Analysis and Assessment of Cross-Language Question Answering Systems

Juncal Gutiérrez-Artacho
University of Granada, Spain

María-Dolores Olvera-Lobo
University of Granada, Spain

INTRODUCTION

Within the sphere of the Web, the overload of information is more notable than in other contexts. Thus, too often, on planning a search with tools from the Web (search engines, directories, or meta-search engines), the number of web pages found proves excessive and not all of them are relevant or useful for the objectives of the user.

Cross-language information retrieval (CLIR) is an active sub-domain of Information Retrieval (IR). Like IR, CLIR is centered upon the search for documents, reconciling queries and documents which are written in different languages. Cross-lingual information access covers a wide range of tasks that enable users to access information in languages other than their own, including IR, Question Answering, information extraction and summarization. CLIR enables users to find information in languages they do not know, but CLIR search results are not immediately useful because a separate translation techniques must be applied before the user can read the results. This type of systems has opened a new research field that examines the most effective methods for IR and investigates which resources are required for a correct translation.

Question answering systems (QAS) are presented as an alternative to the traditional Information Retrieval (IR) systems, seeking to offer precise and understandable answers to factual questions instead of showing the user a list of documents related to a given search (Jackson and Schilder, 2005). The functioning of QAS is based on short-answer models (Blair-Goldensohn et al., 2004), and the main advantage that it offers is that the user does not have to consult complete documents to collect the information needed, as the system provides the correct answer in the form of a number, a noun, a short phrase, or a brief fragment of text.

In relation to the types of QAS, the present work focuses on the analysis and assessment of multilingual and cross-language QAS. These systems need to include some type of linguistic translation resource, tool, or technique for the proper retrieval of the result, since the QAS can retrieve the answer from a collection of documents written in languages differing from the one in which the question is formulated. Given that the QAS is presented as a substantial advance in the improvement of IR (Kolomiyets and Moens, 2011), it becomes necessary to determine its effectiveness for the final user. With this aim, 7 studies were undertaken to evaluate: a) in the first two, the linguistic resources and tools used in these systems for multilingual retrieval (Research 1; Research 2); and b) the performance and quality of the answers of the main monolingual and multilingual QA of general domain and specialized domain in the Web (QuALiM, SEMOTE, START, TrueKnowledge, and HONqa) in response to different types of questions (of definition, of facts, and of lists) and subjects (e.g. art, literature, biology, medicine, names, history, economy, or sports), so that different evaluation means can

be applied (Research 3, Research 4, Research 5, Research 6, Research 7).

BACKGROUND

In the field of CLIR tools are being created that can greatly assist specialists in their work; as well as helping other users find a wide variety of information. These tools are evolving but several years of study and research are still needed to improve implementations. One of the main difficulties facing these tools is the task of translating queries made by users and the documentary sources found in response (Diekema, 2003). Given the current expansion in research, development, and the creation of CLIR systems, it was considered worthwhile analysing and evaluating the resources used by one type of these systems: multi-lingual QAS.

Frequently, a keyword query entered into a web search tool (search engine or meta-search engine) to satisfy a user's information need, provides too many result pages – many of which are useless or irrelevant to the user. In effect, modern IR systems allow us to locate documents that might have the associated information, but the majority of them leave it to the user to extract the useful information from an ordered list (Dwivedi & Singh, 2013). In contrast to the IR scenario, a QAS processes questions formulated into Natural Language instead of keyword based queries, and retrieves answers instead of documents (Peñas et al., 2012). Therefore, the usefulness of these types of systems for quickly and effectively finding specialized information has been widely recognized (Diekema et al., 2004).

While the development of QAS represents progress, the systems nevertheless suffer restrictions. Many were only developed as prototypes, or demonstration versions, and few were marketed. Some researchers have designed and created systems that were presented and discussed at various forums and conferences. However, because the usefulness of the systems was limited to very specific contexts, or because of problems of implementation, only a few of these systems were later developed for end users.

These circumstances have fuelled academic interest in CLIR, and the techniques of natural language processing. This interest led to many conferences dealing exclusively, or partially, with CLIR – such as TREC, the Cross-Language Evaluation Forum (CLEF), the NII Text Collection for IR Systems (NTSIR), the Language Resources and Evaluation Conference (LREC), among others. However, research on CLIR, which mostly began in 1996, has not led to commercial success and so dissemination was limited.

METHODOLOGY

For the purposes of in-depth analysis of QAS, a series of studies has been proposed which are gathered together in seven research articles published in international magazines with peer review and a chapter in a book. The following table provides a summary of the aspects addressed in the analysis performed.

Table 1. Summary table of the contents of the publications

A1: Analysis Oriented at Resources and Linguistic Tools in Multilingual QAS		
Research 1: Analysis of the linguistic resources used in CLEF		Research 2: General analysis of linguistic resources used in QAS
A2: Evaluation Oriented at the System		
Research 3: Identification and analysis of the evaluation measures used for the testing of QAS		
Research 4: Evaluation of QAS as efficient sources of terminological information.	Research 5: Evaluation of general domain QAS against those of specialist domain in the biomedical field.	Research 6: Evaluation of the efficacy of the functioning of general domain QAS.
Research 7: Final evaluation. The objective and subjective evaluation of the multilingual search for answers		

7 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/analysis-and-assessment-of-cross-language-question-answering-systems/184155

Related Content

Hybrid Data Mining Approach for Image Segmentation Based Classification
Mrutyunjaya Panda, Aboul Ella Hassanienand Ajith Abraham (2016). *International Journal of Rough Sets and Data Analysis (pp. 65-81).*
www.irma-international.org/article/hybrid-data-mining-approach-for-image-segmentation-based-classification/150465

Efficient Cryptographic Protocol Design for Secure Sharing of Personal Health Records in the Cloud
Chudaman Devidasrao Sukte, Emmanuel Markand Ratnadeep R. Deshmukh (2022). *International Journal of Information Technologies and Systems Approach (pp. 1-16).*
www.irma-international.org/article/efficient-cryptographic-protocol-design-for-secure-sharing-of-personal-health-records-in-the-cloud/304810

Australian Users' Interactions with E-Services in a Virtual Environment
Kamaljeet Sandhu (2012). *Virtual Work and Human Interaction Research (pp. 115-126).*
www.irma-international.org/chapter/australian-users-interactions-services-virtual/65318

An Effective Emotional Analysis Method of Consumer Comment Text Based on ALBERT-ATBiFRU-CNN
Mei Yang (2023). *International Journal of Information Technologies and Systems Approach (pp. 1-12).*
www.irma-international.org/article/an-effective-emotional-analysis-method-of-consumer-comment-text-based-on-albert-atbifru-cnn/324100

A Protocol for Evaluating Mobile Applications
Clare Martin, Derek Floodand Rachel Harrison (2013). *Information Systems Research and Exploring Social Artifacts: Approaches and Methodologies (pp. 398-414).*
www.irma-international.org/chapter/protocol-evaluating-mobile-applications/70726