

Bookmarking, thesaurus, and cooperation in bilingual question answering

Jussi Karlgren, Preben Hansen, Magnus Sahlgren
Swedish Institute of Computer Science
jussi,preben,mange@sics.se

1 Three simultaneous experiments

This year's interactive cross-lingual information access experiment was designed to measure three parameters we expected could affect the performance of users in cross-lingual tasks in languages in which the users are less than fluent. Firstly, introducing new technology, we measure the effect of topic-tailored term expansion on query formulation. Secondly, introducing a new component in the interactive interface, we investigate - without measuring by using a control group - the effect of a bookmark panel on user confidence in the reported result. Thirdly, we ran subjects pair-wise and allowed them to communicate, to investigate the effect of cooperation on task completion.

2 Subjects and data

The target language was French. No translation service was provided. All sixteen queries were formulated in French, all documents were displayed in French, and answers were given in French. All eight subjects were primarily Swedish speakers and moderately competent in the target language: typically they had taken French as their second foreign language requirement in high school. No fluent French readers were accepted.

3 System

The system was based on a standard retrieval system developed at SICS which allowed users to enter French words into a search window and then displayed search results in a standard ranked list. By clicking on a document title in the ranked list, the document itself was displayed in a document display window next to the ranked list.

4 Thesaurus and term expansion

The term expansion experiment was the most important experiment made in this set of user tests. Subjects were first given eight queries without term expansion capabilities and then eight with an added window where a French word could be entered to give up to five suggestions of related terms. The thesaurus used for expansion was generated automatically from parallel corpora of EU legislation by GSDM methodology (as described in previous year's CLEF publications by us) which relates terms by their distributional characteristics. As it turned out, subjects were frustrated by its inherent unpredictability and its patchy coverage. Most subjects tested it once or twice and did not use it thereafter.

5 Bookmark panel

Another feature introduced into this year's system by us was a bookmark panel. All users were given this feature in all queries. Users could mark an arbitrary selection from a displayed document and click a save button, to place a bookmark in a visible bookmark list next to the document display screen. Any number of selections in any document could be bookmarked for display in the bookmark panel. The first bookmarked selection was copied into the answer display field by default and answer reporting for each topic was then done by clicking "done" whereby the answer displayed in the answer display field was logged. Here the user did have the option of instead manually entering an answer into the answer display field or to edit the displayed default answer.

6 Cooperation

Subjects performed the experiments in pairs. The sequence of queries were kept different within pairs: the subjects never had the same query simultaneously. Subject communication was logged by encoding communication in one of a limited set of categories such as "vocabulary question" "system operation question".

7 Results and conclusions

The thesaurus was not useful due to its limited coverage. Users were happy about the bookmark panel and commented on it favourably: a comparative experiment without bookmarking will be performed shortly. Data analysis is still in process to investigate the effects of cooperation.

Acknowledgements

The work reported here is partially funded by the European Commission under contracts IST-2000-29452 (DUMAS) which is hereby gratefully acknowledged.