

The same would obviously apply to P

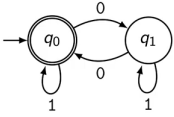
Myhill-Nerode

Adding a word to a word, to make it compatible with a language

$L(w) = \{w' | ww' \in L\}$  including:  $L(\epsilon)!$

w	L(w)	Q
$\epsilon$	$L(\epsilon) = L$	$q_0$
0	$L(0) = \{w \in \Sigma^* \mid  w _0 \text{ uneven}\}$	$q_1$
1	$L(1) = \{w \in \Sigma^* \mid  w _0 \text{ even}\} = L$	$q_0$
$\vdots$	$\vdots$	$\vdots$

even and uneven amounts of 0s  
mod 2 zero's, 1's don't matter



Detecting Nonregular Languages with Myhill

The examples before always had a specific amount of words/characters that one had to add, in order to accept the word.

However, there are languages that would need infinite states

in order to find the entire language of a DFA

A good example for this is the language  $1^n 0^n$

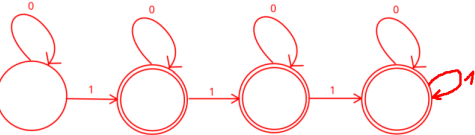
w	L(w)	Q
$\epsilon$	$\{0^n 1^n \mid n \geq 0\}$	$q_0$
0	$\{0^{n+1} 1^n \mid n \geq 0\}$	$q_1$
00	$\{0^{n+2} 1^n \mid n \geq 0\}$	$q_2$
000	$\{0^{n+3} 1^n \mid n \geq 0\}$	$q_3$
$0^k$	$\{0^{n+k} 1^n \mid n \geq 0\}$	$q_k$
$\vdots$	$\vdots$	$\vdots$
01	$\{\epsilon\}$	
001	$\{1\}$	
0001	$\{11\}$	
$\vdots$	$\vdots$	$\vdots$
1	$\emptyset$	e
10	$\emptyset$	e
$\vdots$	$\vdots$	$\vdots$

for every 0 that we add, we need a 1  
this means that for n+k 0's we need k 1's  
as  $\lim_{k \rightarrow \infty}$  we need  $\infty$  states!  
not possible with a **Deterministic** automaton!

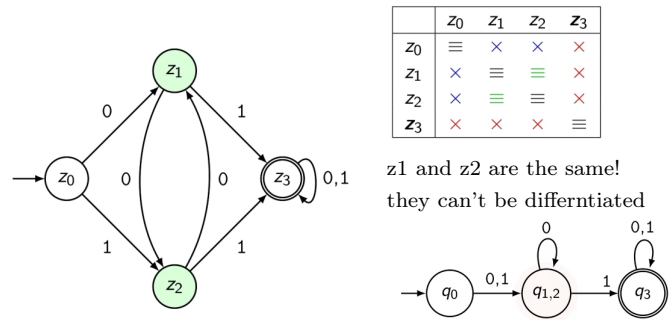
also note: we have clear error states  
anything starting with 1 is an error.

Differentiation of States

To get a minimal DFA, we eliminate all states that are superfluous.



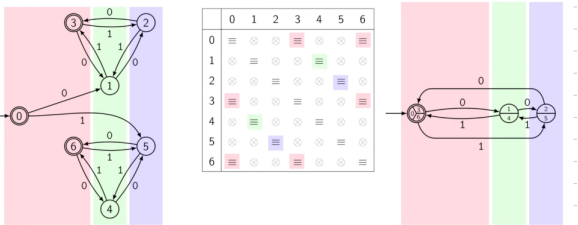
Here the 3 acceptable states can be put together, they are the same!



New minimal Automaton!

This algorithm makes it easy to see whether or not states are the same!

