June, 1982

ASK FOR INFORMATION RETRIEVAL: PART I. BACKGROUND AND THEORY

Robert N Oddy
N J Belkin
H M Brooks

ASK FOR INFORMATION RETRIEVAL: PART I. BACKGROUND AND THEORY

N.J. BELKIN, R.N. ODDY, H.M. BROOKS

Emerald Article: Journal of Documentation, Vol. 38 Issue #: 2 pp. 61 - 71

Permanent link to this document:
http://dx.doi.org/10.1108/eb026722

We report the results of a British Library Research and Development Department funded design study for an interactive information retrieval system which will determine structural representations of the anomalous states of knowledge (ASKs) underlying information needs, and attempt to resolve the anomalies through a variety of retrieval strategies performed on a database of documents represented in compatible structural formats. Part I discusses the background to the project and the theory underlying it, Part II (next issue) presents our methods, results and conclusions. Basic premises of the project were: that information needs are not in principle precisely specifiable; that it is possible to elicit problem statements from information system users from which representations of the ASK underlying the need can be derived; that there are classes of ASKs; and, that all elements of information retrieval systems ought to be based on the user's ASK. We have developed a relatively freeform interview technique for eliciting problem statements, and a statistical word co-occurrence analysis for deriving network representations of the problem statements and abstracts. Structural characteristics of the representations have been used to determine classes of ASKs, and both ASK and information structures have been evaluated by, respectively, users and authors. Some results are: that interviewing appears to be a satisfactory technique for eliciting problem statements from which ASKs can be determined; that the statistical analysis produces structures which are generally appropriate both for documents and problem statements; that ASKs thus represented can be usefully classified according to their structural characteristics; and, that of thirty-five subjects, only two had ASKs for which traditional 'best match' retrieval would be intuitively appropriate. The results of the design study indicate that at least some of our premises are reasonable, and that an ASK-based information retrieval system is at least feasible.

INTRODUCTION

IMPROVEMENTS IN THE performance of information retrieval (IR) systems as presently designed seem to be limited to only marginal gains in terms of complete recall and precision or complete user satisfaction (see, e.g. Robertson and Sparck Jones1). In these two papers *(Part II: Results of a design study,* to appear in *Journal of Documentation,* vol. 38 no. 3) we report on a design study2 for an experimental IR system based on radically different hypotheses than those underlying present systems, which we think may allow the design of IR systems which produce significantly better performance than currently offered.

Recent work by Belkin3-6 and others (see, e.g. Harbo and Kajberg7 and Hollnagel3) has called into question some traditional assumptions of IR, in particular those concerning: the relationship of the request put to the IR system to the information need underlying the request; the basis for text and request representation in IR systems; and, the retrieval mechanisms suitable for IR. This new approach recognizes that a fundamental element in the IR situation is the development of an information need out of an inadequate state of knowledge. From this realization it goes on to say that for IR to be successful, that information need must be represented in terms appropriate for just that task, with the remaining elements of the system (i.e. document representation, retrieval mechanism) being represented or constructed on the basis of that representation. One means to an appropriate representation is consideration of the information need as an 'anomalous state of knowledge' (ASK).6,9

The ASK hypothesis is that an information need arises from a recognized anomaly in the user's state of knowledge concerning some topic or situation and that, in general, the user is unable to specify precisely what is needed to resolve that anomaly. Thus, for the purposes of IR, it is more suitable to attempt to describe that ASK, than to ask the user to specify her/his need as a request to the system.

Oddy's10-12 experimental interactive IR system, THOMAS, was designed to allow information retrieval without query formulation, relying on the system to construct an 'image' of the user's need. The program operates upon a graph whose points represent documents, subjects and authors; the lines stand for associations. A tructural image of the subject area of interest to the user is maintained during the online dialogue. Its main component is a subgraph of the program's document collection structure. Strictly speaking, this image is not asserted to be a representation of the user's need, but rather a formal context within which documents satisfying the need might be found. Thus it is the structural properties of the image, and not merely matching terms, which determine the program's choice of documents to show the user. The user may react to the documents, and their descriptions, and her/his response is used by the program to determine modifications to the image. This work has strong affinity to the ASK hypothesis, but is limited in its means for representation of document and user to traditional document descriptors, although in explicitly structural terms. In the design study, we investigated means of combining and extending these two developments to deal with complex IR situations in an entirely new way, which the ASK hypothesis predicts would produce better results than traditional IR systems. The design study attempted to resolve a number of problems raised by the hypotheses, and to provide a preliminary specification for a system which incorporates these new assumptions.

