

## Expert Retrieval Assistance Development and Experimentation

Richard S. Marcus  
Laboratory for Information and Decision Systems  
Massachusetts Institute of Technology

### Abstract

The development of enhancements to the CONIT retrieval assistant system leading toward an expert system version are described. Enhancements included a combined menu and command interface allowing users to get assistance in selecting appropriate search modification tactics and have recall estimates calculated by CONIT. Experiments on a partially completed version of this system are summarized. Our conclusion from the development, experimentation, and analyses is that expert retrieval assistance via techniques we have proposed does provide a potential for vastly improved search effectiveness for a wide class of users. However, the development effort to realize this potential is large and, as yet, only partly completed. The particular efforts most needed to realize the potential are included in our design for the stage 2 Expert CONIT system.

### 1. Background

We and others have pioneered in research ([MARC83b], [WILL86]) to assist retrieval system users, especially inexperienced end users, by way of the mechanism known as an *intermediary* (or *front end* or *gateway*) system (our experimental intermediary system was named CONIT). Despite the apparent success of this research, including the operational systems developed by the new field created therefrom ([LISA84], [HUSH86], [WEIN86]), we realized that much more powerful retrieval assistance effectiveness was possible. Preliminary explorations along these lines were carried out (see e.g., [MARC81] and [MARC83a]) and led to our investigation of what we have termed an "expert retrieval assistant." Our recent work in this investigation, described in this paper, has focussed on three main areas: (1) detailed design and partial implementation of an expert version of our CONIT retrieval assistant; and (2) preliminary testing and analysis of CONIT in its "standard" and "expert" modes; and (3) partial conversion of CONIT to a workstation environment.

### 2. Expert CONIT Development

The expert retrieval assistant, while seeking a new level of intelligence and sophistication, depends for its success on a number of kinds of functionality that we have already developed in standard CONIT including (1) a common command language; (2) a knowledge base for the distributed, heterogeneous retrieval systems built from production rules; and (3) a search methodology based on automatic generation and execution of search strategy derived from a user's natural language topic expressions. Expert CONIT design has, then, been conceived as a series of encapsulating elaborations of standard CONIT considered as a basic core. The first elaboration consisted of adding a "construction" or "meta" level in which individual searches and whole search strategies could be constructed *prior* to their actual execution in one or more files. The construction level was also coordinated with the ability to save search strategies from session to session in *search catalogs* for individual or shared community use over time.

The second elaboration consisted of adding a "meta-meta" level, labeled "ASSIST," which assists users with a question-and-answer, menu-oriented mode in performing

appropriate construction and execution operations. One unique feature of ASSIST is the explication for the user of the commands that are implied by his menu selection actions and answers to questions. Explicating the commands is intended to help the user understand what, in fact, is being done for him and ease the way to user-directed command operations if and when he chooses to take such initiatives. Each menu or response request point in ASSIST is considered a *stage* and the progress of assistance in a user's session is guided by a *stage transition table* which specifies what assistance stage is to follow based on the current stage and the user's menu choice. As of this writing, there are approximately 1000 stages.

Besides these structural design elaborations, we have sought in a different dimension to design and implement various *intelligent agents* to enable aspects of our goals. These agents are associated with the elements of our model of the search process which has five major components: (1) formalized problem representation; (2) search strategy formulation; (3) execution of search strategies; (4) review and evaluation of search results; (5) search modification. (See [MARC83a] for details.) We have completed an initial, partial version of each of these five components as summarized below.

Expert CONIT mimics some important aspects of what human expert intermediaries do in gathering the background of a client's problem. Through a computer-user dialog a problem description is evolved including a problem title, paragraph-length description and six (so far) problem qualifiers such as maximum money and time to be spent, recall estimates and desires, and desired document types. The next two aspects of problem representation in the ASSIST mode also include elements of search strategy formulation. First a series of assistance stages is available by which a user can learn about what files are available for searching. The culmination of assistance in this area is the generation of a *search file list*, a list of files to be used for searching once the terms and structure of the search strategy are developed sufficiently.

One of the unique features of our expert retrieval system investigations has been to clearly separate, and have distinct data structures for, the *conceptual topic formulation* on the one hand and the *search strategy* on the other hand. The construction (meta) mode permits an arbitrarily complex Boolean formulation of both the topic conceptualization and the search strategy; the experienced user can directly command such formulations. However, we have determined that the beginning user should be encouraged to develop a first conceptualization from a keyword phrase which gives a succinct expression of his topic (assuming topic and not author or citation search) by "including one key word for each important concept of the problem." Since the problem title assigned by a user often fits this criterion, ASSIST gives the user the option of using his problem title as previously entered or another "keyword phrase."