$L = \{w \in \{0, 1\}^* \mid |w|_0 \equiv |w|_1 \pmod 3\}$        $L = \{w \in \{0, 1\}^* \mid |w|_0 \equiv |w|_1 \pmod 3\}$



Beauty of a language / Pumping Lemma

a language L can be pumped if the following is valid:

$\exists N > 0$  where  $w \in L \wedge |w| \geq N$

If this word can be divided into 3 parts x,y,z while:

$|xy| \leq N \wedge |y| > 0 \wedge |x| \geq 0 \wedge |z| \geq 0 \wedge xy^kz \in L \rightarrow \forall k \in \mathbb{N}$

Find a number N that is bigger than 0 but less than the length of xy

while the length of y is greater than 0 and  $xy^kz$  is still part of the language

Note: The power of  $xy^kz$  is NOT a power, but an indicator how many y's!!

Any language that can be pumped is a regular language, and is therefore "schön"

Consider the following 2 examples:

$L(s1) = \{0, 1\}^*$  ending with 1

$z = 1, xy =$  any combination of 1 and 0

try:  $x=0, y=1010111, z=1, \text{True} \rightarrow N \geq 7$

try:  $x=0, y=1, z=1, \text{True} \rightarrow N \geq 1$

This can be done with any y, x

It will always satisfy all requirements

of the pumping lemma.

this language is regular.

$L(s1) = \{0^n, 1^n \mid n \geq 0\}$

if  $z=1$ , more 1's than 0's, FALSE

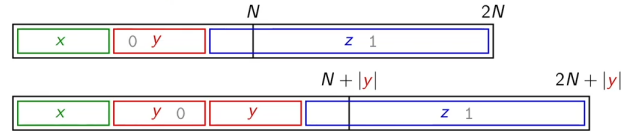
if  $x=0$ , more 0's than 1's, FALSE

This means we need  $x=0, z=0$

If we do that, then y is empty

No matter what we do, it is FALSE

$0^n 1^n$  is therefore not regular



Pumping Lemma usage guide !!FOLLOW THIS!!

1. Claim that L is regular

2. According to pumping lemma  $\exists N$

Don't make claims about the size of N!

3. Choose a word  $w \in L \mid |w| \geq N$

Definition with N has to be written!

4. Division into parts according to Pumping Lemma

$w = xyz, |xy| \leq N, |y| > 0, \text{etc}$

5. Check if word is in language

min 1 word not in language:  $xy^kz \notin L \mid k \in \mathbb{N}$

Explain why this word is not in the language

6. Contradiction, aka explain that this language is not regular.

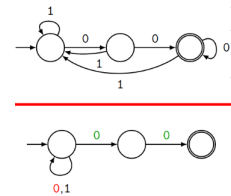
Non-Deterministic Finite Automaton (NFA / NEA)

Before every step was clear, there was no other determinism

other than the input and the current state. A non-deterministic automaton

Can have other things, like only accept the last 2 0's

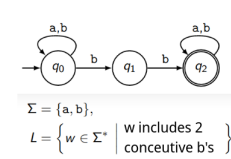
This is illustrated as both a DFA and an NFA:



Both have the same goal, only the last 2 zero's lead to the acceptance state.

However one is obviously easier, while not giving clear info on in what state it is, it is hence non-deterministic.

Formal Definition:



Machine A:  $\{Q, \Sigma, \delta, q_0, F\}$

State =  $Q \rightarrow \{q_1, q_2, \dots, q_n\}$

Alphabet =  $\Sigma$

Transitioning-Function =  $\delta : Q \times \Sigma \rightarrow P(Q)$

Starting State =  $q_0 = L(\epsilon) = L$

Acceptable Endstates =  $F \subset Q$

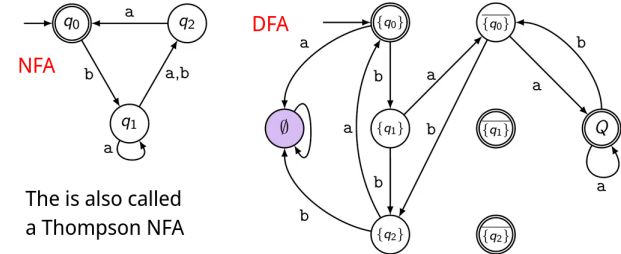
P(Q) is the Potence Quantity!!

