

Examen

Théorie des langages (INF105)

Pierre SENELLART

pierre.senellart@telecom-paristech.fr

L'examen du module INF105 dure une heure et demie. Tous les documents sont autorisés, mais calculatrices, ordinateurs et objets communicants sont interdits. L'énoncé comporte un unique problème noté sur 24 (les points au-dessus de 20 seront tronqués). Il est précédé d'un rappel de notions mathématiques utiles.

Rappels

On rappelle qu'une *relation d'équivalence sur un ensemble* X est une relation \equiv vérifiant :

- (réflexivité) $\forall x \in X, x \equiv x$;
- (symétrie) $\forall x, y \in X, x \equiv y \Rightarrow y \equiv x$;
- (transitivité) $\forall x, y, z \in X$, si $x \equiv y$ et $y \equiv z$ alors $x \equiv z$.

La *classe d'équivalence pour \equiv* d'un élément $x \in X$ est l'ensemble $\{y \in X \mid y \equiv x\}$.

Étant donné une relation d'équivalence \equiv sur un ensemble X , le *quotient* de X par \equiv , noté $X_{/\equiv}$ est l'ensemble des classes d'équivalence pour \equiv des éléments de X .

Étant donné un mot $u = \alpha_1 \dots \alpha_n$, le *miroir* de u , noté u^R , est le mot $u^R = \alpha_n \dots \alpha_1$.

Problème : Théorème de Myhill–Nerode

Dans tout le problème, on considère l'alphabet $\Sigma = \{a, b\}$.

1. Dans cette question, on s'intéresse au langage L_0 défini par l'expression rationnelle $(a^*b) + (ab^*)$ (+ dénote ici la disjonction, parfois aussi notée \mid).
 - a) (1,5 points) En appliquant la construction de Thompson vue en cours, construire un automate non déterministe à transitions spontanées reconnaissant L_0 . Il est inutile de détailler la construction, mais l'automate donné doit être exactement le résultat de l'application de la construction de Thompson, sans simplification possible. *Indication* : l'automate de Thompson pour L_0 a 14 états.
 - b) (1,5 points) Éliminer les transitions spontanées de l'automate obtenu à la question précédente à l'aide d'un algorithme vu en cours. Il est inutile de détailler la construction.
 - c) (1,5 points) En appliquant une méthode du cours, déterminer l'automate obtenu à la question précédente. Il est inutile de détailler la construction. On s'assurera que le résultat est un automate déterministe *complet*.
 - d) (1,5 points) En déduire l'automate canonique A_0 reconnaissant le langage L_0 ; on détaillera la construction et on justifiera que l'automate obtenu est bien minimal.

Tourner la page.

Deux mots de Σ^* sont dits *syntactiquement équivalents à droite* par rapport à un langage L , ce que l'on note $u \sim_L v$, s'ils ont les mêmes continuations dans L :

$$\forall x \in \Sigma^*, ux \in L \Leftrightarrow vx \in L$$

2. (2 points) Montrer que, pour tout langage L , \sim_L est une relation d'équivalence sur Σ^* .

Pour un langage L donné, quand $\Sigma^*_{/\sim_L}$ est fini, on définit l'automate fini déterministe complet $A_{\sim_L} = (\Sigma, \Sigma^*_{/\sim_L}, q_0, F, \delta)$ de la manière suivante :

- q_0 est la classe d'équivalence pour \sim_L de ε ;
- F est l'ensemble des classes d'équivalence pour \sim_L des mots de L ;
- Si u est un mot avec pour classe d'équivalence x alors $\delta(x, a)$ est la classe d'équivalence de ua .

3. (2 points) Il y a exactement 6 classes d'équivalence de Σ^* par \sim_{L_0} . Donner une description (sous la forme d'une expression rationnelle ou de description en français) de chacune de ces classes d'équivalence. Aucune justification n'est exigée.
4. (2 points) Construire l'automate $A_{\sim_{L_0}}$.
5. (1 point) Montrer que si un langage L est tel que $\Sigma^*_{/\sim_L}$ est fini, alors L est rationnel.
6. (3 points) Montrer que si un langage L est rationnel, alors $\Sigma^*_{/\sim_L}$ est fini.

Les deux questions précédentes forment le *théorème de Myhill–Nerode*. Cette caractérisation permet en particulier de démontrer que certains langages ne sont pas rationnels, et est plus puissante que le lemme de pompage (qui n'est pas caractéristique des langages rationnels) : nous allons montrer qu'il existe des langages non rationnels vérifiant toutes les conséquences du lemme de pompage pour les langages rationnels.

7. On considère maintenant le langage $L_1 = \{ss^Rt \mid s, t \in \Sigma^*, s \neq \varepsilon\}$ des mots commençant par un palindrome (non nul) de longueur paire.
- a) (2 points) Proposer une grammaire hors-contexte pour L_1 . On ne demande pas de justification.
- b) (3 points) L'objectif de cette question est de montrer que L_1 vérifie l'ensemble des conséquences du lemme de pompage pour les langages rationnels donné dans le polycopié (Théorème 4.25). Plus précisément, montrer qu'il existe un entier k tel que pour tout mot x de L_1 plus long que k , x se factorise en uvw avec (i) $|v| \geq 1$; (ii) $|uv| \leq k$; et (iii) pour tout $i \geq 0$, $uv^i w \in L_1$. *Indication* : on pourra prendre $k = 4$.
- c) (3 points) En utilisant le théorème de Myhill–Nerode, montrer que L_1 n'est pas rationnel.