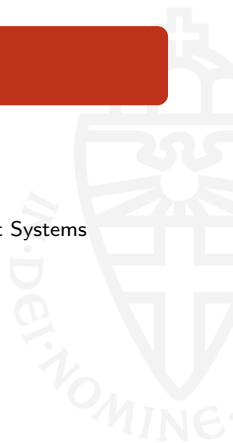


# Non-Regular Languages

H. Geuvers and T. van Laarhoven

Institute for Computing and Information Sciences – Intelligent Systems  
Radboud University Nijmegen

Version: fall 2014



# Outline

The Class of Regular Languages

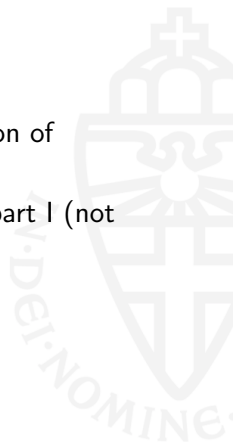
The Pumping Lemma for Regular Languages



# Organisational

Next week is different:

- Hours 1-2 : recap of part I of FLGA and discussion of assignment4.
- Hours 3-4 (starting **10:30** sharp): Test-exam on part I (not mandatory).
- Locations: HG00.310, HG00.065 and HG00.206



# Equivalence of definitions

**Theorem.** Let  $L \subseteq \Sigma^*$ . Then the following are equivalent

- 1  $L$  is “**machine-regular**”, i.e.  $L = \mathcal{L}(M)$  for some DFA (or NFA,  $\text{NFA}_\lambda$ )
- 2  $L$  is **regular**, i.e.  $L = \mathcal{L}(e)$  for some regular expression

(**Proof.** See previous lectures.)

So:

- To show that a language is regular we can give a regular expression or a (non-)deterministic automaton (with  $\lambda$ -steps).
- To show closure properties of the class of regular languages, we can use regular expressions, deterministic automata, non-deterministic automata, ...

# Closure properties of the Class of regular languages

If  $L_1$  and  $L_2$  over  $\Sigma$  are regular then so are

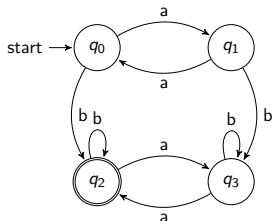
- $\overline{L_1}$ , (NB.  $\overline{L_1} = \{w \in \Sigma^* \mid w \notin L_1\}$ .)
- $L_1^*$
- $L_1 \cup L_2$
- $L_1 \cap L_2$
- $L_1 L_2$
- $L_1^R$  (NB.  $L_1^R = \{w \in \Sigma^* \mid w^R \notin L_1\}$ .)
- $\{w \in \Sigma^* \mid \exists v \in L_1 (w \text{ is a prefix of } v)\}$   
(NB.  $w$  is a prefix of  $v$  if  $v = wu$  for some  $u \in \Sigma^*$ .)



# How do we show that a language is *not* regular?

Regular languages can be **pumped!**

Example: Consider  $\Sigma = \{a, b\}$  and the automaton



accepting

$$\{w \in \Sigma^* \mid |w|_b \geq 1 \wedge |w|_a \text{ is even}\}$$

What happens if a word of length 4, 5, 6, 7, ... is accepted?

It has made a 'cycle' which can be repeated arbitrarily often!

For example,  $baaaa$  is accepted, and also all  $baa(aa)^n$  are accepted.

We say  $aa$  is a substring that *can be pumped*.

# Non regular languages

Let  $\Sigma = \{a, b\}$ . We will develop a technique that can be used to show that languages are **not regular**.

This technique will be applied to show that

$$\{a^n b^n \in \Sigma^* \mid n \geq 0\}$$

is not regular  
and to show that

$$\{w \in \Sigma^* \mid w \text{ is a palindrome}\}$$

is not regular.

A *palindrome* is a word  $w$  such that  $w^R = w$ .

Remember that  $w^R$  is the *reverse of*  $w$ , defined by

$$\begin{aligned}\lambda^R &:= \lambda \\ (s w)^R &:= w^R s\end{aligned}$$



## Pumping Lemma for Regular Languages

4

**Pumping Lemma.** Let  $L \subseteq \Sigma^*$  be a regular language  
Then there **exists a number**  $k \geq 1$  (pumping number) such that  
for **every**  $w \in L$  **with**  $|w| \geq k$  one has the following

- 1  $w$  can be split in three parts,  $w = uvz$ ,
- 2 with  $|uv| \leq k$  and  $|v| \geq 1$ ,
- 3 such that for all  $n \geq 0$  one has  $uv^n z \in L$ .

**Corollary**  $L = \{a^n b^n \mid n \geq 0\}$  **is not regular**

Proof. Suppose  $L$  is regular. (Towards a contradiction.)

Let  $k \geq 1$  be as in the Pumping Lemma

Take  $w = a^k b^k$ . Then  $w \in L$  and  $|w| \geq k$

Therefore there are  $u, v, z$  such that we can write  $a^k b^k = uvz$ ,  
with  $|uv| \leq k$ ,  $|v| \geq 1$  and  $uv^n z \in L$  for all  $n \geq 0$ .

