

# CS 301

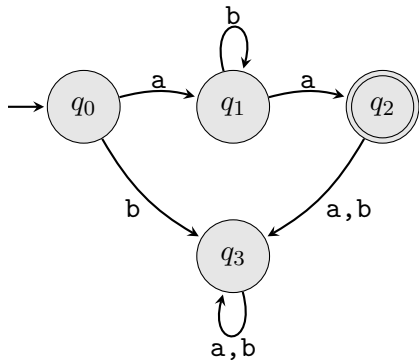
## Lecture 06 – Nonregular languages and the pumping lemma

Stephen Checkoway

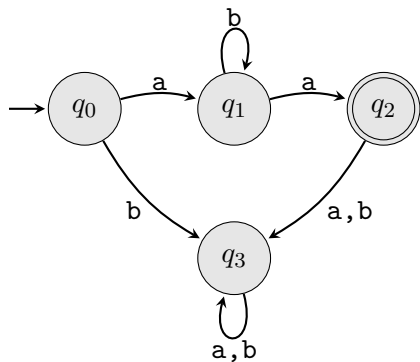
February 5, 2018



# DFA $M_1$



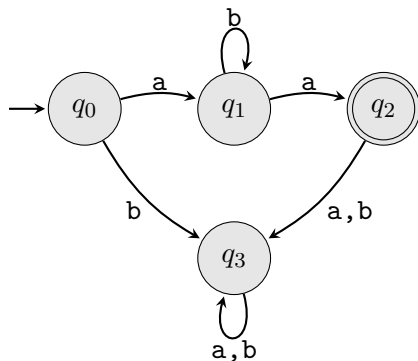
# DFA $M_1$



## Strings in the language

- aa
- aba
- abba
- abbba
- $ab^k a$  for all  $k \geq 0$

# DFA $M_1$



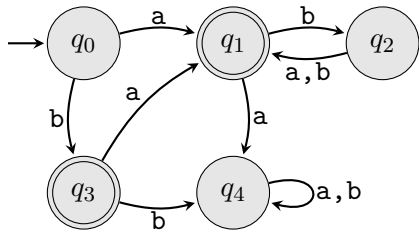
## Strings in the language

- aa
- aba
- abba
- abbba
- $ab^k a$  for all  $k \geq 0$

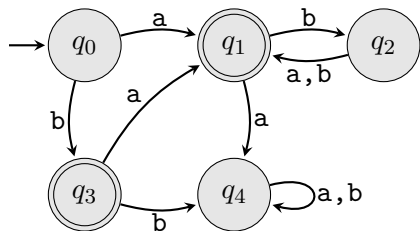
All of the strings  $w \in L(M_1)$  s.t.  $|w| \geq 3$  have a curious property:  $w$  can be written as  $w = xyz$  where

- 1  $|y| > 0$  and
- 2  $xy^i z \in L(M_1)$  for all  $i \geq 0$

# DFA $M_2$



## DFA $M_2$



Strings in the language include

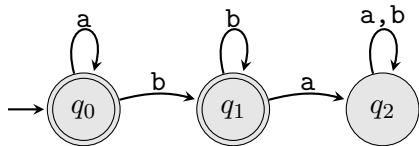
- a
- b
- ba
- aba
- abb
- baba
- abbba
- bababb

Again, strings  $w \in L(M_2)$  s.t.  $|w| \geq 3$  can be written as  $w = xyz$  with  $|y| > 0$  and  $xy^iz \in L(M_2)$ .

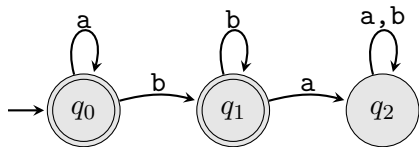
E.g.,  $x = ba$ ,  $y = ba$ ,  $z = \varepsilon$

- $xy^0z = ba$
- $xy^1z = baba$
- $xy^2z = bababa$
- ...

