

Grammaires de graphes et logique : Exercices et problèmes

Bruno Courcelle
UFR de Maths et Info, Université Bordeaux 1

Automne 2004

Les exercices marqués * sont "difficiles". Je n'ai pas de réponse aux questions marquées **.

Sauf indication contraire, tous les arbres, graphes et ensembles ordonnés sont finis.

Version du 28 novembre 2004.

1 : Grammaires de graphes

1-1) Rédiger complètement la preuve que, dans le cas général, la plus petite solution du système d'équations associé à une grammaire algébrique est formée des langages engendrés par les non-terminaux.

1-2) Montrer que le système d'équations $\{X = aXY + b ; Y = bXYc + X\}$ a une seule solution dans le domaine des langages de mots finis sur n'importe quel alphabet contenant a, b, c . Est-ce toujours vrai dans le domaine des langages de mots finis et infinis ?

1-3) Donner des conditions suffisantes pour qu'un système d'équations ait une solution unique dans le domaine des langages de mots finis.

1-4)* Donner une condition nécessaire et suffisante pour qu'un système d'équations ait une solution unique dans le domaine des langages de mots finis.

1-5) Construire une grammaire qui engendre les cographes dont le nombre de sommets est un multiple de 3. Même question pour les graphes série-parallèles.

1-6) Construire une grammaire qui engendre les graphes série-parallèles 2-coloriables.

1-7)* En utilisant la notion de sur-solution d'une équation, montrer que tout graphe série-parallèle est 3-coloriable. Montrer que tout graphe série-parallèle est planaire.

1-8) Quelles propriétés algébriques (associativité, commutativité, distributivité, ...) satisfont les opérations qui engendrent les co-graphes et les graphes série-parallèles ?

1-9) Démontrer que les systèmes :

$$\{X = X \otimes (X \oplus X) + \mathbf{1}\} \text{ et } \{X = Y \otimes X + \mathbf{1} ; Y = X \oplus X\}$$

définissent le même ensemble X de cograves.

1-10) Tout mot peut être considéré comme un graphe orienté dont les arcs sont étiquetés par des lettres. Le mot vide correspond à un sommet isolé. Tout langage algébrique $\subseteq \{a, b, c\}^*$ est un ensemble de graphes HR-équationnel. Montrer que les langages non algébriques $\{a^n b^n c^n / n > 0\}$ et $\{ww/w \in \{a, b, c\}^*\}$ sont HR-équationnels.

1-11)** Est-ce que l'intersection de deux langages algébriques est toujours HR-équationnelle ?

1-12) Construire un système d'équations qui définit l'ensemble des *arbres sans racine* (graphes simples, connexes, non orientés, sans cycles) dont le nombre de sommets est un multiple de 3.

1-13) Le *complémentaire* \overline{G} d'un graphe simple nonorienté sans boucle G est le graphe de même type obtenu en supprimant tous les arcs de G et en ajoutant les arcs $x - y$ pour tous $x \neq y$, non adjacents dans G . Montrer que le complémentaire d'un cographe est un cographe. Transformer un système d'équations définissant des cograves en un système définissant leurs complémentaires.

1-14) Quelle condition syntaxique sur un terme t à k variables est nécessaire et suffisante pour que $t^{\mathcal{P}\mathcal{A}}(X_1, \dots, X_k) = \{t^{\mathcal{A}}(a_1, \dots, a_k) \mid a_i \in X_i\}$ pour toute algèbre \mathcal{A} et pour toutes parties X_1, \dots, X_k du domaine de \mathcal{A} ?

2 : Structurations des graphes et largeurs correspondantes

2-1) Montrer que la largeur arborescente d'un graphe est égale au maximum des largeurs arborescentes de ses composantes connexes, ainsi qu'au maximum des largeurs arborescentes de ses composantes biconnexes. Le résultat analogue est-il vrai pour les composantes fortement connexes ?

