5'. Automates finis (deuxième partie)

UL2IN005
Mathématiques discrètes

Nathalie Sznajder
Antoine Genitrini

Sorbonne Université

2025 – 2026

Langages rationnels

Définition

Soit A un alphabet. L'ensemble $Rat(A^*)$ des langages rationnels sur A est défini par induction structurelle par :

(B) $\emptyset$, $\{\varepsilon\}$ et $\{a\}$, pour tout $a \in A$ sont des langages rationnels,

(I) Si L_1 et L_2 sont des langages rationnels, alors $L_1 \cup L_2$ et $L_1.L_2$ sont aussi des langages rationnels, et si M est un langage rationnel, alors M^* est aussi un langage rationnel.

Expressions rationnelles

Définition

Soit A un alphabet. L'ensemble des expressions rationnelles sur A est défini par induction structurelle par

(B) 0 est une expression rationnelle, et, pour tout $a \in A$, a est une expression rationnelle.

(I) Si E est une expression rationnelle, alors $(E)^*$ est une expression rationnelle, et si E_1 et E_2 sont deux expressions rationnelles, alors $(E_1 + E_2)$ et $(E_1 \cdot E_2)$ sont des expressions rationnelles.

Définition

Le langage associé à une expression rationnelle sur A est défini par induction structurelle par

(B) $L(0) = \emptyset$, $L(a) = \{a\}$, pour tout $a \in A$.

(I) $L(E_1 + E_2) = L(E_1) \cup L(E_2)$, $L(E_1 \cdot E_2) = L(E_1).L(E_2)$ et $L(E^*) = L(E)^*$.

Expressions rationnelles

Les langages rationnels sont donc les langages associés à une expression rationnelle :

Exemple

- $\emptyset = L(0)$
- $\{a\} = L(a)$
- $\{\varepsilon\} = (L(0))^* = L(0^*)$

Expressions rationnelles

Les langages rationnels sont donc les langages associés à une expression rationnelle :

Exemple

- $\emptyset = L(0)$
- $\{a\} = L(a)$
- $\{\varepsilon\} = (L(0))^* = L(0^*)$

Exemple

$((a+b)\cdot(c+a))^*$

Expressions rationnelles

Les langages rationnels sont donc les langages associés à une expression rationnelle :

Exemple

- $\emptyset = L(0)$
- $\{a\} = L(a)$
- $\{\varepsilon\} = (L(0))^* = L(0^*)$

Exemple

$((a + b) \cdot (c + a))^*$
$L(((a + b) \cdot (c + a))^*)$

Expressions rationnelles

Les langages rationnels sont donc les langages associés à une expression rationnelle :

Exemple

- $\emptyset = L(0)$
- $\{a\} = L(a)$
- $\{\varepsilon\} = (L(0))^* = L(0^*)$

Exemple

$((a+b)\cdot(c+a))^*$
$L(((a+b)\cdot(c+a))^*) = L((a+b)\cdot(c+a))^*$

Expressions rationnelles

Les langages rationnels sont donc les langages associés à une expression rationnelle :

Exemple

- $\emptyset = L(0)$
- $\{a\} = L(a)$
- $\{\varepsilon\} = (L(0))^* = L(0^*)$

Exemple

$((a+b)\cdot(c+a))^*$

$L(((a+b)\cdot(c+a))^*) = (L(a+b).L(c+a))^*$

Expressions rationnelles

Les langages rationnels sont donc les langages associés à une expression rationnelle :

Exemple

- $\emptyset = L(0)$
- $\{a\} = L(a)$
- $\{\varepsilon\} = (L(0))^* = L(0^*)$

Exemple

$((a+b) \cdot (c+a))^*$

$L(((a+b) \cdot (c+a))^*) = ((L(a) \cup L(b)).(L(c) \cup L(a)))^*$

Expressions rationnelles

Les langages rationnels sont donc les langages associés à une expression rationnelle :

Exemple

- $\emptyset = L(0)$
- $\{a\} = L(a)$
- $\{\varepsilon\} = (L(0))^* = L(0^*)$

Exemple

$((a+b)\cdot(c+a))^*$
$L(((a+b)\cdot(c+a))^*) = (\{a\} \cup \{b\}).(\{c\} \cup \{a\})^*$

