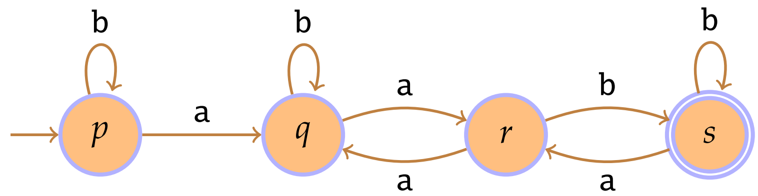

	w_1		w_2		w_3		w_4		w_5		w_6		w_7		w_8		w_9	
	a		a		b		a		a		b		b		a		b	
p		q		r		s		r		q		q		q		r		s
q_0		q_1		q_2		q_3		q_4		q_5		q_6		q_7		q_8		q_9

So far we have seen that for all $w \in L$

$$|w| \geq p\,,\ \exists\, x, y, z \in \Sigma^* \text{ s.t. } w = xyz \text{ with } xy^*z \in L$$

Let us enlist a few of the decompositions

	w_1		w_2		w_3		w_4		w_5		w_6		w_7		w_8		w_9	
	a		a		b		a		a		b		b		a		b	
p		q		r		s		r		q		q		q		r		s
q_0		q_1		q_2		q_3		q_4		q_5		q_6		q_7		q_8		q_9

$\underbrace{\texttt{aa}}_{x}\underbrace{\texttt{ba}}_{y}\underbrace{\texttt{abbab}}_{z}$ $\quad q_2 = q_4$

$\underbrace{\texttt{a}}_{x}\underbrace{\texttt{abaabb}}_{y}\underbrace{\texttt{ab}}_{z}$ $\quad q_1 = q_7$

$\underbrace{\texttt{a}}_{x}\underbrace{\texttt{abaa}}_{y}\underbrace{\texttt{bbab}}_{z}$ $\quad q_1 = q_5$

$\underbrace{\texttt{aa}}_{x}\underbrace{\texttt{baabba}}_{y}\underbrace{\texttt{b}}_{z}$ $\quad q_2 = q_8$

$\underbrace{\texttt{a}}_{x}\underbrace{\texttt{abaab}}_{y}\underbrace{\texttt{bab}}_{z}$ $\quad q_1 = q_6$

$\underbrace{\texttt{aab}}_{x}\underbrace{\texttt{aabbab}}_{y}\underbrace{\varepsilon}_{z}$ $\quad q_3 = q_9$

Pumping Lemma

For every regular language L there exists some constant p such that, for every string $w \in L$ with $|w| \geq p$, there exist $x, y, z \in \Sigma^*$ with $w = xyz$, $|y| \geq 1$, $|xy| \leq p$, and for all $i \in \mathbb{N}$, $xy^i z \in L$.

Proof

Let L be a regular language.

Let $M = (R, \Sigma, \delta, r_0, F)$ be a DFA recognizing L

Let

$$p \tag{1}$$

be the number of states in M

Consider $w \in L$ with $|w| \geq p$

That is

$$w = w_1 w_2 \cdots w_n \quad \text{s.t.} \quad n \geq p$$

Here $w_j \in \Sigma$, $j = 1, 2, 3, \ldots, n$

Let

$$q_0 q_1 q_3 \cdots q_n \tag{2}$$

that M enters while processing w.

Here

$$q_0 = r_0$$

and

$$q_j = \delta(q_{j-1}, w_j) \quad \text{for} \quad j = 1, 2, 3, \ldots, n$$

The length of sequence (2) is at least $p + 1$

Among the first $p + 1$ elements in the sequence, two must be the same, by the pigeonhole principle.

We call the first of these q_k and the second q_l.

That is

$$q_k = q_l \quad \text{with} \quad 0 \leq k < l \leq p \tag{3}$$

Let

$$x = w_1 w_2 \cdots w_k$$

$$y = w_{k+1} w_{k+2} \cdots w_l$$

$$z = w_{l+1} w_{l+2} \cdots w_n$$

Obviously $w = xyz$.

From (3) we have

$$k \neq l \implies y \neq \varepsilon \implies |y| > 0 \implies |y| \geq 1 \tag{4}$$

Furthermore, from the decomposition of w we have $|xy| = l$.

And from (3) it can be inferred that

$$\left|xy\right| \leq p \tag{5}$$

As

$$x \text{ takes } M \text{ from } q_0 \text{ to } q_k$$

$$y \text{ takes } M \text{ from } q_k \text{ to } q_k (= q_l)$$

$$\text{and} \quad z \text{ takes } M \text{ from } q_k \text{ to } q_n \in F$$

Therefore, M must accept

$$xy^i z \quad \text{for} \quad i \geq 0$$

That is

$$xy^iz \in L\,,\ \forall\, i \in \mathbb{N} \tag{6}$$

From (1), (3), (4), and (6) we have established the truth of the pumping lemma.

Succinctly the pumping lemma states

$$L \text{ is regular} \implies \Big(\exists\, p\,\forall\, w\,,\ |w| \geq p, \exists\, x, y, z\,,\ |xy| \leq p\,,\ |y| \geq 1\,,\ \forall\, i\,,\ xy^iz \in L\Big) \tag{7}$$

The contrapositive of (7)

$$\Big(\forall\, n\,\exists\, w\,,\ |w| \geq n, \forall\, x, y, z\,,\ |xy| \leq n\,,\ |y| \geq 1\,,\ \exists\, i\,,\ xy^iz \notin L\Big) \implies L \text{ is not regular}$$