LF – Théorie des langages formels

Sylvain Brandel

2022 – 2023

sylvain.brandel@univ-lyon1.fr

CM 6

CARACTÉRISATION

Lien avec les langages rationnels
Stabilité

- Propriété de stabilité des langages reconnus par des automates finis.

- Théorème

  La classe des langages acceptés par les automates finis est stable par :
  - *Union*
  - *Concaténation*
  - *Fermeture itérative* *ordre de la démonstration*
  - *Complément*
  - *Intersection*

- Preuve constructive
  - Construction de l'automate pour chaque opération

Lien avec les langages rationnels
Stabilité

- Union

$L_1 = L(M_1) \qquad M_1 = (K_1, \Sigma, \Delta_1, s_1, F_1)$ → Même Σ

$L_2 = L(M_2) \qquad M_2 = (K_2, \Sigma, \Delta_2, s_2, F_2)$ → Hyp. $K_1 \cap K_2 = \varnothing$

Posons s tel que $s \notin K_1$ et $s \notin K_2$.

→ $M_\cup$: $(\{s\} \cup K_1 \cup K_2, \Sigma, \Delta_1 \cup \Delta_2 \cup \{(s, \varepsilon, s_1), (s, \varepsilon, s_2)\}, s, F_1 \cup F_2)$

$$w \in L(M_\cup) \Leftrightarrow (s, w) \vdash_M (s_1, w) \vdash_M^* (f_1, \varepsilon)\ (f_1 \in F_1) \qquad \leftarrow L_1$$

ou

$$(s, w) \vdash_M (s_2, w) \vdash_M^* (f_2, \varepsilon)\ (f_2 \in F_2) \qquad \leftarrow L_2$$

$$\Leftrightarrow w \in L_1 \cup L_2$$

→ $L(M_\cup) = L_1 \cup L_2$

Lien avec les langages rationnels
Stabilité

- Concaténation

$L_1 = L(M_1) \qquad M_1 = (K_1, \Sigma, \Delta_1, s_1, F_1)$ → Même Σ

$L_2 = L(M_2) \qquad M_2 = (K_2, \Sigma, \Delta_2, s_2, F_2)$ → Hyp. $K_1 \cap K_2 = \varnothing$

→ $M_c : \ (K_1 \cup K_2, \Sigma, \Delta_1 \cup \Delta_2 \cup \{(f_i, \varepsilon, s_2) \mid f_i \in F_1\}, s_1, F_2)$

→ $L(M_c) = L_1 L_2$

Lien avec les langages rationnels
Stabilité

- Etoile de Kleene

$L_1 = L(M_1) \qquad M_1 = (K_1, \Sigma, \Delta_1, s_1, F_1)$

Posons s tel que $s \notin K_1$

→ $M_K : (K_1 \cup \{s\}, \Sigma, \Delta_1 \cup \{(s, \varepsilon, s_1)\} \cup \{(f_i, \varepsilon, s_1) \mid f_i \in F_1\}, s, F_1 \cup \{s\})$

→ $L(M_K) = {L_1}^*$

Lien avec les langages rationnels

Stabilité

- Complément

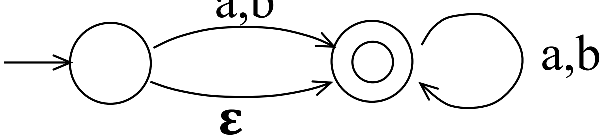
$$L_1 \;=\; L(M_1) \qquad\qquad M_1 = (K_1, \Sigma, \Delta_1, s_1, F_1)$$

$$\rightarrow M_\neg : \;(K_1, \Sigma, \Delta_1, s_1, \neg F_1)\; avec\; \neg F_1 \;=\; K_1 - F_1$$

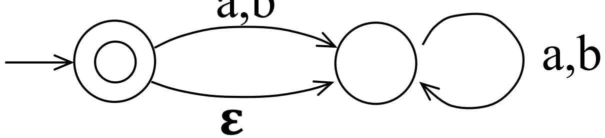
$$\rightarrow L(M_\neg) = \neg L_1$$

Attention M_1 doit être déterministe et complet

ex non det. :



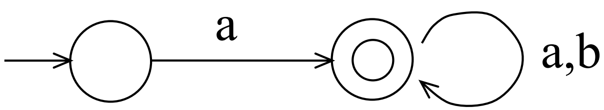
➔



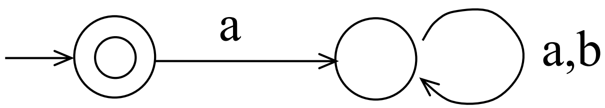
L(M) = (a ∪ b)*

L(M¬) = {ε}

ex non complet :



➔



L(M) = a (a ∪ b)*

L(M¬) = {ε}

Lien avec les langages rationnels
Stabilité

- Intersection

$L_1 = L(M_1) \qquad M_1 = (K_1, \Sigma, \Delta_1, s_1, F_1)$ → Même Σ

$L_2 = L(M_2) \qquad M_2 = (K_2, \Sigma, \Delta_2, s_2, F_2)$ → Hyp. $K_1 \cap K_2 = \varnothing$

- Deux méthodes :

  – $L(M_1) \cap L(M_2) = \overline{\overline{L(M_1)} \cup \overline{L(M_2)}}$

  – Automate produit, avec M_1et M_2 déterministes et complets

    - $M_\cap : (K_1 \times K_2, \Sigma, \{((p_1, p_2), \sigma, (q_1, q_2)) \mid (p_1, \sigma, q_1) \in \Delta_1\ et\ (p_2, \sigma, q_2) \in \Delta_2\}, (s_1, s_2), F_1 \times F_2)$
    - Quadratique en le nombre d'états

→ $L(M_\cap) = L_1 \cap L_2$

Lien avec les langages rationnels

Stabilité

- Retour sur l'union

$L_1 = L(M_1) \qquad M_1 = (K_1, \Sigma, \Delta_1, s_1, F_1)$ → Même Σ

$L_2 = L(M_2) \qquad M_2 = (K_2, \Sigma, \Delta_2, s_2, F_2)$ → Hyp. $K_1 \cap K_2 = \emptyset$

Posons s tel que $s \notin K_1$ et $s \notin K_2$.

