Applying AI Models to the Design of Exploratory Hypermedia Systems

Ray Bareiss
and Richard Osgood
Institute for the Learning Sciences
Northwestern University
1890 Maple Avenue
Evanston, IL 60201
bareiss@ils.nwu.edu,
osgood@ils.nwu.edu

ABSTRACT

Hypermedia systems offer great promise for capturing expertise and subsequently providing multifaceted access in support of a user engaged in a complex task. A primary issue in building such systems is how to structure the knowledge contained in them such that a user with a problem can find the most appropriate knowledge easily and naturally. Our artificial intelligence-based research has produced two approaches to structuring knowledge that show promise for organizing hypermedia knowledge bases. The first is the use of abstract models of a user's problem solving task to provide a global structure for the knowledge base and to reflect that structure in a meaningful human interface to the system. The second is the use of a conversational model to provide local coherence of the links among pieces of information in the knowledge base. These two types of models have been applied in the construction of Trans-ASK, a large hypermedia system in the domain of military transportation planning. This paper discusses the theory underlying the models, their application within Trans-ASK, and finally provides a preliminary evaluation of the resulting system.

KEYWORDS

Hypermedia, Task Models, Indexing, Exploratory Systems

INTRODUCTION

Artificial Intelligence research has concentrated on the design of autonomous problem solving systems. Although there have been successes, AI systems, to date, have proven to be of little utility in assisting human users in the solution of complex, novel problems. This is, in part, because AI systems are limited in the assistance they are designed to provide. The typical system expects input in terms of situational features (e.g., red spots on skin and low-grade fever) and produces a terse solution (e.g., your patient has measles with a probability of 87%) with little explanation of how the solution was derived. Furthermore, such solutions can be produced only when the system builders have incorporated appropriate knowledge of the problem in the system's memory.

In contrast, a human expert can provide much more problem solving assistance. The expert not only can offer advice and, perhaps, recall relevant experiences, but also can provide a rich context in which to interpret that advice and experience. He or she can advise which questions are best to ask in a problem solving situation, why to ask those questions, which questions not to ask, the likely results of asking questions, how to interpret data, alternatives to a recommended approach, *et cetera.* Even if he or she cannot provide a specific solution, useful general guidance can be provided.

Hypermedia systems offer great promise in providing such multifaceted access to expertise. A primary issue in building such systems is how to structure the knowledge contained in them such that a user with a problem can find the most appropriate knowledge easily and naturally. As Spiro and Jehng (1990) have pointed out, most research in hypermedia has been atheoretical, aimed at increasing the power of the technology rather than at providing approaches to appropriate organization of information to meet a user's needs.

Although AI research has had limited success at providing useful problem solving systems, AI research methodologies have produced ways of structuring knowledge that show promise for organizing hypermedia knowledge bases. This paper discusses two complementary model-based approaches to organizing hypermedia systems that have been developed and are in use at the Institute for the Learning Sciences. The first is the use of abstract models of a user's problem solving task to provide a global structure for the knowledge base and to reflect that structure in a meaningful human interface to the system. The second approach is to use a conversational model to provide local coherence of the links among pieces of information in the knowledge base.

Application of the ILS approach has resulted in the production of a class of hypermedia systems that we call ASK systems. In contrast to "traditional" approaches which are based on the metaphor of a well-indexed book, ASK systems are based on the metaphor of conversation with an expert. An ASK system provides flexible access to a database of short video cases extracted from interviews with experts, as well as archival video and textual material. Our goal is to construct systems that enable users to gain access to information which answers their questions as they arise and to structure their interaction with a system in a way which realizes the most important benefits of conversation with a human expert. In particular, an ASK system provides user-directed browsing of information within a framework which provides a coherent model of the domain under investigation. This model enables them to assimilate answers to their questions in such a way that they can then use those answers to solve problems. When using an ASK system, the user selects a topic of interest (via a task model reified in the system's graphical human interface), views initial information (typically in video form), and then can ask any of a large number of contextually relevant follow-up questions for which the system has good answers.

ILS has built a number of ASK systems in domains as diverse as trust bank consulting, industrial development, contemporary American history, change management consulting, the English water business, and social services for Mexican immigrants. To date, the largest ASK system produced, Trans-ASK, operates in the domain of military transportation planning. Trans-ASK is designed to serve as a job aid, training tool, and reference for officers assigned to the United States Transportation Command (TRANSCOM). TRANSCOM is a joint military command which has the responsibility for planning, coordinating, and scheduling military transportation in wartime and, more recently, in peacetime. During a fast-breaking crisis, such as the recent Desert Shield and Desert Storm (D/S/S), a tremendous amount of planning and coordination must be accomplished effectively in a very short period of time. Unfortunately, because of the short term of many military assignments and other factors, many of the action officers who are assigned have little actual experience, and, in the frantic atmosphere of a crisis, they cannot always be given adequate assistance by their co-workers.

