

Nous allons dans ce chapitre introduire le système mathématique qui nous servira comme modèle abstrait des langages.

1. Mots

Un mot se définit sur un alphabet qui est constitué d'éléments atomiques. Dans ce qui suit nous définissons ce qu'est un alphabet et un mot.

Définition 1. Un *alphabet* est un ensemble fini non-vide de lettres ou symboles que l'on notera dans ce cours X .

Exemples d'alphabet:

$X_1 = \{a, b, \dots, z\}$ Alphabet latin.
 $X_2 = \{0, 1\}$ Alphabet Binaire
 $X_3 = \{a, b, c\}$.

Définition 2. Soit X un alphabet, on appelle mot sur X une suite finie d'éléments (lettres) de X .

Exemples de mots: 0, 01, 110 sont des mots sur $X = \{0, 1\}$ un alphabet

Définition 3. Soit X un alphabet, on appelle mot vide un mot particulier qui ne contient aucune lettre de X et que l'on note ε . Ce mot ne contient aucune lettre.

Définition 4. La longueur d'un mot w sur un alphabet X , notée $|w|$, est le nombre de lettres de X composant le mot w .

Exemple:

Soit $X = \{a, b\}$, prenons un mot sur X $w = abab$. On a $|w| = 4$
 Pour $w = \varepsilon$ alors $|w| = 0$

Définition 5. L'occurrence d'une lettre $x_i \in X$ dans un mot w est le nombre d'apparition de la lettre x_i dans w que l'on note $|w|_{x_i}$.

Exemple:

Soit $X = \{a, b\}$ prenons un mot sur X $w = abaabaaa$. On a $|w|_a = 6$. et $|w|_b = 2$

1.1. Concaténation des mots

Soit X un alphabet. On définit, la concaténation notée \cdot , une opération sur les mots de X que l'on définit comme suit:

$w = w_1 w_2 \dots w_n$ $|w| = n$ avec $w_i \in X$
 $v = v_1 v_2 \dots v_m$ $|v| = m$ et $v_j \in X$
 on a $w \cdot v = w_1 w_2 \dots w_n v_1 v_2 \dots v_m$

Montrons que \cdot est une loi de composition interne.

$$w \cdot v = w_1 w_2 \dots w_n \cdot v_1 v_2 \dots v_m = u_1 u_2 \dots u_{n+m} \quad \text{avec } u_i = w_i \quad \text{si } 1 \leq i \leq n \\ \text{et} \quad u_i = v_{i-n} \quad \text{si } n+1 \leq i \leq n+m$$

On a $w \cdot \varepsilon = \varepsilon \cdot w = w$

Définition 6. (Définition récursive d'un mot) Soit X un alphabet, w est un mot sur X si et seulement si :

- $w = \varepsilon$
- ou $w = x_i \cdot u$ avec $x_i \in X$ et u un mot sur X .

1.2. Ensemble de tous les mots (X^*)

X^* est exactement l'ensemble de tous les mots construits sur un alphabet. Ces mots sont de longueur 0, 1, 2, X^* est le monoïde libre engendré par X . $\forall w \in X^*$, w s'écrit en fonction des éléments de X .

Nous allons montrer à travers l'exemple suivant comment construire l'ensemble X^* :

Exemple : Soit $X = \{a, b\}$

$X^0 = \{\varepsilon\}$ Mots de longueur 0

$X^1 = \{a, b\}$ Mots de longueur 1

$X^2 = \{aa, ab, ba, bb\} = X \cdot X$ Mots de longueur 2

$X^3 = \{aaa, aab, aba, abb, baa, bab, bba, bbb\} = X^2 \cdot X$ Mots de longueur 3.

...

$X^i = X^{i-1} \cdot X$ Mots de longueur i où \cdot est défini comme étant la concaténation de deux ensembles : $X^{i-1} \cdot X = \{w \cdot x_k / w \in X^{i-1} \text{ et } x_k \in X\}$

...

$X^* = X^0 \cup X^1 \cup \dots \cup X^n \cup \dots$

$X^+ = X^1 \cup X^2 \cup \dots \cup X^n \cup \dots$

$X^n \cdot \{\varepsilon\} = \{w \cdot \varepsilon / w \in X^n \text{ et } \varepsilon \text{ est le mot vide}\}$

$X^n \cdot \Phi = \Phi$ où Φ représente l'ensemble vide.

