Théorie des langages rationnels : THLR

CM 3

Uli Fahrenberg

EPITA Rennes

S3 2022

Aperçu

Programme du cours

1. Mots, langages
2. Langages rationnels, expressions rationnelles
3. Automates finis
4. Langages non-rationnels
5. Langages reconnaissables, minimisation

Dernièrement : L'algèbre de mots

Soit Σ un ensemble fini.

- on appelle les éléments $a, b, \ldots \in \Sigma$ des symboles

On dénote Σ^* l'ensemble de tous les suites finies d'éléments de Σ.

- donc $\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \cdots = \bigcup_{n \geq 0} \Sigma^n$
- on appelle les éléments $u, v, w, \ldots \in \Sigma^*$ des mots

La concaténation de deux mots $a_1 \ldots a_n$ et $b_1 \ldots b_m$ est le mot $a_1 \ldots a_n b_1 \ldots b_m$.

- ε : le mot vide
- l'opération « . » sur mots est associative et a ε comme élément neutre de deux côtés

La longueur $|u|$ d'un mot $u \in \Sigma^*$: le nombre de symboles de u.

- $|\varepsilon| = 0$ et $|uv| = |u| + |v|$
- u^n : la concaténation de n copies de u
- $|u^n| = n\,|u|$

Dernièrement : L'algèbre de langages

Un langage est un sous-ensemble $L \subseteq \Sigma^*$.

- opérations ensemblistes : $L_1 \cup L_2$, $L_1 \cap L_2$, $\overline{L}$
- concaténation : $L_1\, L_2 = \{u_1\, u_2 \mid u_1 \in L_1, u_2 \in L_2\}$
- $L^n = L \cdots L$ (n copies de L)
- étoile de Kleene : $L^* = L^0 \cup L^1 \cup L^2 \cup \cdots = \bigcup_{n \geq 0} L^n$

L'opération « . » sur langages est associative et a $\{\varepsilon\}$ comme élément neutre de deux côtés.

- $L.\emptyset = \emptyset.L = \emptyset$
- $\emptyset^* = \{\varepsilon\}^* = \{\varepsilon\}$

5 minutes de réflexion

Vrai ou faux ?

1. $\{a\}^n = \{a^n\}$
2. $\{a, b\}^n = \{a^n, b^n\}$
3. L^* est un ensemble infini pour tout $L \subseteq \Sigma^*$
4. pour tout $L \subseteq \Sigma^*$ et $n \in \mathbb{N}$, l'ensemble $\{u \in L \mid |u| \leq n\}$ est un ensemble fini
5. $\{a, b\}^* = \{a\}^*\{b\}^*$
6. $\{a, b\}^* = \big(\{a\}^*\{b\}\big)^*\{a\}^*$
7. $(L_1 \cup L_2)^* = \big(L_1^* L_2\big)^* L_1^*$

5 minutes de réflexion

Vrai ou faux ?

1. $\{a\}^n = \{a^n\}$ ✓
2. $\{a, b\}^n = \{a^n, b^n\}$ ✗
3. L^* est un ensemble infini pour tout $L \subseteq \Sigma^*$ ✗
4. pour tout $L \subseteq \Sigma^*$ et $n \in \mathbb{N}$, l'ensemble $\{u \in L \mid |u| \leq n\}$ est un ensemble fini ✓
5. $\{a, b\}^* = \{a\}^*\{b\}^*$ ✗
6. $\{a, b\}^* = \big(\{a\}^*\{b\}\big)^*\{a\}^*$ ✓
7. $(L_1 \cup L_2)^* = \big(L_1^* L_2\big)^* L_1^*$ ✓

Langages rationnels

Opérations rationnelles

Soit Σ un alphabet, on travaille avec des langages dans $\mathcal{P}(\Sigma^*)$.

Définition

Les opérations rationnelles dans $\mathcal{P}(\Sigma^*)$ sont $\cup$, . et *.

- donc union, concaténation et étoile de Kleene

Théorème (pour plus tard) : Toutes les autres opérations sont exprimables par $\cup$, . et *.