Trans-ASK captures the expertise of 33 experienced transportation planners at TRANSCOM in the context of their experiences during D/S/S, Somolia, and Hurricane Andrew. Their expertise is most often conveyed in the form of first person stories which are vivid and memorable to a viewer and, by analogy, provide the viewer with more effective guidance in the appropriate application of that expertise than would a set of decontextualized general principles, such as the "lessons learned" documents typically generated by military commands after each operation. (We have shot 68 hours of (raw) video interviews, and the system currently contains over 21 hours of indexed material.)

The remainder of this paper discusses the use of task and conversational models to structure hypermedia as exemplified by the design of Trans-ASK.

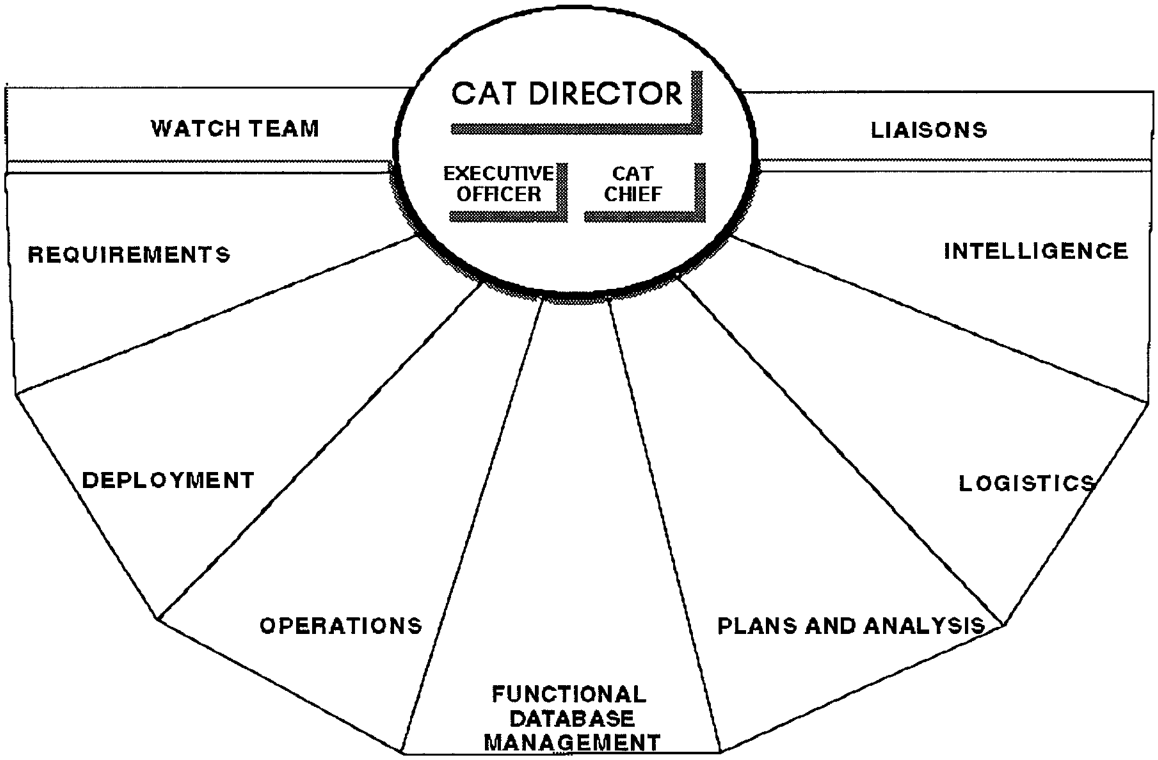

Figure 1: The Functional Roles of the Crisis Action Team

THE USE OF TASK MODELS TO PROVIDE GLOBAL STRUCTURE

A great deal of research in AI and cognitive science has established the existence and utility of abstract knowledge of intellectual tasks, such as planning and diagnosis. This work has resulted in the codification of a wide range of models, including Sussman's and Sacerdoti's planning critics [Suss75; Sace77], Wilensky's meta-plans [Wile82], Schank's thematic memory structures [Scha82], Chandrasekaran's generic tasks [Chan83], and the KADS models [Wiel92]. We believe that much human problem solving knowledge is organized around abstract task models (cf. situated cognition [Brow89]) and hence the memory of a hypermedia instructional system should have a similar organization.

