

Session : 2

EPREUVE : Langages Formels et Compilation

Durée : 2 h 00 – (documents papiers - sauf livres - autorisés ; appareils électroniques interdits)

Les exercices sont indépendants. Le barème est donné à titre indicatif.

Exercice 1 – 2.5 points

Soit la grammaire sous forme de Chomsky suivante :

$G1 = (\{S, X, Y, U, D, T\}, \{1, 2, 3\}, S, \{S \rightarrow UX|DY|TX, U \rightarrow 1, D \rightarrow 2, T \rightarrow 3, X \rightarrow 2|DY, Y \rightarrow 1|3|UX|TX\})$

- 1) Analysez le mot 23212 en utilisant l'algorithme de Cocke-Kasami-Younger
- 2) D'après ce que vous avez obtenu, quels sous-mots de 23212 sont reconnus ?

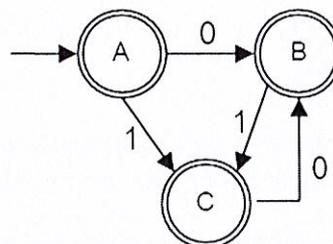
Exercice 2 – 3.5 points

Soit la grammaire $G2 = (\{S, A\}, \{0, 1, \&\}, S, \{S \rightarrow 1|1\&A, A \rightarrow 1|0|A\&A\})$

- 1) Montrez que $G2$ est ambiguë.
- 2) Transformez $G2$ en une grammaire équivalente sous forme normale de Greibach. Soit $G2'$ la grammaire ainsi obtenue.
- 3) Construisez l'automate à pile associé à $G2'$.

Exercice 3 – 7 points

Soit l'automate fini AF3 suivant :



- 1) AF3 est-il déterministe ou non déterministe ? Justifiez votre réponse.
- 2) Donnez les étapes de l'analyse du mot 101010101 par AF3.
- 3) Expliquez, en français, et le plus précisément possible, comment sont constitués les mots du langage reconnu par AF3.
- 4) Calculez l'expression régulière associée à AF3.
- 5) Quel est le type du langage reconnu par AF3 ?
- 6) Donnez la grammaire $G3$ qui engendre les mots du langage reconnu par AF3.
- 7) Expliquez pourquoi il n'est pas possible de trouver une grammaire équivalente à $G3$ sous forme normale de Greibach ou sous forme normale de Chomsky.

Exercice 4 – 3 points

Soit le langage L4 constitué des mots ayant les caractéristiques suivantes :

- Les caractères constituant les mots appartiennent à l'ensemble {a, b}
- La longueur minimale des mots est 1
- Le premier caractère des mots peut être a ou b
- Le dernier caractère des mots peut être a ou b, indépendamment du premier caractère
- Dans un mot, on ne peut pas avoir 2 a qui se suivent et les b vont toujours par 2 exactement (alternance de « a » et de « bb », on ne peut pas non plus avoir deux « bb » de suite).

Exemples de mots corrects : a, bb, abb, bba, abba, bbabb, abbabb, bbabba

Exemples de mots incorrects : aa, b, aba, bab, abbb, abbbb, baa

Donnez un automate à mémoire linéairement bornée qui permet de reconnaître les mots de ce langage.

Exercice 5 – 4 points (à rendre sur une feuille à part)

- 1) Donnez l'expression régulière représentant les mots de longueur supérieure ou égale à 1 comportant n'importe quel caractère numérique ou alphabétique, et des espaces.
- 2) Écrivez les programmes Lex et Yacc qui permettent de calculer le nombre de lettres total des mots reconnus. Dans la chaîne d'entrée on ne doit trouver que des mots correspondant à l'expression de la question 1, séparés par des #. La chaîne se termine par \n. Attention, il faut respecter :
 - a. pour la partie Lex, donnez des définitions régulières des mots reconnus, la ou les règles Lex, l'envoi des informations nécessaires au programme Yacc. Le programme n'effectue pas de calcul, la somme des nombres de caractères se fera dans les règles de Yacc. Il faut afficher une erreur si un symbole n'est pas reconnu ;
 - b. pour la partie Yacc, donnez le ou les token, le symbole de départ, les règles de grammaire et les expressions du calcul du nombre total de caractères. Il faut afficher une erreur si la chaîne n'est pas conforme à la grammaire ;

Exemples :

abcd#efgh# abcd1234 #reconnu\n → chaîne valide, 3 mots reconnus, $4+4+10+7 = 25$ lettres

abcd#3efgh\n → 2 mots reconnus, 9 lettres

lreconnu\n → chaîne valide, un seul mot de 8 lettres

#\n → pas de mots à séparer, chaîne invalide

mot correct \n\n autre → chaîne invalide, deux \n à la suite, et mal placés