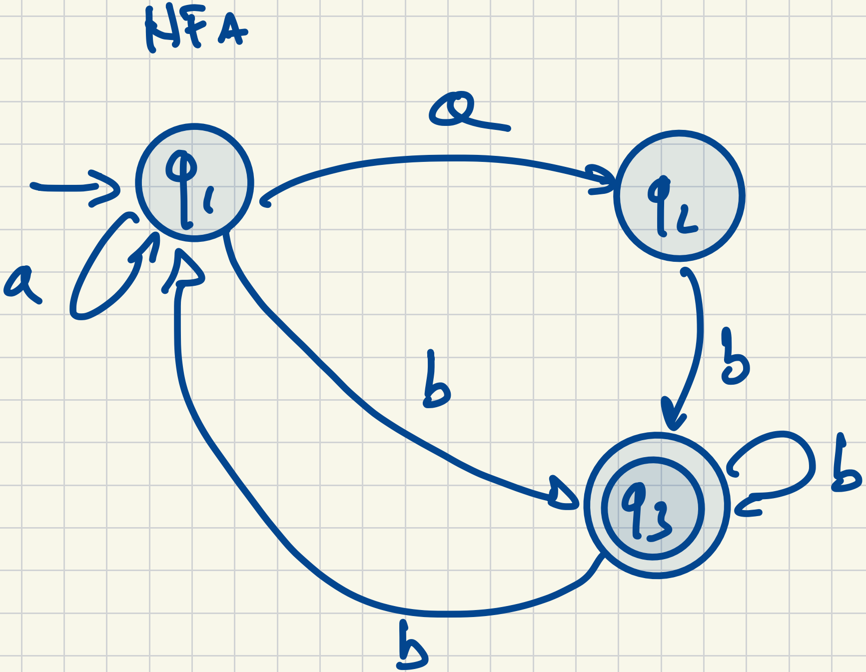
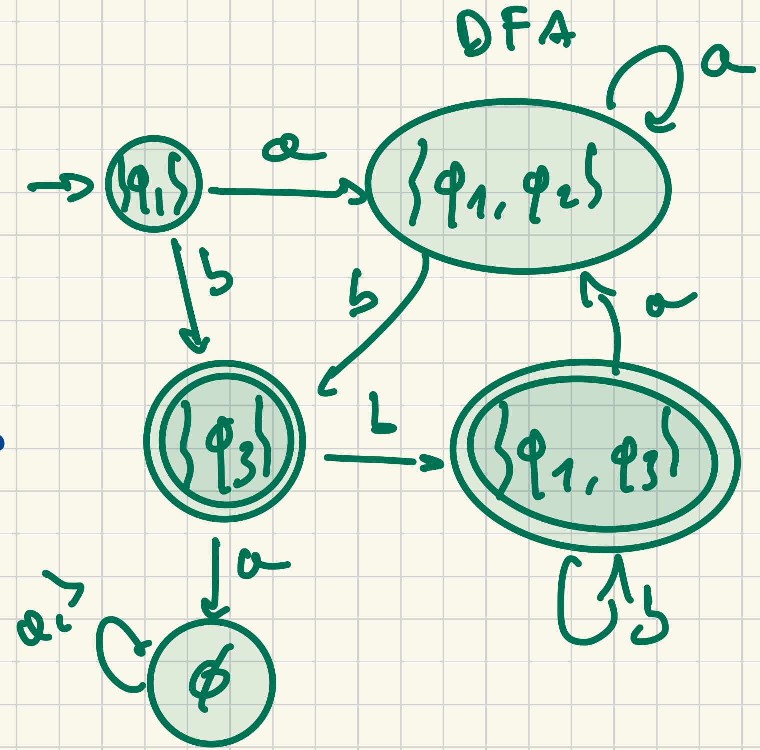
Sept, 24th