Expressions rationnelles

Les langages rationnels sont donc les langages associés à une expression rationnelle :

Exemple

- $\emptyset = L(0)$
- $\{a\} = L(a)$
- $\{\varepsilon\} = (L(0))^* = L(0^*)$

Exemple

$((a+b)\cdot(c+a))^*$

$L(((a+b)\cdot(c+a))^*) = (\{a,b\}.\{c,a\})^*$

Expressions rationnelles

Les langages rationnels sont donc les langages associés à une expression rationnelle :

Exemple

- $\emptyset = L(0)$
- $\{a\} = L(a)$
- $\{\varepsilon\} = (L(0))^* = L(0^*)$

Exemple

$((a+b)\cdot(c+a))^*$
$L(((a+b)\cdot(c+a))^*) = \{aa, ac, ba, bc\}^*$

Expressions rationnelles

Théorème

Un langage $L \subseteq A^*$ est rationnel si et seulement si il existe une expression rationnelle E telle que $L(E) = L$.

Expressions rationnelles

Théorème

Un langage $L \subseteq A^*$ est rationnel si et seulement si il existe une expression rationnelle E telle que $L(E) = L$.

Définition

Deux expressions rationnelles E et F sont *équivalentes* si et seulement si $L(E) = L(F)$.

Équivalence des langages rationnels et reconnaissables

Théorème (Kleene (1956))

Un langage est rationnel si et seulement si il est reconnaissable.

Proposition

Tout langage rationnel est reconnaissable.

Proposition

Tout langage reconnaissable est rationnel.

Équivalence des langages rationnels et reconnaissables

Proposition

Tout langage rationnel est reconnaissable.

Démonstration On le montre par induction structurelle :

- **Base.** $L = \emptyset$ est accepté par :

  $L = \{\varepsilon\}$ est accepté par :

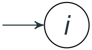
  et $L = \{a\}$ est accepté par :

Équivalence des langages rationnels et reconnaissables

Proposition

Tout langage rationnel est reconnaissable.

Démonstration On le montre par induction structurelle :

- **Base.** $L = \emptyset$ est accepté par : 

  $L = \{\varepsilon\}$ est accepté par :

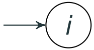
  et $L = \{a\}$ est accepté par :

Équivalence des langages rationnels et reconnaissables

Proposition

Tout langage rationnel est reconnaissable.

Démonstration On le montre par induction structurelle :

- **Base.** $L = \emptyset$ est accepté par : 

  $L = \{\varepsilon\}$ est accepté par : 

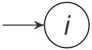
  et $L = \{a\}$ est accepté par :

Équivalence des langages rationnels et reconnaissables

Proposition

Tout langage rationnel est reconnaissable.

Démonstration On le montre par induction structurelle :

- **Base.** $L = \emptyset$ est accepté par : 

  $L = \{\varepsilon\}$ est accepté par : 

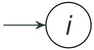
  et $L = \{a\}$ est accepté par : 

Équivalence des langages rationnels et reconnaissables

Proposition

Tout langage rationnel est reconnaissable.

Démonstration On le montre par induction structurelle :

- **Base.** $L = \emptyset$ est accepté par : 

  $L = \{\varepsilon\}$ est accepté par : 

  et $L = \{a\}$ est accepté par : 

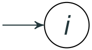
- **Induction.**

Équivalence des langages rationnels et reconnaissables

Proposition

Tout langage rationnel est reconnaissable.

Démonstration On le montre par induction structurelle :

- **Base.** $L = \emptyset$ est accepté par : 

  $L = \{\varepsilon\}$ est accepté par : 

  et $L = \{a\}$ est accepté par : 

- **Induction.** Soient E_1, E_2 et F des expressions rationnelles. Supposons que $L(E_1)$, $L(E_2)$ et $L(F)$ soient reconnaissables. Alors, par les propriétés de clôture des langages reconnaissables, $L(E_1) \cup L(E_2)$, $L(E_1).L(E_2)$ et $L(F)^*$ sont reconnaissables.

