Algorithme CYK & Lemme de la pompe

Corrigé partiel de la feuille de travaux dirigés n^{o}10

1.

		0	1	0	2	0	1	1
j=1		Z	U	Z	S	Z	U	U
j=2	k=1	-	-	-	W	-	-	
j=3	k=1	-	-	S	-	-		
	k=2	-	-	-	-	-		
		-	-	S	-	-		
j=4	k=1	-	-	-	-			
	k=2	-	-	-	-			
	k=3	-	-	V	-			
		-	-	V	-			
j=5	k=1	-	S	-				
	k=2	-	-	-				
	k=3	-	-	-				
	k=4	-	-	-				
		-	S	-				
j=6	k=1	-	-					
	k=2	-	-					
	k=3	-	-					
	k=4	-	-					
	k=5	-	V					
		-	-					
j=7	k=1	-						
	k=2	-						
	k=3	-						
	k=4	-						
	k=5	-						
	k=6	-						
		-						

D'où : $0102011 \notin L(G)$.

Remarque : Ce résultat n'est pas surprenant ; la grammaire engendre $L(G) = \{w2w^{-1} | w \in (0+1)^\star\}$.
Ainsi, les variables désignent :

— S - les palindromes impairs sur $\{0, 1\}$ avec 2 comme séparateur de milieu.
— Y - les palindromes impairs sur $\{0, 1\}$ avec 2 comme séparateur de milieu, suivis d'un 1.
— V - les palindromes impairs sur $\{0, 1\}$ avec 2 comme séparateur de milieu, suivis d'un 0.
— Z - 0 (comme zéro).
— U - 1 (comme un).

On peut vérifier ainsi les résultats du tableau de l'algorithme CYK.

2.

		0	1	2	1	1
j=1		L,S,Y,X,Z	L,S,U,V,W	L,D,W,X	L,S,U,V,W	L,S,U,V,W
j=2	k=1	L,S,V,W	L,S,W,X	L,S,W	L,S,V,W	
j=3	k=1	L,S	L	L,S		
	k=2	W,X	V,W	V,W		
		L,S,W,X	L,S,V,W	L,S,V,W		
j=4	k=1	L,S	L,S			
	k=2	-	-			
	k=3	V,W	W			
		L,S,V,W	L,S,W			
j=5	k=1	L,S				
	k=2	-				
	k=3	-				
	k=4	V,W				
		L,S,V,W				

D'où : $01211 \in L(G)$.

Remarque : La grammaire engendre le complémentaire du langage de l'exercice précédent.
Ainsi, les variables désignent :
- S - les mots non palindromes impairs sur $\{0, 1\}$ avec 2 comme séparateur de milieu.
- Y - les mots de S, suivis d'un 1.
- V - les mots de S, suivis d'un 0.
- W - mots quelconqes, qui ne se terminent pas par un 0.
- X - mots quelconqes, qui ne se terminent pas par un 1.
- L - mots quelconqes, non vides.
- Z - 0 (comme zéro).
- U - 1 (comme un).
- D - 2 (comme deux).

On peut vérifier ainsi les résultats du tableau de l'algorithme CYK.

3.
a) Supposons $L_1 = \{a^i b^j c^k \mid i < j < k\}$ algébrique, donc il existe n du lemme. Soit $z = a^n b^{n+1} c^{n+2}$. Par le lemme z s'écrit sous la forme $uvwxy$. Soit p (resp. q et r) le nombre de a (resp. b et c) dans vx.
Fait : $p = 0$ ou $r = 0$ (à cause de la longueur de vwx).
On distingue les quatre cas suivants :
- $q \neq 0, p = 0$: $|uv^0wx^0y|_a = n$ et $|uv^0wx^0y|_b \leq n$ donc $uv^0wx^0y \notin L_1$.
- $q \neq 0, p \neq 0$: $r = 0$ et donc $|uv^2wx^2y|_c = n + 2$ et $|uv^2wx^2y|_b \geq n + 2$ donc $uv^2wx^2y \notin L_1$.
- $q = 0, p \neq 0$: $|uv^2wx^2y|_a \geq n + 1$ et $|uv^2wx^2y|_b = n + 1$ donc $uv^2wx^2y \notin L_1$.
- $q = 0$, $p = 0$: $r \neq 0$ (sinon $vx = \varepsilon$) donc $|uv^0wx^0y|_c \leq n + 1$ et $|uv^0wx^0y|_b = n + 1$ donc $uv^0wx^0y \notin L_1$.