THEORETICAL BASIS OF THE ASK IR SYSTEM

The basis of the project we describe here is a combination of what we have rather loosely termed the 'ASK hypothesis' with the principles underlying Oddy's THOMAS system. In this section, we expand somewhat upon the theory underlying these two approaches to IR, and on how they can be combined into a framework suitable for what we think of as a 'second-generation' IR system. We begin by discussing what we see as the basic weakness of current IR systems. The typical IR system now available, either operational or experimental, depends on what we call the 'best-match' principle. This principle, briefly stated, is that, given a representation of a request for information (e.g. a 'query' or set of index terms), the best possible system response will be the text whose representation most closely matches it. Leaving aside, for the moment, the problem of representation, one can identify two assumptions basic to the best-match principle: that it is possible for the user to specify precisely the information that she/he requires; and, that information needs (or at least expressions of them) are functionally equivalent to document texts.

Let us make clear how the best-match principle depends upon these two assumptions. It depends upon the first in that only those concepts and/or relations which are explicitly stated by the user as being significant are significant in the matching procedure. The user must be able to specify all of the relevant aspects of the problem, in order for the system to work optimally. If they are not stated, then documents treating these concepts (relations) cannot be retrieved (since they do not match what was specified), or will be ranked only very low in terms of probability or degree of relevance (since they match the specification only poorly). The best-match principle depends upon " assumption of equivalence between expression of need and document text in that it treats the representation of need as a representation of the document ideal for resolving that need.13 The best-match principle looks first for a document which is just like the expression of need; that is, which is functionally equivalent to it.

Whether the best-match principle is a fundamental weakness in current IR systems is a function of the reasonableness of the two assumptions on which it is based. Simplifying assumptions, of course, are basic to the scientific enterprise. Later, we make some that may seem rather simple-minded, as well as merely simple. Nevertheless, such assumptions ought not stray too far away from what we conceive as the more complicated reality. Our contention, of course, is that the assumptions underlying the 'best-match' principle are sufficiently divorced from reality to make them quite untenable.

The specifiability assumption is by no means an obvious characteristic of the typical IR situation, which Wersig14 has characterized as rooted in a 'problematic situation'. The problematic situation, and other related suggestions (e.g.8,15) emphasize that information need is in fact not a need in itself, but rather a means toward satisfying some more basic need, typically, in the situations with which information science is concerned, the resolution of a problem. The most general thing that one can say about such a circumstance is that the user, faced with a problem, recognizes that her/his state of knowledge is inadequate for resolving that problem, and decides that obtaining information about the problem area and its circumstances is an appropriate means towards its resolution. There are certainly occasions when one might be able to specify precisely what information is required to bring the state of knowledge to a structure adequate for resolution of the problem, but it seems obvious that the more usual situation will be that in which what is appropriate for the purpose is not known in advance. In such a situation, the best-match strategy does not seem a reasonable first choice for IR purposes.

The assumption that expression of information need and document text are functionally equivalent also seems unlikely, except in the special case in which the user is able to specify that which is needed as a coherent or defined information structure. A document, after all, is supposed to be a statement of what its author knows about a topic, and is thus assumed to be a coherent statement of a particular state of knowledge. The expression of an information need, on the other hand, is in general a statement of what the user does not know. Thus, the document is a representation of a coherent state of knowledge, while a query or other text related to an information need will be a representation of an anomalous, or somehow inadequate or incoherent state of knowledge. To assume that these two kinds of texts are alike in the characteristics useful for IR appears unwarranted and at the very least ought to be investigated before being applied. Despite some work in philosophy, linguistics and artificial intelligence which seems to indicate that questions are different in kind from expository statements (e.g.16-18), IR based on the best-match principle uncritically accepts their equivalence.

