Chaîne de traitements linguistiques pour l'acquisition automatique de lexique

Programmation d'une mini grammaire TAG pour le français

Lionel Clément

Projet Atoll

INRIA Rocquencourt

Lionel.Clement@inria.fr

Journée RLT – le jeudi 12 septembre 2002

Utilisation du compilateur de grammaires de Bertrand Gaiffe

- Motivations :
  - Produire une petite grammaire du français (quelques centaines de schèmes)
  - Faire évoluer cette grammaire très facilement
  - L'adapter aisément à la chaine de traitement
- Moyens :
  - Reprise du travail de thèse de Marie-Hélène
  - Utilisation de l'outil de Bertrand Gaiffe
  - Quelques semaines de travail, pas plus

Outils

- Feuille de style pour visualiser les arbres produits (XSLT)
- Outil plus souple pour éditer la *Méta-Grammaire* (YAMGE)
- Et aussi : nouveau compilateur de grammaires (YAMGC) (inachevé)
  - Pour produire des grammaires autres que TAG (RCG, LFG)
  - Pour explorer d'autres choix méthodologiques

YAMGE

- Script XSLT pour visualiser les arbres produits (XML → LaTeX)
- Script Perl pour éditer le document XML via une interface graphique

  daVinci de Univ. Bremen : http://www.informatik.uni-bremen.de/daVinci/
- Petit script tcl/tk pour compiler les fichiers, visualiser, etc.

Mini Grammaire du Français

- 4 dimensions : place, sous-cat initiale, redistribution et réalisation
  - 4 dimensions : place, sous-cat initiale, redistribution et réalisation
  - Place : ordre des mots (envisager l'usage pour autre chose que clitiques)
  - Sous-cat initiale : dépend du nombre d'arguments. Doit-on distinguer les arguments des actants ?
  - Redistribution des fonctions initiales
  - Réalisation des fonctions finales
- Description partielles d'arbres
  - place, sous-cat initiale (auxiliaires, montée, causatif, verbes pleins), réalisation

Remarques pour Bertrand

- Je n'utilise pas la puissance de l'unification (mais seulement du filtrage)
- Je pourrais ne pas utiliser l'opérateur de dominance large (.. >). Est-il coûteux ?

YAMGC

- Quelques choix envisagés
  - Besoins/ressources : liste d'ensembles
  - Quasis-noeuds : variables ou constantes (les variables libres sont éliminées)
  - Ajout de contraintes négatives de croisement de classes
  - Modification de l'organisation de la MG en dimensions (graphes distincts)

demo de yamge