

Université Bordeaux I
Master d'Informatique 2^e année
Traitement Symbolique des Langues
Christian Retoré

Examen du 29 janvier 2004
Seuls les documents manuscrits sont autorisés

Le sujet est sans doute un peu long. On traitera ce qu'on pourra.

Exercice A

(A.i) Appliquer l'algorithme RG d'apprentissage des grammaires AB rigides aux deux exemples suivants:

- $[_e[_e \text{ Pierre aime}]] \text{ Marie}$
- $[_e \text{ Pierre} [_e \text{ aime Marie}]]$

(A.ii) Qu'en concluez-vous?

Exercice B

On considère la grammaire AB G définie par le lexique suivant:

$$\begin{aligned} a & : S / B, (S / B) / S \quad [\text{deux types pour } a] \\ b & : B \end{aligned}$$

(B.i) Donnez une grammaire algébrique (hors-contexte) en forme normale de Greibach G' qui engendre le même langage que G . Il n'est pas demandé de démontrer que les langages engendrés sont les mêmes.

(B.ii) Donnez une grammaire algébrique (hors-contexte) en forme normale de Chomsky G'' qui engendre le même langage que G . Il n'est pas demandé de démontrer que les langages engendrés sont les mêmes.

(B.iii) En appliquant l'algorithme d'analyse de Cocke Kasami Younger à G'' montrez que $aabb$ appartient au langage engendré et pas $abab$.

(B.iv) Quel est le langage engendré par ces grammaire? Il n'est pas demandé de le démontrer.

Exercice C

On considère la phrase suivante:

(1) Une secrétaire inscrit chaque étudiant.

(C.i) Combien y a-t-il de formules logiques non-équivalentes exprimant en logique du premier ordre le sens de cette phrase?

(C.ii) Donner un lexique associant à chaque mot de la phrase un type syntaxique, et un lambda terme sémantique qui lui corresponde.

(C.iii) Trouver une analyse syntaxique pour chacune des lectures proposées à la première question. Pour chaque analyse, on explicitera le calcul de la représentation sémantique correspondante.

Exercice D

(D.i) Donner un type syntaxique à la préposition « à » et au verbe « *présenter* » de sorte à ce qu'on puisse analyser des phrases du genre:

(2) X présente Y à Z.

(D.ii) On considère la phrase suivante:

(3) Une secrétaire présente chaque étudiant à un professeur.

Combien cette phrase a-t-elle de lectures différentes? Donner une formule logique pour chacune de ses lectures.

(D.iii) (difficile) Construire un lexique syntaxique et sémantique contenant les mots « *un* », « *chaque* », « *une* », « *étudiant* », « *professeur* », « *secrétaire* », « *présenter* », « *à* » de manière à pouvoir analyser la phrase quantifiée précédente.

(D.iv) (difficile) Pensez-vous que toutes les lectures que vous avez proposées correspondent à des analyses dans le calcul de Lambek, Justifiez votre réponse. Eventuellement, démontrer que votre réponse est correcte.

Exercice E

Etant donnée une formule F du calcul de Lambek, on note $|F|$ son nombre d'occurrences de variables propositionnelles. Par exemple $|(S / B) / S| = 3$. On admet le résultat suivant, qu'on essaiera de bien comprendre:

Lemme: *Tout séquent $A_1, \dots, A_n \vdash A_{n+1}$ avec $|A_i| \leq m$ est démontrable dans le calcul de Lambek si et seulement si il est démontrable dans le calcul de Lambek*

- en n'utilisant que la règle de coupure,
- à partir de séquents démontrables dans le calcul de Lambek de la forme (a) ou (b):
 - (a) $X_1, X_2 \vdash X_3$ avec $|X_j| \leq m$ pour $j \in \{1, 2, 3\}$.
 - (b) $Y_1 \vdash Y_2$ avec $|Y_j| \leq m$ pour $j \in \{1, 2\}$.

(E.i) Pourquoi n'y a-t-il qu'un nombre fini de séquents démontrable de la forme (a) ou (b)?

(E.ii) (cours) Soit Lex le lexique d'une grammaire de Lambek, et soit m la taille maximum d'une formule du lexique. Soit $Z_1, \dots, Z_p \vdash Z_{p+1}$ un séquent apparaissant dans une démonstration d'un séquent $A_1, \dots, A_n \vdash S$ tel que pour tout $i \leq n$ il existe un mot w tel que $A_i \in \text{Lex}(w)$. Pourquoi a-t-on $|Z_i| \leq m$ pour tout $i \in \{1, \dots, p+1\}$.

(E.iii) (demande une bonne intuition) Etant donnée une grammaire de Lambek définie par le lexique Lex , en s'inspirant du lemme ci-dessus, donnez une grammaire hors-contexte (algébrique) qui engendre le même langage.

(E.iv) (difficile) Démontrez que la grammaire de Lambek et la grammaire hors-contexte (algébrique) que vous avez proposée à la question précédente engendrent effectivement le même langage.