After the user supplies a keyword phrase (either title or otherwise) the system transforms that into a Boolean Topic Representation by excluding insignificant words and treating each significant word as representative of one of a set of intersecting Boolean conceptual factors. An initial (default) search strategy is then (automatically) developed by creating an all-fields, truncated (subject) search on the *stem* (automatically derived) of each significant word and intersecting the component word stem searches. Execution of this or other (previous) searches in one or more files is specifiable by user selection from menu options. Once a search has been run, menu assistance is available to direct users to specify the parameters for commands to display search results.

An initial version of a module which assists a user in modifying his search strategy has been completed. When the user chooses between the general goals of broadening or narrowing a search (strategy) the system presents a menu of techniques appropriate to that goal. For example, the broadening techniques include: (1) supplying alternate terms; (2) dropping search factors; (3) changing Boolean ANDs to ORs; (4) replacing a term with a better one; (5) changing exact or close proximity (phrase) match requirements to looser word, stem or truncation requirements; (6) use other files; and (7) creating a new search entirely (from existing searches and/or new terms). For any technique the user can find assistance in how to develop that technique for his current problem. For example, sample searches in which changing ANDs to ORs are given and particular searches of the current session to which this could apply are indicated. Also, ways of finding additional files (by revisiting and delving more deeply into the help-files assistance stages) or terms (e.g., by scanning the index for a file or displaying document information) can be pursued with the detailed assistance given as needed.

These various assistance stages lead the user to choose specific terms or searches to operate on and to define the exact nature of the change. ASSIST then issues the appropriate CONSTRUCT commands to generate the desired modified searches and search strategy. Note that new searches are built, where possible, by *combining* (not *removing*) existing searches. This permits two related capabilities: (1) search execution at the remote retrieval system need only be incremental and thus can avoid costly repeating of searches; and (2) previous searches remain to be used for later display, comparison, or modification purposes.

CONIT was also enhanced by providing for user-specified truncation and proximity searching (by specified maximum number of intervening words or by within-field limitation). These new functionalities are available in the common command language, in the construction mode, and in assistance (menu) mode. *Operators* have been defined and implemented in the command mode which permit not only the *setting* of the desired specification within a *new* search but also the *modification* of *existing* searches (e.g., changing proximity or truncation specifications in particular ways).

We have also implemented several facets of recall estimation based on two distinct methodologies. In one methodology the number of relevant documents found, and the subset of those which were *previously known* by the user, are taken as a fraction of the *a priori* estimates of these numbers given by the user as problem qualifiers to arrive at two (somewhat independent) measures of recall and precision. Numbers of relevant documents found in a given set are extrapolated, if necessary, from the sample actually viewed by the searcher. If the user gives responses enabling CONIT to estimate the fractional recall based on previously known document figures, CONIT computes the newly derived recall base, compares it with the one based on the user's *a priori* estimation of the number of relevant documents, and apprises the user of any discrepancy.

The second methodology of recall estimation is based on a model of search comprehensiveness we have developed ([OVER74], [MARC83a]) which states that if terms (and files) are incrementally added in order of relevancy (most relevant first) -- a condition one would expect to be approximated for reasonably subject-knowledgeable searchers, then recall would be increased by steadily decreasing increments. To make this model useful in a computational setting the additional assumption has been made that, in fact, the number of unretrieved relevant documents is reduced by a constant factor for each added term. This leads to the formula for fractional recall [MARC83a]:

$$r = 1 - (1-k)^n \quad (1)$$

where  $k$  is the constant factor and  $n$  is the number of terms searched. We have extended this basic formula to cover the situation where a search has several conceptual factors to be searched in several databases. Under suitable independence assumptions (discussed in [CHON86]) one can account separately for all of these components in a multiplicative formula -- one factor (given by an expression of the kind shown above) for each conceptual factor and one additional factor for the number of databases searched:

$$r = (r_d)(r_{cf1})(r_{cf2}) \dots (r_{cfn}) \quad (2)$$

where

$r$  = overall fractional recall

$r_d$  = fractional recall attributable to database component

$r_{cfi}$  = fractional recall of documents related to conceptual factor  $i$   
( $i=1,2,\dots,nf$ )

Each of the component recall factors in (2) is derived from the general formula (1); for  $r_d$   $n$  is the number of databases searched and  $k = k_d$  (for which we have chosen 0.8) and for  $r_{cfi}$   $n$  is the number of terms searched and  $k = k_c$  (for which we have chosen the constant 0.75).