The contradictions between 'reality' and the assumptions underlying best-match IR have indeed not gone unnoticed. Associative retrieval[19] and relevance feedback,[1] for instance, are significant attempts to resolve the problems inherent in the best match principle. To some extent these techniques are undeniably successful. Relevance feedback, for instance, can be shown to improve retrieval performance in a variety of situations.[1] Nevertheless, one should note that neither of these approaches takes explicit account of the user's inability to express the information need in terms of the need itself. Both techniques attempt to deal with this problem on the basis of relationships among the documents in the collection, primarily those of word co-occurrence. Relevance feedback seems more responsive to the need issue, in that it uses information gathered from relevance judgements in modifying query formulations, and is therefore not solely dependent upon the characteristics of the document collection. It remains within the best-match paradigm, however, since it assumes that the eventual query formulation is equivalent to the ideal document. Associative retrieval extends the search without knowledge of the individual user's requirements on the basis of document collection characteristics alone. In a sense, this is a genuine recognition of the specification problem, but its distance from the user makes it somewhat unrealistic. Furthermore, it also assumes that the text and need representations are functionally equivalent.

These techniques have resulted in improvements in performance of IR systems, but improvements which, in terms of total recall and precision, or complete user satisfaction, are only marginal. In the end, the results of such modifications of the best-match principle should, we believe, encourage one either to look for theoretical limits on the performance of IR systems, or, to look for principles that take better account of the 'reality' of the IR situation. We opt for the second strategy as the more optimistic, although we recognize that the two are not mutually exclusive.

The assumptions underlying best-match IR mean that such systems cannot use information from the user about doubt, uncertainty or suspicion of inadequacy in the user's state of knowledge, since these factors will not be specified in the usual document, and are difficult for the user to specify. These, however, are the factors which prompt people to go to IR systems, and which ought therefore to be taken explicitly into account by any realistic IR procedure. One reason that best-match IR does not do this is because the problem is, in principle, extremely difficult. We suggest that one approach to taking account of these factors, which we call 'anomalies', is to try to represent them in terms of the user's larger-scale intentions and goals, without asking the user to specify the information needed to resolve the anomaly. Oddy's work on IR systems without query specification, combined with Belkin's speculations about the knowledge structures underlying information needs, seem to us appropriate means to this end.

Our basic orientation to the problem of combining these two approaches is the cognitive viewpoint, which suggests that interactions of humans with one another, with the physical world and with themselves are always mediated by their states of knowledge about themselves and about that with which or whom they interact. Additionally, we look at the IR situation as a recipient-controlled communication system, aimed at resolving the expressed information needs of humans, primarily via texts produced by other human beings. This communication system is outlined in Figure 1.

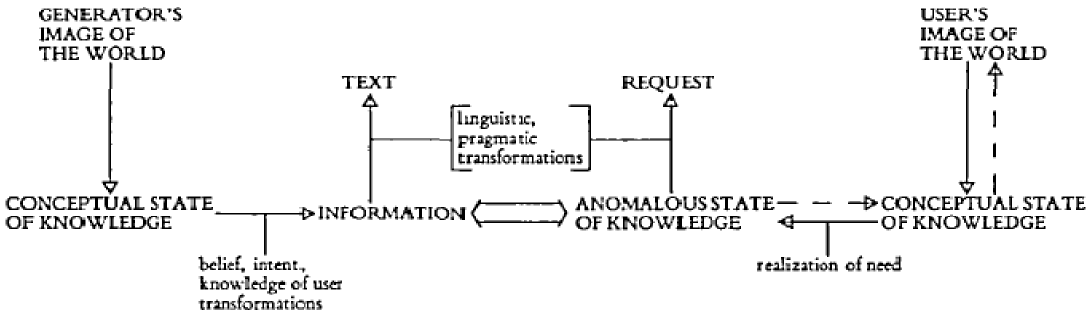

FIG. 1. *A cognitive communication system for information retrieval* (from Belkin,[9] p. 135)

