DM2 : Corrigé

1. $KK^\star L + L = K^+L + L = (K^+ + \varepsilon)L = K^\star L$ donc $K^\star L$ est bien solution de l'équation (E).

2. Soit X un langage solution de l'équation (E).

   Montrons par récurrence sur $n \in \mathbb{N}$ la propriété $P(n)$ suivante : $K^nL \subset X$. On a $K^0L = L \subset KX + L = X$ puisque X est solution de (E), donc on a bien $P(0)$. De plus, si $n \in \mathbb{N}$ vérifie la propriété $P(n)$, on a :

   $$K^{n+1}L = K(K^nL) \subset KX \subset KX + L = X$$

   la première inclusion étant garantie par l'hypothèse de récurrence. On a donc bien $P(n+1)$.

   On déduit de ce résultat que X contient le langage $\sum\limits_{n\in\mathbb{N}} K^nL = \left(\sum\limits_{n\in\mathbb{N}} K^n\right) L = K^\star L$.

3. Soit X un langage solution de (E) et supposons par l'absurde que $X \setminus K^\star L$ est non vide.

   Alors $\{|m| \,|\, m \in X \setminus K^\star L\}$ est un ensemble non vide et minoré donc admet un élément minimal et il fait donc sens de considérer l'un des mots m de $X \setminus K^\star L$ de longueur minimale. Par définition, $m \in X = KX + L$ donc $m \in KX$ ou $m \in L$. Mais cette deuxième possibilité est exclue car dans ces conditions, on aurait $m = \varepsilon m \in K^\star L$, ce qui n'est pas.

   Par conséquent, $m \in KX$ et il existe $k \in K$ et $m' \in X$ tel que $m = km'$. On a d'une part $m' \notin K^\star L$ sinon on aurait $m \in K^\star L$ aussi ce qui est une contradiction. D'autre part, $\varepsilon \notin K$ donc $k \neq \varepsilon$ et donc $|m'| = |m| - |k| < |m|$. Ainsi m' est un élément de $X \setminus K^\star L$ dont la taille est strictement plus petite que celle minimale pour un élément de cet ensemble et c'est la contradiction recherchée.

4. La question 1 montre que $K^\star L$ est solution de $X = KX + L$, d'où l'existence. De plus si, X est solution de cette équation, $K^\star L \subset X$ d'après 2 et $K^\star L \subset X$ d'après 3, donc toute solution X de (E) vérifie $X = K^\star L$ ce qui garantit l'unicité demandée.

5. $\Sigma^\star$ est en général différent de $K^\star L$ et est solution de (E) dès que $\varepsilon \in K$. En effet, dans ces conditions, $\Sigma^\star \subset K\Sigma^\star \subset K\Sigma^\star + L$ d'une part et $K\Sigma^\star + L \subset \Sigma^\star$ (tout langage et inclus dans $\Sigma^\star$) d'autre part ce qui garantit que $\Sigma^\star = K\Sigma^\star L$. L'unicité de la question 4 ne tient donc plus.

6. Un mot de L_1 est soit vide, soit commence par a, soit commence par b. Dans le premier cas, il vaut ε, dans le second, il est égal à la lettre a concaténée à un mot ayant un nombre pair de b, dans le dernier, il est égal à la lettre b concaténée à un mot ayant un nombre impair de b. D'où $L_1 = \varepsilon + aL_1 + bL_2$.

   Un raisonnement similaire donne la seconde équation du système.

7. L'équation (2) se réécrit $L_2 = KL_2 + L$ avec $K = \{a\}$ et $L = bL_1$. Comme $\varepsilon \notin K$, on applique le lemme d'Arden (question 4), qui nous assure que $L_2 = a^\star bL_1$. On substitue cette expression dans (1) et on obtient :

   $$L_1 = aL_1 + ba^\star bL_1 + \varepsilon = \underbrace{(a + ba^\star b)}_{\text{ne contient pas } \varepsilon} L_1 + \varepsilon$$

   On réapplique le lemme d'Arden et on obtient $\begin{cases} L_1 = (a + ba^\star b)^\star \varepsilon = (a + ba^\star b)^\star \\ L_2 = a^\star bL_1 = a^\star b(a + ba^\star b)^\star \end{cases}$

8. A la place d'appliquer le lemme d'Arden sur l'équation (2) puis substituer L_2 dans (1), on applique le lemme d'Arden à (1) puis on substitue l'expression obtenue pour L_1 dans (2) :

   Comme $\varepsilon \notin \{a\}$, on peut appliquer le lemme d'Arden à (1) et ainsi obtenir $L_1 = a^\star(bL_2 + \varepsilon)$. Puis :

   $$L_2 = aL_2 + bL_1 = aL_2 + ba^\star(bL_2 + \varepsilon) = \underbrace{(a + ba^\star b)}_{\text{ne contient pas } \varepsilon} L_2 + ba^\star$$

   En appliquant le lemme d'Arden à l'équation précédente, on obtient : $\begin{cases} L_2 = (a + ba^\star b)^\star ba^\star \\ L_1 = a^\star b(a + ba^\star b)^\star ba^\star + a^\star \end{cases}$

   Les unicités garanties par le lemme d'Arden assurent donc que les deux expressions rationnelles de l'énoncé sont bien équivalentes, et leur sémantique est L_2 dans les deux cas.

