

Cartouche du document

Année : ING 1
 Matière : Théorie des langages
 Activité : Examen

Objectifs

cet examen porte sur :

- les grammaires régulières et hors contexte
- les automates finis déterministes ou non
- langages rationnels : les quotients gauches et les systèmes d'équations des automates
- langages algébriques : l'algorithme de cKY et les automates à pile
- les machines de Turing.

Tous les documents sont autorisés.

Durée : 1h30

les ordinateurs sont autorisés mais seulement en fonctionnement local.

Vous devez rendre des copies papiers.

Sommaire des exercices

1 - Plaques d'immatriculation

Corps des exercices

1 - Plaques d'immatriculation

Énoncé :

ci-joint un descriptif des règles de calcul des anciennes et des nouvelles (à partir de juin 2009) plaques d'immatriculation en France métropolitaine :

Anciennes : (source Wikipedia) composées d'un maximum de 8 caractères (autres que les espaces), elles sont découpées en trois parties séparées par des espaces, le premier espace pouvant être supprimé (Annexe I de l'arrêté du 1er juillet 1996) :

- 1^{ère} partie de un (deux depuis octobre 1976) à quatre chiffres
- 2^{ème} partie de une à trois lettres
- 3^{ème} partie de deux chiffres. Pour la corse, 2A et 2B (il reste cependant toujours quelques plaques 20).

la numérotation s'établit de la manière suivante : les chiffres de la première partie évoluent en premier et lorsqu'ils arrivent à leur valeur maximale (999 ou 9999 selon le système), la seconde partie composée de une à trois lettres est incrémentée, c'est-à-dire que la dernière lettre est remplacée par sa lettre suivante dans l'alphabet. On voit donc apparaître les numéros suivants (cas de la plupart des départements avec un passage à quatre chiffres à la série QA). la notation 00 correspond au numéro du département et n'influe pas sur la méthode de numérotation) :

- de 1 A 00 à 999 Z 00
- de 1 AA 00 à 999 PZ 00

- de 1 QA 00 à 9999 ZZ 00
- de 11 AAA 00 à 999 ZZZ 00.

Nouvelles : (source **Wikipedia**) chaque véhicule possède un numéro « à vie ».

ce numéro est constitué de 9 caractères dont :

- 1^{ère} partie : deux lettres
- séparation : un tiret "-"
- partie : trois chiffres
- séparation : un tiret "-"
- 3e partie : deux lettres.

Principe de numérotation :

- de AA-001-AA à AA-999-AA (les chiffres évoluent en premier)
- de AA-001-AB à AA-999-AZ (2e élément lettres de droite)
- de AA-001-BA à AA-999-ZZ (1er élément lettres de droite)
- de AB-001-AA à AZ-999-ZZ (2e élément lettres de gauche)
- de BA-001-AA à ZZ-999-ZZ (1er élément lettres de gauche).

le but de l'exercice est de détecter si une plaque est ancienne ou nouvelle (ou fausse).

Question 1)

Énoncé de la question

Barème 2 points

Proposer un découpage lexicale/syntaxique pour ce problème. On vous demande simplement de donner les terminaux et non terminaux des deux grammaires issues du découpage, sans écrire les règles, mais en précisant le rôle de chaque élément via une phrase descriptive.

Exemple :

- 1..9 : terminaux de la partie lexicale correspondant aux chiffres
- P1A : non terminal de la partie syntaxique correspondant à la 1^{ère} partie des anciennes plaques.
- ...

Solution de la question

les éléments terminaux de l'analyse lexicale sont les différents symboles utilisés pour constituer une plaque ancienne ou nouvelle.

- les chiffres { 0, ..., 9 } ;
- les lettres majuscules { A, ..., Z } ;
- le symbole -.

les non terminaux de l'analyse lexicale sont :

- un groupement de 1 chiffre noté 1CHIFFRE ;
- un groupement de 2 chiffres noté 2CHIFFRES ;
- un groupement de 3 chiffres noté 3CHIFFRES;
- un groupement de 4 chiffres noté 4CHIFFRES;
- un groupement de 1 lettre majuscule noté 1LETTRE;
- un groupement de 2 lettres majuscules noté 2LETTRES;
- un groupement de 3 lettres majuscules noté 3LETTRES;

- un groupement de 1 tiret noté TIRET;
- un groupement de 1 espace noté ESPACE;
- les deux groupements 2A et 2B espace notés DEPTCORSE;

les terminaux de l'analyse syntaxique sont :

- 1CHIFFRE ;
- 2CHIFFRES ;
- 3CHIFFRES;
- 4CHIFFRES;
- 1LETTRE;
- 2LETTRES;
- 3LETTRES;
- TIRET;
- DEPTCORSE;
- ESPAcE .

les non terminaux de l'analyse syntaxique sont :

- Plaque (axiome) ;
- PlaqueAncienne ;
- PlaqueNouvelle ;

