COL 352 Introduction to Automata and Theory of Computation

Nikhil Balaji

Bharti 420
Indian Institute of Technology, Delhi
nbalaji@cse.iitd.ac.in

February 1, 2023

Lecture 9: Non-regularity: Pumping Lemma

Limitations of Finite Automata

$$L_{0,1} := \{0^n 1^n \mid n \geq 0\}$$

$$PAL := \{ww^R \mid w \in \Sigma^*\}$$

Generalise?

These arguments seem to be example specific.. can they be generalised?

Pumping Lemma

From the previous argument, we filter out a property of regular languages .

Pumping Lemma

Lemma

Let L be a regular language. Then

- there exists a constant n such that
- for every word $w \in L$ such that $|w| \geq n$

Pumping Lemma

Lemma

Let L be a regular language. Then

- there exists a constant n such that
- for every word $w \in L$ such that $|w| \geq n$
- we can break w in to three parts $w = xyz$, such that

Pumping Lemma

Lemma

Let L be a regular language. Then

- there exists a constant n such that
- for every word $w \in L$ such that $|w| \geq n$
- we can break w in to three parts $w = xyz$, such that
  1. $y \neq \epsilon$
  2. $|xy| \leq n$ and
  3. For all $k \geq 0$, $xy^k z \in L$.

Pumping Lemma

Lemma

Let L be a regular language. Then

- there exists a constant n such that
- for every word $w \in L$ such that $|w| \geq n$
- we can break w in to three parts $w = xyz$, such that
  1. $y \neq \epsilon$
  2. $|xy| \leq n$ and
  3. For all $k \geq 0$, $xy^k z \in L$.
- Lets write it using symbols

Pumping Lemma

Lemma

Let L be a regular language. Then

- there exists a constant n such that
- for every word $w \in L$ such that $|w| \geq n$
- we can break w in to three parts $w = xyz$, such that
  1. $y \neq \epsilon$
  2. $|xy| \leq n$ and
  3. For all $k \geq 0$, $xy^k z \in L$.
- Lets write it using symbols

$$\exists n \forall w \in L \, (|w| \geq n \implies \exists xyz.(xyz = w \wedge y \neq \epsilon \wedge |xy| \leq n \wedge (\forall k \ \ xy^k z \in L))$$

Understanding the lemma

It is called pumping lemma because we can "pump" y any number of times including 0.

Understanding the lemma

It is called pumping lemma because we can "pump" y any number of times including 0.

- How many quantifier alternations?

Understanding the lemma

It is called pumping lemma because we can "pump" y any number of times including 0.

- How many quantifier alternations?
- Exercise: Prove it holds for finite languages.

Understanding the lemma

It is called pumping lemma because we can "pump" y any number of times including 0.

- How many quantifier alternations?
- Exercise: Prove it holds for finite languages.
- What happens if $y = \epsilon$?

Proof of Pumping Lemma

- If L is regular, then there exists DFA A s.t. $L = L(A)$.

Proof of Pumping Lemma

- If L is regular, then there exists DFA A s.t. $L = L(A)$.
- Suppose A has n states.

Proof of Pumping Lemma

- If L is regular, then there exists DFA A s.t. $L = L(A)$.
- Suppose A has n states.
- Consider word $w = w_1 \dots w_m$, where $m >> n$.

Proof of Pumping Lemma

- If L is regular, then there exists DFA A s.t. $L = L(A)$.
- Suppose A has n states.
- Consider word $w = w_1 \dots w_m$, where $m >> n$.
- $q_s = q_0 \xrightarrow{w_1} q_1 \xrightarrow{w_2} \dots \xrightarrow{w_m} q_m = q_f$ is a run of A.