# DFA $M_3$



## DFA $M_3$



$$L(M_3) = \{a^m b^n \mid m, n \geq 0\}$$

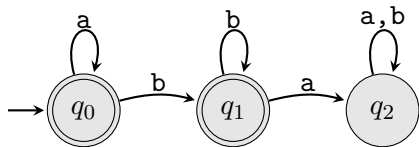
Strings  $w \in L(M_3)$  s.t.  $|w| \geq 1$  have the same property.

E.g.,  $x = \varepsilon$ ,  $y = a$ ,  $z = abb$

- $xy^0z = abb$
- $xy^1z = aabb$
- $xy^2z = aaabb$
- $xy^iz = a^{i+1}bb$



## DFA $M_3$



$$L(M_3) = \{a^m b^n \mid m, n \geq 0\}$$

Strings  $w \in L(M_3)$  s.t.  $|w| \geq 1$  have the same property.

E.g.,  $x = \varepsilon$ ,  $y = a$ ,  $z = abb$

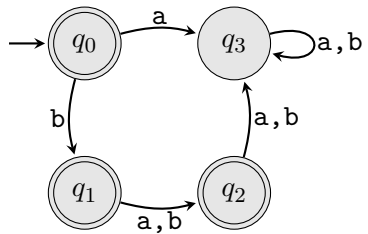
- $xy^0z = abb$
- $xy^1z = aabb$
- $xy^2z = aaabb$
- $xy^iz = a^{i+1}bb$

Not every way we split the strings works

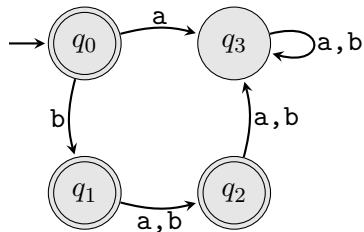
$x = a$ ,  $y = ab$ ,  $z = b$

- $xy^0z = ab \in L(M_3)$
- $xy^1z = aabb \in L(M_3)$
- $xy^2z = aababb \notin L(M_3)$

# DFA $M_4$



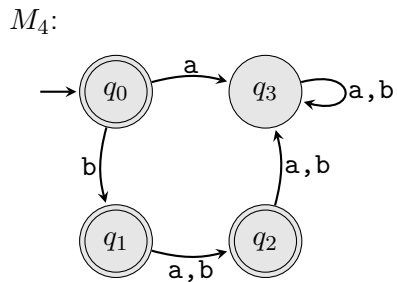
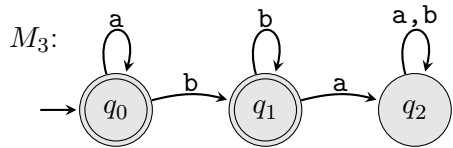
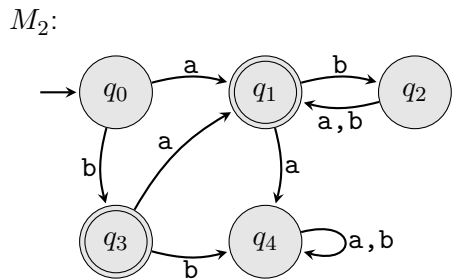
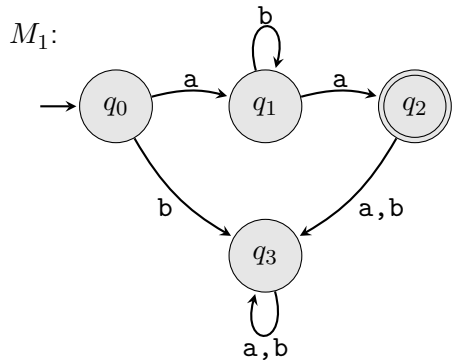
## DFA $M_4$



$$L(M_4) = \{\varepsilon, b, ba, bb\}$$

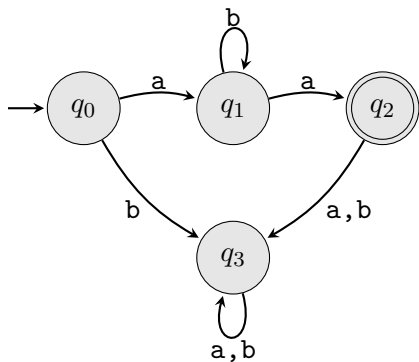
$L(M_4)$  doesn't appear to have this property (unless we say it holds for all strings in  $L(M_4)$  with length at least 3 because there are no such strings)

What do  $M_1$ ,  $M_2$ , and  $M_3$  have that  $M_4$  lacks?



