

Supporting Multiple Information-Seeking Strategies in a Single System Framework

Xiaojun Yuan

School of Communication, Information and Library Studies
Rutgers, The State University of New Jersey
New Brunswick, NJ 08901, USA
yuanxj@rci.rutgers.edu

Abstract

This study investigates the support of multiple information seeking strategies (ISSs) within a single system, and the relation between varieties of ISSs and system design. It proposes to construct and evaluate an interactive information retrieval system which can adaptively support multiple ISSs, and allow change from one ISS to another within an ISS space. It is conducted in a series of steps: iterative designing -evaluating of several systems supporting different ISSs; specifying an interaction structure for multiple ISSs; and, implementing and evaluating a dynamically adaptive system supporting multiple ISSs. The study aims to make a contribution to interactive information retrieval drawing attention to user interface design, and to HCI, in integration of multiple support techniques within a single system framework.

Keywords

Information-seeking strategy, interaction structure, user interface design, evaluation, information retrieval

1 Introduction

Traditional information retrieval (IR) systems ask people to specify queries using terms to select documents from the selected databases. Current research indicates some problems of such systems. (1) Human information-seeking behavior is more complicated than just query formulation and term selection. For instance, people need to get an idea of which domain or genre of information they

need, then familiarize themselves with the content and structure of various databases. They need to learn about a domain of interest and extend their knowledge of this domain so that they can formulate more effective queries. They need to change their searching and seeking strategies to improve their queries. (2) Human information-seeking behaviors are not discrete processes. These behaviors interact with one another during information-seeking processes (Belkin, 1996). (3) Interaction, not representation or comparison, is the central process of IR (Belkin, 1993). (4) Users with different goals need to use different information-seeking strategies (ISSs) to conduct searches. However, traditional IR systems only support one ISS, that is, formulating queries using terms to select documents from some databases (Belkin, 1993, 1995; Xie, 2000). (5) It is known that different ISSs can be optimally supported by different combinations of IR techniques (Oddy, 1977). The existing diversity of ISSs indicates that a system which provides good support for one ISS is unlikely to provide good support for the others. A system trying to support all ISSs with only one technique will be able to support them at only mediocre levels.

Therefore, the goal of the research is to design an interactive IR system which incorporates different IR techniques to adaptively support different ISSs. Specifically, our solution to these problems focuses mainly on the following two goals.

- (1) Constructing a single IR system in which support techniques are explicitly represented so that it is possible to shift from one combination of support techniques to another in real time, and appropriate support techniques are suggested to the user by the system.
- (2) Evaluating the effectiveness and usability of the system within controlled experiments.

2 Research Problems

We aim to investigate the following research problems:

- (1) Implementing and evaluating several systems which are tailored to scanning or searching.
- (2) Developing a structure for guiding and controlling sequences of different support techniques.
- (3) Constructing and evaluating a single system which supports scanning and searching through integration of different support techniques within a single system framework.

3 Methodology

3.1 Research Problem 1

SYSTEMS

Using the Lemur toolkit (LEMUR), we implemented and evaluated several different prototype IR systems designed to support scanning (situation 1) or searching (situation 2). Table 1 describes the tasks, as well as features and support techniques for each system.

HYPOTHESES

Hypothesis 1: The system summarizing each database is more effective in supporting scanning tasks than the baseline system which provides a ranked list of documents with descriptions about which databases these documents are in. (E1.1/B1.1, situation1-task1)

Hypothesis 2: The system providing table of contents navigation is more effective in supporting scanning tasks than the baseline system which lists ranked paragraphs. (E1.2/B1.2, situation1-task2)

Hypothesis 3: The system presenting clustered retrieval results is more effective in supporting searching tasks than the baseline system which presents a ranked list of retrieval results. (E2.1/B2.1, situation2-task1)

Hypothesis 4: The system supporting fielded queries is more effective in supporting searching tasks than the baseline system which provides a generic query search. (E2.2/B2.2, situation2-task2)

Situations	Tasks	Experimental Systems	Baseline Systems	Support Techniques
1 Scanning	1 Identify best databases	E1.1 Alphabetically ordered databases (showing summary for each)	B1.1 Ranked documents (showing name of the database with the document)	Summary of each database
	2 Find comments or quotations from an electronic book	E1.2 Table of Contents navigation within documents	B1.2 Ranked paragraphs	Table of contents navigation
2 Searching	1 Find relevant documents	E2.1 Ranked clusters	B2.1 Ranked documents	Clustered retrieval results
	2 Find the name of an electronic book	E2.2 Field search	B2.2 Generic query search	Fielded query