Note the P(Q), it means that we have more complex transitions!

no arrow, multiple arrows for a certain character. See b in example

General tipp for NFA, only write what you need to accept the word!

NFA To DFA



The is also called a Thompson NFA

$\bar{q}_1$  is the complement Quantity. It means, any state other than  $q_1$

Q is the full Quantity, in here the NFA can be in any state.

$\emptyset$  is the error state.

## Formal Definition of the Transition

DFA is marked with ', the regular expression is for NFA

Given  $\delta$  of an NFA and Transitions  $M \subset Q$

$\delta' : Q \times \Sigma \rightarrow P(Q) : (M, a) \mapsto \delta'(M, a) = \bigcup_{q \in M} \delta(q, a)$

$Q' = P(Q)$

$\Sigma' = \Sigma$

$q'_0 = \{q_0\}$

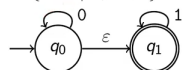
$F' = \{M \in P(Q) | F \cap M \neq \emptyset\}$

M is the union of all possible endstates.

Note: a Thompson NFA needs at least 1 acceptable endstate!

## $\varepsilon$ Transitions in NFA's

$L = \{0^k 1^l \mid k, l \geq 0\}$



$\delta : Q \times (\Sigma \cup \{\varepsilon\}) \rightarrow P(Q)$

This means that  $\varepsilon$  signifies a transition without using a character!!

## Conversion from $\varepsilon$ -NFA to regular NFA

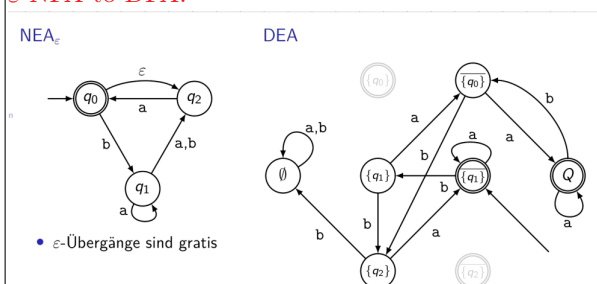
Any  $\varepsilon$ -NFA can be converted to a regular NFA!

$E(q)$  = Quantity of all  $\varepsilon$ -Transitions from q

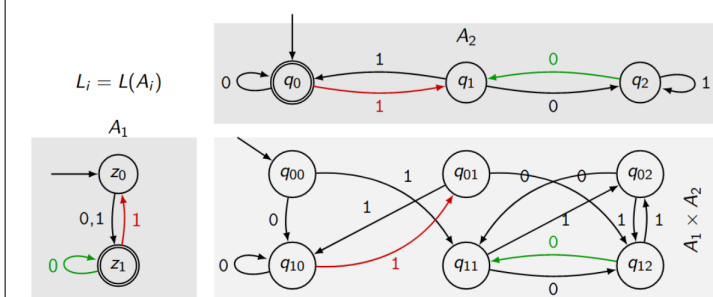
$E(M) = \bigcup_{q \in M} E(q)$  Quantity of all  $\varepsilon$ -Transitions

$\delta = Q \times (\Sigma \cup \{\varepsilon\}) \rightarrow P(Q) : (q, a) \mapsto E(\delta(q, a))$

## $\varepsilon$ -NFA to DFA:

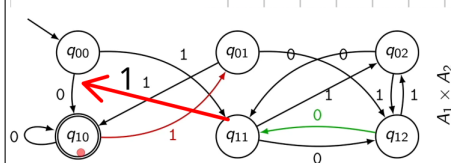


## Set-Operations with Automaton



## Intersection $L_1 \cap L_2$

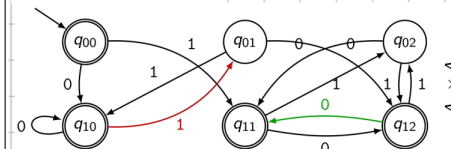
$F = F_1 \times F_2$



cartesian product -> acceptable endstate which both share

## Union $L_1 \cup L_2$

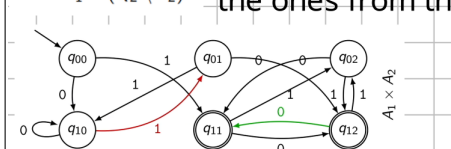
$F = F_1 \times Q_2 \cup Q_1 \times F_2$



