TD Algorithmique du texte

1 Notations et définitions de base

Exercice 1 ***Facteurs, préfixes, suffixes***

1. *Donner tous les facteurs du mot abbbaaa.*
   Correction
   ε, a, b, aa, ab, ba, bb, aaa, abb, baa, bba, bbb, abbb, baaa, bbaa, bbba, abbba, bbaaa, bbbaa, abbbaa, bbbaaa, abbbaaa.
2. *Donner la liste des préfixes de abbaa.*
   Correction
   ε, a, ab, abb, abba, abbaa.
3. *Donner la liste des suffixes de abcd.*
   Correction
   ε, d, cd, bcd, abcd.
4. *Combien de préfixes a un mot de longueur n ?*
   Correction
   $n + 1$ *(autant que de positions possibles pour l'indice de fin)*
5. *Combien de facteurs a un mot de longueur n ?*
   Correction
   $n + 1$ *(autant que de positions possibles pour l'indice de début)*
6. *Combien de facteurs (distincts) possède le mot* a^n *?*
   Correction
   $n + 1$ *(autant que de nombres de a possibles dans le mot)*
7. *Combien de facteurs (distincts) possède le mot* $a^m b^n$ *?*
   Correction
   Distinguons le mot vide, et puis les mots qui ne contiennent que des a (au moins un) : il y en a m, les mots qui ne contiennent que des b (au moins un) : il y en a n, et les mots qui contiennent au moins un a et un b (on imagine un curseur de début entre la position 0 et la position $m-1$*, et un curseur de fin entre la position m et la position* $n+m-1$*) : nm, ce qui donne un total de* $nm+n+m+1 = (n+1)(m+1)$ *facteurs distincts.*

Exercice 2

1. *Compter les occurrences des lettres a et b dans les mots suivants :* a^3cbbca*, aabgjdd, titi, babc.*
   Correction

w	a^3cbba	$aabgjdd$	$titi$	$babc$
$\|w\|_a$	4	2	0	1
$\|w\|_b$	2	1	0	2

2. *Donner l'ensemble des couples* (u, v) *tels que* $uv = abaac$*.*
   Correction

$$\{(\varepsilon, abaac), (a, baac), (ab, aac), (aba, ac), (abaa, c), (abaac, \varepsilon)\}$$

3. *Calculer LM pour les ensembles suivants :*
   (a) $L = \{a, ab, bb\}$ *et* $M = \{\varepsilon, b, a^2\}$.
   Correction
   $$LM = \{a, ab, a^3, abb, aba^2, bb, b^3, b^2a^2\}$$
   (b) $L = \emptyset$ *et* $M = \{a, ba, bb\}$.
   Correction
   $$LM = \emptyset$$
   (c) $L = \{\varepsilon\}$ *et* $M = \{a, ba, bb\}$.
   Correction
   $$LM = M$$
   (d) $L = \{aa, ab, ba\}$ *et* $M = \{a, b\}^*$.
   Correction
   $$LM = \{aax, abx, bax | x \in \{a, b\}^*\} = aaM \cup abM \cup baM$$

Exercice 3 *Soit* $L = \{ab, ba\}$. *Parmi les mots suivants, lesquels sont dans* L^* *: abba, ababa, aab, ababab,* ε*, baab, bbaabb ?*

Correction
abba, ababab, ε*, baab.*

2 Palindromes

Soit $\mathcal{P}$ l'ensemble des langages ne contenant que des palindromes sur l'alphabet $A = \{a, b, c\}$.

Exercice 4

1. *Donner un exemple de langage qui est dans* $\mathcal{P}$.
   Correction
   $$L = \{aa, b, ccc\}$$
2. *Est-ce que les langages suivants sont dans* $\mathcal{P}$ *?*
   (a) $L_1 = \{a^n \mid n \in \mathbb{N}\}$
   Correction
   Oui, tous les mots sont constitués uniquement du caractère a, ils sont donc des palindromes.
   (b) $L_2 = \{a^n b^n \mid n \in \mathbb{N}\}$
   Correction
   Non, par exemple le mot ab est dans L_2 *mais n'est pas un palindrome.*

(c) $L_3 = \{a^n b a^m \mid n, m \in \mathbb{N}\}$
Correction
Non, par exemple le mot ab est dans L_3 mais n'est pas un palindrome.