## Repeated state for some string in the language

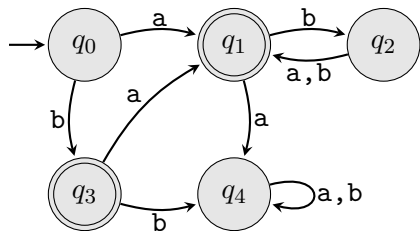
$M_1$ ,  $M_2$ , and  $M_3$  all have a repeated state in some accepting computation



On input  $aba$ ,  $M_1$  goes through states  $q_0$ ,  $q_1$ ,  $q_1$ ,  $q_2$

State  $q_1$  is repeated so we can repeat it 0 or more times by following the loop on  $b$

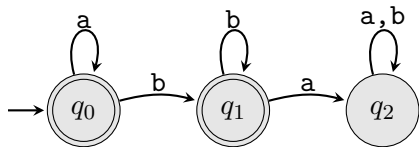
$M_2$



On input baba,  $M_2$  goes through states  $q_0, q_3, q_1, q_2, q_1$

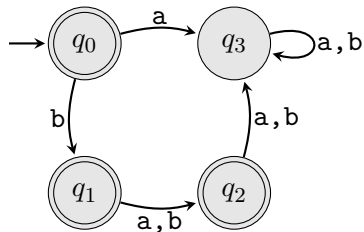
State  $q_1$  is repeated so we can perform the  $q_1 \rightarrow q_2 \rightarrow q_1$  sequence corresponding to input ba 0 or more times

$M_3$



On input aabb,  $M_2$  goes through states  $q_0$ ,  $q_0$ ,  $q_0$ ,  $q_1$ ,  $q_1$

State  $q_0$  is repeated so we can perform the  $q_0 \rightarrow q_0$  sequence corresponding to input a 0 or more times

$M_4$ 

None of the strings in  $L(M_4)$  lead to a repeated state

As mentioned, we can “cheat” and say that the property holds for strings of length at least 3 since  $L(M_4)$  has no strings of length at least 3



## Pumpable languages

A language  $A$  is said to be **pumpable** if there exists an integer  $p > 0$  s.t. for all strings  $w \in A$  with  $|w| \geq p$ , there exist strings  $x, y, z \in \Sigma^*$  with  $w = xyz$  s.t.

- 1  $xy^iz \in A$  for all  $i \geq 0$
- 2  $|y| > 0$
- 3  $|xy| \leq p$

The integer  $p$  is called the **pumping length**

## Pumpable languages

A language  $A$  is said to be **pumpable** if there exists an integer  $p > 0$  s.t. for all strings  $w \in A$  with  $|w| \geq p$ , there exist strings  $x, y, z \in \Sigma^*$  with  $w = xyz$  s.t.

