

Chapitre 3

Langage

Version avec preuve.

1 Généralités sur les langages

1.1 Un peu de vocabulaire.

**Définition 1 :**

- Un **alphabet** est un ensemble fini ; ces éléments sont des **lettres**.
- Un **mot** sur l'alphabet A est une suite finie de lettres de A : $w = a_1a_2...a_n$.
- La **longueur du mot** précédent, notée $|w|$ est n (nombre de lettres).
- On note $|w|_a$ le nombre d'occurrences de la lettre a dans w . (on a donc $|w| = \sum_{a \in A} |w|_a$)
- On définit un **mot vide** noté généralement ϵ , constitué d'aucune lettre. Sa longueur est nulle.
- L'ensemble des mots de longueur n est noté A^n ; l'ensemble des mots est noté A^* . L'ensemble des mots non vides est noté A^+ .
- On définit la concaténation de deux mots $w_1 = a_1a_2...a_n$ et $w_2 = b_1b_2...b_m$ par $w_1.w_2 = a_1a_2...a_nb_1b_2...b_m$ de sorte que $|w_1.w_2| = |w_1| + |w_2|$.
- a^n désigne la concaténation de n lettres a (si $n = 0$, alors $a^0 = \epsilon$).
- Par convention, $\epsilon.w = w.\epsilon = w$, de sorte que $\cdot$ définit sur A une loi de composition interne associative, non commutative (sauf si A est un singleton), de neutre ϵ . Ainsi, $(A^*, \cdot)$ constitue un monoïde. Une autre convention consiste à confondre une lettre et le mot constitué uniquement de cette lettre.
- Un langage sur A est une partie de A^* .
- Si $w = a_1a_2...a_n$, les **préfixes**(resp. **suffixes**) de w sont les mots de la forme $a_1a_2...a_k$ (resp. $a_ka_2...a_n$), avec $1 \leq k \leq n$. De plus, par convention, ϵ est préfixe et suffixe de tout mot. Un préfixe (resp. suffixe) propre de w est un préfixe (resp. suffixe) distinct de ϵ .
- Un **facteur** de $w = a_1a_2...a_n$ est un mot de la forme $a_i...a_j$, avec $1 \leq i \leq j \leq n$. Un **sous-mot** est un mot de la forme $a_{i_1}...a_{i_k}$, avec $1 \leq i_1 < ... < i_k \leq n$. Ici encore, ϵ est un facteur et un sous-mot de tout mot.
- Le **miroir** de $w = a_1a_2...a_n$ est $w' = a_n...a_1$. Si $w = w'$, on dit que w est un palindrome.

1.2 Propriétés.

Propriété 1 :

■ $\forall u, v, x, y \in A^*$, $uv = xy \Rightarrow \exists t \in A^*$ tel que soit $u = xt$ et $tv = y$, soit $x = ut$ et $v = ty$

Preuve :

Preuve

Supposons par exemple $|u| > |x|$. x est un préfixe de uv donc de u , ce qui justifie l'existence d'un mot t tel que $u = xt$. Dans ce cas, l'égalité $uv = xy$ peut encore s'écrire $xtv = xy$ puis en simplifiant $tv = y$. Le cas $|u| < |x|$ se traite de la même façon.

Théorème 1 :

■ Si $xy = yz$ avec $x \neq \epsilon$, alors $\exists u, v \in A^*$ et un entier $k \geq 0$ tel que :

$$x = uv \quad , \quad y = (uv)^k u = u(vu)^k \quad , \quad z = vu$$

Preuve :

Preuve

- Si $|x| > |y|$:

x	y
y v	z

Il existe v tel que $x = yv$ et $z = vy$ en posant $y = u$ et $k = 0$, ça marche...

- Si $|x| < |y|$: Démonstration par récurrence sur la longueur de y : Pour $y = 2$, on a $y = aa$ et $x = a \dots$ Supposons la relation vraie pour tout y de longueur au moins n . Soit y tel que $|y| = n + 1$, il existe t tel que $y = xt$ donc on a $txt = xtz$ soit $xt = tz$ avec $|t| \leq n$, donc on utilise l'hypothèse de récurrence et on a $x = uv, t = (uv)^k u = u(vu)^k, z = vu$, donc $y = (uv)(uv)^k u = (uv)^{k+1} u$.