→ $M_\cup : (\{s\} \cup K_1 \cup K_2, \Sigma, \Delta_1 \cup \Delta_2 \cup \{(s, \varepsilon, s_1), (s, \varepsilon, s_2)\}, s, F_1 \cup F_2)$

- Autre méthode :
  - Automate produit, avec M_1et M_2 déterministes et complets
    - $M_\cup : (K_1 \times K_2, \Sigma, \{((p_1, p_2), \sigma, (q_1, q_2)) \mid (p_1, \sigma, q_1) \in \Delta_1 \; et \; (p_2, \sigma, q_2) \in \Delta_2\},$
      $(s_1, s_2), F_1 \times K_2 \cup K_1 \times F_2)$
    - Quadratique en le nombre d'états

→ $L(M_\cup) = L_1 \cup L_2$

Lien avec les langages rationnels
Stabilité

- Automate produit – intersection

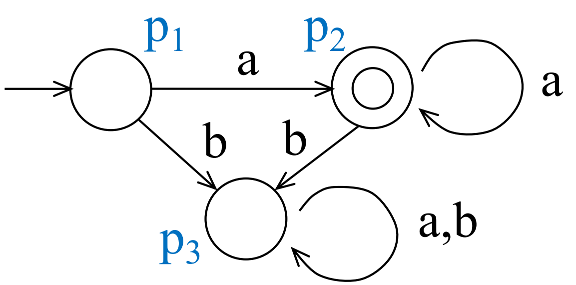
$$L_1 = L(M_1) \qquad M_1 = (K_1, \Sigma, \Delta_1, s_1, F_1)$$
$$L_2 = L(M_2) \qquad M_2 = (K_2, \Sigma, \Delta_2, s_2, F_2)$$

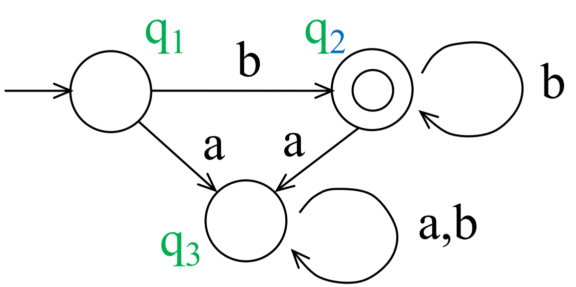
→ Même Σ

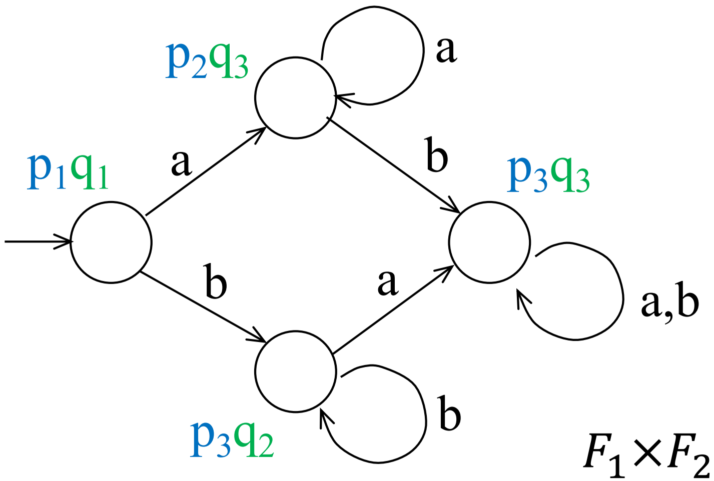
→ Hyp. $K_1 \cap K_2 = \varnothing$

– Automate produit, avec M_1 et M_2 déterministes et complets

- $M_\cap$: $(K_1 \times K_2, \Sigma, \{((p_1, p_2), \sigma, (q_1, q_2)) \mid (p_1, \sigma, q_1) \in \Delta_1 \ et \ (p_2, \sigma, q_2) \in \Delta_2\}, (s_1, s_2), F_1 \times F_2)$







$$F_1 \times F_2 = \emptyset$$

Lien avec les langages rationnels
Stabilité

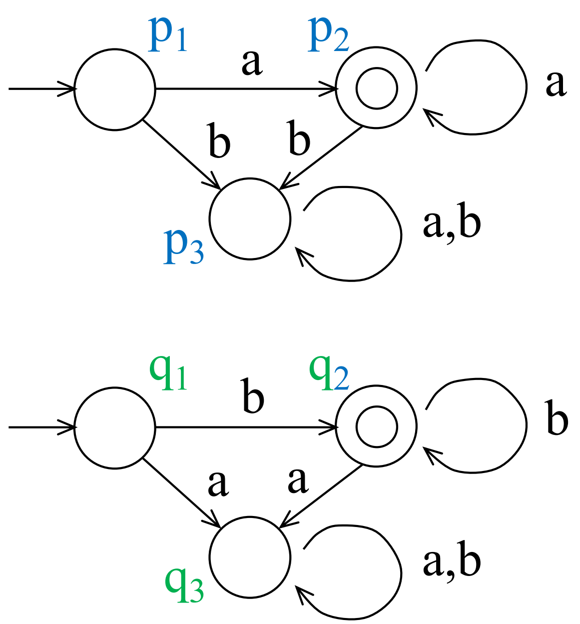
- Automate produit – union

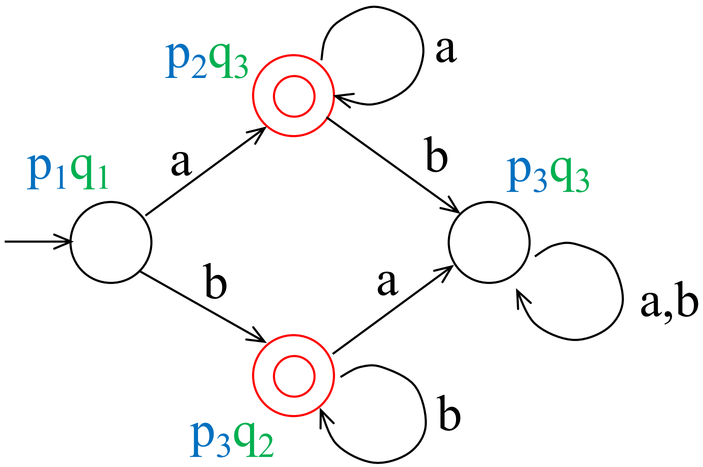
$$L_1 = L(M_1) \qquad M_1 = (K_1, \Sigma, \Delta_1, s_1, F_1)$$
$$L_2 = L(M_2) \qquad M_2 = (K_2, \Sigma, \Delta_2, s_2, F_2)$$

