

Responsable : Patrick SAINT-DIZIER

stdizier@irit.fr ☎ 05 61 55 62 44

| ECTS | COURS | TD | TP | |
|------|-------|----|----|--|
| 3 | 20 | | | |

Objectifs

La communication Homme-Machine en langage naturel, sous une grande diversité de formes (langue simplifiée, schémas, langue standard, interrogation, les modes textuels, en situation de dialogue, dans le WEB, dans les autoroutes de l'information et le multi-media, etc.) se trouve au coeur de nombreuses préoccupations (interrogation de services, dialogue homme-machine, traductions, aide à la rédaction technique, systèmes tutoriels). Elle fait appel à des outils conceptuels ou techniques largement répandus (ontologies, terminologies, systèmes de traits,...).

Ce cours propose une analyse des fondements théoriques du traitement automatique du langage naturel écrit. Il propose les éléments fondamentaux linguistiques, syntaxiques et sémantiques, cognitifs et conceptuels, de représentation des connaissances et de traitement informatique.

Description

1. Objectifs de la communication en langue naturelle :
les enjeux, les exigences, familles d'applications
le langage et son environnement

2. Rappel d'éléments d'informatique théorique :
grammaires, automates, systèmes de traits, stratégies d'analyse et de génération, unification, lambda-calcul

3. Typologie des applications incluant des éléments de langue :
familles d'applications et leurs principales caractéristiques.

4. Fondements linguistiques :
Lexique : morphologie, sous-catégorisation, rôles thématiques, relations lexicales
fondements de la syntaxe, techniques d'analyse syntaxique et de construction de représentations.
fondements de la sémantique : notions de sens, principe de compositionnalité, formalismes sémantiques

5. Méthodes d'analyse et d'évaluation :
Méthodes d'analyse de problèmes (par ex. sur corpus), dépouillement de données langagières, annotations.
Evaluation des coûts et des besoins.
Techniques d'évaluation des ressources, protocoles pour l'évaluation des productions (analyse, génération).

Ouvrages Conseillés

Mathematical Methods in Linguistics. B. Partee et ali. Kluwer academic, 1993.