1.4 Non regular languages

How can we prove that a language is not regular?

$$L_1 = \{ w \in \Sigma^* \mid |w| = 27 \}$$

$$L_2 = \{ w \in \Sigma^* \mid |w| > 27 \}$$

$$L_3 = \{ a^n b^n \mid n \geq 0 \}$$

$$L_4 = \{ a^n b^n \mid n \geq 28 \}$$

$$L_5 = \{ w \in \{0,1\}^* \mid \#_0(w) = \#_1(w) \}$$

$$L_6 = \{ w \in \{0,1\}^* \mid \#_{01}(w) = \#_{10}(w) \}$$

L_1 is regular

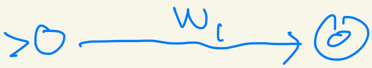
$$\to q_0 \xrightarrow{a \in \Sigma} q_1 \to \qquad \to q_{27} \xrightarrow{a \in \Sigma} q_{28} \xrightarrow{a \in \Sigma} t$$

accepting

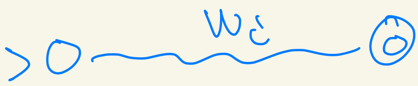
L_2 is regular

NB: Every finite language is regular!

$L = \{w_1, w_2, \ldots, w_k\}$ k fixed

M_1 

$\Leftarrow$ accept only w_1

M_i 

$\Leftarrow$ accept of w_i

Theorem 1.70 (pumping lemma)

Let A be a regular language over Σ

Then $\exists p \in \mathbb{N}$ such that

$\forall s \in A$ with $|s| \geq p$ $\exists x, y, z \in \Sigma^*$ such that $s = xyz$

1. $xy^iz \in A \quad \forall i \geq 0$

2. $|y| > 0$

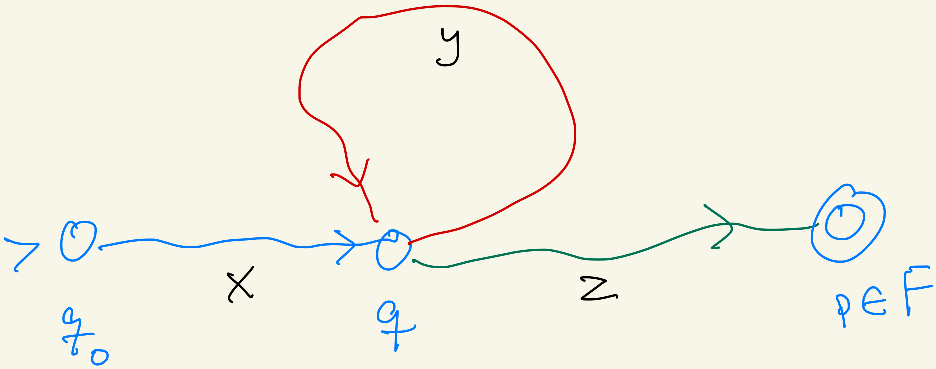
3. $|xy| \leq p$

Proof: Let M be a DFA s.t $L(M) = A$

and let $p = \#$states in M.

- Since $|s| \geq p$ some state of M must be repeated while M reads s

- let q be the first state which is repeated as M reads s



1. holds as we can repeat y 0 or more times

2. holds as we read at least one symbol between the first two occurences of q

3. holds by the choice of q (and the pigeonhole principle)

$\square$

How to use the pumping lemma to show that a language L is <u>not</u> regular:

Note that L must be infinite and hence contain arbitrarily long strings as otherwise it is regular

Proof by contradiction (seen as a game against an adversary)

- Suppose L <u>is</u> regular (*)
- Then $L = L(M)$ for some DFA M

  Let p be #states in M
- We choose a string $s \in L$ with $|s| \geq p$ (designed to get a contradiction)
- The adversary gives a partition $s = xyz$ satisfying 2. and 3. in Theorem 1.70
- We find a value i s.t. 1. does not hold $\rightarrow$ contradicting (*)

Example 1.73

$B = \{0^n 1^n \mid n \geq 0\}$

- Suppose B is regular. Let M be a DFA with $L(M) = B$ and let p be # states of M.

- We choose $s = 0^p 1^p$

- The adversary proposes a partition $s = xyz$ satisfying 2. and 3. that is, $|y| > 0$ and $|xy| \leq p$

- As $|xy| \leq p$ we must have $y = 0^r$ for some $r > 0$.

