Languages
- Identifying Nonregular Languages
Identifying Nonregular Languages

Recall the *Pigeonhole Principle*

**Example 6:**
Is the language $L = \{a^n b^n : n \geq 0\}$ regular?

**Observation:**
In a transition graph with $n$ vertices, any walk of length $n$ or longer must repeat some vertex. *i.e.*, has a cycle.
Pumping Lemma

Let $L$ be an infinite regular language. Then, there exists some positive integer $m$ such that any $w \in L$ with $w \geq m$ can be decomposed as

$$w = xyz$$

with

$$|xy| \leq m,$$

and

$$|y| \geq 1,$$

such that

$$w_i = xy^i z,$$

is also in $L$ for all $i = 0, 1, 2, \ldots$. 
Using Pumping Lemma

Four moves when using the pumping lemma

- Select an integer $m$.
- Given $m$, pick a string $w \in L$ with $|w| \geq m$.
- Need a decomposition as $w = xyz$ with $|xy| \leq m$ and $|y| \geq 1$.
- Then, we need to choose an $i$ such that the pumped string $w_i = xyz^i$ is not in $L$. 
Example 7

$L = \{a^n b^n : n \geq 0\}$ is not regular.

**Solution:**
Suppose $L$ is regular. Let $m = n$. Consider

$$w = a^n b^n = xyz$$

with $|xy| \leq m = n$ and $|y| \geq 1$. Then,

$1 \leq |y| = k < n$. So, $y = a^k$, $n > k \geq 1$. However, when $i = 0$,

$$w_0 = a^{m-k} b^m \notin L.$$
Example 8

Show that $L = \{ww^R : w \in \{a, b\}^*\}$ is not regular.
Example 9

Let $\Sigma = \{a, b\}$. The language

$$L = \{w \in \Sigma^* : n_a(w) < n_b(w)\}$$

is not regular.
Example 10

The language $L = \{(ab)^n a^k : n > k, k \geq 0\}$ is not regular.
Example 11

Show that $L = \{a^n : n \text{ is a perfect square}\}$ is not regular.
Example 12

Show that the language

\[ L = \{ a^n b^k c^{n+k} : n \geq 0, k \geq 0 \} \] is not regular.
Example 13

Show that $L = \{a^n b^k : n \neq k\}$ is not regular.
Pitfalls for Pumping Lemma

- using the pumping lemma to show that a language is regular
- to start with a string not in $L$
- making some assumptions about the decomposition $xyz$