Important Cognitive Components of Domain-Specific Search Knowledge

Suresh K. Bhavnani
School of Information
University of Michigan
Ann Arbor, MI 48109-1092
Tel: +1-734-615-8281
bhavnani@umich.edu

ABSTRACT

Many users have acquired a sophisticated understanding of searching the Web in specific domains. For example, we often hear of users who can get amazing deals for electronic products on the Web. What knowledge do such users have, and how does it affect their search behavior? To address this question, we observed information retrieval experts in the domains of healthcare and online shopping, while they performed tasks within and outside their domains of expertise. When performing tasks within their domains of expertise, experts used declarative and procedural components of domain-specific search knowledge that enabled them to perform effective searches. In contrast, when they performed tasks outside their domains of expertise, they used a range of general-purpose search methods leading to comparatively less effective search results. The study demonstrates the role of domain-specific search knowledge, and pinpoints its cognitive components. The paper concludes by suggesting approaches that should make the components of domain-specific search knowledge explicit and available to many users.

Keywords

Domain-specific search knowledge, information retrieval.

INTRODUCTION

Despite huge advances in making information accessible to vast numbers of users, the effective retrieval of relevant information remains a challenge. Numerous user studies of different information retrieval (IR) systems repeatedly show that despite knowledge of basic search techniques, many users do not acquire strategic knowledge to find relevant information effectively [8, 9, 11].

To address this problem, several studies have attempted to identify effective IR strategies [1, 2, 5, 14], in addition to understanding the complex processes involved in search [4]. For example, early studies identified the *Building Block* strategy that specifies to break searches into smaller queries and then build it back to a larger one. More recent work attempts to predict browsing behavior such as when users continue to search within a site, and when they move on to another site [4].

The focus of such studies is on understanding *domain-general knowledge*, which is useful to perform tasks in many domains, whether one is searching for prices of a digital camera, or searching for health-related information. However, while domain-general knowledge may be important, is it sufficient for effective search?

Prior research has shown that searchers with *subject* knowledge in a domain know how to select terms that make them effective information searchers within specific domains [18]. However, evidence from Internet surveys [15], and from everyday experience suggest that some users have acquired domain-specific *search* knowledge that goes beyond knowing the subject-specific terms to enter in a query. For example, many university students often buy electronic gadgets at bargain prices on the Web because they know which sites to visit for different products, and in what order. This knowledge therefore appears to have declarative and procedural components that need a closer examination. What are the cognitive components of domain-specific search knowledge that such users have, and how does it affect their search behavior within their domains of expertise?

This paper attempts to identify the components of domain-specific search knowledge used by IR experts while they perform searches on the Web. In an exploratory study, IR experts in the domains of healthcare and online shopping were observed while they performed tasks within and outside their domains of experience. A fine-grained analysis of their interactions (using problem behavior graphs) and post-task interviews revealed declarative and procedural components of domain-specific search knowledge that enabled them to perform effective searches within their domains of expertise. In contrast, when they performed tasks outside their domain of expertise, they used a range of general-purpose search methods leading to comparatively less effective search results. The study therefore demonstrates the critical importance of domain-specific search knowledge, and identifies the specific components of such knowledge.

The paper concludes by suggesting approaches that make domain-specific search knowledge explicit and available to users in order to assist them in performing effective searches in unfamiliar domains.

EXPLORATORY STUDY TO IDENTIFY DOMAIN-SPECIFIC SEARCH KNOWLEDGE

The goal of our study was to identify the cognitive components of domain-specific search knowledge and their effects on search behavior. We therefore focused on recruiting participants who were experienced searchers in one of two domains (but not both), and observed them perform tasks within and outside their domains of experience.

Five healthcare search experts were recruited from three medical libraries on the University of Michigan campus. All healthcare search experts had six or more years experience accessing medical information and used Web browsers on a daily basis. Similarly, four online shopping experts were recruited from the student and recently graduated student community. All had three or more years of experience in shopping on the Web.

The participants were asked to perform eight tasks in two domains: four tasks related to healthcare, and four tasks related to online shopping. The tasks were adopted from the Text Retrieval Conference (TREC)[1], organized by the National Institute of Standards and Technology. (The entire set of TREC tasks is available from http://www.itl.nist.gov/iaui/894.02/projects/t10i/guidelines.html, and those selected for detailed analysis in this paper are described in the next section.)