Proof of Pumping Lemma

- If L is regular, then there exists DFA A s.t. $L = L(A)$.
- Suppose A has n states.
- Consider word $w = w_1 \dots w_m$, where $m >> n$.
- $q_s = q_0 \xrightarrow{w_1} q_1 \xrightarrow{w_2} \dots \xrightarrow{w_m} q_m = q_f$ is a run of A.
- By PHP, there exists $i \neq j \in [m]$ s.t. $q_i = q_j = q'$

Proof of Pumping Lemma

- If L is regular, then there exists DFA A s.t. $L = L(A)$.
- Suppose A has n states.
- Consider word $w = w_1 \dots w_m$, where $m >> n$.
- $q_s = q_0 \xrightarrow{w_1} q_1 \xrightarrow{w_2} \dots \xrightarrow{w_m} q_m = q_f$ is a run of A.
- By PHP, there exists $i \neq j \in [m]$ s.t. $q_i = q_j = q'$
- Break $w = xyz$

Proof of Pumping Lemma

- If L is regular, then there exists DFA A s.t. $L = L(A)$.
- Suppose A has n states.
- Consider word $w = w_1 \dots w_m$, where $m >> n$.
- $q_s = q_0 \xrightarrow{w_1} q_1 \xrightarrow{w_2} \dots \xrightarrow{w_m} q_m = q_f$ is a run of A.
- By PHP, there exists $i \neq j \in [m]$ s.t. $q_i = q_j = q'$
- Break $w = xyz$
-
    - $x = w_1 \dots w_i$
    - $y = w_{i+1} \dots w_j$
    - $z = w_{j+1} \dots w_m$
- By construction, $y \neq \epsilon$, $|xz| \leq n$

Proof of Pumping Lemma

- If L is regular, then there exists DFA A s.t. $L = L(A)$.
- Suppose A has n states.
- Consider word $w = w_1 \dots w_m$, where $m >> n$.
- $q_s = q_0 \xrightarrow{w_1} q_1 \xrightarrow{w_2} \dots \xrightarrow{w_m} q_m = q_f$ is a run of A.
- By PHP, there exists $i \neq j \in [m]$ s.t. $q_i = q_j = q'$
- Break $w = xyz$
-
    - $x = w_1 \dots w_i$
    - $y = w_{i+1} \dots w_j$
    - $z = w_{j+1} \dots w_m$
- By construction, $y \neq \epsilon$, $|xz| \leq n$
- For any $k > 0$, xy^kz takes A to q_f, hence accepting.

Proof of Pumping Lemma

- If L is regular, then there exists DFA A s.t. $L = L(A)$.
- Suppose A has n states.
- Consider word $w = w_1 \dots w_m$, where $m >> n$.
- $q_s = q_0 \xrightarrow{w_1} q_1 \xrightarrow{w_2} \dots \xrightarrow{w_m} q_m = q_f$ is a run of A.
- By PHP, there exists $i \neq j \in [m]$ s.t. $q_i = q_j = q'$
- Break $w = xyz$
-
    - $x = w_1 \dots w_i$
    - $y = w_{i+1} \dots w_j$
    - $z = w_{j+1} \dots w_m$
- By construction, $y \neq \epsilon$, $|xz| \leq n$
- For any $k > 0$, $xy^k z$ takes A to q_f, hence accepting.
- Therefore, $\forall k > 0$, $xy^k z \in L$

Pumping Lemma

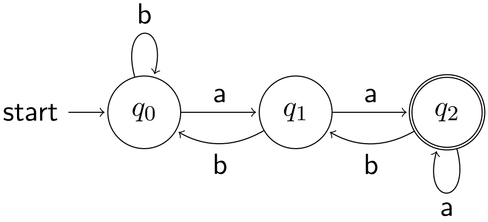
Lemma

Let L be a regular language. Then

- there exists a constant n such that
- for every word $w \in L$ such that $|w| \geq n$
- we can break w in to three parts $w = xyz$, such that
  1. $y \neq \epsilon$
  2. $|xy| \leq n$ and
  3. For all $k \geq 0$, $xy^k z \in L$.

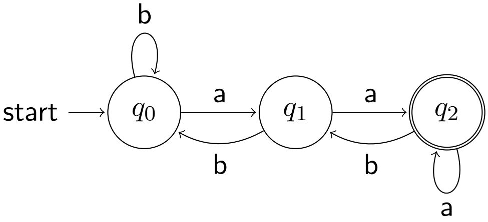
Example

Consider DFA A below.