Exercice 1.

Montrer que X^i est exactement l'ensemble des mots de longueur i construit sur l'alphabet X . (Démonstration par récurrence sur la longueur des mots)

Exercice 2.

Vérifier la validité de la proposition suivante $X^{i-1} \cdot X = X \cdot X^{i-1}$.

Théorème 1.

$(X^*, \cdot, \varepsilon)$ a la structure d'un monoïde.

Exercice 3.

Démontrer le théorème 1.

Théorème 2.

La fonction longueur est un morphisme de monoïde.

Démonstration:

On définit la fonction longueur f tq:

$$f: (X^*, \cdot, \varepsilon) \rightarrow (\mathbb{N}^+, +, 0)$$

Cette fonction associe à chaque mot de X^* sa longueur qui appartient à \mathbb{N} . On note $f ||$.

| Soient x et y deux mots quelconques appartenant à X^* tel que :

$$x = x_1 x_2 \dots x_n \quad |x| = n$$

$$y = y_1 y_2 \dots y_m \quad |y| = m$$

montrons que $L(x \cdot y) = L(x) + L(y)$

$$1. \text{ On a } x \cdot y = x_1 x_2 \dots x_n \cdot y_1 y_2 \dots y_m y_1 y_2 \dots y_m$$

$$|x \cdot y| = |x_1 x_2 \dots x_n \cdot y_1 y_2 \dots y_m y_1 y_2 \dots y_m| = n+m = |x| + |y|$$

$$2. L(\varepsilon) = 0$$

1.3. Mot Miroir

Définition 7. Soit X an alphabet et soit w un mot de X^* , on appelle mot miroir de w , noté w^R , le mot obtenu en inversant les lettres de w .

Soit $w \in X^*$ tq $w = w_1 w_2 \dots w_n$ on a $w^R = w_n w_{n-1} \dots w_1$, $|w| = |w^R|$

Exemple:

$$X = \{a, b\} \quad w = \text{abbab} \quad w^R = \text{babba}$$

Définition 8. On appelle palindrome tout mot w de X^* tel que $w = w^R$

Exemple:

$$X = \{a, b\} \quad w = \text{ababa} \quad w^R = \text{ababa}$$

1.4. Relations sur les mots

Définition 9. Soit X an alphabet et X^* son monoïde libre engendre, et soient u et $v \in X^*$:

- u est dit ***facteur gauche*** de v s'il existe un mot $h \in X^*$ tel que $v = u \cdot h$
Exemple : $X = \{a, b\}$ $v = \text{abaa}$. Les facteurs gauches de v sont : $\varepsilon, a, ab, aba, abaa$.

- u est dit ***facteur propre gauche*** de v s'il existe un mot $h \in X^+$ tel que $v = u \cdot h$
Exemple : $X = \{a, b\}$ $v = \text{abaa}$. Les facteurs propres gauches de v sont : ε, a, ab, aba .

- u est dit **facteur droit** de v s'il existe un mot $h \in X^*$ tel que $v = h \cdot u$
Exemple : $X = \{a, b\}$ $v = abaa$. Les facteurs droits de v sont : $\varepsilon, a, aa, baa, abaa$.
- u est dit **facteur gauche** de v s'il existe un mot $h \in X^*$ tel que $v = h \cdot u$
Exemple : $X = \{a, b\}$ $v = abaa$ ε, a, aa, baa , sont des facteurs propres droits de v .

Exercice 3: Montrer que la relation Facteur Gauche est une relation d'ordre partiel. Est-elle une relation d'ordre total?

Lemme de Levi

Ce lemme montre des propriétés remarquables sur les mots: Soit X un alphabet et X^* le monoïde libre engendré par X et soient w, u_1, u_2, v_1 et v_2 5 mots de X^* tels que $w = u_1 \cdot v_1 = u_2 \cdot v_2$, nous avons les trois relations suivantes:

1. Si $|u_1| < |u_2| \Rightarrow u_2 = u_1 \cdot h$ et $v_1 = h \cdot v_2$
2. Si $|u_1| = |u_2| \Rightarrow u_1 = u_2$ et $v_1 = v_2$.
3. Si $|u_1| > |u_2| \Rightarrow u_1 = u_2 \cdot h$ et $v_2 = h \cdot v_1$

Démonstration:

On a par définition que $|w| = |u_1| + |v_1| = |u_2| + |v_2|$. $u_1 \cdot v_1$ et $u_2 \cdot v_2$ sont deux décompositions possible de w . Nous allons établir les relations qui existent entre les facteurs gauches, respectivement les facteurs droits, de w .

