Interface Techniques for Making Searching for Information More Effective

Nicholas J. Belkin
School of Communication, Information and Library Studies, Rutgers University
4 Huntington Street
New Brunswick, NJ 08901-1071 USA
+1 732 932 7500 x8271
nick@belkin.rutgers.edu

WORKSHOP TOPIC AREAS

Input, output, and iterative search; user psychology and behavior

EXTENDED ABSTRACT

In this presentation, I describe three different techniques that our research group has studied for making interactive information retrieval (IR) systems more usable and more effective, with particular reference both to methods that we have employed, and results that we have obtained. We have carried out most of this research within the context of the Interactive Track of the Text REtrieval Conferences (TREC) (see, e.g. Hersh & Over, 2002), which has been conducted annually since 1994. These have been experimental studies conducted under strict conditions, which allow us to draw statistically significant conclusions, and to some extent make our results comparable to those of other groups participating in the track.

One of the major problems facing users of interactive IR systems is finding the "right" words to use in the queries that they put to the systems. There is a well-known technique for addressing this problem, *relevance feedback*, which has been extensively tested in experimental, non-interactive IR system evaluation, and has been shown to be extremely effective in such environments (e.g. Salton & Buckley, 1990). Relevance feedback starts from the premise that a user in an IR system is quite unlikely to be able to come up with an explicit specification of what s/he wants to find, and that an interactive, iterative information seeking episode is necessary to achieve a "good" query to the system. The current best IR systems are those which provide search results ranked according to some prediction of the retrieved items' relevance; this prediction is typically based upon complex term-weighting formulae and matching algorithms. Because of the complexity (not to say inscrutability) of these mechanisms, it is thought that users cannot judge how best to modify their initial queries to make them better. However, users are usually able to make judgments of whether items which are retrieved are *relevant*, or not, to their interests or information problems. Relevance feedback takes advantage of this ability, by asking users to make such judgments, and then modifying the initial query on the basis of characteristics of the documents which have been judged relevant, or not. The typical modification is to increase the weight of query terms which occur in relevant documents, to decrease the weight of query terms which occur in non-relevant documents, and, most significantly, to add new terms to the query which are "important" in the relevant documents. All of this is typically understood to be accomplished without the user's intervention, or even knowledge.

There has, unfortunately, been relatively little investigation of relevance feedback techniques in interactive IR system contexts, and of the few studies, most have had negative results. That is, they have shown that users tend not to take up the opportunity to use relevance feedback when it is offered, and that when they do, they are not terribly pleased with the results (see, e.g. Hancock-Beaulieu & Walker, 1992). The exception to this trend is in a series of studies which were conducted by our research group between 1994 and 1999, which showed that particular ways of implementing relevance feedback (especially in terms of interface design) could result in usable and effective interactive IR systems. The first of these studies (Koenemann, 1996; Koenemann & Belkin, 1996) demonstrated that: (a) an interactive IR system using relevance feedback was more effective than one which did not offer this feature; and, (b) user control over the new terms which were added to the query led to better search results, and increased satisfaction than for versions of relevance feedback which did not offer such control.

Following these results, we ran a series of studies to investigate how best to implement user controlled relevance feedback in the IR system interface (Belkin, et al., 2001). The results of these studies were that systems which suggested terms for users to add to a query (with either positive or negative weights) based on relevance feedback were reasonably effective and usable, but, that a system which suggested terms to be added without asking for relevance judgments (using a *pseudo-relevance feedback* technique, which assumes that the top *n* retrieved documents are relevant) was better accepted, led to increased satisfaction with the search results, and to increased performance. Figure 1 is a screen shot of one version of our interface which offers term suggestions. Taken together, these results suggest specific ways in which term suggestion for supporting query modification can be implemented in interface design to make searching more effective.

A characteristic of best-match IR systems, which rank documents roughly based on the degree of match to the query, is that they achieve better performance as query length increases. However, queries in operational interactive IR systems are typically on the order of only two words or so. In order to deal with this mismatch, a good deal of research has been done on automatically increasing the length of the initial user's query, without the user's knowledge or intervention. Most such work has used some version of pseudo-relevance feedback. Our research group has taken another course, investigating interface techniques which might encourage users to start with longer queries in the first place. In Belkin, et al. (2002), we found that using a query box with room for five lines (figure 1) led to longer queries than using a standard, single-line query input mode. We also found that asking searchers to enter their queries as complete sentences or questions, as opposed to lists of keywords and/or phrases, led to significantly longer queries (even after non-content words were removed). This study demonstrated a positive relationship between query length and performance in the search task, but the results were only indicative. In addition, we saw a negative relationship between the extent of interaction that a searcher engaged in, and satisfaction with the search results.