The primary issue is, of course, what model(s) the user possesses. This is the sort of question that motivates research in student/user modeling (see, e.g., [vanL88]). However, the problems inherent in such an approach and the research community's slow progress in solving them, militate against its use in systems intended for deployment in the short term. Our alternative approach is to search for an abstract level of task knowledge that is nearly universal and can therefore serve as common ground between a system and its human user. If the memory of the system is structured around the appropriate task model, the user will have the ability to locate specific instantiations of portions of the model which are likely to be relevant to his/her problem. It is important to note, however, that the goal of the system is not to convey this abstract knowledge to the user but rather to present task-specific knowledge that the user is presumed to be able to locate and interpret because of his or her prior possession of a similar (implicit) model.

Rather than using abstraction task models in an attempt to build autonomous problem solving or instructional systems, we are using them to structure the knowledge base of hypermedia systems to provide a coherent basis for user-directed exploration. Through this approach, we believe that the user will be able to understand the organization of the system's knowledge intuitively, keep his or her mental process "in sync" with the state of the system, and, consequently, understand available navigational choices in the context of a dialogue with the system.

The Trans-ASK knowledge base and its human interface is organized around a model of the job roles in the TRANSCOM crisis action team (CAT), the tasks carried out by people in those roles, and typical problems in successfully accomplishing them. This mirrors the situation in which an action officer is likely to consult Trans-ASK: he/she is in a particular role, carrying out a task, and having a problem doing so.[1]

When an officer first begins to use Trans-ASK, he is asked to identify his role within the CAT. He does so via graphical selection on a screen that graphically displays the functional roles of the CAT (see Figure 1). While similar to an organization chart, this screen reflects general roles rather than the specific organization of the CAT at some point in time (which is subject to frequent change). The roles that a new officer is most likely to play are:

- requirements: identifying and verifying transportation requirements
- operations: determining the availability of transportation
- deployment: coordinating between requirements and available transportation

Other roles are more specialized and are not typically assigned to non-specialists. Let's assume that our hypothetical user is a requirements action officer. After identifying this role to the system, he is asked to identify the task on which he is working via menu selection. Example tasks are:

- Identifying transportation requirements.
- Verifying the accuracy of requirements.
- Monitoring sustainment requirements (e.g., food and other supplies for deployed troops).
- Ensuring that valid and prioritized requirements are entered into JOPES (the computerized Joint Operations Planning and Execution System currently used by the military).

Space does not permit in-depth discussion of what these tasks are.[2] The menu presents an additional choice, "What does requirements do?" which is intended to offer general background to a user of Trans-ASK at a service school or another command rather than to a TRANSCOM officer.

Suppose the user selects "Ensure that valid and prioritized requirements are entered into JOPES." He is then asked to identify his particular problem. He may have the general problem of not knowing how to accomplish the task at all or a more specific problem, such as:

- JOPES data is inaccurate.
- The situation has changed making requirements outdated.
- Priorities are changing very rapidly.

After making this selection, he is presented with initial information about the problem and then is free to explore the knowledge base to understand its implications and possible solutions. (The exploration, or browsing, process is described in the next section.)

---

[1]ASK systems built by ILS have used other abstract models as well. For example, ASK Tom [Ferg92] was organized around a model of the significant entities in the world of trust banking and the relationships among them.

[2]They were gleaned from an analysis of crisis action procedures produced for TRANSCOM by SRA Corporation.

This model was formulated to provide an initial orientation and high-level assistance to new requirements officers. We have since been asked by TRANSCOM to provide greater detail to enable the user to more easily locate possible solutions to specific problems. Two types of changes were made to the original enumeration of tasks: First, some of the tasks were decomposed to provide a greater level of detail. For example, "verifying requirements" was decomposed into verifying that the requestor has the command authority to make the transportation request and validating the accuracy (e.g., the size) of the requirement. Second, additional, secondary tasks were added which were invisible in the original enumeration. For example, the task of determining which transportation component command is most appropriate to handle a requirement was added. The associated lists of problems have been similarly expanded.

In summary, by identifying his job role, task, and problem, the user is selecting an area of expertise in the Trans-ASK database. He is then free to explore that area via the conversational structure described in the next section.

USING A CONVERSATIONAL MODEL FOR LOCAL COHERENCE

The local organization of an ASK system is derived from a simple theory of human memory organization. As a new case is acquired during problem solving, it is indexed in human memory relative to recalled cases that raise questions answered by the new case. Conversely, questions the new case raises that are answered by existing cases, provide reverse indices.