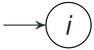
Équivalence des langages rationnels et reconnaissables

Proposition

Tout langage rationnel est reconnaissable.

Démonstration On le montre par induction structurelle :

- **Base.** $L = \emptyset$ est accepté par :



  $L = \{\varepsilon\}$ est accepté par :

  et $L = \{a\}$ est accepté par :

- **Induction.** Soient E_1, E_2 et F des expressions rationnelles. Supposons que $L(E_1)$, $L(E_2)$ et $L(F)$ soient reconnaissables. Alors, par les propriétés de clôture des langages reconnaissables, $L(E_1) \cup L(E_2)$, $L(E_1).L(E_2)$ et $L(F)^*$ sont reconnaissables. Donc $L(E_1 + E_2)$, $L(E_1 \cdot E_2)$ et $L(F^*)$ sont reconnaissables.

Équivalence des langages rationnels et reconnaissables

Proposition

Tout langage reconnaissable est rationnel.

Notations :

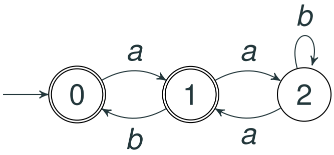
Soit L un langage reconnaissable et $\mathcal{A} = (S, T, I, F)$ un automate tel que $L(\mathcal{A}) = L$.

état de S	Automate	Langage
s	$\mathcal{A}_s = (S, T, \{s\}, F)$	$L_s = L(\mathcal{A}_s)$

Soit $\{s \xrightarrow{a_1} s_1, s \xrightarrow{a_2} s_2, \ldots, s \xrightarrow{a_p} s_p\}$, l'ensemble des transitions sortant de s, alors :

$$L_s = \begin{cases} \{a_1\}.L_{s_1} \cup \cdots \cup \{a_p\}.L_{s_p} \cup \{\varepsilon\} & \text{si } s \text{ est un état final} \\ \{a_1\}.L_{s_1} \cup \cdots \cup \{a_p\}.L_{s_p} & \text{sinon} \end{cases}$$

Équivalence des langages rationnels et reconnaissables



$$\begin{cases} L_0 = \{a\}.L_1 \cup \{\varepsilon\} & (1) \\ L_1 = \{a\}.L_2 \cup \{b\}.L_0 \cup \{\varepsilon\} & (2) \\ L_2 = \{b\}.L_2 \cup \{a\}.L_1 & (3) \end{cases}$$

Équivalence des langages rationnels et reconnaissables

Proposition (Lemme d'Arden)

Soient K et M deux langages tels que $\varepsilon \notin K$. L'équation $X = K.X \cup M$ a pour unique solution $X = K^*.M$

- Montrons que $K^*.M$ est une solution de l'équation :
  $K.(K^*.M) \cup M = (K^+.M) \cup M = (K^+ \cup \{\varepsilon\}).M = K^*.M$
- Montrons que, pour tout L solution de l'équation, $K^*.M \subseteq L$. On montre par récurrence sur n : $P(n)$: "pour tout $u \in K^n.M$, $u \in L$". On sait que $L = K.L \cup M$. Pour $n = 0$, si $u \in M$, alors $u \in K.L \cup M = L$. Soit $n \geq 0$, et supposons $P(n)$. Soit $u \in K^{n+1}.M$, alors $u = vw$ avec $v \in K$ et $w \in K^n.M$. Par hypothèse de récurrence, $w \in L$, donc $u = vw \in KL \subseteq L$.
- Montrons que, pour tout L solution de l'équation, $L \subseteq K^*.M$. Supposons, par l'absurde, que ce n'est pas le cas et prenons u le plus petit mot de $L \setminus \{K^*.M\}$.Si $u \in M$, alors $u \in K^*.M$, impossible. Donc $u \in K.L$. Alors $u = vw$ avec $v \in K$ et $w \in L$. Comme $\varepsilon \notin K$, $|w| < |u|$ et $w \in K^*.M$. Donc $u = vw \in K.K^*.M \subseteq K^*.M$. Impossible.

