Separating the Knowledge Layers: Cognitive Analysis of Search Knowledge Through Hierarchical Goal Decompositions

Suresh K. Bhavnani
School of Information, University of Michigan, Ann Arbor MI, 48109-1092
Email: bhavnani@umich.edu

Marcia J. Bates
Department of Information Studies, Graduate School of Education and Information Studies, University of California at Los Angeles (UCLA), Los Angeles, CA 90095-1520
Email: mjbates@ucla.edu

Hierarchical goal decompositions have proved to be a useful method to make explicit the knowledge required by users to perform tasks in a wide range of applications such as computer-aided drafting (CAD) systems. This analysis method progressively decomposes a given task starting from the task layer on the top of the decomposition, to the keystroke layer at the bottom. The analysis enables a close inspection of the knowledge required to perform the task at each layer of the decomposition. In this paper we show how the method of hierarchical goal decomposition can be used to understand more precisely the knowledge that is required to perform information search tasks. The analysis pinpoints: (1) the critical strategies in the intermediate layers of knowledge that are known by experts searchers; (2) why such knowledge is difficult to acquire by novice searchers; (3) how the analysis provides testable predictions of behavior based on the acquisition of different types of knowledge. We conclude by discussing the advantages provided by hierarchical goal decompositions, and how such an approach can lead to the design of systems and training.

Introduction

Numerous studies have attempted to analyze the nature of search knowledge that searchers have acquired. These studies have used a variety of methods to identify effective strategies including self-reflection (e.g. Bates, 1979; Drabenstott, 2000), systematic observations of experts performing complex tasks (e.g. Fidel, 1991; Xie, 2000), expert-novice comparisons to understand differences in search knowledge (e.g. Holscher & Strube, 2000; Hsieh-Yee, 1993; Lazonder et al., 2000; Sutcliffe et al., 2000; Shute & Smith, 1993), and analysis of query logs of undifferentiated users to understand broad general trends (e.g. Jansen et al., 2000).

While these studies have begun to shed light on many different facets of search knowledge, there is a surprising lack of cumulative research leading to a deeper understanding of search phenomena from a cognitive perspective. For example, numerous studies continue to report the difficulty of acquiring effective search strategies in a wide range of IR systems. However, such studies have not led to a deeper understanding of why effective search is so elusive, and how to address the problem in a systematic way.

This is in sharp contrast to the rapid developments in human-computer interaction (HCI) that have developed and refined powerful and general representations to explain psychological phenomena ranging from the essential knowledge required for the transfer of knowledge (Singley & Anderson, 1989), to explanations of why efficient strategies to use complex computer applications such as computer-aided drafting (CAD) systems are difficult to acquire (Bhavnani & John, 2000). We believe such cumulative developments have benefited from the use of well-accepted methods of cognitive analysis that focus on making the knowledge to perform such tasks explicit.

In this paper we argue that similar to the field of HCI, the field of library and information science (LIS) should equally benefit by the use of analysis methods that require explicit descriptions of the knowledge involved in search competence. We demonstrate the benefits of explicit descriptions of knowledge by focusing on one such approach called *hierarchical goal decomposition*. This approach is an integral part of task analysis techniques such as hierarchical task analysis (Kirwan & Ainsworth, 1993), and GOMS (Goals, Operators, Methods, and Selection Rules) first demonstrated by Card et al. (1983), and more recently reviewed by John & Kieras (1996). The method of hierarchical goal decomposition exploits the fact that goals can be described at different levels of detail depending on what is being analyzed. In this paper we show the utility of using this method to progressively decompose a given task starting from the task layer at the top of the decomposition, to the keystroke layer at the bottom. The analysis enables a close inspection of the knowledge required to perform the task at each layer of the decomposition. The hallmark of this method is that it separates the layers of knowledge at different levels of abstraction, revealing how tasks map to high-level strategies, and how those strategies map to the actual operations to complete a task in particular system.

We begin by describing the steps to perform an effective strategy to perform a CAD task. We argue that while this description is sufficient to explain the procedure to use the strategy, it cannot explain why such a strategy is difficult to acquire despite many years of experience using CAD systems. In contrast, we show that a hierarchical goal decomposition of the CAD task provides a cognitively informed explanation to address this question.

We then show how the same method of analysis can be used to study tasks involving the search for information on the Web. The analysis reveals: (1) the critical strategies in the intermediate layers of knowledge that are used by search experts; (2) why such knowledge is difficult to acquire by novice searchers; (3) how the analysis provides testable predictions of behavior based on the acquisition of different types of knowledge. We conclude by discussing how insights derived from the hierarchical goal decomposition can lead directly to the design of training, and to new forms of websites that make such knowledge available to large numbers of users.

Analysis of a CAD strategy

Consider the task of drawing three identical arched windows in a CAD system. As shown in Figure 1A, one way to perform this task is to draw all the arcs across the windows, followed by drawing all the vertical lines, followed by drawing all the horizontal lines. This strategy can be called *Sequence-by-Operation.* An alternate way to do the same task (as shown in Figure 1B) is to draw all the elements of the first shape (Detail), group these elements (Aggregate), and then to make multiple copies of the aggregate to create the other shapes (Manipulate). Both these methods allow a user to complete the task. Such *non-obligatory* and *goal-directed* methods have been called "strategies" (Bhavnani & John, 2000; Siegler, & Jenkins; 1989). The *Sequence-by-Operation* and *Detail-Aggregate-Manipulate* methods described above are prime examples of strategies that can be used in complex computer systems.