This general theory argues that coherence in a conversation is memory based [Chaf79]. Our ASK systems are based on a model of the general memory organization that might underlie a conversation about problem solving [Scha77] After hearing a piece of information in such a conversation, we hypothesize there are only a few general categories of follow-up information that represent a natural continuation of possible threads of the conversation rather than a major topic shift. The categories can be thought of poles along four axes or dimensions:

1. Refocusing: Adjustments to the specificity of topic under consideration (fine tuning a "zoom" to a topic, which was originally accomplished via the task model) and slight digressions like clarifying of the meanings of terms or describing situations in which the topic arises.

    context: the big picture within which a piece of information fits.

    specifics: an example of a general principle or details of a situation.

2. Causality: Explanations and outcomes---that is, the need to flesh out causal/temporal relationships as a means of gaining understanding. (We group temporal order and the causal chain because people typically collapse the distinction.)

    causes (or earlier events): how a situation developed.

    results (or later events): the outcome of a situation.

3. Comparison: Questions of similarity and difference, analogy and alternative at the same level of abstraction as the current information.

    analogies: similar situations from other contexts or from the experiences of other experts.

    alternatives: different approaches to take in a situation or differing expert opinions.

4. Advice: The application of knowledge or carrying away a lesson, either negative or positive, for use in the problem solver's situation.

    opportunities: advice about things to capitalize on in a situation.

    warnings: advice about things that can go wrong .

These eight categories can also be thought of as the most general kinds of questions that a user is likely to have in a conversation about problem solving.[3] The browsing interface of an ASK system reifies this model of conversation about problem solving in the form of eight graphically displayed categories of relative links between stories [Ferg92]. A user can find his or her specific question in one of eight categories. If the user has only a very vague idea of what question to ask, he or she can simply browse through the questions in a general category that looks promising.

To understand relative indexing within Trans-ASK, consider the case of "Working on Christmas" told by the TRANSCOM Operations Chief during D/S/S:

> The best example of personality driven agreements or lack of agreements was Christmas. ... The people had planned not to move large numbers of troops and equipment out of Europe between the 24th and 26th of December, not because they wanted Americans to have the day off but because host-nation support (German bus drivers, German truck drivers) would not be readily available over the Christmas holiday. ... This was perceived as TRANSCOM not wanting to provide transportation over the Christmas holiday even though we had not planned it that way. And consequently, discussions were conducted at very high level, ... it became an attitude of "Well my people are tough enough to work over Christmas and show up at the airport. Why aren't your people tough enough to be there with the airplane to pick them up and take them to Saudi Arabia." In that kind of environment, a decision was made that the airplanes would be there and that those people would be there. Now this was completely without consideration for the planning or the long term effect of those decisions. ... Problems arose in that the host-nation support in fact was not there. ... Consequently, airplanes arrived "no load available." Very ineffective use of airlift. Again, agreements or disagreements or perceptions among senior leaders based on phone calls and handshakes in some cases had a negative effect on the operation.

This story provides an engaging introduction to the impact of image-driven decision making on the war effort. It is likely to raise a number of questions in the mind of an interested viewer. A key feature of Trans-ASK is its browsing interface which anticipates the viewer's most immediate questions and provides answers to all of those for which answers are available.

The system can provide answers to a number of questions raised in some of the eight general conversational categories. Example questions are (6 of the 14 answered in the system):

Context: Who had final control over transportation planning?

Specifics: How much control does TRANSCOM leadership exercise over day-to-day operations?

Results: What was the effect of informal coordination on transportation planning?

Analogies: What is another example of planning failure due to informal coordination during D/S/S?

Opportunities: How can a leader be made to change his mind when he's wrong?

Warnings: What are the disadvantages of informal coordination?

There are no questions answered in the causes and alternatives categories. Again, questions are only displayed in the interface if the system can provide answers. The graphical layout of the browsing interface is depicted in Figure 2.

---

[3]We now believe that with minor adjustments the model we present applies to decision making, teaching, design, and interest-based browsing tasks as well.