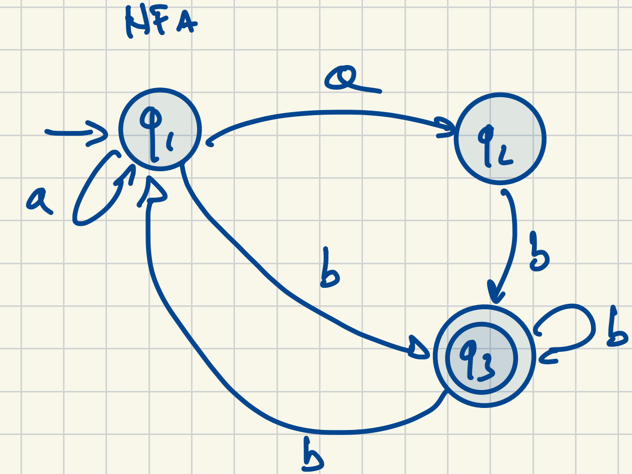
Kleene's theorem

$$\begin{tikzcd}
 & \text{Reg. language} & \\
 & \text{Reg. expression} \arrow[u, Leftrightarrow] \arrow[dr] & \\
\text{DFA} \arrow[ur] & & \varepsilon\text{-NFA} \arrow[ll] \arrow[dl, dashed] \\
 & \text{NFA} \arrow[ul] &
\end{tikzcd}$$

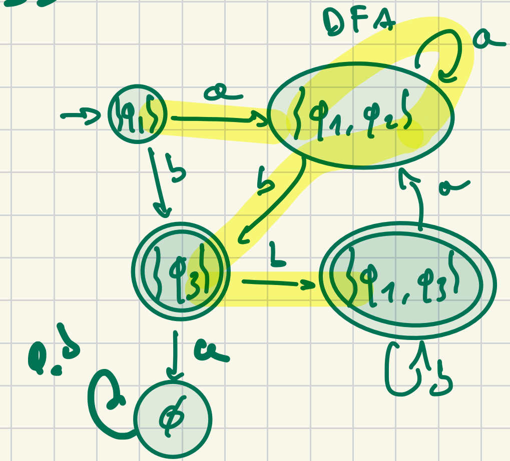
Determinization of NFAs

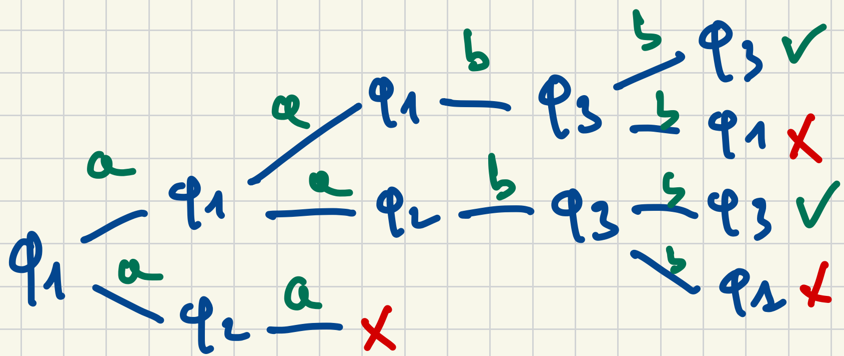






input: aabb



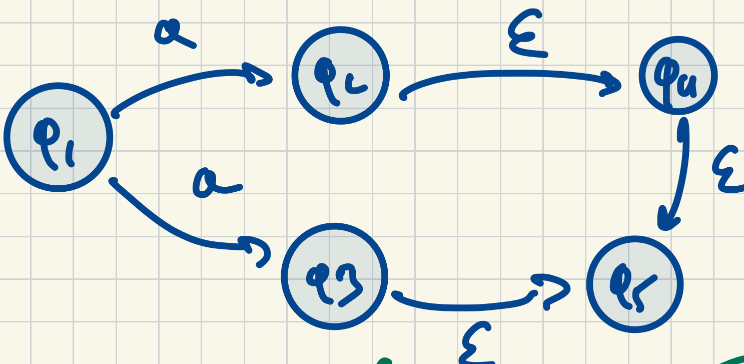


This is the run of the DFA!

$\{q_1\}$ $\{q_1, q_2\}$ $\{q_1, q_2\}$ $\{q_3\}$ $\{q_1, q_3\}$

How to deal with ε-transitions?

