Identifying Nonregular Languages

[Linz § 4.1]

Pigeonhole Principle (PHP)

If we put n items in m boxes (pigeonholes) and $n > m$, then at least 1 box has > 1 item in it.

Based on this:

Pumping Lemma for Regular Languages (PL)

Let L be an (infinite) regular language.

Then $\exists m > 0$ s.t. $\forall w \in L$ with $|w| \geq m$,

w can be **decomposed** as

$$w = xyz$$

with $$y \neq \lambda$$

and $$|xy| \leq m$$

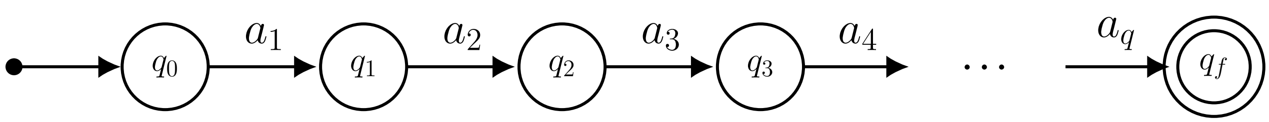
s.t. $\forall i = 0, 1, 2, ...$

$$w_i = xy^i z \in L$$

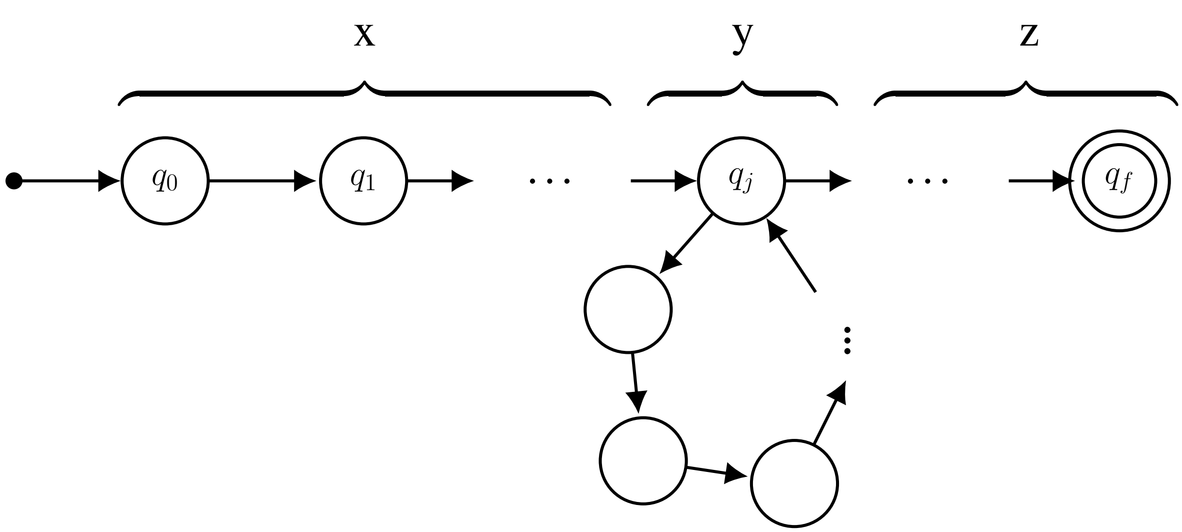
$\boldsymbol{Proof}$: Since L is regular, ,$\exists$ **dfa** M s.t. $L = L(M)$.

Suppose M has p states. Take $m = p + 1$
Let w be any word of L with $|w| = q \geq m > p$.
Say $w = a_1 a_2 a_3 ... a_q$ $(q > p)$.

Take any path of w in M, from a **start state** q_0 to a **final state** q_f :



Since there are only $p < q$ states,
by the **PHP**, the path must **visit some state** q_j **more than once**,
producing a **loop**



Let $x =$ segment of w up to the **first loop**

$y =$ segment of w in the loop

$z =$ the rest of w.

(Assume this is the **first loop** on the path of w.)

Note:

- $y \neq \lambda$
- $|xy| \leq m$
- $w_i = xy^i z \in L \qquad \forall i = 0, 1, 2, \ldots$

Also, taking

$i = 0:$	$w_0 = xz \in L$	(no loop)
$i = 1:$	$w_1 = w \in L$	(loop once)
$i > 1:$	$w_i \in L$	(loop more than once)

Example 1: Show $L = \{a^n b^n \mid n \geq 0\}$ is not regular

Proof : (by contradiction)

Suppose L is regular.