The tasks were randomized within each domain, as was the order between the domains. Participants were told to perform the tasks, as they would normally do for themselves. They were asked to think aloud while performing the tasks and were reminded to keep talking if they stopped. After each set of tasks within a domain, they were asked questions in a structured interview regarding how they performed the task. The protocols, interactions, and interviews were recorded using screen capture tools.

ANALYSIS AND RESULTS

Our analysis and results focus on the interactions of all nine participants each performing the following two tasks:

> *1. Tell me three categories of people who should or should not get a flu shot and why?*
>
> *2. Get two price quotes for a new digital-camera (3 or more megapixel and 2x zoom). Stop when you feel you have found the lowest prices.*

The above *flu-shot task*, and *camera task* were selected based on the following criteria: (1) Tasks in which experts in each domain had the most experience; (2) Tasks that took the longest average time for the experts in each domain. These criteria were designed to choose those tasks that took a long time for experts to complete despite their expertise. The *flu-shot task* was within the domain of expertise for the healthcare search experts, but outside the domain of expertise for the online shopping experts. The opposite was true for the *camera task*.

Searching within and outside domains of expertise

Our initial observations revealed a difference in behaviors between participants performing IR tasks within and outside their domains of expertise. For example, when an IR expert in healthcare performed the *camera task*, her overall approach was to use a general-purpose search engine to find websites for cameras and their prices. She began by using the search engine alltheweb.com through which she found the site megapixel.net. Here she found a camera meeting the task criteria but found no prices. She therefore continued searching for websites containing cameras and prices in alltheweb, and after three more queries found CNET.com (a price-comparison and product-review site). She then returned to megapixel.net to retrieve the name and model of the camera, and searched for its price in CNET.com, where she found two prices for the camera. She ended the task by picking the lower price ($619).

The above search method relies on general-purpose knowledge or "weak methods" [12] such as entering and modifying queries in a general search engines. While such methods are useful to perform tasks in a wide range of domains, they are not as powerful as methods that are tailored to a specific domain.

In contrast to the above approach, an online shopping expert performed the same task using less general but more powerful methods that were specific to online shopping. His overall plan consisted of: (1) identifying cameras and prices from sites that provide reviews and prices such as Epinions.com; (2) comparing prices across vendors through mySimon.com that specializes in price comparisons; (3) searching for coupons from techbargains.com that apply to online stores such as STAPLES.com. He repeated the last two steps until he found a low price for a camera ($389) with features that exceeded the task requirements. Throughout the task, he accessed websites by directly typing their addresses or uniform resource locators (URLs) in the browser address field, and never used search engines such as Google.

The shopping expert's search behavior exhibits knowledge components ranging from how to sequence stages of the search, to knowledge of specific URLs. Such domain-specific search knowledge allowed him to perform an effective search with high-quality results. To understand more precisely the nature of these domain-specific components, we codified the interactions and developed formal descriptions of the behavior.

Codification of Interactions

The interactions of nine experts each performing two tasks were analyzed as working in a problem space [13]. A problem space is defined by a set of states, a set of

[1] TREC is a premier IR conference that specifies tasks that are researched by all participants of the conference. The goal is to set a baseline comparison across different research groups. We piloted and adopted the TREC tasks and guidelines for our research.

Operator	Definition	Example	
		Formal Description	**Informal Description**
Find-websites	Searches for one or more websites using a search engine.	**Find-websites** (Engine=Google Query="Digital Camera")	User accesses Google to enter the query "Digital Camera".
Scan-websites	Scans one or more websites with the same goal(s).	**Scan-websites** (Camera-models=? Review=? Feature>=2X, 3 MP Price<$400 TU=Epinions.com, CNET.com)	User has the goal of finding camera models and their reviews that meet the criteria of 2 times zoom, 3 megapixel, and a price less than $400 by visiting Epinions.com and CNET.com. Both sites are visited by typing in the URL (TU).
Compare	Compares information within a website or across websites.	**Compare** (Price<526? Vendor=? Vendor-reputation=high Camera-model=Olympus C-3030 TU=PRICEWATCH.com)	User has the goal of finding a camera price lower than $526 sold by a vendor whose reputation is high for the camera Olympus C-3030. He visits PRICEWATCH.com by typing in the URL.
Verify	Verifies information already found	**Verify** (Confirm=? List=9 categories of people CL=WYETH.com)	User has the goal of verifying 9 categories of people [who should get a flu shot]. She does this by clicking on a link to visit WYETH.com.
End-task	Ends task either out of frustration, satisficing, or complete satisfaction. In some cases the experimenter informed the user that the task was completed.	**End-task** (Statement="That's the one to go for…I know this is a reputable site")	User has decided to buy a camera from a reputable site and ends the task.