1. Si $|u_1| = |u_2| \Rightarrow u_1 = u_2$ et $v_1 = v_2$ 2 facteurs gauches d'un même mot de même longueur.
 $\Rightarrow |v_1| = |v_2| \Rightarrow v_1 = v_2$
2. si $|u_1| < |u_2| \Rightarrow u_1$ est un facteur gauche propre de $u_2 \Rightarrow u_2 = u_1 \cdot h$
 On remplace u_2 par sa décomposition dans $u_1 \cdot v_1 = u_2 \cdot v_2$. On obtient :
 $u_1 \cdot v_1 = u_1 \cdot h \cdot v_2 \Rightarrow v_1 = h \cdot v_2$
3. Ce cas est laissé en exercice.

2. Langages

Définition 10. Soit X un alphabet. On appelle langage sur X toute partie L de X^* .

Exemple: $L = X^*$, $L = \{\varepsilon\}$, $L = \{w \in \{a, b\}^* \text{ tq } w = w_1 a w_2, w_1 \text{ et } w_2 \in \{a, b\}^*\}$

Définition 11. Un langage est fini s'il contient un nombre fini de mots.

Exemple: $X = \{a, b, c\}$ $L_f = \{w \in X^* / |w| \leq 3\}$

Définition 12. Un langage est infini s'il contient un nombre infini de mots.

Exemple: $X = \{a, b, c\}$ $L_i = \{w \in X^* / |w| \geq 3\}$

Définition 13. Un langage est vide s'il ne contient aucun mot $L = \Phi$

2.1. Opérations sur les Langages

Les langages sont des ensembles de X^* , on peut donc exécuter sur les langages toutes les opérations définies sur les ensembles. Soient L_1, L_2 deux langages:

- Union :

$$L_1 \cup L_2 = \{w \in X^* / w \in L_1 \text{ ou } w \in L_2\}$$
- Intersection:

$$L_1 \cap L_2 = \{w \in X^* / w \in L_1 \text{ et } w \in L_2\}$$
- Différence:

$$L_1 - L_2 = \{w \in X^* / w \in L_1 \text{ et } w \notin L_2\}$$
- Complément:

$$\overline{L_1} = \{w \in X^* / w \notin L_1\}$$
- Concaténation:

$$L_1 \cdot L_2 = \{w_1 \cdot w_2 \in X^* / w_1 \in L_1 \text{ et } w_2 \in L_2\}$$
- Image Miroir:

$$L_1^R = \{w \in X^* / w^R \in L_1\}$$
- Puissance d'un langage:
 Elle est définie par récurrence sur la longueur des mots de la manière suivante:

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

$$L^1 = L \cdot \{\epsilon\}$$

$$L^2 = L \cdot L$$

$$\vdots$$

$$\vdots$$

$$\vdots$$

$$L^{n+1} = L^n \cdot L = L \cdot L^n$$
- Itération:

$$L^* = L^0 \cup L^1 \cup \dots \cup L^n \cup \dots$$

$$L^* = \cup L^i \quad i \geq 0$$
- Itération Positive: $L^+ = L^1 \cup L^2 \cup \dots \cup L^n \cup \dots$

$$L^+ = \cup L^i \quad i > 0 \quad (\epsilon \notin L^+)$$

Résidu par rapport à un mot:

