Formal languages and automata

DFAs to regular expressions

https://t.me/fla_uog

mh.olyaee@gmail.com

Summary

- Regular Expression (RE) define regular sets
- RE $\Rightarrow$ NFA $\Rightarrow$ DFA

Proof Idea

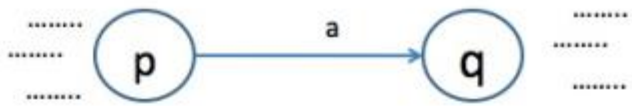
- Now we want to extract RE from NFA

PROOF IDEA

- Generalized transitions: label transitions with regular expressions
- Generalized NFAs (GNFA)
- Iteratively eliminate states of the GNFA one by one, until only two states and a single generalized transition is left.

Generalized transitions

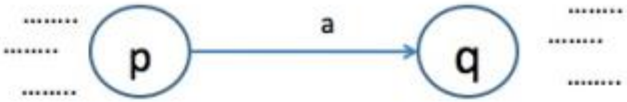
- DFAs have single symbols as transition labels



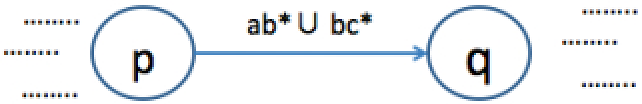
  - If you are in state *p* and the next input symbol matches *a*, go to state *q*

Generalized transitions

- DFAs have single symbols as transition labels



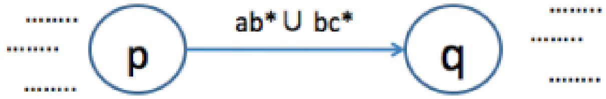
- Now consider



  - If you are in state p and a prefix of the remaining input matches the regular expression $\mathbf{ab}^* \cup \mathbf{bc}^*$ then go to state q

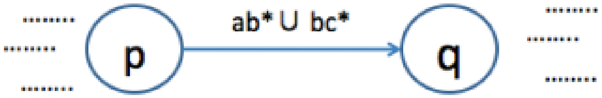
Generalized transitions

- A generalized transition is a transition whose label is a regular expression

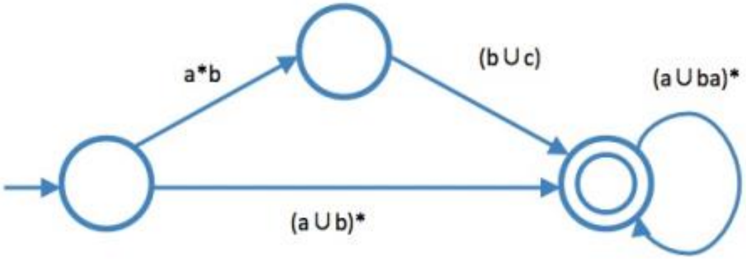


Generalized transitions

- A generalized transition is a transition whose label is a regular expression



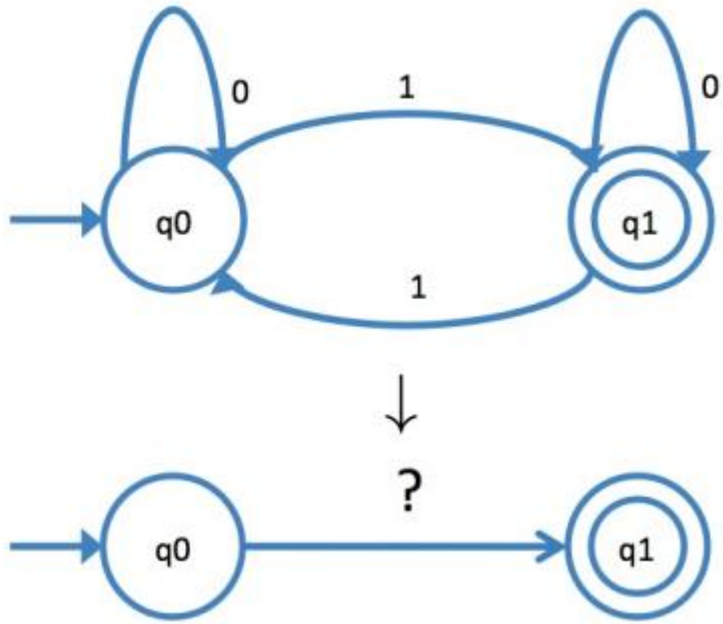
- A Generalized NFA is an NFA with generalized transitions.



  - In fact, all standard DFA transitions are generalized transitions with regular expressions of a single symbol!

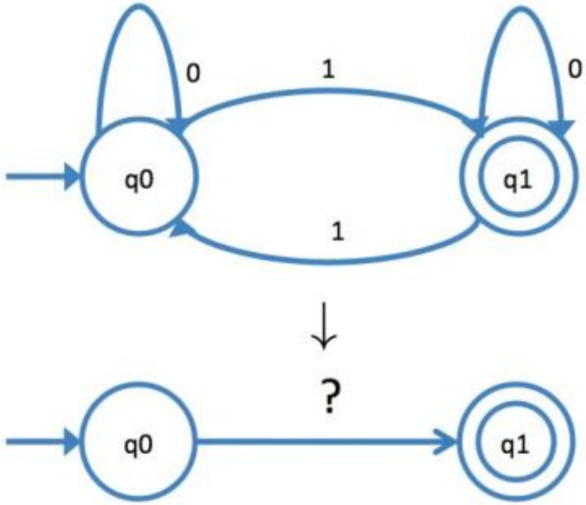
Generalized transitions

- Consider the 2-state DFA



Generalized transitions

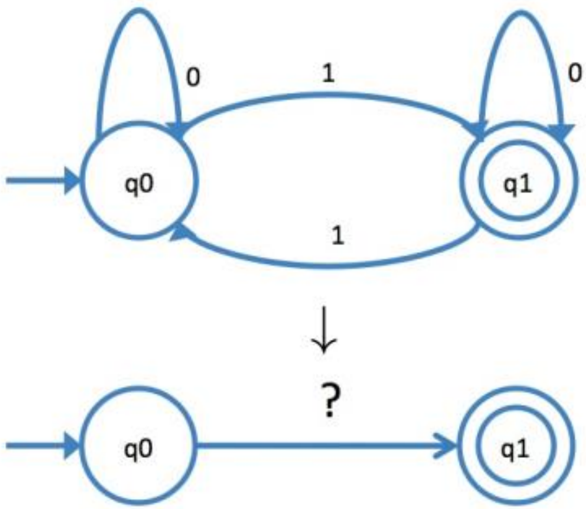
- Consider the 2-state DFA



- $\mathbf{0}^*\mathbf{1}$ takes the DFA from state q_0 to q_1

Generalized transitions

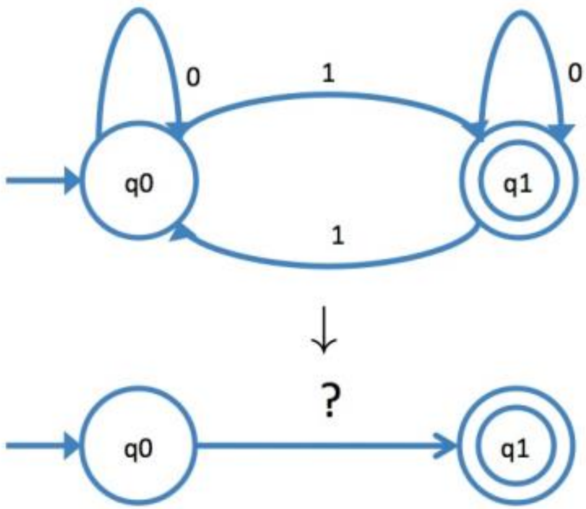
- Consider the 2-state DFA



- $\mathbf{0}^*\mathbf{1}$ takes the DFA from state q_0 to q_1
- $(\mathbf{0} \cup \mathbf{10}^*\mathbf{1})^*$ takes the machine from q_1 back to q_1