Donc L_1 n'est pas algébrique.
Autre preuve. Supposons $L_1 = \{a^i b^j c^k \mid i < j < k\}$ algébrique, donc il existe n du lemme de l'étoile. Soit $z = a^n b^{n+1} c^{n+2}$. Par le lemme z s'écrit sous la forme $uvwxy$. Soit p (resp. q et r) le nombre de a (resp. b et c) dans vx. $|uv^iwx^iy|_a = n + (i-1)p$, $|uv^iwx^iy|_b = n + 1 + (i-1)q$ et $|uv^iwx^iy|_c = n + 2 + (i-1)r$. Pour $i = 0$, on a $p \geq q \geq r$ et, pour $i = 2$, $p \leq q \leq r$. On a donc $p = q = r$. Mais on doit avoir $p = 0$ ou $q = 0$, sinon $|vwx| \geq n + 3$. Donc $p = q = r = 0 \Rightarrow |vx| = 0$, une contradiction.

b) Supposons $L_2 = \{a^i b^j \mid j = i^2\}$ algébrique, donc il existe n du lemme de l'étoile. Soit $z = a^n b^{n^2}$. Par le lemme z s'écrit sous la forme $uvwxy$. Soit k (resp. l) le nombre de a (resp. b) dans vx. Comme on doit avoir $uv^iwx^iy \in L_2$, on doit avoir $\forall i \geq 0 : ((n-k) + ik)^2 = (n^2 - l) + il$, ce qui implique $\forall i \geq 0$:

$2kn + (i-1)k^2 = l$ ce qui est impossible. Donc L_2 n'est pas algébrique.

c) Supposons $L_3 = \{a^k b^k c^l | k \leq l \leq 2k\}$ algébrique, donc il existe n du lemme de l'étoile. Soit $z = a^n b^n c^n$. Par le lemme z s'écrit sous la forme $uvwxy$. Soit p (resp. q et r) le nombre de a (resp. b et c) dans vx.
Fait : $p = 0$ ou $r = 0$ (à cause de la longueur de vwx).

— $r = 0$: $|uv^2wx^2y|_a > n$ ou $|uv^2wx^2y|_b > n$ alors que $|uv^2wx^2y|_c = n$ donc $uv^2wx^2y \notin L_3$.
— $r \neq 0$: donc $p = 0$ et ainsi $|uv^0wx^0y|_a = n$ et $|uv^0wx^0y|_c < n$ donc $uv^0wx^0y \notin L_3$.

Donc L_1 n'est pas algébrique.
d) $L_5 = \bar{L_1} \cap a^\star b^\star c^\star$. On observe que $L_5 = \{a^i b^j c^k \mid i \geq j \text{ ou } j \geq k\} = \{a^i b^j \mid i \geq j\}c^\star \cup a^\star\{b^j c^k \mid j \geq k\}$ ce qui prouve (c.f. propriétés de clôture) que ce langage **est algébrique** (l'intersection avec $a^\star b^\star c^\star$ ne sert que pour le tri des lettres).

e) Supposons $L_6 = \{ww^{-1}w \mid w \in (a+b)^\star\}$ algébrique. Soit R le langage rationnel $a^\star b^\star a^\star b^\star$. Soit $L_6' = L_6 \cap R$, i.e. le langage $\{a^i b^{2j} a^{2i} b^j \mid i \geq 0,\ j \geq 0\}$. On montre donc que L_6' n'est pas algébrique. Supposons L_6' algébrique, donc il existe n du lemme de l'étoile. Soit $z = a^n b^{2n} a^{2n} b^n$. Par le lemme, $z = uvwxy$ ou $z = ABCD$ avec $A = a^n$ $B = b^{2n}$ $C = a^{2n}$ $D = b^n$.

— vx contient un a : comme $|vwx| \leq n$, vwx ne peut pas intersecter à la fois A et C. Ainsi $uv^0wx^0y \notin L_6'$ car on supprime au moins un a du coté A ou du coté C.
— vx contient un b : comme $|vwx| \leq n$, vwx ne peut pas intersecter à la fois B et D. Ainsi $uv^0wx^0y \notin L_6'$ car on supprime au moins un b du coté B ou du coté D.

Donc L_6' n'est pas algébrique et L_6 n'est pas algébrique.

f) $L_7 = \{w \in (a+b+c)^\star : |w|_a = |w|_b = |w|_c\}$. Soit $R = a^\star b^\star c^\star$ rationnel. $L_7 \cap R = \{a^n b^n c^n : n \geq 0\}$ est non algébrique. Donc L_7 n'est pas algébrique.