ε-closure (q) = all the states you can reach by taking 0 or more ε-transitions.



① $\delta(q_1, a) = \{q_2, q_3\}$

② ε-closure $(\{q_2, q_3\})$
$= \{q_2, q_3, q_4, q_5\}$

$$\{q_1\} \xrightarrow{a} \{q_2, q_3, q_4, q_5\}$$

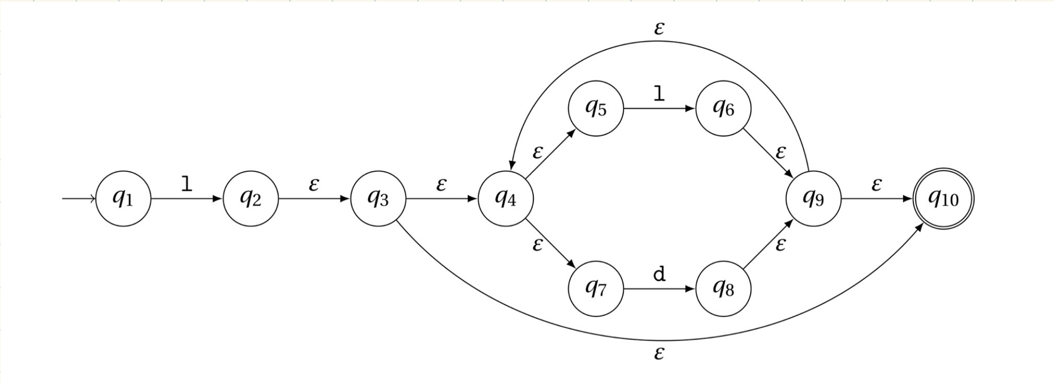


Figure 2.8: The ε-NFA built from the r

$l(l+d)^*$

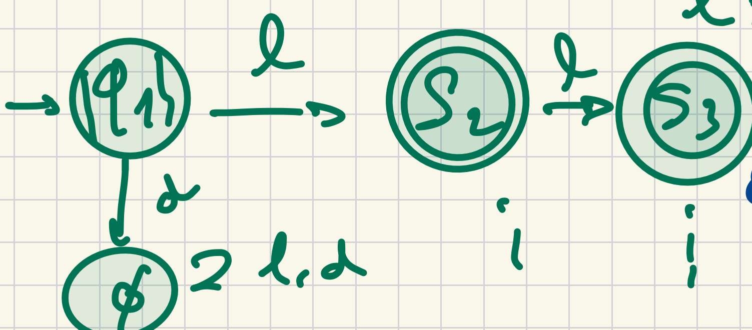
$\varepsilon\text{-closure}(\{q_1\}) = \{q_1\}$ = initial state

$\varepsilon\text{-closure}(\delta(\{q_1\}, l)) = \varepsilon\text{-closure}(\{q_2\}) =$

$\{q_2, q_3, q_4, q_5, q_7, q_{10}\} = S_2$

l:

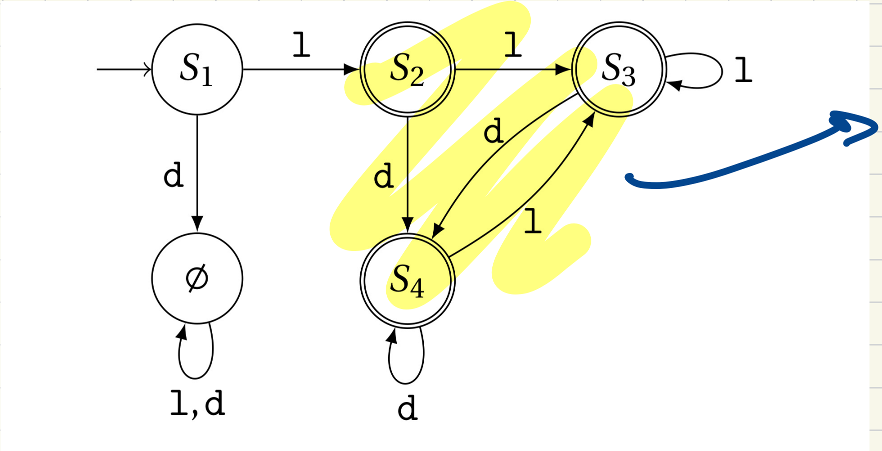
q_2	q_3	q_4	q_5	q_7	q_{10}
x	x	x	q_6	x	x



