Langages formels, calculabilité et complexité

Examen du 4 mai 1998

durée 3 heures

Magistère MMFAI

Il va sans dire que la rigueur des raisonnements, la clarté des explications, mais aussi la qualité de la présentation influeront sensiblement sur la note.

1 Minimisation

Soit $\mathcal{A} = (Q, \delta, i, F)$ un automate déterministe dont tous les états sont accessibles à partir de l'état initial i.

a) On considère $\tilde{\mathcal{A}} = (Q, T, F, \{i\})$ l'automate miroir de $\mathcal{A}$ où

$$T = \{(q, a, p) \mid \delta(p, a) = q\}$$

Montrer que $\tilde{\mathcal{A}}$ reconnaît le miroir du langage reconnu par $\mathcal{A}$.

b) On note $\mathcal{B}$ le déterminisé de l'automate $\tilde{\mathcal{A}}$ (par la procédure de déterminisation classique) réduit aux états accessibles à partir de l'état initial F de l'automate $\mathcal{B}$. Montrer que $\mathcal{B}$ est minimal.

c) En déduire une procédure de minimisation d'un automate et comparer sa complexité à celle de la procédure de minimisation utilisant l'équivalence de Nérode.

2 Fonctions récursives

a) Soit $f : \mathbb{N} \longrightarrow \mathbb{N}$ une fonction primitive récursive. Montrer que la fonction

$$\begin{array}{rccl} g: & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & n & \longmapsto & \max\{f(k) \mid 0 \leq k \leq n\} \end{array}$$

est primitive récursive.

b) Soit $f : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ une fonction primitive récursive. Montrer que la fonction

$$\begin{array}{rccl} g: & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & n & \longmapsto & \max\{f(k, n-k) \mid 0 \leq k \leq n\} \end{array}$$

est primitive récursive.

3 MAX2SAT

Soient x, y, z, w quatre variables booléennes. On considère l'ensemble $\mathcal{E}(x, y, z, w)$ formé des 10 clauses suivantes :

$$(x), (y), (z), (w)$$
$$(\neg x \vee \neg y), (\neg y \vee \neg z), (\neg z \vee \neg x)$$
$$(x \vee \neg w), (y \vee \neg w), (z \vee \neg w)$$

a) Quel est le nombre maximum de clauses de $\mathcal{E}(x, y, z, w)$ satisfaisables simultanément ?

Le problème *MAX2SAT* est le suivant :

Données : Un ensemble $\mathcal{E}$ de clauses contenant chacune au plus deux littéraux et un entier k.

Question : Existe-t-il une assignation des variables de $\mathcal{E}$ qui permette de satisfaire au moins k clauses de $\mathcal{E}$?

b) Montrer que le problème *MAX2SAT* est dans la classe **NP**.

c) Montrer que le problème *3SAT* se réduit à *MAX2SAT*.
Indication : On pourra associer à une clause $C = \alpha \vee \beta \vee \gamma$ l'ensemble $R(C) = \mathcal{E}(\alpha, \beta, \gamma, w_C)$ où w_C est une nouvelle variable liée à la clause C.

d) Montrer que le problème *MAX2SAT* est **NP**-complet.

4 Segments rationnels infinis

On dit qu'un langage $L \subseteq \Sigma^*$ admet un segment rationnel infini (IRS) s'il existe un mot non vide $u \in \Sigma^+$ et deux mots $\alpha, \beta \in \Sigma^*$ tels que $\alpha u^* \beta \subseteq L$.

On considère le problème *IRS* :

Données : Un langage L d'une classe de langage $\mathcal{C}$ (rationnels, algébriques, ...).

Question : Est-ce que le langage L admet un segment rationnel infini ?

a) Montrer qu'un langage rationnel est sans segment rationnel infini si et seulement si il est fini. En déduire que le problème *IRS* est décidable pour les langages rationnels.

b) Montrer que le langage de Lukasiewicz Ł engendré par la grammaire

$$S \longrightarrow aSS + b$$

admet un segment rationnel infini. Montrer que Ł $\cap\, a^+b^+$ est sans segment rationnel infini.

c) Montrer que le langage L engendré par la grammaire

$$S \longrightarrow aSb + Sb + b$$

admet un segment rationnel infini. Montrer que, pour tout langage rationnel K, $L \cap K$ est soit fini soit sans segment rationnel infini.

On veut maintenant montrer que le problème *IRS* est indécidable pour les langages algébriques.

Etant donnés deux langages G et D sur l'alphabet Σ on définit le langage $L(G, D)$ sur l'alphabet $\Sigma \cup \{\sharp\}$ ($\sharp \notin \Sigma$) par :

$$L(G, D) = \{g_1\sharp \cdots \sharp g_k \sharp d_k \sharp \cdots \sharp d_1 \sharp \mid k \geq 1, g_1, \ldots, g_k \in G \text{ et } d_1, \ldots, d_k \in D\}$$

d) Montrer que si les langages G et D sont algébriques, il en est de même du langage $L(G, D)$.

e) Montrer que $L(G, D)$ est sans segment rationnel infini si et seulement si

- G et D sont sans segment rationnel infini, et
- $G \cap D = \emptyset$.

f) On considère n couples de mots $(u_1, v_1), \ldots, (u_n, v_n)$ sur l'alphabet Σ et n nouvelles lettres $x_1, \ldots, x_n \notin \Sigma$. Montrer que les langages G et D engendrés par les grammaires

- $S \longrightarrow u_1 S x_1 + \cdots + u_n S x_n + u_1 x_1 + \cdots + u_n x_n$ pour G, et
- $S \longrightarrow v_1 S x_1 + \cdots + v_n S x_n + v_1 x_1 + \cdots + v_n x_n$ pour D

sont sans segment rationnel infini.

g) Montrer que le problème *IRS* est indécidable pour les langages algébriques.

Une grammaire algébrique $G = (\Sigma, V, P)$ est dite simple si

- $P \subseteq V \times \Sigma V^*$, et
- pour tout couple $(x, a) \in \Sigma \times V$ il existe au plus une règle de la forme $x \longrightarrow am$.

Un langage est simple s'il peut être engendré par une grammaire simple.

On va établir que le problème *IRS* est décidable pour les langages simples.

h) Montrer qu'un langage L engendré par une grammaire simple est préfixe : si $u, uv \in L$ alors $v = \varepsilon$.

i) Soit $G = (\Sigma, V, P)$ une grammaire simple. Montrer que si une variable $S \in V$ engendre tous les mots du langage $\alpha u^* \beta \subseteq \Sigma^*$ alors il existe un mot $p \in V^*$ tel que $p \xrightarrow{*} u^k p$ pour un entier $k \geq 1$.

j) Soient $p \in V^+$ et $u \in \Sigma^+$. Montrer que si $p \xrightarrow{*} up$ alors $p = p_1 x p_2$ avec $p_1 \xrightarrow{*} u_1$, $x \xrightarrow{*} u_2 p_1 x$ et $u = u_1 u_2$.

k) Montrer que le langage engendré par une grammaire simple réduite $G = (\Sigma, V, P)$ d'axiome S admet un segment rationnel infini si et seulement si il existe $\alpha \in \Sigma^*$, $u \in \Sigma^+$, $x \in V$ et $m \in V^*$ tels que

$$S \xrightarrow{*} \alpha x m \quad \text{et} \quad x \xrightarrow{*} ux$$

l) En déduire que le problème *IRS* est décidable pour les langages simples.