Figure 1. High-level operators identified from the interactions and think-aloud protocols. Question marks denote a goal, TU and CL stand for *Type URL*, and *Click Link* [4]. The number and type of arguments for each of the above high-level operators vary depending on the interaction.

operators for moving between states, an initial state, a goal state, and a current state.

Identification of Operators

Through the analysis of protocols and interactions, we identified five operators that were sufficient to describe the search behaviors of the participants. Figure 1 shows definitions and examples for each of these operators. Because our goal was to analyze the overall search interaction, we defined operators at a higher level of abstraction than those used by Card et al., [4], in their development of similar graphs. We did, however, include their low-level operators as arguments of our high-level operators as different ways to visit websites. For example, the operator *Scan-websites* uses *Click Link* (CL), and *Type URL* (TU) to visit websites. Because the study by Card et al. restricted users to a single window, we added the argument *Click Existing Window* (CEW) to fully describe the behavior of our users.

Development of Problem Behavior Graphs

The operators were used to create descriptions of the search behaviors using problem behavior graphs (PBGs) [13]. Figures 2 shows an example of a PBG constructed from our data. The nodes of the graph (shown as boxes) represent knowledge states (knowledge that the user had acquired regarding the search task at a particular stage of the search process). The arcs of the graph (shown as arrows) represent the operators used to reach a particular knowledge state. Vertical lines represent backtracking to an earlier state (such as returning to the page of results in a search engine after following an unproductive link). Time therefore runs from left to right, and then from top to bottom. The information contained in the states contains the same arguments as those defined for the operators shown in Figure 1.

The PBG in Figure 2 describes the search behavior of an online shopping expert (S-1) who regularly shops for electronic gadgets on the Web. The PBG shows S-1's interactions at three levels. At the highest level of abstraction, his search had the following stages: *Review, Compare, Discount*.

At the next level of detail, the PBG describes the states of knowledge he passed through, in addition to the operators he used to move from one state to another. For example, within the *Review* stage, he scanned three sets of websites each with different goals (as shown by the operators between states 2, 3, and 4 in Figure 2).

Finally, the lowest level of detail is expressed by the arguments of the operators. These arguments show the user's goals, and the operator inputs. The states have the same arguments and represent the outputs of the operators. These arguments allow us to reconstruct the behavior of the user. For example, the first four states show how S-1 identified highly reviewed cameras and their prices to get a general understanding of the features available and their price range. He began with the goal of finding the model and price of a camera that has 2 times zoom, and 3 or more megapixel resolution (State-1). He then scanned Epinions.com looking for reviews of a camera less than $400 (based on prior knowledge of such gadgets) that meets