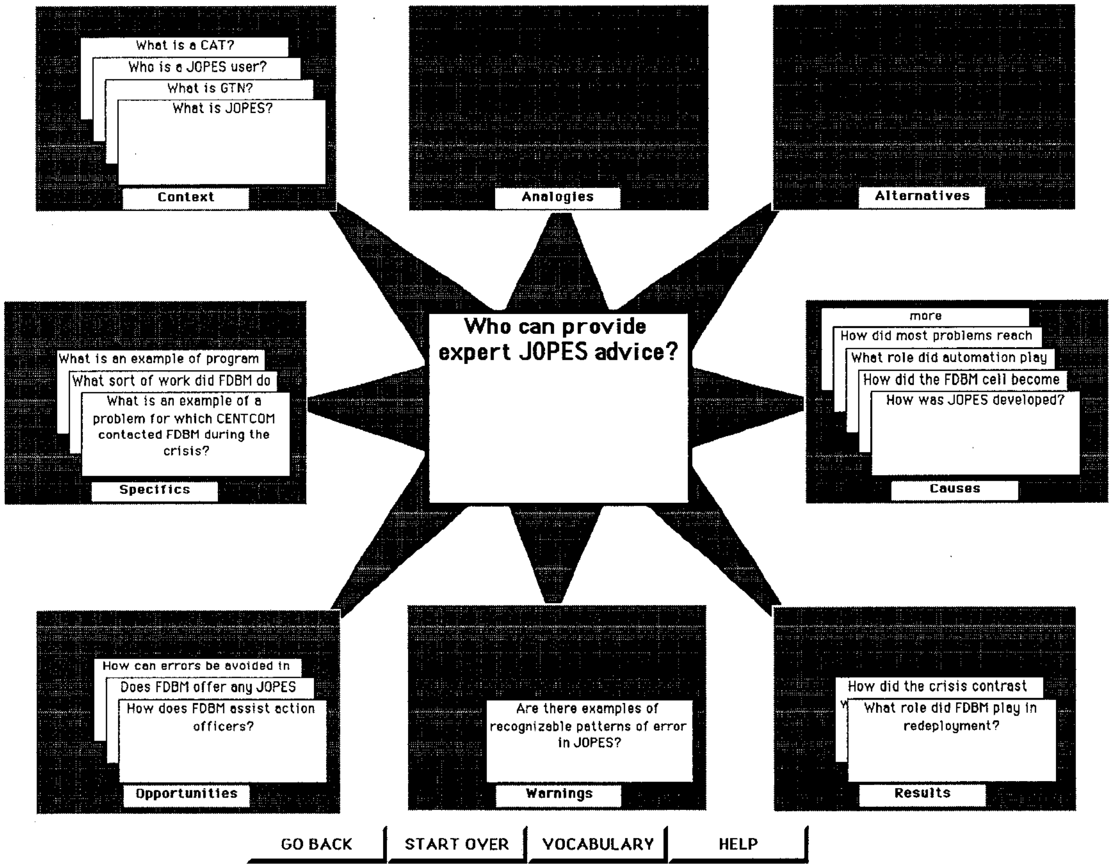

Figure 2: The Browsing Interface of Trans-ASK

INDEXING THE SYSTEM

In this section we describe how Trans-ASK was indexed, then we address some the issues for relative indexing that the experience has raised. The first step in the indexing process was to catalog the raw interview video, identifying good material and segmenting it into short clips (i.e., 1-2 minutes) that seemed to have single primary topics. The clips are also edited and pressed to videodisc at this point in the process.[4]

After bookkeeping information about a set of clips has been entered into an indexing tool, a portion of the ASK system is constructed by the process of "question-based indexing." Indexers doing this must adopt the job role and, hopefully, the degree of sophistication of the intended user of the system. They view each clip in isolation and enumerate the questions that the clip is likely to raise in the mind of prospective users working on particular tasks. The general question categories of the ASK theory (illustrated in the previous section) provide guidance in this process. Next, indexers enumerate questions for which the clip is likely to provide a good answer. Each question raised and each question answered is categorized by topic (e.g., high priority shipments) and by general conversational category (e.g., causes).

[4]The current version of Trans-ASK employs multiple videodiscs in a Pioneer "video jukebox." We believe that the system must be converted to digital video before deployment to enhance its usability. (We have built two later systems, in other domains, employing DVI video on magnetic disks.)

During relative indexing, links are made between stories to enable user-directed browsing of the ASK system. The questions (raised and answered) provide an abstraction of story content that is easier to manipulate during the relative indexing process than the stories themselves. Indexers use the indexing tool to compose a set of questions raised, identifying them by topics and, optionally, by conversational category. They similarly compose a set of questions answered. Links between stories are made by graphically selecting a question raised and a semantically equivalent question answered. The process is performed manually, rather than via natural language processing (i.e., parsing the questions and matching the resulting computer-generated representations), because of the tremendous variability possible in the statement of semantically similar questions.