Generalized transitions

- Consider the 2-state DFA



- $\mathbf{0}^*\mathbf{1}$ takes the DFA from state q_0 to q_1
- $(\mathbf{0} \cup \mathbf{10}^*\mathbf{1})^*$ takes the machine from q_1 back to q_1
- So $\mathbf{?} = \mathbf{0}^*\mathbf{1}(\mathbf{0} \cup \mathbf{10}^*\mathbf{1})^*$ represents all strings that take the DFA from state q_0 to q_1

Generalized NFA

- Take any DFA and transform it into a GNFA
  - with only two states: one start and one accept
  - with one generalized transition
- then we can “read” the regular expression from the label of the generalized transition (as in the example above)

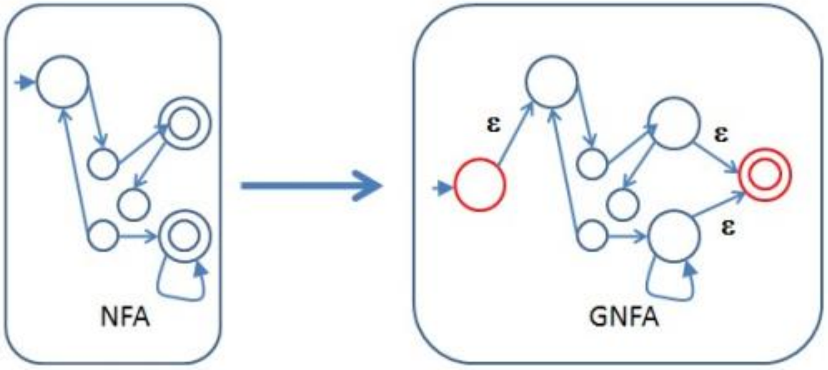
DFA to GNFA

- We will add two new states to a DFA:
  - A new start state with an ϵ-transition to the original start state, but with no transitions from any other state
  - A new final state with an ϵ-transition from all the original final states, but with no transitions to any other state

DFA to GNFA

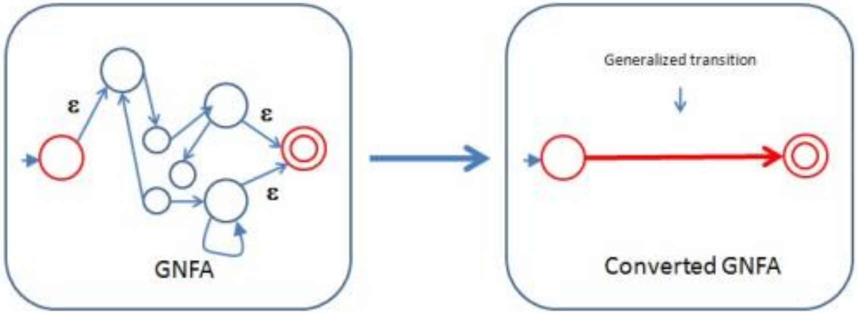
- We will add two new states to a DFA:
  - A new start state with an ϵ-transition to the original start state, but with no transitions from any other state
  - A new final state with an ϵ-transition from all the original final states, but with no transitions to any other state

- The previous start and final states are no longer!



Reducing GNFA

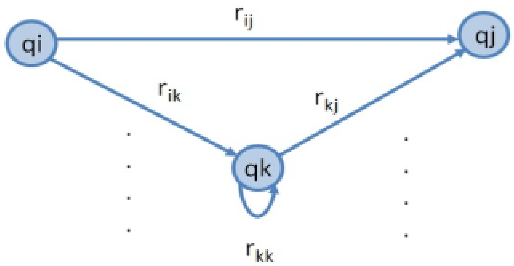
- We eliminate all states of the GNFA one-by-one leaving only the start state and the final state.



- When the GNFA is fully converted, the label of the only generalized transition is the regular expression for the language accepted by the original DFA.

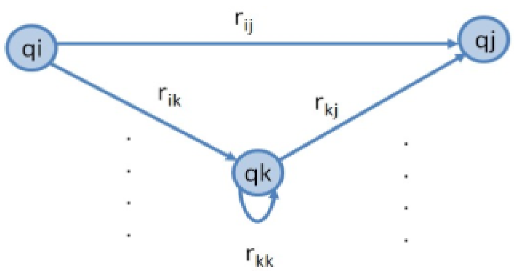
Eliminating states

- Suppose we want to eliminate state q_k, and q_i and q_j are two of the remaining states ($i = j$ is possible).



Eliminating states

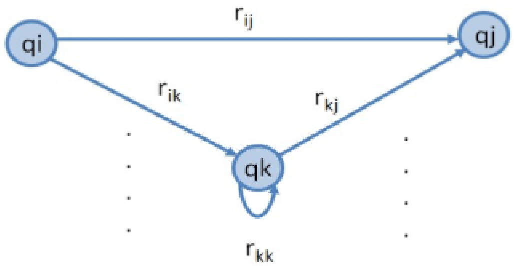
- Suppose we want to eliminate state q_k, and q_i and q_j are two of the remaining states ($i = j$ is possible).



- How can we modify the transition label between q_i and q_j to reflect the fact that q_k will no longer be there?

Eliminating states

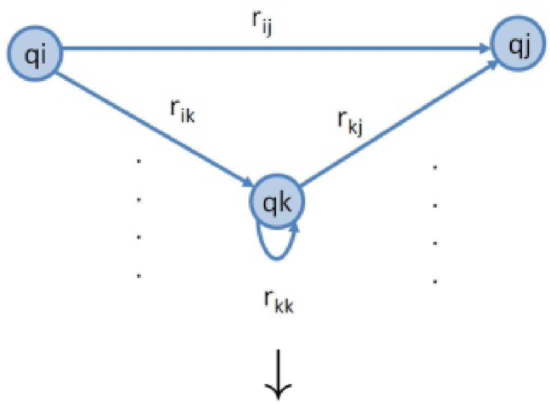
- Suppose we want to eliminate state q_k, and q_i and q_j are two of the remaining states ($i = j$ is possible).