En considérant le graphe P_4 (chemin non orienté à 3 arêtes et 4 sommets), montrer que la largeur de clique d'un graphe n'est pas forcément le maximum de celles de ses composantes biconnexes.

Majorer la largeur de clique d'un graphe en fonction de celles de ses composantes biconnexes.

2-2) Soit $G_{n \times m}$ la grille rectangulaire à $n \times m$ sommets. Donner le meilleur majorant possible de sa largeur arborescente.

2-3) En fonction de quels paramètres (degré, largeur arborescente, diamètre,...) relatifs à un graphe G peut-on majorer la largeur arborescente de son "line graph" $Line(G)$ (graphe des incidences entre les arêtes) ?

2-4) Un graphe *planaire extérieur* (*outerplanar*) est un graphe connexe dont tous les sommets sont incidents à la face infinie pour un plongement dans le plan. Montrer que ces graphes sont de largeur arborescente bornée. Quel est le maximum ?

2-5) Soit H un graphe obtenu à partir d'un graphe G par subdivisions itérées d'arêtes. Peut-on majorer la largeur arborescente de H en fonction de celle de G ?

2-6)* Un ensemble d'entiers A est semi-linéaire si le langage $\{a^n / n \in A\}$ est rationnel. Montrer que cela équivaut à dire que A est une union finie de progressions arithmétiques. Le *Théorème de Parikh* montre que l'ensemble des nombres d'occurrences de la lettre a dans les mots d'un langage algébrique est semi-linéaire. Montrer que si a est une étiquette d'arc, l'ensemble des nombres d'arcs étiquetés a dans les graphes d'un ensemble HR-équationnel est semi-linéaire. Montrer que si a est une étiquette de sommet, l'ensemble des nombres de sommets étiquetés a dans un ensemble de graphes VR-équationnel est semi-linéaire. En déduire que l'ensemble des grilles carrées $G_{n \times n}$ n'est ni HR-équationnel, ni VR-équationnel.

Indiquer d'autres manières de prouver ces faits.

2-7) Le graphe $G(f)$ d'une fonction partielle $f : V \rightarrow V$ où V est un ensemble fini a pour sommets les éléments de V et pour arcs les couples $x \rightarrow f(x)$ pour tous $x \in V$. Montrer que ces graphes ont une largeur arborescente bornée indépendante de f et V . Préciser son maximum.

2-8)* Montrez que l'ensemble des graphes $G(f)$ est HR-équationnel et VR-équationnel.

2-9) Montrer que tout graphe planaire est mineur d'une grille carrée $G_{n \times n}$ suffisamment grande.

2-10)* Montrer que l'ensemble des "line"-graphes des graphes série-parallèles est un ensemble VR-équationnel. Donner un majorant de la largeur de clique de ces graphes.

2-11)* Soient G et H deux graphes simples, sans boucles, non orientés, et v un sommet de G . On note $G[H/v]$ le résultat L de la substitution de H à v dans G , défini ainsi, en supposant G et H disjoints : Les sommets de L sont ceux de G et de H sauf v , les arêtes de L sont celles de G non incidentes à v , celles de H et les arêtes $u-w$ pour tout u dans G adjacent à v et tout w dans H . Montrer que la largeur de clique de L est le maximum de celles de G et de H .

2-12) Montrer que si L est un ensemble VR-équationnel de graphes simples, sans boucles, non orientés, alors l'ensemble M de leurs *complémentaires* ("edge complements", voir 1-13) est VR-équationnel.

2-13)** Est-il vrai que pour G graphe simple, sans boucles, non orienté, sa largeur arborescente est majorée par $a \cdot \text{Card}(E_G)^{1/2} + b$ pour des constantes a et b ?

C'est bien le cas pour les cliques et pour les grilles.

2-14) La *Propriété de Helly* pour les arbres (graphes non orientés, connexes, sans cycles) énonce que si k sous-arbres d'un arbre sont 2 à 2 d'intersection non vide, alors l'intersection de ces k sous-arbres est non vide.