Then $L = L(M)$ for some **dfa** M.

Let $m = (\textbf{\# states in } M) + 1$

Let $w = a^m b^m \in L$. (1)

So $$|w| \geq m$$

By PL: $\exists x,\ y \neq \lambda,\ z:\quad w = xyz,$ (2)

$$|xy| \leq m, \tag{3}$$

and $\forall i \geq 0$ $$w_i = xy^i z \in L \tag{4}$$

Note:

by (1), (3): $xy \preceq a^m$

Let $$\ell = |y| > 0$$

Taking $i = 0$: $w_0 = \underline{xz = a^{m-\ell} b^m \notin L \quad ⨳\ (4)}$

OR

Taking $i = 2$: $w_2 = \underline{xy^2 z = a^{m+\ell} b^m \notin L \quad ⨳\ (4)}$

Note:

We can **choose** w as in (1) (provided $|w| \geq m$)

but we **cannot choose** x, y, z as in (2).

Note 1:

L is generated by a simple grammar!

Note 2:

Why (intuitively) is $\{a^n b^n \mid n \geq 0\}$ not regular

but $\{(ab)^n \mid n \geq 0\}$ is?

Example 2: Show $L = \{a^m b^n \mid m < n\}$ is not regular

$Proof$:

Suppose L is regular.

Then $L = L(M)$ for some **dfa** M.

Let $m >$ **# states in** M

Let $w = a^m b^{m+1} \in L$.

So $|w| \geq m$.

By PL: $\exists x,\ y \neq \lambda,\ z:$ $w = xyz,$

and $|xy| \leq m,$

and $\forall i \geq 0$ $w_i = xy^i z \in L$

Note: $xy \preceq a^m$

Let $$\ell = |y| > 0$$

Taking $i = 2$: $w_2 = xy^2z = a^{m+\ell}b^{m+1} \notin L$, $\times\!\!\times$

Example 3: Show $L = \{u \in \Sigma^* \mid n_a(u) = n_b(u)\}$ is not regular

<u>Method 1</u> Suppose L is regular

Then $L = L(M)$ for some **dfa** M.

Let $m >$ **# states in** M

Let $w = a^m b^m \in L$

Etc., ***exactly*** as in Example 1.

<u>Method 2</u> Let $L' = \{a^m b^n \mid m, n > 0\}$

<u>Note</u>:

L' is **regular** (WHY?)

$$\underline{L' = L(a^* b^*)}$$

Suppose L is regular.

Then $L \cap L'$ is regular (by result (5) on p. 4-4)

But $L \cap L' = \underline{\{a^n b^n \mid n \geq 0\}}$, $\times\!\!\times$ Ex. 1.

Example 4: Prove:

$$L = \{a^k \mid \textbf{k is a perfect square}\}$$
$$= \{a^{n^2} \mid n \geq 0\}$$
$$= \{\lambda,\ a,\ a^4,\ a^9,\ \ldots\}$$ **is not regular.**

$Proof$: Suppose L is regular.

Then $L = L(M)$ for some **dfa** M.

Let $m >$ **# states in** M

Let $w = a^{m^2} \in L$.

Note: $|w| = m^2 > m$

By PL: $\exists x,\ y \neq \lambda,\ z:\quad w = xyz,$

and $\quad |xy| \leq m, \quad$ (1)

and $\quad \forall i \geq 0, \quad w_i = xy^i z \in L.$

Let $\quad \ell = |y| > 0$

Then by (1): $0 < \ell \leq m$ (2)

Take $i = 2$: $\quad w_2 = xy^2 z \in L$

But $|w_2| = m^2 + \ell > m^2$ (3)

Also,

$$\begin{aligned}(m+1)^2 &= m^2+2m+1 \\ &> m^2+\ell \qquad \textbf{(by (2))} \\ &= |w_2| \qquad\qquad (4)\end{aligned}$$

So by (3) and (4): $m^2 < |w_2| < (m+1)^2$

So $|w_2|$ ***cannot be a perfect square!***