- 1  $xy^i z \in A$  for all  $i \geq 0$
- 2  $|y| > 0$
- 3  $|xy| \leq p$

The integer  $p$  is called the **pumping length**

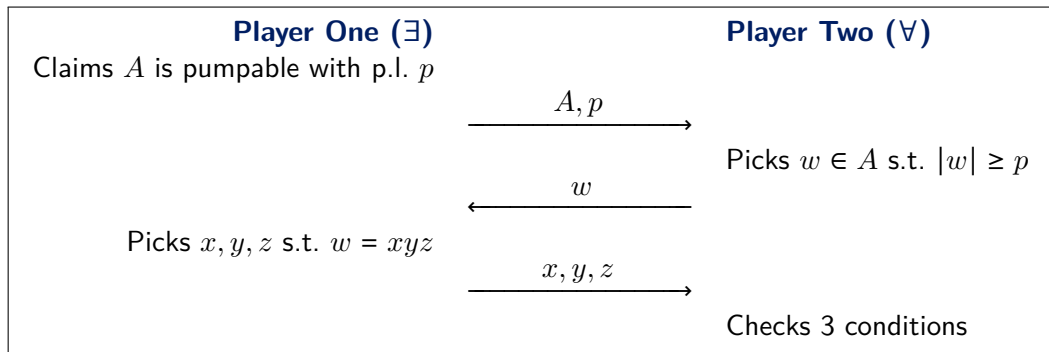
Almost certainly the most complicated mathematical definition you've seen:

$$\exists p > 0. \forall w \in A. \exists x, y, z \in \Sigma^*. \forall i \geq 0. [\dots]$$

Contrast with the definition of a continuous function  $f : \mathbb{R} \rightarrow \mathbb{R}$  from calculus

$$\forall \varepsilon > 0. \exists \delta > 0. [\dots]$$

## A two-player game



Player One “wins” the game if

- 1  $xy^iz \in A$  for all  $i \geq 0$
- 2  $|y| > 0$
- 3  $|xy| \leq p$

Player One can win if and only if  $A$  is pumpable

# Pumping lemma for regular languages

Theorem (Pumping lemma)

*Regular languages are pumpable.*

Note: The converse is *not* true! There are pumpable languages that are not regular

## Proof

Let  $M = (Q, \Sigma, \delta, q_0, F)$  be a DFA with  $L(M) = A$  and set  $p = |Q|$ .

If  $A$  contains no strings of length at least  $p$ , then we're finished since  $A$  is pumpable with pumping length  $p$ .

## Proof

Let  $M = (Q, \Sigma, \delta, q_0, F)$  be a DFA with  $L(M) = A$  and set  $p = |Q|$ .

If  $A$  contains no strings of length at least  $p$ , then we're finished since  $A$  is pumpable with pumping length  $p$ .

Otherwise, let  $w$  be a string in  $A$  of length  $n \geq p$ .

Write  $w = w_1 w_2 \cdots w_n$  where each  $w_i \in \Sigma$ .

Let  $r_0, r_1, \dots, r_n$  be the accepting computation of  $M$  on  $w$ .

By the pigeonhole principle, in the first  $p + 1$  states  $(r_0, r_1, \dots, r_p)$ , there are states  $r_j = r_k$  s.t.  $0 \leq j < k \leq p$ .

# Proof

Set

$$x = w_1 w_2 \cdots w_j$$

$$y = w_{j+1} w_{j+2} \cdots w_k$$

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

$$\begin{array}{l} \text{input:} \quad \overbrace{w_1 \ w_2 \ \cdots \ w_j}^x \ \overbrace{w_{j+1} \ w_{j+2} \ \cdots \ w_k}^y \ \overbrace{w_{k+1} \ w_{k+2} \ \cdots \ w_n}^z \\ \text{states:} \quad r_0 \ r_1 \ r_2 \ \cdots \ r_j \ r_{j+1} \ r_{j+2} \ \cdots \ r_k \ r_{k+1} \ r_{k+2} \ \cdots \ r_n \end{array}$$

Remember  $\delta(r_{m-1}, w_m) = r_m$  for all  $1 \leq m \leq n$

- ②  $|y| = k - j > 0$
- ③  $|xy| \leq p$  because  $k \leq p$

# Proof

①  $xy^iz \in A$

$$\delta(r_{m-1}, w_m) = r_m \quad \forall m$$



# Proof

$$\textcircled{1} \ xy^i z \stackrel{?}{\in} A$$

$i = 0$

$$\begin{array}{cccccccc} & & \underbrace{x} & & \underbrace{z} & & & \\ & & w_1 & w_2 & \cdots & w_j & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & & & & & \end{array}$$

