Supporting Complex Tasks in a Spoken Language Interface

Xiaojun Yuan
College of Computing and Information
University at Albany, SUNY
Albany, NY 12222, USA
+1 518 591 8746

xyuan@albany.edu

Nicholas J. Belkin
School of Communication & Information
Rutgers University
New Brunswick, NJ 08901, USA
+1 732 932 7500

belkin@rutgers.edu

ABSTRACT

Current search engines do a fine job in assisting users with simple and direct tasks, but need more improvement in coping with difficult user tasks. Users of information systems typically carry out searches with very short queries, on the order of two words or so. This makes it very difficult for the systems to disambiguate their queries and identify potentially relevant documents, and leads to sub-optimal retrieval performance. We hypothesize that users will provide better and more useful descriptions of their information problems if they are able to speak to the system and to easily indicate through speech and gesture, those documents and aspects of documents which they find useful, and not useful. Therefore, spoken language interfaces would be able to better assist users with difficult tasks.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *search process.*

General Terms

Design, Human Factors.

Keywords

Searching, spoken language interface, user performance

1. INTRODUCTION

There is no doubt that when dealing with simple and easy tasks, the existing search engines do a fine job. For example, "Where is the capital of China?", users can simply go to a search engine site, and type in "capital China." The answer can be found out from the snippet of the top ranked search results. However, current search engines do not do a good job on the complex situations because of the complexity of human information behavior and needs.

Situation 1 -- "Supporting simple and common requests that express complex and dynamic needs."

Assuming an attendee of SIGIR 2011 would like to find some social events or activities to enjoy in a night of the stay in Beijing, He types in the keywords "Entertain me in Beijing" in his favorite search engine. This task could be complex and challenging because the task itself being not specific but ambiguous and amorphous in goal, the language and culture difference, and required knowledge with China.

Situation 2 -- "Doing a task through a mobile environment." Assuming a SIGIR USA attendee is driving to the airport for SIGIR 2011 conference in Beijing, and needs to find a reasonably

rated parking lot near the airport. This task itself is not complex, but the information system needs to provide good support to accommodate the user's information needs in a mobile environment.

Users of information systems typically carry out searches with very short queries, on the order of two words or so [6]. This makes it very difficult for the systems to disambiguate their queries and identify potentially relevant documents, and leads to sub-optimal retrieval performance. Instead of simply returning a ranked list of documents to respond to this simple query, a better search system or interface is needed to assist users locate needed information in completing complex tasks. This system should assist users during their entire search process and reduce the degree of user perceived task complexity, by iteratively constructing a complex query or search strategy in each searching stage, and by progressively integrating the partial answers into a coherent one at the later stages. To achieve the above-mentioned goals, a spoken language interface which guides users in the user-computer dialogues, and iteratively accepts and aggregates the accumulated query results is necessary and appropriate. More specifically, in response to situation 1, a spoken query interface would allow the user to further extend the original query, and talk more about what the user would like to do to be entertained in Beijing. Through the user's spoken queries, the system would be able to elicit detailed and clear information needs from the user and to produce meaningful retrieval results for the user to choose. In situation 2, a spoken language interface is very important because the user may not be able to type in queries while driving the car. Such an interface would enable the user to articulate spoken queries by talking to the interface, and to respond to retrieval results or reformulate new queries without worrying about typing. Again, the system would respond more satisfactorily by the iteratively collected spoken queries.

In this paper, we propose that a spoken language interface or system that can allow users to talk about their information needs and use gesture to point out what they would like to view is appropriate and effective in supporting difficult tasks with complex and dynamic needs and should be addressed in the related field.

2. PROBLEM BACKGROUND

The tendency of users of information systems to begin their searches with brief queries is probably due to two factors: the general inability of people to specify precisely what documents they require in order to resolve their information problems (cf. [1]); and, the difficulty that people have in finding terms appropriate both for describing their information problems, and matching the terms which have been used to describe the documents in the database with which they are interacting.

To address these two problems, a variety of ways have been proposed and investigated. One approach has been to devise interface techniques which encourage searchers to input longer

1

queries (e.g. [7]); another has been to automatically enhance the initial query without the searchers' intervention, or through query expansion based on thesauri or similar tools (cf. [5]); a third to offer to searchers, based on their initial queries, terms which could be used to enhance their initial queries (e.g. [2]). Although each of these approaches has been shown to afford some benefit in retrieval effectiveness, none of them has involved searchers in developing and understanding their information problems, finding better ways to express their information "needs", nor succeeded in substantially improving either retrieval effectiveness or searcher satisfaction with the interaction [7].

We propose to address the problem of encouraging effective interaction of the searcher with information systems by moving from keyboard-based interaction to spoken language and gestural interaction of the searcher with the information system.