More recently (Belkin, et al., 2003), we built on these results, and compared the efficacy of an interface to a Web search engine which asked people to describe their information problem to one which asked them simply to enter their query. The former system resulted in substantially and significantly longer queries, significantly increased satisfaction with search results, and significantly fewer query iterations per search. This rather simple difference in the interface to the search engine resulted in quite dramatic changes in behavior and performance. Figure 2 is a screen shot of our information problem description interface. In the same study, we also compared a system which we predicted would reduce interaction effort (one which displayed the full text of retrieved documents in four scrollable panes at a time), with a standard system which displayed twenty retrieved titles and descriptive snippets at a time. We found that the former system reduced the amount of interaction required to obtain comparable results, and led to significantly increased satisfaction with search results (figure 2).

It is of interest to note that the success of each of the techniques described above depends upon close coupling of interface design to underlying system capabilities, as well as to the cognitive tasks in which the user is engaged while searching for information. This suggests that considering only one of these two factors is insufficient, and that effective interface design for supporting information seeking should be based upon deep understanding of the underlying system, as well as of the problems which users face in interacting effectively with the system.

What might be the next steps to take in system design to support information seeking? I think that there are two approaches that should be investigated. One is to design systems which can support more than type of interaction with information within a single framework. Belkin (1996) and Cool & Belkin (2002 propose that an information seeking episode can be modeled as a sequence of different kinds of interaction with information, each requiring its own specific combination of support techniques (Figure 3). For instance, a searcher may begin an information seeking episode by getting an overview of what some databases cover, then move to browsing through some documents in the database, then to evaluating a few, then specifying a query, then saving some documents, and perhaps inserting part of one into another document. Designing a framework and interface to support this variety of interactions is clearly a substantial challenge.

The second approach is one that is often mentioned, but not yet often put into practice. That is personalization of search systems to their specific users and specific contexts and uses, based on implicit sources of evidence gathered from observation of user behavior, both past and during the specific search episode (e.g. Kelly & Belkin, 2002). Although this clearly requires substantial research with respect to the underlying system algorithms, such a system could not be implemented without substantial research in interface design. These two issues, supporting multiple interactions with information within a single interface framework, and personalizing the support of information interaction, are the next grand challenges for information search system design.

REFERENCES

1. Belkin, N.J. (1996) Intelligent information retrieval: whose intelligence? In J. Krause, M. Herfurth, & J. Marx (Eds.) *Herausfordurungen an die Informations-wirtschaft. Informationsverdichtung, Information-bewertung und Datenvisualisirung. Proceedings des 5. Internationalen Symposiums für Informations-wissenschaft (ISI '96)* (pp. 25-32). Konstanz: Universitätsverlag Konstanz.
2. Belkin, N.J., et al. (2001) Iterative exploration, design and evaluation of support for query reformulation in interactive information retrieval. *Information Processing and Management*, *37*, 403-434.
3. Belkin, N.J., et al. (2002) Rutgers' TREC 2001 Interactive track experience. In E.M. Voorhees & D.K. Harman (Eds.) *The tenth text retrieval conference, TREC 2001* (pp. 465-472). Washington, D.C.: GPO.
4. Belkin, N.J., et al. (2003) Rutgers Interactive track at TREC 2002. In E.M. Voorhees & D.K. Harman (Eds.) *The text retrieval conference (TREC 2002)* (in press). Washington, D.C.: GPO.
5. Cool, C. & Belkin, N.J. (2002) A classification of interactions with information. In H Bruce, R. Fidel, P. Ingwersen & P. Vakkari (eds.) *Emerging frameworks and methods. Proceedings of the Fourth International Conference on Conceptions of Library and Information Science (CoLIS4).* (pp.1-15). Greenwood Village, CO: Libraries Unlimited.
6. Hancock-Beaulieu, M. & Walker, S. (1992) An evaluation of automatic query expansion in an online library catalogue. *Journal of Documentation*, *48*, 4, 406-421.
7. Hersh, W. & Over, P. (2002) TREC-2001 interactive track report. In E.M. Voorhees & D.K. Harman (Eds.) *The tenth text retrieval conference, TREC 2001* (pp. 38-41). Washington, D.C.: GPO.
8. Kelly, D. & Belkin, N.J. (2002) A User Modeling System for Personalized Interaction and Tailored Retrieval in Interactive IR. In *Proceedings of ASIST 2002*. Medford, NJ: Information Today.
9. Koenemann, J. (1996) *Relevance feedback: usage, usability, utility*. Ph.D. Dissertation, Rutgers University. New Brunswick, NJ.
10. Koenemann, J. & Belkin, N.J. (1996) A case for interaction: A study of interactive information retrieval behavior and effectiveness. In: *Proceedings of CHI 96* (pp. 205-212). New York: ACM
11. Salton, G. & Buckley, C. (1990) Improving retrieval performance by relevance feedback. *Journal of the American Society for Information Science*, 41, 4, 288-297.