$$\delta(r_{m-1}, w_m) = r_m \quad \forall m$$

# Proof

$$\textcircled{1} \ xy^i z \stackrel{?}{\in} A$$

$i = 0$

$$\begin{array}{cccccccc} & \underbrace{\hspace{10em}}_x & & \underbrace{\hspace{10em}}_z & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & & & & \end{array}$$

$$\begin{aligned} \delta(r_{m-1}, w_m) &= r_m \quad \forall m \\ \delta(r_j, w_{k+1}) &= \delta(r_k, w_{k+1}) = r_{k+1} \end{aligned}$$

# Proof

$$\textcircled{1} \quad xy^i z \stackrel{?}{\in} A$$

$i = 0$

$$\begin{array}{cccccccc} & \underbrace{\hspace{1.5cm}}_x & & \underbrace{\hspace{1.5cm}}_z & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$$\begin{aligned} \delta(r_{m-1}, w_m) &= r_m \quad \forall m \\ \delta(r_j, w_{k+1}) &= \delta(r_k, w_{k+1}) = r_{k+1} \end{aligned}$$

# Proof

①  $xy^iz \stackrel{?}{\in} A$

$$\delta(r_{m-1}, w_m) = r_m \quad \forall m$$
$$\delta(r_j, w_{k+1}) = \delta(r_k, w_{k+1}) = r_{k+1}$$

$i = 0$

$$\begin{array}{cccccccc} & \underbrace{\hspace{1.5cm}}_x & & \underbrace{\hspace{3.5cm}}_z & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$i = 1$

$$\begin{array}{cccccccccccc} & \underbrace{\hspace{1.5cm}}_x & & \underbrace{\hspace{3.5cm}}_y & & \underbrace{\hspace{3.5cm}}_z & & & & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{j+1} & w_{j+2} & \cdots & w_k & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{j+1} & r_{j+2} & \cdots & r_k & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

# Proof

①  $xy^iz \in A$

$$\delta(r_{m-1}, w_m) = r_m \quad \forall m$$
$$\delta(r_j, w_{k+1}) = \delta(r_k, w_{k+1}) = r_{k+1}$$

$i = 0$

$$\begin{array}{cccccccc} & \underbrace{x} & & \underbrace{z} & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$i = 1$

$$\begin{array}{cccccccccccc} & \underbrace{x} & & \underbrace{y} & & \underbrace{z} & & & & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{j+1} & w_{j+2} & \cdots & w_k & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{j+1} & r_{j+2} & \cdots & r_k & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$i = 2$

$$\begin{array}{cccccccccccc} & \underbrace{x} & & \underbrace{y} & & \underbrace{y} & & \underbrace{z} & & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{j+1} & w_{j+2} & \cdots & w_k & w_{j+1} & w_{j+2} & \cdots & w_k & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{j+1} & r_{j+2} & \cdots & r_k & & & & & & & & \end{array}$$

# Proof

①  $xy^iz \in A$

$$\begin{aligned} \delta(r_{m-1}, w_m) &= r_m \quad \forall m \\ \delta(r_j, w_{k+1}) &= \delta(r_k, w_{k+1}) = r_{k+1} \\ \delta(r_k, w_{j+1}) &= \delta(r_j, w_{j+1}) = r_{j+1} \end{aligned}$$

$i = 0$

$$\begin{array}{cccccccc} & \underbrace{x} & & \underbrace{z} & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$i = 1$

$$\begin{array}{cccccccccccc} & \underbrace{x} & & \underbrace{y} & & \underbrace{z} & & & & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{j+1} & w_{j+2} & \cdots & w_k & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{j+1} & r_{j+2} & \cdots & r_k & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$i = 2$

$$\begin{array}{cccccccccccc} & \underbrace{x} & & \underbrace{y} & & \underbrace{y} & & \underbrace{z} & & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{j+1} & w_{j+2} & \cdots & w_k & w_{j+1} & w_{j+2} & \cdots & w_k & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{j+1} & r_{j+2} & \cdots & r_k & & & & & & & & \end{array}$$