The critical difference between the above two strategies is that the Detail-Aggregate-Manipulate strategy exploits the iterative power of the computer through aggregation commands. In order to exploit this capability, the user must complete drawing all the elements of the first window before grouping them and copying them. By grouping before applying operations, the user exploits the iterative power of the computer because the computer performs the iteration over all the elements in the group. In contrast, the Sequence-by-Operation strategy in Figure 1A does not exploit this capability as the user performs the iteration by drawing each element.

Several studies have shown that strategies such as Detail-Aggregate-Manipulate: (1) can save time and reduce errors (Bhavnani & John, 1996, 1997, 1998; Bhavnani et al., 1999; Nilsen et al., 1993); (2) are often not acquired by many users even after many years of experience of using commands in an application (Rosson, 1983; Nilsen et al., 1993; Bhavnani et al., 1996).

Reasons for this difficulty are not apparent by the description of the strategy in Figure 1B. After all, most users of computers know how to group many objects, in addition to knowing how to copy them. Still, this strategy is often not used. To understand why strategies such as Detail-Aggregate-Manipulate are difficult to acquire, we need a more precise understanding of the knowledge underlying its execution. One way to do that is through a hierarchical goal decomposition of a specific CAD task using the Detail-Aggregate-Manipulate strategy.

Hierarchical goal decomposition of the CAD task

Figure 2 provides a cognitive analysis of the knowledge required to perform the 3-window drawing task using the Detail-Aggregate-Manipulate strategy. The left part of the figure shows the goal decomposition in four layers of knowledge: (1) the *task layer* that describes the task

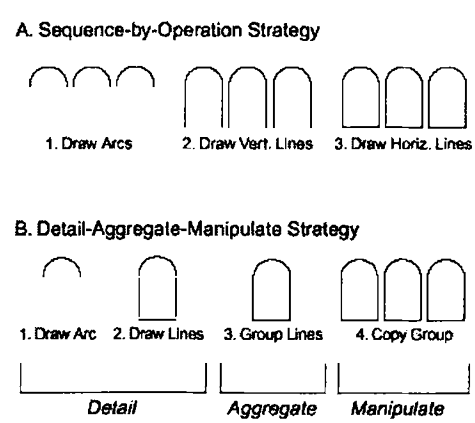

Figure 1. Two strategies to perform the 3-window drawing task.

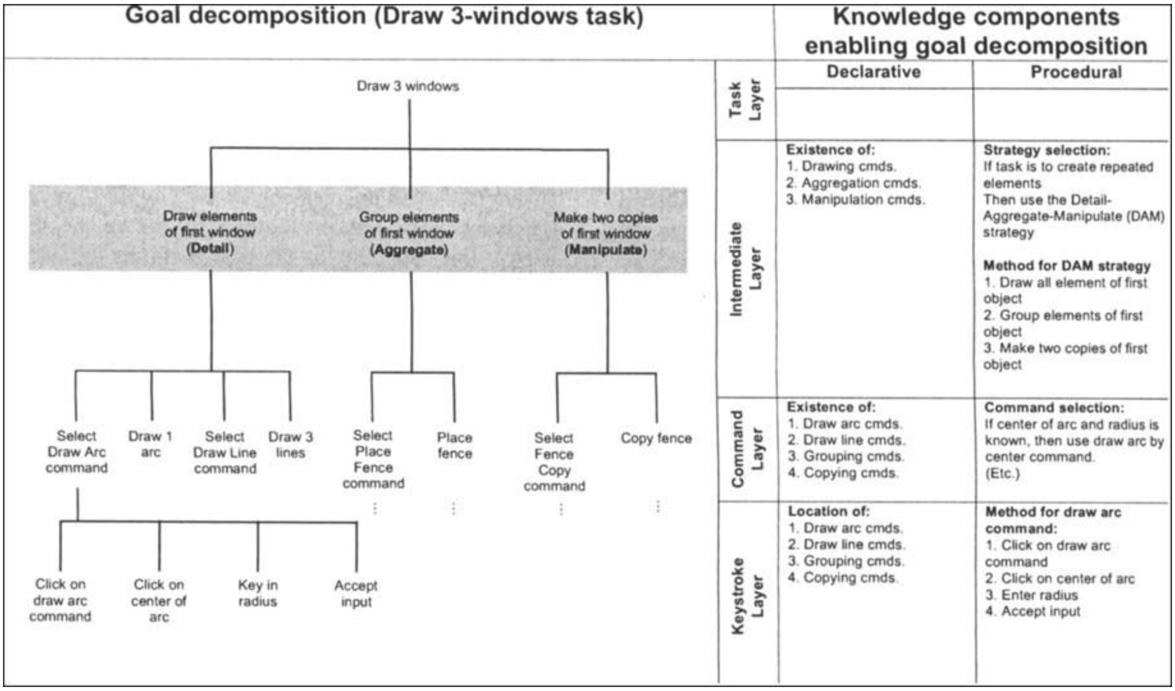

Figure 2. Cognitive analysis of the 3-window CAD task using the Detail-Aggregate-Manipulate strategy. The left part of the figure shows the hierarchical goal decomposition of the 3-window drawing task with the intermediate layer of knowledge marked in grey. The right part of the figure shows examples of the procedural and declarative components of knowledge at each layer of the decomposition that are required to complete the task.

performed by the user; (2) the *intermediate layer* that decomposes the task with knowledge of how to organize the different kinds of commands provided by the system; (3) the *command layer* that decomposes each of the stages in the intermediate layer into specific commands; (4) the *keystroke layer* that specifies the operations needed to execute the commands.

