Série de TD N° : 01 - Corrigé

Langages Formels

Exercice 1.1 Déterminer les facteurs, les préfixes et les suffixes du mot $u = abac$.

Corrigé.

Facteurs : ε, a, b, c, ab, ba, ac, aba, bac, abac
Préfixes : ε, a, ab, aba, abac
Suffixes : ε, c, ac, bac, abac

Exercice 1.2 Soit $u = a_1 \dots a_n$ un mot de longueur n, avec $a_i \neq a_j$ pour tous $i \neq j$. Combien u comporte-t-il de préfixes ? de suffixes ? de facteurs ?

Corrigé.

Les préfixes de u distincts de ε sont entièrement déterminés par leur dernière lettre, qui peut valoir $u[1], u[2], \dots, u[n]$. Il y en a donc n, à quoi se rajoute le mot vide. Soit $n+1$ préfixes en tout pour un mot u de longueur n.

On montre de manière similaire que le nombre de suffixes de u est $n+1$.

Pour déterminer le nombre de facteurs de u, il suffit de constater que les facteurs non-vides de u sont exactement les préfixes de $u[i..n]$ quand i parcourt $\{1,\dots,n\}$. D'où découle que le nombre de ces facteurs (non vides) vaut :

$$\sum_{i=1}^{n} (\text{nombre de préfixes non-vides de } u[i..n]) = n + (n-1) + \cdots + 1 = \frac{n(n+1)}{2}$$

À quoi se rajoute le mot vide, ce qui donne finalement $1 + \frac{n(n+1)}{2}$ facteurs pour u.

Exercice 1.3

1. Compter les occurrences des lettres a et b dans les mots suivants : a^3cbbca, *aabg jdd*, *titi*, *babc*.
2. Donner l'ensemble des couples (u, v) tels que $uv = abaac$.
3. Calculer LM pour les ensembles suivants :
   - $L = \{a,ab,bb\}$ et $M = \{\varepsilon,b,a^2\}$;
   - $L = \emptyset$ et $M = \{a,ba,bb\}$;
   - $L = \{\varepsilon\}$ et $M = \{a,ba,bb\}$;
   - $L = \{aa,ab,ba\}$ et $M = \{a,b\}^*$.

Corrigé

1.

u	$\mid u \mid_a$	$\mid u \mid_b$
a^3cbbca	4	2
$aabgjdd$	2	1
$titi$	0	0
$babc$	1	2

2.

u	v
ε	$abaac$
a	$baac$
ab	aac
aba	ac
$abaa$	c
$abaac$	ε

3. $\{a, ab, bb\} \cdot \{\varepsilon, b, a^2\} = \{a, ab, bb, ab^2, b^3, a^3, aba^2, bba^2\}$;

- $\varnothing \cdot \{a,ba,bb\} = \varnothing$;
- $\{\varepsilon\} \cdot \{a,ba,bb\} = \{a,ba,bb\}$;
- $\{aa,ab,ba\} \cdot \{a,b\}^*$ est l'ensemble de tous les mots sur l'alphabet $\{a,b\}$ qui débutent par le préfixe *aa*, *ab*, ou *ba*.

Exercice 1.4 On considère l'alphabet $\Sigma = \{a,b\}$, et les langages L_1 et L_2 suivants :

$$L_1 = \{a^n b^n \mid n \in \mathbf{N}\}$$
$$L_2 = \{b^n a^n \mid n \in \mathbf{N}\}$$

Calculez : $L_1 \cup L_2$, $L_1 \cap L_2$, $L_1.L_2$, $L_2.L_1$, L_1^2.

