

Lecture 9 – The Pumping Lemma for Regular Languages

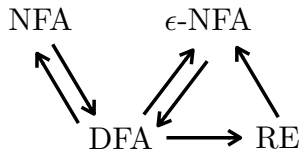
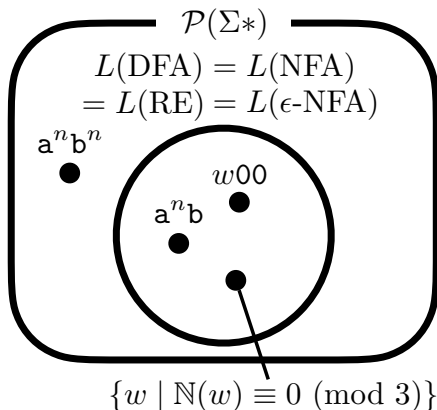
COSE215: Theory of Computation

Jihyeok Park



2023 Spring

- Not all languages are regular: e.g., $L = \{a^n b^n \mid n \geq 0\}$.



- How to prove that a language is **NOT** regular? **Pumping Lemma!**

1. Pumping Lemma for Regular Languages

Pumping Lemma

Proof of Pumping Lemma

Application: Proving Languages are Not Regular

Examples

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

Example 2: $L = \{ww^R \mid w \in \{a, b\}^*\}$

Example 3: $L = \{a^l b^m c^n \mid l + m \leq n\}$

Example 4: $L = \{a^{n^2} \mid n \geq 0\}$

Example 5: $L = \{a^n b^k c^{n+k} \mid n, k \geq 0\}$

Lemma (Pumping Lemma for Regular Languages)

For a given regular language L , **there exists** a *positive integer* n such that **for all** $w \in L$, if $|w| \geq n$, **there exists** $w = xyz$ such that

- ① $|y| > 0$
- ② $|xy| \leq n$
- ③ $\forall i \geq 0. xy^iz \in L$

$A = \quad L \text{ is regular}$



$B = \exists n > 0. \forall w \in L. |w| \geq n \Rightarrow \exists w = xyz. \textcircled{1} \wedge \textcircled{2} \wedge \textcircled{3}$

- Let L be a regular language.
- Then, \exists DFA $D = (Q, \Sigma, \delta, q_0, F)$. s.t. $L(D) = L$. Let $n = |Q| > 0$.
- Take any $w = a_1 a_2 \cdots a_m \in L$ s.t. $|w| = m \geq n$.
- Let $p_i = \delta^*(q_0, a_1 \cdots a_i)$ for all $0 \leq i \leq m$. Then, $p_0 = q_0 \wedge p_m \in F$.
- Consider the first $n + 1$ states: $p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} p_n$.
- By Pigeonhole Principle, there exists $i < j \leq n$ s.t. $p_i = p_j$.
- Split $w = xyz$:

$$\begin{array}{lll} x = a_1 \cdots a_i & y = a_{i+1} \cdots a_j & z = a_{j+1} \cdots a_m \\ |x| = i & |y| = j - i > 0 & |xy| = j \leq n \\ \delta^*(q_0, x) = p_i & \delta^*(p_i, y) = p_j & \delta^*(p_i, z) = p_m \end{array}$$

- Then, $\forall i \geq 0$. $\delta^*(q_0, xy^i z) = p_m$ (by induction on i).
- Finally, $\forall i \geq 0$. $xy^i z \in L$.

- Let L be a regular language.
- Then, \exists DFA $D = (Q, \Sigma, \delta, q_0, F)$. s.t. $L(D) = L$. Let $n = |Q| > 0$.
- Take any $w = a_1 a_2 \cdots a_m \in L$ s.t. $|w| = m \geq n$.
- Let $p_i = \delta^*(q_0, a_1 \cdots a_i)$ for all $0 \leq i \leq m$. Then, $p_0 = q_0 \wedge p_m \in F$.
- Consider the first $n + 1$ states: $p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} p_n$.
- By Pigeonhole Principle, there exists $i < j \leq n$ s.t. $p_i = p_j$.
- Split $w = xyz$:

$$\begin{array}{lll}
 x = a_1 \cdots a_i & y = a_{i+1} \cdots a_j & z = a_{j+1} \cdots a_m \\
 |x| = i & \textcircled{1} - |y| = j - i > 0 & |xy| = j \leq n - \textcircled{2} \\
 \delta^*(q_0, x) = p_i & \delta^*(p_i, y) = p_j & \delta^*(p_i, z) = p_m
 \end{array}$$

- Then, $\forall i \geq 0$. $\delta^*(q_0, xy^i z) = p_m$ (by induction on i).
- Finally, $\forall i \geq 0$. $xy^i z \in L - \textcircled{3}$.

