Feuille de TD n° 1 : Mots, Langages, Automates

A désigne un alphabet fini. Le mot vide est noté ε.

Exercice 1 : Généralités

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 $u \cdot v = abaac$.
3. Calculer $\mathcal{L} \cdot \mathcal{M}$ pour les ensembles suivants :
   - $\mathcal{L} = \{a, ab, bb\}$ et $\mathcal{M} = \{\varepsilon, b, a^2\}$;
   - $\mathcal{L} = \emptyset$ et $\mathcal{M} = \{a, ba, bb\}$;
   - $\mathcal{L} = \{\varepsilon\}$ et $\mathcal{M} = \{a, ba, bb\}$;
   - $\mathcal{L} = \{aa, ab, ba\}$ et $\mathcal{M} = A^*$.
4. Un mot u est un *facteur* d'un mot v si u apparaît à l'intérieur de v : v s'écrit $w_1 \cdot u \cdot w_2$ pour certains mots w_1 et w_2. Un mot u est un *sous-mot* d'un mot v si on peut obtenir u à partir de v par « effacement » de certaines lettres (pas forcément consécutives) de v. Le nombre d'*occurrences* d'un facteur (resp. sous-mot) u dans le mot v est le nombre de façons de voir u comme facteur (resp. sous-mot) de v.

   Donner le nombre d'occurrences du facteur aba dans le mot $v = ababab$. Donner le nombre d'occurrences du sous-mot aba dans le même mot v.
5. Montrer que le produit est une opération distributive par rapport à l'union, c'est-à-dire que, pour tous langages $\mathcal{L}$, $\mathcal{M}$ et $\mathcal{N}$, on a : $\mathcal{L} \cdot (\mathcal{M} \cup \mathcal{N}) = (\mathcal{L} \cdot \mathcal{M}) \cup (\mathcal{L} \cdot \mathcal{N})$. Montrer que le produit n'est pas distributif par rapport à l'intersection.

Exercice 2 : Un peu de `grep`

La commande `grep -E 'regexp' fichier` retourne les lignes d'un fichier où l'on reconnaît l'*expression régulière* `regexp`. Une expression régulière est formée à partir de lettres de l'alphabet (chaque lettre constituant un *élément*) et des symboles :
- $*$ qui indique que l'élément précédent apparaît un nombre quelconque de fois
- $+$ qui indique que l'élément précédent apparaît au moins une fois
- (et) qui entourent une expression régulière pour former un élément
- | qui indique une disjonction entre deux expressions régulières (on reconnaît l'une ou l'autre)
- ? qui indique que l'élément précédent est facultatif (i.e. il apparaît au plus une fois).

Une expression régulière représente donc sous forme condensée un ensemble de mots (c'est-à-dire un *langage*).
Soit tutu le fichier contenant :
caa
cbbab
cabab
caaabba

1. Que répond la commande `grep -E 'regexp' tutu` lorsque `regexp` vaut respectivement :
   - ca^+
   - $c(ab)^+$
   - $ca^+(a^*b^*)$
   - $(aa^+)|(bb^+)$
   - $baa?b$
   - a^*

   et pourquoi ?
2. Décrire sous forme ensembliste le langage correspondant aux expressions régulières précédentes.

(a) Automate $\mathcal{A}_1$ (b) Automate $\mathcal{A}_2$

Exercice 3 : Calcul par un automate.
L'alphabet considéré est $\{a, b\}$.

1. Dans quel état se trouve l'automate $\mathcal{A}_1$ après lecture des mots a, ab, abb, $abba$? Après lecture du mot ε ?
2. Lesquels de ces mots sont reconnus par l'automate $\mathcal{A}_1$?
3. Que se passe-t-il quand on donne le mot aab à lire à l'automate $\mathcal{A}_1$?
4. Les mots aba^2b, a^2ba^2b, ab^4 et b^3a^2 sont-ils reconnus par l'automate $\mathcal{A}_1$?
5. Décrire les mots reconnus par l'automate $\mathcal{A}_1$.
6. Après lecture du mot b^3a^2, dans quel état se trouve l'automate $\mathcal{A}_2$?
7. Y a-t-il des mots que l'automate $\mathcal{A}_2$ ne peut pas lire jusqu'au bout ?
8. S'il n'a lu aucun a, dans quel état se trouve l'automate $\mathcal{A}_2$?
9. Dans quels cas l'automate $\mathcal{A}_2$ se trouve-t-il dans l'état 1 ?
10. Dans quels cas arrive-t-il à l'état final 2 ? Quels mots reconnaît-il ?

Exercice 4 : Commutation
Soient u et v deux mots. On dit que u et v *commutent* si $u \cdot v = v \cdot u$.
Montrer que u et v commutent si et seulement s'il existe un mot w et deux entiers positifs ou nuls m et n tels que $u = w^m$ et $v = w^n$. Pour le sens $\Rightarrow$, on pourra procéder par récurrence sur $|u| + |v|$.

Exercice 5 : Conjugaison
Deux mots u et v sont dits *conjugués* s'il existe deux mots w_1 et w_2 tels que $u = w_1 \cdot w_2$ et $v = w_2 \cdot w_1$. 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 $u \cdot w = w \cdot v$.