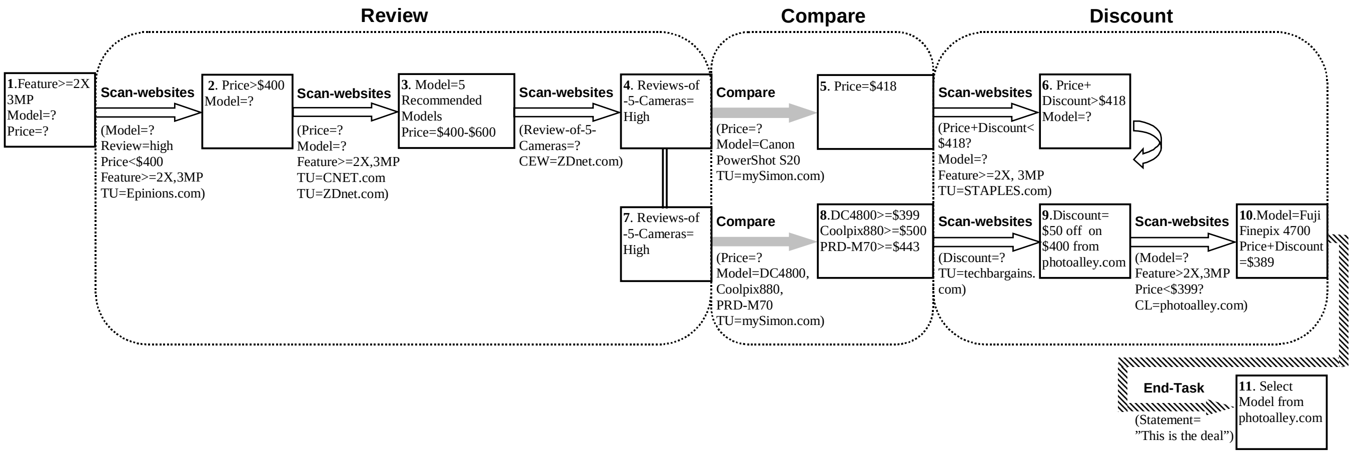

Figure 2. Problem behavior graph of S-1, an expert online shopper, performing the *camera task*. The graph shows his search behavior at three levels: (1) overall stages of the search method (Review, Compare, Discount); (2) knowledge states (boxes) and operators (arrows); (3) operator arguments specifying goals and interactions (text in boxes, and parentheses). The arguments TU, CL, and CEW stand for *Type URL*, *Click Link*, and *Click Existing Window* respectively as described in Figure 1. The back arrow represents the end of an abandoned search path, and the vertical lines represent a return to an earlier state.

those requirements but did not find any (State-2). Next, he scanned CNET.com and ZDnet.com to find a model meeting the same criteria and its price, and found 5 recommended models ranging in price from $400-$500 (State-3). Finally, he scanned ZDnet.com by clicking on an existing window in order to read reviews of the 5 models and found that all cameras had high ratings (State-4).

The PBG therefore makes salient, in a compact representation, the overall structure of search, in addition to the details of the interactions. Furthermore, the PBG provides a more precise understanding of the knowledge used during search. Such details are easily overlooked when analysis is based solely on observations or transcripts. Therefore, to identify the components of domain-specific search knowledge across the 9 participants, we developed and analyzed 18 PBGs in addition to analyzing transcripts of the interviews.

COMPONENTS OF DOMAIN-SPECIFIC SEARCH KNOWLEDGE

Analysis of 18 PBGs and transcripts of the interviews provided evidence for the declarative and procedural components of domain-specific search knowledge.

Declarative Components

Experts knew three types of declarative knowledge: (1) Classification knowledge consisting of classes of websites within a domain; (2) URL knowledge of specific websites; (3) Content knowledge consisting of the nature and type of information within a website.

Classification Knowledge

Search experts in both domains had well-formed classifications of websites within their domains. For example, the following quote is typical of a healthcare expert's classification of health-related websites:

> "I classify websites by the type of audience they are designed for, so if they are designed for the general population, consumers, and their families you know, that certainly is one big category. … And then sites like MEDLINE, which bring you – which take you deep into the journal literature and very specific."

Shopping experts similarly classified shopping sites into review, comparison, discount, and product sites. However, they had far less clarity in their classification of healthcare sites:

> "I think there are two main [categories], one is just providing general information, or answer to questions, or stuff like that, and one is just about like hospitals - how you can make appointments, and other stuff. That's not information about uh diseases, but how to get medical care, how to schedule appointments and stuff like that. Other than that, I have no experience on medical sites".

URL Knowledge

Experts knew specific URLs of collections and sources when performing tasks within their domains. For example, all the healthcare experts directly typed in the URL of MEDLINEplus.gov, a collection that indexes a large number of reliable healthcare sources for consumers. Similarly, online shopping experts knew the URLs of sources such as Epinions.com that contain reviews and prices of mainly electronic products. Because experts knew the URLs of most of the sites they visited, there were relatively fewer sites visited by clicking on links. However, the reverse was true when they performed tasks outside their domain. Figure 3 shows this relationship by comparing the occurrences of *Type URL*s (TUs) and *Click Link*s (CLs) for tasks within and outside their domains of expertise[2].

Content Knowledge

Besides knowing the URLs of sites, experts also knew the nature of information contained in the sites. For example,

[2] TUs and CLs that were used to revisit a site were not counted. The count is therefore a measure of the unique sites visited.

healthcare experts performed the *flu-shot task* by visiting 14 dot-org and dot-gov sites, compared to 3 dot-com sites that were commercial pharmaceutical websites that provided detailed information about side effects from the flu shot. The following exemplifies the healthcare experts' distinction between reliable government sources, and not so reliable sources that provide healthcare information:

> "I do not trust Adam.com. I would want other information to confirm that information before I would trust it. And I would be personally uncomfortable giving that information to a patient."

In contrast, the shopping experts performed the *flu-shot task*, by visiting 25 dot-com sites (out of a total of 29 total unique sites). Only 4 of those visited were dot-org and dot-gov sites.

Similarly online shopping experts knew which price comparison engines indexed small vendors ("mom and pop vendors") that typically had lower prices but were not that reliable. On the other hand, they also knew which price engines had more reputable vendors.

