Université Paris Cité – UFR d'Informatique
L2 Informatique, Math-Info, Bio-Info, Japonais-Info

Automates et analyse lexicale
Année 2022–2023

TD n°8

Lemme de l'Étoile & Propriétés de Clôture

Exercice 1 (Lemme de l'étoile) *On rappelle le lemme de l'étoile :*

> *Soit $\mathcal{L}$ un langage reconnaissable. Il existe un entier N tel que tout mot $u \in \mathcal{L}$ de taille supérieure ou égale à N admet une factorisation $u = xyz$ satisfaisant :*
> — $y \neq \epsilon$ *et* $|xy| \leq N$
> — $xy^kz \in \mathcal{L}$ *pour tout entier* $k \geq 0$.

Pour chacun des langages suivants, montrer s'il est reconnaissable ou non.

1. $\{a^m b^n : m, n \in \mathbb{N}\}$ rationnel
2. $\{a^m b^n : m < n\}$
3. $\{a^n b^n\}$
4. $\{u^2 : u \in \{a, b\}^*\}$
5. $\{a^{2n} : n \in \mathbb{N}\}$ rationnel
6. $\{a^{n^2} : n \in \mathbb{N}\}$
7. $\{a^p : p \text{ premier}\}$

Exercice 2 (Clôture par préfixes, suffixes...) *La* clôture sous préfixe *d'un langage L est définie comme*

$$Pref(L) = \{u \mid \textit{il existe } v \textit{ tel que } u \cdot v \in L\}$$

Il s'agit de l'ensemble des préfixes des mots de L.

1. *Montrer que si L est reconnaissable, alors $Pref(L)$ est également reconnaissable. On pourra par exemple donner un algorithme pour transformer un automate pour L en un automate pour $Pref(L)$.*
2. *La réciproque est-elle est vraie ?*
3. *Montrer que si L est reconnaissable, alors l'ensemble des suffixes de L, l'ensemble des facteurs de L et l'ensemble des sous-mots de L sont également reconnaissables. On pourra procéder de manière similaire à la question 1, ou également utiliser de manière astucieuse les propriétés de clôture de Rec.*

Exercice 3 ((*) Langage miroir) *Le langage miroir d'un langage $\mathcal{L}$ est le langage $\widetilde{\mathcal{L}} = \{\widetilde{u} \mid u \in \mathcal{L}\}$, avec $\widetilde{u} = x_{n-1} \cdots x_0$ pour $u = x_0 \cdots x_{n-1}$. Par exemple pour le langage*

$$\mathcal{L}_1 = \{w \in \{a, b\}^* \mid \textit{la troisième lettre de } w \textit{ est un } b\}.$$

on a

$$\widetilde{\mathcal{L}_1} = \{w \in \{a, b\}^* \mid \textit{la troisième lettre de } w \textit{ à partir de la fin est un } b\}.$$

1. *Donner une expression rationnelle pour $\mathcal{L}_1$ et une pour $\widetilde{\mathcal{L}_1}$.*
2. *Décrire un procédé permettant de construire en général une expression rationnelle pour $\widetilde{\mathcal{L}}$ à partir de l'expression rationnelle pour $\mathcal{L}$.*
3. *Décrire un procédé permettant de construire l'automate qui reconnaît le langage $\widetilde{\mathcal{L}}$ étant donné celui de $\mathcal{L}$. Est-ce qu'en commençant avec un automate déterministe pour $\mathcal{L}$ on obtient toujours un automate déterministe pour $\widetilde{\mathcal{L}}$?*

Lemme de l'étoile

Tous les langages rationels satisfait la propriété d'itération

Pour montrer qu'un langage <u>n'est pas rationnel</u> avec le lemme de l'étoile, on fait un raisonnement par l'absurde

C'est à dire, on suppose qu'il satisfait la propriété d'itération et on arrive à une contradiction

Propriété d'itération de $\mathcal{L}$

$\exists N \in \mathbb{N} \setminus \{0,1\}$

$\forall u,\ u \in \mathcal{L}$ et $|u| \geq N$

$\exists xyz,\ y \neq \varepsilon$ et $|xy| \leq N$ et $u = xyz$

$\forall k,\ xy^kz \in \mathcal{L}$

Ex 1 :

2) $\{a^m b^n : m < n\}$

Prenons un entier N suffisamment grand

Pour tout mot u du langage tel que $|u| > N$, il existe x, y, z tel que $u = xyz$

Par l'absurde, on suppose que $\mathcal{L}$ est rationnel
D'apres le lemme de l'étoile, il satisfait la propriété d'itération
Soit N la constante associés

On pose $u = a^N b^{N+1}$, $u \in \mathcal{L}$ et $|u| \geq N$
Soient x, y, z associés, tel que $y \neq \varepsilon$
$|xy| \leq N$ et $u = xyz$