Corrigé

- $L_1 \cup L_2 = \{a^n b^n \mid n \in \mathbf{N}\} \cup \{b^n a^n \mid n \in \mathbf{N}\} = \{ a^n b^n \text{ ou } b^n a^n \mid n \in \mathbf{N} \}$
- $L_1 \cap L_2 = \{a^n b^n \mid n \in \mathbf{N}\} \cap \{b^n a^n \mid n \in \mathbf{N}\} = \{\varepsilon\}$
- $L_1.L_2 = \{ a^n b^n b^m a^m \mid n,m \in \mathbf{N}\} = \{ a^n b^{n+m} a^m \mid n,m \in \mathbf{N} \}$
- $L_2.L_1 = \{ b^n a^n\, a^m b^m \mid n,m \in \mathbf{N}\} = \{ b^n a^{n+m} b^m \mid n,m \in \mathbf{N} \}$
- $L_1^2 = \{ a^n b^n\, a^m b^m \mid n,m \in \mathbf{N} \}$

Exercice 1.5

Prouver les assertions suivantes, où Σ est un alphabet, $a,b \in \Sigma$ et $u, v, x, y \in \Sigma^*$.

1. $au = bv \Rightarrow (a = b \text{ et } u = v)$;
2. $xu = xv \Rightarrow u = v$;
3. $(xu = yv \wedge |x| = |y|) \Rightarrow u = v$;
4. $(xu = yv \wedge |x| \leq |y|) \Rightarrow$ *(x est préfixe de y et v est suffixe de u).*

Corrigé

1. Deux mots sont égaux s'ils sont la même longueur et si leurs lettres coïncident en chaque position. Les mots *au* et *bv* s'écrivent sous la forme $au_1 \ldots u_p$ et $bv_1 \ldots v_n$, avec u_i , $v_i \in \Sigma$ pour tout *i*. Puisque $au = bv$, on peut tirer de ce qui précède : $p = n$, $a = b$ et $u_i = v_i$ pour tout $i \leq n$. D'où en particulier $a = b$ et $u = v$.
2. On établit ce résultat par récurrence sur $|x|$.
   - Si $|x| = 0$, alors $x = \varepsilon$ et la conclusion est immédiate.
   - Sinon, On suppose que pour $|x| = n$, $xu = xv \Rightarrow u = v$. On montre la propriété pour $|x| = n+1$.

     Dans ce cas $x = ay$ avec $a \in \Sigma$, $y \in \Sigma^*$ et $|y| = n$ et l'on a : $ayu = ayv$. D'après le résultat 1 précédant, ceci impose $yu = yv$, et par hypothèse de récurrence : $u = v$.
3. Récurrence sur $|x| = |y|$:
   - Si $|x| = 0$, alors $x = y = \varepsilon$ et la conclusion est immédiate.
   - Sinon, $x = ax'$ et $y = by'$ avec $a,b \in \Sigma$ et $x', y' \in \Sigma^*$ et $|x'| = |y'|$. On a : $ax'u = by'v$. D'après 1, ceci impose $a = b$ et $x'u = y'v$. Comme $|x'| < |x|$, on peut conclure, grâce à l'hypothèse de récurrence, que $u = v$.

Exercice 1.6 Montrer que :

1. Il n'existe pas de mot $x \in \{a,b\}^*$ tel que $ax = xb$.
2. Il n'existe pas de mots $x, y \in \{a,b\}^*$ tel que $xay = ybx$.

Corrigé

1. Rappelons que l'on note $|u|_a$ le nombre d'occurrences de la lettre *a* dans le mot *u*. Clairement, si *u* et *v* sont deux mots, $|uv|_a = |u|_a + |v|_a$. En particulier, si un mot *x* satisfait $ax = xb$, on aurait $|ax|_a = |xb|_a$ et donc $|x|_a + 1 = |x|_a$, contradiction.
2. L'argument est le même que pour la question précédente : si $xay = ybx$, alors $|xay|_a = |ybx|_a$ et donc $|x|_a + |y|_a + 1 = |x|_a + |y|_a$, contradiction.