Procedural Components

While the declarative components provide critical concepts within a domain, procedural components provide the methods to use such concepts to perform effective searches. Our analysis revealed that experts knew two types of procedural knowledge: (1) Sequencing knowledge that allowed them to order classes of websites in an overall search plan; (2) Termination knowledge that specified when to end a search.

Sequencing Knowledge

As shown earlier in Figure 2, the online shopping expert S-1 performed the *camera task* by sequencing different classes of websites (Review, Compare, Discount) to perform an effective search. All the shopping experts used variations of this sequence. Figure 4 shows two other online shopping experts S-4 and S-3 who also used similar sequencing for the camera task. S-4 first scanned websites such as CNET.com to review cameras, then visited sites such as PRICEWATCH.com to review cameras and compare prices, and finally looked for coupons and discounts at sites like myCoupons.com. S-3 explicitly demonstrated only two of the stages by going to review sites like CNET.com, and comparison sites like mySimon.com. An analysis of the post-task interview revealed that S-3 had performed a quick mental calculation about a discount from amazon.com (based on his prior knowledge of discounts from Amazon), and decided that the discount was not worth it as the base price from that site was already too high. The *Review-Compare-Discount* sequence therefore appears to be an important domain-specific strategy for such online shopping tasks.

All the experts in healthcare also used sequencing knowledge. As shown in Figure 4, H-1 demonstrated the stages of first accessing a reliable collection like

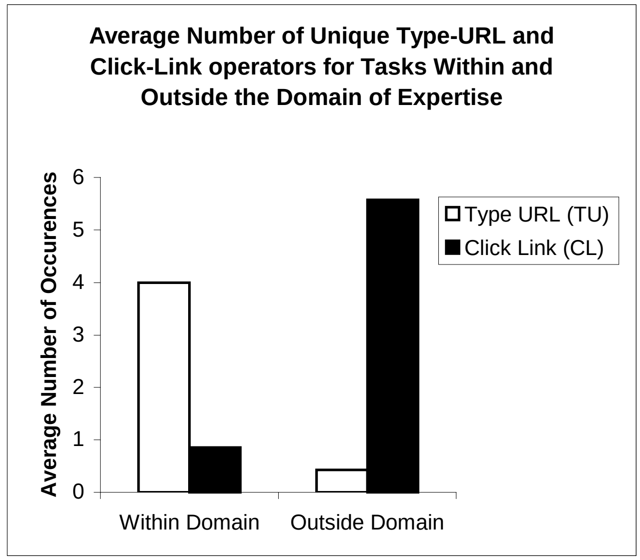

Figure 3. Reversal in the average number of TU and CU occurrences between search tasks within and outside domain of expertise. The differences between TUs and CLs are significant in both types of tasks ($p<0.05$ based on a two tailed t-test).

MEDLINEplus, and then finding sources for information within that collection. H-3 also followed these two stages but, in addition, verified the information by visiting a commercial pharmaceutical company that produces the flu vaccine. There she found a comprehensive and detailed list of people who should and who should not get a flu shot.

Such sequencing knowledge was absent in the search behavior of these very same experts when they performed tasks outside their domains. As shown by the solid black arrows (which represent finding a website by typing a query in a general-purpose search engine) in Figure 4, three participants used general-purpose search engines to perform tasks outside their domains. The healthcare expert H-1 aborted the task, as she could not find a price comparison engine through Google; H-4, with similar expertise, used the subject index in Yahoo to find digital cameras but because she did not look anywhere else for a price, ended up with an incorrect price due to an error in Yahoo's price of the camera. Online shopping expert S-4 used three general search engines (Google, AskJeeves, and YAHOO) to perform the *flu-shot task*, and S-3 used Google with many different queries to perform the same task. These variations demonstrate a continuum ranging from general and weak methods, to specific and strong methods as first described by Newell [12].