- How can we modify the transition label between q_i and q_j to reflect the fact that q_k will no longer be there?
    - There are two paths between q_i and q_j
        - Direct path with regular expression $\mathbf{r_{ij}}$
        - Path via q_k with the regular expression $\mathbf{r_{ik} r_{kk}^* r_{kj}}$

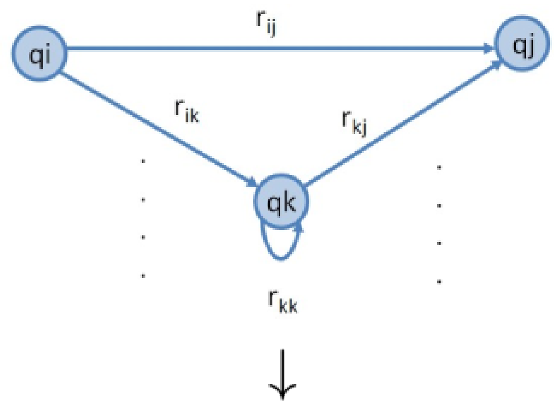
Eliminating states

- There are two paths between q_i and q_j
  - Direct path with regular expression $\mathbf{r_{ij}}$
  - Path via q_k with the regular expression $\mathbf{r_{ik} r_{kk}^* r_{kj}}$



Eliminating states

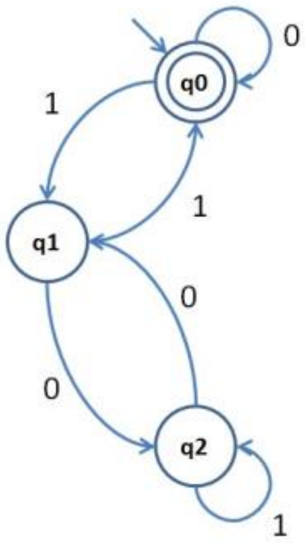
- There are two paths between q_i and q_j
  - Direct path with regular expression $\mathbf{r_{ij}}$
  - Path via q_k with the regular expression $\mathbf{r_{ik}r_{kk}^*r_{kj}}$



- After removing q_k, the new label would be $\mathbf{r'_{ij}} = \mathbf{r_{ij}} \cup \mathbf{r_{ik}r_{kk}^*r_{kj}}$