The right part of Figure 2 specifies two types of knowledge components required at each layer of the decomposition: (1) *declarative* knowledge that refers to knowledge of facts and relationships; (2) *procedural* knowledge that refers to knowledge of the steps needed to select and execute different methods.

As shown in Figure 2, the task layer is decomposed into the intermediate layer of knowledge that contains the three steps of the Detail-Aggregate-Manipulate strategy. This decomposition is critically dependent on declarative and procedural components of knowledge. For example, the user must have declarative knowledge of the existence of different classes of commands such as grouping, and copying commands. Furthermore, the user must know the procedural knowledge of how to sequence these classes of commands to complete the task. In this case, it is to first complete drawing all the components of the first window, followed by grouping those elements, and only then to copy the first window to create the other two.

Below the intermediate layer of knowledge is the command layer. This layer contains knowledge of how to decompose each of the nodes in the intermediate layer, into nodes representing the use of specific commands. This layer also requires declarative knowledge such as the existence of specific CAD commands like the Draw Arc command, in addition to the procedural knowledge of how to select between different types of commands.

The nodes in the command layer are finally decomposed into nodes in the keystroke layer, which represent the observable motor actions of a user (such as the clicks of the mouse, and key-ins on the keyboard) to perform the entire task. This layer requires declarative knowledge of where the commands are located, and the procedural knowledge of executing such commands.

An analysis of the knowledge in the different layers of the decomposition reveals that they contain qualitatively different types of knowledge. For example, knowledge to use commands is qualitatively different from knowledge to decompose a task in order to efficiently use command categories provided by the system. Furthermore, the decomposition reveals the relative independence of acquiring the procedural knowledge in the intermediate layer. A user might know how to select and use commands (knowledge in the command and keystroke layers), but this knowledge does not lead users to spontaneously acquire the Detail-Aggregate-Manipulate strategy. As described earlier, there is nothing preventing a user from using the same command knowledge to perform the task using the Sequence-by-Operation strategy. The procedural knowledge to decompose the task in the intermediate layer is therefore neither acquired spontaneously by knowing the task (top layer), nor by knowing how to select and use commands in the command and keystroke layers.

This explanation has led to the testable hypothesis that strategies in the intermediate layer of knowledge are difficult to acquire spontaneously just by knowing commands, and therefore such strategies must be *explicitly* taught. This hypothesis has been extensively tested with different populations of students (Bhavnani et al., 2001). The experiments have helped to pinpoint those strategies that are indeed difficult to acquire, and those that can be automatically learned just from learning commands. This research is leading to a deeper understanding of how to teach the strategic use of complex authoring applications more effectively.

The method of hierarchical decomposition therefore provides two distinct advantages over simpler strategy descriptions such as the one shown in Figure 1. First, it provides a deeper understanding of why critical strategies are neither spontaneously acquired from knowledge of tasks, nor from knowledge of tools. Second, it provides a hypothesis that can be tested by controlling the knowledge that is taught at different layers of the decomposition.

While the goal decomposition method has provided valuable insights for understanding the learnability of efficient strategies in the use of authoring applications, can it provide the same advantages in the analysis of search behavior?

Analysis of search strategies

We illustrate the value of hierarchical goal decompositions by analyzing the behavior of an expert medical reference librarian searching for healthcare information on the Web. This task was performed as part of an exploratory study (Bhavnani, 2001, 2002) that focused on the identification of effective search strategies.

The above expert was observed while she performed the following task: *Tell me three categories of people who should or should not get a flu shot and why?* To complete the above task, the reference librarian first accessed the reliable[1] healthcare collection called MEDLINEplus[2]. Next, she used the query "flu shot" to search within MEDLINEplus and got several hits. She visited two of those links and retrieved categories of people who should and should not get a flu shot. Not being satisfied with the sources she had visited, she retrieved the name of a flu shot ("Flushield") from a third link because she had the explicit goal of verifying the information she had obtained from MEDLINEplus.

She then attempted to verify the information obtained through MEDLINEplus by visiting a pharmaceutical company that sold Flushield. She did this by first visiting rxlist.com (a reliable site for drug-related information) but failed to find the information. She then used Google[3] to find the pharmaceutical company that sold Flushield (wyeth.com), and verified the information she had obtained from MEDLINEplus by reading the indications and contraindications for the vaccine.

Although she did not explicitly search for more categories than required by the task, she completed the task by having access to a comprehensive list of 9 categories of people who should get a flu shot, and 5 categories of people who should not. She took approximately 7 minutes to complete the task, and visited 3 sources for medical information, all of which were reliable.

When the same task was given to a user with equivalent experience in searching for information on the Web, but no experience in searching for healthcare information, he did not exhibit the strategy used by the expert. Instead, he went directly to Google and typed in "who should or should not recieve (sic) flu shots" in the query box. He visited numerous hits provided by Google *in roughly the same order of the displayed hits.* He took approximately 20 minutes to complete the task during which time he used a total of 5 queries, and visited 13 sites (plus two dead links), none of them high-quality healthcare sites recommended by Consumer and Patient Health Information Section[4] (CAPHIS). The strategy of relying on Google led him to retrieve fragmented pieces of information from a variety of unreliable sites in almost thrice the amount of time.

The above task descriptions do not provide an explanation of why the novice did not acquire the strategy known by the expert, despite having many years of experience in using search engines and browsers. To

[1] Reliability was determined by the presence or absence of a site in the recommended list of the Consumer and Patient Health Information Section (CAPHIS) of the Medical Library Association that analyzes and lists trustworthy healthcare sites.