The origins of this approach are based on Taylor's research on question negotiation between user and librarian in special libraries [10], and on the experience of elicitation of verbal descriptions of searchers' information problems in studies of Anomalous State of Knowledge (ASK)-based information systems (e.g. [1]). Taylor found that, in the types of interactions that he studied, librarians engaged in conversations aimed at eliciting a number of different aspects of the searchers' information problems, and that the searchers were indeed able to address these different aspects. Belkin and his colleagues found that, when suitably prompted, searchers were able to provide search intermediaries with extended verbal descriptions of their information problems. Subsequently, Saracevic et al. [8], in their analysis of searcher and intermediary interaction with information systems, showed that there was substantial direct commentary by both searcher and intermediary on results retrieved with respect to a query put to the system, and, more recently, Crestani & Du [4] have shown that asking for expression of search need in verbal terms results in significantly longer queries than those expressed through a keyboard interface.

Crestani has led a group which has considered spoken language queries and their effectiveness in a variety of contexts [4]. Some of this work has investigated the effectiveness of spoken queries, as well as their length, but in simulated rather than real interaction. Zue, et al. [11]'s work is perhaps the most complete in terms of spoken language query understanding, but it has been applied in limited domains.

The main arguments against taking the spoken language and gesture approach to query input and interaction have been that: there has not been strong evidence that such interaction will actually result in more effective results; it is unclear that searchers will willingly engage in such interaction; and, most importantly, that speech understanding technology is not robust enough to support such interaction. Our position is that: there is some evidence that longer queries and more extensive response to search results that would be afforded by this mode of interaction does improve retrieval effectiveness (e.g. [7]); that when encouraged to describe their information problems more fully, searchers will do so ([3]; [7]); and, that spoken language interaction with information systems appears to be either doable right now [11] or in the very near future, with commercially available speech understanding systems (e.g. Dragon). There is also evidence that speech recognition technology is already in place in a mobile environment [9]. For instance, Google outlined developments in voice search, which allows users to search the Internet from a mobile phone by speaking their requests or queries to Google in Japanese, in addition to Chinese and English. Google is planning to add new languages next year.

3. DISCUSSION AND CONCLUSIONS

How to design a system that can iteratively provide assistance to users in different information-seeking stages during their complex task searching process is an important issue in interactive information retrieval. We believe our proposal makes an important step toward better understanding users' information needs, and investigating different ways to elicit users' information needs and thus in turn improve user performance and satisfaction, as well as reducing the perceived user task complexity.

4. ACKNOWLEDGEMENTS

5. REFERENCES

[1] Belkin, N.J. (1980) Anomalous States of Knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, v. 5: 133-143.

[2] Belkin, N.J., Marchetti, P.G. & Cool, C. (1993) BRAQUE: Design of an interface to support user interaction in information retrieval. *Information Processing and Management*, vol. 29: 325-344.

[3] Belkin, N.J., Cool, C., Kelly, D., Kim, G., Kim, J.-Y., Lee, H.-J., Muresan, G., Tang, M.-C., Yuan, X.-J. (2003) Query Length in Interactive Information Retrieval. In SIGIR '03. Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 205-212). New York: ACM.

[4] Crestani, F. & Du, H. (2006) "Written versus spoken queries: A qualitative and quantitative comparative analysis." Journal of the American Society for Information Science and Technology, 57(7): 881-890.

[5] Efthimiadis, E.N. (1996) Query Expansion. In: Williams, Martha E., ed. Annual Review of Information Systems and Technology, v31, pp 121-187, 1996.

[6] Jansen, B. J., Spink, A., and Saracevic, T. (2000). Real life, real users, and real needs: A study and analysis of user queries on the web. Information Processing and Management. 36(2), 207-227.

[7] Kelly, D., Dollu, V. J., & Fu, X. (2005). The loquacious user: A document-independent source of terms for query expansion. In Proceedings of the 28th Annual ACM International Conference on Research and Development in Information Retrieval (SIGIR '05), Salvador, Brazil, 457-464.

[8] Saracevic, T., Spink, A. & Wu, M-M. (1997). Users and intermediaries in information retrieval: What are they talking about? User modeling. Proceedings of the Sixth International Conference, UM97. New York: Springer, 43-54.

[9] Stone, B. (2009). Google Adds Live Updates to Results. New York Times, December 8 Issue. Retrieved from http://www.nytimes.com/2009/12/08/technology/companies/08google.html

[10] Taylor, R.S. (1968) Question negotiation and information seeking in libraries. College and Research Libraries, v. 29, 178-194.

[11] Zue, V., Seneff, S., Glass, J.R., Polifroni, J., Pao, C., Hazen, T.J. & Hetherington, L. (2000). JUPITER: A Telephone-Based Conversational Interface for Weather Information," *IEEE Transactions on Speech and Audio Processing*, Vol. 8, No. 1, January 2000.

2