→ Même Σ

→ Hyp. $K_1 \cap K_2 = \varnothing$

- – Automate produit, avec M_1 et M_2 déterministes et complets
  - $M_\cup$: $(K_1 \times K_2, \Sigma, \{((p_1,p_2),\sigma,(q_1,q_2)) \mid (p_1,\sigma,q_1) \in \Delta_1 \; et \; (p_2,\sigma,q_2) \in \Delta_2\}, (s_1,s_2), F_1 \times K_2 \cup K_1 \times F_2)$





Lien avec les langages rationnels

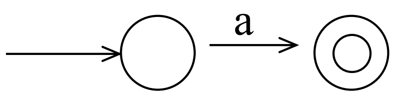
Inclusion des classes de langages

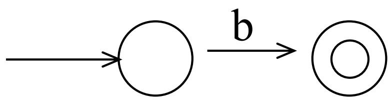
- Théorème

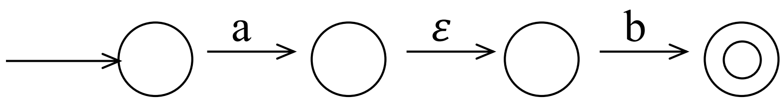
  La classe des langages acceptés par les automates finis contient les langages rationnels.

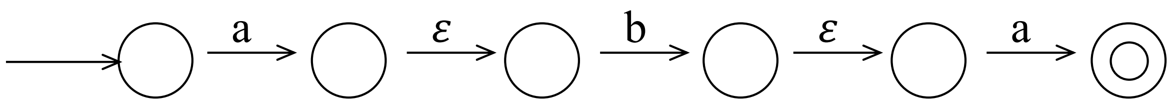
- Exemple

  $(ab \cup aba)^*$

a 

b 

ab 

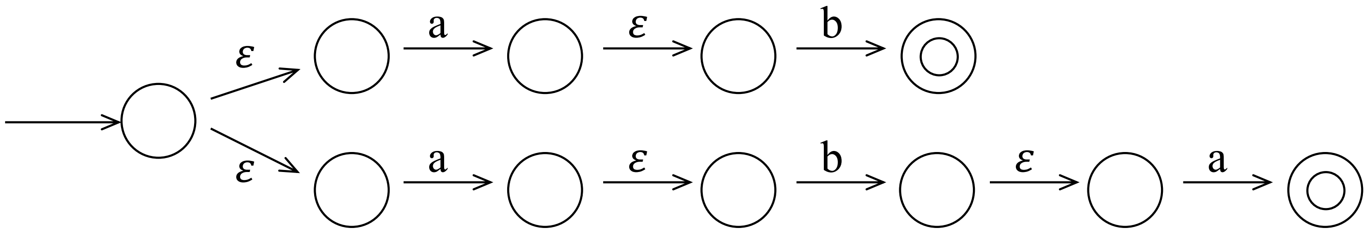
aba 

Lien avec les langages rationnels
Inclusion des classes de langages

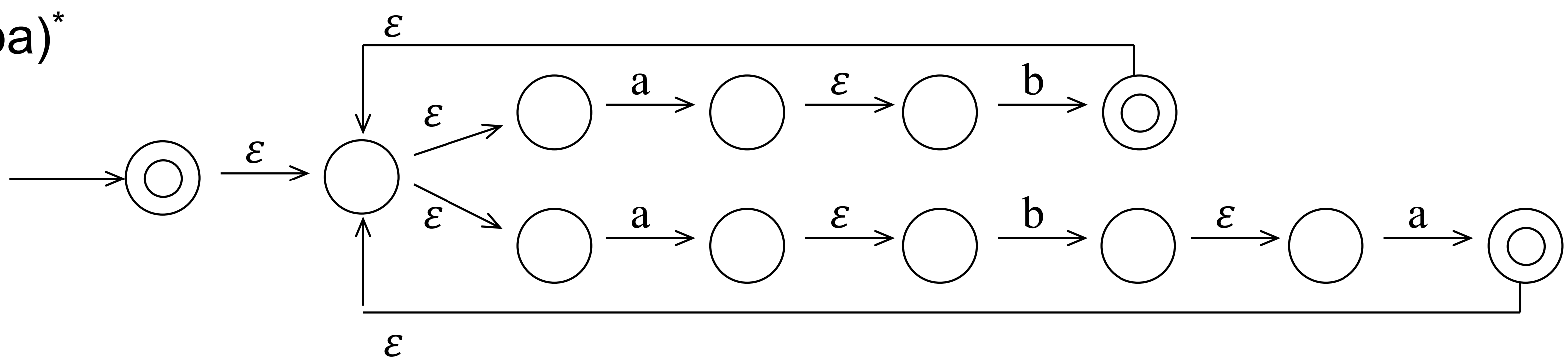
- Exemple

  $(ab \cup aba)^*$

$ab \cup aba$



$(ab \cup aba)^*$



Lien avec les langages rationnels

Caractérisation des langages rationnels

- Théorème

  Un langage est rationnel ssi il est accepté par un automate fini.

- Preuve

  On suppose qu'on a numéroté (ordonné) les états.