Exercice 1.7 Soient A,B,C trois langages sur un même alphabet. Prouver les propriétés suivantes :

1. $(AB)C = A(BC)$
2. $A(B \cup C) = AB \cup AC$
3. $A(B \cap C) \subseteq AB \cap AC$ et l'inclusion réciproque est fausse en général.
4. $A \subseteq B \Rightarrow A^* \subseteq B^*$
5. $(A^*)^* = A^*$
6. $(A \cup B)^* = (A^*B^*)^*$

Corrigé

1. Se déduit immédiatement de l'associativité de la concaténation des mots.
2. $u \in A(B \cup C) \Leftrightarrow \exists a \in A, \exists x \in B \cup C : u = ax$

   $\Leftrightarrow (\exists a \in A, \exists x \in B : u = ax)$ *ou* $(\exists a \in A, \exists x \in C : u = ax)$

   $\Leftrightarrow u \in AB$ *ou* $u \in AC$

   $\Leftrightarrow u \in AB \cup AC$
3. Tout mot u de $A(B \cap C)$ s'écrit $u = ax$ avec $a \in A$ et $x \in B \cap C$ et appartient donc à AB (puisque $a \in A$ et $x \in B$) et AC (puisque $a \in A$ et $x \in C$). Par contre, pour $A = \{1,10\}$, $B = \{01\}$ et $C = \{1\}$, on a $101 \in AB \cap AC$, mais clairement $101 \notin A(B \cap C)$ car $B \cap C = \varnothing$
4. Si $A \subseteq B$, alors tout mot de la forme $u_1 \dots u_n$ avec $n \in \mathbf{N}$ et $u_1,\dots,u_n \in A$ est dans B^n, puisque chaque u_i est dans B. Autrement dit, $A^n \subseteq B^n$ pour tout n. D'où s'ensuit que $A^* = \bigcup_{n \in N} A^n \subseteq B^* = \bigcup_{n \in N} B^n$.
5. De $A \subseteq A^*$ on tire, grâce à la question précédente : $A^* \subseteq (A^*)^*$.

   Montrons maintenant que $(A^*)^* \subseteq A^*$. Soit $u \in (A^*)^*$. Donc il existe $n \geq 0$ tel que $u \in (A^*)^n$. Cela veut dire que u peut s'écrire $u = u_0, \dots, u_n$ avec $u_i \in A^*$ (pour tout i entre 0 et n). Puisque chaque u_i est une concaténation de mots de A, il s'en suit que u est lui-même une concaténation de mots de A, c-à-d $u \in A^*$. D'où l'inclusion $(A^*)^* \subseteq A^*$.
6. On montre que : a) $(A \cup B)^* \subseteq (A^*B^*)^*$ et b) $(A^*B^*)^* \subseteq (A \cup B)^*$

   a) On a $A \subseteq A^*$ et $A^* \subseteq A^*B^*$, puisque $\varepsilon \in B^*$. D'où $A \subseteq A^*B^*$, et on prouve de manière similaire que $B \subseteq A^*B^*$. L'inclusion $A \cup B \subseteq A^*B^*$ en découle et entraîne $(A \cup B)^* \subseteq (A^*B^*)^*$, par la question 4.

   b) De $A \subseteq A \cup B$ (resp. $B \subseteq A \cup B$) on tire, par la question 4, $A^* \subseteq (A \cup B)^*$ (resp. $B^* \subseteq (A \cup B)^*$), puis $A^*B^* \subseteq (A \cup B)^*$, puisque $(A \cup B)^*$ est clos pour la concaténation. Il découle $(A^*B^*)^* \subseteq ((A \cup B)^*)^*$ (toujours par la question 4), puis $(A^*B^*)^* \subseteq (A \cup B)^*$ puisque $((A \cup B)^*)^* = (A \cup B)^*$ (question 5).

Exercice 1.8 Soient A, B deux langages sur un même alphabet.

1. Comparer $(A \cup B)^*$ et $A^* \cup B^*$.
2. Comparer $(A \cap B)^*$ et $A^* \cap B^*$.
3. Comparer $(AB)^*$ et A^*B^*.