Montrer que si l'intersection de 2 sous-arbres est non vide, c'est un arbre.

Montrer que si 3 sous-arbres sont 2 à 2 d'intersection non vide, alors leur intersection est un sous-arbre non vide.

Montrer la propriété de Helly par récurrence sur k .

On rappelle que dans un arbre, deux noeuds sont reliés par un unique chemin.

2-15)* A faire en utilisant 2-14) Soit G un graphe non orienté et (T, f) une décomposition arborescente de ce graphe. On écrit $K(A, B) \subseteq G$ si A et B sont 2 ensembles disjoints de sommets tels que tout sommet de A est adjacent à tout sommet de B .

Montrer que si $K(A, B) \subseteq G$ alors l'un au moins de A et B est contenu dans une boîte de (T, f) , par récurrence sur la cardinalité de $A \cup B$. On montrera d'abord cela pour A et B chacun à 2 éléments.

3 : Calculs inductifs

3-1) Déterminer les propriétés inductives permettant de tester sur les arbres syntaxiques des graphes de largeur arborescente au plus 3 si ces graphes :

- sont connexes,
- sont sans circuit,
- ont un circuit Hamiltonien (qui passe par chaque sommet une fois et une seule ; un 8 n'est pas Hamiltonien).

3-2) Déterminer inductivement le nombre de chemins orientés qui relient les deux sommets distingués d'un graphe série-parallèle.

3-3)* On considère les arbres binaires finis formalisés comme les termes de $T(\{f, a, b\})$ où f est un symbole binaire, a et b sont deux constantes. Soit L l'ensemble des arbres tels que toutes les occurrences de la constante b sont à égale distance de la racine. Montrer qu'il n'existe pas d'automate fini dont L est l'ensemble associé. (On pourra utiliser des lemmes de l'ouvrage en ligne TATA). En conclure que l'appartenance à L n'est pas une propriété inductive.

3-4) Avec les notations de 3-3), on définit le *nombre de Strahler* de $t \in T(\{f, a, b\})$ par l'induction suivante :

$$\begin{aligned} NS(a) &= NS(b) = 1 \\ NS(f(t_1, t_2)) &= NS(t_1) + 1 \text{ si } NS(t_1) = NS(t_2), \\ NS(f(t_1, t_2)) &= \text{Max}\{NS(t_1), NS(t_2)\} \text{ si } NS(t_1) \neq NS(t_2). \end{aligned}$$

Construire un système d'équations Σ_k qui définit l'ensemble des arbres de nombre de Strahler au plus k .

On donnera une construction telle la taille de Σ_k soit $O(k)$. (La *taille d'un système* est la somme des longueurs de ses équations, tous les symboles étant comptés pour 1, même dans le cas de symboles x_k , pour k grand, que l'on pourrait compter comme de longueur $\log(k)$.)

3-5) Généraliser les définitions et les méthodes de calcul concernant les fonctions et les ensembles de fonctions inductivement calculables au cas des algèbres à plusieurs sortes et pour les fonctions de toutes arités.

4 : Logique du second-ordre monadique et propriétés de graphes

4-1) a) Dans la définition du degré d'un sommet d'un graphe non orienté, on compte pour une boucle pour 2. Ainsi la somme du degré sortant et du degré entrant d'un sommet d'un graphe orienté est égale au degré de ce sommet dans le graphe non orienté associé en "oubliant les orientations". Ecrire une formule

du premier ordre qui exprime qu'un graphe non orienté est de degré au plus k , k fixé (c'est à dire que chaque sommet est de degré au plus k).

On précisera pour quels types de graphes G la représentation $|G|_1$ suffit. On écrira des formules pour les 2 types de représentation considérés, $|G|_1$ et $Inc(G)$.

b) Ecrire une formule de la LSOM (logique du second-ordre monadique) qui exprime qu'un graphe orienté possède deux sommets s et t tels que tout sommet est sur un chemin qui va de s à t . Laquelle des deux représentations $|G|_1$ et $Inc(G)$ convient ?

