
NLAISE

NATURAL LANGUAGE ACCESS TO INTERNET SEARCH ENGINES

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE
IN
COMPUTER SCIENCE
UNIVERSITY OF REGINA

BY
Gayathri Mahalingam
Regina, Saskatchewan
September 1997

Abstract

Natural language access to selected Internet search engines is presented. Searching for relevant documents using the existing search engines poses certain problems that include finding the most appropriate search term and scanning through a large number of potentially relevant documents. We have provided a natural language access which enables the user to present the query in English without any need to transform it to suit the individual search engines. The user's query is semantically analyzed using a HPSG parser to generate appropriate search terms. Through the semantic analysis, we eliminate the contribution of non-keywords to the search and thus help to reduce the number of sites returned.

We have evaluated the performance of our system with regard to the quality of search results. The results demonstrate the ease of expression by which we frame the query and consistency of responses. The total number of sites returned is also less compared t O the result s returned from the existing search engines. We have thus presented an application of natural language processing techniques to improve the performance of the existing search engines.

The full thesis can be found at

http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp04/mq30514.pdf