Progress in search interfaces requires vigorous inquiry into how search features can be embedded into application environments such as those for decision-making, personal information collecting, and designing. Progress can be made by focusing on mid-level descriptions of how search components can draw upon and update workspace content and structure. The immediate goal is to advance our understanding of how to shape and exploit context in search. The long-term goal is to develop an interdisciplinary design resource that enables stakeholders in the computing, social, and information sciences to more richly impact each others’ work.

The Problem

Over a period of 3 years, while working at a large Internet company, I observed hundreds of consumers as they used a variety of Web products, including search, content destinations, e-commerce, and community products. My colleagues and I would often ask participants in our studies, “What do you like the most and the least about the Web?” The typical participant would reply, “Everything I need is available.” Later, after completing some self-directed search tasks, this same person would say, “It’s hard to find things.” Few, if any, participants were able to resolve this paradox.

For many, information immediacy is a defining characteristic of contemporary culture. With our computers and Internet connections, our perception is that all the information we need is nearby. Yet, we cope with information overload. As we slog through our information trails to find answers, we seem to encounter an enormous amount of nonrelevant material. Simon (1996), in Sciences of the Artificial, resolved this paradox by observing that our attention is limited: Information technology will not make the days longer nor will it speed-up our brains. Thus, the solution, Simon argued, is to focus on the suitability, quality, and presentation of information for the task at hand.

Today, search engines are remarkable for giving us access to information. Rarely, however, do they lead us to clear directions for finding answers; rather, they return starting points. As part of a study, I once visited a wealthy, semi-retired businessman. His computer had a very small screen with low resolution, $640 \times 480$, and a cable modem. At the end of each day, he would enter a stock symbol into a search box, press enter, scan the display for the closing price and volume, and enter this data with a pencil onto a paper grid. The businessman spent about one hour each day repeating this task for approximately 50 stocks that were on his “watch” list. This process of lookup-and-record was compelling to the businessman because it was a vast improvement over using the morning newspaper. When I demonstrated an online portfolio to automate the process of search, he was intrigued but not immediately enthusiastic. His caution may have been due mainly to the “quality time” he had to reflect upon his stocks while carrying out this repetitive sequence of lookups.

This report of an isolated search episode, of the most basic and uncomplicated type, illustrates three general points. First, the businessman seemed to consider search to be a narrow communication process as if he were using his telephone 50 times to ask an operator for quotes. Access to the data was sufficient. Second, the retrieval system had no connection with the businessman’s broader goal of tracking stocks. Once data was received, it was put to use in a decoupled environment. The system had no memory of past search episodes, and it had no representation for prospective events. Third, the system did not offer the businessman a path from conceptualizing search as “communication” (i.e., the search dialog) to conceptualizing search as a “workspace” or representation of his context (i.e., the portfolio). In short, it lacked adaptability.

Many cognitive and social models of the search process have been proposed to account for descriptions of such information seeking phenomena. Pettigrew, Fidel, and Bruce (2001) surveyed many of these models. The models that follow a cognitive perspective characterize search as a series of problem-solving phases in which the whole process is embedded in a context of rich communication and domain knowledge. These stage models, however, are ineffective for design, at least partly because they abstract away the external representations, search infrastructure, and domain features used in the search process. Indeed, Pettigrew et al. (2001) say, “For the field of information behavior, the challenge remains to provide concrete guidance for system design” (p. 68). One direction for meeting this challenge is...
to work at a level of description that lies between these general models of search, on the one hand, and rich but unconnected descriptions of search activity on the other. Two candidates for this mid-level of description are pattern languages for abstracting system regularities (Alexander, Ishikawa, & Silverstein, 1977) and user scenarios for abstracting situations of use (Carroll, 2000).

The Solution

Research in user interfaces for information seeking should seek to discover how search can be embedded within narrowly defined problem-solving contexts. In the 1990s, several research projects examined the idea of information workspaces (e.g., InfoGrid, Rao, Card, Jellinek, Mackinlay, & Robertson, 1992; SketchTrieve, Hendry & Harper, 1997; DLITE, Cousins, Paepcke, Winograd, Bier, & Pier, 1997; and NaviQue, Furnas & Rauch, 1998). Using direct manipulation interfaces, these information workspaces aimed to enable searchers:

1. To access a variety of different kinds of search technology and document types through unified presentation and interaction styles.
2. To represent work artifacts in the search interface.
3. To capture the history of search activity for retrospective analysis and to represent future search plans.
4. To annotate search activity.
5. To share search activity with others.

