



Travaux Dirigés d'informatique linguistique n°6

Cours d'Introduction à l'informatique linguistique

—Licence Informatique 3ème année—

Unification

Dans ce TP, nous allons voir la notion d'unification de structures de traits.

1 Prise en main de l'unification avec NLTK

Lire la documentation de la section 9.2.6. `Feature Structures in NLTK and Unification` dans le document :

<http://nltk.sourceforge.net/lite/doc/en/featgram.html>

Implémenter les exemples.

2 bases de données et unification

Soit la base de données représentée par les trois tables suivantes :

np	nom	prenom	ville
1	Dupond	Luc	Paris
2	Dupond	Leon	Marseille
3	Martin	Luc	Lyon
4	Cresus	Louis	Paris

TAB. 1 – Personne

ns	nom	type
1	football	collectif
2	rugby	collectif
3	judo	individuel
4	tennis	individuel

TAB. 2 – Sport

np	ns
1	1
2	2
3	2
4	1
4	4

TAB. 3 – Pratiquer

1. Représenter chaque table par une liste de structures de traits (chaque ligne est une structure de traits)
2. À l'aide de l'opération d'unification, calculer la liste des personnes habitant à Paris.
3. À l'aide de l'opération d'unification, calculer la liste des personnes qui pratiquent un sport collectif.
4. À l'aide de l'opération d'unification, calculer la liste des personnes qui s'appellent Dupond et qui pratiquent un sport collectif.

3 Levee d'ambiguite

Les contraintes d'accords peuvent servir à lever l'ambiguïté lexicale de certains mots. Par exemple, le verbe **viens** est ambigu : le verbe **venir** au présent soit à la première personne du singulier soit à la deuxième personne du singulier. Le contexte "je viens" et la contrainte d'accord permet de lever cette ambiguïté.

1. Écrire une fonction **desamb** qui prend comme paramètre une liste de structures de traits ambigus (ex. les deux structures de traits de la forme **viens**), une structure de traits exprimant le contexte de levée d'ambiguïté (ex. la structure de traits associée à la forme **je**) et une contrainte (ex. **accord**). Cette fonction retournera la bonne analyse parmi l'ensemble des analyses possibles.

Par exemple, `desamb([viens1,viens2],je,"accord")` retournera **viens1**, si :

- **viens1** est la structure de traits associée à la forme **viens** à la 1ère personne du singulier
- **viens2** est la structure de traits associée à la forme **viens** à la 2ème personne du singulier
- **je** est la structure de traits associée à la forme **je** à la 1ère personne du singulier
- **"accord"** est l'attribut dont la valeur est une structure de traits représentant l'accord (cf. cours).

2. Tester l'exemple ci-dessus.
3. Trouver d'autres exemples et appliquer les à l'aide de **desamb**.

4 Implantation de l'algorithme d'unification

Implémenter en Python l'algorithme d'unification vu en cours. Vous définirez votre propre représentation des structures de traits.