### 3.0 Experimentation

We have conducted a number of evaluative studies of the above system which we call stage-1 Expert CONIT. Besides our own testing as system developers, there have been (1) demonstrations to Information Scientists and others at our laboratory and at conferences; (2) a few uses by the M.I.T. Laboratory for Computer Science Medical Group; and (3) two formal studies: one completed as part of a Bachelor's thesis effort [YIP87] and a second undertaken as a project for the Defense Technical Information Center (DTIC) [MARC88].

The MIT LCS Medical Group usages were conducted by users who, mainly, had previously learned the command mode of standard CONIT (see [MARC85] for discussion of previous experiments). These users, and a few new users who were taught the necessary commands by their colleagues in the group, continued to use the efficient command mode in an effective way. Several of the more frequent users appeared to attain proficiencies that matched or surpassed what is attainable by expert human information specialists acting as search intermediaries. (Of course, this success is achieved, in part, due to their being able to perform their own searching without the complications of a human intermediary.) In a few cases the users took advantage of the ASSIST menu mode through use of the HELP command in order to remind themselves of command operations that they had forgotten. Their success in using CONIT in this manner lends support to the contention that the integrated command and menu modes can be helpful to experienced users.

The two formal experimental studies were aimed at new users of CONIT who would try out the newly added features of the stage-1 expert CONIT. In the first study there were 5 experimental users (primarily graduate, post graduate, and undergraduate students at several colleges in the Boston area) none of whom had previously used retrieval systems. These 5 users ran 11 CONIT sessions involving 14 different problems. The second study involved 10 users who were professionals at DTIC and other U.S. government agencies or

contractors. These 10 users, who ran 33 sessions on 12 topics, had previous experience with retrieval systems ranging from end use through human intermediaries to being information specialists themselves. Subjects searched by the various users ranged over a variety of topics.

Despite the relatively small number of searches we have obtained very useful information on the utility of our new techniques by virtue of the extensive nature of the sessions and their attendant analyses. For example, the average session length in study 1 was 66 minutes and the median time to search one topic in study 2 was 58 minutes. In study 2 it was observed that for each minute of connection online users averaged about 3 interactions with CONIT (i.e., 3 inputs to the system); the 33 sessions lasted just over 18 hours and thus involved 3,000 inputs from users (and, of course, a like number of responses from the CONIT system). Users were given a pre-use briefing and a pre-search questionnaire as well as post session debriefing and questionnaires. Additional information was collected through spontaneous comments given during sessions with the COMMENT command or as annotations on the printed typescripts of the sessions. Also, user relevance judgments on retrieved documents were obtained. Parameters derived include amount of time spent in various aspects of the search, number and kind of commands used, and numbers of documents retrieved, shown (online and offline), and judged relevant.

All users in study 1 agreed that CONIT was a very effective retrieval tool. They found an average of 9 relevant documents for each of the 14 topics. Menu assistance permitted reasonably rapid learning and achievement of successful results -- about 50 minutes per problem -- especially in light of time required to prepare the detailed problem description, review and select files, and learn about and make search strategy modifications. Compared to previous CONIT experiments, it appeared that users achieved higher precision by virtue of the fact that in 5 out of the 14 problems users successfully navigated the assistance menus to select and execute appropriate search "narrowing" operations. The construction features and reconnect and re-use of dropped DIALOG searches enabled saving of online retrieval system connect time and costs. On the negative side, a number of system bugs were uncovered and most users complained that the amount of instructional information provided by CONIT was at least somewhat daunting.

In study 2 we found that 7 of the 10 users advanced in their sessions to the point where searches were actually performed. Four of the users performed searches on two topics each making a total of 11 topics searched among those 7 users. For all 11 topics documents were retrieved; for 8 of the 11 topics we have relevance judgments. In 3 cases very significant recall results were obtained ( $R > 100$  in 2 cases,  $R > 50$  in a third). One case had moderate recall ( $R = 12$ ) and the other 4 cases had low recall (from 1 to 3 relevant documents).

In interpreting these results we see indications that range from positive to cautionary to negative. Positive indications include satisfactory recall and session time, generally positive user reactions, and some indications of actual or potential success for the new techniques introduced into expert CONIT. The cautionary indication is that analysis of these results do not show, on the whole, appreciably better results than were obtained with standard CONIT. The negative indications include user complaints that excessive reading of text was required, a large number of non-productive sessions, and the limited use of many of the CONIT features.

