Module Langages Formels
TD 1 : Mots et langages

Exercice 1 `Mots multiplicativement dépendants`

Deux mots u et v sont dits **multiplicativement dépendants** s'ils sont puissances d'un même troisième, c'est à dire s'il existe un mot w et deux entiers m et n tels que

$$u = w^n \text{ et } v = w^m$$

Deux mots u et v sont dits **commutatifs** si $uv = vu$.

1.1. Donner un exemple de deux mots commutatifs de longueur supérieure à 2

1.2. Prouver la proposition suivante :

> **Proposition** :
> Deux mots u et v commutent si et seulement si ils sont multiplicativement dépendants.

Exercice 2 `Codes`

On appelle **code** sur un alphabet Σ tout langage X sur Σ tel que $x_1x_2\dots x_p = y_1y_2\dots y_q$ et $x_i \in X$ pour tout i et $y_j \in X$ pour tout j entraînent $p = q$ et $x_i = y_i$ pour tout i. Dire que X est un code revient donc à dire que tout élément de X^* se factorise de manière unique sur X.

2.1. Les langages suivants sont-ils des codes ?
- $X_1 = \{ab, baa, abba, aabaa\}$
- $X_2 = \{b, ab, baa, abaa, aaaa\}$
- $X_3 = \{aa, ab, aab, bba\}$
- $X_4 = \{a, ba, bba, baab\}$

2.2. Soit u un mot de Σ^*, montrer que la partie $\{u\}$ est un code si et seulement si $u \neq \epsilon$.

2.3. Soient u et v deux mots distincts de Σ^*, montrer que la partie $\{u, v\}$ est un code si et seulement si u et v ne commutent pas.

2.4. Soit X une partie de Σ^* ne contenant pas ϵ et telle qu'aucun mot de X ne soit préfixe propre d'un autre mot de X. Montrer que X est un code (un tel code est appelé code préfixe).

Exercice 3 `Résiduels`

Soit L un langage sur un alphabet Σ, et soit u un mot sur Σ. On appelle **résiduel à gauche** de L par rapport à u, et on note $u^{-1}L$ l'ensemble des mots v sur Σ tels que $uv \in L$.

3.1. Calculer le résiduel de L par rapport à tout mot u sur $\Sigma = \{a, b\}$ dans les exemples suivants :
- $L = a^*b^*$
- $L' = \{a^n b^n \mid n \geq 0\}$

3.2. Si x est une lettre de Σ, que valent $x^{-1}(L_1 \cup L_2)$, $x^{-1}(L_1 L_2)$ et $x^{-1}L_1^*$ où L_1 et L_2 sont deux langages sur X ?

Exercice 4 `Mots de lyndon`

Soit Σ un alphabet ayant au moins deux lettres. Deux mots u et v sur Σ sont dits **conjugués** s'il existe des mots s et t sur Σ tels que $u = st$ et $v = ts$. Ils sont **conjugués propres** lorsque ni s ni t ne sont vides.

4.1. Vérifier que la relation $\mathcal{R}$ définie sur Σ^* par $u \mathcal{R} v$ si u et v sont conjugués est une relation d'équivalence.

On suppose que Σ est totalement ordonné par une relation $\leq$.

On munit Σ^* de l'ordre total lexicographique en posant $u \leq v$ si l'une des clauses suivantes est satisfaite :
- u est un préfixe de v ;
- $u = xay$, $v = xbz$ avec $a, b \in \Sigma$, $a < b$ et $x, y, z \in \Sigma^*$.

4.2. Soient $u, v \in \Sigma^*$, tels que $u < v$ et u n'est pas un préfixe de v. Montrer que

$$\forall w, z \in \Sigma^*, uw < vz.$$

On dit d'un mot qu'il est **primitif** s'il ne peut s'écrire sous la forme u^k avec $k \geq 2$.

4.3. Soit $u \in \Sigma^+$. Montrer que les assertions suivantes sont équivalentes :

(i). u est un mot primitif et il est le plus petit de sa classe de conjugués.

(ii). u est strictement inférieur à chacun de ses conjugués propres.

(iii). u est strictement inférieur à chacun de ses suffixes propres.

Un mot satisfaisant l'une de ces propriétés est appelé **mot de Lyndon**. On note L l'ensemble des mots de Lyndon sur Σ.

4.4. Soient u et v deux mots de Lyndon. Montrer que si $u < v$, alors soit u est préfixe de v, soit u n'est pas facteur de v (i.e. u n'apparait pas dans v).

4.5. Soient u et v deux mots de Lyndon , montrer que si $u < v$ alors uv est aussi un mot de Lyndon.

4.6. Soit $w \in L - \Sigma$, si $w = uv$ où v est le plus long suffixe de w appartenant à L, prouver que $u \in L$.

4.7. Montrer que tout mot de Σ^+ s'écrit sous la forme $u = u_1 u_2 \ldots u_n$ où les u_i sont dans L et $u_1 \geq u_2 \geq \ldots \geq u_n$.

4.8. Soit $u = u_1 u_2 \ldots u_n$ la factorisation de $u \in \Sigma^+$ comme produit décroissant de mots de Lyndon, prouver que u_n est le plus petit suffixe de u pour l'ordre lexicographique.

4.9. En déduire que la décomposition de Lyndon est unique.