Théorème 2 :

■ Si $xy = yx$, avec $x \neq \epsilon$, et $y \neq \epsilon$ alors $\exists u \in A^*$ et deux entiers i, j tels que $x = u^i$, $y = u^j$.

Preuve :

Preuve

Récurrence sur la longueur de xy :

- pour $xy = 2$, évident.
- Supposons la relation vraie pour tout xy de longueur au plus n . Soit xy de longueur $n + 1$, on utilise le théorème précédent, ce qui donne $x = uv = vu$ avec uv de longueur inférieur à n (sinon c'est trivial) donc $\exists t \in A^*$ et deux entiers i, j tels que $u = t^i, v = t^j$, donc $x = t^{i+j}$ et $y = t^{i+k(i+j)}$. La propriété est donc vraie.

2 Opérations sur les langages .

Les langages sur A étant des parties de A^* , on dispose naturellement des notions de réunion, d'intersection, de différence et du complémentaire de langages.

$$L_1 \cup L_2 = \{w \in A^* | w \in L_1 \text{ ou } w \in L_2\}$$

$$L_1 \cap L_2 = \{w \in A^* | w \in L_1 \text{ et } w \in L_2\}$$

$$\overline{L_1} = \{w \in A^* | w \notin L_1\}$$

On peut de plus définir la concaténation (ou le produit) de deux langages L_1 et L_2 par :

$$L_1 L_2 = \{w_1 w_2 \in A^* | w_1 \in L_1 \text{ et } w_2 \in L_2\}$$

On définit alors de façon naturelle L^n comme l'ensemble des mots constitués de n mots de L concaténés :

$$L^0 = \{\epsilon\}$$

$$\forall n \in \mathbb{N}, \quad L^{n+1} = L^n . L$$

Définition 2 :

La fermeture de Kleene (ou **itération** ou encore **étoile**) d'un langage L est la réunion des L_n , pour $n \in \mathbb{N}$. Elle est notée L^* :

$$L^* = \bigcup_{n \in \mathbb{N}} L^n$$

elle contient toujours le mot vide.

L^+ désigne la réunion des L^n pour $n \geq 1$:

$$L^+ = \bigcup_{n \in \mathbb{N}^*} L^n$$

Ce langage contient le mot vide si et seulement si $\epsilon \in L$. On a $L^+ = L.L^*$.

Exercice 1 : Lemme d'Arden.

Soient A et B des langages (avec $\epsilon \notin A$) sur l'alphabet Σ .

Montrer que l'équation de langage suivante :

$$L = A.L \cup B$$

admet pour unique solution le langage $L = A^*B$.

Preuve

On vérifie que $L = A^*B$ est solution de l'équation $L = A.L \cup B$ en démontrant l'égalité.

Pour cela on montre la double inclusion entre les langages .

Preuve de l'unicité de la solution si $\epsilon \notin A$ (sinon Σ^* est aussi une solution...)

Soient L et L' deux solutions distinctes.

Soit w le plus petit mot tel que $w \in L$ et $w \notin L'$. (Si $L \subset L'$, on change les notations...)

Étant donné que $L = A.L \cup B$, alors $B \subset L$, de même $B \subset L'$ donc $w \notin B$. Ce qui impose que $w \in A.L$ et alors $w = au$, avec $a \in A$, et $u \in L$. On a donc construit un mot u plus petit que w .

3 Langages et expressions rationnelles .

3.1 Langage rationnel :

Définition 3 :

Soit l'alphabet A . Les langages rationnels sur A sont définis inductivement par :

- $\{\epsilon\}$ et $\emptyset$ sont des langages rationnels.
- $\forall a \in A, \{a\}$ est un langage rationnel.
- Si L_1 et L_2 sont des langages rationnels, alors $L_1 \cup L_2$ et $L_1.L_2$ également.
- Si L est rationnel, alors L^* l'est également.