[2] http://www.medlineplus.gov

[3] http://www.google.com

[4] http://caphis.mlanet.org/consumer/consumerAll.html

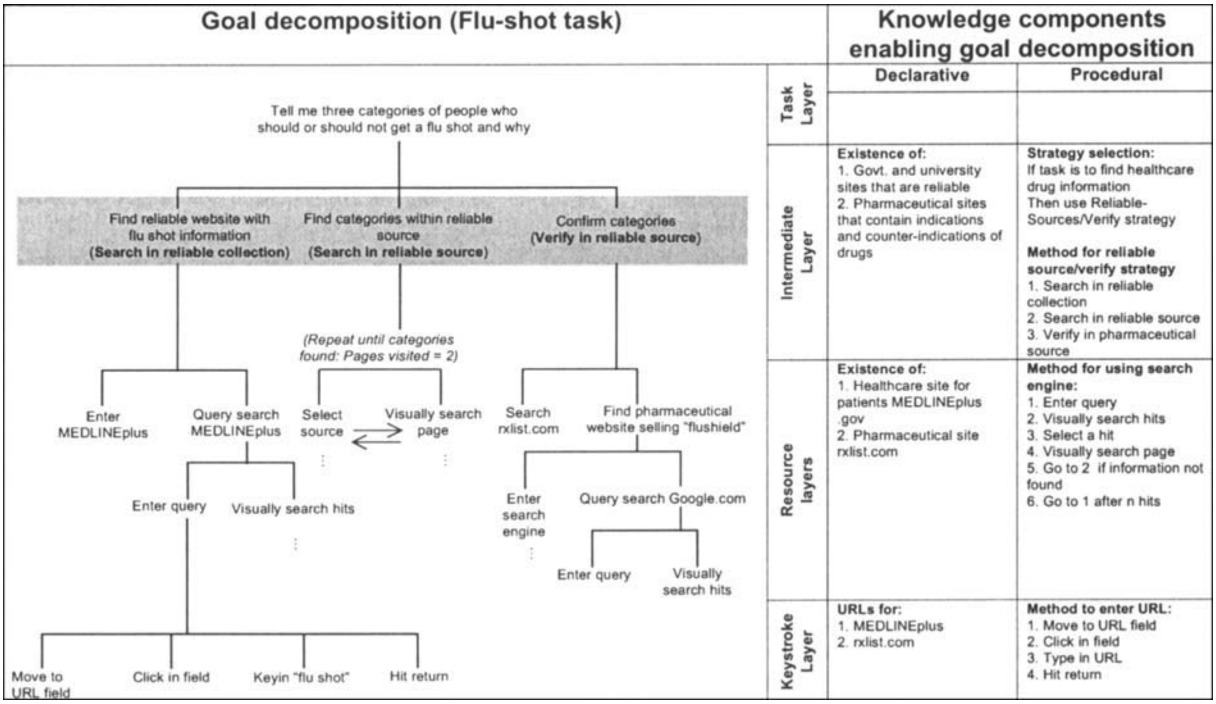

Figure 3. Cognitive analysis of how an expert performed the flu-shot task. The left part of the figure shows the hierarchical goal decomposition of the flu-shot task with the intermediate layer of knowledge marked in grey. Horizontal arrows represent how the user revisited nodes in the hierarchy described by the italicised text above. The right part of the figure shows examples of the procedural and declarative components of knowledge at each layer of the decomposition that are required to complete the task.

understand this, we performed a hierarchical decomposition of the task, and analyzed the knowledge components at each layer of the decomposition.

Hierarchical goal decomposition of the flu-shot task

Figure 3 shows a hierarchical goal decomposition of the flu-shot task. Similar to the goal decomposition of the CAD task shown in Figure 2, this task is also decomposed into the intermediate, command, and keystroke layers.

The intermediate layer of knowledge contains the critical strategy that the healthcare expert used. To use this strategy the user must know the distinction between reliable healthcare collections such as those sponsored by governments and universities, and unreliable sites such as personal pages. Furthermore, the user must also know that pharmaceutical sites provide information about indications and contraindications about drugs, knowledge that is useful to verify who should or should not take a particular drug. Finally the user must know how to sequence this declarative knowledge into a procedure: First search in a reliable collection, then select only reliable sources within that collection, and then verify that information through a reliable pharmaceutical source.

The resource layer (similar to the command layer in Figure 2) contains the knowledge required to search within specific websites such as in MEDLINEplus and in Google. This knowledge includes knowledge of the existence of specific sites, and the procedure of searching within the sites (e.g. enter a query, review the hits, etc.). Finally, the keystroke layer provides the motor actions to interact with the computer such as keying in the URL of a site.

The decomposition reveals the difficulty of obtaining knowledge in the intermediate layer of knowledge. Similar to the CAD task in Figure 2, neither the task description, nor knowledge of the existence of specific sites such as MEDLINEplus in the resource layer, provide the critical declarative and procedural knowledge required by the strategy. It is crucial to understand that knowledge of specific sites such as MEDLINEplus is also *precisely* what search engines like Google provide. Such engines are *not* designed to provide the declarative knowledge of which sites are reliable, nor are they designed to provide the procedural knowledge of how to sequence the different stages of the strategy in the intermediate layer. Users who rely on Google therefore can obtain knowledge represented in the resource layer, but are not provided knowledge in the

