Langages formels, calculabilité et complexité

Magistère MMFAI

Examen du 22 mai 2000
durée 3 heures

Les documents sont autorisés.

Toutes les réponses devront être correctement justifiées.

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.

Le barème sur 40 n'est donné qu'à titre indicatif et pourra être modifié.

1 Degré d'une grammaire (20 points)

Soit $G = (A,V,P,S)$ une grammaire. On appelle degré d'une dérivation $S \xrightarrow{*} u$ le nombre de règles différentes utilisées au cours de cette dérivation. Pour chaque entier p on définit

$$L(G,p) = \{u \in A^* \mid \text{ il existe une dérivation } S \xrightarrow{*} u \text{ de degré } p\}$$
$$L(G,\leq p) = \{u \in A^* \mid \text{ il existe une dérivation } S \xrightarrow{*} u \text{ de degré } \leq p\}.$$

a) Montrer que $L(G,\leq p)$ est algébrique.

b) On commence par traiter un exemple. La grammaire G est définie par les règles

$$\begin{aligned} S &\longrightarrow S_1S_2S_3 \\ S_1 &\longrightarrow aS_1b + ab \\ S_2 &\longrightarrow aS_2b^2 + ab^2 \\ S_3 &\longrightarrow aS_3b^3 + ab^3 \end{aligned}$$

Calculer les langages $L(G,p)$ pour $p > 0$.
Pour quelle valeur minimale de p a-t-on $L(G) = L(G,\leq p)$?

On revient au cas général et on veut montrer que le langage $L(G,p)$ est algébrique. On numérote les règles de G de 1 à n. On considère un nouvel alphabet $R = \{r_1,\dots,r_n\}$ disjoint de A et on considère la grammaire $H = (A \cup R,V,P',S)$ dont les règles sont définies comme suit : si $x \longrightarrow m$ est la i^{e} règle de G alors $x \longrightarrow r_i m$ est la i^{e} règle de H.

c) Montrer que pour tout p, le langage $K_p = \{u \in (A \cup R)^* \mid |\text{alph}(u) \cap R| = p\}$ est rationnel.

d) Montrer qu'il existe un morphisme $\varphi : (A \cup R)^* \to A^*$ tel que pour tout p, $L(G,p) = \varphi(L(H) \cap K_p)$.

e) Montrer que pour tout p, le langage $L(G,p)$ est algébrique.

On définit le degré d'une grammaire G comme le plus petit entier p tel que $L(G) = L(G,\leq p)$.

f) On suppose que G est une grammaire non ambiguë. Montrer que pour tout $u \in L(G)$ il existe un unique entier p tel que $u \in L(G,p)$. En déduire que $d(G)$ est le plus petit entier p tel que $L(G,q) = \emptyset$ pour tous $q > p$.
Montrer que si G est une grammaire non ambiguë alors son degré est calculable.

On veut maintenant montrer que le degré n'est pas calculable pour les grammaires linéaires. On fixe l'alphabet $A = \{a,b\}$ et une lettre $c \notin A$.

On rappelle que le problème de Post ci-dessous est indécidable.

Données : Des couples de mots $(u_1,v_1),\ldots,(u_k,v_k)$ sur un alphabet A.
Question : Existe-t-il $n \geq 1$ et $i_1,\ldots,i_n \in \{1,\ldots,k\}$ tels que $u_{i_1}\cdots u_{i_n} = v_{i_1}\cdots v_{i_n}$?

Soit $n > 1$ un entier et $C = \{(u_1,v_1),\ldots,(u_k,v_k)\}$ un ensemble de couples de mots. On définit la grammaire $G_{C,n}$ par les règles

$$\begin{aligned} S &\longrightarrow S_1 \\ &\vdots \\ S_{n-1} &\longrightarrow S_n \\ S_n &\longrightarrow u_1 S_n \tilde{v_1} + \cdots + u_k S_n \tilde{v_k} + c \end{aligned}$$

où $\tilde{v}$ dénote le miroir du mot v.

g) Montrer que le degré de $G_{C,n}$ est au moins $n+1$ et que le langage engendré par $G_{C,n}$ est

$$L_C = \{uc\tilde{v} \mid \exists p > 0, \exists i_1,\ldots,i_p \text{ tels que } u = u_{i_1}\cdots u_{i_p} \text{ et } v = v_{i_1}\cdots v_{i_p}\}.$$

h) On considère le langage $M = \{ucv \mid u,v \in A^* \text{ et } v \neq \tilde{u}\}$. Donner une grammaire linéaire H qui engendre M. Quel est le degré de la grammaire H ?
On rappelle que les réponses doivent être justifiées.

i) Montrer que le problème de Post associé à l'ensemble de couples C n'a pas de solution si et seulement si $L_C \subseteq M$. En déduire que le degré d'une grammaire linéaire n'est pas calculable.