The absence of sequencing strategies (and the associated declarative knowledge) had a direct effect on the search results. The shopping experts found cameras on average at $60 less than the healthcare search experts. In addition, although the healthcare search experts did not explicitly search for more categories than required by the task, they found a comprehensive list of 9 categories of people who

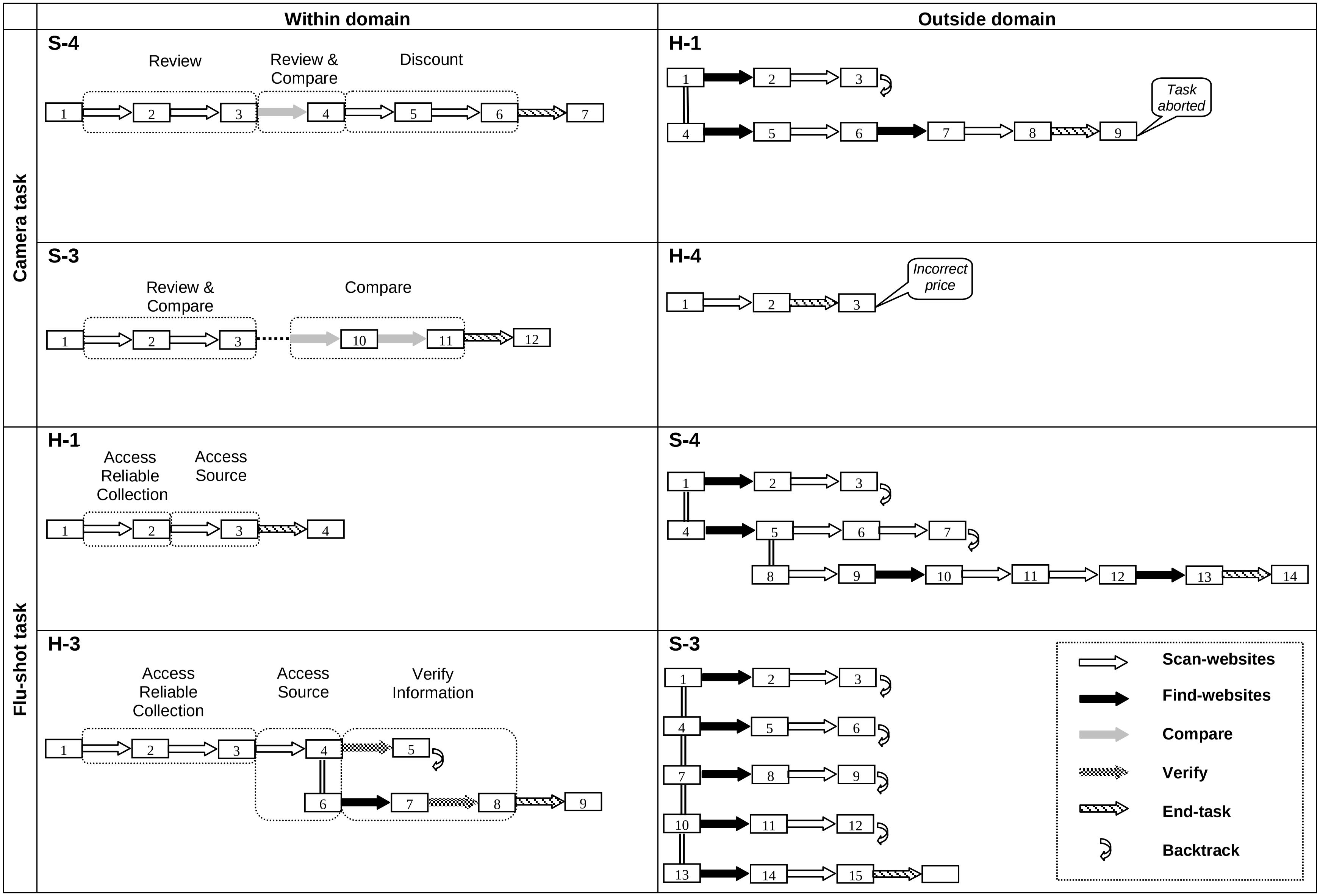

Figure 4. Eight problem behavior graphs of tasks performed within and outside domains of expertise. The above IR experts in shopping (S) and in healthcare (H) were chosen to show the variations we observed in sequencing knowledge. The within domain tasks have clear sequencing of classes of sites, whereas the tasks performed outside the domains of expertise are mainly performed using general-purpose search engines as shown by the solid black arrows.

should get a flu shot by going to an average of 3.7 reliable sites in very few steps. In contrast, none of the shopping experts found all the categories despite the fact that they visited on average 12 sites[3].

Experts, when performing tasks within their domains, therefore used sequencing knowledge to perform effective searches. However, these same experts used a range of weak methods to perform searches outside their domains leading to less effective search results.