Équivalence des langages rationnels et reconnaissables

Proposition

Tout langage reconnaissable est rationnel.

On remarque que $L = \bigcup_{i \in I} L_i$. On résout donc le système donné par les équations

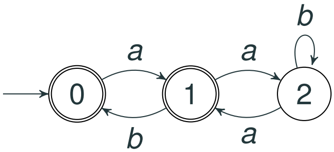
$$L_s = \begin{cases} \{a_1\}.L_{s_1} \cup \cdots \cup \{a_p\}.L_{s_p} \cup \{\varepsilon\} & \text{si } s \text{ est un état final} \\ \{a_1\}.L_{s_1} \cup \cdots \cup \{a_p\}.L_{s_p} & \text{sinon} \end{cases}$$

en utilisant un pivot de Gauss et le lemme d'Arden. On peut montrer par induction que les solutions de ce système sont des langages rationnels. On s'autorise donc à manipuler des expressions rationnelles et on résout plutôt le système :

$$L_s = \begin{cases} a_1.L_{s_1} + \cdots + a_p.L_{s_p} + \varepsilon & \text{si } s \text{ est un état final} \\ a_1.L_{s_1} + \cdots + a_p L_{s_p} & \text{sinon} \end{cases}$$

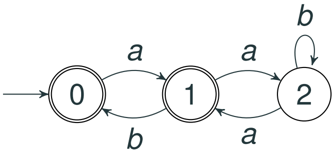
On donne ensuite l'expression rationnelle $\Sigma_{i \in I} L_i$

Équivalence des langages rationnels et reconnaissables



$$\begin{cases} L_0 & = aL_1 + \varepsilon & (1) \\ L_1 & = aL_2 + bL_0 + \varepsilon & (2) \\ L_2 & = bL_2 + aL_1 & (3) \end{cases}$$

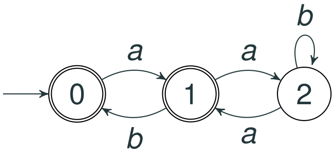
Équivalence des langages rationnels et reconnaissables



$$\begin{cases} L_0 &= aL_1 + \varepsilon & (1) \\ L_1 &= aL_2 + bL_0 + \varepsilon & (2) \\ L_2 &= bL_2 + aL_1 & (3) \end{cases}$$

On applique le lemme d'Arden sur l'équation (3).

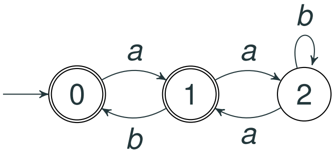
Équivalence des langages rationnels et reconnaissables



$$\begin{cases} L_0 &= aL_1 + \varepsilon & (1) \\ L_1 &= aL_2 + bL_0 + \varepsilon & (2) \\ L_2 &= b^* aL_1 & (3) \end{cases}$$

On applique le lemme d'Arden sur l'équation (3). On remplace L_2 dans (2).

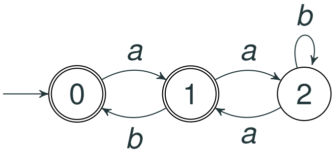
Équivalence des langages rationnels et reconnaissables



$$\begin{cases} L_0 & = aL_1 + \varepsilon \qquad (1) \\ L_1 & = ab^* aL_1 + bL_0 + \varepsilon \qquad (2) \\ L_2 & = b^* aL_1 \qquad (3) \end{cases}$$

On applique le lemme d'Arden sur l'équation (3). On remplace L_2 dans (2). On applique le lemme d'Arden sur l'équation (2).

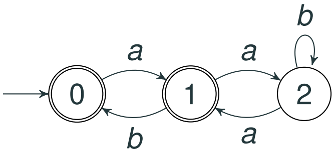
Équivalence des langages rationnels et reconnaissables



$$\begin{cases} L_0 = aL_1 + \varepsilon & (1) \\ L_1 = (ab^*a)^*(bL_0 + \varepsilon) & (2) \\ L_2 = b^*aL_1 & (3) \end{cases}$$

On applique le lemme d'Arden sur l'équation (3). On remplace L_2 dans (2). On applique le lemme d'Arden sur l'équation (2). On remplace L_1 dans (1).