Example

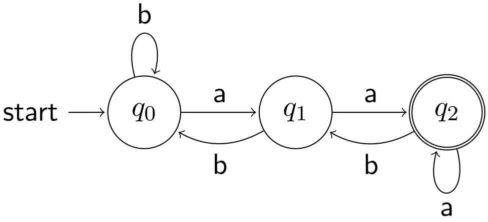
Consider DFA A below.



The size of DFA is $n = 3$.

Example

Consider DFA A below.

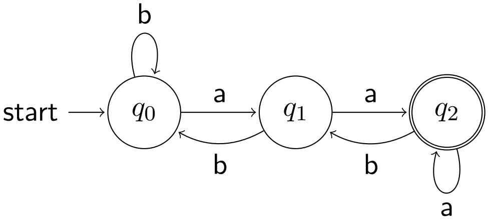


The size of DFA is $n = 3$.

Choose $w = abaaba$. $|w| = 6 \geq 3$ and w is in $L(A)$.

Example

Consider DFA A below.



The size of DFA is $n = 3$.

Choose $w = abaaba$. $|w| = 6 \geq 3$ and w is in $L(A)$.

The run is $q_0 \xrightarrow{a} q_1 \xrightarrow{b} q_0 \xrightarrow{a} q_1 \xrightarrow{a} q_2 \xrightarrow{b} q_1 \xrightarrow{a} q_2$

q_1 is repeated in the first 4 states of the run.

Choose $x = a$, $y = ba$, $z = aba$

Therefore, $a(ba)^k aba \in L(A)$

Using the Pumping Lemma

- Pumping lemma gives a useful property of regular languages. How do we use it?

Using the Pumping Lemma

- Pumping lemma gives a useful property of regular languages. How do we use it?
- What does the contrapositive of the lemma say?

Using the Pumping Lemma

- Pumping lemma gives a useful property of regular languages. How do we use it?
- What does the contrapositive of the lemma say?
- $A \implies B$ iff $\neg B \implies \neg A$.

Using the Pumping Lemma

- Pumping lemma gives a useful property of regular languages. How do we use it?
- What does the contrapositive of the lemma say?
- $A \implies B$ iff $\neg B \implies \neg A$.
- L is regular $\implies$ Pumping Conditions.
  $\neg$(Pumping Conditions) $\implies$ L is not regular.

Using the Pumping Lemma

- Pumping lemma gives a useful property of regular languages. How do we use it?
- What does the contrapositive of the lemma say?
- $A \implies B$ iff $\neg B \implies \neg A$.
- L is regular $\implies$ Pumping Conditions.
  $\neg$(Pumping Conditions) $\implies$ L is not regular.

Recall the conditions of the lemma.

Using the Pumping Lemma

- Pumping lemma gives a useful property of regular languages. How do we use it?
- What does the contrapositive of the lemma say?
- $A \implies B$ iff $\neg B \implies \neg A$.
- L is regular $\implies$ Pumping Conditions.
  $\neg$(Pumping Conditions) $\implies$ L is not regular.

Recall the conditions of the lemma.

$$\exists n \forall w \in L\,(|w| \geq n \implies$$
$$\exists xyz.(xyz = w \land y \neq \epsilon \land |xy| \leq n \land (\forall k \quad xy^k z \in L)))$$

Using the Pumping Lemma

- Pumping lemma gives a useful property of regular languages. How do we use it?
- What does the contrapositive of the lemma say?
- $A \implies B$ iff $\neg B \implies \neg A$.
- L is regular $\implies$ Pumping Conditions.
  $\neg$(Pumping Conditions) $\implies$ L is not regular.

Recall the conditions of the lemma.

$$\exists n \forall w \in L \, (|w| \geq n \implies$$
$$\exists xyz.(xyz = w \wedge y \neq \epsilon \wedge |xy| \leq n \wedge (\forall k \quad xy^k z \in L)))$$

What is the contrapositive??

Using the Pumping Lemma

- Pumping lemma gives a useful property of regular languages. How do we use it?
- What does the contrapositive of the lemma say?
- $A \implies B$ iff $\neg B \implies \neg A$.
- L is regular $\implies$ Pumping Conditions.
  $\neg$(Pumping Conditions) $\implies$ L is not regular.