Remarques :

- Tous les langages finis sont rationnels, puisqu'ils se déduisent des singletons par un nombre fini d'application des opérations d'union et de concaténation.
- Par définition, l'ensemble des langages rationnels est clos pour les trois opérations rationnelles. (on dit qu'il est rationnellement clos.)
- La famille des langages rationnels correspond précisément au plus petit ensemble de langages qui contient tous les langages finis, et est rationnellement clos.

3.2 Expression rationnelle (ou régulière) :

Définition 4 :

L'ensemble E des expressions rationnelles sur un alphabet A est le plus petit ensemble tel que :

- E contient les éléments de A , ainsi que deux éléments distincts des précédents ϵ et $\emptyset$.
- Si e_1 et e_2 sont deux éléments de E , alors $e_1 + e_2$ et $e_1.e_2$ également.
- Si e est un élément de E , alors e^* est un élément de E .

Remarques :

- On note parfois $e_1|e_2$ à la place de $e_1 + e_2$.
- Cette définition permet les preuves par induction : pour montrer une propriété $P(e)$ pour toute expression e , il suffit d'établir $P(\emptyset)$, $P(\epsilon)$, $P(a)$ pour tout $a \in A$, puis de montrer que si $P(e_1)$ et $P(e_2)$ sont vérifiées, alors $P(e_1 + e_2)$, $P(e_1.e_2)$ et $P(e_1^*)$ également.

Définition 5 :

On définit de façon inductive le langage associé à une expression rationnelle e , noté $L(e)$:

- $L(\epsilon) = \{\epsilon\}$; $L(\emptyset) = \emptyset$; $L(a) = \{a\}$, $\forall a \in A$.
- $L(e_1 + e_2) = L(e_1) \cup L(e_2)$.
- $L(e_1 \cdot e_2) = L(e_1) \cdot L(e_2)$.
- $L(e^*) = L(e)^*$

Définition 6 :

Un langage est dit rationnel si c'est le langage associé à une expression rationnelle. On note $\text{Rat}(A)$ l'ensemble des langages rationnels : $\text{Rat}(A) \subset \mathcal{P}(A^*)$.

3.3 Équivalence et réduction :**Définition 7 :**

Deux expressions rationnelles sont équivalentes si elles dénotent le même langage.

Exemple d'expressions équivalentes

$\emptyset a = a \emptyset = \emptyset$	$\epsilon a = a \epsilon = a$
$\emptyset^* = \epsilon$	$\epsilon^* = \epsilon$
$a + b = b + a$	$a + \emptyset = a$
$a + a = a$	$a^* = (a^*)^*$
$c(a + b) = ca + cb$	$(a + b)c = ac + bc$
$(ab)^* a = a(ba)^*$	
$(a + b)^* = a^*(a + b)^*$	$(a + b)^* = (a^* + b)^*$
$(a + b)^* = (a^* b^*)^*$	$(a + b)^* = (a^* b)^* a^*$

Exercice 2 : Soit $A = \{a, b\}$.

1. Montrer que le langage formé des mots comportant le facteur ab est rationnel.

Preuve

$$L = (a + b)^* ab (a + b)^*$$

2. Montrer que le langage formé des mots ne comportant pas le facteur ab est rationnel.

Preuve

$$L = b^* a^*$$

3. Montrer que le langage formé des mots comportant le facteur ab mais pas le facteur aa est rationnel.

Preuve

$L = b^*ab(b + ab)^*$ (le premier b^*ab est présent pour être sûr d'avoir le facteur ab .)

Exercice 3 : Soit A un alphabet, L un langage et $u \in A^*$. On définit le quotient (ou résiduel) gauche de L par u par :

$$u^{-1}L = \{v \in A^* | uv \in L\}$$

Soit $a \in A$ une lettre et soient K et L deux langages.