(d) $L_4 = \{c a^n b a^n c \mid n \in \mathbb{N}\}$
Correction
Oui, tous les mots sont identiques à leur mot miroir (car c'est le même n en exposant du a) ils sont donc des palindromes.

Exercice 5 *Est-ce que $\mathcal{P}$ est stable par union, intersection, concaténation et le passage au carré $(L.L)$?*

Correction
On dit qu'un ensemble est stable par une opération binaire si pour tout couple d'éléments dans cet ensemble, le résultat de l'opération sur ce couple est également dans l'ensemble. Donc pour montrer une stabilité, il faut le montrer en toute généralité pour tout couple d'éléments, et pour montrer qu'il n'y a pas stabilité, il suffit de donner un contre-exemple, c'est-à-dire un couple d'éléments de l'ensemble tel que le résultat de l'opération sur ces éléments n'est pas dans l'ensemble.

Union Si L_1 et L_2 sont deux langages de $\mathcal{P}$, alors $L_1 \cup L_2$ est constitué de mots qui sont soit dans L_1, soit dans $L2$, mais dans les deux cas ce sont des palindromes, dont $L_1 \cup L_2$ est également dans $\mathcal{P}$.

Intersection Si L_1 et L_2 sont deux langages de $\mathcal{P}$, alors $L_1 \cap L_2$ est constitué de mots qui sont dans L_1 et dans $L2$, ce sont donc des palindromes, dont $L_1 \cap L_2$ est également dans $\mathcal{P}$.

Concaténation On peut construire un contre-exemple : $L_1 = \{a^n | n \in \mathbb{N}\}$ et $L_2 = \{b^m | m \in \mathbb{N}\}$ sont deux langages de $\mathcal{P}$, mais $L_1 L_2 = \{a^n b^m | n, m \in \mathbb{N}\}$ n'est pas constitué que de palindromes, par exemple il contient le mot ab qui n'en est pas un.

Carré On peut également trouver un contre-exemple. Le langage L_4 de l'exercice précédent est constitué uniquement de palindromes, mais son carré $L_4^2 = \{c a^n b a^c c a^m b a^m c | n, m \in \mathbb{N}\}$ contient le mot $cbccabac$ qui n'est pas un palindrome.

3 Conjugaison

Deux mots u et v sont dits *conjugués* s'il existe deux mots w_1 et w_2 tels que $u = w_1 w_2$ et $v = w_2 w_1$. En d'autres termes, v s'obtient à partir de u par permutation cyclique de ses lettres.

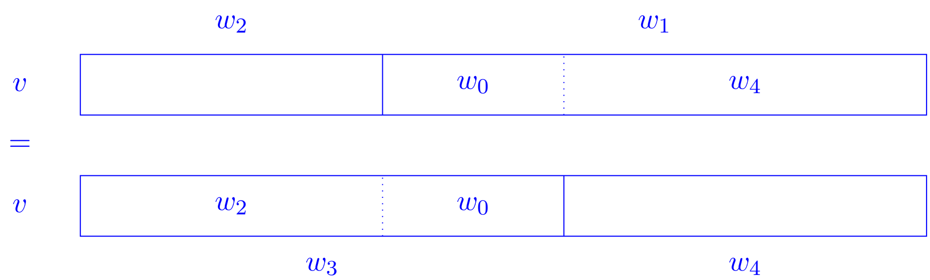
Exercice 6 *Montrer que la conjugaison est une relation d'équivalence.*

Correction
On rappelle les propriétés qui définissent une relation d'équivalence : il s'agit d'une relation binaire réflexive, symétrique et transitive.

Réflexivité Une relation binaire est réflexive si tout élément est en relation avec lui-même. Ici, on doit montrer que tout mot est conjugué avec lui-même. Soit u un mot, il existe w_1 et w_2 tels que $u = w_1 w_2$ et $u = w_2 w_1$, il suffit de prendre $w_1 = u$ et $w_2 = \varepsilon$, par exemple.