Figure 1 suggests that the driving force of the IR situation is the user's problem, which leads to recognition of an inadequacy in her/his state of knowledge (an anomaly with respect to the problem). Whether an anomaly is resolved or not is evaluated in terms of the problem. It should be noticed that the problem determines not only the conceptual requirements of an appropriate response, but also the situational requirements such as, for instance, source of information, mode of information etc. (see Cooper22 and Wilson15 for more on such factors). It should also be noted that the anomaly, and the user's perception of the problem, will probably change with each instance of communication between user and mechanism. This dynamism implies that information systems ought to be highly iterative, and interactive.

Hollnagel has suggested a cognitive communication model3 which is more active than that of Figure 1, and which perhaps takes better account of the interactive nature of this situation. His model suggest that the parties to the situation, the user and the mechanism, can be considered as partners in a dialogue, each, among other things, adapting its model of the other, and of the world, and therefore its response, according to what it has learned via communication with the other. Accurate models, and helpful dialogue then can in general be best arrived at through a series of interactions, rather than just an initial statement and subsequent response. This view of the IR situation leads to the conclusion that IR systems ought to be designed to be iterative and interactive. Relevance feedback is an example of a step in this direction.

Thus, the cognitive view of the IR situation leads us to some design principles for IR systems which we think incorporate the THOMAS and ASK approaches. These design principles require representation of the user's anomalies, evaluation in terms of the problem the user faces and iteration and interaction in retrieval.

Rather than to attempt to test the theoretical position underlying the ASK/THOMAS approach to IR by generating hypotheses from the theory which lead to falsifiable predictions, we decided to incorporate the theory in a general system design, and then to test the system against some well-specified goals. This general approach to understanding of a theory seems to us to be especially appropriate in our situation, where the theory itself is only partially developed in well-specified form, and where the problem area has a number of highly interactive components which may change significantly between experimental and 'real' settings. In such a situation, it seems reasonable to adopt an iterative methodology, in which the general principles lead to a specific system design, which is tested by reference to its feasibility in a real environment and its efficacy in solving the problem which it is set. At each stage of testing, the design can be modified according to results of the test, and thus the method is iterative, but in the end it is the system as a whole which is tested, rather than some particular aspect of it. This sort of program should begin with a general system based on theoretically motivated design principles, and with some study aimed at discovering whether the design is feasible. It is to these two points, in particular the latter, that the Design Study2 was directed.

THE STRUCTURE OF THE ASK IR SYSTEM

Our first problem then was to design an IR system which would incorporate our theoretical ideas. Again, our basic premises arise from what we consider to be one of the central difficulties of IR: that people who use IR systems typically do so because they have recognized an anomaly in their state of knowledge of some topic, but they are unable to specify precisely what is necessary to resolve that anomaly. This can be seen as a restatement and perhaps extension of ideas proposed by Taylor.23 Thus, we presume that it is unrealistic (in general) to ask the user of an IR system to say exactly what it is that she/he needs to know, since it is just the lack of that knowledge which has brought her/him to the system in the first place. This premise leads us to conclude that IR systems should be designed with the non-specifiability of information need as a major parameter. Furthermore, we assume that we can get closer to appropriate responses in an IR system via iteration and interaction. What sort of IR system could this be?

We consider that IR systems, in general, consist of: a mechanism for representing information need; a text store; a mechanism for representing and organizing texts; a mechanism for retrieving texts appropriate to particular information needs and, usually, a mechanism for evaluating the effectiveness of the retrieval. Figure 2 indicates these components, and some other features, and the relationships among them.

From Figure 2, and experience, one can see that the starting points for IR system design are at either text or need representation, and that which of these one chooses, and the chosen method of representation, will strongly influence all the other elements of the IR system. Most previous systems have begun from the text representation end, with not very much influence from the need end. We have noted that need representation appears to be the fundamental problem of IR, and so we suggest that a good IR system should be one which begins with need representation and designs the rest of the system about a mechanism and formalism specifically designed for that purpose.