Lemma (Pumping Lemma for Regular Languages)

For a given regular language L , **there exists** a *positive integer* n such that **for all** $w \in L$, if $|w| \geq n$, **there exists** $w = xyz$ such that

- ① $|y| > 0$
- ② $|xy| \leq n$
- ③ $\forall i \geq 0. xy^iz \in L$

$A =$

L is regular



$B = \exists n > 0. \forall w \in L. |w| \geq n \Rightarrow \exists w = xyz. \textcircled{1} \wedge \textcircled{2} \wedge \textcircled{3}$

Lemma (Pumping Lemma for Regular Languages)

$$A = L \text{ is regular}$$



$$B = \exists n > 0. \forall w \in L. |w| \geq n \Rightarrow \exists w = xyz. \textcircled{1} \wedge \textcircled{2} \wedge \textcircled{3}$$

$$A \Rightarrow B \quad (0)$$

$$B \Rightarrow A \quad (X)$$

$$\neg B \Rightarrow \neg A \quad (0)$$

$$\begin{aligned} \neg B &= \forall n > 0. \neg(\forall w \in L. |w| \geq n \Rightarrow \exists w = xyz. \textcircled{1} \wedge \textcircled{2} \wedge \textcircled{3}) \\ &= \forall n > 0. \exists w \in L. \neg(|w| \geq n \Rightarrow \exists w = xyz. \textcircled{1} \wedge \textcircled{2} \wedge \textcircled{3}) \\ &= \forall n > 0. \exists w \in L. |w| \geq n \wedge \neg(\exists w = xyz. \textcircled{1} \wedge \textcircled{2} \wedge \textcircled{3}) \\ &= \forall n > 0. \exists w \in L. |w| \geq n \wedge \forall w = xyz. \neg(\textcircled{1} \wedge \textcircled{2} \wedge \textcircled{3}) \\ &= \forall n > 0. \exists w \in L. |w| \geq n \wedge \forall w = xyz. \neg(\textcircled{1} \wedge \textcircled{2}) \vee \neg\textcircled{3} \\ &= \forall n > 0. \exists w \in L. |w| \geq n \wedge \forall w = xyz. (\textcircled{1} \wedge \textcircled{2}) \Rightarrow \neg\textcircled{3} \end{aligned}$$

To prove a language L is **NOT** regular, we need to show that

$$\forall n > 0. \exists w \in L. |w| \geq n \wedge \forall w = xyz. (\textcircled{1} \wedge \textcircled{2}) \Rightarrow \neg \textcircled{3}$$

$$\textcircled{1} |y| > 0$$

$$\textcircled{2} |xy| \leq n$$

$$\textcircled{3} \forall i \geq 0. xy^i z \in L$$

Note that $\neg \textcircled{3} = \exists i \geq 0. xy^i z \notin L$.

We can prove this by following the steps below:

① Assume **any** positive integer n is given.

② **Pick** a word $w \in L$.

③ Show that $|w| \geq n$.

④ Assume **any** split $w = xyz$ is given, and $\textcircled{1} |y| > 0 \wedge \textcircled{2} |xy| \leq n$.

⑤ $\neg \textcircled{3}$ Pick $i \geq 0$, and show that $xy^i z \notin L$ using $\textcircled{1}$ and $\textcircled{2}$.

Let's prove that L is **NOT** regular using the Pumping Lemma:

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

- ① Assume **any** positive integer n is given.
- ② Let $w = a^n b^n \in L$.
- ③ $|w| = n + n = 2n \geq n$.
- ④ Assume **any** split $w = xyz$ is given, and ① $|y| > 0 \wedge$ ② $|xy| \leq n$.
- ⑤ Then, \neg ③ $xy^0z \notin L$ because:
 - Since ② $|xy| \leq n$,

$$x = a^p \quad y = a^q \quad z = a^r b^n$$

where $0 \leq p, q, r \leq n$ and $p + q + r = n$.

- Since ① $|y| > 0$, we know $q > 0$.
- Finally, $xy^0z = xz = a^p a^r b^n = a^{n-q} b^n$ ($\because p + q + r = n$).
But, $a^{n-q} b^n \notin L$ ($\because q > 0$).