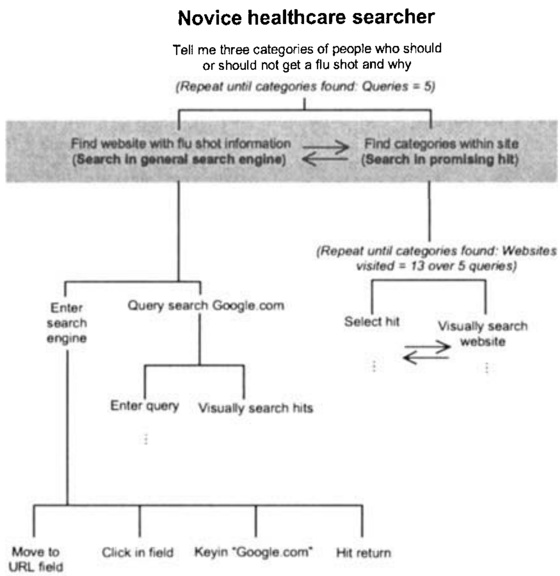

Figure 4. Hierarchical decomposition of the flu-shot task as performed by a novice. The intermediate layer of knowledge marked in grey, and horizontal arrows represent how the user revisited nodes in the hierarchy described by the italicised tex above.

intermediate layer. This knowledge is also difficult to infer just by visiting such sites.

This prediction was confirmed by the behavior of the novice performing the flu-shot task described earlier. The novice had many years of experience in using Google, but not much experience in searching for healthcare information. Figure 4 shows the goal decomposition of this novice searcher performing the flu-shot task. As shown, his intermediate layer of knowledge contains the strategy of using a general-purpose search engine. Because Google does not provide the strategic knowledge in the intermediate layer, he cannot infer the more sophisticated strategy demonstrated by the expert.

The above explanations lead to the prediction that despite experience with using general-purpose search engines, users may never discover strategies such as those known by the expert reference librarians. The knowledge to use such strategies needs to be *explicitly* taught. Furthermore, the decomposition and associated explication of the declarative and procedural components pin-points the knowledge that needs to be transmitted to users. Such explicit modeling can lead directly to testable educational interventions such as those that have been done to teach authoring applications (Bhavnani et al., 2001).

The above prediction is of course not unique to the flu-shot task. We have observed the same phenomenon for other search tasks as well. We observed an expert with many years of experience shopping for electronic gadgets while he searched for price information related to a new digital camera. His overall strategy was to: (1) identify cameras and prices from sites that provide reviews and prices such as Epinions.com; (2) compare prices across vendors through sites such as mySimon.com that specialize in price comparisons; (3) search for coupons from sites such as 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.

Figure 5 shows the goal decomposition approach to analyze his search strategy. The decomposition shows once again how the strategy in the intermediate layer of knowledge is difficult to spontaneously acquire from knowledge lower down in the decomposition such as knowing how to search within a site, or knowing the existence of sites such as Epinions.com that Google would provide. As can be predicted, a novice with little experience in searching for price information did not discover this strategy despite being proficient in using Google. She used exactly the strategy shown in Figure 4.

The difficulty of acquiring strategies in the intermediate layer of knowledge is therefore a general phenomenon occuring in computer applications ranging from authoring applications to information retrieval systems on the Web. The reason why this problem occurs is also explained by using the general approach of hierarchical goal decomposition that provides a cognitively based explanation of this phenomenon. The representation also provides the impetus to design and test different approaches to make such knowledge available to large numbers of users.

Discussion and implications for future research

The objective of this paper has been to present a method that is suitable as a basis for testing and research on the cognitive aspects of information searching. As Ingwersen has stated:

> The task of information retrieval (IR) and IR systems design is to bring cognitive structures of authors, systems designers and indexers into accord with those of the information worker, and the user—at the *event of searching*. (Ingwersen, 1992, p. 39)

Much work has been done in information science on both the design of information retrieval systems thought to be optimally supportive of the cognitive processes of searching (Bates, 1989, 1990, 1994, 2002; Belkin et al., 1993; Chen & Dhar, 1991; Ingwersen, 1992, 1996; Pejtersen, 1984; and others), and on the cognitive processes involved in searching itself (Bates, 1979; Belkin et al., 1982; Ellis, 1989; Ingwersen, 1996; Kuhlthau, 1993; Marchionini, 1995; Saracevic et al., 1988; Xie, 2000), though, on the