Corrigé

1. Nous avons $(A^* \cup B^*) \subseteq (A \cup B)^*$ mais l'inverse n'est pas toujours vrai.

   Montrons que $(A^* \cup B^*) \subseteq (A \cup B)^*$. De $A \subseteq A \cup B$ (resp. $B \subseteq A \cup B$) on tire, par la question 4 de l'exercice 7, $A^* \subseteq (A \cup B)^*$ (resp. $B^* \subseteq (A \cup B)^*$), d'où $(A^* \cup B^*) \subseteq (A \cup B)^*$.

   Montrons maintenant qu'il existe A et B tel que $(A \cup B)^* \not\subset (A^* \cup B^*)$. Prenons $A = \{a\}$ et $B = \{b\}$. On a $ab \in (A \cup B)^*$ mais $ab \notin A^* \cup B^*$ (les mots de $A^* \cup B^*$ sont formés soit uniquement de a ou uniquement de b).

2. Nous avons $(A \cap B)^* \subseteq A^* \cap B^*$ mais l'inverse n'est pas toujours vrai.

   Montrons que $(A \cap B)^* \subseteq A^* \cap B^*$. Soit w $\in (A \cap B)^*$ donc w peut s'écrire $u = u_0, \ldots, u_n$ avec $n \geq 0$ et $u_i \in A \cap B$ (pour tout i entre 0 et n) *c-à-d,* $u_i \in A$ et $u_i \in B$ et donc $u_i \in A^*$ *et* $u_i \in B^*$ *d'où* $u_i \in A^* \cap B^*$.

   Montrons maintenant qu'il existe A et B tel que $A^* \cap B^* \not\subset (A \cap B)^*$. Prenons $A = \{a\}$ et $B = \{aa\}$. On a $aa \in A^* \cap B^*$ mais $aa \notin (A \cap B)^*$ *(car* $A \cap B = \emptyset$ *et donc* $(A \cap B)^* = \{\varepsilon\}$ *).*

3. Nous avons $(AB)^* \neq A^*B^*$.

   Prenons $A = \{a\}$ et $B = \{B\}$.

   $AB = \{ab\}$ *et* $(AB)^* = \{\varepsilon, ab, abab, ababab, \ldots\} = \{(ab)^n \mid n \geq 0\}$

   $A^* = \{\varepsilon, a, aa, aaa, \ldots\}$, $B^* = \{\varepsilon, b, bb, bbb, \ldots\}$ *et*

   $A^*B^* = \{\varepsilon, b, bb, bbb, \ldots, a, ab, abb, abbb, \ldots., aa, aab, aabb, aabbb, aaa, aaab, aaabb, aaabbb, \ldots..\} = \{a^nb^m \mid n, m \geq 0\}$

   Nous avons par exemple, $abab \in (AB)^*$ mais $abab \notin A^*B^*$ (car les mots de A^*B^* sont formés d'une suite de a suivi d'une suite de b, c-à-d une fois qu'on a un b on ne peut plus avoir un a à nouveau), d'où $(AB)^* \not\subset A^*B^*$.

   De plus par exemple, $aa \in A^*B^*$ *(prendre* $n = 2$ *et* $m = 0$*)* mais $aa \notin (AB)^*$ (car dans $(AB)^*$, on ne peut jamais avoir deux a consécutifs), d'où $A^*B^* \not\subset (AB)^*$.

Exercice 1.9 Trouver l'ensemble des facteurs gauches et l'ensemble des facteurs droits des langages sur l'alphabet $\Sigma = \{a,b\}$ suivants :

1. $L_1 = \{a^n.b^n \mid n \geq 0\}$.
2. $L_2 = \{a^n.b^m \mid 0 \geq n \geq m\}$.
3. $L_3 = \{w \in \Sigma^* \mid |w|_a = |w|_b\}$.

Corrigé

1. $L_1 = \{a^n.b^n \mid n \geq 0\} = \{\varepsilon, ab, aabb, ...\}$

   a) Facteurs gauches (préfixes). On peut montrer que

$$Préfixes(L_1) = \{ a^n.b^m \mid n \geq m \geq 0\}$$

   Notons l'ensemble $\{ a^n.b^m \mid n \geq m \geq 0\}$ *par* X. Pour cela, on montre que tout préfixe d'un élément de L_1 est dans X et tout élément de X est dans $Préfixes(L_1)$ *i.e.,* un préfixe d'un élément de L_1 .