In short, these information workspaces sought to include many of the properties associated with spreadsheets. Rather than emphasize the communication aspects of search systems, these systems explored the idea that search interfaces can be environments for ongoing, reflective problem solving. Thus, for example, searches can be designed and reused within these environments and representations for prospective plans and features of the domain can be expressed. Incidentally, these research projects do not seem to have influenced the development of search interfaces for the Web.

Meanwhile, toward the end of the 1990s a variety of client-side search tools became available (e.g., Copernic—Copernic, 2003, and FirstStop WebSearch—Brush Systems Group, Inc., 2003). These commercial tools allow a person to search multiple search engines from the desktop and provide ways of processing the results. Searches can be saved, organized into folders, and scheduled to run periodically. Unfortunately, there appear to be no published studies in the research literature on how these kinds of tools are actually being used—this is certainly an area for research. In sum, while these research and commercial applications enable a fuller range of search activities than did their predecessors, they do so in a relatively domain-neutral fashion.

Example Scenario

Now, consider this scenario, which focuses on a very narrow context: the selection and annotation of references.

Work context. John needs to write a technical report. To support his work, he manages a list of references (see Figure 1A). EndNote (Thompson ResearchSoft, 2003) allows John to store, organize, and annotate references. John’s major motivator for using EndNote is that it is a convenient application for selecting, importing, and formatting references in reports. Indeed, EndNote is tightly coupled to his word processor.

Search. John uses a search interface to find references from multiple sources around the world (see Figure 1B). Scanning the results page (see Figure 1C), John selectively imports references into his list. Items already in the list are highlighted in the search results page; thus, John can readily identify duplicates. Authors in the result are also highlighted, allowing John to be sensitive to new and familiar sources.

Taking advantage of context. In the background, search agents are searching for authors’ home pages and upcoming conference home pages for papers to be published in proceedings. These attribute values are added to items in the reference list. In addition, a citation network is created. The visualization highlights existing sources in the network (see Figure 1D), allowing John to investigate new material and consider it for inclusion. With this visualization, John is able to identify new groups of authors for further investigation. An important side effect is that John learns about citation searching, a hitherto unknown technique.

This scenario emphasizes the workspace and how search functionality can improve the way people update, organize,
annotate, and prioritize items in the workspace. The scenario shows how a range of search functions can be embedded in the workspace to improve John’s experience. John’s reference material is enriched and, as a side effect, John learns of a new searching technique.

Using this very simple and familiar scenario, we can propose a general three-step pattern: (a) Enable users to create, revise, and annotate a structure (e.g., list of references); (b) design search to update and maintain this structure by treating it as the problem-solving context (e.g., highlight duplicates and add authors’ home pages); and (c) use search functionality to present alternative structures, which reveal different types of information (e.g., show reference list within citation network). This analysis suggests that when search is framed as a secondary aspect of a work context it may become, somewhat unexpectedly, more powerful.

Towards Best Practices

Progress in search requires that we discover ways to embed search components within content applications and within users’ work contexts. On the evidence of current search interfaces, this is not an easy task. We can make progress towards better search interfaces in the following ways:

1. Investigate how narrow, well-defined domains, such as information collection, decision making, and designing, can be represented as structured workspaces.
2. Develop scenarios that illustrate how search can be embedded in these structured workspaces.
3. Seek to show how models of information seeking behavior explain the scenarios.

This project of mid-level description of search functions, interactions, and structures should lead to a variety of possibilities for innovation and research. It aims to focus on users’ information seeking in a concrete yet flexible and general fashion. At the same time, it must make explicit connections with models of information behavior. Working at this mid-level may provide the foundation for an interdisciplinary design resource for improving search. To be successful, however, this effort must enable stakeholders in the computing, social, and information sciences to equally represent their questions, findings, and controversies.

References