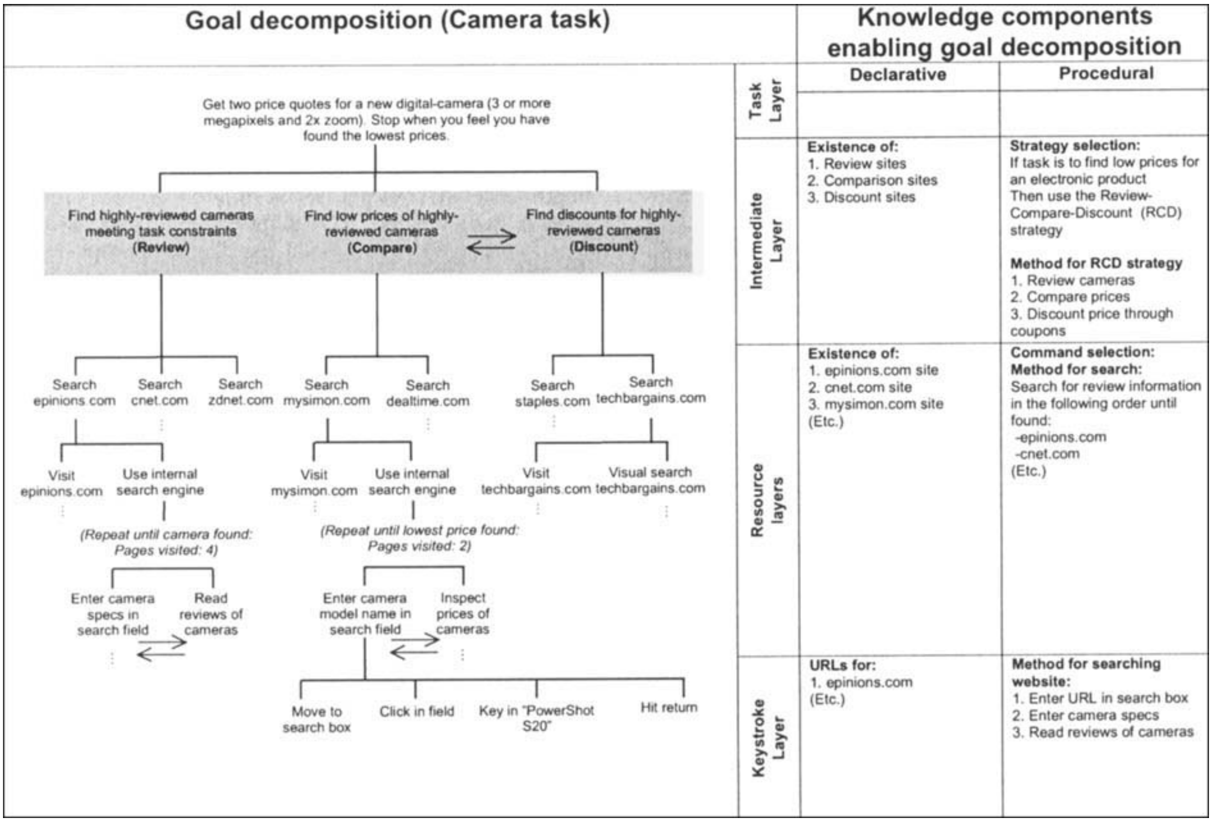

Figure 5. Cognitive analysis of how an expert performed the camera task. The left part of the figure shows the hierarchical goal decomposition of the camera task with the intermediate layer of knowledge marked in grey. Horizontal arrows represent how the user revisited nodes in the hierarchy described by the italicised text above. The right part of the figure shows examples of the procedural and declarative components of knowledge at each layer of the decomposition that are required to complete the task.

whole, probably less has been done on the latter. The goal decomposition method has been demonstrated to have great value in extensive cognitive science literature, where it has been shown to be an excellent approach for understanding a variety of human activities. In this paper we have demonstrated that the goal decomposition approach can also be effectively applied to the analysis of information search strategies.

The application of this method to analyze search strategies has a number of advantages that should enable research on the topic to move forward rapidly in the future. Research and theory on search strategy in information science has been fragmented, and has seldom cumulated in such a way that we could build on past discoveries. By drawing on a well-established psychological method, we may finally be able to begin that cumulative effort.

There are a number of features of this analysis method that may be particularly productive for research in search strategy and other cognitive functions in information science. It is important to recognize that the four levels of the hierarchical decomposition are each qualitatively different; each represents distinct and specifiable aspects of human cognitive processing around any given activity. By analyzing search using these levels, we may be able to gain a better understanding of the sticking points in learning how to search, and of the specific kinds of learning and operating that must take place to complete a successful search. Furthermore, analyzing and distinguishing declarative and procedural knowledge at each level permits a much more rigorous analysis of just what searching consists of, and just what kinds of learning is required of the successful searcher.

Another promising direction for research is to distinguish the declarative and procedural knowledge for subject content, from the declarative and procedural knowledge for information structure and organization. Thus, there might conceivably be four columns of knowledge types in the right part of the goal decompositions: Declarative and procedural subject knowledge (first and second columns) and declarative and procedural information structure and organization knowledge (third and fourth columns). Such structuring would illustrate the uniquely information-skills-related knowledge that an effective searcher must also have, along with some content knowledge of a subject area (Cf. Bates, 1999).

We believe that further interesting work can be done on the intermediate layers in relation to information searching. There are additional sub-activities within the intermediate layers that need to be unpacked. What are the choices that the searcher makes within a given information resource? In the flu-shot example, MEDLINEplus contains several on-screen options for the searcher from the home page. What strategies does the searcher possess with regard to this site, or like sites, and how are search moves chosen?

We have emphasized the intermediate layer of analysis, because it is at this layer where the search strategy must develop. For a query of even modest complexity, a search strategy is not obvious from the nature of the task, nor is it obvious from the design of the resources that may solve the task query. The searcher thus must construct a path between the task and the resources. The path can be hapless and random, where the searcher has few related conceptual resources to draw upon. The path can be successful eventually, but slow and wasteful along the way. Or, the path can be effective and efficient, with considerable generalizability to other comparable situations.

By conceptualizing information searching in this manner, we are in a position to study: (1) the stages of strategy development, as the searcher gradually improves the directness and power of the strategy; (2) general differences between expert and novice searchers; (3) the kinds of strategies that are particularly hard or easy to develop; (4) how people move from the task statement to strategy formulation, and from strategy formulation to resource selection and use; (5) optimal levels of generality of strategies for various kinds of information needs, and other testable hypotheses arising directly from the analysis of the decomposition.

A deeper understanding of these various aspects of searching could also dramatically improve training and design of educational materials and programs. For example, the last point regarding levels of generality in searching behaviors goes to the heart of the nature of search expertise. In the flu-shot example, is it best to teach new students a specific sequence of actions for common types of queries, or are there general search capabilities that the expert selectively mixes and matches as needed in the process of searching?