  Soit $M = (K, \Sigma, \Delta\ s, F)$ un automate fini (déterministe ou non). $|K| = n$.

  $$K = \{q_1, q_2, \dots, q_n\}, \quad s = q_1$$

  Le langage reconnu par M est la réunion de tous les langages reconnus en parcourant tous les chemins possibles dans le graphe.

  → À chaque chemin allant de s à f $(\in F)$, on associe le langage trouvé.

Lien avec les langages rationnels
Caractérisation des langages rationnels

- On pose $R(i,j,k)$ = l'ensemble des mots obtenus par lecture de l'automate M
  - en partant de l'état q_i,
  - en arrivant dans l'état q_j (avec le mot vide),
  - en ne passant que par des états intermédiaires dont le numéro est ≤ k.

- $R(i,j,k)$ est un langage

- $R(i,j,k)\ \{w\ |(q_i,w) \vdash_M^* \ (q_j,\varepsilon)\ \textit{sans passer par des états intermédiaires}$
$\textit{dont le numéro est} > k\}$

$$R(i,j,n) \ = \ \{w \mid (q_i,w) \vdash_M^* (q_j,\varepsilon)\}$$

$$L(M) = \bigcup_{i|q_i \in F} R(1,i,n)$$

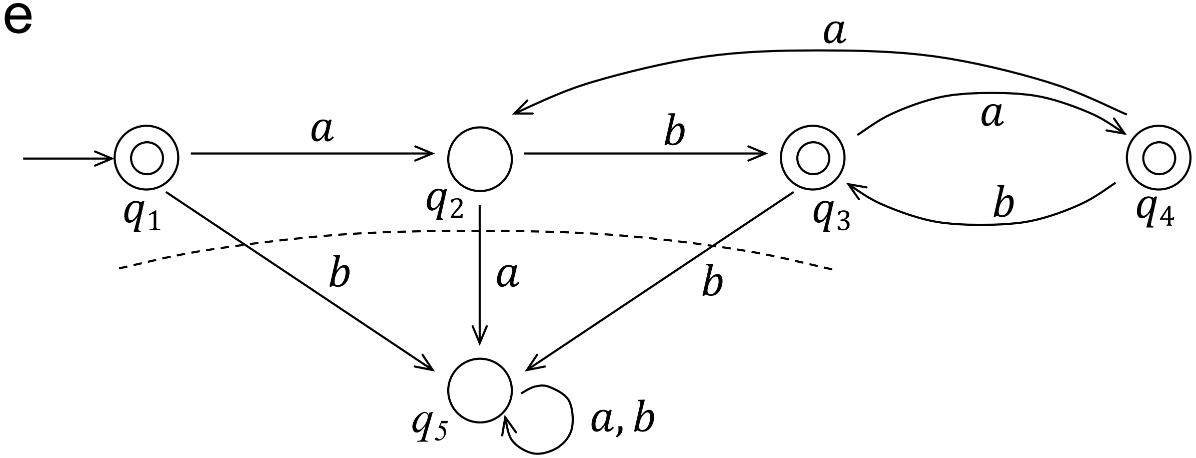
Lien avec les langages rationnels
Caractérisation des langages rationnels

- $R(i,j,k)$ est un langage rationnel dont on peut calculer la représentation par récurrence sur k.

- Preuve

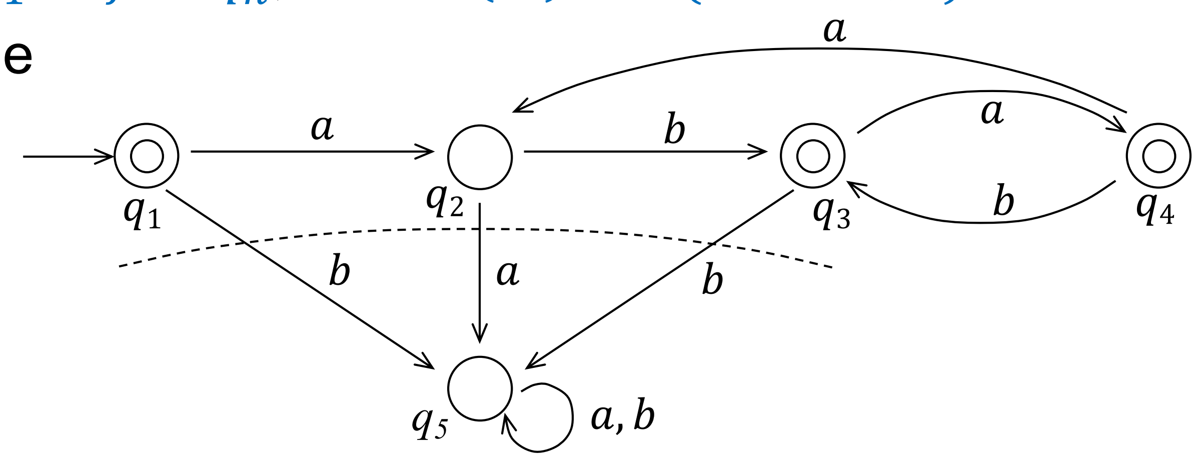
$$R(i,j,k) \;=\; R(i,j,k-1) \cup R(i,k,k-1).\big(R(k,k,k-1)\big)^{*}.R(k,j,k-1)$$

