

Recherche d'Information et Traitement Automatique des Langues Naturelles

Josiane Mothe^{1,2}

(1) IRIT, UMR 5505, Université de Toulouse, 118 Route de Narbonne, 31062 Toulouse Cedex

(2) IUFM, Ecole interne, Université de Toulouse, 56 av. de l'URSS, 31079 Toulouse

Josiane.mothe@irit.fr

RÉSUMÉ

La recherche d'information s'intéresse à l'accès aux documents et une majorité de travaux dans le domaine s'appuie sur les éléments textuels de ces documents écrits en langage naturel. Les requêtes soumises par les utilisateurs de moteurs de recherche sont également textuelles, même si elles sont très pauvres d'un point de vue linguistique. Il paraît donc naturel que les travaux en recherche d'information cherchent à s'alimenter par les avancées et les résultats en traitement automatique des langues naturelles. Malgré les espoirs déçus des années 80, l'engouement pour l'utilisation du traitement du langage naturel en recherche d'information reste intact, poussé par les nouvelles perspectives offertes.

Dans cette conférence, nous balayerons les aspects de la recherche d'information qui se sont le plus appuyés sur des éléments du traitement automatique des langues naturelles. Nous présenterons en particulier quelques résultats relatifs à la reformulation automatique de requêtes, à la prédiction de la difficulté des requêtes, au résumé automatique et à la contextualisation de textes courts ainsi que les perspectives actuelles offertes en particulier par les travaux en linguistique computationnelle.

ABSTRACT

Information Retrieval and Natural Language Processing

Information retrieval aims at providing means to access documents. Most of current work in the domain relies on the textual elements of these documents which are written in natural language. Users' queries are also generally textual, even if the queries are very poor from a linguistic point of view. As a results information retrieval field aimed at feeding on advances and results from natural language processing field. In spite of the disappointed hopes of the 80s, the enthusiasm for using natural language processing in information retrieval remains high, pushed by the new perspectives.

In this talk, we will mention the various aspects of information retrieval which rely, at various levels, on natural language processing components. We will present in particular some results regardless automatic query reformulation, query difficulty prediction, automatic summarization and short text contextualization as well as some perspectives offered in particular considering computational linguistics.

MOTS-CLÉS : Recherche d'information, traitement automatique des langues, reformulation de requêtes, difficulté des requêtes, résumé automatique

KEYWORDS: Information retrieval, natural language processing, query reformulation, query difficulty, automatic summarization