1. Montrer que :

$$a^{-1}(K \cup L) = a^{-1}K \cup a^{-1}L$$

Preuve

$$\begin{aligned} m \in a^{-1}(K \cup L) &\Leftrightarrow am \in K \cup L \\ &\Leftrightarrow am \in K \vee am \in L \\ &\Leftrightarrow m \in a^{-1}K \vee m \in a^{-1}L \\ &\Leftrightarrow m \in a^{-1}K \cup a^{-1}L \end{aligned}$$

2. Montrer que :

$$a^{-1}(KL) = \begin{cases} (a^{-1}K)L & \text{si } \epsilon \notin K \\ (a^{-1}K)L \cup a^{-1}L & \text{si } \epsilon \in K \end{cases}$$

Preuve

- Si $\epsilon \notin K$:
 $m \in a^{-1}(KL) \Leftrightarrow am \in KL$
 $\Leftrightarrow \exists u \in K, \exists v \in L, am = uv$
 $\Leftrightarrow \exists u \in K, \exists v \in L, \exists t \in A^* am = uv \wedge at = u \wedge tv = m$
 $\Leftrightarrow \exists v \in L, \exists t \in a^{-1}K, am = atv$
 $\Leftrightarrow \exists v \in L, \exists t \in a^{-1}K, m = tv$
 $\Leftrightarrow m \in (a^{-1}K)L$
- Si $\epsilon \in K$:
 $m \in a^{-1}(KL) \Leftrightarrow am \in KL$
 $\Leftrightarrow \exists u \in K, \exists v \in L, am = uv$
 $\Leftrightarrow (\exists u \in K \setminus \{\epsilon\}, \exists v \in L, am = uv) \vee (\exists v \in L, am = v)$
 $\Leftrightarrow (\exists u \in K \setminus \{\epsilon\}, \exists v \in L, \exists t \in A^* am = uv \wedge at = u \wedge tv = m) \vee (\exists v \in L, am = v)$
 $\Leftrightarrow (\exists v \in L, \exists t \in a^{-1}K, am = atv) \vee (m \in a^{-1}L)$
 $\Leftrightarrow (\exists v \in L, \exists t \in a^{-1}K, m = tv) \vee (m \in a^{-1}L)$
 $\Leftrightarrow m \in (a^{-1}K)L \vee (m \in a^{-1}L)$
 $\Leftrightarrow m \in (a^{-1}K)L \cup a^{-1}L$

3. Montrer que :

$$a^{-1}L^* = (a^{-1}L)L^*$$

Preuve

$$\begin{aligned}
m \in a^{-1}L^* &\Leftrightarrow am \in L^* \\
&\Leftrightarrow \exists u \in L \setminus \{\epsilon\}, \exists v \in L^*, am = uv \\
&\Leftrightarrow \exists u \in L \setminus \{\epsilon\}, \exists t \in A^*, \exists v \in L^*, am = uv \wedge at = u \\
&\Leftrightarrow \exists t \in a^{-1}L, \exists v \in L^*, am = atv \\
&\Leftrightarrow \exists t \in a^{-1}L, \exists v \in L^*, m = tv \\
&\Leftrightarrow m \in (a^{-1}L)L^*
\end{aligned}$$

4. En déduire que si L est un langage rationnel alors $a^{-1}L$ est aussi un langage rationnel.

Preuve

La démonstration se fait par induction structurelle sur les langages rationnels :

- Si $L = \{\epsilon\}$, ou si $L = \emptyset$, alors on a clairement $a^{-1}L = \emptyset$. Donc il est rationnel.
- Si $L = \{a\}$, alors $a^{-1}L = \{\epsilon\}$. Donc il est rationnel.
- Supposons L_1 et L_2 deux langages rationnels tel que $a^{-1}L_1$ et $a^{-1}L_2$ soit rationnels. On a $a^{-1}(K \cup L) = a^{-1}K \cup a^{-1}L$, donc $a^{-1}(K \cup L)$ est l'union de deux rationnels, il est donc rationnel.
- On fait de même pour $a^{-1}(KL)$ et pour $a^{-1}L^*$.
Conclusion : Par induction structurelle sur l'ensemble des langage rationnel, $a^{-1}L$ est bien un langage rationnel.

5. Montrer que si L est un langage rationnel et si $u \in A^*$ alors $u^{-1}L$ est aussi un langage rationnel.

Preuve

Soit L un langage rationnel. Prouvons la propriété pour tout $u \in A^*$.