Table 1. Situations, Tasks and Systems

EXPERIMENTAL DESIGN

Participants conducted four searches on four different topics that are suitable for scanning or searching. This is a within-subjects design. Each subject searched half of the topics in one system, then half of the topics in the other system. Within the topic block, the topic order was randomly assigned. No two subjects used the same order of topics and the same order of systems. The experiments were replicated by exchanging the order of the systems.

TEXT COLLECTIONS

There are two text collections: one is TREC HARD 2004 collection (HARD) which is suitable for situation1-task1 and situation2-task1, the other is a book database which is good for situation1-task2 and situation2-task2. This database is composed of books downloaded from Project Gutenberg (Gutenberg).

TASKS

In this study, we used the simulated work task situation model (Borlund, 1997) to make subjects' behavior as true-to-life as possible, hoping this will

make our results robust. Here is an example for situation1-task1.

Topic: As a graduate student, you are asked to write an essay about air pollution for one of your courses. You are supposed to get information you need from a system that is composed of several databases. Each database has lots of documents on a variety of topics. You believe it would be interesting to discover factors that cause air pollution, but you have no idea which databases are good on this topic.

Task: Please find out which databases are good for this particular topic, and rank the databases in order of likelihood of being good. Put your answer in the given space.

3.2 Research Problem 2 (Future Work)

In order to guide the presentation of specific support techniques during the information seeking process, we need to specify an interaction structure. This interaction structure is equivalent to a dialogue manager, and can be used to control the interactions between the system and the user. We will employ the idea of interaction structure developed in the MERIT system (Belkin, 1995). This structure models human-computer interaction as dialogues and particular dialogue structures are associated with different ISSs. This structure will be incorporated into the system at the user interface level and act as the dialogue manager.

3.3 Research Problem 3 (Future Work)

SYSTEM

The integrated system will allow the user to use a variety of ISSs and to seamlessly switch from one ISS to another in the information-seeking process. The user will be able to choose which ISS to use at any time. ISSs will be classified according to the goal of the interaction, the topic or task, and the information-seeking stage, etc. The system should be able to suggest to the user appropriate ISSs at the appropriate times, given the current state of the information-seeking process.

HYPOTHESIS

Hypothesis 5: The integrated system purposely designed for supporting both scanning and

searching is more effective in supporting tasks requiring scanning and searching than the generic baseline system.

EXPERIMENTAL DESIGN

This will be a within-subject experimental design. The subjects will search the integrated system and then the baseline system. The experiment will be replicated by changing the order of the systems.

4 Conclusion

Our aim is to contribute to the field of interactive information retrieval drawing attention to the user interface design and HCI. The systems in research problem 1 have been implemented and the user studies were conducted. Future work will focus on the interaction structure and construction and testing of the integrated system. Through this we hope to improve information retrieval and human-computer interaction.

References

- Belkin, N. J., Marchetti, P. G., & Cool, C. BRAQUE: Design of an interface to support user interaction in information retrieval. 1993. *Information Processing & management*, 29(3): 325-344.
- Belkin, N.J., Cool, C., Stein, A., Theil, U., Cases, scripts and information seeking strategies: on the design of interactive information retrieval systems. 1995. *Expert Systems with Applications*, 9(3): 379-395.
- Belkin, N. J. Intelligent Information Retrieval: Whose Intelligence? 1996. In *Proceedings of the Fifth International Symposium for Information Science (ISI-96)*, 25-31.
- Borlund, P. & Ingwerson, P. The development of a method for the evaluation of interactive information retrieval systems. 1997. *Journal of Documentation*, 53(3): 225-250.
- GUTENBERG. <http://www.gutenberg.org/>
- HARD. <http://projects ldc.upenn.edu/HARD/>
- LEMUR. <http://www.lemurproject.org/>
- Oddy, R.N. Information retrieval through man-machine dialogue. 1977. *Journal of Documentation*, 33(1): 1-14.
- Xie, H. Shifts of interactive intentions and information-seeking strategies in interactive information retrieval. 2000. *Journal of the American Society for Information Science*, 51(9): 841-857.