Our detailed analyses of the experiments leads us to offer explanations for the various indications listed above and tentative conclusions about the potential utility of the expert

retrieval assistance techniques we have been investigating. Several factors appear to be at work in diminishing the relative force of the positive indications. An underlying factor is the preliminary character of the CONIT system tested and, especially, its computer-human interface component. In particular, bugs in the system often caused session abortions or, at least, forced users to drop potentially productive lines and seeks ways around the failures. A typical bug, for example, resulted from a mistake in the stage transition table causing the user to be thrust into an unintended path on the assistance network for a given option choice. In a number of situations, especially for the far-flung usages in study 2, telecommunications failures caused line drops which halted sessions (although users *could* dial back into Multics and reconnect to their CONIT process, they often did not understand that option or felt it too onerous). It is also clear that "smoother" assistance logic needs to be developed in many cases -- for example, simplifying the wording of textual explanations for easier understanding. It is evident, also, that the perceived excessive *amount* of textual explanation is a strong negative factor for most users.

....Users resist learning about the system specifics, learning about searching in general, and describing their problem in detail. System explanations on why these activities could be helpful to the user may overcome users' reticence but, perhaps more often, may simply confound the problem of perceived unwanted system verbosity.

As our previous studies have shown [MARC83b], there is a need for a whole spectrum of teaching aids and methodologies. Users expressed a desire for printed (offline) instructional material of different depths and varieties and for direct human assistance in the form of demonstrations, guidance (when online or after sessions) via telephone and/or computer messaging. In fact, many of these modalities *were* available in one form or another to many of the users but not in as comprehensive, coordinated, and debugged (smooth) a manner as desirable. Some users felt, and we tend to concur, that a full-fledged *course* of demonstrations and guided and experimental usage of the *current* CONIT could be an excellent method for teaching new searchers about search methodology and practice.

Clearly, as we have previously discovered ([MARC85], [MARC83b]), motivation and context are critical to a user's persistence and, hence, to the success of the search. The users in study 1 appeared to have greater motivation than those in study 2; they were, in general, serious students with a critical need for the information they sought whereas the users in study 2 in several cases appeared to be merely "trying out the system," without a very pressing need, partly to please the experiment organizers. Study 2 users also had greater accessibility to other mechanisms for accessing the online databases (e.g., company librarians) and they lacked the personal presence during their sessions of human experimental monitors who, in study 1, were available to insure initial connection to CONIT, recover in case of a system crash, and (subliminally, at least) encourage users to persist in their searching (though, of course, *explicit* help with searching was excluded). The experimental results appeared clearly influenced by these environmental factors: study 2 users as a group had (1) more aborted or abbreviated (non-productive) sessions, (2) more complaints or reservations about the system, and (3) made fewer attempts to switch from menu to commands and utilize search modification assistance modules.

Other socially conditioned attitudes also seemed to affect results. Users' concern with costs, which CONIT prominently discusses and displays, inhibits users' persistence. In the opinion of this author, at least, users are conditioned to believe bibliographic searching is *not* intellectually challenging and should be easy to perform and, in any case, a comprehensive bibliographic search is not critical to one's professional efforts. Such

attitudes clearly have a negative impact when the computer system is trying to get the searcher to put an effort into learning about and pursuing search processes.

Another reason for the lack of comparative positiveness in the current results is the fact that a number of the factors of greatest importance for assisting users achieve effective retrieval were already included in standard CONIT. These include the keyword-stem approach to searching natural language phrases in heterogeneous databases; the facility to translate automatically search strategies to any system and database; and the facility to generate, record, and automatically regenerate searches as needed. Thus, previous experiments with users of standard CONIT already demonstrated recall effectiveness that matched that achieved by human expert searchers (see [MARC83b]).

#### 4. Conclusions and Future Directions

Our conclusion from the development, experimentation, and analyses we have reported on in this paper is that expert retrieval assistance via techniques we have proposed does provide a potential for vastly improved search effectiveness for a wide class of users. However, the development effort to realize this potential is large (a confirmation of conclusions previously enunciated [MARC86]) and, as yet, only partly completed. The particular efforts most needed to realize the potential are included in our design for the stage 2 Expert CONIT system: (1) a human-computer interface based on modern workstation facilities; (2) completion of our plans for intelligent agents for enhanced assistance (see below); (3) greater automaticity of assistance -- i.e., more computer control to balance the current emphasis on human control of search evaluation and modification and provision for the desired *mixed initiative* capability [MARC83b]; (4) smoothing of explanations based on user reactions uncovered in experimental usage; and (5) full debugging of system to insure workability and reliability. An outline of our planned efforts in these directions follows.