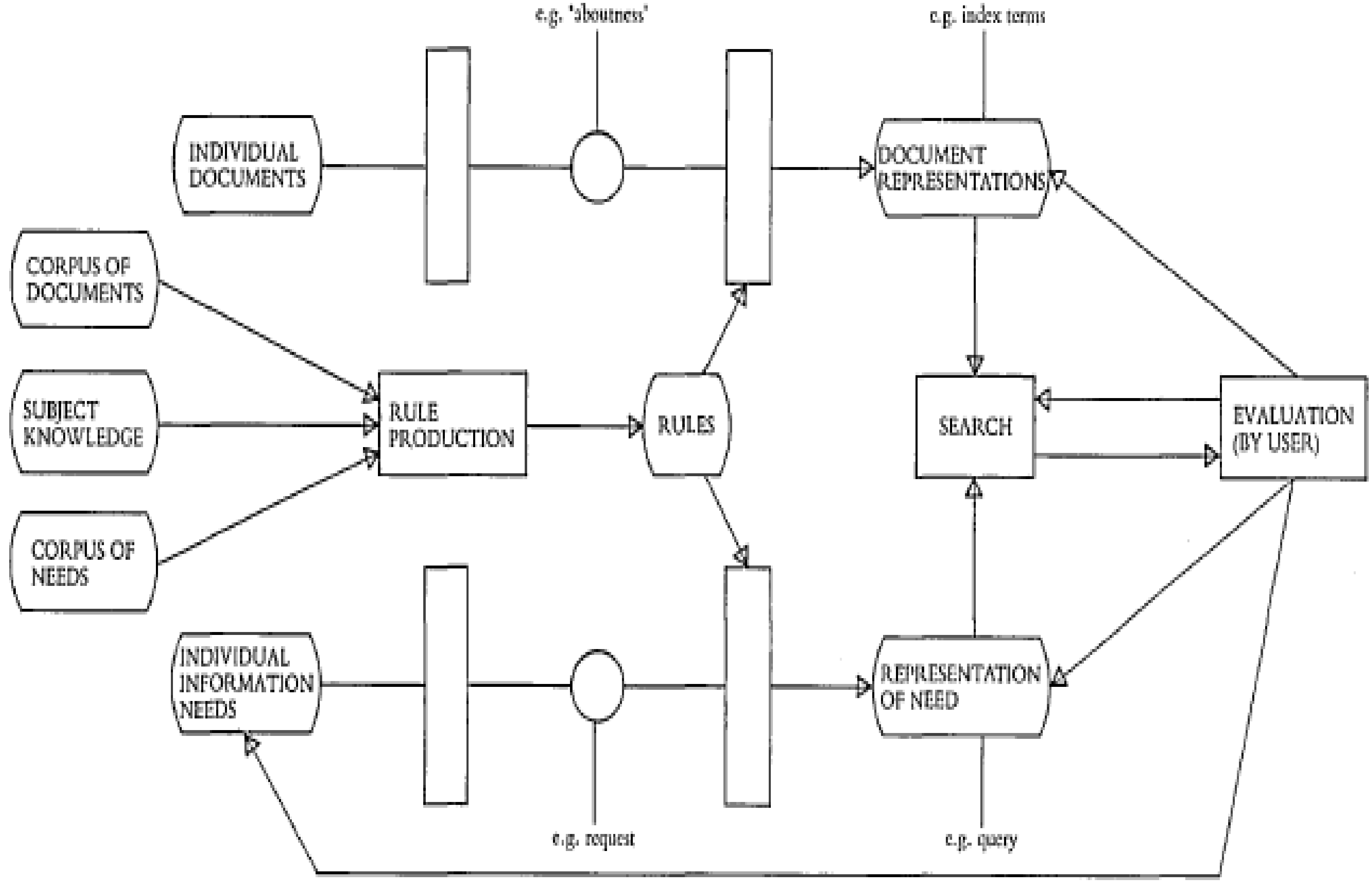

FIG. 2. *A model of information retrieval* (after Robertson[24])