Recall the conditions of the lemma.

$$\exists n \forall w \in L\, (|w| \geq n \implies$$
$$\exists xyz.(xyz = w \wedge y \neq \epsilon \wedge |xy| \leq n \wedge (\forall k \quad xy^k z \in L)))$$

What is the contrapositive??

$$\forall n \exists w \in L\, (|w| \geq n \wedge$$
$$\forall xyz.(xyz = w \wedge y \neq \epsilon \wedge |xy| \leq n \implies (\exists k \quad xy^k z \in L)))$$

In words...

Theorem

If

- for every n
- there is a word $w \in L$ where $|w| \geq n$.
- for every breakup of w into three words $w = xyz$ such that
  1. $y \neq \epsilon$
  2. $|xy| \leq n$

  there is a $k \geq 0$ such that $xy^k z \notin L$, then

  L is not regular.

In words...

Theorem

If

- for every n
- there is a word $w \in L$ where $|w| \geq n$.
- for every breakup of w into three words $w = xyz$ such that
  1. $y \neq \epsilon$
  2. $|xy| \leq n$

  there is a $k \geq 0$ such that $xy^k z \notin L$, then

  L is not regular.

How do we apply this lemma?

How to use the lemma?

Consider language L

- Let n be an arbitrary number (pumping length).
- (Cleverly) Find a representative string w of L of size $\geq n$.
- Try out all ways to break the string into xyz triplet satisfying that $|y| > 0$ and $|xy| \leq n$. There will be finitely many cases to consider.
- For every triplet show that for some i the string xy^iz is not in L, and hence it yields contradiction with pumping lemma.

Applying Pumping Lemma: Example 1

Consider

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

Applying Pumping Lemma: Example 1

Consider

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

- For each n, we need a word. Let $w = a^n b^n$.

Applying Pumping Lemma: Example 1

Consider

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

- For each n, we need a word. Let $w = a^n b^n$.
- The first n characters of w are a^n. The breaks x and y are to be from within a^n.

Applying Pumping Lemma: Example 1

Consider

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

- For each n, we need a word. Let $w = a^n b^n$.
- The first n characters of w are a^n. The breaks x and y are to be from within a^n.
- Let $x = a^i, y = a^j$ where $i + j \leq n$ and $j \neq 0$.

$$w = a^i a^j a^{n-i-j} b^n$$

Applying Pumping Lemma: Example 1

Consider

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

- For each n, we need a word. Let $w = a^n b^n$.
- The first n characters of w are a^n. The breaks x and y are to be from within a^n.
- Let $x = a^i, y = a^j$ where $i + j \leq n$ and $j \neq 0$.

$$w = a^i a^j a^{n-i-j} b^n$$

- Choose $k = 0$ for each i, j. The corresponding word is

$$a^i (a^j)^0 a^{n-i-j} b^n = a^{n-j} b^n \notin L_{a,b}$$

Example 2

Consider

$$L = \{w \in \{a,b\}^* \mid w \text{ has equal number of a's and b's}\}$$

Example 2

Consider

$$L = \{w \in \{a,b\}^* \mid w \text{ has equal number of a's and b's}\}$$

- For each n, we need a word. Let $w = a^n b^n$.

Example 2

Consider

$$L = \{w \in \{a,b\}^* \mid w \text{ has equal number of a's and b's}\}$$

- For each n, we need a word. Let $w = a^n b^n$.
- The first n characters of w are a^n. The breaks x and y are to be from within a^n.
- Let $x = a^i, y = a^j$ where $i + j \leq n$ and $j \neq 0$.

$$w = a^i a^j a^{n-i-j} b^n$$

- Choose $k = 0$ for each i, j. The corresponding word is

$$a^i (a^j)^0 a^{n-i-j} b^n = a^{n-j} b^n \notin L_{a,b}$$

Example 2 again

Consider

$$L = \{w \in \{a,b\}^* \mid w \text{ has equal number of a's and b's}\}$$

Example 2 again

Consider

$$L = \{w \in \{a,b\}^* \mid w \text{ has equal number of a's and b's}\}$$

Exercise: What is $L \cap a^*b^*$?