4-2) Soit V un ensemble fini (partiellement) ordonné par \leq . Son *diagramme de Hasse* est le plus petit graphe orienté H dont V est l'ensemble de sommets et la fermeture transitive est \leq . Montrer l'existence et l'unicité de H . Sans l'hypothèse de finitude de V , H peut ne pas exister : donner un contre-exemple. Montrer que l'on peut définir les arcs de H par une formule de la LSOM évaluée dans la structure $\langle V, \leq \rangle$.

4-3)* Construire une formule de la LSOM qui exprime qu'un arbre de degré au plus k a un nombre pair de nœuds.

4-4) Construire une formule de la LSOM qui exprime qu'un graphe donné G est biparti complet (isomorphe à $K_{n,m}$, pour quelques n et m non donnés). Construire une formule de la LSOM utilisant des quantifications sur les ensembles d'arcs qui exprime que G est isomorphe à $K_{n,n}$ pour quelque n non donné.

4-5)* On rappelle qu'un langage est l'ensemble des modèles d'une formule de la LSOM si et seulement si il est rationnel. Le langage $\{a^n b^n / n > 0\}$ n'est pas rationnel. Donner de ce fait une preuve simple. En construisant un graphe biparti sur chaque mot de $a^* b^*$, montrer l'on ne peut pas construire une formule de la LSOM n'utilisant pas de quantification sur les ensembles d'arcs qui exprimerait qu'un graphe donné G est isomorphe à $K_{n,n}$.

4-6) Construire une formule de la LSOM utilisant des quantifications sur les ensembles d'arcs qui exprime qu'un graphe fini est Hamiltonien. En utilisant 4-5), montrer que l'on ne peut pas faire cette construction sans utiliser de quantification sur les ensembles d'arcs.

4-7) Un arbre binaire fini ou infini est donné comme une structure relationnelle $A = \langle N, lsuc(.,.), rsuc(.,.) \rangle$ (ensemble des nœuds, $lsuc(x,y)$ ssi y est fils gauche de x , $rsuc(x,y)$ ssi y est fils droit de x). Construire une formule de la LSOM qui exprime qu'une structure représente bien un arbre (tout nœud accessible depuis la racine par un unique chemin etc...). Construire des formules de la LSOM qui définissent sur un arbre binaire les ordres totaux de parcours en "preorder", "postorder", "inorder".

4-8)* Suite du précédent ; on peut utiliser la méthode de 4-5) : montrer que l'ordre hiérarchique sur N défini par $x < y$ ssi $d(r,x) < d(r,y)$ ou $d(r,x) =$

$d(r, y)$ et $x <_{in} y$ n'est pas définissable en LSOM (où $d(r, x)$ est la distance de x à r , la racine, et $<_{in}$ est l'ordre du parcours "inorder").

4-9)* Suite de 4-7) L'objectif est de construire une formule de la LSOM $\varphi(X)$ qui exprime que dans un arbre binaire, un ensemble de nœuds est fini.

Soit A un arbre binaire (cf 4-7)). Une *branche* est chemin orienté issu de la racine qui est *maximal*, c'est à dire qui est infini ou bien aboutit à une *feuille* (un nœud sans successeurs). Un ensemble de nœuds C est une *coupe* si :

- c'est une *anti-chaîne* (deux éléments sont incomparables pour la relation ancêtre)

- et il rencontre (forcément en un seul nœud) toute branche infinie.

Montrer en utilisant Lemme de Koenig (voir cours de I. Walukiewicz) que toute coupe est finie.

Utiliser cette notion pour construire $\varphi(X)$.