Our mechanism is based on some experimental work by Belkin,4 which indicated that networks constructed from constrained word associations yield reasonable representations of individuals' states of knowledge about the subject to which the associations are constrained. In the same experiment, a network generated from word co-occurrence analysis of a 3,000 word text (using the stimulus words of the word association test as major nodes) gave a reasonable representation of the 'information structure' of that text. The argument for using these techniques is based on the assumptions that concepts (represented by words) which are closely associated in an individual's state of knowledge will (a) be recalled close to one another in tasks such as word association; and, (b) occur in close proximity to one another in a text by that person on the specific topic.

The formalism used for representation, a network of concepts represented by words, depends upon our concept of a state of knowledge as a multi-dimensional structure. We have discussed this concept, as it applies to our work,3-4,10 but the basic idea seems to us unexceptionable, except perhaps for the substitution of words for concepts. A network is certainly an effective way of representing this type of structure, and has the advantage, in our case, of being relatively easily derived from the sort of association data we use.

There is no question that this sort of representation of a state of knowledge (or of the information structure underlying a text) is simplistic and naive, if one is attempting to obtain detailed representations for such purposes as natural language understanding, machine translation or retrieval from memory. On the other hand, it has the advantages of being fairly easily determined and reasonably machine manipulable, important considerations in an IR context, where one needs to represent actual information needs and to manipulate large amounts of data. Our strategy for choice of formalism and mechanism for representation depends upon our assumption that there is no 'real' or 'true' representation of knowledge or information, but rather many possible representations, each appropriate to particular problems. We decided upon networks derived from association data as a potentially useful representation for the purposes of IR because of the advantages of this type of representation mentioned above, and because there has already been some experience with similar formalisms in IR. But we bear in mind that this is indeed a very simple model, and so treat it only as preliminary: if it works, we will be pleased, if not, then we will amend it in ways the evidence seems to indicate, or discard it for different models.

To this point, we have discussed means of representing information, or what people know about a topic. Yet our problem is to represent what someone does not know; that is, to represent an ASK. We obviously cannot use word association tests for obtaining the data for this purpose, since we would need to represent the ASK corresponding to a different topic for each user of the IR system, and thus would have no standard set of stimulus words. Our solution in this case derived, to some extent, from the suggestion by Wersig14 that a 'problematic situation' underlies an information need. In terms of the ASK hypothesis, this means that the user, attempting to realize a goal or solve a problem, recognizes that her/his state of knowledge is inadequate (anomalous) in terms of the goal; resolution of the anomaly will help to solve the problem. Thus, if we are able to obtain a statement of the problem, this may give us the data from which we can construct a representation of the ASK. This, of course, is old hat to reference librarians. So the basic data for our representation of ASKs are narrative statements by the users of the IR system, of the problems which brought them to the system. We assume that some sort of associative analysis of those statements will yield representations of the ASKs inherent in the problematic situation, which will be useful within the context of an IR system aiming to resolve the ASKs.

We have now discussed representation of need and text. A second major mechanism is that of search, or retrieval. Here, our arguments against the best-match principle become important. We assume that one needs mechanisms which can recognize unspecified anomalies, and which can resolve, rather than match them. We recognize that there are ASKs which can be resolved by best-match type mechanisms, but we also recognize, on the basis of IR system experience, as well as theoretical speculation, that there must be ASKs for which this type of mechanism is not suitable. Therefore, our IR system design assumes that there exist different types of ASK which require different retrieval mechanisms.

