# Automata Theory and Formal Languages
# Class 3

Marcin Luckner, PhD
mluckner@mini.pw.edu.pl

Version 1.2
October 28, 2021

# Regular language

- A regular language is a language generated by a regular expression.
- To show that a language is regular one can:
  - create a regular expression,
  - use the Myhill-Nerode lemma.
- To show that a language is not regular one can:
  - use the pumping lemma contraposition,
  - use the Myhill-Nerode lemma.

# Myhill-Nerode lemma

## Myhill-Nerode lemma

Language $L$ is regular if and only if relation $R_L$ induced by language $L$ has the finite number of equivalence classes.

# Myhill-Nerode lemma - example I

- Use Myhill-Nerode theorem to prove that the following language is regular
  - The language $L$ over alphabet $\Sigma = \{a, b\}$ without three adjacent identical letters.
    $aabbaa \in L$, $aaab \notin L$

# Myhill-Nerode lemma - example II

- Let us name the equivalence classes:

| | |
|---|---|
| $A_a$ | words in $L$ that end with a single $a$ |
| $A_{aa}$ | words in $L$ that end with a double $a$ |
| $A_b$ | words in $L$ that end with a single $b$ |
| $A_{bb}$ | words in $L$ that end with a double $b$ |
| $A_\epsilon$ | $\{\epsilon\}$ |
| $A_R$ | words that are not in $L$ |

- The union of the classes covers $\Sigma^*$:
  - A word from the language must end with $a, aa, b, bb$ or be the empty word $\epsilon$
  - All other words are in $A_R$.

# Myhill-Nerode lemma - example III

- Elements from the same class are in the relation:

$$\begin{aligned}
A_a &\ \forall w \in A_a wz \in L \iff \\
&\ z \in L \wedge z \text{ starts with not more than one a} \\
A_{aa} &\ \forall w \in A_{aa} wz \in L \iff \\
&\ z \in L \wedge z \text{ does not start with a} \\
A_b &\ \forall w \in A_b wz \in L \iff \\
&\ z \in L \wedge z \text{ starts with not more than one b} \\
A_{bb} &\ \forall w \in A_{bb} wz \in L \iff \\
&\ z \in L \wedge z \text{ does not start with b} \\
A_\epsilon &\ \forall w \in A_\epsilon wz \in L \iff z \in L \\
A_R &\ \forall w \in A_R wz \in L \iff z \in \emptyset
\end{aligned}$$

# Myhill-Nerode lemma - example IV

- Elements from other classes are not in the relation.
- It is enough to find $z$ that attached to a word from one class creates a word from $L$ and attached to a word from another class creates a word outside $L$.

- $A_a$,$A_{aa}$: $z = a$
- $A_a$,$A_b$: $z = aa$
- $A_a$,$A_{bb}$: $z = aa$
- $A_a$,$A_\epsilon$: $z = aa$
- $A_a$,$A_R$: $z = a$

- $A_{aa}$,$A_b$: $z = a$
- $A_{aa}$,$A_{bb}$:$z = a$
- $A_{aa}$,$A_\epsilon$:$z = a$
- $A_{aa}$,$A_R$: $z = \epsilon$

- $A_b$,$A_{bb}$: $z = b$
- $A_b$,$A_\epsilon$: $z = bb$
- $A_b$,$A_R$: $z = b$
- $A_{bb}$,$A_\epsilon$:$z = b$
- $A_{bb}$,$A_R$: $z = \epsilon$
- $A_\epsilon$,$A_R$: $z = \epsilon$

# Myhill-Nerode lemma - example V

$A_a$ words in $L$ that end with a single $a$

$A_{aa}$ words in $L$ that end with a double $a$

$A_b$ words in $L$ that end with a single $b$

$A_{bb}$ words in $L$ that end with a double $b$

$A_\epsilon$ $\{\epsilon\}$

$A_R$ words that are not in $L$

- The number of classes is 6, which is finite.
- Therefore, $L$ is regular.

# Pumping lemma contraposition

## Pumping lemma contraposition

If for any constant $n_L$ there exists a word $z \in L$ such that

$$(|z| \geq n_L) \wedge [(\forall_{u,v,w} z = uvw \wedge |uv| \leq n_L \wedge |v| \geq 1) \exists_{i=0,1,2\ldots} z_i = uv^i w \notin L]$$

then language $L$ is not regular.

# Pumping lemma contraposition - example I

- Use the pumping lemma contraposition to prove that the following language is not regular.
    - The language $L = \{a^i b^j : i \geq j \geq 1\}$ over alphabet $\Sigma = \{a, b\}$

# Pumping lemma contraposition - example II

- Use the pumping lemma contraposition we will show that exists such $z = uvw$ from $L$ that pumping segment $v$ we can create a word not in $L$.
  1. For each $n \in N$ word
     $$z = a^n b^n$$
     belongs to language $L$.
  2. Because $|uv| \leq n$ then $v = a^k$ where $k \leq n$.
  3. Let us take $i = 0$. then word
     $$z_0 = a^{n-k} b^n \notin L$$
     Therefore, language $L$ is not regular.

# Pumping lemma contraposition - discussion

- The main aspect of the proof is a selection of $z = uvw$ word.
- Part $uv$ should be homogeneous. Then it is easier to describe all variants of $v$.
- The contraposition allows us to multiply $(i > 1)$ or remove $(i = 0)$ $v$.
- The word should be *on the edge* of the language. Then it is easy to modify the word in a way that creates a word not in $L$.

# Assignments I

1. Which of the following languages are regular? Prove your answer
   1.1 The language $L = \{a^i b^j a^k : i, j, k \in N \wedge i + k \leq j \leq 100\}$ over alphabet $\Sigma = \{a, b\}$
   1.2 The language $L = \{a^i b^j c^k : i, j, k \in N \wedge i + k \leq j\}$ over alphabet $\Sigma = \{a, b, c\}$
   1.3 The language $L$ over alphabet $\Sigma = \{a, b, c\}$, such that each sequence of identical letters is shorter than the previous one $aaabbc \in L$, $aabbaac \notin L$

# Assignments II

2. Use Myhill-Nerode theorem to prove that the following languages are regular
    2.1 The language $L$ over alphabet $\Sigma = \{0, 1\}$ contains words, which have at least two 1's after a pair of adjacent 0's and before the next pair of adjacent 0' or the end of a word. A sequence of three 0's is threatened as a single pair. $10100101 \in L$, $10100100 \notin L$
    2.2 The language $L$ over alphabet $\Sigma = \{0, 1\}$ contains words, which are a binary representation of odd numbers without useless zeros.

# Assignments III

3. Use the pumping lemma contraposition to prove that the following languages are not regular.
   3.1 The language $L$ over alphabet $\Sigma = \{0, 1\}$, such that for each position of 1 in the word a sum of 1's before position (including observed one) is bigger than a number of preceding pairs of adjacent 0's
   $100100101 \in L$, $0010000100 \notin L$
   3.2 The language $L$ over alphabet $\Sigma = \{0, 1\}$, a number of 1's has a common factor, different than 1, with the number of 0's (different than 1)
   $011000 \in L$, $00101 \notin L$