The desire to enhance interface capabilities, provide a more exportable system, and reduce costs -- together with the coming demise of our Multics mainframe -- led to the plan to convert our system to a workstation environment. Although, there is a natural association of current CONIT menus with windows, we intend to go beyond a simple one-for-one association; explanations will be much abbreviated with details, when needed, easily callable, by the user and/or system, as pull-down menus or pop-up windows. For reasons of compatibility, cost, and availability -- as well as functionality -- we decided to move to a MicroVax workstation in a UNIX/C environment (see Section 5 for further rationale and plans). We have converted about 75 percent of the total lines of code from PL1 to C. This represents over 90 percent of the code we need for a "core" system for running CONIT. Approximately 60 percent of the converted code has been debugged.

Search evaluation will be made much more precise with the introduction of new estimation procedures based on user's relevance judgements on retrieved documents. These same relevance judgements, expanded by virtue of having the user select from pre-defined *reasons* for *irrelevance*, will be used to automate the selection of search strategy modifications for improving retrieval results. Furthermore, a model of search precision as a function of degree of exactness of retrieval match will enable *ranking* the documents retrieved by their likely relevance. In addition, more automated techniques for selecting potentially relevant files based on previously tested CONIT techniques not incorporated in the current system ([MARC81] and [MARC83a]) are planned to be re-introduced.

## ACKNOWLEDGMENTS

The research described in this paper was supported by the National Library of Medicine under Grant LM-03210, by the National Science Foundation under grant IST-84-14485, and by the Logistics Management Institute in conjunction with the Defense Technical Information Center. The author also acknowledges the many vital contributions of his project co-workers including students Man-Kam K. Yip, William T. Lee, Steven H. Schwartz, Monica R. Gerber, Ricardo A. Cardenas, Hing-Fai Louis Chong, Aleks Gollu, and Man-Wah Colina Yip and Visiting Scientist Ann I. Johansson.

## REFERENCES

[CHON86] Chong, Hing Fai Louis. "Recall Estimation for Information Retrieval Assistance." Bachelor of Science thesis in Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA.; September, 1986.

[HUSH86] Hushon, Judith M. "How Micro-CSIN, a New Gateway to Online Systems, Stacks Up." *Proceedings of the Seventh National Online Meeting*. 203-210. New York; May, 1986.

[LISA84] Lisanti, Suzana. "The On-line Search." *BYTE*. 9(13):215-230; December, 1984.

[MARC81] Marcus, R.S. "An Automated Expert Assistant for Information Retrieval". *Proceedings of the 44th ASIS Annual Meeting*. 18:270-273; October, 1981.

[MARC83a] Marcus, R.S. "Computer-Assisted Search Planning and Evaluation." *Proceedings of the 46th Annual Meeting of the American Society for Information Science*. 20:19-21; October, 1983.

[MARC83b] Marcus, R.S. "An Experimental Comparison of the Effectiveness of Computers and Humans as Search Intermediaries." *Journal of the American Society for Information Science*. 34 (6):381-404. November, 1983.

[MARC85] Marcus, R.S. "Development and Testing of Expert Systems for Retrieval Assistance". *Proceedings of the 48th ASIS Annual Meeting*. 22:289-292; October, 1985.

[MARC86a] Marcus, R.S. "Issues for Expert Retrieval Assistance." *Abstracts of Papers, Second Conference on Computer Interfaces and Intermediaries for Information Retrieval*. 2:38. Boston; May, 1986. (Defense Technical Information Center Report)

[MARC86b] Marcus, R.S. "Design Questions in the Development of Expert Systems for Retrieval Assistance." *Proceedings of the 49th Annual Meeting of the American Society for Information Science*. 23:185-189; September, 1986.

[MARC88] Marcus, R.S. Experimental Evaluation of CONIT in DGIS Environment. MIT Laboratory for Information and Decision Systems Report LIDS-R-1746. February, 1988.

[WEIN86] Weinberger, Marvin. "Knowledge Gateways: A State-of-the-Art in Information Integration." *Beyond Databases: Knowledge Delivery Systems Come of Age*. Information Industry Association Seminar, Cambridge, MA, August, 1986.



[WILL86] Williams, Martha E. "Transparent Information Systems Through Gateways, Front Ends, Intermediaries, and Interfaces." *Journal of the American Society of Information Sciences*. 37(4):204-214; July, 1986.

[YIP81] Yip, Man-Kam. An Expert System for Document Retrieval. Master of Science Thesis in Electrical Engineering and Computer Science, Massachusetts Institute of Technology. Cambridge, MA. February, 1981.

[YIP87] Yip, Man-Wah Colina. Testing and Evaluation of the CONIT Retrieval Assistance System. Bachelor of Science thesis in Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA; June, 1987.