Soit L un langage de X ($L \subseteq X^*$) et soit u un mot de X^*

1. Résidu Gauche

$$L u^{-1} = \{v \in X^* / v \cdot u \in L\}$$

2. Résidu Droit

$$u^{-1} L = L // u = \{v \in X^* / u \cdot v \in L\}$$

Propositions:

$$(L_1 \cup L_2) // u = (L_1 // u) \cup (L_2 // u)$$

$$(L_1 \cap L_2) // u = (L_1 // u) \cap (L_2 // u)$$

$$(L_1 - L_2) // u = (L_1 // u) - (L_2 // u)$$

$$L_1 // (u \cdot v) = (L_1 // u) // v$$

$$(L_1 \cdot L_2) // u_i = (L_1 // u_i) \cdot L_2 \cup (L_2 // u_i) \text{ si } \varepsilon \in L_1 \quad u_i \in X$$

$$= (L_1 // u_i) \cdot L_2 \text{ si } \varepsilon \notin L_1$$

$$L^* // u_i = L^* // u_i \cdot L^*$$

Résidu par rapport a un langage

$$L K^{-1} = \{ v \in X^* / \exists u \in K \text{ tq } u \cdot v \in L \}$$

Série d'exercices

1.1. Vérifier la validité des propositions suivantes :

- $(L^*)^* = L^*$
- $(L_1 \cup L_2)^* = L_1^* \cup L_2^*$
- $(L_1 \cdot L_2)^* = L_1^* \cdot L_2^*$
- $(L_1 \cup L_2)^* = L_1^* (L_1 \cup L_2)^*$
- $(L_1^* \cdot L_2^*)^* = (L_1 \cup L_2)^*$
- $(L_1 \cup L_2)^* = (L_1^* L_2^*)^*$
- $L(L_1 \cap L_2) = L \cdot L_1 \cap L \cdot L_2$

1.2. Montrer que la fonction longueur d'un mot est un morphisme de monoïde.

1.3. Montrer que la relation Facteur gauche est une relation d'ordre partiel.

1.4. Montrer que si $u^2 v^2 = w^2$ alors $uv = vu$, avec u, v et $w \in X^*$.

1.5. Soit l'équation $uw^R = wv$, u, v et $w \in X^*$, existe-il une relation entre u et v ?

1.6. Comparer les langages suivants sur $X = \{a, b\}$:

- $L_1 = \{w_1 ab w_2, w_1, w_2 \in X^*\}$
- $L_2 = \{b^i a^j b w, i \geq 0, j > 0, w \in X^*\}$
- $L_3 = \{b^i a^j, i, j \geq 0\}$

1.7. Comparer les deux langages suivants sur $X = \{a, b\}$:

- $L_1 = \{a^i b^j, i, j \geq 0\}$ et $L_2 = \{w_1 b a^i b w_2, i > 0, w_1 \text{ et } w_2 \in X^*\}$
- Comparer les langages dénotés par les deux expressions régulières.

1.8. Comparer les langages suivants sur $X = \{a, b\}$:

- $L_1 = \{a^i b^j (ab)^k a a w t q i, k \geq 0, j > 0, w \in X^*\}$
- $L_2 = \{w a (ab)^i a b^j a^k, i, k \geq 0, j > 0\}$

1.9. Comparer les trois langages suivants :

- $L1 = \{ww^R / w \in \{0, 1\}^*\}$
- $L2 = \{(01)^i (10)^j (01)^k (10)^m / i, j, k, m \geq 0\}$
- $L3 = \{(01)^i (10)^j (01)^j (10)^i / i, j \geq 0\}$

1.10. Soit $X = \{x, y\}$, Montrer que $xy = yx$ ssi $\exists z \in X^*$ tel que $x \in z^*$ et $y \in z^*$

1.11. Soit $X = \{a, b\}$ et $w \in X^*$, Montrer que si $wa = bw$ alors $a = b$ et $w \in \{a\}^*$

1.12. Soit X un alphabet, on dit que $w \in X^*$ est un palindrome

$$\text{si } w = \varepsilon \text{ ou } w = w_1 w_2 \dots w_n = w_n \dots w_2 w_1 \text{ avec } w_i \in X$$

Soit $X = \{a, b\}$, on définit sur X la suite de mots $(f_n)_{n > 0}$ (Suite de Fibonacci) de la façon suivante : $f_1 = a$, $f_2 = ab$ et $f_{n+2} = f_{n+1} f_n$, $n > 0$.

Montrer que pour tout $i \geq 2$,

$$f_i = \begin{cases} uab & \text{si } i \text{ pair} \\ uba & \text{si } i \text{ impair} \end{cases} \quad u \text{ étant un palindrome}$$