Not all links between clips made via this method actually seem to maintain conversational coherence when the clips are viewed in juxtaposition. This is due to a combination of incompletely elaborated questions and faulty question matches. So, indexers often shift to a browsing mode to evaluate the quality of links. Like expert system building, ASK system indexing relies on rapid prototyping and iterative refinement.

A problem we encountered during the construction of Trans-ASK was that the sheer magnitude of the system (hundreds of clips and several thousand questions) made question matching cumbersome. To deal with this problem we constructed Trans-ASK as a set of single expert, mini-ASK systems, each of which contained the results of a single interview. The content of such a single expert ASK system necessarily raises many interesting questions which are not answered by clips from the interview. Some of these are questions which could only be answered by the original interviewee (for example, "What were you thinking when you did that?") and, hence in our case, were discarded. However other questions, which call for general background or advice, could be answered satisfactorily by any planning expert. The indexer marked such questions as "exportable." After the set of single expert ASK systems was constructed, they were merged and a second question linking process was carried out using only the exportable questions. The full Trans-ASK system, containing multiple interconnected experts, is the result of this process.

ISSUES IN QUESTION-BASED INDEXING

Our experience in Trans-ASK and other ASK Systems has begun to yield some practical guidance for the content analysis of stories by the question-based indexing method. The clip segmentation problem and the question matching problem are discussed here.

Before relative indices are built, the interview footage must be segmented into short video clips. Each clip is chosen so that it makes a single main point. In solving the granularity problem this way, a careful balance must be struck. The smaller the clip size the more the local context is narrowed such that more questions are raised and fewer questions are answered. The limiting case on the small end is of course simply the piece of video that directly answers a question raised. Because users are skipping around in the content as they navigate the system, such terse answers do not preserve enough local context for most users to make sense of the answer. Conversational coherence is lost, because users will have to pursue too many background follow-up questions just to interpret the answer. At the other end of the spectrum, if a clip is too large it provides too much context information and takes too long to get to the point where it answers the user's question. So there is a trade off between small clip size in which local context questions are answered by traversing a link to another clip and larger clip sizes where more are answered right within the clip itself. Indexers are asked to use their knowledge of the prospective user's need for local context when segmenting footage into clips.

Question vagueness, superfluity, parallelism, applicability, similarity, and insufficiency have profound effects on the quality of links found by question matching. The "vagueness" issue arises in many question matching situations. Should a link be made when a question raised in one story is about general transportation problems while the question answered by another is about one specific kind of transportation problem, i.e., how should matching be conducted when one question is vague or general and the other is specific? The "superfluity" issue arises when questions answered by one story are never raised in any other stories. Was the content analysis incomplete or is the material in these stories actually unnecessary? The "parallelism" issue arises when questions raised in one topic area (e.g., Under the D/S/S topic: Who was responsible for transportation planning?) are also raised in parallel topic areas (e.g.,Vietnam War). How should this parallelism be represented in the browsing interface? Next, the "applicability" issue arises when certain questions answered are raised in virtually every story within a given wider context. For example, "What was the role of TRANSCOM in D/S/S?" provides background for nearly every story in Trans-ASK. How should these globally raised questions be distinguished in the browsing interface? The "similarity" issue arises when a question raised in one story is matched by similar questions answered from several different stories. Should one or all of these alternative answers be linked? The "insufficiency" issue arises when a question raised in a story is not answered by any single story but requires multiple stories to construct a good answer, if indeed a good answer exists at all. How many outstanding questions raised with no answer can be tolerated before the utility of the network of stories is compromised? When an answer can be constructed, how are all these pieces of the answer identified and linked? Space does not permit addressing these issues here. See [Osgo93] for more detail.

We single out the problem of question vagueness during matching for further discussion because it illustrates the power of the question-based indexing method. There are two kinds of vagueness: "Topical vagueness" and "conversational vagueness." We deal with each type of vagueness for both questions raised and questions answered.

The first kind of vagueness in a question concerns the "topical" span of coverage of a question. There are many interesting stories an expert could tell that would answer a question raised like "What is another example of a planning failure during DSS?" Such a question typically defines a space of acceptable answers. In the example story, that question was matched to the question answered, "Did TRANSCOM ever move things by air that should have been moved by sea?" Finding a specific question answered to match a vague question raised depends on both appearing in the same categories for potential matching. Two means are available to help with this. First questions must be put into "all" of the general conversational categories that make sense. This insures that the question will appear everywhere it belongs, so linking will be complete. Second, the topical classification groups must be either very inclusive or be relaxable so that general questions and the specific questions will be in the same matching category. This approach provides indexers, who must ultimately make the matching judgments, with opportunities to find appropriate ones.