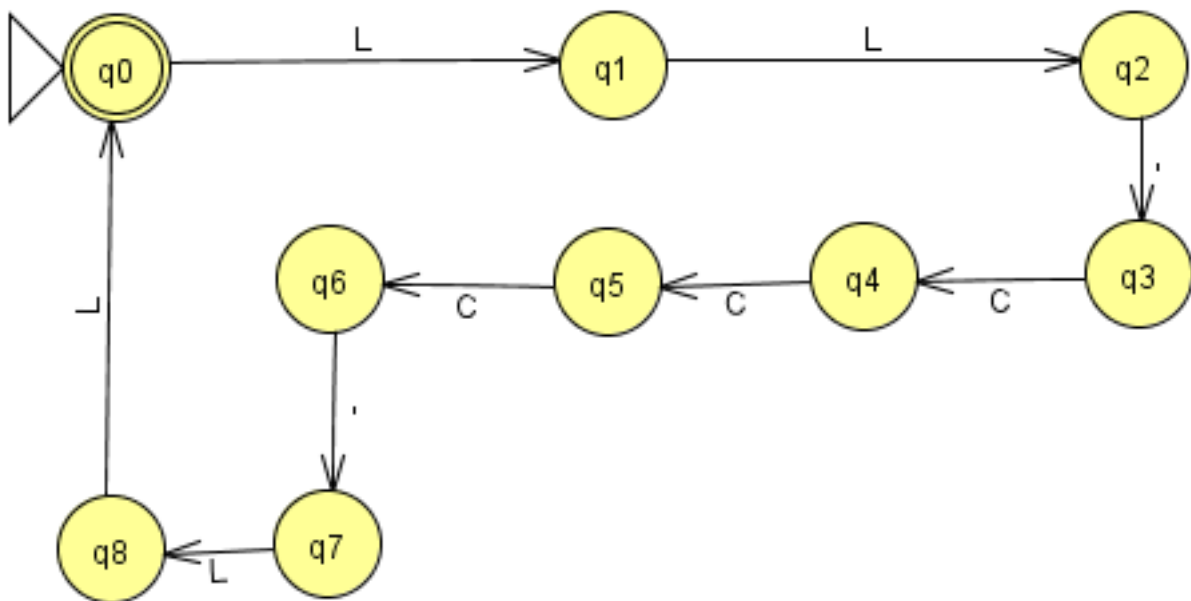
Question 2)

Énoncé de la question

Barème 2 points

construire un automate permettant de reconnaître une liste de nouvelles plaques à partir de l'alphabet constitué des chiffres $c = \{0..9\}$, des lettres $l = \{A..Z\}$ et du symbole $t = \{-\}$.

Solution de la question



Question 3)

Énoncé de la question

Barème 2 points

Appliquer la mise en équation et le lemme d'Arden afin de trouver l'expression régulière correspondant à une liste de nouvelles plaques.

Solution de la question

$$Q0 = l.Q1 + \varepsilon$$

$$Q1 = l.Q2$$

$$Q2 = -.Q3$$

$$Q3 = c.Q4$$

$$Q4 = c.Q5$$

$$Q5 = c.Q6$$

$$Q6 = -.Q7$$

$$Q7 = l.Q8$$

$$Q8 = l.Q0$$

En opérant des substitutions dans les sept dernières équations, on obtient $Q1 = (l-ccc-l).Q0$

en remplaçant dans la première équation, on obtient : $Q0 = ll-ccc-ll.Q0 + \varepsilon$.

On applique le lemme d'Arden ($X = A.X + B$ on prend $A = \{ll-ccc-ll\}$ et $B = \{\varepsilon\}$), on obtient $Q0 = (ll-ccc-ll)^*$.

Question 4)

Énoncé de la question

Barème 2 points

Appliquer la méthode des quotients gauches pour vérifier votre résultat.

Solution de la question

On pose $la_0 = (ll-ccc-ll)^*$.

$$la_1 = l^{-1} la_0 = (l-ccc-ll)(ll-ccc-ll)^*$$

On a un nouveau langage la_1 et la transition (la_0, l, la_1) .

$$la_2 = l^{-1} la_1 = (-ccc-ll)(ll-ccc-ll)^*$$

On a un nouveau langage la_2 et la transition (la_1, l, la_2) .

$$la_3 = -^{-1} la_2 = (ccc-ll)(ll-ccc-ll)^*$$

On a un nouveau langage la_3 et la transition (la_2, l, la_3) .

$$la_4 = c^{-1} la_3 = (cc-ll)(ll-ccc-ll)^*.$$

On a un nouveau langage la_4 et la transition (la_3, l, la_4) .

$$la_5 = c^{-1} la_4 = (c-ll)(ll-ccc-ll)^*.$$

On a un nouveau langage la_5 et la transition (la_4, l, la_5) .

$$la_6 = c^{-1} la_5 = (-ll)(ll-ccc-ll)^*.$$

On a un nouveau langage la_6 et la transition (la_5, l, la_6) .

$$la_7 = -^{-1} la_6 = (ll)(ll-ccc-ll)^*.$$

On a un nouveau langage la_7 et la transition (la_6, l, la_7) .

$$la_8 = l^{-1} la_7 = (l)(ll-ccc-ll)^*.$$

On a un nouveau langage la_8 et la transition (la_7, l, la_8) .

$$la_9 = l^{-1} la_8 = (ll-ccc-ll)^* = la_0.$$

On n'a pas de nouveau langage mais simplement la transition (la_9, l, la_0) .

Question 5)

Énoncé de la question

Barème 2 points

Ecrire une grammaire permettant de vérifier si une liste de plaques contient moins d'anciennes plaques que de nouvelles. On considère qu'une analyse a déjà été faite et que les listes sont triées sous la forme $a^* n^*$, où a désigne un terminal pour une ancienne plaque et n un terminal pour une nouvelle plaque.

Solution de la question

On définit :

- AN le non terminal qui représente les groupements de la forme $a^p b^p$;
- N le non terminal qui représente les groupements de la forme n^+
- S le non terminal qui représente les groupements de la forme $a^p b^q$ avec $p < q$;

$$AN \rightarrow \varepsilon$$

$$AN \rightarrow a AN n$$

$$N \rightarrow n$$

$$N \rightarrow N n$$

$$S \rightarrow AN N$$