$\varepsilon\text{-closure}(\{q_6\}) =$

$\{q_6, q_4, q_5, q_7, q_9, q_{10}\} = S_3$

Resulting DFA:



Could be merged.

$l(l+d)^*$

Size of the DFA

The states of the DFA are subsets of the states of the ε-NFA.

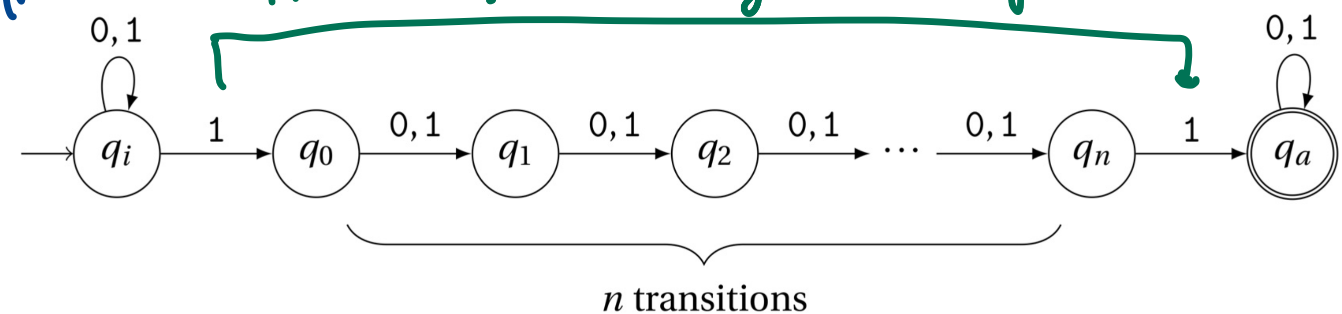
So, potentially there are exponentially many states in the DFA.

Can it happen for real? Yes!

We will introduce an infinite family of automata (NFA) A_n s.t. the deterministic counterpart has size 2^{n+1} $O(2^n)$

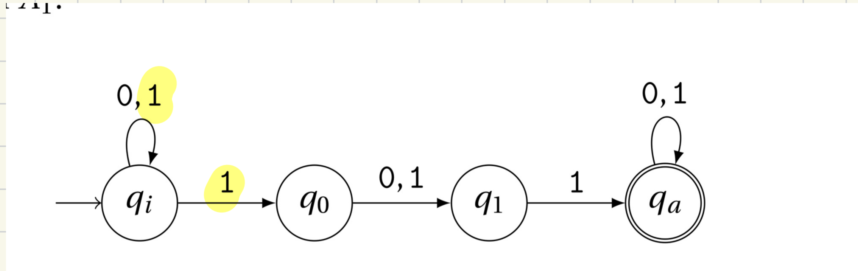
A_n

A '1' separated by n sets from another '1'



$n+3$ states
$= O(n)$

Example: A_1



0 1 1 0 1 0 ✓

If I remember the two last bits, I can check the word

$\underbrace{0 \quad 1}_{\text{Memory}} \quad 1 \quad 0 \quad 1 \quad 0$

$\to$ I read a 1 : I check the first bit of the memory. $= 0 \to$ I don't accept yet.

$0 \quad \underbrace{1 \quad 1} \quad 0 \quad 1 \quad 0$

$\to$ I read a 0 : I ignore.