Donc $xyz = a^N b^{N+1}$, a, comme $|xy| \leq N$
$\exists i\ j$ tels que

$$x = a^i \quad y = a^j$$

$i + j \leq N$ et $j \neq 0$ car $y \neq \varepsilon$

a	b
N	$N+1$

$x \mid y \mid z$

$a^i \quad a^j \quad a^{N-i-j} b^{N+1}$

Brouillon

$$xy^0z = a^i (a^j)^0 a^{N-i-j} b^{N+1}$$
$$= a^{N-j} b^{N+1}$$ Marche pas

$$xy^2z = a^i (a^j)^2 a^{N-i-j} b^{N+1}$$
$$= a^{N+j} b^{N+1} \quad \checkmark$$

où $m < n$

$N+j \geq N+1$ donc on peut quitter le langage

On pose $k=2$

$$xy^kz = a^i a^{2j} a^{N-i-j} b^{N+1}$$
$$= a^{N+j} b^{N+1}$$

Or, $j \neq 0$, car $y \neq \varepsilon$

D'après la propriété d'itération $xy^kz \in L$

c'est absurde

4) $\{u^2 : u \in \{a,b\}^*\}$

$u \in \Sigma^*$

u	u
$b^N a$	$b^N a$

$x \mid y \mid z$

Par l'absurde, on suppose que $\mathcal{L}$ est rationel
D'apres le lemme de l'étoile, il satisfait la propriété d'itération
Soit N la constante associés

On pose $z = b^N a b^N a$, $z \in \mathcal{L}$ et

$|z| \geq N$

Soient x, y, z associés tels que $y \neq \varepsilon$
$|xy| \leq N$ et $z = xyz$

Donc $xyz = b^N a b^{N+1} a$, comme $|xy| \leq N$

$\exists i, j$ tel que $x = b^i$ $y = b^j$ $z = b^{N-i-j} a b^N a$

$i + j \leq N$ et $j \neq 0$ car $y \neq \varepsilon$

$xy^k z = b^i b^j b^{N-i-j} a b^N a$

$xy^0 z = b^i (b^j)^0 b^{N-i-j} a b^N a$

$= b^{N-j} a b^N a$ Marche ✓

Or $j \neq 0$, car $y \neq \varepsilon$

D'après la propriété d'itération

$xy^k z \in \mathcal{L}$

c'est absurde

3) $\{a^n b^n\} \mathcal{L}_3$

Par l'absurde, supposons que $\mathcal{L}_3$ est rationnel

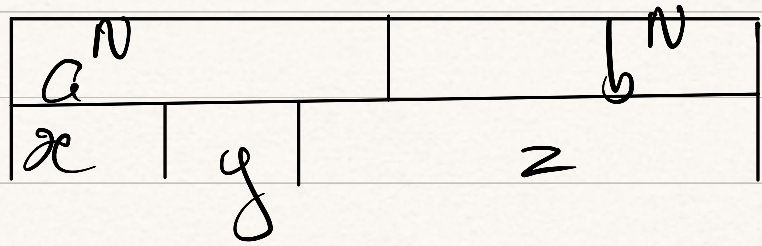
Il satisfait donc la propriété d'itération

Soit N la constante associée

On pose $u = a^N b^N$, $u \in \mathcal{L}_3$ et $|u| \geq N$

Soit x, y, z associées tel que $y \neq \varepsilon$

On sait que $|xy| \leq N$, $u = xyz$ et $y \neq \varepsilon$



De plus, comme $|xy| \leq N$, il existe i et j tel que $x = a^i$ $\quad y = a^j$

et $z = a^{N-j-i} b^N$

On pose $k = 0$

D'après la propriété d'itération :

$$xy^0z = xz \in \mathcal{L}_3, \text{ or}$$

$$xz = a^i a^{N-i-j} b^N = a^{N-j} b^N$$

et $j \neq 0$ ($y \neq \varepsilon$) contradiction

6) $\{a^{n^2} : n \in \mathbb{N}\}$

Par l'absurde, supposons que $\mathcal{L}_3$ est rationnel
Il satisfait donc la propriété d'itération
Soit N la constante associée

On pose $u = a^{N^2}$ $\quad u \in \mathcal{L}_6$ et $|u| \geq N$
Soit x, y, z le découpage associé à u

$$|xy| < N \qquad y \neq \varepsilon \qquad u = xyz$$

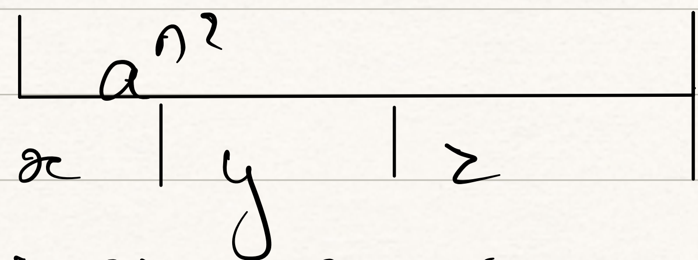
On en déduit qu'il existe i et j tels que

$$x = a^i \qquad y = a^j$$

$$z = a^{N^2-i-j}$$

On pose $k = 2$
D'après la propriété d'itération

$$xy^2z \in \mathcal{L}_6$$



Or $xy^2z = a^i a^{2j} a^{N^2-i-j}$
$= a^{N^2+j}$
$=$

et $N^2 < N^2 + j < (N+j)^2$

car $y \neq \varepsilon$ (donc $j \neq 0$)