The evaluative mechanism in an IR system is now recognized to have an importance equivalent to the representation and retrieval mechanisms, and this is the point at which we see the interactive component become most visible. There are many difficulties in establishing this mechanism, most of which we attempt to avoid here. But the basis of any evaluation must be the satisfaction of the user of the IR system with the system's performance in resolution of her/his ASK. Oddy10 built his system THOMAS about this realization by making evaluation its driving force. THOMAS is truly interactive in that each retrieval depends upon the user's evaluation of the results of the previous retrieval. Thus, the evaluation mechanism feeds back directly to the need-representation mechanism, modifying it according to the user's immediate, perceived judgement of the value of the mechanisms for resolving her/his information need. This sort of evaluation can thus cater for initial misrepresentation by the IR system and for changes in the anomaly and the problem which arise during the course of the search. In the IR system which we propose here, an additional factor for evaluation, multiple retrieval mechanisms, has been introduced. We assume, therefore, that an effective ASK-based system must be interactive in such a way as to incorporate direct evaluation by the user of retrieval output in terms suitable for modification of both need representation and choice of retrieval mechanism.

Thus, we propose that an IR system based on the representation of information need as an ASK look something like that outlined in Figure 3. We do not mean to suggest that this is the only possible such system, but rather only that this is a realization of these ideas, based on further assumptions about representation and on practicalities of IR.

1. USER'S PROBLEM STATEMENT
2. STRUCTURAL ANALYSIS OF PROBLEM STATEMENT
3. CHOICE OF RETRIEVAL STRATEGY ACCORDING TO TYPE OF ASK
4. ABSTRACT PRESENTED TO USER SIMULTANEOUSLY WITH EXPLANATION OF WHY TEXT WAS CHOSEN (STRATEGY AND SIGNIFICANT FEATURES)
5. STRUCTURED DIALOGUE BETWEEN SYSTEM AND USER FOR SYSTEM TO INFER USER'S EVALUATION OF:
(A) METHOD OF CHOICE
(B) SUITABILITY OF DOCUMENT TO PROBLEM
AND/OR
(C) WHETHER NEED HAS CHANGED
6. MODIFICATIONS ACCORDING TO EVALUATION OR FINISH
7. RETURN TO 2 OR 3 AS NECESSARY

FIG. 3. ***An ASK-based information retrieval system design***

The system would work as follows:
1. The user discusses her/his information problem in an unstructured statement (say, 2-3 paragraphs long).
2. The problem statement is converted to a structural representation of the user's ASK by the text analysis program.
3. According to the type of problem structure (PS), one of several available retrieval mechanisms is chosen to interrogate the database (each member of which is represented by a structural representation of the information associated with the text). In this context, a retrieval mechanism is a strategy for resolving the anomalous aspects of a PS.
4. The abstract (i.e. the text) is printed out for the user to read. Simultaneously, the user is presented with a brief explanation of why that particular text was chosen (explanation of the retrieval mechanism), indicating aspects of the text structure which the system finds significant in that choice.
5. The system then initiates a structured dialogue with the user, based on the information presented to her/him, inferring from the response the user's attitude toward:
(a) the method of choice;
(b) the suitability of the text to the problem; and,
(c) whether her/his information need has changed.
6. (a) the system changes retrieval mechanism if necessary, and/or
(b) the system modifies the problem structure if necessary, or
(c) the system stops if the user is satisfied.
7. The system returns to step 2 or 3.

Our design study intended to investigate the following aspects of the system, which we felt needed to be clarified before going on to actually implementing the system.
1. The text analysis program. This was based upon an existing algorithm4 which determines statistically based structures. The suitability of the eventual algorithm must be tested on real problem statements (see 2 below) and on abstracts (see 3 below).

2. The feasibility of obtaining problem statements, the possible types of problem statement, and the analysis of problem statements. To these ends, a number of real problem statements were collected and analysed, and the analyses verified.
3. The feasibility of using abstracts as the basic documents in the system, and the analysis of abstracts. To these ends, real abstracts were collected and analysed, and the analyses verified.
4. Some specific retrieval mechanisms should be proposed, based in large part on the types of problem structures encountered, although the ASK hypothesis implies some strategies.
5. The interactive system (especially the dialogue with the user) should be specified (although not necessarily fixed).

In the event, we were unable to do much with points 4 and 5 in terms of empirical investigation, so that the Design Study aimed finally to consider the first three issues in detail, as necessary preliminaries to the design of an ASK-based IR system. In Part II, (*Journal of Documentation*, vol. 38, no. 3) we describe the methods and present the results and conclusions of the Design Study.

