

Traitement Automatique du Langage Syntaxe

TP LFG

Travaux pratiques avec l'analyseur LFG de A. Andrews

Vous allez utiliser aujourd'hui un analyseur LFG dont le nom est PC-LFG. Comme son nom l'indique, il fonctionne sur toutes les machines de type PC, et est très peu gourmand : un simple environnement DOS est suffisant et ses sources (compilées) tiennent sur une disquette 720Ko. Il est distribué à des fins de recherche uniquement, et peut être obtenu à l'adresse suivante : <http://arts.anu.edu.au/linguistics/software/lfgpc.html>. Vous pouvez donc le copier pour votre usage personnel, mais en aucun cas vous ne devez en faire un usage commercial. La version actuelle est la 2.0. et fonctionne sous DOS (et sous Windows en mode DOS).

Comment écrit-on une grammaire avec PC-LFG ?

D'un point de vue méthodique, PC-LFG se révèle facilement utilisable et fait bien la distinction entre la grammaire et le lexique, même si ces derniers figurent dans le même fichier. Ainsi, on écrit d'abord la grammaire, puis le lexique.

Pour analyser une phrase, il suffit de choisir le menu *Parsing*, de saisir la phrase que l'on souhaite étudier, puis d'appuyer sur *F10*, puis *espace*. Si une analyse est trouvée, le programme fournit la structure-c, puis la structure-f. Si d'autres analyses sont possibles, il fournira alors une autre structure-c puis une autre structure-f.

Quelles sont les conventions d'écriture ?

On peut utiliser les parenthèses pour noter une catégorie ou une séquence optionnelle. Les pipes (|, i.e. en général *Alt Gr 6*) servent à marquer une disjonction (elle peut être marquée au niveau des catégories ou des séquences, ou bien au niveau des traits (ex : MODE ind|subj)). La négation de la valeur d'un trait est notée /= (ainsi MODE /=ppart indique que l'on souhaite tout type de mode sauf le participe passé). Enfin l'étoile de Kleene est naturellement représentée par *.

Au niveau des équations, la flèche vers le haut est notée par l'accent circonflexe (i.e. ^) et la flèche vers le bas est notée par un v (suivi d'une espace). Quelques abréviations sont utilisées. Ainsi ^=v n'est pas inscrit, mais est tacitement présent. Les noms de fonctions sont simplement indiqués par des deux-points (i.e. :)

Plutôt que d'être écrites au-dessous des non-terminaux des règles syntagmatiques, les équations fonctionnelles sont écrites à côté de chaque non-terminal.

Par exemple :

$$\begin{array}{ccccc} \text{SV} & \rightarrow & \text{V} & & \text{SN} \\ & & \uparrow = \downarrow & \uparrow \text{OBJ} = \downarrow & \end{array}$$

s'écrira :

SV -> V (^)=v SN (^ OBJ)= v)

ou plus simplement :

SV -> V SN:OBJ

Pour plus de renseignements, n'hésitez pas à consulter les fichiers d'aide disponibles (rédigés en anglais), spécialement le fichier `hothelp.doc`.

Manipulations avec PC-LFG

1- Sous Windows, ouvrez un terminal (Menu Démarrer -> Exécuter. Taper la commande « `cmd` ».)

2- Déplacez-vous dans le répertoire contenant le programme (commande « `dir` » pour voir la liste des répertoires, « `cd <nom_du_répertoire>` » pour changer de répertoire).

3- Tapez « `lfg` » pour entrer dans l'analyseur PC-LFG (vous pouvez lancer d'abord la commande « `hothelp` », qui chargera en mémoire une aide dynamique. Il suffira alors de presser Alt-Shift-D pour visualiser l'aide. Mais le mieux est de visualiser cette aide dans un autre terminal, lorsqu'on est sous Windows.)

4- Chargez alors la grammaire : choisissez le menu Grammar. Attention : la souris n'est pas active ! Le menu Edit Load vous permet de visualiser la grammaire que vous chargez. Tapez `fr.lfg`

5- Vous devez voir s'afficher la grammaire. Observez-la un peu, puis tapez F10 : le programme vous propose de renommer le fichier contenant la grammaire. Appelez le donc `fr1.lfg`

6- Allez maintenant dans le menu Parsing. Vous allez analyser quelques phrases... Les phrases ne commencent pas par une majuscule (vous n'utiliserez les majuscules que pour les noms propres), mais se terminent par un point. Il faut alors appuyer sur F10, puis espace. Analysez au moins les phrases suivantes :

1. Max dort.
2. le singe dort.
3. le singe mange une banane.

7- Les phrases suivantes sont-elles reconnues ? S'il y a un échec, quelle en est la raison ?

1. Max dorment.
2. le chat mange une banane.
3. le singe mange.
4. le chien dort le chat.

8- Corrigez la grammaire de façon à traiter l'accord entre le sujet et le verbe.

Pour cela, retournez dans le menu Grammar, puis faites Edit & Load. Il ne vous reste alors plus qu'à compléter la grammaire.

Lorsque le résultat vous paraît satisfaisant, parsez les phrases de votre choix et notez les analyses obtenues (vous pouvez déjà reprendre les phrases données en 6 et 7)

Faites de même pour les questions suivantes.

9- Elargissez alors la grammaire de manière à pouvoir traiter toute la conjugaison de **manger** et de **dormir** au présent. Pensez à ajouter les pronoms personnels manquants.

10- Elargissez la grammaire de manière à pouvoir traiter les objets indirects introduits par *à* et *de*. Ajoutez également les verbes « *parler de ... à ...* » et « *donner ... à ...* ».

11- Introduire les contractions de « *du, des* » et « *au, aux* » et s'arranger pour interdire « *de le, de les, à le, à les* ».

12- Elargissez la grammaire de manière à pouvoir traiter les adjectifs épithètes et leur accord. (*le petit chat, les petits chats, les gentils petits chats noirs*).

13- Elargissez la grammaire de manière à pouvoir traiter les attributs du sujet (*les chats sont noirs, les chats sont des animaux*).

14- Traitez les compléments circonstanciels de temps : modifiez la grammaire de manière à ce que la phrase *le chat dort la nuit* soit reconnue, mais que *le chat dort le chien* ne le soit pas.