$0 \quad 1 \quad \underbrace{1 \quad 0} \quad 1 \quad 0$

$\to$ I read a 1 : I check : ok ✓

Conclusion: for the automaton A_1

we need 2 bits of memory

———— A_2

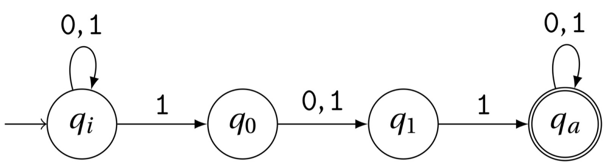
———— 3 bits ————

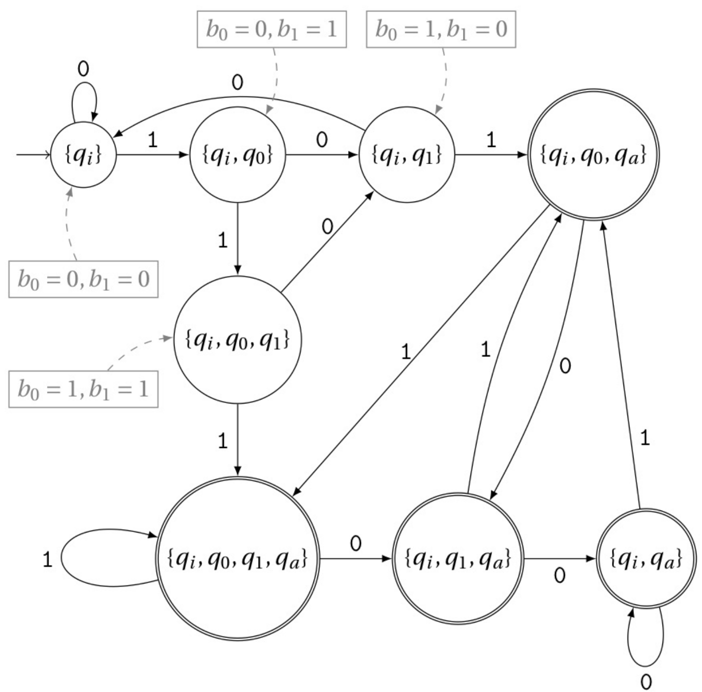
⋮

$A_n \longrightarrow n+1$ bits of memory.

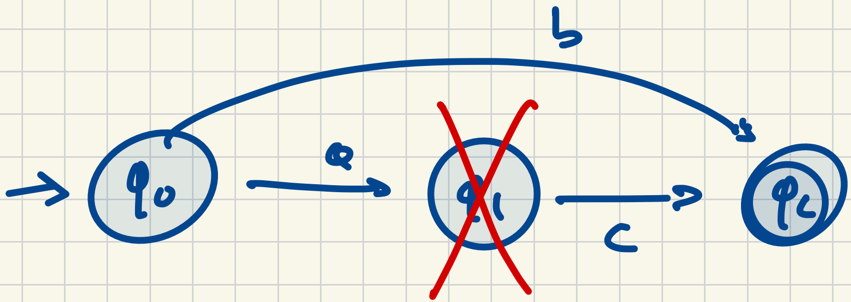
The content of the memory
will be stored in the slots!

$n+1$ bits of memory $\rightarrow$ you need at least
2^{n+1} slots.

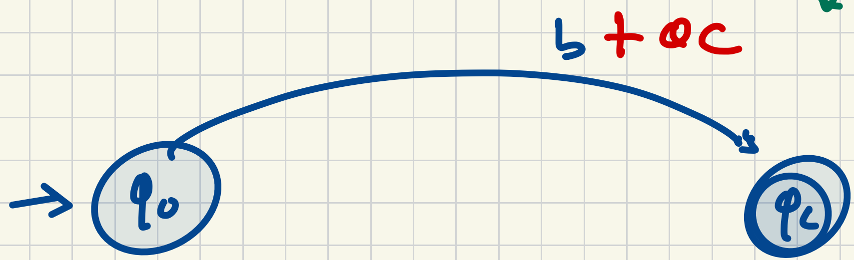

0, 1
1
0, 1
1
0, 1
q_i
q_0
q_1
q_a


$b_0 = 0, b_1 = 1$
$b_0 = 1, b_1 = 0$
$b_0 = 0, b_1 = 0$
$b_0 = 1, b_1 = 1$
$\{q_i\}$
$\{q_i, q_0\}$
$\{q_i, q_1\}$
$\{q_i, q_0, q_a\}$
$\{q_i, q_0, q_1\}$
$\{q_i, q_0, q_1, q_a\}$
$\{q_i, q_1, q_a\}$
$\{q_i, q_a\}$