On procède par récurrence sur $|u|$:

- Si $u = \epsilon$, on a $u^{-1}L = L$.
- Supposons que pour tout mot de longueur $n - 1$, la propriété soit vraie. Soit u un mot de longueur n . On a donc l'existence de $a \in A$, et $v \in A^*$, tel que $u = va$. Montrons tout d'abord que $u^{-1}L = a^{-1}(v^{-1}L)$:

$$\begin{aligned}
m \in u^{-1}L &\Leftrightarrow um \in L \\
&\Leftrightarrow v am \in L \\
&\Leftrightarrow am \in v^{-1}L \\
&\Leftrightarrow m \in a^{-1}(v^{-1}L)
\end{aligned}$$

Exercice 4: On considère l'alphabet $\Sigma = a, b$. On note L_1 le langage des mots ayant un nombre pair de b et L_2 le langage des mots ayant un nombre impair de b .

1. Justifier que l'on a les relations :

$$\begin{aligned}
L_1 &= aL_1 + bL_2 + \epsilon \\
L_2 &= aL_2 + bL_1
\end{aligned}$$

Preuve

$$\begin{aligned}
m \in L_1 &\Leftrightarrow m = \epsilon \vee (m = a.x \wedge x \in L_1) \vee (m = b.x \wedge x \in L_2) \\
&\Leftrightarrow m \in aL_1 + bL_2 + \epsilon
\end{aligned}$$

De même pour l'autre relation.

2. Résoudre le système à l'aide du lemme d'Arden.

Preuve

Puisque $\epsilon \notin \{a\}$ on en déduit d'après le lemme d'Arden que $L_2 = a^*bL_1$ donc la première équation s'écrit aussi $L_1 = aL_1 + ba^*bL_1 + \epsilon = (a + ba^*b)L_1 + \epsilon$ et le lemme d'Arden donne cette fois $L_1 = (a + ba^*b)^*$. On a donc aussi $L_2 = a^*b(a + ba^*b)^*$.

4 Langages locaux .

Définition 8 :

On considère un langage L sur un alphabet A . On définit :

- $P(L) = \{a \in A \mid a.A^* \cap L \neq \emptyset\} = \{a \in A \mid \exists u \in A^*, a.u \in L\}$
 $P(L)$ correspond à l'ensemble des premières lettres des mots de L .
- $S(L) = \{a \in A \mid A^*.a \cap L \neq \emptyset\} = \{a \in A \mid \exists u \in A^*, u.a \in L\}$
 $S(L)$ correspond à l'ensemble des dernières lettres des mots de L .
- $F(L) = \{(a, b) \in A^2 \mid A^*.ab.A^* \cap L \neq \emptyset\} = \{(a, b) \in A^2 \mid \exists (u, v) \in (A^*)^2, u.a.b.v \in L\}$
 $F(L)$ correspond à l'ensemble des facteurs de 2 lettres des mots de L .
- $N(L) = \{(a, b) \in A^2 \mid A^*.ab.A^* \cap L = \emptyset\} = \{(a, b) \in A^2 \mid \exists (u, v) \in (A^*)^2, u.a.b.v \notin L\}$
 $N(L)$ correspond à l'ensemble des mots de 2 lettres non facteurs des mots de L .

Exemple 1 :

Pour $L = L(a^*bb^*)$, sur l'alphabet $A = \{a, b\}$, on a :

- | | |
|---------------------|---------------------------|
| • $P(L) = \{a, b\}$ | • $F(L) = \{aa, ab, bb\}$ |
| • $S(L) = \{b\}$ | • $N(L) = \{ba\}$ |

Définition 9 :

Un langage L sur un alphabet A est local s'il existe deux parties P et S de A et une partie N de A^2 telles que w est un mot de L si et seulement si :

- la première lettre de w est dans P ;
- la dernière lettre de w est dans S ;
- Aucun facteur de longueur 2 de w n'est dans N .