Exemples :

- $\text{Pref}(\{a\}\{b\}^*) = \{\varepsilon\} \cup \{a\} \cup \{a\}\{b\}^*$
- $\text{Suff}(\{a\}\{b\}^*) =$
- $\{a\}\{b\}^* \cap \{a\}^*\{b\} =$

Opérations rationnelles

Soit Σ un alphabet, on travaille avec des langages dans $\mathcal{P}(\Sigma^*)$.

Définition

Les opérations rationnelles dans $\mathcal{P}(\Sigma^*)$ sont $\cup$, . et *.

- donc union, concaténation et étoile de Kleene

Théorème (pour plus tard) : Toutes les autres opérations sont exprimables par $\cup$, . et *.

Exemples :

- $\text{Pref}(\{a\}\{b\}^*) = \{\varepsilon\} \cup \{a\} \cup \{a\}\{b\}^*$
- $\text{Suff}(\{a\}\{b\}^*) = \{b\}^* \cup \{a\}\{b\}^*$
- $\{a\}\{b\}^* \cap \{a\}^*\{b\} =$

Opérations rationnelles

Soit Σ un alphabet, on travaille avec des langages dans $\mathcal{P}(\Sigma^*)$.

Définition

Les opérations rationnelles dans $\mathcal{P}(\Sigma^*)$ sont $\cup$, . et *.

- donc union, concaténation et étoile de Kleene

Théorème (pour plus tard) : Toutes les autres opérations sont exprimables par $\cup$, . et *.

Exemples :

- $\text{Pref}(\{a\}\{b\}^*) = \{\varepsilon\} \cup \{a\} \cup \{a\}\{b\}^*$
- $\text{Suff}(\{a\}\{b\}^*) = \{b\}^* \cup \{a\}\{b\}^*$
- $\{a\}\{b\}^* \cap \{a\}^*\{b\} = \{a\}\{b\}$

Langages rationnels

Définition (3.1)

Les langages rationnels sur Σ sont définis inductivement comme suite :

1. $\emptyset$ et $\{\varepsilon\}$ sont des langages rationnels
2. pour tout $a \in \Sigma$, $\{a\}$ est un langage rationnel
3. si L_1 et L_2 sont des langages rationnels, alors $L_1 \cup L_2$, $L_1 . L_2$ et L_1^* le sont également

- $\{\varepsilon\} = \emptyset^* \Rightarrow$ on peut enlever $\{\varepsilon\}$ de la définition

Lemme

L est rationnel si et seulement si

- *$L = \emptyset$ ou $L = \{a\}$ pour un $a \in \Sigma$ ou*
- *$L = L_1 \cup L_2$, $L = L_1 L_2$ ou $L = L_1^*$ pour L_1 et L_2 rationnels.*

(En quoi ce lemme est-il différent de la définition ?)

Rationalité

Théorème

Si L_1 *et* L_2 *sont des langages rationnels, alors* $L_1 \cap L_2$, $\overline{L}_1$, $\mathrm{Pref}(L_1)$, $\mathrm{Suff}(L_1)$ *et* $\mathrm{Fact}(L_1)$ *le sont aussi.*

- pour la démonstration faut attendre quelques semaines

5 minutes de réflexion

Rationnel ou pas rationnel, sur alphabet $\Sigma = \{a, b, c\}$?

1. $\{a, b, abcba\}$
2. $\{a^n \mid n \geq 0\}$
3. $\{w \in \Sigma^* \mid w \text{ contient au moins trois } a\}$
4. $\{w \in \Sigma^* \mid |w| \geq 5\}$
5. $\{a^{2n} \mid n \geq 0\}$
6. $\{a^{n^2} \mid n \geq 0\}$
7. $\{a^m b^n \mid m, n \geq 0\}$
8. $\{a^n b^n \mid n \geq 0\}$

5 minutes de réflexion

Rationnel ou pas rationnel, sur alphabet $\Sigma = \{a, b, c\}$?