Symétrie Une relation binaire est symétrique si pour tout couple (x, y) d'éléments qui sont en relation, alors le couple (y, x) est également en relation. Ici, on doit montrer que pour tous mots u et v, si u est conjugué avec v, alors v est conjugué avec u. C'est presque trivial, il suffit d'échanger les rôles de w_1 et w_2 dans la définition pour conclure.

Transitivité Une relation binaire est transitive si, à chaque fois que l'on a un couple (x, y) et un couple (y, z) dans la relation, alors le couples (x, z) est aussi dans la relation. On doit montrer que pour tous mots u, v, w, si u et v sont conjugués, et v et w sont conjugués, alors u et w sont conjugués aussi. Soient u, v, w des mots tels que u et v sont conjugués, et v et w sont conjugués. Alors il existe w_1, w_2, w_3, w_4 tels que $u = w_1w_2$ et $v = w_2w_1$ d'une part, et $v = w_3w_4$ et $w = w_4w_3$ d'autre part. On remarque que $v = w_2w_1 = w_3w_4$. On suppose en toute généralité que w_2 est plus petit que w_3 et donc est un préfixe de w_3. Donc w_3 s'écrit sous la forme w_2w_0. On remarque alors que $v = w_2w_1 = w_2w_0w_4$. Donc $w_1 = w_0w_4$. On a alors : $u = w_0w_4w_2$ et $w = w_4w_2w_0$. En posant $a = w_0$ et $b = w_4w_2$, on a $u = ab$ et $w = ba$, donc u et w sont conjugués.



Exercice 7 *Montrer que u et v sont conjugués si et seulement s'il existe un mot w tel que $uw = wv$.*

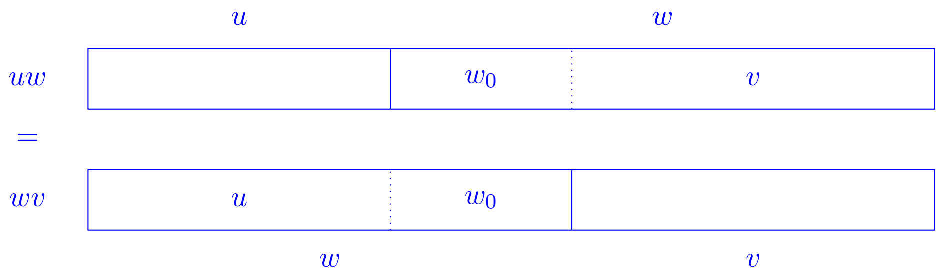
Correction

Il s'agit de montrer une équivalence, on va donc démontrer le sens direct et le sens réciproque.

$\Rightarrow$: Soient u et v deux mots conjugués. On veut démontrer qu'il existe un mot w tel que $uw = wv$. On sait que u et v sont conjugués, donc par définition il existe des mots w_1 et w_2 tels que $u = w_1w_2$ et $v = w_2w_1$. Si on prend $w = w_1$, on a alors $uw = w_1w_2w_1 = w_1v = wv$, ce qui prouve la propriété voulue.

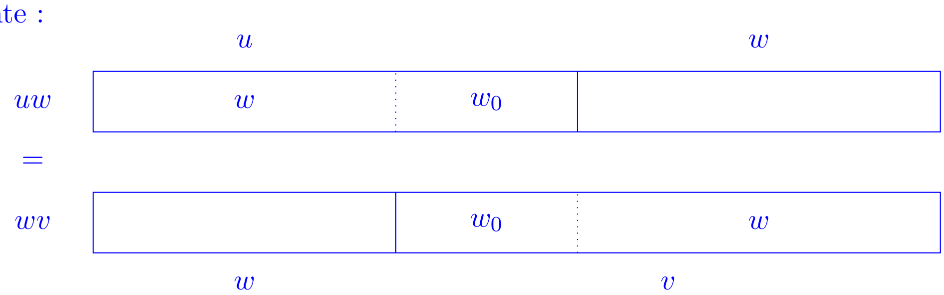
$\Leftarrow$: Soient deux mots u et v tels qu'il existe un mot w avec $uw = wv$. Pour se simplifier la tâche, on va avancer un argument de minimalité : on suppose que w est le mot le plus petit tel que $uw = wv$. Examinons maintenant ce mot uw : il y a deux cas possibles : soit w est plus grand que u, soit w est de taille strictement inférieure à u.