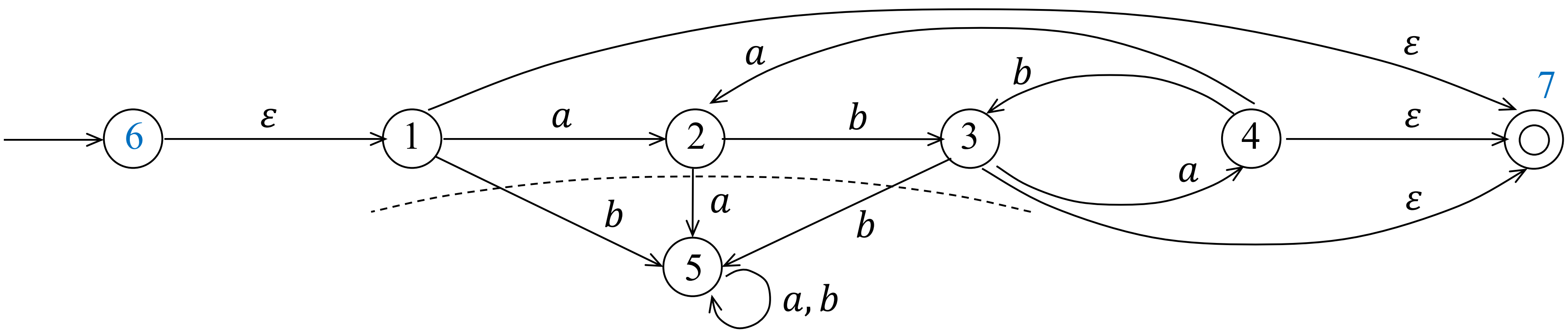
- Exemple



Lien avec les langages rationnels
Caractérisation des langages rationnels

- Forme spéciale de M ($L(M)$ inchangé) :
  - Un seul état final f
  - Si $(p, \sigma, q) \in \Delta$, alors $p \neq f$ et $q \neq s$ (pas de retour de f ou vers s)
- $s = q_{n-1}$ et $f = q_n$, alors $L(M) = R(n-1, n, n)$
- Exemple





Lien avec les langages rationnels
Caractérisation des langages rationnels

- Exemple
  - $R(i,j,0)$ → étiquettes sur les flèches.
  - Principe : calculer $R(6,7,7)$

$R(6,7,7)$ $= R(6,7,6) \cup R(6,6,6)\,R(6,6,6)^*\,R(6,7,6)$

$= R(6,7,5) \cup R(6,5,5)\,R(5,5,5)^*\,R(5,7,5) \cup$ ...

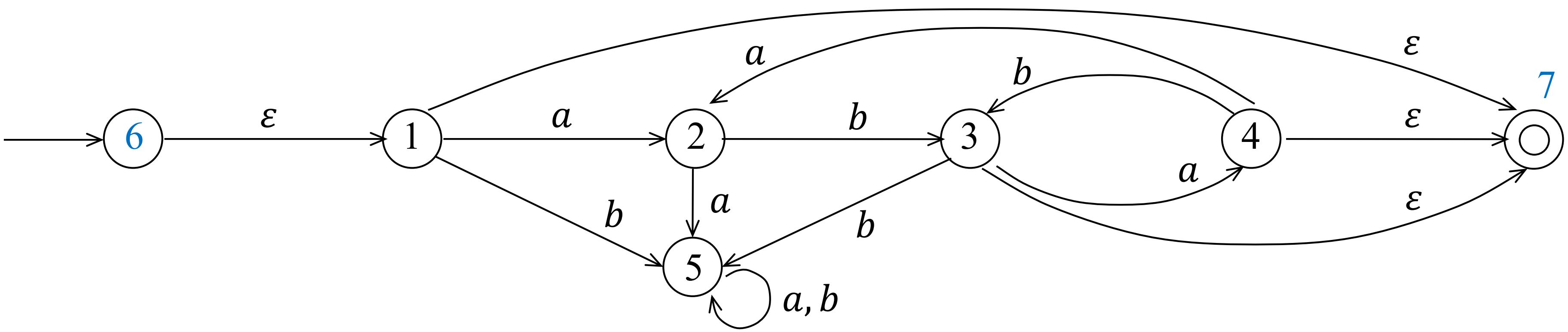
$= \ldots$

Fastidieux $\Rightarrow$ on calcule les $R(i,j,k)$ de proche en proche.

→ On supprime q_1 (ce qui revient à calculer $R(i,j,1)$)

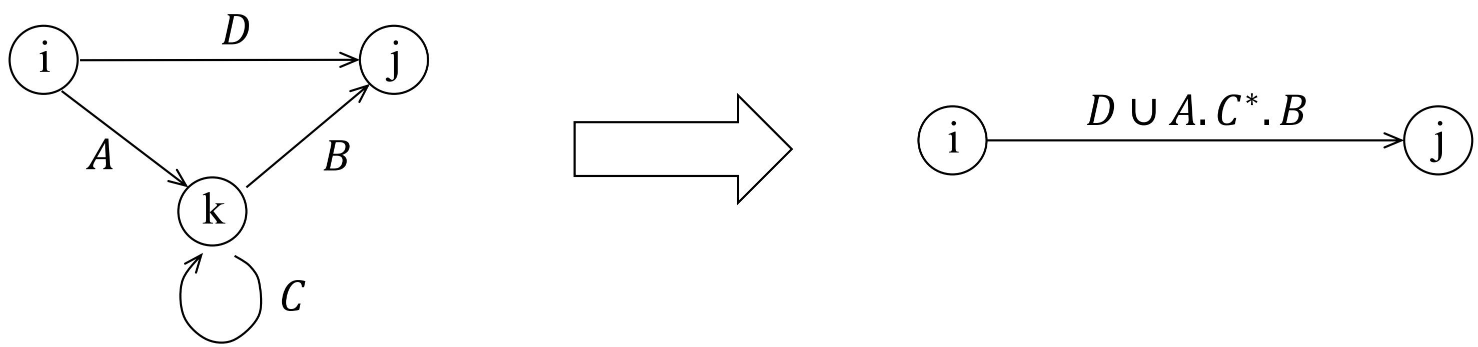
→ On supprime q_2 (ce qui revient à calculer $R(i,j,2)$)