On a alors :

$$L \setminus \{\epsilon\} = (PA^* \cap A^*S) \setminus (A^*NA^*)$$

Exemple 2 :

Sur l'alphabet $A = \{a, b\}$, les langages suivants sont locaux :

- $L = L(ab)$ car $ab = aA^* \cap A^*b \setminus A^*\{aa, bb, ba\}A^*$.
- $L = L((ab)^*)$ car $(ab)^+ = aA^* \cap A^*b \setminus A^*\{aa, bb\}A^*$.
- $L = L((a)^*)$ car $a^+ = aA^* \cap A^*a \setminus A^*\{ab, ba, bb\}A^*$.

Exercice 5 : Montrer que, sur l'alphabet $A = \{a, b\}$, les langages suivants :

- $L_1 = L((a)^*) \cup L((ab)^*) = L(a^* + (ab)^*)$.
- $L_2 = L((a)^*).L((ab)^*) = L((a)^*. (ab)^*)$.

ne sont pas locaux.

Preuve

On montre dans les deux cas que aba est dans le langage $(PA^* \cap A^*S) \setminus (A^*NA^*)$, mais pas dans L .

Théorème 3 :

I L'intersection de deux langages locaux est un langage local.

Preuve : **Preuve**

On pose $L_1 \setminus \{\epsilon\} = (P_1.A^* \cap A^*.S_1) \setminus (A^*N_1A^*)$ et $L_2 \setminus \{\epsilon\} = (P_2.A^* \cap A^*.S_2) \setminus (A^*N_2A^*)$. On démontre que, pour $P = P_1 \cap P_2$, $S = S_1 \cap S_2$ et $N = N_1 \cup N_2$, on a donc :

$$\begin{aligned} (P_1 \cap P_2) \setminus \{\epsilon\} &= ((P_1.A^* \cap A^*.S_1) \setminus (A^*N_1A^*)) \cap ((P_2.A^* \cap A^*.S_2) \setminus (A^*N_2A^*)) \\ &= (PA^* \cap A^*S) \setminus (A^*NA^*) \end{aligned}$$

Théorème 4 :

I Si L_1 et L_2 sont deux langages locaux définis sur des alphabets disjoints, alors $L_1 \cup L_2$ est encore un langage local.

Preuve : **Preuve**

Notons A_1 et A_2 les deux alphabets, avec $L_1 \subset A_1^*$ et $L_2 \subset A_2^*$. On pose $P(L_1 \cup L_2) = P(L_1) \cup P(L_2)$, $S(L_1 \cup L_2) = S(L_1) \cup S(L_2)$ et $F(L_1 \cup L_2) = F(L_1) \cup F(L_2)$.

On pose $L = (P(L_1 \cup L_2)A^* \cap A^*S(L_1 \cup L_2) \setminus (A^*N(L_1 \cup L_2)A^*))$. On a déjà l'inclusion : $L_1 \cup L_2 \subset L$.

Soit $u \in L$, notons $u = a_1a_2...a_n$.

- $a_1 \in P(L_1) \cup P(L_2)$. Supposons que $a_1 \in P(L_1)$, donc $a_1 \in A_1$.
- $a_1a_2 \in F(L_1) \cup F(L_2)$, avec le jeu des deux alphabets, on a forcément $a_1a_2 \in F(L_1)$, donc $a_2 \in A_1$.
- En poursuivant de proche en proche, on prouve que toutes les lettres sont dans A_1 .

Or L_1 est un langage local, donc $u \in L_1$, soit $u \in L_1 \cup L_2$

Théorème 5 :

I Si L_1 et L_2 sont deux langages locaux définis sur des alphabets disjoints, alors $L_1.L_2$ est encore un langage local.

Preuve :**Preuve**

Posons :

$$P(L_1L_2) = \begin{cases} P(L_1) & \text{si } \epsilon \notin L_1 \\ P(L_1) \cup P(L_2) & \text{sinon} \end{cases} \quad S(L_1L_2) = \begin{cases} S(L_2) & \text{si } \epsilon \notin L_2 \\ S(L_1) \cup S(L_2) & \text{sinon} \end{cases}$$

$$F(L_1L_2) = F(L_1) \cup F(L_2) \cup S(L_1)P(L_2)$$

Notons $L = (P(L_1L_2)A^* \cup A^*S(L_1L_2)) \setminus (A^*N(L_1L_2)A^*)$. On a donc, par construction, $L_1L_2 \subset L$.