Dans le premier cas, on a la configuration suivante :



Donc w s'écrit sous la forme uw_0, mais aussi sous la forme w_0v, ou w_0 est un mot strictement plus petit que w. Comme on a $uw_0 = w_0v$, avec w_0 strictement plus petit que w, cela contredit la minimalité de w pour cette propriété et on aboutit donc à une contradiction.

Seul le deuxième cas reste donc possible, et l'on a la configuration suivante :



On a alors w qui est préfixe de u, qui s'écrit donc sous la forme $u = ww_0$. De même, on a $v = w_0w$. En prenant $w_1 = w$ et $w_2 = w_0$, on a bien trouvé deux mots w_1 et w_2 tels que $u = w_1w_2$ et $v = w_2w_1$, donc u et v sont conjugués.

4 Mots de Fibonacci

On considère l'alphabet $\Sigma = \{a, b\}$. On définit les mots de Fibonacci par :

$$\begin{cases} Fib_0 & = & \varepsilon \\ Fib_1 & = & b \\ Fib_2 & = & a \\ Fib_n & = & Fib_{n-1}Fib_{n-2} \text{ pour tout } n \geq 2 \end{cases}$$

Exercice 8 *Donner les mots de Fibonacci jusqu'à $n = 8$. Démontrez par récurrence sur $n \geq 0$ que la longueur de Fib_n est F_n, le nombre de Fibonacci d'ordre n.*

Correction

$$\begin{array}{lcl} Fib_0 & = & \varepsilon \\ Fib_1 & = & b \\ Fib_2 & = & a \\ Fib_3 & = & ab \\ Fib_4 & = & aba \\ Fib_5 & = & abaab \\ Fib_6 & = & abaababa \\ Fib_7 & = & abaababaabaab \\ Fib_8 & = & abaababaabaababaababa \end{array}$$

Rappel : les nombres de Fibonacci sont définis par récurrence par

$$\begin{cases} F_0 & = & 0 \\ F_1 & = & 1 \\ F_n & = & F_{n-1} + F_{n-2} \text{ pour tout } n \geq 2 \end{cases}$$

Pour la récurrence, on pose la propriété P_n : "$|Fib_n| = F_n$".
Base : $|Fib_0| = |\varepsilon| = 0 = F_0$, $|Fib_1| = |a| = 1 = F_1$, et $|Fib_2| = |b| = 1 = F_2$.
Récurrence : Soit $n \geq 2$. On suppose que $\forall k \in \{0, \dots n\}$, la proposition P_k est vraie (il s'agit donc d'une récurrence généralisée). On veut montrer qu'on en déduit la propriété au rang $n+1$. Par définition des mots de Fibonacci, on a $Fib_{n+1} = Fib_n Fib_{n-1}$. Donc $|Fib_{n+1}| = |Fib_n| + |Fib_{n-1}| = F_n + F_{n-1} = F_{n+1}$. Donc P_{n+1} est vraie.
Conclusion : Pour tout $n \geq 0$, P_n est vraie.

Exercice 9

1. *Montrer que pour $n \geq 3$, Fib_n est un préfixe de tous ses successeurs.*
   Correction

   Soit $n \geq 3$. On pose P_k la propriété "Fib_n est préfixe de Fib_{n+k}". Montrons par récurrence que pour tout $k \geq 0$, P_k est vraie.
   Base : Fib_n est préfixe de lui-même.
   Récurrence : Soit $k \geq 0$. On suppose que la proposition P_k est vraie. On veut montrer qu'on en déduit la propriété au rang $k+1$. Par définition des mots de Fibonacci, on a $Fib_{n+k+1} = Fib_{n+k} Fib_{n+k-1}$. Par hypothèse de récurrence, Fib_n est préfixe de Fib_{n+k}, donc aussi de Fib_{n+k+1}.
   Conclusion : Pour tout $k \geq 0$, P_k est vraie.
2. *Montrer que pour $n \geq 4$, le carré de Fib_n est un préfixe de tous ses successeurs à partir de Fib_{n+2}.*

