Non-Regular Languages

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

(i) L is "machine-regular", i.e. $L = L(M)$ for some DFA (NFA, NFA$_\lambda$)

(ii) L is regular, i.e. $L = L(e)$ for some regular expression

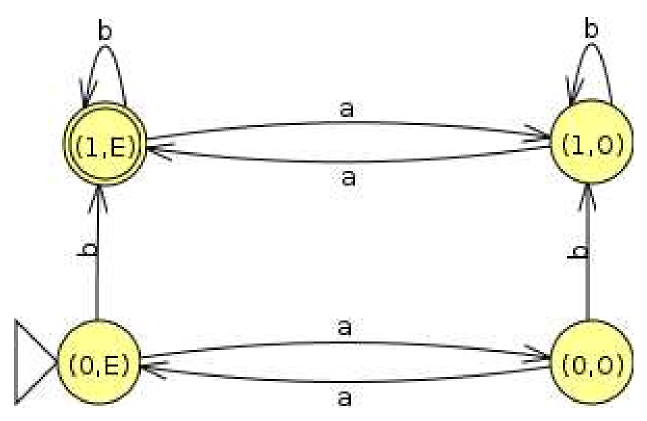
Proof. See previous lectures.

So:

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

How to show that a language is *not* regular?

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



accepting

$$\{w \in \Sigma^* \mid \#_b(w) \geq 1 \land \#_a(w) \text{ is even}\}$$

What happens if a word of length $4, 5, 6, 7, \ldots$ 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*.

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

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

Theorem. Let $L \subseteq \Sigma^*$ be a regular language

Then there exists a number $p \geq 1$ (pumping number) such that

for every $w \in L$ with $|w| \geq p$ one has the following

(i) w can be split in three parts, $w = xyz$,

(ii) with $|xy| \leq p$ and $|y| \geq 1$,

(iii) such that for all $n \geq 0$ one has $xy^n z \in L$.

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

Proof. Suppose L is regular. Let $p \geq 1$ be as in the pumping lemma

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

Therefore there are x, y, z such that we can write $a^p b^p = xyz$, with $|xy| \leq p$ and $xy^n z \in L$ for all $n \geq 0$.

Then $y = a^q$, for some $q \geq 1$. But then $xy^2 z \notin L$. Contradiction. ■

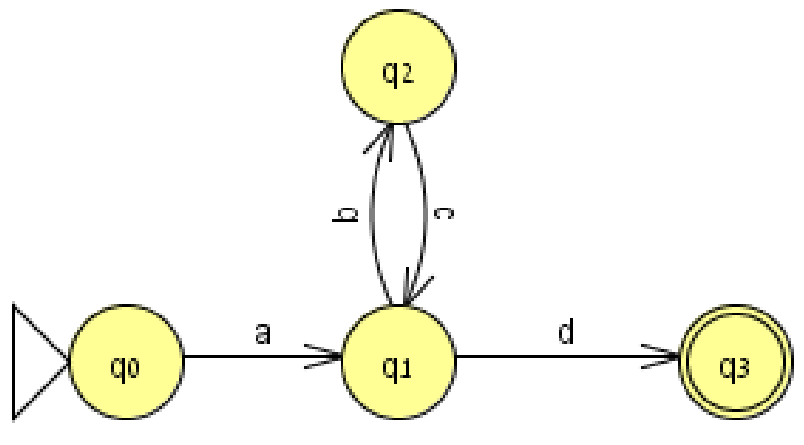
Let L be regular. Then L is m-regular. Let L be accepted by M.

Let M have p states

Then a word w of length $\geq p$ must pass twice a state q

Then $w = xyz$, where we read x to go to q, read y to loop at q, read z to go to a final node. But then $xy^n z$ is accepted for all n. ∎

Example $abcd \in L(M)$



Since q_1 is visited twice we can pump: $a(bc)^n d \in L(M)$ for all $n \geq 0$.

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

Pumping lemma. For L a language, L is regular $\Rightarrow$ L can be pumped

We use this as follows: For L a language,

$$L \text{ cannot be pumped } \Rightarrow L \text{ is not regular}$$

L can be pumped means:

$\exists p \geq 1 \forall w \in L.(|w| \geq p \;\Rightarrow$

$\exists xyz.[w = xyz \;\&\; |xy| \leq p \;\&\; |y| \geq 1 \;\&\; \forall n \in \mathbb{N}\,.xy^n z \in L])$

L cannot be pumped means:

$\forall p \geq 1 \exists w \in L.(|w| \geq p \;\&$

$\forall xyz.[w = xyz \;\&\; |xy| \leq p \;\&\; |y| \geq 1 \;\&\; \exists n \in \mathbb{N}\,.xy^n z \notin L])$

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

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

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

- for every way of splitting up w as $w = x\,y\,z$,
- with $|x\,y| \leq p$ and $|y| \geq 1$,
- you can find an $n \geq 0$ for which $x\,y^n\,z$ 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 $p \geq 1$ (arbitrary)

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

Let x, y, z (arbitrary) be so that $a^p b a^p = xyz$, with $|xy| \leq p$ and $|y| \geq 1$.

Take $n = 0$. Then $xy^n z = xy^0 z \notin L$ (check).

So, L is not 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 $\overline{L'} = L$ would also be regular, but 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. $L = L'' \cap L(a^* b^*)$. If L'' is regular, then also L is regular, but it is not!

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