Exercices d'algorithmique du texte

Préfixes, suffixes et facteurs

EXERCICE 1. – Facteurs, préfixes et suffixes d'un mot.

Un mot u est un *facteur* d'un mot v si il existe deux mots w_1 et w_2 tels que $v = w_1 u w_2$. Un mot u est un *préfixe* d'un mot v si il existe w tel que $v = uw$. De même, un mot u est un *suffixe* d'un mot v si il existe w tel que $v = wu$.

Attention : ne pas oublier que w, w_1 et w_2 peuvent être égal au mot vide ε. Ainsi, un mot u est un facteur (ainsi qu'un préfixe et un suffixe) de lui-même.

a) Donner tous les facteurs du mot $abbbaaa$.
b) Donner la liste des préfixes de $abbaa$.
c) Donner la liste des suffixes de $abcd$.
d) Combien de préfixes a un mot de longueur n ?
e) Combien de facteurs (distincts) possède le mot a^n ?
f) Combien de facteurs (distincts) possède le mot $a^m b^n$?

Autres définitions et propriétés de base

Exercice 1.2

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

► **Exercice 2** ◄ Soit $L = \{ab, ba\}$. Parmi les mots suivants, lesquels sont dans L^* : $abba$, $ababa$, $aabb$, $ababab$, ε, $baab$, $bbaabb$?

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 $xay = ybx$.

Palindromes

► **Exercice 5** ◄ Soit $\mathcal{P}$ l'ensemble des langages ne contenant que des palindromes sur l'alphabet $A = \{a, b, c\}$. Est-ce que les langages suivants sont dans $\mathcal{P}$?

- $L_1 = \{a^n \mid n \in \mathbb{N}\}$
- $L_2 = \{a^n b^n \mid n \in \mathbb{N}\}$
- $L_3 = \{a^n b a^m \mid n, m \in \mathbb{N}\}$
- $L_4 = \{c a^n b a^n c \mid n \in \mathbb{N}\}$

Est ce que $\mathcal{P}$ est stable pour l'union, l'intersection, la concaténation et le passage au carré $(L \cdot L)$?

Conjugaison

Exercice 1.9 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.

1. Montrer que la conjugaison est une relation d'équivalence, c'est-à-dire :
   – tout mot u est conjugué à lui-même ;
   – si u est conjugué à v, alors v est conjugué à u ;
   – si u est conjugué à v et v est conjugué à w, alors u est conjugué à w.
2. Montrer que u et v sont conjugués si et seulement s'il existe un mot w tel que $uw = wv$.

Mots de Fibonacci

Exercice 12 *Donner les mots de Fibonacci jusqu'à $n = 8$.*

Exercice 13 *Montrer que pour $n \geq 2$, Fib_n est un préfixe de Fib_{n+1}.*

Exercice 15 *Pour $n > 2$, on note g_n le préfixe de Fib_n de longueur $|Fib_n| - 2$. Montrer que pour tout $n > 5$, g_n est un préfixe de Fib^2_{n-1} et de Fib^3_{n-2}.*

Codes

EXERCICE 6. – Codes.

Un langage C est appelé un *code* si et seulement si tout mot de C^* se décompose de manière unique comme une concaténation d'éléments de C.

a) Soit $\Sigma = \{a, b\}$. Les langages suivants sont-ils des codes ?
- $C_1 = \{ab, ba, a\}$,
- $C_2 = \{aa, ab, ba, b\}$,
- $C_3 = \{ab, baa, abba, aabaa\}$.

b) Donner un exemple de code sur l'alphabet $\{a, b\}$ possédant au moins 6 mots.

c) Montrer que si C est un code, alors $\varepsilon \notin C$.

d) On suppose $\varepsilon \notin C$. Montrer que si C n'est pas un code, alors il existe $u, v, x, y \in C$ tels que u est un préfixe propre de v et x est un suffixe propre de y (on rappelle qu'un préfixe *propre* de u est un préfixe de u différent du mot vide ε et de u lui-même ; on définit de même la notion de suffixe propre).

e) La réciproque à la question d) est-elle vraie ?

f) Existe-t-il un algorithme qui, étant donné un code C composé d'un nombre fini de mots, répond *oui* si C est un code et *non* sinon ?

Bord et période

► **Exercice 10** ◄ Soit u un mot sur l'alphabet A, un bord de u est un mot, différent de u lui-même, qui est à la fois préfixe et suffixe de u. Quels sont tous les bords de $u_1 = aabaaba$ et de $u_2 = baabaabab$? Montrer que la relation "est un bord de" est transitive.