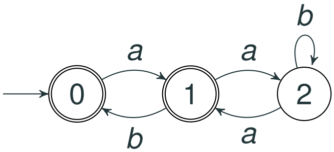
Équivalence des langages rationnels et reconnaissables



$$\begin{cases} L_0 &= a(ab^*a)^*(bL_0+\varepsilon)+\varepsilon & (1)\\ L_1 &= (ab^*a)^*(bL_0+\varepsilon) & (2)\\ L_2 &= b^*aL_1 & (3)\end{cases}$$

On applique le lemme d'Arden sur l'équation (3). On remplace L_2 dans (2). On applique le lemme d'Arden sur l'équation (2). On remplace L_1 dans (1).

Équivalence des langages rationnels et reconnaissables

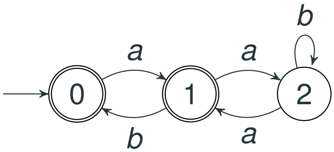


$$\begin{cases} L_0 &= a(ab^*a)^*(bL_0+\varepsilon)+\varepsilon & (1) \\ L_1 &= (ab^*a)^*(bL_0+\varepsilon) & (2) \\ L_2 &= b^*aL_1 & (3) \end{cases}$$

On applique le lemme d'Arden sur l'équation (3). On remplace L_2 dans (2). On applique le lemme d'Arden sur l'équation (2). On remplace L_1 dans (1). On applique une dernière fois le lemme d'Arden à

$$L_0 = a(ab^*a)^*(bL_0+\varepsilon)+\varepsilon = a(ab^*a)^*bL_0 + a(ab^*a)^* + \varepsilon$$

Équivalence des langages rationnels et reconnaissables



$$\begin{cases} L_0 = a(ab^*a)^*(bL_0+\varepsilon)+\varepsilon & (1) \\ L_1 = (ab^*a)^*(bL_0+\varepsilon) & (2) \\ L_2 = b^*aL_1 & (3) \end{cases}$$

On applique le lemme d'Arden sur l'équation (3). On remplace L_2 dans (2). On applique le lemme d'Arden sur l'équation (2). On remplace L_1 dans (1). On applique une dernière fois le lemme d'Arden à

$$L_0 = a(ab^*a)^*(bL_0+\varepsilon)+\varepsilon = a(ab^*a)^*bL_0 + a(ab^*a)^*+\varepsilon$$

$$L_0 = (a(ab^*a)^*b)^*(a(ab^*a)^*+\varepsilon)$$

Langages non reconnaissables

Théorème

Il existe des langages non reconnaissables.

Langages non reconnaissables

Théorème

Il existe des langages non reconnaissables.

Démonstration Le langage $L = \{a^n b^n \mid n \in \mathbb{N}\}$ n'est pas reconnaissable. Supposons que L est reconnaissable. Alors il existe un automate $\mathcal{A}$ tel que $L(\mathcal{A}) = L$. Soit N le nombre d'états de $\mathcal{A}$. Le mot $w = a^N b^N \in L$ est accepté par $\mathcal{A}$. Soit $s_0 \xrightarrow{a} s_1 \xrightarrow{a} s_2 \cdots \xrightarrow{a} s_N \xrightarrow{b} \cdots \xrightarrow{b} s_{2N}$ une exécution acceptante de $\mathcal{A}$ sur w, où s_0 est un état initial et s_{2N} un état acceptant. Parmi les $N+1$ états $s_0, \ldots, s_N$, il y en a (au moins) deux égaux. On note $s_i = s_j$ pour $i < j$, ce qui correspond à une boucle dans l'automate. Donc on peut réécrire l'exécution ci-dessus en : $s_0 \xrightarrow{a^{p_1}} s_i \xrightarrow{a^{p_2}} s_i \xrightarrow{a^{p_3}} s_N \xrightarrow{b^N} s_{2N}$, avec $p_1 + p_2 + p_3 = N$ et $p_2 > 0$, et on peut construire une autre exécution acceptante : $s_0 \xrightarrow{a^{p_1}} s_i \xrightarrow{a^{p_3}} s_N \xrightarrow{b^N} s_{2N}$. L'étiquette de cette exécution est le mot $w' = a^{p_1} a^{p_3} b^N$, avec $w' \in \mathcal{A}$, mais $p_1 + p_3 < N$, donc $w' \notin L$. CONTRADICTION $\square$

Minimisation d'automates

Combien d'automates reconnaissent le langage a^* ?