One possible approach, currently being tested (Bhavnani, 2001, Bhavnani et al., in press), is to provide new forms of websites called *Strategy Hubs*. These websites suggest strategies for searching common information needs. A typical example is the instance where an individual or a family member has been newly diagnosed with a disease. A range of common types of information is needed under these circumstances—what to expect in a doctor's examination, what is the prognosis of my disease if caught at a specific stage, what are the conventional and experimental treatments, etc. The Strategy Hub can suggest sequences of search moves for each such common type of need. Research on the Strategy Hub will lead to greater insights into the intermediate layer of knowledge, including the extent to which specific search strategies can be generalized within and across domains.

As search strategy research has moved from the manual environment (Bates 1979, 1981) to online, and across the various online technologies of databases (Fenichel, 1981; Fidel, 1984; Saracevic et al., 1988; Siegfried et al., 1993), online catalogs (Drabenstott & Weller, 1996; Hildreth, 1989; Matthews et al., 1983), and the World Wide Web (Bates, 1998; Jansen et al., 2000; Wang et al., 2000; Yang, 1997), we continue to learn and re-learn some basics—searchers use short, simple queries; build little on previous experience; learn slower with age, etc., but we do not really advance our understanding of the cognitive processes in searching; we do not penetrate to why these searching problems persist. Information seeking research sometimes does hint at the cognitive processes, but usually in the process of carrying out studies that are more sociologically than cognitively oriented—so once again the deeper understanding of the mental processes fails to develop.

The method advanced in this paper does not just show sequential steps in search, nor does it cross-tabulate search success with various sociological or technological variables, as many studies have. Instead, the method decomposes the *cognitively distinctive* elements and activities in the mind of the searcher during the process. These distinctive layers and cognitive features are often intermingled indiscriminately in other search strategy research. Commonly, even very rigorously designed empirical studies (see, e.g., *Journal of the American Society for Information Science and Technology* special issue on Web research, Spink, 2002) are testing and comparing elements of these four layers in an undifferentiated way. As a consequence, the research may tell us many things, but not shed much light on the cognitive process of searching.

It should also be said that much remains to be tested. Information search strategies are fundamentally heuristic, not algorithmic, and many of the processes studied in cognitive science are simpler and more algorithmic in nature compared to information searching. Furthermore, as suggested above, many other questions remain to be tested on the effectiveness of this analysis method for search strategy. But we posit that use of this goal decomposition approach provides a much more promising approach to search strategy research than is evident in much of the previous fragmented work on the subject.

Acknowledgements

This research was partly supported by the National Science Foundation, Award# EIA-9812607. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of NSF or the U. S. Government. The authors thank F. Peck, F. Reif, G. Vallabha for their contributions.

References

Bates, M. J. (1979). Information search tactics. *Journal of the American Society for Information Science*, 30, 205-214.

Bates, M. J. (1981). Search techniques. *Annual Review of Information Science and Technology*, 16, 139-169.

Bates, M. J. (1989). Design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13, 407-424.

Bates, M. J. (1990). Where should the person stop and the information search interface start? *Information Processing & Management*, 26, 575-591.

Bates, M. J. (1994). The design of databases and other information resources for humanities scholars: The Getty Online Searching Project Report No. 4. *Online & CDROM Review*, 18, 331-340.

Bates, M. J. (1998). Indexing and access for digital libraries and the Internet: Human, database, and domain factors. *Journal of the American Society for Information Science*, 49, 1185-1205.

Bates, M. J. (1999). The invisible substrate of information science. *Journal of the American Society for Information Science*, 50, 1043-1050.

Bates, M.J. (2002). The cascade of interactions in the digital library interface. *Information Processing and Management*, 38, 381-400.

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982). ASK for information retrieval: Part I. Background and theory. *Journal of Documentation*, 38, 61-71.

Belkin, N. J., Marchetti, P. G., & Cool, C. (1993). BRAQUE: Design of an interface to support user interaction in information retrieval. *Information Processing & Management*, 29, 32-344.

Bhavnani, S. K. (1998). *How Architects Draw with Computers: A Cognitive Analysis of Real-World CAD Interactions*, unpublished Ph.D. dissertation, Carnegie Mellon University, Pittsburgh.

Bhavnani, S. K. (2001). Important cognitive components of domain-specific search knowledge. *Proceedings of TREC' 2001*, 571-578.

Bhavnani, S. K. (2002). Domain-specific search strategies for the effective retrieval of healthcare and shopping information. *Proceedings of CHI'2002*, 610-611.

Bhavnani, S.K., Bichakjian, C.K., Schwartz, J.L., Strecher, V.J., Dunn, R.L., Johnson, T.M., & Lu, X. (in press). Getting patients to the right healthcare sources: From real-world questions to Strategy Hubs. *Proceedings of AMIA'2002*.

Bhavnani, S.K., Flemming, U., Forsythe, D.E., Garrett, J.H., Shaw, D.S., & Tsai, A. (1996). CAD Usage in an Architectural Office: From Observations to Active Assistance. *Automation in Construction*, 5, 243-255.

Bhavnani S.K., & John, B.E. (1996). Exploring the Unrealized Potential of Computer-Aided Drafting, *Proceedings of CHI'96*, 332-339.

Bhavnani, S.K., & John, B.E. (1997). From Sufficient to Efficient Usage: An Analysis of Strategic Knowledge. *Proceedings of CHI'97*, 91-98.