Correction

Soit $n \geq 4$. Puisque Fib_{n+2} est préfixe de tous ses successeurs, il suffit de montrer que Fib_n^2 est un prefixe de Fib_{n+2}. On a $Fib_{n+2} = Fib_{n+1} Fib_n = Fib_n Fib_{n-1} Fib_n = Fib_n Fib_{n-1} Fib_{n-1} Fib_{n-2}$. On décompose le deuxième Fib_{n-1} :
$Fib_{n+2} = Fib_n Fib_{n-1} Fib_{n-2} Fib_{n-3} Fib_{n-2} = Fib_n Fib_n Fib_{n-3} Fib_{n-2}$ donc F_n^2 est préfixe de Fib_{n+2}, ce qui clôt la preuve.

5 Bords et périodes

Exercice 10

1. *Soit x un mot non vide. Soit u le plus petit mot tel que x est préfixe de ux. Montrer que $|u| = period(x)$.*
   Correction

   Montrons que $|u|$ est une période de x : x est préfixe de ux, donc on a $x[i] = x[i + |u|]$ pour tout $i \in \{0, \dots, |x| - |u| + 1\}$, ce qui est la définition d'une période. Montrons que c'est la plus petite. D'après la *proposition du cours sur la caractérisation des périodes, si p est une période, alors x peut s'écrire sous la forme $w = yw = wz$ avec $|y| = p$. Donc x est préfixe de $xz = ywz = yx$. Donc y vérifie l'hypothèse. Comme u est de taille minimale, $|u|$ est la plus petite période de x, d'où $period(x) = |u|$.*

2. *Soit x un mot non vide. Montrer que les trois propositions suivantes sont équivalentes :*
   (a) *$period(x^2) = |x|$,*
   (b) *x est primitif, c'est-à-dire ne peut être écrit sous la forme u^k pour $k > 1$,*
   (c) *x^2 contient seulement 2 occurrences de x.*

Correction

Pour montrer toutes ces équivalences, il suffit de montrer que $(a) \Rightarrow (b)$, $(b) \Rightarrow (c)$ et $(c) \Rightarrow (a)$, toutes les autres implications pouvant se déduire de ces trois là. Chacune de ces implications se montre assez rapidement par contraposée (c'est-à-dire que l'on contredit la conclusion, et on en déduit que l'hypothèse ne peut pas être vraie).

- $(a) \Rightarrow (b)$. *On suppose que x n'est pas primitf, donc qu'il s'écrit sous la forme $x = u^k$ avec $k > 1$. Alors $x^2 = u^{2k}$ et il admet donc $|u|$ comme période. Sa période la plus petite est donc $\geq |u| < |x|$.*
- $(b) \Rightarrow (c)$. *On suppose que x^2 contient strictement plus de deux occurrences de x. On considère la deuxième occurrence de x dans x^2 et u le préfixe précédent cette occurrence. Alors on a $|u| > 0$ (puisque c'est la deuxième occurence) et $|u| < |x|$ (puisque ce n'est pas la dernière). Alors en effectuant un aller-retour entre la première et la deuxième occurrence de x, et en exploitant leur chevauchement, on montre que x s'écrit sous la forme $u^k v$ avec $|v| < |u|$ et v préfixe de u. Si $v \neq \varepsilon$, alors $|v|$ est une période de x et on aurait une occurrence de x après le préfixe v dans x^2, ce qui contredit sa minimalité. Donc $v = \varepsilon$, nécessairement $k > 1$ (car $|u| < |x|$) et x est primitif.*
- $(c) \Rightarrow (a)$. *On suppose que $period(x) < |x|$. Alors on a aussi une occurrence de x à la position p de x^2, ce qui constitue une troisième occurrence de x dans x^2.*