1. $\{a, b, abcba\}$ ✓
2. $\{a^n \mid n \geq 0\}$ ✓
3. $\{w \in \Sigma^* \mid w \text{ contient au moins trois } a\}$ ✓
4. $\{w \in \Sigma^* \mid |w| \geq 5\}$ ✓
5. $\{a^{2n} \mid n \geq 0\}$ ✓
6. $\{a^{n^2} \mid n \geq 0\}$ ✗
7. $\{a^m b^n \mid m, n \geq 0\}$ ✓
8. $\{a^n b^n \mid n \geq 0\}$ ✗

Expressions rationnelles

Expressions rationnelles

Une notation pratique pour des langages rationnels :

Définition (3.2)

Les expressions rationnelles sur Σ sont définis inductivement comme suite :

1. $\emptyset$ et ε sont des expressions rationnelles
2. pour tout $a \in \Sigma$, a est une expression rationnelle
3. si e_1 et e_2 sont des expressions rationnelles, alors $e_1 + e_2$, $e_1 . e_2$ et e_1^* le sont également

Expressions rationnelles

Une notation pratique pour des langages rationnels :

Définition (3.1, *recall*)

Les langages rationnels sur Σ sont définis inductivement comme suite :

1. $\emptyset$ et $\{\varepsilon\}$ sont des langages rationnels
2. pour tout $a \in \Sigma$, $\{a\}$ est un langage rationnel
3. si L_1 et L_2 sont des langages rationnels , alors $L_1 \cup L_2$, $L_1.L_2$ et L_1^* le sont également

- presque la même chose ! mais
- 3.1 introduit une classe de sous-ensembles de Σ^*,
- 3.2 définit des expressions syntaxiques

Expressions rationnelles

Une notation pratique pour des langages rationnels :

Définition (3.2)

Les expressions rationnelles sur Σ sont définis inductivement comme suite :

1. $\emptyset$ et ε sont des expressions rationnelles
2. pour tout $a \in \Sigma$, a est une expression rationnelle
3. si e_1 et e_2 sont des expressions rationnelles, alors $e_1 + e_2$, $e_1 . e_2$ et e_1^* le sont également

- presque la même chose ! mais
- 3.1 introduit une classe de sous-ensembles de Σ^*,
- 3.2 définit des expressions syntaxiques

On va relier les deux en donnant une sémantique aux expressions rationnelles.

Sémantique

Définition

Le langage dénoté par une expression rationnelle e sur Σ est $L(e) \subseteq \Sigma^*$ définit inductivement comme suite :

1. $L(\varnothing) = \emptyset$, $L(\varepsilon) = \{\varepsilon\}$
2. $L(a) = \{a\}$ pour tout $a \in \Sigma$
3. $L(e_1 + e_2) = L(e_1) \cup L(e_2)$, $L(e_1 . e_2) = L(e_1) . L(e_2)$, $L(e^*) = (L(e))^*$

Théorème

$L \subseteq \Sigma^*$ est rationnel ssi il existe une expression rationnelle e telle que $L = L(e)$.

Démonstration.

Par induction structurelle ...

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\mathrm{Pref}(L)$ *l'est aussi.*

Démonstration.

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\mathrm{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\mathrm{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.
2. Nous construirons une expression rationnelle $\mathrm{pref}(e)$ telle que $L(\mathrm{pref}(e)) = \mathrm{Pref}(L)$.

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\mathrm{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.
2. Nous construirons une expression rationnelle $\mathrm{pref}(e)$ telle que $L(\mathrm{pref}(e)) = \mathrm{Pref}(L)$.
3. ... par induction structurelle :

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\mathrm{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.
2. Nous construirons une expression rationnelle $\mathrm{pref}(e)$ telle que $L(\mathrm{pref}(e)) = \mathrm{Pref}(L)$.
3. ... par induction structurelle :
4. $\mathrm{pref}(\emptyset) =$, $\mathrm{pref}(\varepsilon) =$, $\mathrm{pref}(a) =$

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\mathrm{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.
2. Nous construirons une expression rationnelle $\mathrm{pref}(e)$ telle que $L(\mathrm{pref}(e)) = \mathrm{Pref}(L)$.
3. … par induction structurelle :
4. $\mathrm{pref}(\varnothing) = \varnothing$, $\mathrm{pref}(\varepsilon) = \varepsilon$, $\mathrm{pref}(a) = a + \varepsilon$