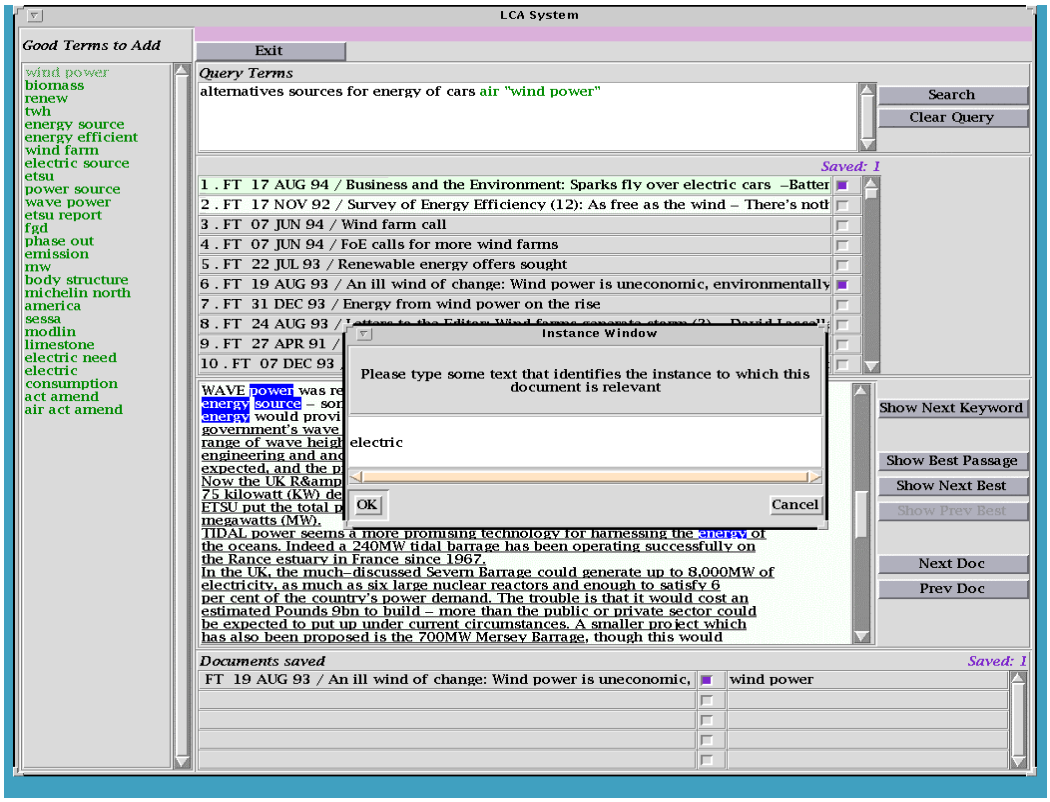

Figure 1. Interface which offers good terms to add based on pseudo-relevance feedback, with large query entry box.