Termination Knowledge

Experts in both domains demonstrated knowledge of when to terminate a search based on coverage. For example the shopping expert ended his search after checking three search engines:

> "So I had those prices to look at - three independent [price] search engines… It seems like a pretty good deal"

In contrast, all inexperienced shoppers terminated searches too early based on what appeared to be a perception of how long it would take to continue the search. None of them used more than one price comparison engine and were satisfied after finding a low price for just one camera. For example, after finding and comparing the price for a single camera a novice shopper stated:

> "Probably couldn't find anything lower… I wouldn't spend any more time fooling around with this, even for my own personal use".

DISCUSSION

The importance of domain-specific search knowledge should be no surprise to psychologists who have known the limited utility of domain-general knowledge or "weak methods" [6, 12] when solving non-trivial tasks within specific domains. These results should also be no surprise to librarians who have taken many classes during their education on sources and collections. However, we were still surprised to see the powerful effect such knowledge has on search behavior. One of the five reference librarians aborted the *camera task* because she lacked online shopping knowledge, while the other librarians performed largely ineffective searches compared to their expert counterparts. Similarly, the online shopping experts, who reportedly spend hours each day surfing the Web had strong skills in using browsers and search engines. However, despite this knowledge they performed relatively ineffective searches in a critical area such as healthcare. These results are therefore at odds with the growing confidence of large numbers of users who increasingly rely on general-purpose engines like Google to perform searches in many domains.

Appropriate to the nature of studies that do fine-grained analysis of interactions, our sample size is small. However we have found corroborative evidence for our results from a survey of Internet users [16]. This survey reports that most users, who search for health information on the Web, do so by searching directly through general-purpose search engines like Google. This is a critical problem given the prevalence of inaccurate, out-of-date, and often wrong information provided by a large number healthcare sites [3]. While this issue may not be as critical in online shopping, unreliable information can have serious consequences in the domain of healthcare. The identification and dissemination of domain-specific search knowledge should therefore be an important research effort in such domains.

There have been numerous attempts at organizing and making domain-specific knowledge available to users [7]. For example, sites like *InvisibleWeb.com* and *SearchEngineGuide.com* index numerous domain-specific search engines ranging from architecture to healthcare. However, such sites typically ignore the procedural components and the user is left to decipher first how to navigate the hierarchies to reach an appropriate search engine, and then how to sequence appropriate engines to perform a task. For example, we were unable to find reliable healthcare collections such as MEDLINEplus through either of the above sites, making it difficult to perform the *flu-shot task* effectively. Furthermore, even if we did find the appropriate search engines, there is no instruction on how to sequence them. Such tasks require much more knowledge than a search engine's URL.

Besides the above attempts to organize search engines across many domains, there are numerous portals at universities and organizations that provide lists of "helpful" links that are organized by expert librarians. For example, HealthWeb.org provides links to many different healthcare sources. Such sites provide the URL knowledge, but again do not make the procedural knowledge explicit to enable the selection and sequencing of sources and collections to perform specific tasks. The components of domain-specific search knowledge therefore help to pinpoint where such designs are lacking. This should lead to approaches to make the procedural components in addition to declarative components explicitly available to users so that they can make effective use of Web resources.

SUMMARY AND FUTURE RESEARCH

To understand the components and role of domain-specific search knowledge in information retrieval, we observed IR experts perform tasks within and outside their domain of experience. We found that experts were more effective when they used domain-specific search knowledge compared to when they had to rely only on domain-general knowledge. We also showed how this search behavior could be modeled adequately using problem behavior graphs at different levels of abstraction. Analysis of these problem behavior graphs and interviews helped us identify the declarative and procedural components of domain-specific search knowledge used by the participants.

Because current general-purpose search engines and portals do not provide all the components of domain-specific

[3] To enable a fair comparison, data from two participants who did not complete the entire task were not included in this calculation.

search knowledge, we are exploring approaches to make such knowledge explicit and available to users. Our identification of the components of domain-specific search knowledge should help to facilitate the design of structured interviews to elucidate the components in various domains. For example, very little is known about when experts decide to terminate a search, and structured interviews could enable the rapid elicitation of such knowledge from experts.

Once domain-specific search knowledge is made explicit, this knowledge could be made available to users through the design of classroom instruction. For example, there is an increasing need for healthcare professionals to understand how to retrieve accurate, reliable, objective, current, and comprehensive information [10] to make effective healthcare decisions. We are actively engaged in including domain-specific and domain-general knowledge in an IR course at the University of Michigan directed to freshman students entering the healthcare professions.