Example 2

Let's prove that L is **NOT** regular using the Pumping Lemma:

$$L = \{ww^R \mid w \in \{a,b\}^*\}$$

- ① Assume **any** positive integer n is given.
- ② Let $w = a^n b^n b^n a^n \in L$.
- ③ $|w| = n + n + n + n = 4n \geq n$.
- ④ Assume **any** split $w = xyz$ is given, and ① $|y| > 0 \wedge$ ② $|xy| \leq n$.
- ⑤ Then, \neg ③ $xy^0z \notin L$ because:
 - Since ② $|xy| \leq n$,

$$x = a^p \quad y = a^q \quad z = a^r b^n b^n a^n$$

where $0 \leq p, q, r \leq n$ and $p + q + r = n$.

- Since ① $|y| > 0$, we know $q > 0$.
- Finally, $xy^0z = xz = a^p a^r b^n b^n a^n = a^{n-q} b^n b^n a^n (\because p + q + r = n)$.
But, $a^{n-q} b^n b^n a^n \notin L (\because q > 0)$. □

Let's prove that L is **NOT** regular using the Pumping Lemma:

$$L = \{a^l b^m c^n \mid l + m \leq n\}$$

- ① Assume **any** positive integer n is given.
- ② Let $w = a^n b^n c^{2n} \in L$.
- ③ $|w| = n + n + 2n = 4n \geq n$.
- ④ Assume **any** split $w = xyz$ is given, and ① $|y| > 0 \wedge$ ② $|xy| \leq n$.
- ⑤ Then, \neg ③ $xy^0z \notin L$ because:
 - Since ② $|xy| \leq n$,

$$x = a^p \quad y = a^q \quad z = a^r b^n c^{2n}$$

where $0 \leq p, q, r \leq n$ and $p + q + r = n$.

- Since ① $|y| > 0$, we know $q > 0$.
- Finally, $xy^0z = xz = a^p a^r b^n c^{2n} = a^{n-q} b^n c^{2n}$ ($\because p + q + r = n$).
But, $a^{n-q} b^n c^{2n} \notin L$ ($\because q > 0$).



Let's prove that L is **NOT** regular using the Pumping Lemma:

$$L = \{a^{n^2} \mid n \geq 0\}$$

- ① Assume **any** positive integer n is given.
- ② Let $w = a^{n^2} \in L$.
- ③ $|w| = n^2 \geq n$.
- ④ Assume **any** split $w = xyz$ is given, and ① $|y| > 0 \wedge$ ② $|xy| \leq n$.
- ⑤ Then, \neg ③ $xy^2z \notin L$ because:
 - Since ① $|y| > 0$ and ② $|xy| \leq n$,

$$y = a^k$$

where $1 \leq k \leq n$. Then,

$$n^2 < n^2 + k \quad (\because 1 \leq k) \qquad n^2 + k < (n+1)^2 \quad (\because k \leq n)$$

- Finally, $xy^2z = xyyz = a^{n^2+k} \notin L \quad (\because n^2 < n^2 + k < (n+1)^2)$. □

Let's prove that L is **NOT** regular:

$$L = \{a^n b^k c^{n+k} \mid n, k \geq 0\}$$

- It is much easier to use **closure properties** under **homomorphisms**.
- Consider a homomorphism $h : \{a, b, c\} \rightarrow \{a, b\}^*$:

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

- Then,

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

- If L is regular, then $h(L)$ must be regular as well.
- However, we know $h(L)$ is **NOT** regular.
- Therefore, L is **NOT** regular. □

1. Pumping Lemma for Regular Languages

Pumping Lemma

Proof of Pumping Lemma

Application: Proving Languages are Not Regular

Examples

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

Example 2: $L = \{ww^R \mid w \in \{a, b\}^*\}$

Example 3: $L = \{a^l b^m c^n \mid l + m \leq n\}$

Example 4: $L = \{a^{n^2} \mid n \geq 0\}$

Example 5: $L = \{a^n b^k c^{n+k} \mid n, k \geq 0\}$

- Please see <https://github.com/ku-plrg-classroom/docs/tree/main/equiv-re-fa>.
- The due date is Apr. 13 (Thu.).
- Please only submit `Implementation.scala` file to **Blackboard**.

- Equivalence and Minimization of Finite Automata

Jihyeok Park

jihyeok_park@korea.ac.kr

<https://plrg.korea.ac.kr>