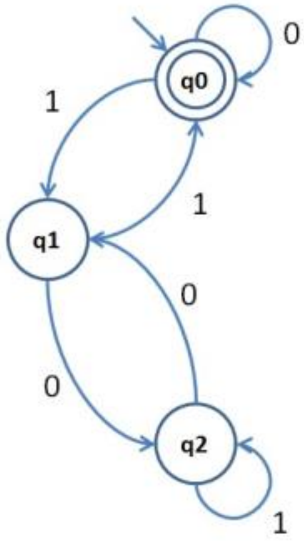
Example

- DFA for binary numbers divisible by 3

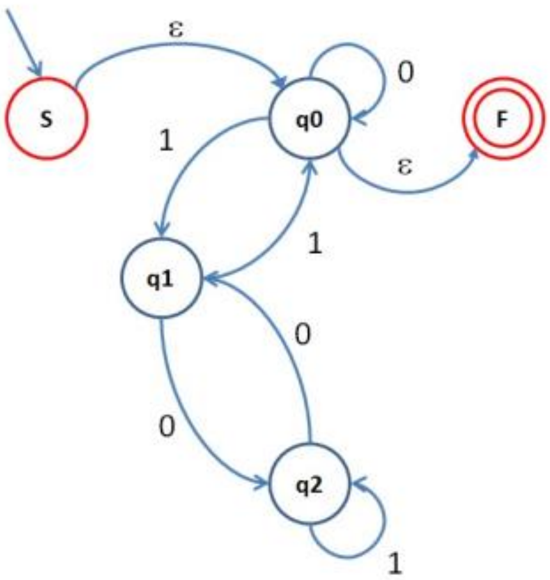


Example

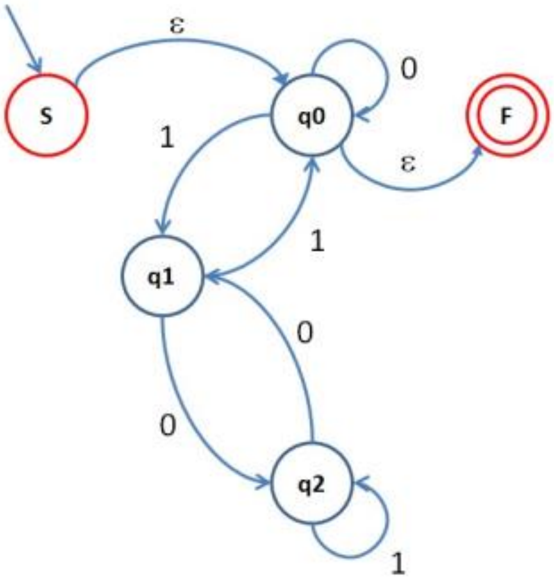
- DFA for binary numbers divisible by 3



- Initial GNFA

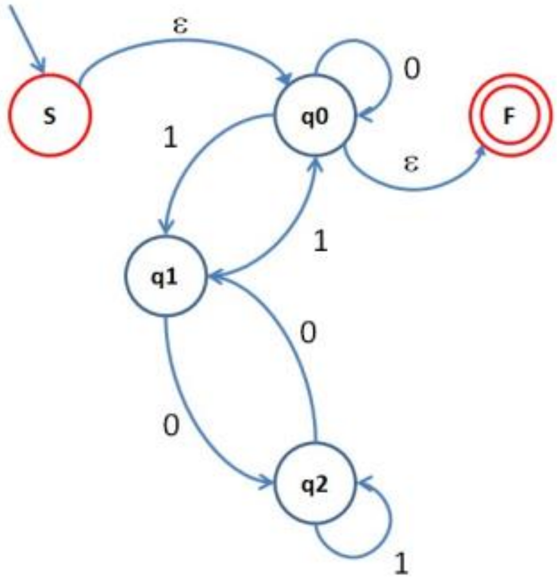


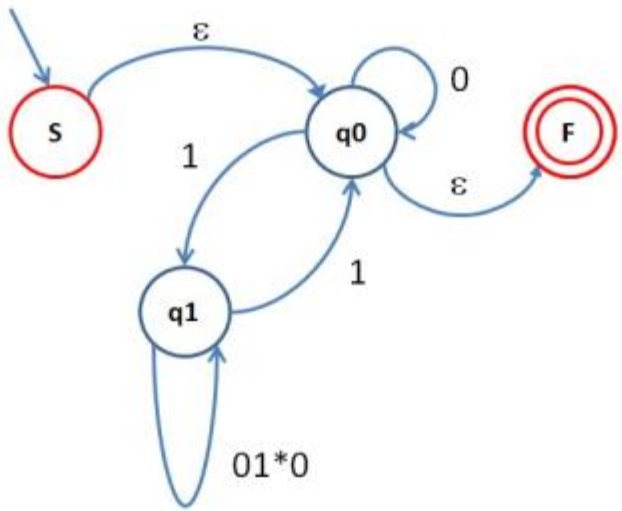
- Let's eliminate q_2



$$q_i = q_1, q_j = q_1, q_k = q_2$$

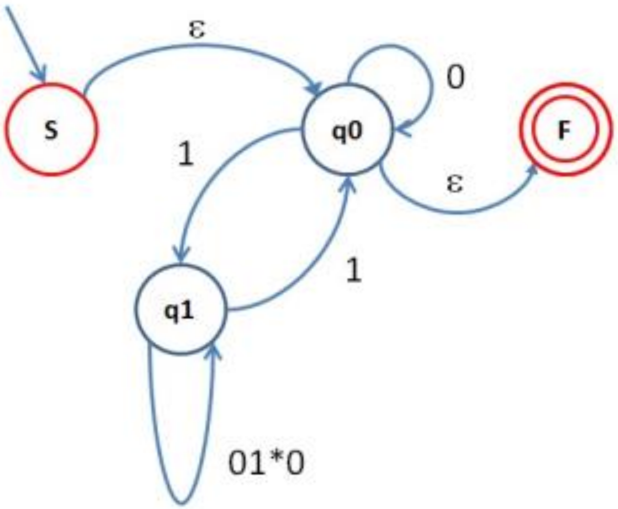
- Let's eliminate q_2



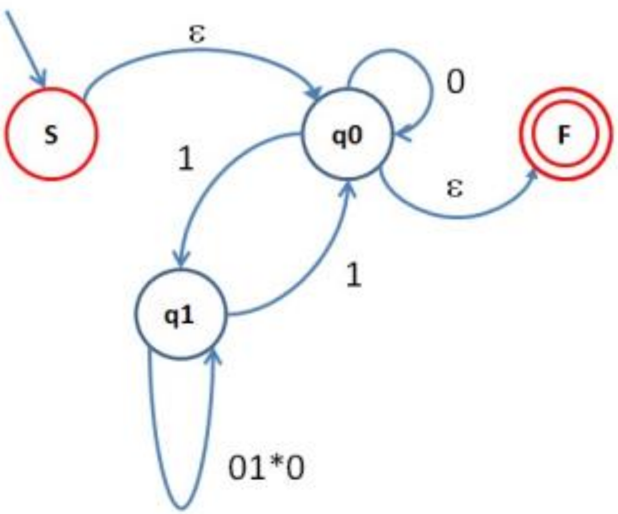


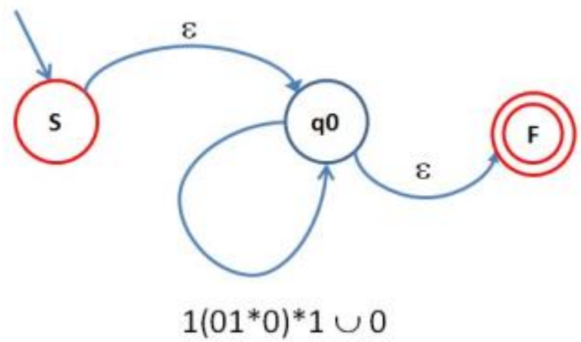
$$q_i = q_1, q_j = q_1, q_k = q_2$$

- Let's eliminate q_1

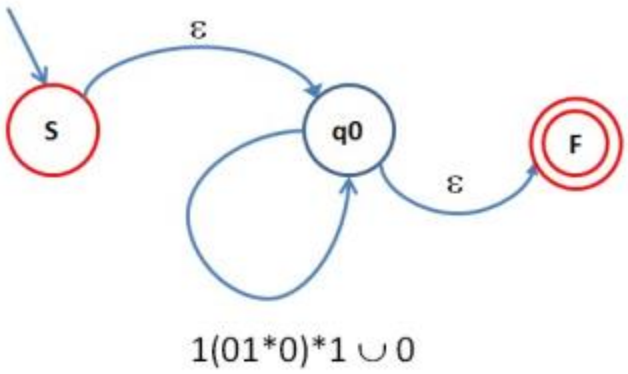


- Let's eliminate q_1

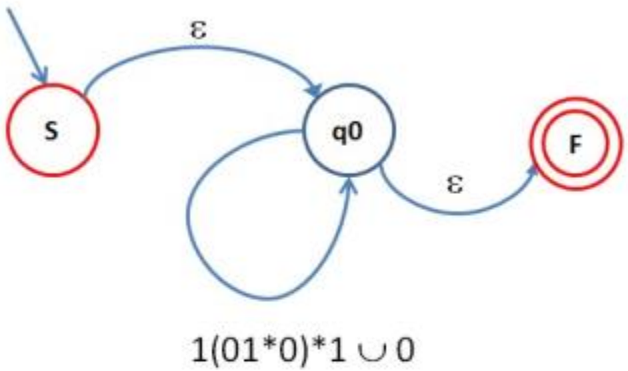


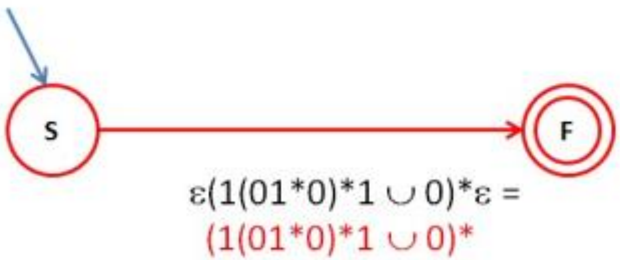


- Let's eliminate q_0



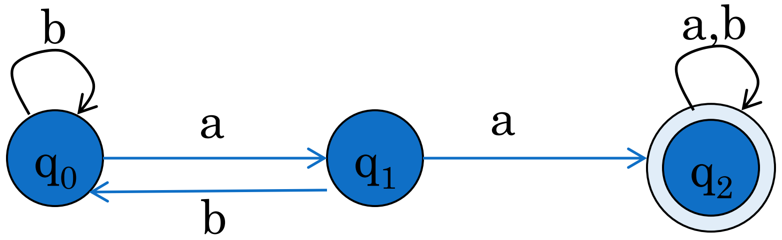
- Let's eliminate q_0



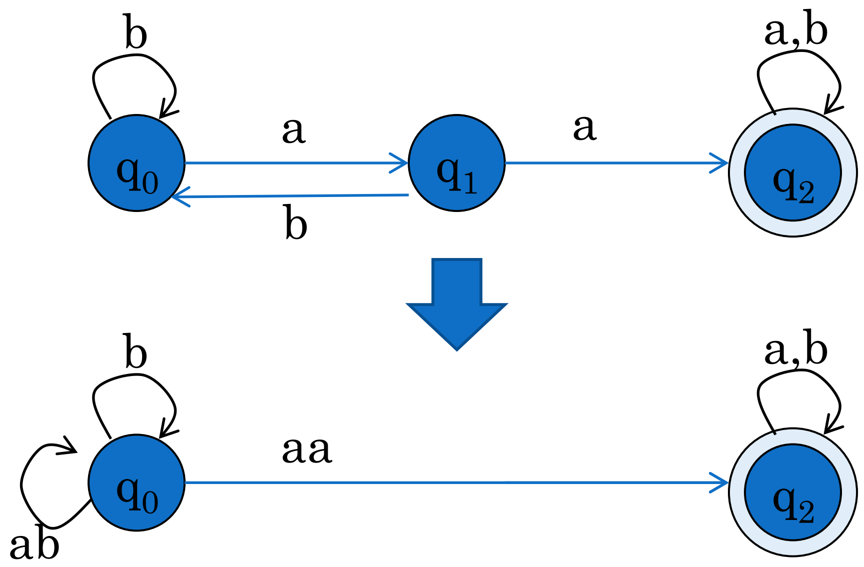


- So the regular expression we are looking for is $(\mathbf{1}(\mathbf{01}^*\mathbf{0})^*\mathbf{1} \cup \mathbf{0})^*$