acceptable endstate of both combined

## Difference $L_1 \setminus L_2$

$F = F_1 \times (Q_2 \setminus F_2)$



acceptable endstate of one without the ones from the other

## Pump-able, but not regular

$$L = \{a^i b^j c^k \mid i = 0 \vee j = k\} = \underbrace{\{b^j c^k \mid j, k \geq 0\}}_{= L_1} \cup \underbrace{\{a^i b^j c^k \mid i > 0 \wedge j = k\}}_{= L_2}$$

The first part L1 is regular, but the second isn't

The only way we can figure that out is by using the Myhill method!

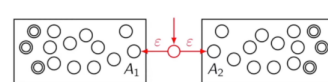
Because L2 is not regular, the composite Language L is not regular

## Regular Operations

### Alternative

$L = L_1 \cup L_2$

$= L(A_1) \cup L(A_2)$

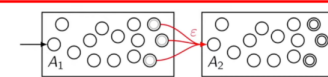


if or branching statement

### Chaining

$L = L_1 L_2 = L(A_1) L(A_2)$

$= \{w_1 w_2 \mid w_1 \in L_1\}$

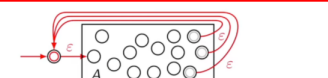


one after another

### \*-Operation

$L^* = \{\varepsilon\} \cup L \cup L^2 \cup \dots$

$= \bigcup_{k=0}^{\infty} L^k$

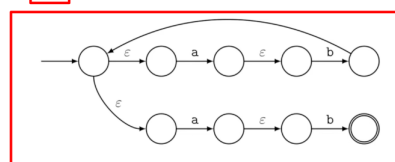


loop

## Regular Expressions

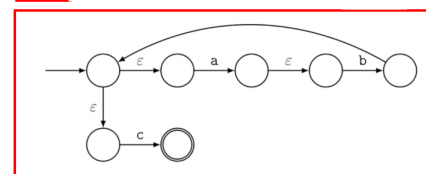
$(ab)^+$

do while



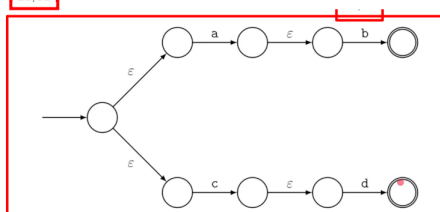
$(ab)^*c$

chain



$ab|cd$

alternative



## Regular expression

String r to describe a language

$L = L(r)$

## Regular Operations

$L(r_1) \cup L(r_2) = L(r_1 | r_2)$  Alternative

$L(r_1) L(r_2) = L(r_1 r_2)$  Chaining

$L(r_1)^* = L(r_1^*)$  \*-Operation

## Primitive regular expressions

expressions for Words with length  $\leq 1$

$L = L(r)$	$r$	NEA
$\emptyset$	$\emptyset$	
$\{\varepsilon\}$	$\varepsilon$	
$\{a\}$	$a$	
$\{0, s, t\}$	$[ost]$	
$\{a, b, \dots, s\}$	$[a-s]$	
$\Sigma$	$.$	

## Contractions

$$\begin{aligned} r+ &= rr^* & r? &= \varepsilon | r \\ r\{2\} &= rr & r\{2,3\} &= rr | rrr \\ r\{3\} &= |r|rr|rrr & r\{3,\} &= rrrr^* \end{aligned}$$

## Generalized NFA/NEA

NEA<sub>ε</sub>, that has it's transitions described NFA<sub>ε</sub>, as regular expressions

There is a DFA for every regular expression!
