Minimal parsing Key Concept based Question Answering System

Sunil Kopparapu*, Akhilesh Srivastava, P.V.S. Rao

The home page of a company is an effective means for show casing their offerings in general and their products and technology in particular. Companies invest major effort, time and money in designing their web pages to enable their users to access information they are looking for as quickly and as easily as possible. In spite of all these efforts, it is not uncommon for a user to spend a sizable amount of time trying to retrieve the particular information that he is looking for. Today, he has to go through several hyperlink clicks or manually search the pages displayed by the site search engine to obtain the information that he is looking for. Much time gets wasted if the required information does not exist on that website. With websites being increasingly used as sources of information about companies and their products, there is need for a more convenient interface.

In this paper we discuss a system based on a set of Natural Language Processing (NLP) techniques which addresses this problem. The system enables a user to ask for information from a particular website in free style natural English. The NLP based system is able to respond to the query by understanding the intent of the query and then using this understanding to retrieve relevant information from its unstructured info base or structured database for presenting it to the user. The interface avoids the user having to click through several hyperlinked pages. The core of system is its ability to understand the question without formally parsing it. The system is based on identifying key concepts and keywords and then using them to retrieve information. The no parse approach enables system framework to be used for different input languages with minimal architectural changes. Further, the key concept-key word approach gives the system an inherent ability to provide approximate answers in case the exact answers are not present in the information database or infobase.

In the full paper we will discuss (a) the need for a minimal parsing system while identifying the disadvantages of a full parsing QA system, (b) describe in greater depth the approach (keyconcept keyword based system) that we have adopted and (c) show how it can be used to create a better interface to access information form a practical web-page.

*Contact surface: Cognitive Systems Research Laboratory, Tata Consultancy Services Limited, Plot No 14, Sector 24, Vashi-Turbe, Navi Mumbai 400705, India. phone: +91-22-56163251, fax: +91-22-27839926, e-mail: SunilKumar.Kopparapu@TCS.Com