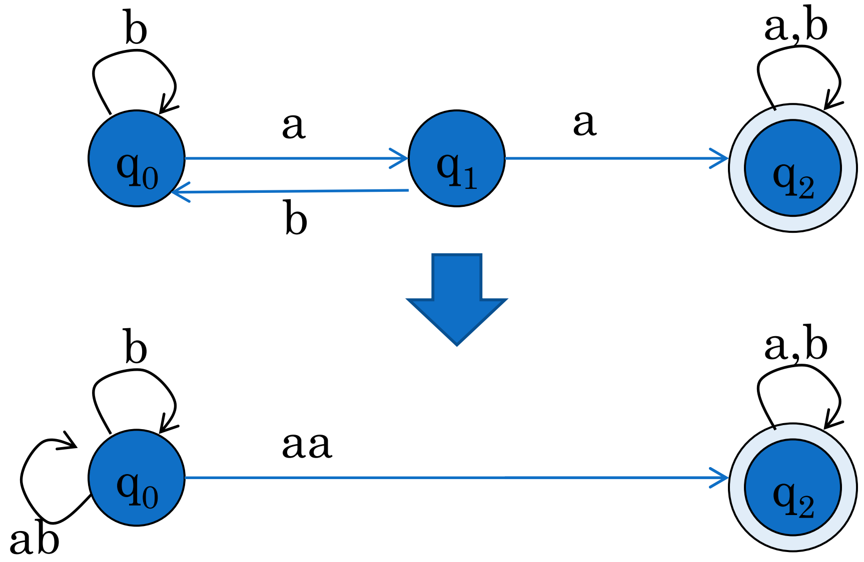
Example



Example

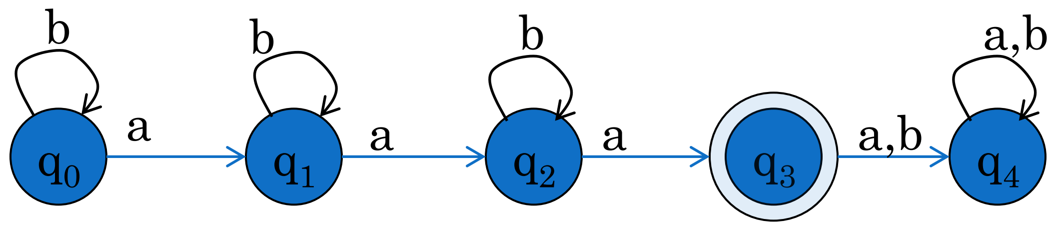


Example



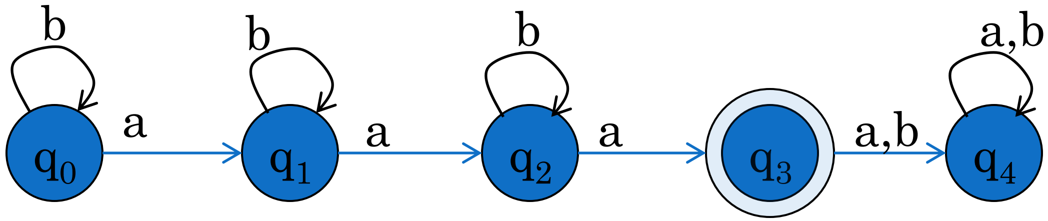
- $(ab + b)^{*}\mathrm{aa}(a + b)^{*}$

Example



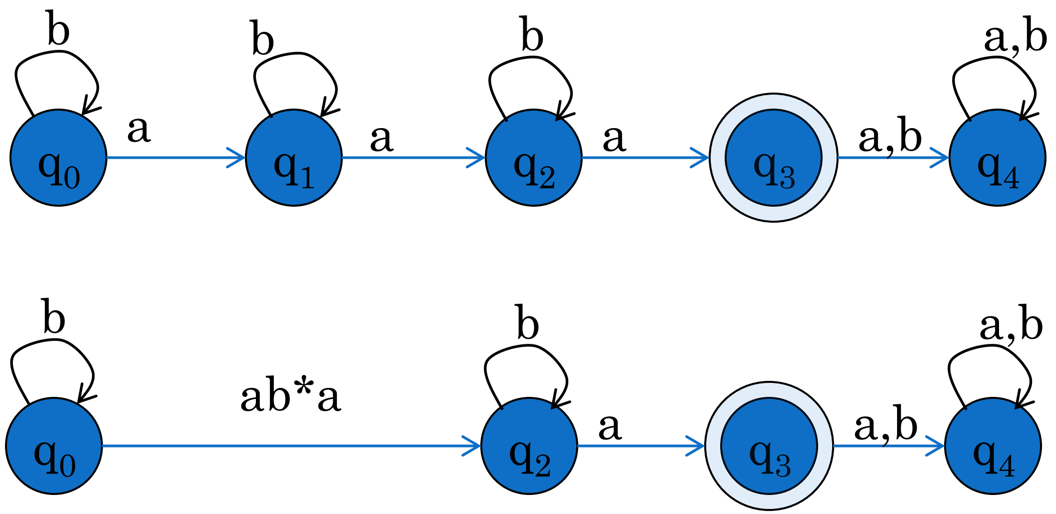
Example

- Eliminating q1



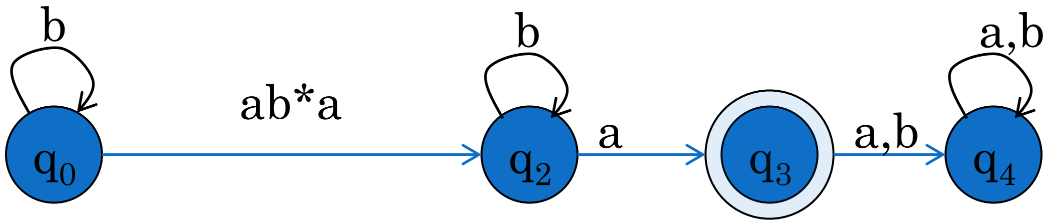
Example

- Eliminating q1



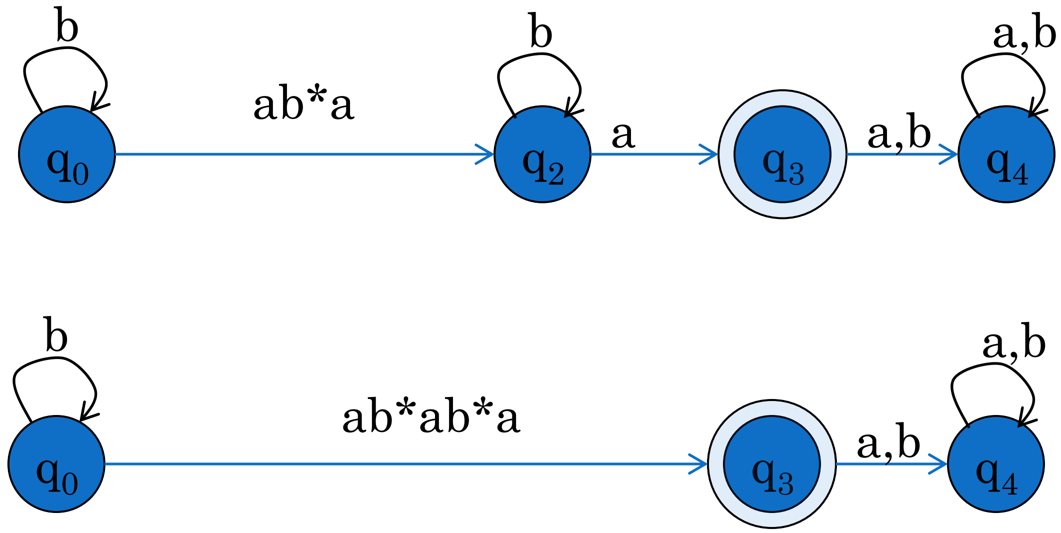
Example