# Proof

$$\textcircled{1} \ x y^i z \stackrel{?}{\in} A$$

$$\delta(r_{m-1}, w_m) = r_m \quad \forall m$$
$$\delta(r_j, w_{k+1}) = \delta(r_k, w_{k+1}) = r_{k+1}$$
$$\delta(r_k, w_{j+1}) = \delta(r_j, w_{j+1}) = r_{j+1}$$

$$i = 0$$

$$\begin{array}{ccccccc} & \underbrace{x} & & \underbrace{z} & & & \\ & w_1 & w_2 & \cdots & w_j & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$$i = 1$$

$$\begin{array}{ccccccc} & \underbrace{x} & & \underbrace{y} & & \underbrace{z} & & \\ & w_1 & w_2 & \cdots & w_j & w_{j+1} & w_{j+2} & \cdots & w_k & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{j+1} & r_{j+2} & \cdots & r_k & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$$i = 2$$

$$\begin{array}{ccccccc} & \underbrace{x} & & \underbrace{y} & & \underbrace{y} & & \underbrace{z} & & \\ & w_1 & w_2 & \cdots & w_j & w_{j+1} & w_{j+2} & \cdots & w_k & w_{j+1} & w_{j+2} & \cdots & w_k & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{j+1} & r_{j+2} & \cdots & r_k & r_{j+1} & r_{j+2} & \cdots & r_k & & & & \end{array}$$

# Proof

$$\textcircled{1} \quad xy^i z \in A$$

$$\begin{aligned} \delta(r_{m-1}, w_m) &= r_m \quad \forall m \\ \delta(r_j, w_{k+1}) &= \delta(r_k, w_{k+1}) = r_{k+1} \\ \delta(r_k, w_{j+1}) &= \delta(r_j, w_{j+1}) = r_{j+1} \end{aligned}$$

$$i = 0$$

$$\begin{array}{cccccccc} & \underbrace{\phantom{w_1 \ w_2 \ \cdots \ w_j}}_x & & \underbrace{\phantom{w_{k+1} \ w_{k+2} \ \cdots \ w_n}}_z & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$$i = 1$$

$$\begin{array}{cccccccccccc} & \underbrace{\phantom{w_1 \ w_2 \ \cdots \ w_j}}_x & & \underbrace{\phantom{w_{j+1} \ w_{j+2} \ \cdots \ w_k}}_y & & \underbrace{\phantom{w_{k+1} \ w_{k+2} \ \cdots \ w_n}}_z & & & & & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{j+1} & w_{j+2} & \cdots & w_k & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{j+1} & r_{j+2} & \cdots & r_k & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$

$$i = 2$$

$$\begin{array}{cccccccccccccccc} & \underbrace{\phantom{w_1 \ w_2 \ \cdots \ w_j}}_x & & \underbrace{\phantom{w_{j+1} \ w_{j+2} \ \cdots \ w_k}}_y & & \underbrace{\phantom{w_{j+1} \ w_{j+2} \ \cdots \ w_k}}_y & & \underbrace{\phantom{w_{k+1} \ w_{k+2} \ \cdots \ w_n}}_z & & & & & & & & & & & \\ & w_1 & w_2 & \cdots & w_j & w_{j+1} & w_{j+2} & \cdots & w_k & w_{j+1} & w_{j+2} & \cdots & w_k & w_{k+1} & w_{k+2} & \cdots & w_n \\ r_0 & r_1 & r_2 & \cdots & r_j & r_{j+1} & r_{j+2} & \cdots & r_k & r_{j+1} & r_{j+2} & \cdots & r_k & r_{k+1} & r_{k+2} & \cdots & r_n \end{array}$$





## Proof

Starting in state  $r_j$ , when  $M$  reads  $y$ , it ends in state  $r_k = r_j$ .