9. a) On remarque que $L_i = \sum\limits_{(q_i,a_j,q_j)\in\delta} a_jL_j$ si q_i n'est pas final et est égal à ce même langage $+\ \varepsilon$ sinon. On en déduit que les langages L_0, L_1, L_2 sont liés par le système suivant :

$$\begin{cases} L_0 = \varepsilon + aL_1 & (1) \\ L_1 = aL_2 + bL_1 & (2) \\ L_2 = \varepsilon + (a+b)L_2 & (3) \end{cases}$$

b) On applique le lemme d'Arden (et on peut) à l'équation (3) et il vient $L_2 = (a+b)^\star = \Sigma^\star$. En substituant dans (2), on obtient $L_1 = bL_1 + a\Sigma^\star$ et le lemme d'Arden donne $L_1 = b^\star a\Sigma^\star$. On en déduit en substituant dans (1) que le langage reconnu par cet automate est $\varepsilon + ab^\star a(a+b)^\star$.

10. On reprend les notations de la question précédente : les langages L_0, L_1, L_2 associés à cet automates sont solutions du système d'équations suivant :

$$\begin{cases} L_0 = aL_0 + aL_1 & (1) \\ L_1 = aL_1 + bL_0 + \varepsilon & (2) \\ L_2 = bL_1 & (3) \end{cases}$$

Le lemme d'Arden s'applique à (1) et assure que $L_0 = a^\star aL_1 = a^+L_1$. En substituant dans (2), on obtient $L_1 = (a+ba^+)L_1+\varepsilon$, d'où on déduit que $L_1 = (a+ba^+)^\star$ puis que $L_2 = b(a+ba^+)^\star$ et que $L_0 = a^+(a+ba^+)^\star$. D'après les définitions de L_0, L_1, L_2, le langage reconnu par l'automate de cette question est $L_0+L_2 = (a^++b)(a+ba^+)^\star$.

11. Pour $n = 1$, montrer la première partie du théorème revient à montrer le lemme d'Arden ce qui a été fait en question 4. De plus, si K et L sont rationnels, $K^\star L$ aussi donc la deuxième partie est aussi avérée.

Si $n \in \mathbb{N}^\star$ satisfait l'hypothèse de récurrence, considérons un système (S) satisfaisant les hypothèses du théorème à $(n+1)$ équations et inconnues. Observons la dernière équation de ce système :

$$X_n = K_{n,n}X_n + \underbrace{\sum_{j=0}^{n-1} K_{n,j}X_j + L_n}_{=L}$$

Comme $\varepsilon \notin K_{n,n}$ par hypothèse, le lemme d'Arden assure que $X_n = K_{n,n}^\star L$. On reporte ce résultat dans les n équations restantes et on obtient ainsi un nouveau système dont les équations sont :

$$X_i = \sum_{j=0}^{n-1} K'_{i,j}X_j + L'_i \text{ pour tout } i \in [\![0, n-1]\!]$$

avec $K'_{i,j} = K_{i,j} + K_{i,n}K_{n,n}^\star K_{n,j}$ et $L'_i = L_i + K_{i,n}K_{n,n}^\star L_n$. Comme ε n'appartient à aucun des $K_{i,j}$ (par hypothèse), ε n'appartient à aucun des $K'_{i,j}$ et donc ce nouveau système strictement plus petit vérifie les hypothèses de notre propriété récurrente. Ainsi, d'après l'hypothèse de récurrence, ce système possède une unique solution $(X_0, X_1, ..., X_{n-1})$ que l'on complète par X_n pour obtenir une unique solution à (S).

De plus, si tous les langages intervenant dans (S) sont rationnels, par hypothèse de récurrence, $X_0, ..., X_{n-1}$ sont rationnels. Mézalors, X_n est lui aussi rationnel puisque c'est la concaténation de l'étoile d'un langage rationnel et du langage L qui est rationnel en tant que somme finie de concaténations de langages rationnels.

12. Soit L un langage reconnaissable. Alors il existe un automate A qui reconnaît L et on peut loisiblement supposer qu'il n'a qu'un seul état initial q_0 et ne fait pas intervenir d'ε-transitions. On note $q_1, ..., q_{n-1}$ les autres états de A.

Pour tout $i \in [\![0, n-1]\!]$, on introduit L_i le langage des mots dont la lecture dans l'état i mène à un état final de A (comme dans la partie 3). On introduit également les langages $B_i = \begin{cases} \{\varepsilon\} \text{ si } q_i \text{ est final} \\ \emptyset \text{ sinon} \end{cases}$.

Alors $(L_0, L_1, ...L_{n-1})$ est solution du système d'équations (S) suivant :

$$X_i = \sum_{j=0}^{n-1} A_{i,j}X_j + B_i \text{ pour tout } i \in [\![0, n-1]\!]$$

avec $A_{i,j} = \{a \in \Sigma \,|\, \delta(q_i, a) = q_j\}$. Or, pour tous $i, j \in [\![0, n-1]\!]$, B_i est rationnel et $A_{i,j}$ est rationnel et ne contient pas ε (puisque que ce langage est soit vide, soit est un ensemble fini de lettres de Σ). Le théorème de la question 10 s'applique donc et on en déduit que l'unique solution de (S), qu'on sait par ailleurs être $(L_0, ..., L_{n-1})$, voit ces composantes être des langages rationnels. En particulier, L_0 est rationnel et par définition de ce langage, $L_0 = L$ ce qui conclut.

Remarque : Ce DM montre plusieurs usages du lemme d'Arden : il permet de montrer l'équivalence d'expressions rationnelles, de calculer le langage reconnu par un automate et au passage de montrer la rationnalité des langages reconnus de manière constructive.