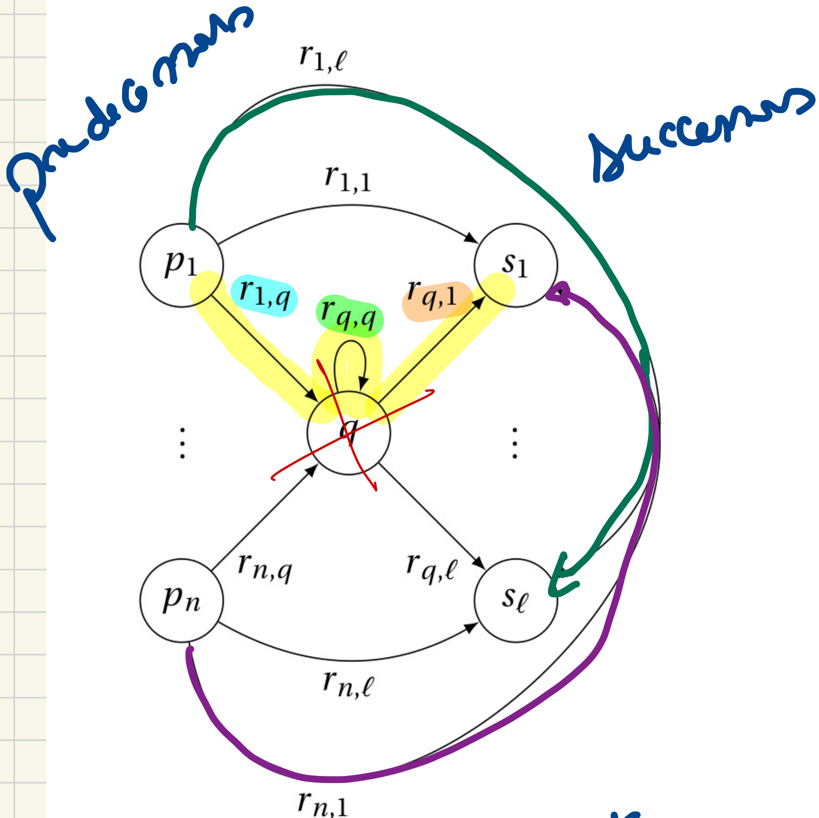
From FA to RE

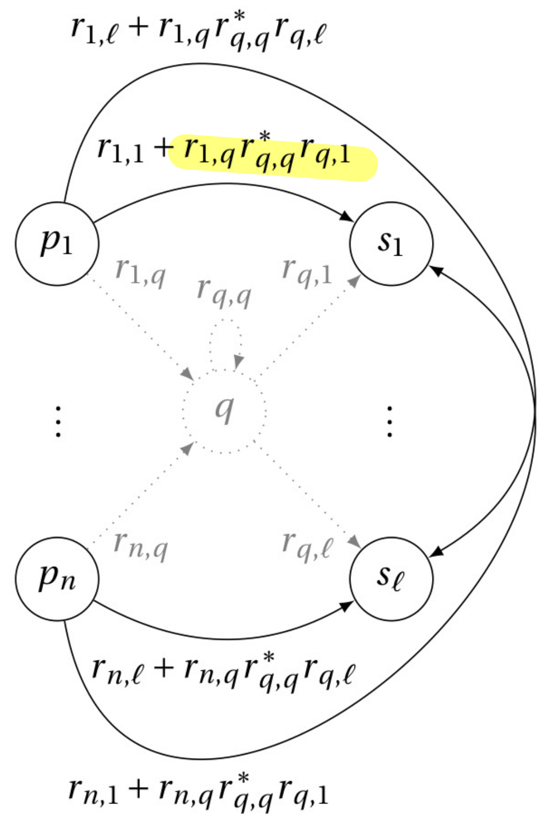


elimination of q_1



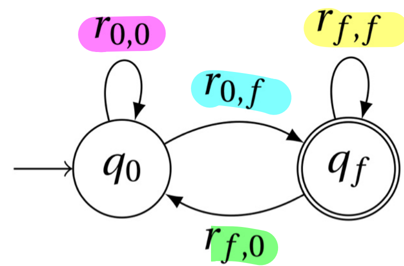
I want to get rid of q.