The components of domain-specific search knowledge could also be made available to users through new kinds of websites. For example, we are exploring the design of a new kind of site called a *Strategy Portal.* Such a portal will be designed to provide users classes of tasks within specific domains. For example, the online shopping domain would contain task categories such as “Shopping for an electronic gadget” and the healthcare domain would contain task categories such as “Prognosis for a disease”. When a user selects a particular task category, the system will provide a sequence of recommended search stages as an overall plan for that task. For example, selecting “Shopping for an electronic gadget” will produce the search stages: *Review, Compare, Discount.* Each stage will provide links to specific sites including tips on how to evaluate information, and how to terminate tasks. The goal of such efforts is to prevent what we saw all too often in our study: users typing terms like “flu shot”, and “digital camera” in Google and getting thousands of hits, and then ending searches more out of exhaustion rather than systematic coverage.

Our study has shown the importance of domain-specific search knowledge and helped to identify its components. Consequently, these results may help to design future systems and training, which should take into account many more components of search knowledge compared to what general-purpose search engines and portals provide today. This could lead users to become more effective when searching for information in unfamiliar domains.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation, Award# EIA-9812607. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of NSF or the U. S. Government. The author thanks P. Anderson, X. Lu, S. Mathan, M. O’Grady, P. Redman, F. Reif, and G. Vallabha, for their contributions.

REFERENCES

1. Bates, M. The design of browsing and berrypicking techniques for the online search interface. *Online Review* 13, 5 (1985), 407–424.
2. Belkin, N., Cool, C., Stein, A., and Thiel, U. Cases, scripts, and information-seeking strategies: on the design of interactive information retrieval systems. *Expert Systems with Applications* 9, 3 (1995), 379-395.
3. Biermann, J.S., Golladay, G.J., Greenfield, M.L. and Baker, L.H. Evaluation of cancer information on the Internet. *Cancer*, 86, 3 (1999), 381-90.
4. Card, S.K., Pirolli, P., Van Der Wege, M., Morrison, J.B., Reeder, R.W., Schraedley, P.K., and Boshart, J. Information scent as a driver of Web behavior graphs: results of a protocol analysis method for Web usability, in *Proceedings of the CHI' 01* (2001), 498 - 505.
5. Drabenstott, K. Web search strategies. In: *Saving the User’s time through subject access innovation*; Papers in honor of Pauline Atherton Caochrane, W. Wheeler (ed.) Champaign, Ill.: University of Illinois (2000), 114–161.
6. Barr A., & Feigenbaum, E. *The Handbook of Artificial Intelligence* (v. 2), William Kaufmann, Los Altos, CA, 1982.
7. Ipeirotis, P., Gravano, L., and Sahami, M. Probe, Count, and Classify: Categorizing Hidden Web Databases in *Proceedings of the 2001 ACM SIGMOD* (2001), 67 – 78.
8. Jansen, B.J., Spink, A., and Saracevic, T. Real life, real users, and real needs: A study and analysis of user queries on the web. *Information Processing and Management*, 36, (2000) 207-227.
9. Kirby, M., and Miller, N. MEDLINE searching in Colleague: Reasons for failure or success of untrained end users. *Medical Reference Services Quarterly* 5 (1986), 17–34.
10. Kapoun, J. Teaching Undergrads Web Evaluation: A Guide for Library Instruction. *College & Research Libraries News* 59 (1998).
11. Marchionini, G. Information seeking in electronic encyclopedias, *Machine-Mediated Learning*, 3, 3 (1989), 21-26.
12. Newell, A. Heuristic programming: ill-structured problems. In J. Aronofsky, (ed.), *Progress in operations research*, New York: Academic Press (1969), 361-414.
13. Newell, A., and Simon, H. *Human Problem Solving*. Prentice Hall, New Jersey, 1972.
14. O'Day, V., and Jeffries, R. Orienteering in an information landscape: how information seekers get from here to there, in *Proceedings of INTERCHI '93* (1993), 438-445.
15. Pastore, M. Discounts, marketing practices working against e-tailers. Available at http://cyberatlas.internet.com/markets /retailing/article/0,,6061_809501,00.html. (July 26, 2001).
16. Pastore, M. Online health consumers more proactive about healthcare. Available at http://cyberatlas.internet.com /markets/healthcare/article/0,,10101_755471,00.html. (April 30, 2001).
17. Pirolli, P., and Card, S. K. Information Foraging. *Psychological Review*, 106 (1999), 643-675.
18. Shute, S., and Smith, P. Knowledge-based search tactics. Information Processing & management, (1993) 29, 1, 29-45.