Exercice 4

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

Exercice 5

1. Montrer que tout mot non vide x peut être écrit de façon unique sous la forme u^k, où u est un mot primitif et $k > 0$. Le mot u est appelé la racine de x, et k son exposant.
2. On note $A = \{a, b\}$. Montrer que l'ensemble $\{y \in A^* : xy = yx\}$ est égal à u^* où u est la racine de x.

Exercice 6 *On rappelle qu'un mot w est dit* sans bord *si ses seuls bords sont lui-même et le mot vide, c'est-à-dire si $period(w) = |w|$. On suppose qu'un mot x a un bord minimal non vide u. Montrer que u est sans bord et que soit $x = u$ soit il existe un mot v tel que $x = uvu$.*

Exercice 10 *Soit $Maxsuf(x)$ le suffixe maximal du mot non vide x, en considérant l'ordre lexico-graphique.*

1. Donner $Maxsuf(x)$ pour $x =$abacabbac.
2. Soit x un mot non vide, et a une lettre. Soit $u = Maxsuf(x)$ et $va = Maxsuf(xa)$. Montrer que v est un bord de u.

Exercice 11 *Soit w un mot non vide, u un bord propre de w et v un bord propre de u. Montrer que $|w| > |u| + |v| + 1$.*

Résultats du cours

Proposition

Soit x un mot non vide et p un entier tel que $0 < p \leq |x|$. Chacune des conditions équivalentes suivantes définit une période :

0. *p est une période de x*
1. *x est un facteur d'un mot y^k avec $|y| = p$ et $k > 0$,*
2. *x peut être écrit sous la forme $(uv)^k u$ avec $|uv| = p$, v non vide et $k > 0$,*
3. *il existe des mots y, z et w tels que $x = yw = wz$ et $|y| = |z| = p$.*

Démonstration :

1. On suppose 0 vrai. Extraire un bord de *x*. En déduire que 0 => 3.
2. On suppose 3 vrai. « Découper » *w* en tronçons de longueur *p*. Prouver que ces tronçons sont identiques puis prouver 2.
3. Prouver 2 => 1.
4. On suppose 1 vrai. En considérant la périodicité de y^k, démontrer 0 (revenir à la définition).

Lemma (Périodicité)

Soient p et q deux périodes d'un mot x. Si $p + q - \mathrm{pgcd}(p, q) \leq |x|$, alors $\mathrm{pgcd}(p, q)$ est aussi une période de x.

1. On procède par induction sur max(p,q). Expliciter la proposition à démontrer pour prouver le lemme.
2. Prouver la base de l'induction.
3. Hérédité :
   a) Expliciter l'hypothèse d'induction et la proposition à démontrer pour prouver l'hérédité.
   b) Pourquoi peut-on supposer p>q ?
   c) Démontrer que x peut s'écrire sous la forme x=uy avec |u|=p et x=vz avec |v|=q.
   d) Démontrer que p-q est une période de z.
   e) Démontrer que q≤|z| puis que q est une période de z.
   f) Démontrer que pgcd(p,q) est une période de z.
   g) Démontrer que q≤|x|/2 puis que v est un préfixe de z.
   h) En appelant t le préfixe de x de longueur pgcd(p,q), démontrer que z est un préfixe d'une puissance de t.
   i) En déduire que |t|=pgcd(p,q) est une période de x.

Retour sur Fibonacci

Soit β le morphisme qui à un mot associe le même mot dont les deux dernières lettres sont inversées. On rappelle la définition des mots de Fibonacci : $f_0=\varepsilon$, $f_1=b$, $f_2=a$, $f_n=f_{n-1}f_{n-2}$ pour $n\geq3$. On rappelle que la longueur de f_i est égale à F_i , terme de rang i de la suite de Fibonacci ($F_0=0$, $F_1=1$ et $F_n=F_{n-1}+F_{n-2}$ pour $n\geq2$). Enfin, on appelle g_n le mot f_n privé de ses deux dernières lettres.

1. Démontrer que, pour tout $n\geq3$, $\beta(f_n)=f_{n-2}f_{n-1}$
2. Démontrer que, pour tout $n\geq6$, $g_n = f_{n-2}^2\, g_{n-3}$.
3. Démontrer que, pour tout $n\geq3$, g_n est un préfixe de f_{n-1}^2 et de f_{n-2}^3 .
4. Démontrer que, pour tout $n\geq5$, F_{n-2} et F_{n-1} sont deux périodes de g_n.
5. Démontrer que, pour tout $n\geq2$, F_{n-2} et F_{n-1} sont premiers entre eux (pgcd = 1).
6. En déduire que la condition « $p+q-\mathrm{pgcd}(p,q)\leq|x|$ » du lemme de périodicité est optimale pour des valeurs de p, q et x que l'on précisera.