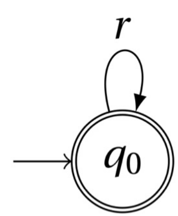




$= r_{1,q} \cdot (r_{q,q})^* \cdot r_{q,1}$

After removing states, we end up in one of the two following cases:



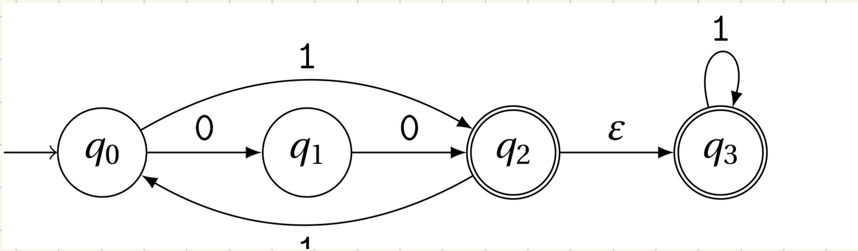


$$\underbrace{\left(r_{0,0} + r_{0,f} \cdot r_{f,f}^* \cdot r_{f,0}\right)^*}_{\text{going from to } q_0 \text{ to } q_0} \cdot r_{0,f} \cdot r_{f,f}^* \qquad\qquad r^*$$

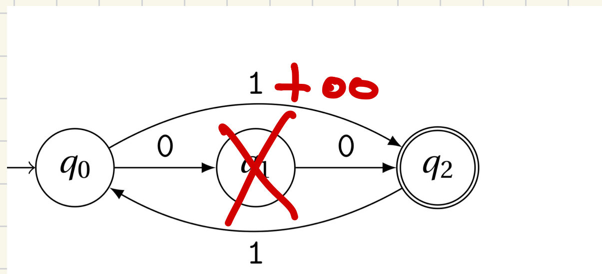
$r_{0,f}$: $q_0 \to q_f$

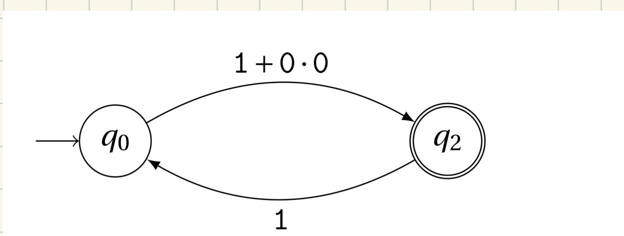
$r_{f,f}^*$: self-loop on q_f

We apply this for each accepting state.



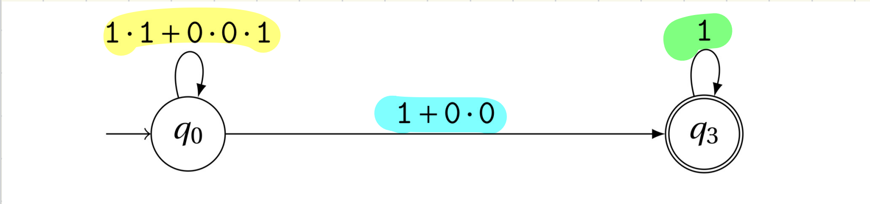
① For q_2:





$\Rightarrow (11 + 001)^* \cdot (1+00)$

For q_3:



$$(11 + 001)^* \cdot (1 + 00)\, 1^*$$

Final result.

$$(11 + 001)^* \cdot (1 + 00)$$

$$+$$

$$(11 + 001)^* \cdot (1 + 00)\, 1^*$$