MOTS ET LANGAGES

Dans toute cette série d'exercices, A est un alphabet.

1) Soit $A = \{a, b\}$. Montrez que tout mot de A^* s'écrit de manière unique sous la forme

$$a^{m_0}ba^{m_1}b\dots a^{m_{p-1}}ba^{m_p}$$

où p, $m_0, \dots, m_p$ sont des entiers naturels.

2) On appelle miroir d'un mot $u = a_1 \dots a_p$ le mot $\overline{u} = a_p \dots a_1$.
Pour tout couple de mots (u, v), comparez $\overline{u.v}$, $\overline{u}$ et $\overline{v}$.

3) Soit $u_1, u_2, u_3, u_4, v_1, v_2, v_3, v_4$ 8 mots de A^* tel que $u_1u_4 = v_1v_4$, $u_1u_2u_4 = v_1v_2v_4$ et $u_1u_3u_4 = v_1v_3v_4$. Montrez que $u_1u_2u_3u_4 = v_1v_2v_3v_4$.

4) Soient u, v, h trois mots de A^* tels que $uh = hv$. Montrez qu'il existe deux mots x, y et deux entiers naturels m, n tels que $u = (xy)^n$, $v = (yx)^n$ et $h = x(yx)^m$.

5) Deux mots u et v sont dits conjugués s'il existe des mots s et t tels que $u = st$ et $v = ts$.

a) Montrez que la relation binaire $\sim$ sur A^* définie par « $u \sim v$ si et seulement si u et v sont conjugués » est une relation d'équivalence.

b) Montrez que pour tout $n \geqslant 1$, $u \sim v \iff u^n \sim v^n$.

c) Montrez que $u \sim v$ si et seulement si il existe un mot w tel que $uw = wv$.

6) Distances sur les mots.
Une application $d : (A^*)^2 \to \mathbb{R}_+$ est appelée une distance sur A^* quand elle vérifie deux axiomes :

— l'axiome de séparation : $\forall (u, v) \in (A^*)^2 \quad d(u, v) = 0 \iff u = v$

— l'inégalité triangulaire : $\forall (u, v, w) \in (A^*)^3 \quad d(u, v) \leqslant d(u, w) + d(w, v)$

a) Montrez que l'application d définie par « $d(u, v) = 0$ si $u = v$ et $d(u, v) = 1$ sinon » est une distance sur A^*, appelée la distance triviale.

b) Si u, v sont deux mots, on appelle $\mathrm{plpc}(u, v)$ le plus long préfixe commun à u et v. On pose alors $d(u, v) = |uv| - 2|\,\mathrm{plpc}(u, v)|$.
Montrez que d est une distance sur A^*, appelée la distance préfixe.

7) Soit u, v deux mots de A^*. Si $v = ux$ où x est un mot, alors on définit le symbole $u^{-1}v$ par $u^{-1}v = x$, sinon le symbole $u^{-1}v$ n'est pas défini.
Montrez que pour tout triplet de mots (u, v, w), $(uv)^{-1}w$ est défini si et seulement si $v^{-1}(u^{-1}w)$ est défini et que dans ce cas, ils sont égaux : $(uv)^{-1}w = v^{-1}(u^{-1}w)$.
Plus généralement, si L est un langage et u un mot, on note $u^{-1}L = \{v \in A^* \ / \ uv \in L\}$ (appelé le résiduel de L par rapport à u).
Si $L = A^*$, quel est son résiduel par rapport à un mot u ? Même question avec $L = \{a^nb^n \ / \ n \in \mathbb{N}\}$.

8) Soit L un langage non vide. Montrez que $\varepsilon \in L \iff L \subset L^2$.

9) Pour L un langage, on note $Pref(L)$ l'ensemble des préfixes des mots de L.

a) Si L, M sont deux langages, déterminez $Pref(LM)$.

b) Si L est un langage, déterminez $Pref(L^*)$.

10) Un langage L est un code si tout mot de L^* s'écrit de manière unique comme concaténation de mots de L.

a) Parmi les langages suivants, lesquels sont des codes ?

— $\{ab, baa, abba, aabaa\}$

— $\{b, ab, baa, abaa, aaaa\}$

— $\{aa, ab, aab, bba\}$

— $\{a, ba, bba, baab\}$

b) Soit u un mot. Montrez que $\{u\}$ est un code si et seulement si u n'est pas le mot vide.

c) Soit u, v deux mots distincts. Montrez que $\{u, v\}$ est un code si et seulement si $uv \neq vu$.

d) Un langage L est dit préfixe quand $\varepsilon \notin L$ et qu'aucun mot de L n'est préfixe propre d'un autre mot de L. Montrez qu'un langage préfixe est un code.