Soit $u \in L$, notons $u = a_1a_2\dots a_n$.

- Si $a_1 \in A_2$ alors $\epsilon \in L_1$ et $a_2 \in A_2$. De proche en proche on prouve que $a_i a_{i+1} \in F(L_2)$. Puisque L_2 est un langage local on en déduit que $u \in L_2$, et puisque $\epsilon \in L_1$, on a aussi $u \in L_1L_2$.
- Si $a_1 \in A_1$, notons $a_1\dots a_k$ le plus long préfixe de u qui soit dans A_1^* . Alors $a_1 \in P(L_1)$ et $\forall i \in [1; k-1] a_i a_{i+1} \in F(L_1)$. Ensuite, de deux choses l'une :
 - Si $k = n$, $a_n \in S(L_1)$, et $\epsilon \in L_2\dots$
 - Si $k < n$, $a_k a_{k+1} \in S(L_1)P(L_2)$, donc $a_{k+1} \in A_2$. On construit encore de proche en proche...

**Théorème 6 :**

La fermeture de Kleene L^* d'un langage local L est aussi un langage local.

Preuve :**Preuve**

On a $P(L^*) = P(L)$, $S(L^*) = S(L)$ et $F(L^*) = F(L) \cup S(L)P(L)$. En effet, considérons un mot $u \in L^* \setminus \{\epsilon\}$. Alors $u = u_1u_2\dots u_p$ avec $u_i \in L$.

La première lettre de u est aussi celle de u_1 donc est dans $P(L)$

La dernière lettre de u est aussi celle de u_p donc est dans $S(L)$

Un facteur de longueur 2 de u est :

- un facteur d'un des u_i auquel cas il est dans $F(L)$;
- ou de la forme xy où x est la dernière lettre de u_i et y la première lettre de u_{i+1} , auquel cas il est dans $S(L)P(L)$;

Posons donc $P = P(L)$, $S = S(L)$ et $N = A^* \setminus (F(L) + S(L)P(L))$. Pour montrer que L^* est local nous devons considérer un mot u de $(PA^* \cap A^*S) \setminus (A^*N(L)A^*)$ et prouver que $u \in L^*$. Parmi les facteurs de longueur 2 de u , considérons uniquement ceux qui appartiennent à $S(L)P(L)$; ils induisent une factorisation de u en mots : $u = u_1u_2\dots u_p$ avec $u_i \in (PA^* \cap A^*S) \setminus (A^*N(L)A^*)$. Or par hypothèse L est local, donc chacun des u_i est un mot de L . On en déduit que $u \in L^*$.

5 Expressions rationnelles linéaires.

**Définition 10 :**

Une expression rationnelle e est dite linéaire lorsque tout caractère de A apparaît au plus une fois dans e :

$$\forall a \in A \quad |e|_a \leq 1$$

**Exemple 3 :**

1. $e_1 = ab$ est une expression rationnelle linéaire.
2. $e_2 = (ab)^*$ est une expression rationnelle linéaire.
3. $e_3 = (ab)^*a$ n'est pas une expression rationnelle linéaire.

**Théorème 7 :**

■ Toute expression rationnelle linéaire dénote un langage local.

Preuve :

Preuve

Raisonnons par induction structurelle sur l'expression rationnelle linéaire e .

- $\emptyset$; et ϵ dénotent respectivement les langages $\emptyset$; et $\{\epsilon\}$ qui sont des langages locaux.
- si $a \in A$, a dénote le langage local $\{a\}$ avec $P = \{a\}$, $S = \{a\}$ et $N = A^2$.
- si $e = e_1 + e_2$ est une expression rationnelle linéaire, il en est de même de e_1 et e_2 , et par hypothèse d'induction ces deux expressions dénotent des langages locaux. Mais e_1 et e_2 peuvent être définis sur des alphabets disjoints, donc e dénote aussi un langage local.
- si $e = e_1 e_2$ ou $e = e^*$ le raisonnement est identique.