When a question answered is posed vaguely, it too is meant to define a span of "topical" applicability. A vague question defines an upper bound on the applicability of a particular story, but it does not necessarily define a lower bound (i.e., how specific a question raised the corresponding story can answer). Therefore, a vague question answered cannot not be used by itself to match a specific question raised. Consider the mismatch between a question answered: "What role did the U.S. military play in the liberation of Kuwait?" versus the question raised: "What role did TRANSCOM play... ?" When a question answered is about a more general topic, e.g. the US Military, indexers cannot judge whether or not the story will cover TRANSCOM, the more specific topic of the question raised. Therefore, they cannot, in general, match a topically specific question raised with a topically general question answered. As indexers develop more skill representing stories as questions answered, two kinds of questions emerge: questions that capture the most general kinds of answers the story can provide and questions that capture the most specific answer the story can provided. Defining these boundaries improves matching efficiency. In situations where this still does not resolve the vagueness of a question answered, the story itself must be examined to determine if the match is acceptable.

The second kind of vagueness is conversational vagueness. Consider the specific question "How was TRANSCOM established" versus the vaguer question "What is the history of TRANSCOM?" If the former question was raised while the latter was answered, the indexer could be reasonably certain, by common sense conversational convention, that the vague question answered would satisfactorily answer the specific question raised, i.e., history should include origins. The converse is not necessarily true: a satisfactory answer to the specific question about the establishment of TRANSCOM could not be necessarily expected to trace the history of the organization. Conversationally vague questions raised cannot be reliably matched to conversationally specific questions answered. However, as the reliability of indexer question-based representation grows, the absence of the more general question answered can be trusted to mean that the more specific one is indeed not answered. Otherwise, indexers must examine the underlying stories to make a linking decision.

Finally, since indexers are (or should be) posing questions at the "right" level of abstraction for the eventual user, it is tempting to believe that the above types of mismatches should rarely occur. This is untrue. On the one hand, indexers are trying to ask naturally occurring questions raised---that is, they are trying to model the user's questions. On the other, they are trying to represent stories with "covering" questions answered, specifying the upper and lower bounds on the kinds of answers the story can provide. Therefore, handling mismatches in question abstraction level is fundamental to the question-based indexing method.

EVALUATION

Trans-ASK is work in progress and, consequently, has yet to be fully evaluated. In particular, we have not conducted formal studies of the effectiveness of structuring content via our models in contrast to other possible methods. The bulk of the feedback we have received on system design came from a two day focus group meeting which included seven current and former TRANSCOM personnel plus six additional people with substantial military planning experience. The meeting included considerable hands-on use of a prototype version of the system.

Overall reaction to Trans-ASK was quite favorable. To paraphrase a typical reaction: Trans-ASK will be a "great benefit" for training new personnel and getting them up to speed. It will be valuable to any of the "individual directorates within TRANSCOM which are responsible for training individuals how you do requirements, how you do deployment, how you do logistics," *et cetera.* "A major or lt. colonel sent to school on joint operations could be going to any of the joint unified commands. Trans-ASK fills the next level of need-to-know for someone coming into TRANSCOM." In particular, they perceived the most valuable aspects of Trans-ASK to be in-context learning of domain concepts and the ability to explore alternative points-of-view and courses-of-action in depth. Their major complaint was that the system (at the time they used it) did not contain enough detail on specific procedures and on how to address specific problems. They also stressed that situating the use of Trans-ASK within performance of an actual job at TRANSCOM (in contrast to using it at a training school) is probably essential to make its content meaningful.

In general, the focus group found the task-based zooming interface to provide an acceptable means of initial entry into the database. They did, however, request that two additional means of access be provided. The first is via "topics" which would, in-effect, provide an on-line encyclopedia of domain concepts.[5] The second is via "speakers" which would provide the opportunity to choose from all clips of a particular expert (e.g., a person who was formerly in the same job role as the user) subindexed by questions answered.

In general, the focus group also found the browsing interface to be intuitive. We performed an experiment in which the focus group members were shown ten diverse clips and asked to enumerate the questions that each raised in their minds. As might be expected from their general comments, their questions were most often requests for additional, specific details.

We also presented the questions that our content analysts had raised in response to each of the ten clips and asked the group members to categorize them using our eight conversational categories. To aid in this, they were given the definitions presented in the previous section of this paper. When the aggregate results of this process were examined by category, "specifics" was always the most frequently chosen category for a set of questions, regardless of our analysts' assigned categories. The second most frequent choice was the analyst-assigned category in all cases. A group member gave us insight into this somewhat surprising phenomenon when he described his strategy for categorizing questions. In most cases, he felt that the *answer* to the questions was providing more detail about the situation and, consequently, assigned the corresponding *question* to "specifics." Informal results suggest that inter-analyst agreement is high for our experienced people, and hence, we believe that users will quickly learn to categorize their information requests consistently with the ASK conversational model.