…



Lien avec les langages rationnels
Caractérisation des langages rationnels

- Règles de suppression d'un état q_k
  - Pour chaque paire d'états $q_i \neq q_k$ et $q_j \neq q_k$ pour lesquels il y a une flèche de q_i à q_k étiquetée par A et une flèche de q_k à q_j étiquetée par B
    - On ajoute une flèche de q_i à q_j étiquetée par $A.C^*.B$, où C est l'étiquette de la flèche de q_k à q_k ; s'il n'y a pas de flèche de q_k à q_k, $C = \emptyset$, donc $C^* = \{\varepsilon\}$, la flèche de q_i à q_j est donc étiquetée par $A.B$
    - S'il y a déjà une flèche de q_i à q_j étiquetée par D, l'étiquette devient $\mathrm{D} \cup A.C^*.B$
  - Pour chaque paire d'états $q_i \neq q_k$ et $q_j \neq q_k$ pour lesquels il y a une flèche de q_i à q_j étiquetée par D et pas de flèche de q_i à q_k ou de q_k à q_j
    - L'étiquette de q_i à q_j reste D
  - On supprime q_k et toutes les flèches entrantes et sortantes

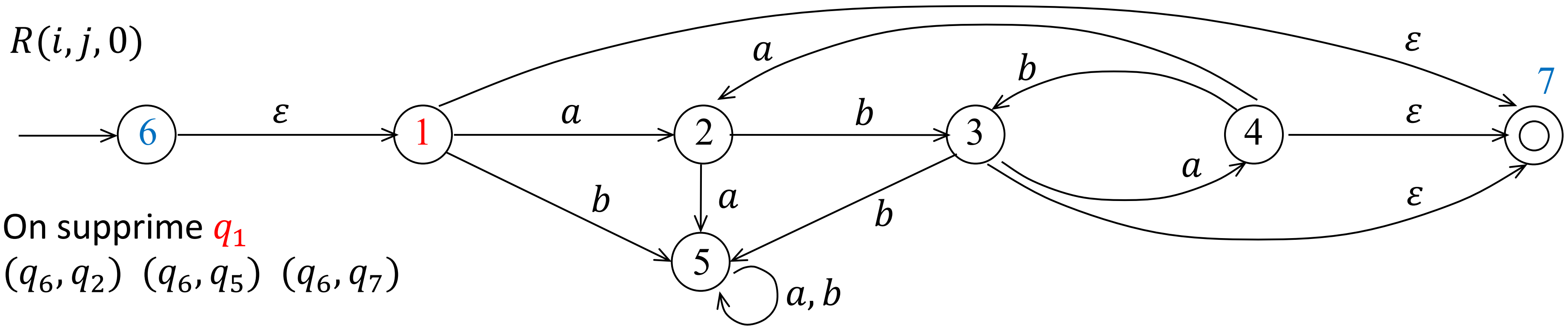


Lien avec les langages rationnels

Caractérisation des langages rationnels

- Exemple

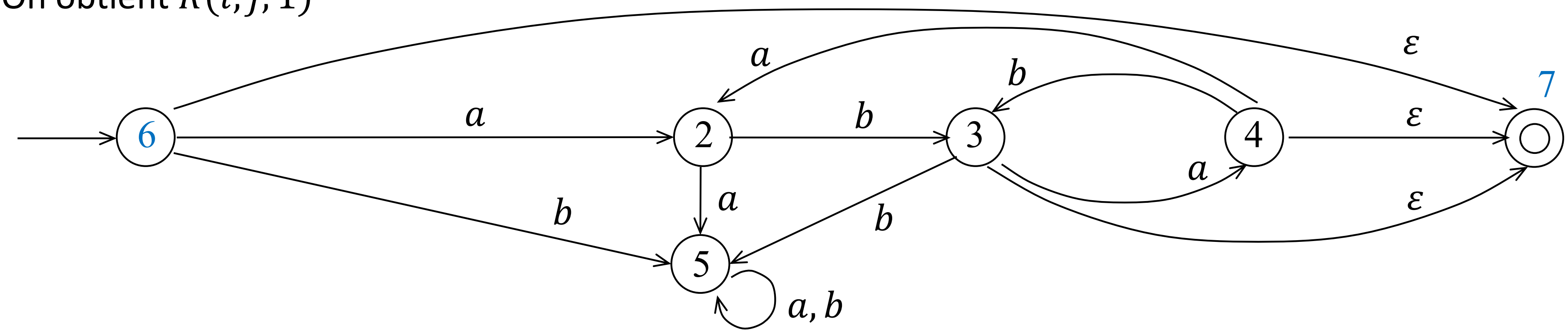
$R(i,j,0)$



On supprime q_1

(q_6, q_2) (q_6, q_5) (q_6, q_7)