  $\underbrace{000\cdots}_{x}\underbrace{00}_{y}11\cdots 1$

  Now $xz = 0^{p-r} 1^p \notin B$

  contradicting 1. of Thm 1.70

Conclusion: B is <u>not</u> regular

Recall the languages L_3 and L_4:

$$L_3 = \{a^n b^n \mid n \geq 0\}$$

$$L_4 = \{a^n b^n \mid n \geq 28\}$$

- L_3 is the same as B if we do $0 \to a$, $1 \to b$

So same proof shows that L_3 is not regular

- $L_4 = L_3 \cap L_2$ $\left(L_2 = \{w \in \{a,b\}^* \mid |w| \geq 27\}\right.$

Can we conclude anything about L_4 from that?

$$L_3 = L_4 \cup \{a^n b^n \mid n \in \underbrace{\{0, 1, 2, \ldots 27\}}_{X}\}$$

$$= L_4 \cup X$$

X is regular

Suppose L_4 is regular

Then L_3 is the union of two regular languages and hence it is regular, contradiction as we have shown that L_3 is not regular

Conclusion: L_4 is not regular

Example 1.74

$C = \{ w \in \{0,1\}^* \mid \#_0(w) = \#_1(w) \}$

Suppose C is regular and let M be a DFA with $L(M) = C$ and let $p = $ #states of M

- We choose $s = 0^p 1^p$
- The adversary proposes a partition $s = xyz$ satisfying 2. and 3. that is, $|y| > 0$ and $|xy| \le p$
- As $|xy| \le p$ we must have $y = 0^r$ for some $r > 0$.

  Now $xz = 0^{p-r} 1^p \notin C$

  contradicting 1. of Thm 1.70

alternative proof using closure properties of regular languages:

If C is regular then $C \cap 0^*1^*$ is regular

but $B = C \cap 0^*1^*$ and we know that

B is not regular. Hence C is not regular.

<u>Example 1.75</u> $F = \{ww \mid w \in \{0,1\}^*\}$

Suppose M is a DFA with p states such that $F = L(M)$

- We choose $s = 0^p 1 0^p 1$
- Adversary proposes a partition $s = xyz$ s.t. 2. and 3. of Theorem 1.70 holds
- As $|xy| \leq p$ and $|y| > 0$, we have $y = 0^i$ for some $i \in \{1, 2, \ldots, p\}$

Now $xz = 0^{p-i} 1 0^p 1 \notin F$ $\}$ so F is <u>not</u> regular.

Alternative version of pumping lemma

Theorem 1.70*

Let A be a regular language over Σ

Then $\exists p \in \mathbb{N}$ such that

$\forall s \in A$ with $|s| \geq p$ $\exists x, y, z \in \Sigma^*$ such that $s = xyz$ and

1. $xy^iz \in A \quad \forall i \geq 0$

2. $|y| > 0$

3. $|yz| \leq p$

Example 1.77 (different proof)

Show that $E = \{0^i 1^j \mid i > j\}$ is not regular

Suppose $E = L(M)$ where M is a DFA with p states

We take $s = 0^{p+1} 1^p$

Adversary gives us $xyz = s$ s.t. $|y| > 0$ and $|yz| \leq p$

Then $y = 1^j$ for some $j \in \{1, 2, \ldots, p\}$ so $xy^2z = 0^{p+1} 1^{p+j} \notin E$ ☹

Hence E is not regular.

$$s = 0^{p+2}1^p \in L$$

$$s = xyz \qquad x = 0^j \quad y = 0^1 \quad z = 0^{p+1-j}1^p$$

$$xz = 0^{p+1}1^p \in L$$

$$xy^iz = 0^{p+1+i}1^p \in L$$

Example not from book: $L = \{0^n \mid n \text{ is a prime}\}$

Suppose M is a DFA with p states s.t $L(M) = L$

We take $s = 0^m$ where $m \geq p$ and m is a prime

The adversary gives us $x, y, z \in 0^*$ s.t

$s = xyz$, $|xy| \leq p$ and $|y| = q > 0$

Note that $|xy^iz| = (m - q) + iq = m + (i-1)q$

So for $i = m + 2q + 3$ we set

$$|xy^iz| = m + (m+2q+2)q$$

$$= m+2q + (m+2q)q$$

$$= \underbrace{(q+1)(m+2q)}_{\text{not a prime}} \qquad \lightning$$