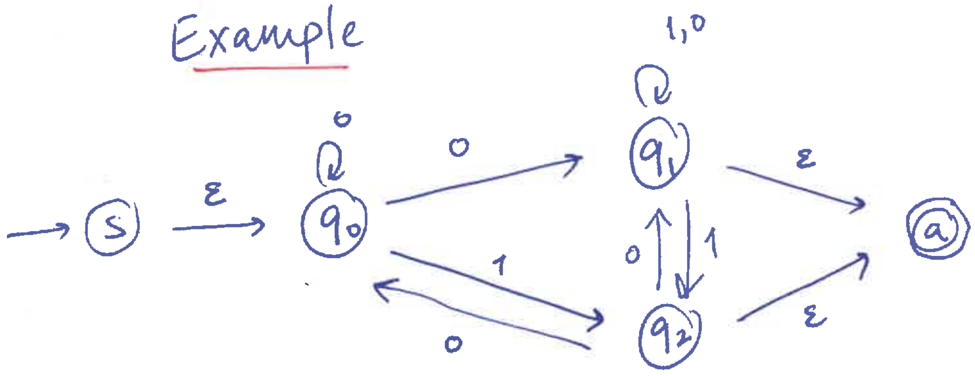
MATH 2301

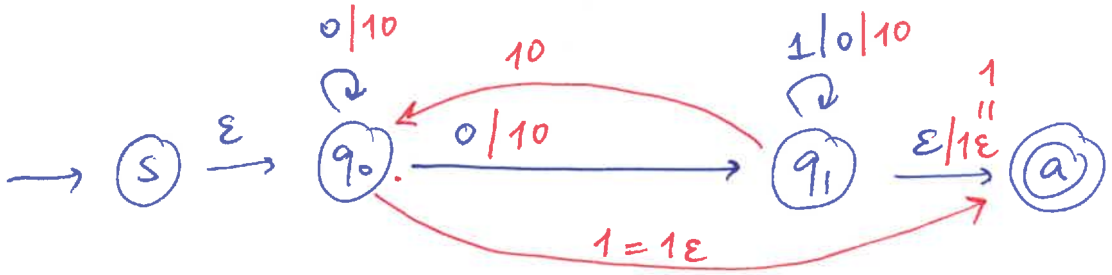
10/10/2022

* **Goal**: NFA (or DFA) $\longrightarrow$ Equivalent regex

Example



(Remove q_2 from this)
↓



Note: At each step, we'll have a machine whose arrow labels are regexes.

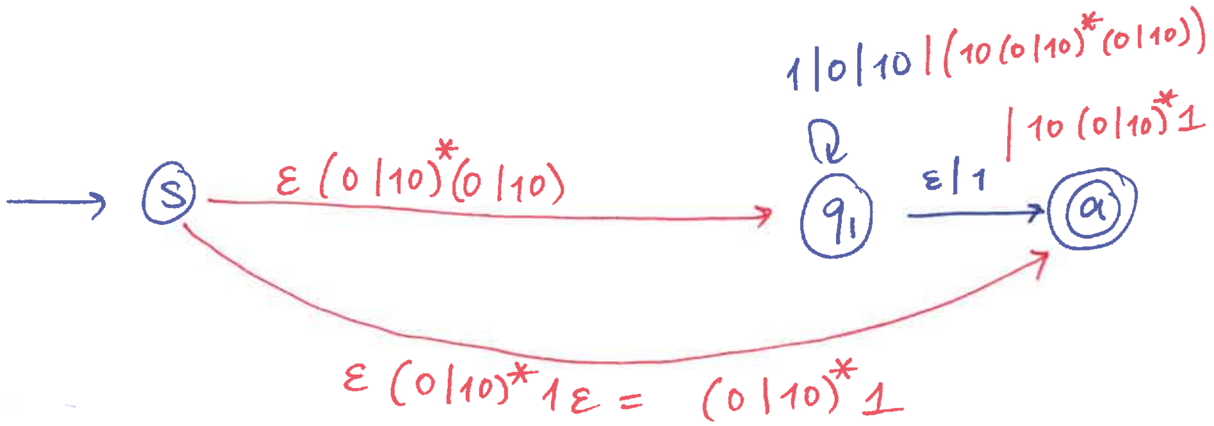
Recap: If removing state q, look at

(1) Every pair of incoming + outgoing arrow,

e.g. $x \xrightarrow{r_1} q \xrightarrow{r_2} y$ $\quad [x \neq q,\ y \neq q]$

If q doesn't have a self-loop, throw in an arrow labelled $r_1 r_2$ from x to y. If there is already an arrow from x to y, then add "$|r_1 r_2$" to its label.

Next, let us remove q_0



Recap

(2) Again, look at

$$\begin{tikzcd} x \arrow[r, "r_1"] & q \arrow[loop above, "r"] \arrow[r, "r_2"] & y \end{tikzcd} \qquad x \neq q \text{ and } y \neq q.$$

Suppose q has a self-loop labelled r.

Add in an arrow (or a label) from x to y labelled $r_1 (r^*) r_2$.

Finally, in our example, we remove q_1.

(Next page)

$$\rightarrow (s) \xrightarrow{((0|10)^*1)\,|\,((0|10)^*(0|10)(1|0|10|(10(0|10)^*(0|10)))^*(\varepsilon|1|10(0|10)^*1))} ((a))$$

The final label from s to a is a regex whose language is exactly the language of the original NFA.

** Summary [Theorem]

1. Delete every state between (s) and ((a)) successively, updating labels after each deletion.
2. Label updates are as written in the "Recap"s.
3. Keep going until there is one arrow from (s) to ((a)). The label on that arrow is an equivalent regex.

** Def: We say that a language L is regular if any of the following equivalent conditions hold:

(1) There is a regex r such that $L = L(r)$.
(2) There is a NFA M such that $L = L(M)$
(3) There is a DFA M' such that $L = L(M')$

** Warning: Not every language is regular!

Example : $L = \{ 0^n 1^n \mid n \geq 0 \}$ is not regular
(Claim)

Aside : Why should we expect non-regular languages to exist?

One answer : Countability / counting.

regexes for a given alphabet is countable.
(can be enumerated in an infinite list)

$\Rightarrow$ # regular languages is also countable.

On the other hand, # of languages (not nec. regular) is the number of subsets of Σ^*.

Thm : The size of the power set of Σ^* is not countable

$\Rightarrow \exists$ non-regular languages.

Second answer : The pumping lemma $\rightarrow$ Wednesday