2 Fonctions récursives (5 points)

Dans cet exercice, on pourra utiliser les fonctions primitives récursives et les schémas de construction de fonctions primitives récursives vus en cours. On rappelle en particulier que la fonction MOD est primitive récursive. La fonction $\mathrm{MOD} : \mathbb{N}^2 \to \mathbb{N}$ est définie par $\mathrm{MOD}(a,0) = 0$ et, si $b \neq 0$, $\mathrm{MOD}(a,b)$ est le reste de la division euclidienne de a par b.

a) Montrer que le prédicat $\mathrm{PR} : \mathbb{N} \to \mathbb{N}$ qui à un entier n associe 1 si n est premier et 0 sinon est primitif récursif.

b) Montrer que la fonction $f : \mathbb{N} \to \mathbb{N}$ qui à un entier n associe le $(n+1)^{\mathrm{e}}$ nombre premier est récursive. Par exemple, $f(0) = 2$, $f(1) = 3$, $f(2) = 5$, $f(3) = 7$, etc ...

c) Montrer que la fonction f définie à la question précédente est en fait primitive récursive.

3 Not All Equal SAT (6 points)

Le problème NAESAT est le suivant :

Donnée : Une formule φ en forme normale conjonctive dont chaque clause contient exactement trois littéraux.

Question : existe-t-il une assignation des variables telle que dans chaque clause de φ les trois littéraux ne s'évaluent ni tous à VRAI ni tous à FAUX?
Dans ce cas, on dit que l'assignation satisfait φ au sens de NAESAT.

Remarque : On accepte qu'un littéral apparaisse plusieurs fois dans une même clause. Ainsi, il est équivalent de demander que chaque clause contienne au plus trois littéraux ou exactement trois littéraux.

a) Donner une réduction de NAESAT à SAT.

b) Soit φ une instance de NAESAT et σ une assignation. On note $\bar{\sigma}$ l'assignation complémentaire définie par par $\bar{\sigma}(x) = \neg\sigma(x)$. Montrer que σ satisfait φ au sens de NAESAT si et seulement si $\bar{\sigma}$ satisfait φ au sens de NAESAT.

c) Montrer que NAESAT est NP-complet.
Indication : Reprendre la réduction de CIRCUIT-SAT à 3-SAT. Introduire une nouvelle variable et l'ajouter aux clauses qui ne contiennent pas déjà trois littéraux.

4 3-Coloriage (9 points)

Soit $\varphi = C_1 \wedge \cdots \wedge C_n$ une formule qui est une conjonction de clauses ayant chacune trois littéraux. Pour chaque $1 \leq j \leq n$, on note $C_{j,1}$,$C_{j,2}$ et $C_{j,3}$ les littéraux qui composent C_j. On a donc $C_j = C_{j,1} \vee C_{j,2} \vee C_{j,3}$. Finalement, on note $x_1, \ldots ,x_p$ les variables de la formule φ.

On construit le graphe non dirigé $R(\varphi) = (S,A)$ de la façon suivante. L'ensemble des sommets est

$$S = \{\top,x_1,\neg x_1, \ldots ,x_p,\neg x_p,C_{1,1},C_{1,2},C_{1,3}, \ldots ,C_{n,1},C_{n,2},C_{n,3}\}.$$

Le graphe contient les arêtes suivantes:

1. $(\top,x_i)$, $(\top,\neg x_i)$ et $(x_i,\neg x_i)$ pour chaque $1 \leq i \leq p$,
2. $(C_{j,1},C_{j,2})$, $(C_{j,2},C_{j,3})$ et $(C_{j,3},C_{j,1})$ pour chaque $1 \leq j \leq n$,
3. $(x_i,C_{j,k})$ si le littéral $C_{j,k}$ est $\neg x_i$, et $(\neg x_i,C_{j,k})$ si le littéral $C_{j,k}$ est x_i, ceci pour chaque $1 \leq i \leq p$, $1 \leq j \leq n$ et $1 \leq k \leq 3$.

a) Dessiner le graphe $R(\varphi)$ pour $\varphi = (x_1 \vee \neg x_2 \vee x_3) \wedge (\neg x_1 \vee x_2 \vee x_3)$.

Le problème du 3-coloriage est le suivant:

Donnée : Un graphe non dirigé $G = (S,A)$.

Question : Existe-t-il un coloriage des sommets avec 3 couleurs tel que deux sommets adjacents n'aient pas la même couleur, i.e. une application $c : S \to \{0,1,2\}$ telle que $\forall i,j \in S$, $(i,j) \in A \Longrightarrow c(i) \neq c(j)$?

b) Montrer que φ vérifie NAESAT si et seulement si le graphe $R(\varphi)$ est 3-coloriable.
Indication : On pourra se ramener au cas où le sommet $\top$ est colorié par 2. Pour les sommets du graphe qui correspondent à des littéraux, on pourra interpréter 0 comme faux et 1 comme vrai tandis que 2 pourra représenter aussi bien vrai que faux.

c) Montrer que R définit une réduction de NAESAT au problème du 3-coloriage. Que peut-on en déduire?