COSE215: Theory of Computation

Lecture 8 — Properties of Regular Languages (2): Pumping Lemma

Hakjoo Oh
2019 Spring

Recap

The following are all equivalent:

- L is a regular language.
- There is a DFA D whose language is L.
- There is an NFA N whose language is L.
- There is a regular expression R whose language is L.

Some Fundamental Questions

- Are all languages regular?
  - No, e.g., $L = \{a^n b^n \mid n \geq 0\}$ is not regular.
- How to prove that a language is non-regular? Two methods:
  1. Direct proof by Pigeonhole principle.
  2. By using the pumping lemma.

Intuition

- Regular languages can be recognized with finite memory.
- Non-regular languages cannot be recognized with finite memory.

Example 1: $L = \{a^n b^n \mid n \geq 0\}$

The intuition behind the proof:

- Suppose there is a DFA D that accepts L.
- Then, when D is run on any two of the strings $\epsilon, a, aa, aaa, \ldots$, D must end in different states.
  - Assume a^n and a^m $(n \neq m)$ lead to the same state.
  - Then $a^n b^n$ and $a^m b^n$ must end up in the same state.
  - This is a contradiction because either $a^n b^n$ is rejected or $a^m b^n$ is accepted.
- This is impossible because there are only finitely many states. We cannot put all these strings into different states.

Example 1: $L = \{a^n b^n \mid n \geq 0\}$

Proof with Pigeonhole principle (If you put more than n pigeons into n holes, then some hole has more than one pigeon.):

- Proof by contradiction.
- Assume L is regular.
- Then there is a DFA $M = (Q, \Sigma, \delta, q_0, F)$ recognizing L.
- Define:
  - Pigeons $= \{a^n \mid n \geq 0\} = \{a, aa, aaa, \ldots\}$
  - Holes $=$ states in Q
- Put pigeon a^n into hole $\delta^*(q_0, a^n)$
  - i.e., the hole corresponding to the state reached by input a^n
- We have $|Q|$ holes but more than $|Q|$ pigeons (actually, infinitely many).
- So, two pigeons must be put in the same hole, say a^i and a^j, where $i \neq j$.
  - That is, a^i and a^j lead to the same state.
- Then, since M accepts $a^i b^i$, it also accepts $a^j b^i$, which is a contradiction.
- Thus, the original assumption that L is regular is false,
- That is, L is non-regular.

Example 2: $L = \{ww \mid w \in \{0,1\}^*\}$ is non-regular

- Show by contradiction, using Pigeonhole principle.
- Assume L is regular, so there is a DFA $M = (Q, \Sigma, \delta, q_0, F)$ recognizing L.
- Define:
  - Pigeons $= \{0^i 1 \mid i \geq 0\} = \{1, 01, 001, \ldots\}$
  - Holes $=$ states in Q
- Put pigeon string $0^i 1$ into hole $\delta^*(q_0, 0^i 1)$
- By Pigeonhole principle, two pigeons share a hole, say $0^i 1$ and $0^j 1$, where $i \neq j$.
- So $0^i 1$ and $0^j 1$ lead to the same state.
- M accepts $0^i 1 0^i 1$, so does $0^j 1 0^i 1$, which is a contradiction.

The Pumping Lemma

Theorem (Pumping Lemma)

For any regular language L there exists an integer n, such that for all $x \in L$ with $|x| \geq n$, there exist $u, v, w \in \Sigma^*$, such that

1. $x = uvw$
2. $|uv| \leq n$
3. $|v| \geq 1$
4. *for all* $i \geq 0$, $uv^iw \in L$.

Proof of Pumping Lemma

- Let M be a DFA for L. Suppose M has n states.
- Take $x \in L$ with $|x| \geq n$, let $m = |x|$:

$$x = a_1 a_2 \ldots a_m$$

- Let $p_i = \delta^*(q_0, a_1 a_2 \ldots a_i)$. Note $p_0 = q_0$ and p_m is a final state.
- Consider the first $n + 1$ states: $p_0 p_1 \ldots p_n$.
- By Pigeonhole principle, two p_i and p_j with $0 \leq i < j \leq n$ share a state, i.e., $p_i = p_j$.
- Break $x = uvw$:
  - $u = a_1 a_2 \ldots a_i$
  - $v = a_{i+1} a_{i+2} \ldots a_j$
  - $w = a_{j+1} a_{j+2} \ldots a_m$
- Note that $\delta^*(p_0, u) = p_i$, $\delta^*(p_i, v) = p_i$, and $\delta^*(p_i, w) = p_m$.
- Thus, $\delta^*(p_0, uw) = p_m$, $\delta^*(p_0, uvw) = p_m$, $\delta^*(p_0, uv^2w) = p_m$, and so on.

Using Pumping Lemma to show non-regularity

- If L is regular, L satisfies pumping lemma?
- If L satisfies pumping lemma, L is regular?
- If L does not satisfy pumping lemma, then L is non-regular?

Pumping lemma can be used only for proving languages not to be regular.

Example 1

Prove that $L = \{0^i 1^i \mid i \geq 0\}$ is not regular.