4-10) Montrer que pour tout graphe fini H simple, non orienté, qui peut avoir des boucles, on peut construire une formule de la LSOM sans quantifications sur les ensembles d'arcs qui exprime qu'un graphe donné contient H comme mineur.

4-11) Construire une formule $\gamma(x, y, X)$ de la LSOM avec quantifications sur les ensembles d'arcs qui exprime que X est l'ensemble des sommets d'un chemin qui relie x à y . (On rappelle qu'un chemin ne passe pas deux fois par le même sommet).

Construire une formule $\delta(x, y)$ de la LSOM sans quantifications sur les ensembles d'arcs qui exprime que dans un graphe orienté fini sans circuit, il existe un chemin orienté de x à y . (On observera que les graphes réduits à un chemin fini peuvent être caractérisés, parmi les graphes finis, par une formule du premier ordre).

5 : Traduction d'une formule LSOM en automate

5-1) Que faut-il ajouter à la preuve du Théorème de Fefermann-Vaught-Shelah pour traiter de la même manière les formules de la logique CMS (= Logique SOM avec comptage) ? En déduire qu'un langage est CMS-définissable si et seulement si il est rationnel.

Donner une autre preuve de ce résultat en montrant que, dans les structures représentant des mots, les prédicats $C_p(X)$ disant que la cardinalité de X est un multiple de p , sont définissables par des formules de la LSOM.

5-2) Montrer le Théorème de Fefermann et Vaught pour les formules du premier ordre et le produit cartésien $S \times T$ de deux structures S et T relatives à la même signature.

5-3) Démontrer que tout langage MS-définissable est rationnel en utilisant le théorème de Fefermann-Vaught-Shelah.

5-4) Soit **EqMS** la logique SOM enrichie de formules atomiques $Eq(X, Y)$ qui disent que deux ensembles X et Y sont de même cardinalité. Construire des formules de **EqMS** qui caractérisent, parmi les structures qui représentent des mots le langage algébrique non rationnel $\{a^n b^n \mid n > 0\}$, les langages non algébriques $\{a^n b^n c^n \mid n > 0\}$ et $\{ww \mid w \in \{a, b\}^*\}$. En conclure que le Théorème de Fefermann-Vaught-Shelah est faux pour les formules de **EqMS**.

5-5)* Montrer que, pour toute instance du Problème de Correspondance de Post, on peut construire par un algorithme une formule de **EqMS** qui est satisfaisable si et seulement si cette instance admet une solution. En déduire que, sur l'ensemble des mots de $\{a, b\}^*$, le problème de satisfaisabilité pour **EqMS** est indécidable.

5-6) Les propriétés " G est planaire" et " G est 4-coloriable" sont toutes les deux exprimables en LSOM. Le Théorème des 4 Couleurs qui dit que tout graphe planaire est 4 coloriable (et dont la preuve est difficile) pourrait-il être démontré comme conséquence de ces observations en utilisant un résultat de décidabilité pour la LSOM ?

5-7) Dans un graphe G on définit la distance de deux ensembles de sommets A et B comme le nombre d'arêtes minimal d'un chemin non-orienté d'un sommet de A à un sommet de B . Cette distance est nulle si A rencontre B . On représente (G, A, B) par la structure logique pour G , comprenant éventuellement les relations unaires représentant les ports, augmentée de deux relations unaires A et B .

a) Exprimer la distance de A et de B comme $Min\varphi$ pour une formule $\varphi(X)$ à construire.

b) De quelles nouvelles constantes faut-il enrichir la signature fonctionnelle VR pour représenter ces graphes avec leurs ensembles A et B ?

c) Soit G fixé, donné par un terme sur VR_k disons t . Pour A et B , ensembles de sommets, comment faut-il transformer t en un terme que l'on notera $t[A, B]$, pour représenter (G, A, B) ?

d) Montrer comment la distance de A et B peut être calculée par un calcul ascendant sur $t[A, B]$.

6 : Transductions de graphes définissables en LSOM