$|xy| \leq N$ donc $j \leq N$

donc $N^2 + j$ n'est pas de la forme n^2, $n \in \mathbb{N}$

Ou $u = a^{(n+1)^2}$

$N = n^2$

$(n+1)^2$

$n^2 + 2n + 1$

$u = a^{(n+1)^2}$
$= a^{2n} a a^{n^2}$

Si $u = xyz$

$x = a^{2n}$ $y = a$ $z = a^{n^2}$

n^2

a^n

x	y	z
a^{2n}	a	a^{n^2}
a^i	a^j	a^{N^2-j-i}

$xy^2z = 2n + 2 + n^2$

$5 \succ \{a^{2n} : n \in \mathbb{N}\}$

Par l'absurde, on suppose que L est rationel
D'apres le lemme de l'étoile, il satisfait
la propriété d'itération
Soit N la constante associés

On pose z

Pour aller plus loin dans la non-rationalité...

Démontrer la non-rationalité d'un langage en utilisant les propriétés de clôture de

Rat = Rec.

Le lemme d'itération n'est pas le seul outil pour montrer qu'un langage n'est pas rationnel.

Pour cela on peut aussi utiliser les propriétés de clôture de la famille Rat = Rec à condition

de connaître déjà quelques langages non rationnels

Vous connaissez déjà beaucoup de ces propriétés dont certaines dérivent de la définition

même de Rat = Rec alors que d'autres ont été démontrées ou indiquées en cours ou en TD.

On sait notamment que Rat = Rec est clos sous : U, ,* , n, C (complémentaire), \ (différence

d'ensembles), A (différence symétrique), ~ (miroir), par préfixes,

Vous connaissez déjà également certains langages non rationnels dont vous pourrez vous

servir, par exemple le langage Lo = {a°br n E N' que l'on a déjà démontré non rationnel par

le lemme de l'étoile (Exol ci-dessus).

Exemple d'application.

On veut montrer que La = {w € {a, b}*||wla = |w|} n'est pas rationnel.

On remarque que : Li Ma** = Lo. Or, si Li était rationnel alors son intersection avec un autre

langage rationnel (dans ce cas a**) le serait aussi car Rat = Rec est clos sous intersection.

Mais

on sait que cette intersection est Lo, qui n'est pas rationnel, donc L ne peut pas l'être.

Note. Pour appliquer cette technique il faut exprimer un langage dont on a déjà montré la

non-rationnalité (dans l'exemple [o) en fonction du langage dont on veut montrer la non-

rationnalité (dans l'exemple

L1), de langages rationnels (dans l'exemple a*b*) et d'opérations sous lesquelles Rat = Rec est fermée (dans l'exemple M, mais parfois on a besoin d'utiliser plusieurs opérations).

Propriété de clôture

Ex 4

2) $\{a^m b^n : m+n \text{ est un carré}\}$ $\mathcal{L}_2$

Par l'absurde, supposons $\mathcal{L}_2$ reconnaissable

Par propriété de clôture,

$\mathcal{L}_2 \cap \mathcal{L}(a^*) = \{a^{n^2} \mid n \in \mathbb{N}\}$ est reconnaissable

c'est absurde

1) $\{a^p : p \text{ non premier}\} = \mathcal{L}_1$

Par l'absurde, supposons $\mathcal{L}_1$ reconnaissable

Par propriétés de clôture

$\{a^p \mid p \text{ premier}\} = \overline{\mathcal{L}_1} \cap \mathcal{L}(a^*)$

c'est absurde

3) $\{a^n b^m \mid n \neq m\}$ $\mathcal{L}_3$

Par l'absurde, supposons $\mathcal{L}_3$ reconnaissable

Par propriétés de clôture

$$\{a^n b^n \mid n \neq m\} = \mathcal{L}_3 \cap \mathcal{L}(a^* b^*)$$

5) $\{a^m b^n c^{m+n} : m, n \in \mathbb{N}\}$ $\mathcal{L}_5$

Par l'absurde, supposons $\mathcal{L}_5$ reconnaissable

Par propriétés de clôture

$$\{a^n c^m \mid n \in \mathbb{N}\} = \mathcal{L}_5 \cap \mathcal{L}(a^* c^*)$$

ou

$$\{b^n c^M \mid n \in \mathbb{N}\} = \mathcal{L}_5 \cap \mathcal{L}(b^* c^*)$$

Rec = Rat

4) $\{u \in \{a, b, c\}^* : |u|_a = |u|_b\}$

Par l'absurde, supposons $\mathcal{L}_4$ reconnaissable

Par propriété de clôture

$$\{a^n b^n \mid n \in \mathbb{N}\} = \mathcal{L}_4 \cap \mathcal{L}(a^* b^*)$$

$\mathcal{L}_f \; \{ uav : u, v \in \{a,b\}^*, |u| = |v| \}$

$\{ b^n a b^n \mid n \in \mathbb{N} \} = \mathcal{L}_f \cap \mathcal{L}(b^* a b^*)$