On obtient $R(i,j,1)$

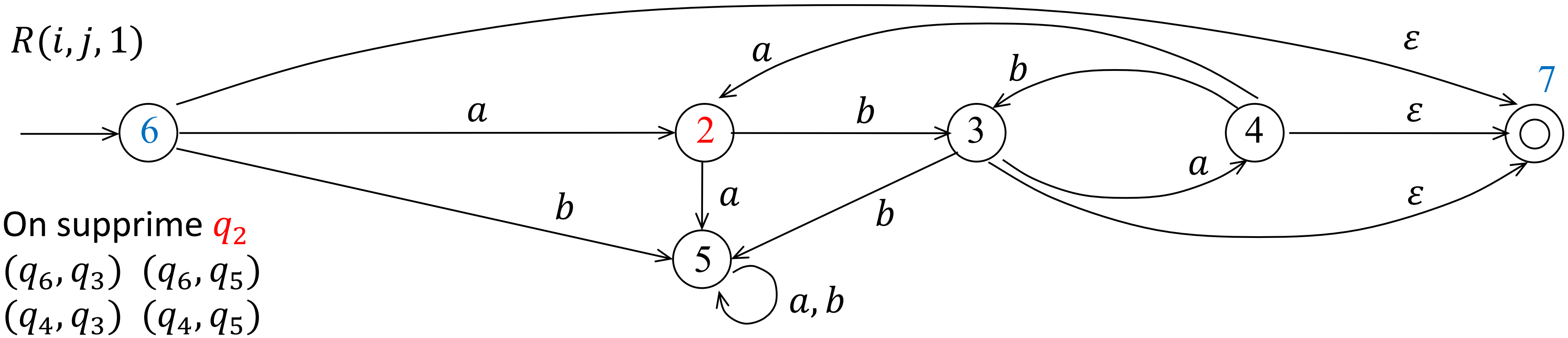


Lien avec les langages rationnels

Caractérisation des langages rationnels

- Exemple

$R(i,j,1)$

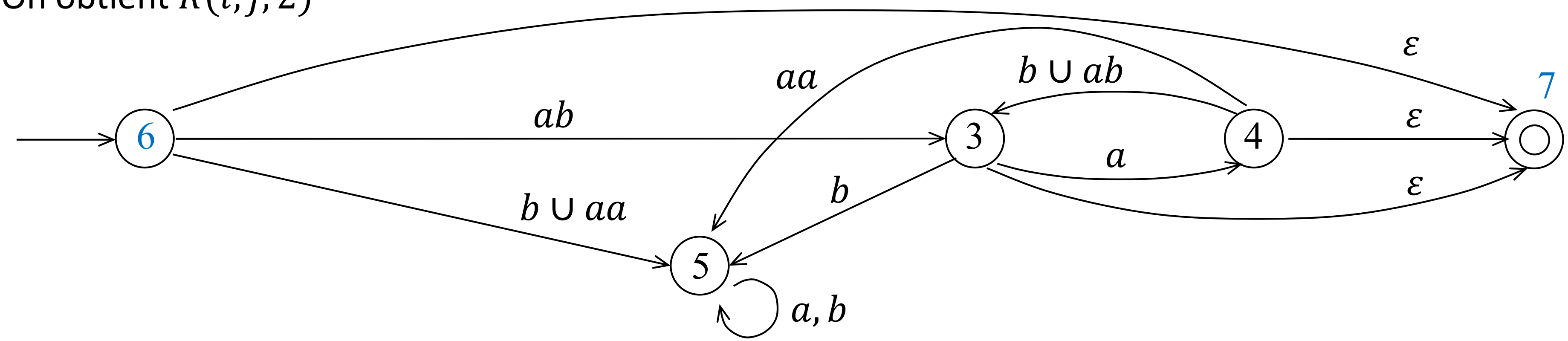


On supprime q_2

(q_6, q_3) (q_6, q_5)

(q_4, q_3) (q_4, q_5)

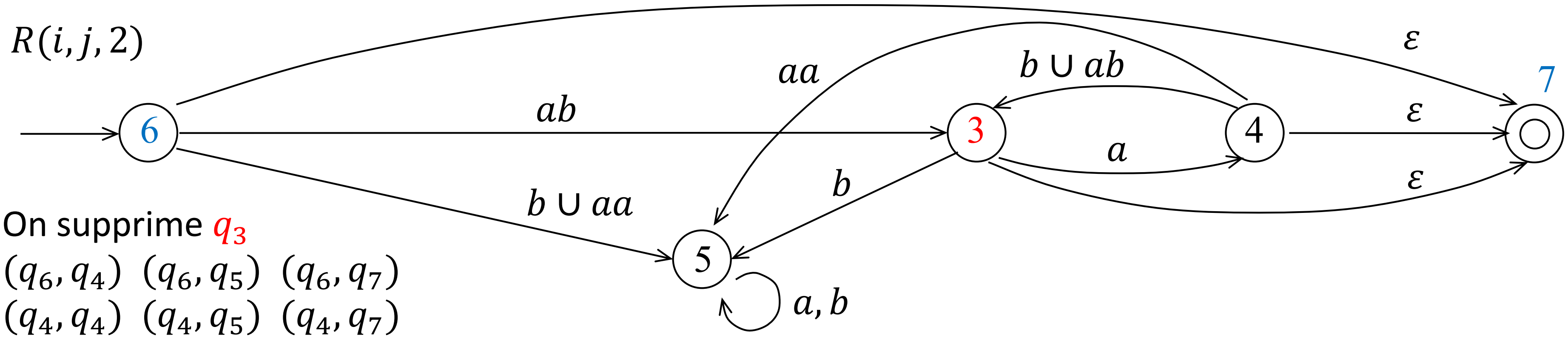
On obtient $R(i,j,2)$



Lien avec les langages rationnels
Caractérisation des langages rationnels

- Exemple

$R(i,j,2)$

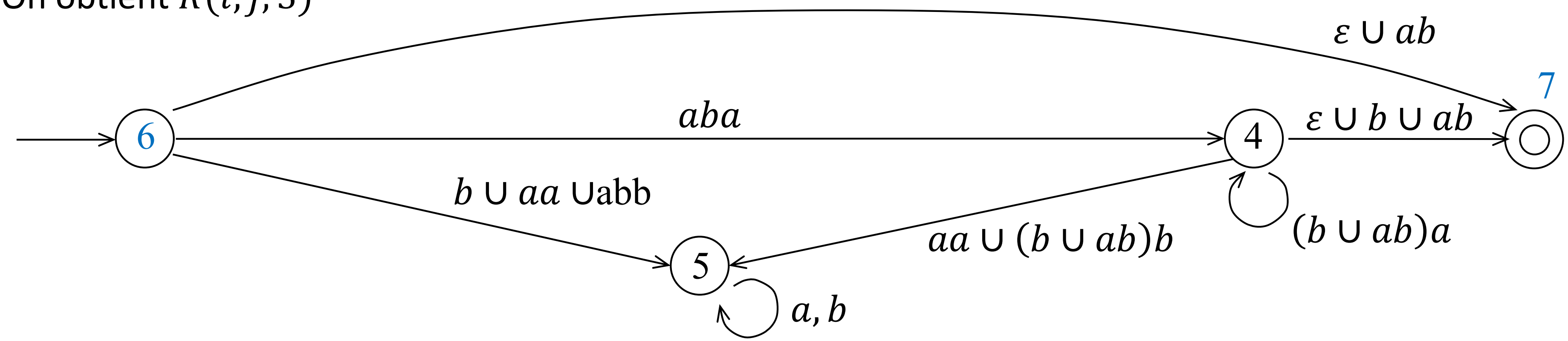


On supprime q_3

(q_6, q_4) (q_6, q_5) (q_6, q_7)