Therefore, when  $M$  runs on  $xy^iz$ , it

- ① starts in state  $r_0 = q_0$  and after reading  $x$  is in state  $r_j$ ;
- ② for each of the  $i$  copies of  $y$ , it is in state  $r_j$ , reads  $y$ , and moves to state  $r_k = r_j$ ;  
and
- ③ from state  $r_k$ , it reads  $z$  and ends in state  $r_n \in F$

Therefore  $M$  accepts  $xy^iz$  so

- ①  $xy^iz \in A$
- ②  $|y| > 0$
- ③  $|xy| \leq p$ .

Therefore,  $A$  is pumpable.



# Proving languages are not regular

If we want to prove language  $A$  is not regular,

- 1 Assume  $A$  is regular

## Proving languages are not regular

If we want to prove language  $A$  is not regular,

- 1 Assume  $A$  is regular
- 2 Since  $A$  is regular, it is pumpable with pumping length  $p$

## Proving languages are not regular

If we want to prove language  $A$  is not regular,

- ① Assume  $A$  is regular
- ② Since  $A$  is regular, it is pumpable with pumping length  $p$
- ③ Construct a string  $w \in A$  of length at least  $p$

## Proving languages are not regular

If we want to prove language  $A$  is not regular,

- 1 Assume  $A$  is regular
- 2 Since  $A$  is regular, it is pumpable with pumping length  $p$
- 3 Construct a string  $w \in A$  of length at least  $p$
- 4 Show that every partition of  $w$  into  $xyz$  such that  $|xy| \leq p$  and  $|y| > 0$  yields some  $i$  such that  $xy^iz \notin A$

## Proving languages are not regular

If we want to prove language  $A$  is not regular,

- 1 Assume  $A$  is regular
- 2 Since  $A$  is regular, it is pumpable with pumping length  $p$
- 3 Construct a string  $w \in A$  of length at least  $p$
- 4 Show that every partition of  $w$  into  $xyz$  such that  $|xy| \leq p$  and  $|y| > 0$  yields some  $i$  such that  $xy^iz \notin A$
- 5 This contradicts the pumping lemma so our assumption must be false, namely  $A$  is not regular

## Example

Let's prove  $A = \{0^n 1^n \mid n \geq 0\}$  is not regular

## Example

Let's prove  $A = \{0^n 1^n \mid n \geq 0\}$  is not regular

Assume  $A$  is regular with pumping length  $p$

Now we need to pick a string  $w \in A$  with length  $|w| \geq p$



## Example

Let's prove  $A = \{0^n 1^n \mid n \geq 0\}$  is not regular

Assume  $A$  is regular with pumping length  $p$

Now we need to pick a string  $w \in A$  with length  $|w| \geq p$

Let  $w = 0^p 1^p$  which has length  $2p \geq p$

Consider  $xyz = w$  such that  $|xy| \leq p$  and  $|y| > 0$

We got to choose  $w$ , but we don't get to choose  $x$ ,  $y$ , and  $z$

We have to consider all possible choices!

What are the possible values of  $x$ ,  $y$ , and  $z$ ?

## Example

Let's prove  $A = \{0^n 1^n \mid n \geq 0\}$  is not regular

Assume  $A$  is regular with pumping length  $p$

Now we need to pick a string  $w \in A$  with length  $|w| \geq p$

Let  $w = 0^p 1^p$  which has length  $2p \geq p$

Consider  $xyz = w$  such that  $|xy| \leq p$  and  $|y| > 0$

We got to choose  $w$ , but we don't get to choose  $x$ ,  $y$ , and  $z$

We have to consider all possible choices!

What are the possible values of  $x$ ,  $y$ , and  $z$ ?

$x$  and  $y$  consist solely of 0s and  $z$  has the rest of the  $p$  0s followed by  $p$  1s:

$x = 0^m$ ,  $y = 0^n$ ,  $z = 0^{p-m-n} 1^p$  where  $n > 0$

## Example

$x = 0^m$ ,  $y = 0^n$ ,  $z = 0^{p-m-n}1^p$  where  $n > 0$

Now we need to find an  $i \geq 0$  such that  $xy^iz \notin A$

What value of  $i$  should we choose?