- Eliminating q2



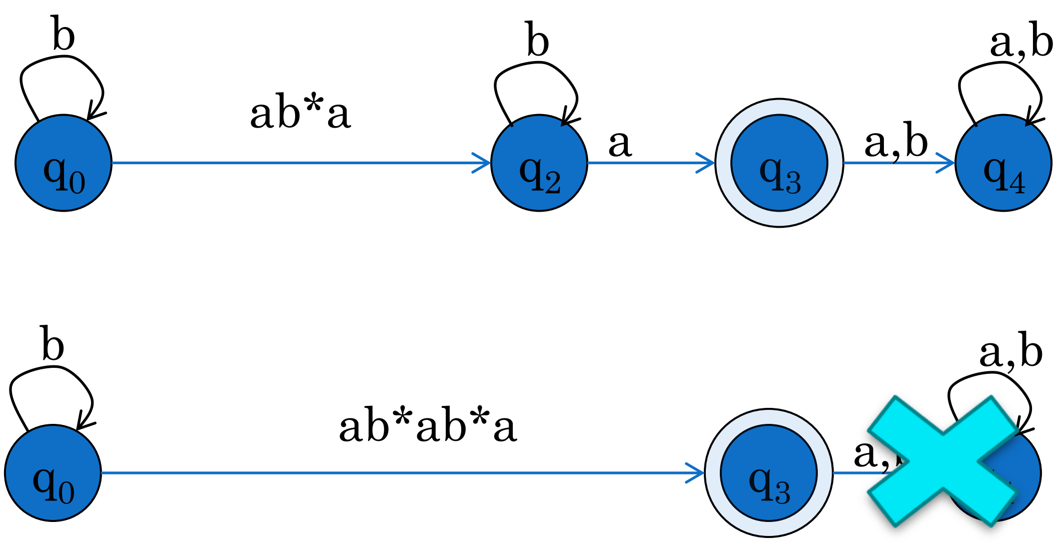
Example

- Eliminating q2



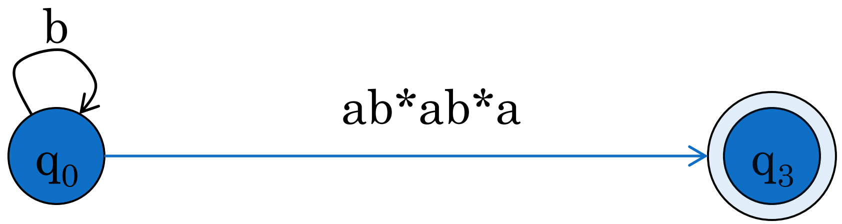
Example

- Eliminating q4



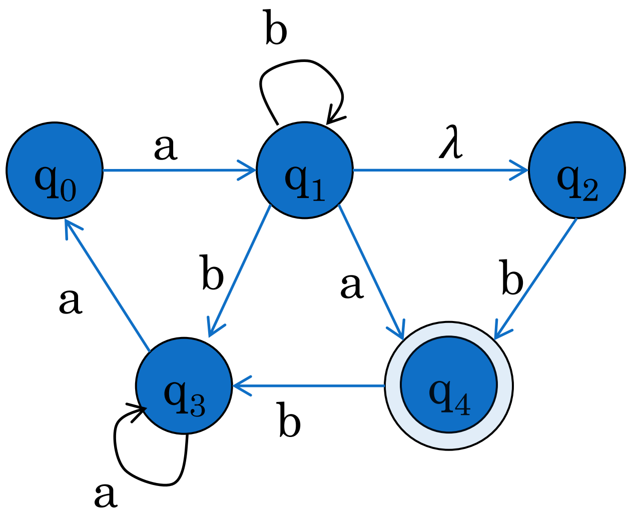
Example

- RE:



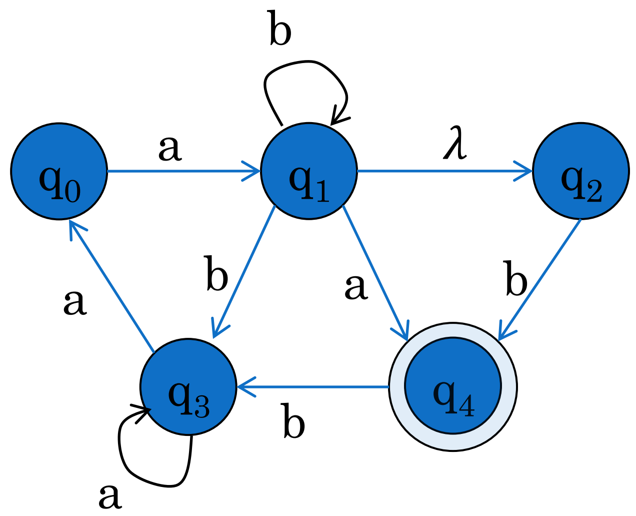
b*ab*ab*a

Example



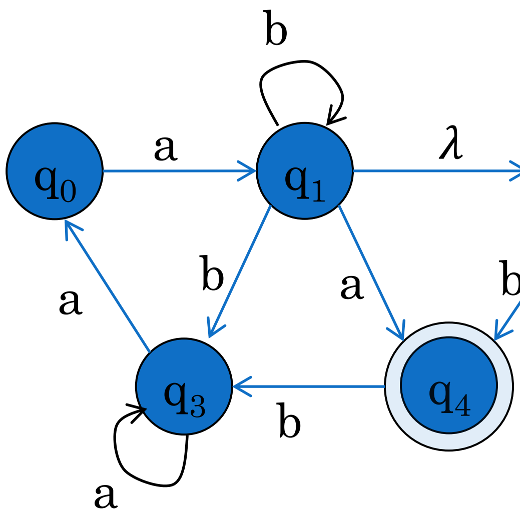
Example

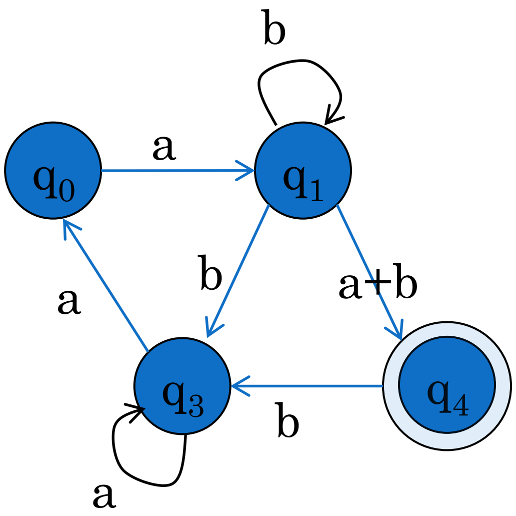
Eliminating q2:



Example

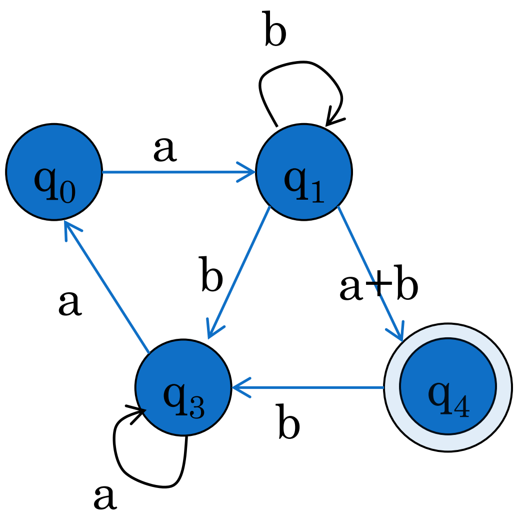
Eliminating q2:





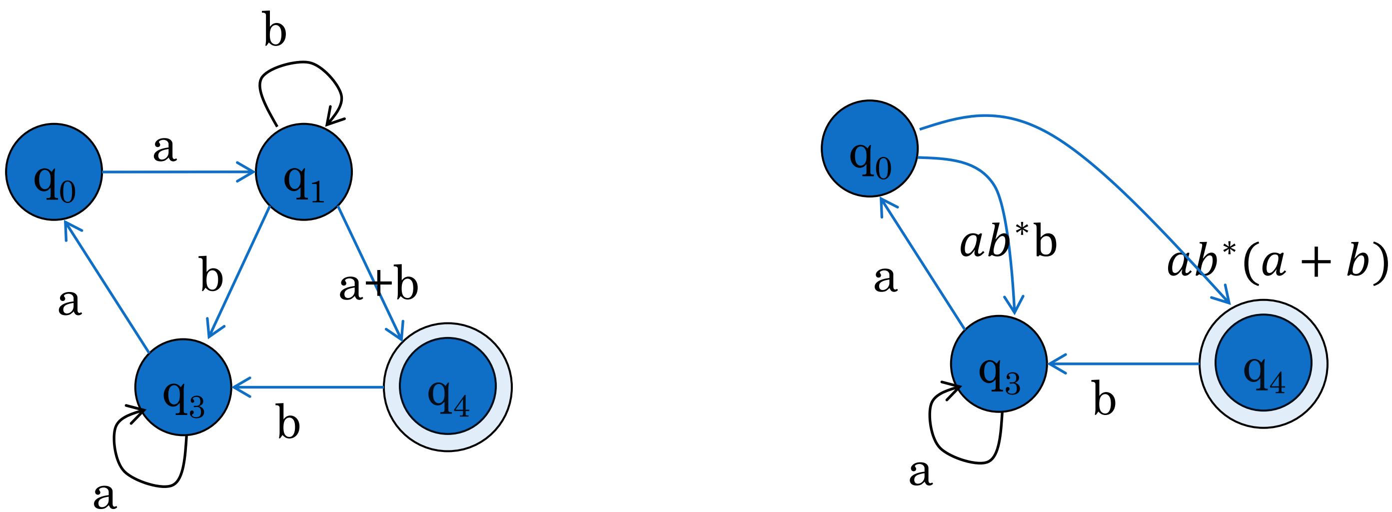
Example

Eliminating q1:



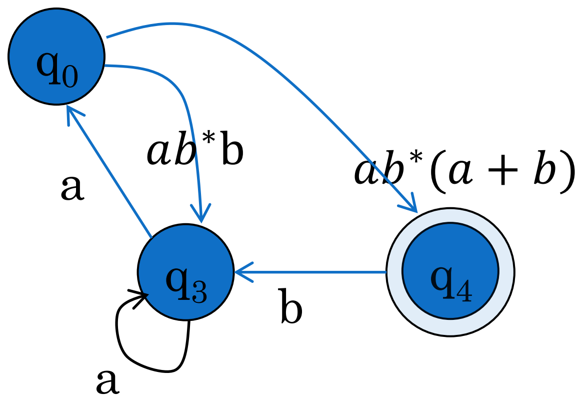
Example

Eliminating q1:



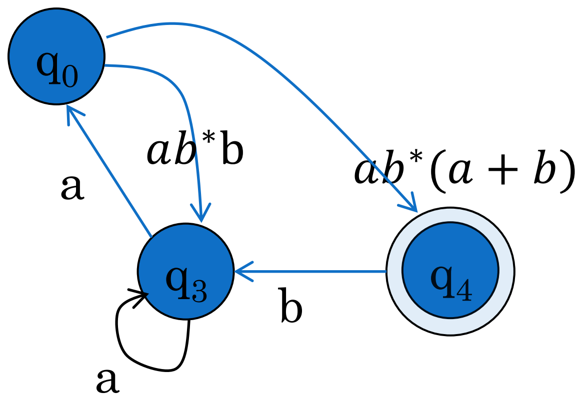
Example

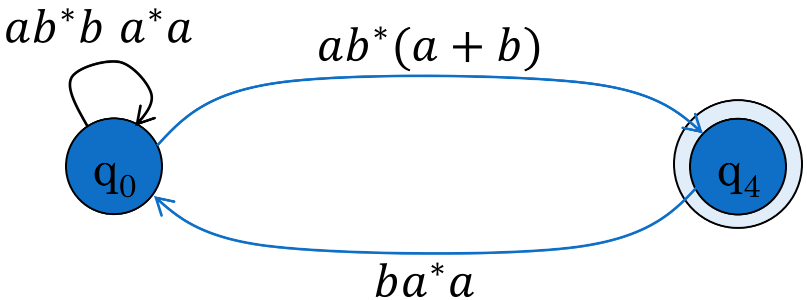
Eliminating q3:



Example

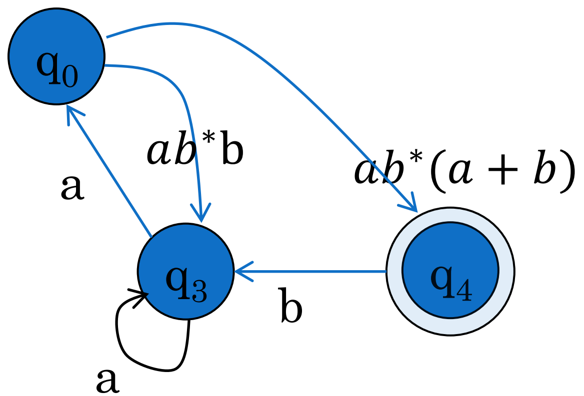
Eliminating q3:

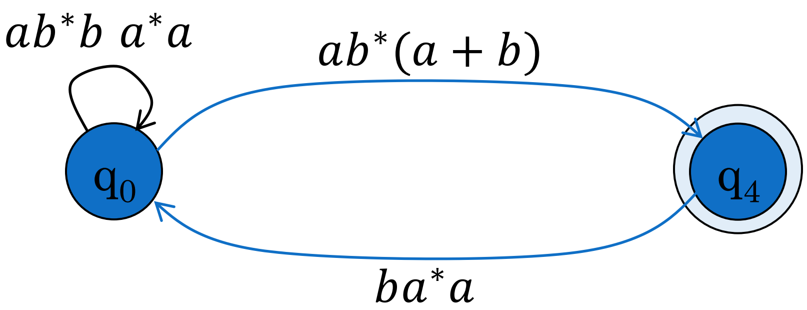




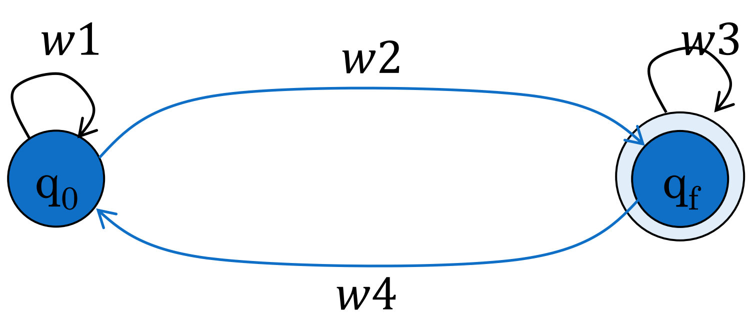
Example

Eliminating q3:

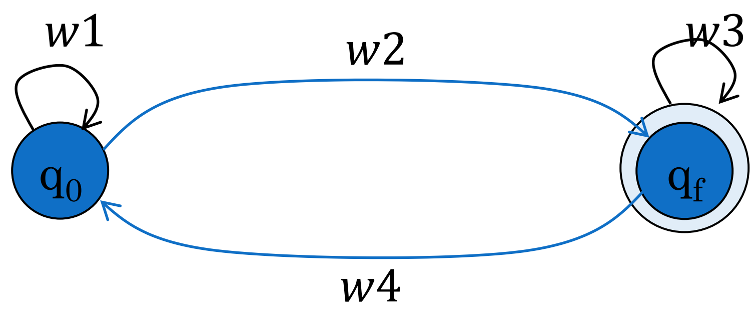




Note:



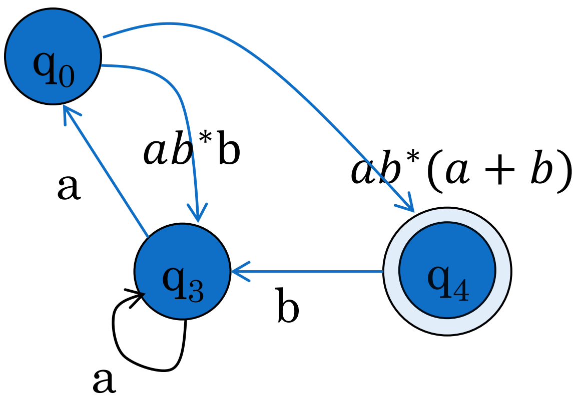
Example

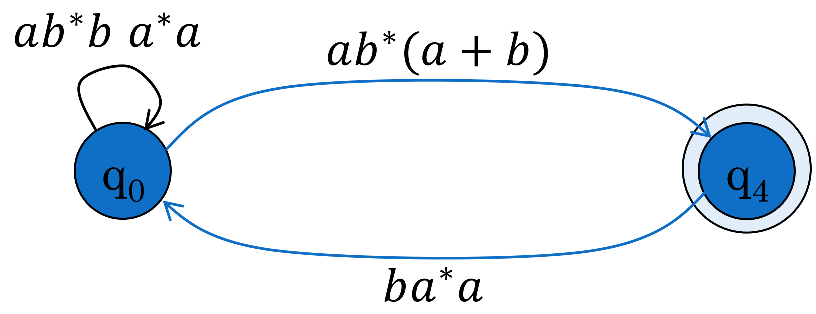


$$w1^{*}w2w3^{*}(w4w1^{*}w2w3^{*})^{*}$$

Example

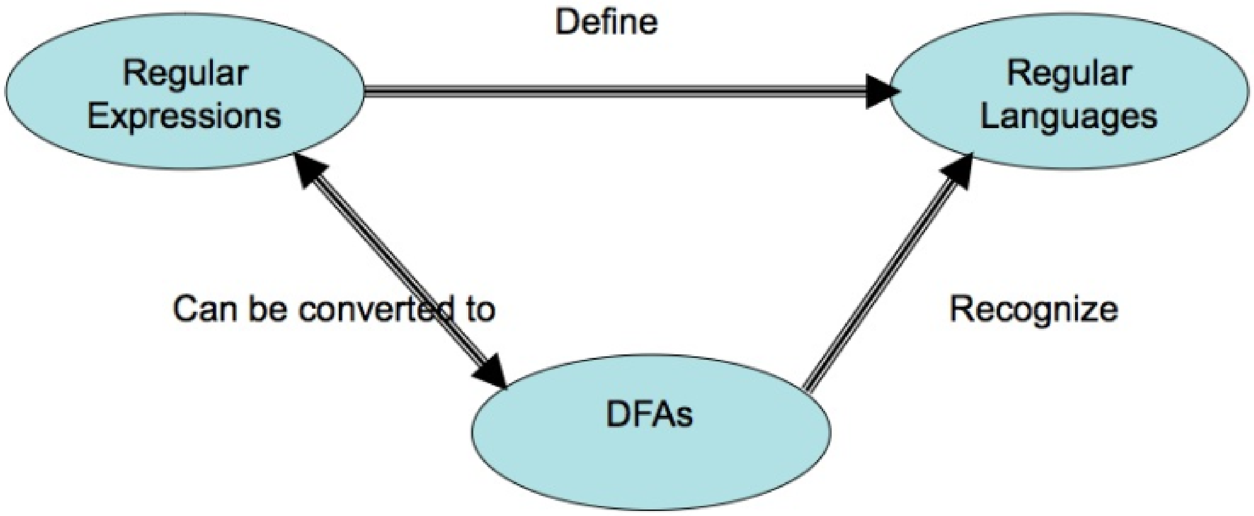
Eliminating q3:



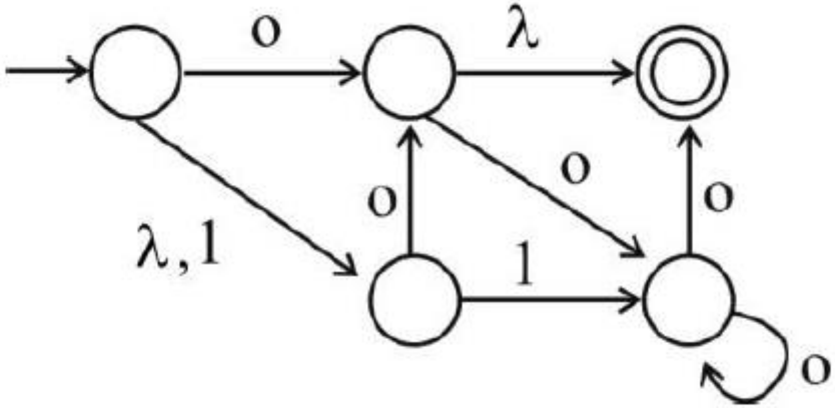


$$(ab^*b\ a^*a\ )^*ab^*(a+b)(ba^*a(ab^*b\ a^*a\ )^*ab^*(a+b)\ )^*$$

The Story So far



Quiz



Homework

- Find RE for the below FAs