Then  $v = a^q$ , for some  $q \geq 1$ . But then  $uv^2 z = a^{k+q} b^k \notin L$ .

Contradiction. 



# Proof of the Pumping Lemma

Let  $L$  be regular. Let  $M$  be an NFA that accepts  $L$ .

Take  $k$  to be the number of states of  $M$ .

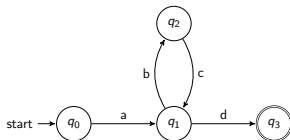
Let  $w \in L$  with  $|w| \geq k$ . Then, reading word  $w$ , we must pass some state more than once.

Say that  $q$  is the first state that we pass twice (reading  $w$ ).

Then  $w = uvz$ , where we read  $u$  to go to  $q$ , read  $v$  to loop at  $q$ , read  $z$  to go to a final node.

But then  $uv^n z$  is accepted for all  $n$ .

**Example**  $abcd \in \mathcal{L}(M)$

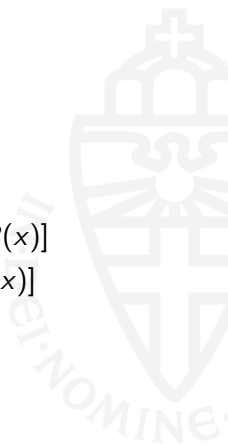


Since  $q_1$  is visited twice we can pump:

$a(bc)^n d \in \mathcal{L}(M)$  for all  $n \geq 0$ .

# Negating formulas

$$\begin{aligned}\neg \exists x. P(x) &\iff \forall x. \neg P(x) \\ \neg \forall x. P(x) &\iff \exists x. \neg P(x) \\ \neg \exists x. [Q(x) \wedge P(x)] &\iff \forall x. [Q(x) \Rightarrow \neg P(x)] \\ \neg \forall x. [Q(x) \Rightarrow P(x)] &\iff \exists x. [Q(x) \wedge \neg P(x)]\end{aligned}$$



# Using the Pumping Lemma to prove non-regularity

**Pumping lemma.**  $L$  is regular  $\Rightarrow L$  can be pumped

**We use this as follows:**

$L$  cannot be pumped  $\Rightarrow L$  is not regular

$L$  can be pumped means:

$$\exists k \geq 1 \forall w \in L. (|w| \geq k \Rightarrow \\ \exists u, v, z [w = uvz \wedge |uv| \leq k \wedge |v| \geq 1 \wedge \forall n \in \mathbb{N} (uv^n z \in L)])$$

$L$  cannot be pumped means:

$$\forall k \geq 1 \exists w \in L. (|w| \geq k \wedge \\ \forall u, v, z [w = uvz \wedge |uv| \leq k \wedge |v| \geq 1 \Rightarrow \exists n \in \mathbb{N} (uv^n z \notin L)])$$

To show that  $L$  is **not regular** it suffices to show it cannot be pumped.

# Using the pumping lemma to show that a language is non-regular

To show that  $L$  is not regular we do the following:

**For each  $k \geq 1$ , find some  $w \in L$  of length  $\geq k$  so that**

- **for every** way of splitting up  $w$  as  $w = uvz$ ,
- with  $|uv| \leq k$  and  $|v| \geq 1$ ,
- you can **find an  $n \geq 0$**  for which  $uv^n z$  is not in  $L$ .

**Application:**  $L = \{w \in \Sigma^* \mid w \text{ is a palindrome}\}$  is not regular.


Proof. We follow the procedure above.

Let  $k \geq 1$  (**arbitrary**)

**Take**  $w = a^k b a^k$ . Then  $w \in L$  (**check**) and  $|w| \geq k$  (**check**)

Let  $u, v, z$  (**arbitrary**) be so that  $a^k b a^k = uvz$ , with  $|uv| \leq k$  and  $|v| \geq 1$ . (Say  $|v| = p \geq 1$ .)

**Take**  $n = 0$ . Then  $uv^n z = uv^0 z = a^{k-p} b a^k \notin L$  (**check**).

So,  $L$  is not regular. 

# Proving that a language is non-regular

Let  $\Sigma := \{a, b\}$ . We know that  $L = \{a^n b^n \mid n \geq 0\}$  is not regular.  
Is  $L' := \{w \in \Sigma^* \mid \forall n \in \mathbb{N} (w \neq a^n b^n)\}$  regular?

**Answer:** No it is not. If  $L'$  is regular, then  $\overline{L'}$  would also be regular, but this is just  $L$  and  $L$  is not regular! So  $L'$  is not regular.

**Lemma** If  $L$  is *not* regular, then also  $\overline{L}$  and  $L^R$  are not regular

Let  $\Sigma := \{a, b, c\}$ .

Is  $L'' := \{a^n c^p b^n \in \Sigma^* \mid n \geq 0, p \geq 0\}$  regular?

**Answer:** No it is not.  $L = L'' \cap \mathcal{L}(a^* b^*)$ . If  $L''$  is regular, then  $L$  would be regular as well, but it is not!

**Lemma** If  $L$  is *not* regular and  $L = L_1 \cap L_2$ , with  $L_1$  regular, then  $L_2$  is not regular.