---

[5]A fundamental issue is whether to provide access to all clips which refer to a topic or only to those which are relevant (by some criteria) to gaining a basic understanding.

The real test of Trans-ASK is planned for late this summer. A group of officers, newly assigned to TRANSCOM, will be given a set of complex problems for which Trans-ASK contains useful information. We will assess their ability to locate relevant areas of the database, their exploratory patterns,[6] and their attitudes with respect to the utility of the system.

CONCLUSION

Trans-ASK owes much of its power to the two models that are used to organize access to the video information about transportation planning it contains. While the content of these videos provides the instruction, the models structure the database so that the user encounters the right videos at the right time. The task model operates on a coarse level to situate the user in an area of video stories likely to be helpful. Because alignment is high between the user's actual task and the task the system assumes, user interest levels and user perceptions of relevance of the information offered will be high as well. Once situated in an area where the system's stories are of high relevance, the conversational model takes over. It keeps relevance high as follow-up questions are asked by only offering as answers stories that have strong conversational connections to the story to which the task model first led the user. We believe that the combination of these models gives ASK systems an advantage over other organizational approaches to hypermedia.

REFERENCES

[Brow89] Brown, J., Collins, A., & Duguid, P. "Situated Cognition and the Culture of Learning." Educational Researcher, Vol. 18, 1989, pp. 32-42.

[Chaf79] Chafe, W. "The Flow of Thought and the Flow of Language." In T. Givon (Ed.), Discourse and syntax. pp. 159-181. (New York) Academic Press, 1979.

[Chan83] Chandrasekaran, B. "Towards a Taxonomy of Problem-Solving Types." AI Magazine, Vol. 4, 1983, pp. 9-17.

[Ferg92] Ferguson, W., Bareiss, R., Birnbaum, L., and Osgood, R. "ASK Systems: An Approach to the Realization of Story-based Teachers." The Journal of the Learning Sciences, Vol. 2 No. 1, 1992, pp. 95-134.

[Osgo93] Osgood, R. "Question-based Indexing." Forth-coming technical report. The Institute for the Learning Sciences, Northwestern University, Evanston, IL., 1993.

[Sace77] Sacerdoti, E. "A Structure for Plans and Behavior." New York: American Elsevier, 1977.

[Scha77] Schank, R. "Rules and Topics in Conversation." Cognitive Science, Vol. 1, 1977, pp. 421-441.

[Scha82] Schank, R. "Dynamic Memory: A Theory of Reminding and Learning in Computers and People." Cambridge, England: Cambridge University Press, 1982.

[Spir90] Spiro, R. and Jehng, J. "Cognitive Flexibility and Hypertext: Theory and Technology for the Nonlinear Traversal of Complex Subject Matter." In D. Nix and R. Spiro (eds.), Cognition, Education, and Multimedia: Exploring Ideas in High Technology, Hillsdale, NJ: Lawrence Erlbaum Associates, 1990.

[Suss75] Sussman, G. "A Computer Model of Skill Acquisition." New York: American Elsevier, 1975.

[vanL88] van Lehn, K. "Student Modeling." In M. Polson & J. Richardson (Eds.), Foundations of Intelligent Tutoring Systems. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.

---

[6] We also plan to implement the requested topic and speaker interfaces and will assess the relative frequencies of access using the three alternative zooming mechanisms.

[Wiel92] Wielinga, B., Schreiber, A. and Breuker, J. "KADS: A Modeling Approach to Knowledge Engineering," Knowledge Acquisition, Vol. 4, 1992, pp. 5-53.

[Wile82] Wilensky, R. "Planning and Understanding." Reading, MA: Addison-Wesley, 1982.

ACKNOWLEDGMENTS

The Trans-ASK system is being constructed by Bob Hooker, Carolyn Majkowski, Jim McNaught, Don Frega, John Welch, and Ann Kolb with assistance from Mark Chung, Rich Lynch, Mike Sijacic, and Don Steul. Other faculty researchers contributing to the project include Roger Schank, Richard Beckwith, Gregg Collins, Alex Kass, and Chris Riesbeck. The people of U.S. Transportation Command and SRA Corporation have provided tremendous help in this project. The Trans-ASK project has been funded by DARPA (contract number MOA972-91-C-0052) as a subcontract to SRA Corporation (subcontract number SC-707-001).