CS 4410

Automata, Computability, and Formal Language

Dr. Xuejun Liang

Chapter 4
Properties of Regular Languages

1. Closure Properties of Regular Languages
   - Closure under Simple Set Operations
   - Closure under Other Operations
2. Elementary Questions about Regular Languages
3. Identifying Nonregular Languages
   - Using the Pigeonhole Principle
   - A pumping Lemma

Identifying Nonregular Languages

Example 4.6 (Using the pigeonhole principle):

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

Using a proof by contradiction

Suppose L is regular. Then some dfa $M = (Q, \{a, b\}, \delta, q_0, F)$ exists for it.

Now look at $\delta^*(q_0, a^i)$ for i = 1, 2, 3,….

There must be some state, say q, such that

$$\delta^*(q_0, a^m) = \delta^*(q_0, a^n) = q \text{ with } m \neq n$$

$$\Rightarrow \delta^*(q_0, a^m b^n) = \delta^*(\delta^*(q_0, a^m), b^n) = \delta^*(\delta^*(q_0, a^n), b^n) = \delta^*(q_0, a^n b^n)$$

$\Rightarrow$ $a^m b^n \in L$ $\Rightarrow$ A contradiction $\Rightarrow$ L is not regular

If we put n objects into m boxes (pigeonholes), and if n > m, then at least one box must have more than one item in it.

A pumping Lemma

Theorem 4.8 (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

$$w = xyz \text{ with } |xy| \leq m, \text{ and } |y| \geq 1,$$

such that

$$w_i = xy^i z \in L, \ i = 0, 1, 2, \ldots$$

A pumping Lemma

Theorem 4.8 (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

$$w = xyz \text{ with } |xy| \leq m, \text{ and } |y| \geq 1,$$

such that

$$w_i = xy^i z \in L, \ i = 0, 1, 2, \ldots$$

$\exists m \forall w \exists xyz \forall i (w_i \in L)$

$m > 0$
$w \in L$
$|w| \geq m$
$w = xyz$
$|xy| \leq m$
$|y| \geq 1$
$w_i = xy^i z$

A pumping Lemma

Theorem 4.8 (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

$$w = xyz \text{ with } |xy| \leq m, \text{ and } |y| \geq 1,$$

such that

$$w_i = xy^i z \in L,\ i = 0, 1, 2, \ldots$$

$$\exists m \forall w \exists xyz \forall i (w_i \in L)$$

$$m > 0$$
$$w \in L$$
$$|w| \geq m$$
$$w = xyz$$
$$|xy| \leq m$$
$$|y| \geq 1$$
$$w_i = xy^i z$$

Idea of using pumping Lemma

Show an infinite regular language L is not regular:
If L violates the pumping lemma, then L is not regular.

L satisfies the pumping lemma

$$\exists m \forall w \exists xyz \forall i (w_i \in L)$$

$$\neg(\exists m \forall w \exists xyz \forall i (w_i \in L)) \Leftrightarrow$$
$$\forall m \exists w \forall xyz \exists i (w_i \notin L)$$

L violates the pumping lemma

$$\forall m \exists w \forall xyz \exists i (w_i \notin L)$$

$$m > 0$$
$$w \in L$$
$$|w| \geq m$$
$$w = xyz$$
$$|xy| \leq m$$
$$|y| \geq 1$$
$$w_i = xy^i z$$

Example 4.7 Using the pumping lemma to show that $L=\{a^n b^n: n \geq 0\}$ is not regular?

$$\forall m \exists w \forall xyz \exists i (w_i \notin L)$$

$$m > 0$$
$$w \in L$$
$$|w| \geq m$$
$$w = xyz$$
$$|xy| \leq m$$
$$|y| \geq 1$$
$$w_i = xy^i z$$

Applying the pumping lemma

$$\forall m \exists w \forall xyz \exists i (w_i \notin L)$$

The correct argument can be visualized as a game we play against an opponent

1. The opponent picks m.
2. Given m, we pick a string w in L of length equal or greater than m.
3. The opponent chooses the decomposition $\mathrm{w} = xyz$, subject to $|xy| \leq m$, $|y| \geq 1$, in a way that makes it hard to establish a contradiction.
4. We try to pick i in such a way that the pumped string $w_i = xy^iz$ is not in L. If we can do so, we win the game.

Example 4.8: Let Σ={a, b}. Show that $L=\{ww^R : w \in \Sigma^*\}$ is not regular.

$$\forall m \exists w \forall xyz \exists i (w_i \notin L)$$

$$m > 0$$
$$w \in L$$
$$|w| \geq m$$
$$w = xyz$$
$$|xy| \leq m$$
$$|y| \geq 1$$
$$w_i = xy^i z$$

Example 4.9: Let $\Sigma=\{a, b\}$. $L=\{w\in\Sigma^*: n_a(w)<n_b(w)\}$ is not regular.

$$\forall m \exists w \forall xyz \exists i (w_i \notin L)$$

$$m > 0$$
$$w \in L$$
$$|w| \geq m$$
$$w = xyz$$
$$|xy| \leq m$$
$$|y| \geq 1$$
$$w_i = xy^iz$$

Example 4.10: $L=\{(ab)^n a^k : n>k, k\geq 0\}$ is not regular.

$$\forall m \exists w \forall xyz \exists i (w_i \notin L)$$

$$m > 0$$
$$w \in L$$
$$|w| \geq m$$
$$w = xyz$$
$$|xy| \leq m$$
$$|y| \geq 1$$
$$w_i = xy^i z$$

Example 4.11: $L=\{a^n : n \text{ is a perfect square}\}$ is not regular.

$$\forall m \exists w \forall xyz \exists i (w_i \notin L)$$

$$m > 0$$
$$w \in L$$
$$|w| \geq m$$
$$w = xyz$$
$$|xy| \leq m$$
$$|y| \geq 1$$
$$w_i = xy^i z$$

Example 4.12: $L=\{a^n b^k c^{n+k} : n \geq 0, k \geq 0\}$ is not regular.

Example 4.13: $L=\{a^n b^k : n \neq k\}$ is not regular.

Applying the pumping lemma

- The pumping lemma says there exist an m as well as the decomposition xyz. But, we do not know what they are.
  - We cannot claim that we have reached a contradiction just because the pumping lemma is violated for some specific values of m or xyz.
- On the other hand, the pumping lemma holds for every $w \in L$ and every and every i.
  - Therefore, if the pumping lemma is violated even for one w or i, then the language cannot be regular.

Some Common Pitfalls

- One mistake is to try using the pumping lemma to show that a language is regular. Even if you can show that no string in a language L can ever be pumped out, you cannot conclude that L is regular. The pumping lemma can only be used to prove that a language is not regular.
- Another mistake is to start (usually inadvertently) with a string not in L.
- Finally, perhaps the most common mistake is to make some assumptions about the decomposition $w = xyz$. The only thing we know is that y is not empty and that $|xy| \leq m$;