(q_4, q_4) (q_4, q_5) (q_4, q_7)

On obtient $R(i,j,3)$

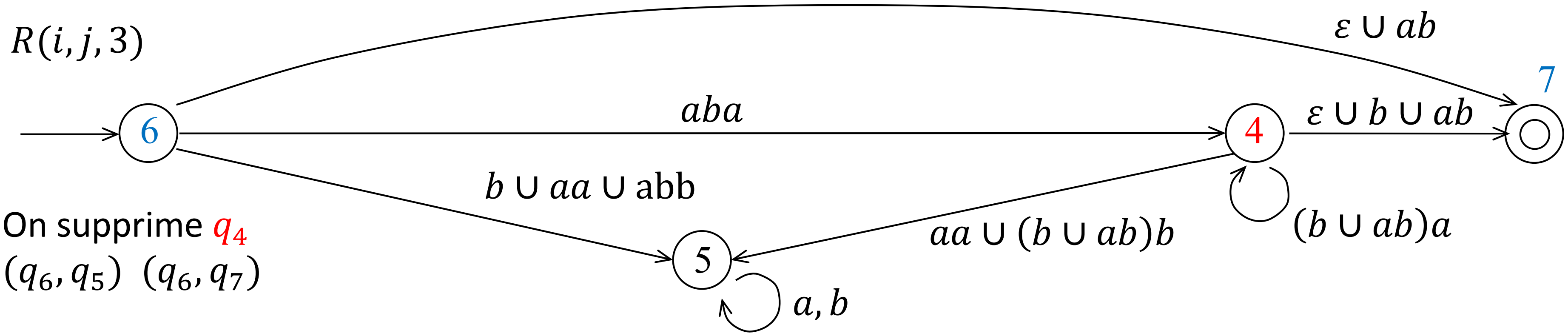


Lien avec les langages rationnels

Caractérisation des langages rationnels

- Exemple

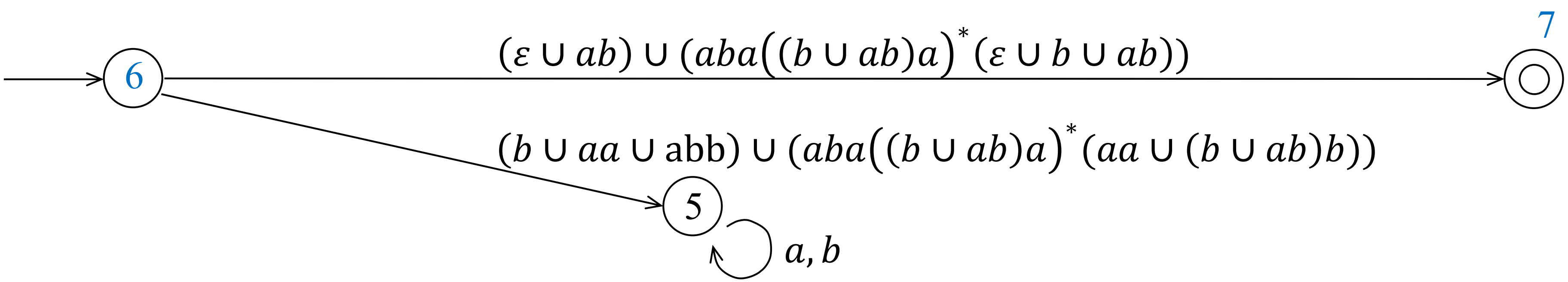
$R(i,j,3)$



On supprime q_4

(q_6, q_5) (q_6, q_7)

On obtient $R(i,j,4)$

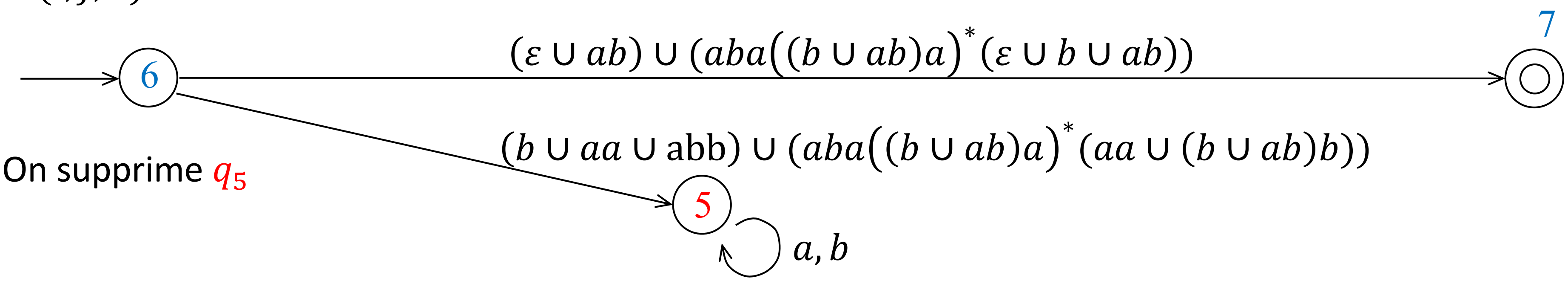


Lien avec les langages rationnels

Caractérisation des langages rationnels

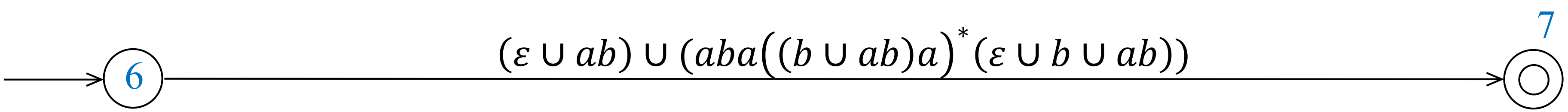
- Exemple

$R(i, j, 4)$



On supprime q_5

On obtient $R(i, j, 5)$



Donc $R(6,7,5) = R(6,7,7) = L(M) = \varepsilon \cup ab \cup aba(ba \cup aba)^*(\varepsilon \cup b \cup ab)$