Minimisation d'automates

Combien d'automates reconnaissent le langage a^* ?

Soit $\mathcal{A} = (S, T, i, F)$ un automate *déterministe* et *complet*. Pour tout $s \in S$, $u \in A^*$, on note $s \cdot u$ l'unique état s' tel que $s \xrightarrow{u} s'$.

Définition

Soit $\mathcal{A} = (S, T, i, F)$ un automate déterministe et complet. Alors, pour tous $s, s' \in S$, on note $s \sim s'$ si et seulement si $L_s = L_{s'}$. Autrement dit, si, pour tout $u \in A^*$, $s \cdot u \in F$ si et seulement si $s' \cdot u \in F$.

Minimisation d'automates

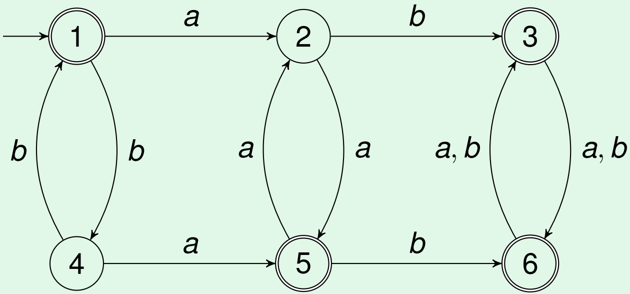
Proposition

La relation $\sim \subseteq S \times S$ *est une relation d'équivalence (équivalence de Nérode). De plus,*

- *pour tout* $a \in A$, *si* $s_1 \sim s_2$ *alors* $s_1.a \sim s_2.a$
- *si* $s_1 \sim s_2$ *alors* $s_1 \in F$ *ssi* $s_2 \in F$

Minimisation d'automates

Exemple



Minimisation d'automates

Définition

Soit $\mathcal{A} = (S, T, i, F)$ un automate déterministe, complet, et monogène. Soit $\sim$ la relation d'équivalence de Nérode. *L'automate quotient* de $\mathcal{A}$ est l'automate $\mathcal{A}_{/\sim} = (S_{/\sim}, \tilde{T}, [i], \tilde{F})$ avec

- $\tilde{T} = \{([s], a, [s']) \mid (s, a, s') \in T\}$
- $\tilde{F} = \{[f] \mid f \in F\}$

Proposition

Soit $\mathcal{A}$ un automate déterministe, complet, et monogène. Alors $L(\mathcal{A}) = L(\mathcal{A}_{/\sim})$ et $\mathcal{A}_{/\sim}$ est l'automate minimal de $\mathcal{A}$.

Minimisation d'automates - algorithme de Moore

On définit des classes d'équivalence de façon itérative :

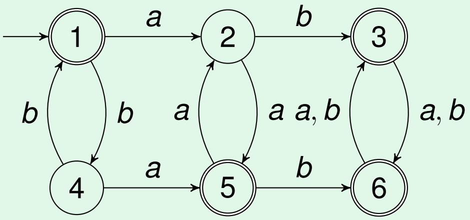
- $s \sim_0 s'$ si et seulement si $s \in F$ ssi $s' \in F$
- $s \sim_{k+1} s'$ si et seulement si $\begin{array}{ll} (i) & s \sim_k s' \\ (ii) & s.a \sim_k s'.a \text{ pour tout } a \in A \end{array}$

Proposition

Il existe $p \in \mathbb{N}$ tel que $\sim_{p+1} = \sim_p$ et l'équivalence de Nérode est $\sim = \sim_p$.

Minimisation d'automates

Exemple



Minimisation d'automates

Exemple