REFERENCES

1. ROBERTSON, S. E. and SPARCK JONES, K. Relevance weighting of search terms. *Journal of the American Society for Information Science, 21,* 1976, 129-46.
2. BEIKIN, N.J. and ODDY, R.N. *Design study for an anomalous state of knowledge based information retrieval system.* British Library R & D Report No. 5547. Birmingham: Computer Centre, University of Aston in Birmingham, 1979.
3. BELKIN, N.J. *and* ROBERTSON, S. E. Information science and the phenomenon of information. *Journal of the American Society for Information Science, 27, 1916,* 197-204.
4. BELKIN, N. J. *A concept of information for information science.* Ph.D. Thesis, University of London, 1977.
5. BELKIN, N. J. Internal knowledge and external information. *In: CC77: The Cognitive Viewpoint.* Proceedings of an international workshop on the cognitive viewpoint, Ghent, 1977. Ghent: University of Ghent, 1977, 187-94.
6. BELKIN, N. J. The problem of 'matching' in information retrieval. *In:* O. Harbo and L. Kajberg, e*ds. Theory and applications of information research.* Proceedings of the Second International Research Forum in Information Science. London: Mansell, 1980, 187-97.
7. HARBO, O. *and* KAJBERG, L., *eds. Theory and applications of information research.* Proceedings of the Second International Research Forum in Information Science. London: Mansell, 1980.
8. HOLLNAGEL, E. The relation between intention, meaning and action. *In:* M. MacCafferty and K. Gray, *eds. The analysis of meaning: Informatics 5.* London: Aslib, 1979, 135-47.
9. BELKIN, N. J. Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science, 5,* 1980, 133-43.
10. ODDY, R. N. *Reference retrieval based on user-induced dynamic clustering.* Ph.D. Thesis, University of Newcastle upon Tyne, 1975.
11. ODDY, R. N. Information retrieval through man-machine *dialogue. Journal of Documentation, 33,* 1977, 1-14.
12. ODDY, R. N. Retrieving references by dialogue rather than by query formulation. *Journal of Informatics, 1, 1911*, 37-53.
13. CLEMENT DAVIES, C. Reference retrieval by user-negotiated term frequency ordering within a dynamically adjusted notional 'document'. *Journal of Informatics, 2,* 1978, 62- 76.
14. WERSIG, G. *Information—Kommunikation—Documentation.* Püllach bei München: Verlag Documentation, 1971.
15. WILSON, T. D. On user studies and information *needs. Journal of Documentation, 31,* 1981, 3-15.
16. BELKNAP, N. D. *and* STEEL, T. B. *The logic of questions and answers.* London: Yale University Press, 1976.
17. HUDSON, R. A. The meaning of questions. *Language, 51,* 1975, 1-31.
18. LEHNERT, w. G. *The process of question answering.* Hillsdale, N.J.: Lawrence Earlbaum Associates, 1978.
19. CROFT, w. B. *Organizing and searching large files of document descriptions.* Ph.D. Thesis, Computer Laboratory, University of Cambridge, 1978.
20. DE MEY, M. The cognitive viewpoint: its development and its scope. *In: CC77: The cognitive viewpoint.* Proceedings of an international workshop on the cognitive viewpoint, Ghent, 1977. Ghent: University of Ghent, 1977, xvi-xxxii.
21. PAISLEY, W. J. *and* PARKER, E. B. Information retrieval as a receiver-controlled communication system. *In: Education for information science.* Washington, D.C.: Spartan Books, 1965, 23-31.
22. COOPER, W. S. On selecting a measure of retrieval effectiveness. *Journal of the American Society for Information Science, 24,* 1973, 87-100.
23. TAYLOR, R. s. Question-negotiation and information seeking in libraries. *College and Research Libraries, 29,* 1968, 178-94.
24. ROBERTSON, S. E. Indexing theory and retrieval effectiveness. *Drexel Library Quarterly, 14,* 1979, 40-56.