## Example

$x = 0^m$ ,  $y = 0^n$ ,  $z = 0^{p-m-n}1^p$  where  $n > 0$

Now we need to find an  $i \geq 0$  such that  $xy^iz \notin A$

What value of  $i$  should we choose?

In this case any  $i \neq 1$  works, so let's go with  $i = 0$  ("pumping down")

$$xy^0z = xz = 0^{p-n}1^p$$

Since  $n > 0$ ,  $p - n \neq p$  so  $xy^0z \notin A$  and thus  $A$  is not regular □

## Palindromes are not regular

Prove  $B = \{w \mid w \in \{0, 1\}^* \text{ and } w = w^{\mathcal{R}}\}$  is not regular

## Palindromes are not regular

Prove  $B = \{w \mid w \in \{0, 1\}^* \text{ and } w = w^{\mathcal{R}}\}$  is not regular

Assume  $B$  is regular with pumping length  $p$

We need to pick  $w$ ; what should we pick?

## Palindromes are not regular

Prove  $B = \{w \mid w \in \{0, 1\}^* \text{ and } w = w^R\}$  is not regular

Assume  $B$  is regular with pumping length  $p$

We need to pick  $w$ ; what should we pick?

Let  $w = 0^p 10^p$

Thus,  $x = 0^m$ ,  $y = 0^n$ , and  $z = 0^{p-m-n} 10^p$

## Palindromes are not regular

Prove  $B = \{w \mid w \in \{0, 1\}^* \text{ and } w = w^R\}$  is not regular

Assume  $B$  is regular with pumping length  $p$

We need to pick  $w$ ; what should we pick?

Let  $w = 0^p 10^p$

Thus,  $x = 0^m$ ,  $y = 0^n$ , and  $z = 0^{p-m-n} 10^p$

Let's "pump up" this time and try  $i = 2$

$$xy^2z = 0^{p+n} 10^p \notin B$$

Therefore,  $B$  is not regular





## Subsets of regular languages

True or false. If  $A$  is a regular language and  $B \subseteq A$ , then  $B$  is regular.

## Subsets of regular languages

True or false. If  $A$  is a regular language and  $B \subseteq A$ , then  $B$  is regular.

**FALSE!** Every language over  $\Sigma$  is a subset of  $\Sigma^*$  which is regular

## Subsets of regular languages

True or false. If  $A$  is a regular language and  $B \subseteq A$ , then  $B$  is regular.

**FALSE!** Every language over  $\Sigma$  is a subset of  $\Sigma^*$  which is regular

What's wrong with this argument?

Since  $A$  is regular, there is a DFA  $M$  that recognizes  $A$

Since  $B \subseteq A$ ,  $M$  accepts every string in  $B$  so  $B$  is regular

# Subsets of regular languages

True or false. If  $A$  is a regular language and  $B \subseteq A$ , then  $B$  is regular.

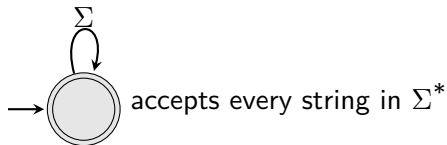
**FALSE!** Every language over  $\Sigma$  is a subset of  $\Sigma^*$  which is regular

What's wrong with this argument?

Since  $A$  is regular, there is a DFA  $M$  that recognizes  $A$

Since  $B \subseteq A$ ,  $M$  accepts every string in  $B$  so  $B$  is regular

It's missing the fact that for  $B$  to be regular there needs to be a DFA  $M'$  that accepts every string in  $B$  and rejects every string not in  $B$



## More nonregular languages

- $C = \{0^m 1^n 0^m \mid m, n \geq 0\}$
- $D = \{0^m 1^n \mid m \leq n\}$
- $E = \{w \mid w \in \{0, 1\}^* \text{ and } w \text{ has the same number of 0s and 1s}\}$