   Il est clair que tous préfixe d'un mot de L_1 est de la forme $a^n.b^m$ avec $n \geq m \geq 0.$

   Tout mot de la forme $a^n.b^m$ *pour* $n \geq m \geq 0$ représente un préfixe du mot $a^n.b^n$ de L_1 ($a^n.b^n$ s'obtient en ajoutant $(n\text{-}m)$ b à la fin de $a^n.b^m$ *)*.

   b) Facteurs droits (suffixes). Avec un raisonnement similaire, on trouve :

$$Suffixes(L_1) = \{ a^n.b^m \mid m \geq n \geq 0\}$$

2. $L_2 = \{a^n.b^m \mid n \geq m \geq 0\}$. .

   a) Facteurs gauches (préfixes). On peut montrer que

$$Préfixes(L_2) = L_2$$

   Pour cela, on montre que tout préfixe d'un élément de L_2 est dans L_2 et tout élément de L_2 est dans $Préfixes(L_2)$ *i.e.,* un préfixe d'un élément de L_2 .

   Soit w est un préfixe d'un élément quelconque $a^n.b^m$ $(n \geq m \geq 0)$ de L_2 , alors soit w est de la forme a^r avec $r \geq 0$ et il est dans L_2, ou bien de la forme $a^n.b^k$ avec $k \leq m$ *et donc* $n \geq k \geq 0$ et w est dans ce cas aussi dans L_2.

   Soit w un élément quelconque de L_2. w est de la forme $a^n.b^m$ $(n \geq m \geq 0)$, mais comme w est préfixe de lui-même alors w est dans *Préfixes(L2).*

   b) Facteurs droits (suffixes). On peut montrer que

$$Suffixe(L_2) = \{ a^n.b^m \mid n, m \geq 0\}$$

   Notons l'ensemble $\{ a^n.b^m \mid n, m \geq 0\}$ *par* X. On montre alors que tout suffixe d'un élément de L_2 est dans X et tout élément de X est un suffixe d'un élément de L_2 .

   Soit w est un suffixe d'un élément quelconque $a^n.b^m$ $(n \geq m \geq 0)$ de L_2 , alors soit w est de la forme b^r avec $r \geq 0$ et il est dans X, ou bien de la forme $a^r.b^m$ avec $r, m \geq 0$ et w est dans ce cas aussi dans X.

Soit w un élément quelconque de X. w est de la forme $a^n.b^m$ avec $n, m \geq 0$. *On a deux cas :*

- *Soit* $n \geq m$ et donc w est dans L_2 mais comme w est suffixe de lui-même alors w est suffixe d'un élément de L_2.
- *Soit* $n < m$ et dans ce cas w est suffixe par exemple de $a^{m-n}\ w = a^{m-n}a^nb^m = a^mb^m$ qui est un élément de L_2.

3. $L_3 = \{ w \in \Sigma^* \mid |w|_a = |w|_b\}$

a) Facteurs gauches (préfixes). On peut montrer que

$$Préfixes(L_3) = \Sigma^* = \{a,b\}^*$$

Pour cela, on montre que tout préfixe d'un élément de L_3 est dans Σ^* et tout élément de Σ^* est dans $Préfixes(L_3)$ *i.e.,* un préfixe d'un élément de L_3.

La première partie est triviale car tout préfixe d'un élément de L_3 est un mot sur l'alphabet $\Sigma = \{a,b\}$.

Soit w un mot quelconque de Σ^*. On peut toujours compléter w par des symboles a et b à droite pour obtenir un mot w' avec $|w'|_a = |w'|_b$ et donc w est un préfixe d'un élément de L_3.

b) Facteurs droits (suffixes). Par un raisonnement similaire, on montre que :

$Suffixe(L_3) = \Sigma^* = \{a,b\}^*$