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\mathrm{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.
2. Nous construirons une expression rationnelle $\mathrm{pref}(e)$ telle que $L(\mathrm{pref}(e)) = \mathrm{Pref}(L)$.
3. . . . par induction structurelle :
4. $\mathrm{pref}(\emptyset) = \emptyset$, $\mathrm{pref}(\varepsilon) = \varepsilon$, $\mathrm{pref}(a) = a + \varepsilon$
5. $\mathrm{pref}(e_1 + e_2) =$
   $\mathrm{pref}(e_1 e_2) =$
   $\mathrm{pref}(e^*) =$

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\text{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.
2. Nous construirons une expression rationnelle $\text{pref}(e)$ telle que $L(\text{pref}(e)) = \text{Pref}(L)$.
3. ... par induction structurelle :
4. $\text{pref}(\emptyset) = \emptyset$, $\text{pref}(\varepsilon) = \varepsilon$, $\text{pref}(a) = a + \varepsilon$
5. $\text{pref}(e_1 + e_2) = \text{pref}(e_1) + \text{pref}(e_2)$
   $\text{pref}(e_1 e_2) =$
   $\text{pref}(e^*) =$

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\mathrm{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.
2. Nous construirons une expression rationnelle $\mathrm{pref}(e)$ telle que $L(\mathrm{pref}(e)) = \mathrm{Pref}(L)$.
3. ... par induction structurelle :
4. $\mathrm{pref}(\emptyset) = \emptyset$, $\mathrm{pref}(\varepsilon) = \varepsilon$, $\mathrm{pref}(a) = a + \varepsilon$
5. $\mathrm{pref}(e_1 + e_2) = \mathrm{pref}(e_1) + \mathrm{pref}(e_2)$
   $\mathrm{pref}(e_1 e_2) = \mathrm{pref}(e_1) + e_1\, \mathrm{pref}(e_2)$
   $\mathrm{pref}(e^*) =$

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\mathrm{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.
2. Nous construirons une expression rationnelle $\mathrm{pref}(e)$ telle que $L(\mathrm{pref}(e)) = \mathrm{Pref}(L)$.
3. ... par induction structurelle :
4. $\mathrm{pref}(\varnothing) = \varnothing$, $\mathrm{pref}(\varepsilon) = \varepsilon$, $\mathrm{pref}(a) = a + \varepsilon$
5. $\mathrm{pref}(e_1 + e_2) = \mathrm{pref}(e_1) + \mathrm{pref}(e_2)$
   $\mathrm{pref}(e_1 e_2) = \mathrm{pref}(e_1) + e_1\,\mathrm{pref}(e_2)$
   $\mathrm{pref}(e^*) = e^*\mathrm{pref}(e)$ (voir tableau)

Pour en finir, une autre démonstration

Théorème

Si L est un langage rationnel, alors $\text{Pref}(L)$ *l'est aussi.*

Démonstration.

1. Soit e une expression rationnelle telle que $L = L(e)$.
2. Nous construirons une expression rationnelle $\text{pref}(e)$ telle que $L(\text{pref}(e)) = \text{Pref}(L)$.
3. ... par induction structurelle :
4. $\text{pref}(\emptyset) = \emptyset$, $\text{pref}(\varepsilon) = \varepsilon$, $\text{pref}(a) = a + \varepsilon$
5. $$\begin{aligned} \text{pref}(e_1 + e_2) &= \text{pref}(e_1) + \text{pref}(e_2) \\ \text{pref}(e_1 e_2) &= \text{pref}(e_1) + e_1\, \text{pref}(e_2) \\ \text{pref}(e^*) &= e^* \text{pref}(e) \qquad (\text{ voir tableau }) \end{aligned}$$
6. Maintenant il faut démontrer que, en fait, $L(\text{pref}(e)) = \text{Pref}(L)$.
7. ... par induction structurelle, encore ...


That's all Folks!