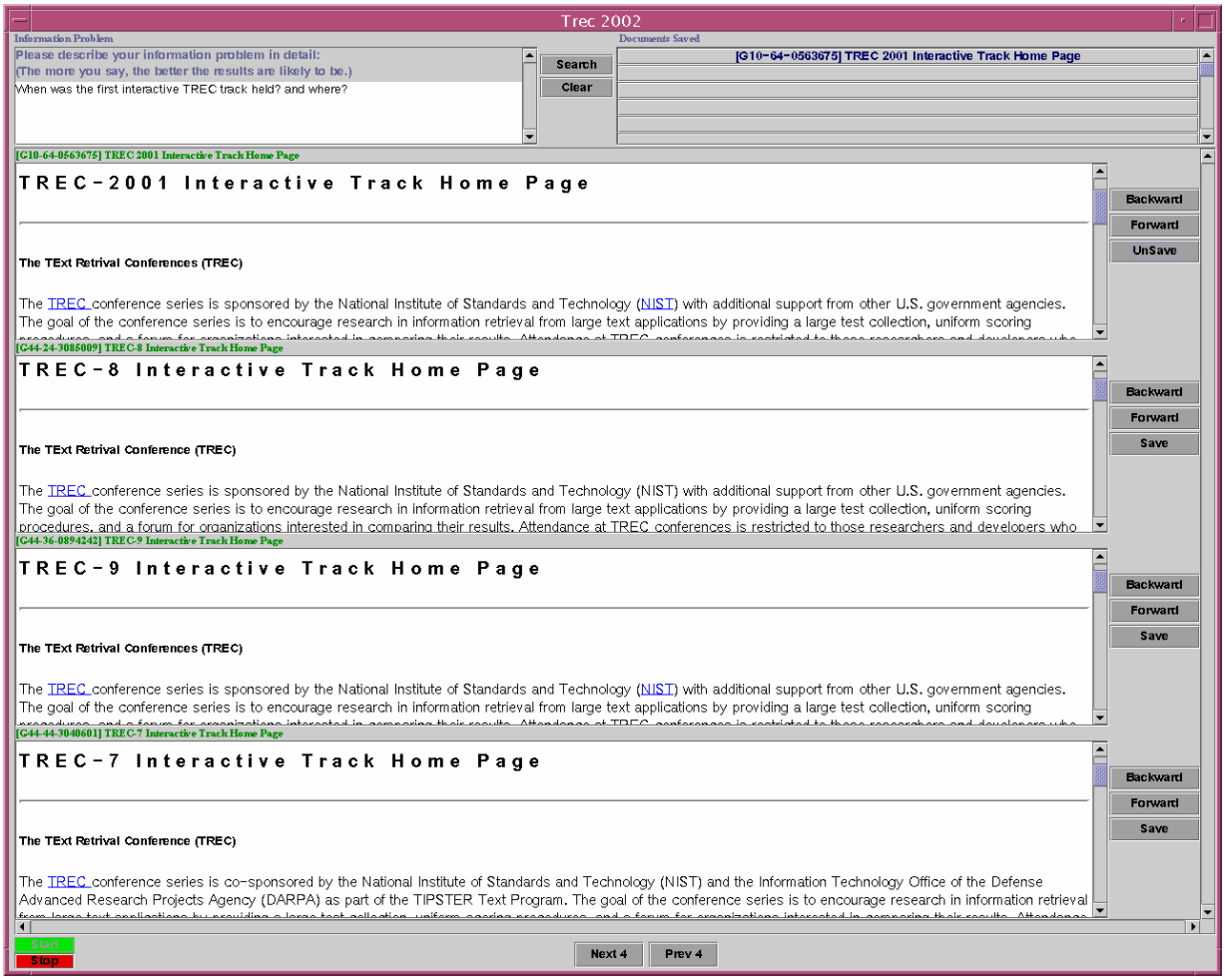

Figure 2. Interface for information problem description, with multiple document display.

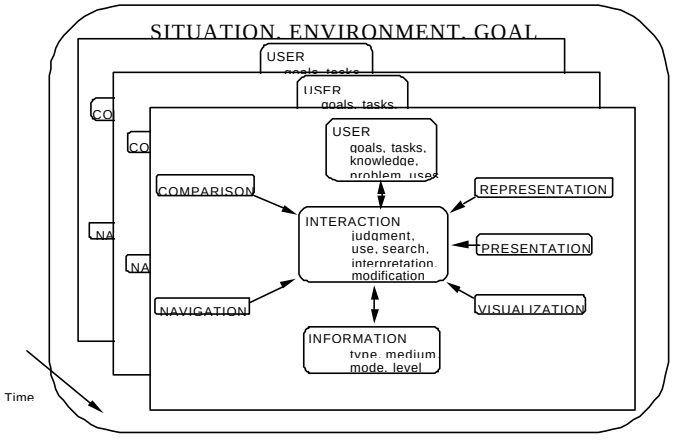

Figure 3. Information retrieval as support for interaction with information. (after Belkin, 1996)