Bhavnani, S.K., & John, B.E. (1998). Delegation and Circumvention: Two Faces of Efficiency. *Proceedings of CHI'98*, 273-280.

Bhavnani, S.K., & John, B.E. (2000). The Strategic Use of Complex Computer Applications. *Human-Computer Interaction*, 15, 107-137.

Bhavnani, S.K., John, B.E., & Flemming, U. (1999). The Strategic Use of CAD: An Empirically Inspired, Theory-Based Course. *Proceedings of CHI'99*, 42-49.

Bhavnani, S.K., Reif, F., & John, B.E. (2001). Beyond Command Knowledge: Identifying and Teaching Strategic Knowledge for Using Complex Computer Applications. *Proceedings of CHI' 2001*, 229-236.

Card, S.K., Moran, T.P., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates,.

Chen, H., & Dhar, V. (1991). Cognitive process as a basis for intelligent retrieval systems design. *Information Processing & Management*, 27, 405-432.

Drabenstott, K. M. (2000). Web search strategies. In: *Saving the user's time through subject access innovation; Papers in honor of Pauline Atherton Caochrane*, William J. Wheeler (Eds.), 114–161. Champaign, Ill.: University of Illinois.

Drabenstott, K. M., & Weller, M. S. (1996). Failure analysis of subject searches in a test of a new design for subject access to online catalogs. *Journal of the American Society for Information Science*, 47, 519-537.

Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation*, 45, 171-212.

Fenichel, C. H. (1981). Online searching: Measures that discriminate among users with different types of experiences. *Journal of the American Society for Information Science*, 32, 23-32.

Fidel, R. (1984). Online searching styles: A case-study-based model of searching behavior. *Journal of the American Society for Information Science*, 35, 211-221.

Fidel, R. (1991). Searchers' Selection of Search keys: 1. The Selection Routine. *Journal of the American Society for Information Science* 42, 7, 490-500.

Holscher, C., & Strube, G. (2000). Web search behavior of Internet experts and newbies. *Computer Networks*, 33, 337-346.

Hildreth, C. R. (1989). *Intelligent Interfaces and Retrieval Methods for Subject Searching in Bibliographic Retrieval Systems*. Washington, D.C.: Library of Congress Cataloging Distribution Service.

Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science*, 44(3) 161-174.

Ingwersen, P. (1992). *Information Retrieval Interaction*. London: Taylor Graham.

Ingwersen, P. (1996). Cognitive aspects of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*, 52, 3-50.

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

John, B., & Kieras, D. (1996). The GOMS Family of User Interface Analysis Techniques: Comparison and Contrast. *Transactions of Computer-Human Interaction*, 3, 320-351.

Kirwan, B., & Ainsworth, L. K. (1993). *A guide to task analysis*. Taylor and Francis, London.

Kuhlthau, C. C. (1993). *Seeking Meaning: A Process Approach to Library and Information Services*. Norwood, NJ: Ablex.

Lazonder A., Biemans, H, & Wopereis, I. (2000). Difference between novice and experienced users in search information on the World Wide Web. *Journal of the American Society for Information Science*, 51(6), 576-581.

Marchionini, G. (1995). *Information Seeking in Electronic Environments*. New York: Cambridge University Press.

Matthews, J. R., et al (Ed.). (1983). *Using Online Catalogs: A Nationwide Survey*. New York: Neal-Schuman.

Nilsen, E., Jong H., Olson J., Biolsi, I., & Mutter, S. (1993). The Growth of Software Skill: A Longitudinal Look at Learning and Performance. *Proceedings of INTERCHI'93*, 149-156.

Pejtersen, A. M. (1984). Design of a computer-aided user-system dialogue based on an analysis of users' search behaviour. *Social Science Information Studies*, 4, 167-183.

Rosson, M. (1983). Patterns of Experience in Text Editing. *Proceedings of CHI '83*, 171-175.

Saracevic, T., et al. (1988). Study of information seeking and retrieving. I. Background and methodology. *Journal of the American Society for Information Science*, 39, 161-176.

Shute, S., & Smith, P. (1993). Knowledge-based search tactics. *Information Processing & Management*, 29, 1, 29-45.

Siegler, R.S., & Jenkins, E. (1989). *How Children Discover New Strategies*. Lawrence Erlbaum Associates, New Jersey.

Siegfried, S., Bates, M. J., & Wilde, D. N. (1993). A profile of end-user searching behavior by humanities scholars: The Getty Online Searching Project Report No. 2. *Journal of the American Society for Information Science*, 44, 273-291.

Singley, M.K., & Anderson, J.R. (1989). *The Transfer of Cognitive Skill*. Harvard University Press, Cambridge, Massachusetts.

Spink, A. (Ed.) (2002). Special issue on Web research. *Journal of the American Society for Information Science and Technology*, 53, 65-196.

Sutcliffe, A., Ennis, M., & Watkinson, S. (2000). Empirical studies of end-user information searching. *Journal of the American Society for Information Science*, 51(13), 1211-1231.

Wang, P., Hawk, W. B., & Tenopir, C. (2000). Users' interaction with World Wide Web resources: An exploratory study using a holistic approach. *Information Processing & Management*, 36, 229-251.

Xie, H. (2000). Shifts of interactive intentions and information-seeking strategies in interactive information retrieval. *Journal of the American Society for Information Science*, 51, 841-857.

Yang, S. C. (1997). Qualitative exploration of learners' information-seeking processes using Perseus hypermedia system. *Journal of the American Society for Information Science*, 48, 667-669.