

T.P. Traitement Automatique des Langues (Retoré)
Université de Montpellier — Master Info 1 – UE TAL

Ce travail sera à déposer sur le Moodle (la date sera précisée ultérieurement) : on rendra un fichier comprenant les réponses aux questions précédées de début quelques lignes d'explications — en utilisant les signes de commentaire % ou /* ... */). Le vocabulaire variera d'un étudiant à l'autre, ainsi que les constructions syntaxiques traitées.

Il s'agit d'illustrer le cours par une grammaire de clauses définies (DCG) à écrire en Prolog (de préférence SWI Prolog). En fait Prolog permet d'écrire la DCG directement comme un ensemble de règles de production (avec le symbole "→"), mais pour comprendre le fonctionnement des grammaires en Prolog, on écrira les clauses correspondantes ; du reste cette syntaxe Prolog pour les grammaires n'est qu'un confort purement syntaxique : lorsqu'on affiche un programme écrit avec "→" ce sont les clauses Prolog habituelle avec ":-" qui apparaissent et non règles de la DCG avec "→".

On souhaite modéliser par des clauses écrites en Prolog la grammaire de phrases simples avec des relatives qui sont aussi bien introduites par le pronom : « qui » (dont l'antécédent est le sujet de la relative) que par le pronom « que » (dont l'antécédent est l'objet de la relative). « *Pierre regarde Marie qui dort. Marie regarde Pierre qui mange une pomme. Marie qui regarde Pierre tombe. Pierre dort. Les pommes tombent.* » Outre les pronoms relatifs mentionnés, les mots considérés sont les articles définis et indéfinis, deux noms propres « Pierre » et « Marie », les noms communs « melon », « pomme », « copain », « femme », les verbes transitifs « regarder », et « manger », les verbes intransitifs « dormir » et « tomber » (les noms commençant par une voyelle complique un peu le traitement des articles : on dit « l'abricot » et non « le abricot »).

A Écrire une grammaire hors-contexte (on dit aussi algébrique) qui produise ces phrases

B En partant de la grammaire hors contexte qui engendre ces phrases, que l'on pourra améliorer et étendre, écrire une DCG qui prenne en compte l'accord en genre et en nombre (attention aux relatives !), et qui produise l'arbre syntaxique comme une liste de listes à peu près comme ceci :

— `analyse(T, [marie, regarde, pierre, qui, mange, une, pomme], [])` répond vrai avec `T` qui représente (il n'y a pas d'indices en prolog, il faudra mettre la catégorie syntaxique comme premier terme de la liste)

`[s[gnmarienp], [gvregardevt, [gn[nppierre[relnquipronmangevt[gnunedet, pommen]]]]]]]`

— `analyse(T, [marie, regarde, pierre, qui, mangent, une, pomme], [])` répond faux (problème d'accord)

On procèdera par différence de listes ($x(L, M)$ est vrai lorsque $L = [w_1, \dots, w_p, w_{p+1}, \dots, w_n]$, $M = [w_{p+1}, \dots, w_n]$ et $[w_1, \dots, w_p]$ est un syntagme x) avec la clause `entre(Mot, [Mot|Reste], Reste)` (où `Mot` et `Reste` sont des variables) et on ajoutera des attributs pour l'accord.

C On programmera ensuite une interface qui permette de saisir une phrase et qui lorsqu'un mot n'est pas présent demande de spécifier sa catégorie et la valeur des attributs de cette catégorie et ajoute ces informations au programme via la commande `assert`.

D Les plus courageux peuvent en outre essayer de traiter certaines des questions suivantes :

- Ajouter des adjectifs.
- Compléter la grammaire avec les pronoms personnels sujets « je tu il/elle nous vous ils/elles ».
- Inclure dans la grammaire l'obligation que le sujet ou l'objet d'un verbe soient animés ou non animés, par exemple « la pomme mange pierre » n'est pas une phrase correcte de ce point de vue.
- Inclure dans l'interface la possibilité de rajouter un mot d'une catégorie nouvelle avec les règles associées à cette nouvelle catégorie, c'est assez difficile car il va falloir ajouter des règles qui produise cette nouvelle catégorie et des règles utilisant cette catégorie pour en produire d'autres.
- Produire des représentations sémantiques, une liste de prédicats appliqués à leurs arguments, le prédicat étant noté en tête de liste : pour la phrase ci-dessus : `[[regarde, pierre, marie], [mange, pierre, pomme]]` on pourra réfléchir au traitement de l'article, mais c'est difficile.