- Show that pumping lemma (P.L.) does not hold.
- If L is regular, then by P.L. there exists n such that ...
- Now let $x = 0^n 1^n$
- $x \in L$ and $|x| \geq n$, so by P.L. there exist u, v, w such that (1)–(4) hold.
- We show that for all u, v, w (1)–(4) do not all hold.
- If (1), (2), (3) hold then $x = 0^n 1^n = uvw$ with $|uv| \leq n$ and $|v| \geq 1$.
- So, $u = 0^s, v = 0^t, w = 0^p 1^n$ with

$$s + t \leq n, \quad t \geq 1, \quad p \geq 0, \quad s + t + p = n.$$

- Then (4) fails for $i = 0$:

$$uv^0 w = uw = 0^s 0^p 1^n = 0^{s+p} 1^n \notin L, \quad \text{since } s + p \neq n$$

Example 2

Prove that $L = \{ww^R \mid w \in \{a,b\}^*\}$ is not regular.

- Show that pumping lemma (P.L.) does not hold.
- If L is regular, then by P.L. there exists n such that ...
- Now let $x = a^n b^n b^n a^n$
- $x \in L$ and $|x| \geq n$, so by P.L. there exist u, v, w such that (1)–(4) hold.
- We show that for all u, v, w (1)–(4) do not all hold.
- If (1), (2), (3) hold then $x = a^n b^n b^n a^n = uvw$ with $|uv| \leq n$ and $|v| \geq 1$.
- So, $u = a^s, v = a^t, w = a^p b^n b^n a^n$ with

$$s + t \leq n, \quad t \geq 1, \quad p \geq 0, \quad s + t + p = n.$$

- Then (4) fails for $i = 0$:

$$uv^0 w = uw = a^s a^p b^n b^n a^n = a^{s+p} b^n b^n a^n \notin L, \text{since } s + p \neq n$$

Example 3

Prove that $L = \{w \in \{a, b\}^* \mid n_a(w) < n_b(w)\}$ is not regular.

- Show that pumping lemma (P.L.) does not hold.
- If L is regular, then by P.L. there exists n such that ...
- Now let $x = a^n b^{n+1}$
- $x \in L$ and $|x| \geq n$, so by P.L. there exist u, v, w such that (1)–(4) hold.
- We show that for all u, v, w (1)–(4) do not all hold.
- If (1), (2), (3) hold then $x = a^n b^{n+1} = uvw$ with $|uv| \leq n$ and $|v| \geq 1$.
- So, $u = a^s, v = a^t, w = a^p b^{n+1}$ with

$$s + t \leq n, \quad t \geq 1, \quad p \geq 0, \quad s + t + p = n.$$

- Then (4) fails for $i = 2$:

$$uv^2w = a^s a^{2t} a^p b^{n+1} = a^{s+2t+p} b^{n+1} \notin L,$$

since $s + 2t + p \geq n + 1$.

Example 4

Prove that $L = \{a^n \mid n \text{ is a perfect square}\}$ is not regular.

- Show that pumping lemma (P.L.) does not hold.
- If L is regular, then by P.L. there exists n such that ...
- Now let $x = a^{n^2}$
- $x \in L$ and $|x| \geq n$, so by P.L. there exist u, v, w such that (1)–(4) hold.
- We show that for all u, v, w (1)–(4) do not all hold.
- If (1), (2), (3) hold then $x = a^{n^2} = uvw$ with $|uv| \leq n$ and $|v| \geq 1$.
- Then, clearly $v = a^k$ with $1 \leq k \leq n$.
- Then (4) fails for $i = 0$:

$$uv^0w = a^{n^2-k} \notin L, \quad \text{since } n^2 - k > (n-1)^2.$$

Example 5

Prove that $L = \{a^n b^k c^{n+k} \mid n \geq 0 \wedge k \geq 0\}$ is not regular.

- It is not difficult to apply the pumping lemma directly, but it is even easier to use closure under homomorphism. Take

$$h(a) = a, \quad h(b) = a, \quad h(c) = c,$$

  then

$$h(L) = \{a^{n+k} c^{n+k} \mid n + k \geq 0\} = \{a^i b^i \mid i \geq 0\}.$$

  We know this language is not regular.

- Also, we know that if a language L_1 is regular, then $h(L_1)$ is regular. Taking its contraposition, we conclude that L is not regular.

cf) The converse of pumping lemma is not true

$$L = \{c^m a^n b^n \mid m \geq 1, n \geq 1\}$$

- L satisfies the pumping lemma.
  - For any $x \in L$ of length ≥ 1, we can take $u = \epsilon$, $v =$ the first letter of x (c), and $w =$ the rest of x.
- However, L is not regular.
  - We can prove this using a general version of pumping lemma: For any regular language L, there exists $n \geq 1$ such that for every string $uvw \in L$ with $|w| \geq p$ such that
    - ⋆ $uwv = uxyzv$
    - ⋆ $|xy| \leq n$
    - ⋆ $|y| \geq 1$
    - ⋆ For all $i \geq 0$, $uxy^izv \in L$.
- Still, the converse of the general lemma is not true.
  - Languages that satisfy the lemma can still be non-regular.
  - For a necessary and sufficient condition to be regular, refer to Myhill-Nerode theorem.