Definition of GNFA I

- Between any two states: a regular expression
- \((Q, \Sigma, \delta, q_{\text{start}}, q_{\text{accept}})\)

\[
\delta : (Q - \{q_{\text{accept}}\}) \times (Q - \{q_{\text{start}}\}) \rightarrow R
\]

- \(R\): all regular expressions over \(\Sigma\)
- **DFA \rightarrow GNFA**
  - Two new states: \(q_{\text{start}}, q_{\text{accept}}\)
  - \(q_{\text{start}} \rightarrow q_0\) with \(\epsilon\)
  - Any \(q \in F\) \(\rightarrow q_{\text{accept}}\) with \(\epsilon\)
In the definition, between any two states there is an expression
But what if in the graph two states are not connected?
\(\emptyset \in R\) so if no connection, we simply consider \(\emptyset\) as the expression between two states
Fig 1.63
$q_i \xrightarrow{(R_1)(R_2)^* (R_3) \cup (R_4)} q_j$

- $q_{\text{rip}}$ is the state being removed
- In the procedure
  - 3-state DFA $\rightarrow$ 5-state GNFA $\rightarrow$ 4-state $\cdots \rightarrow$ 2-state GNFA $\rightarrow$ regular expression
- In the procedure any any $(i,j)$ related to $q_{\text{rip}}$ considered
- Algorithm: convert(G)
  - $k$: $\#$ of G
If $k = 2$

return $R$

If $k > 2$, choose any $q_{\text{rip}} \in Q \setminus \{q_s, q_a\}$ for removal

$$Q' = Q - \{q_{\text{rip}}\}$$

$$\forall q_i \in Q' - \{q_{\text{accept}}\}, q_j \in Q' - \{q_{\text{start}}\}$$

$$\delta'(q_i, q_j) = R_1 R_2^* R_3 \cup R_4,$$

where

$$R_1 = \delta(q_i, q_{\text{rip}}), R_2 = \delta(q_{\text{rip}}, q_{\text{rip}}),$$

$$R_3 = \delta(q_{\text{rip}}, q_j), R_4 = \delta(q_i, q_j)$$
4. Run convert($G'$), where

$$G' = (Q', \Sigma, \delta', q_s, q_a)$$

- Why in the textbook we modify DFA to GNFA?
  - Is it ok to use NFA?
  - Seems ok??