Syracuse University

From the SelectedWorks of Robert Oddy

1977

Information Retrieval through Man-Machine Dialogue.

Robert N Oddy, *Syracuse University*

Available at: https://works.bepress.com/robert_oddy1/21/

Information Retrieval through Man-Machine Dialogue.

Emerald: Journal of Documentation, 33(1), 1-14, 1977.

DOI: 10.1108/eb026631 (Permanent URL)

R. N. ODDY

Computer Centre, University of Aston in Birmingham

The purpose of this paper is to introduce a new method of information retrieval by man-machine interaction. The dialogue supported has more symmetry than most interactive computer systems in that the machine forms an image (rather as a man does) of the view of the human enquirer, without requiring him to ask a precise question, and responds with references according to its image. Initial tests with a prototype program indicate that a performance equal to that obtainable from a more conventional on-line retrieval system is possible without obliging the user to formulate his query.

REFERENCE RETRIEVAL AND BROWSING

RETRIEVING REFERENCES to books, papers, reports, and all the other forms of documentation is part of the job of a library system: prerequisite, in fact, to delivering the actual books, or documents, to the reader. It is a task that may be performed, partly or in whole, by the library user himself, and its nature will depend upon the requirement which prompted him to go to the library, and the type of tools provided for this purpose. I shall describe the prototype of a new tool—an interactive computer program—in the context of one way of thinking about the problems of reference retrieval. The philosophy will be familiar to librarians.

Research workers' requirements for information vary, according to the stage that their work has reached. Sometimes one needs factual information, such as is assembled in reference handbooks. At other times, in contrast, one is nagged by an ill-defined need to find stimulation either from literature or from colleagues. There is a continuous spectrum of requirements between these two. I am concerned here with needs that have an element of ill-definition, and that perhaps includes any that are not at the 'factual' extreme of the spectrum. It is important to try to come to grips with the problem of serving a library user who is not able to formulate a precise query, and yet will recognize what he has been looking for when he sees it. A man, left to his own devices among the bookshelves, accomplishes searches of this sort by browsing. Lancaster describes a type of search which undoubtedly occurs frequently in libraries: 'Personal searches tend to be browsing searches . . . Having found some promising references, [the seeker] locates the documents cited and, from the text and bibliographies of these, may be led to other sources or made aware of additional subject labels that might usefully be consulted in the tools with which he began the search. During this whole process, the "information need" tends to be modified, to a greater or lesser extent, by what is found during the search, and the final set of documents, accepted by the searcher as "useful" in relation to his requirements, may be somewhat different in character from the "kinds" of documents he visualized as useful when the search commenced.' (p. 181).

It seems that the notion of *information* in this context is extremely complicated. The concept of information has been discussed by Belkin[3] and Brookes,[6] and they require that a suitable definition should take account of the state of the recipient's knowledge. It is because the information obtained (somehow) from a document alters the mental state of the reader, that he can conduct the type of browse described above. For the same reason, the 'information content' of a book is very likely to differ from one reader to another. For the time being, therefore, it would seem that we need to read books and other documents to obtain certain types of information; and that fact retrieval from some kind of information machine is not sufficient. In designing a mechanical aid to literature searching, we should take the view expressed by the eminent chemist, Lord Todd:[17] 'We must surely make the maximum use of computers and associated automation, but if we carry it to the point where the scientist no longer browses in the literature without first of all formulating questions then I believe we shall do harm to science.' (p. 9).

For some requirements—and they are not uncommon—the ideal search strategy would appear to consist of a visit to the shelves, and a perusal of the books themselves. The difficulty, of course, is in determining an arrangement of the books which assists the user. The arrangement should bring together literature on similar topics but, for the purposes of browsing,[12] it need not take account of the fine detail in the subject matter. Hierarchical classifications, such as that of Dewey and the Universal Decimal Classification, are frequently used by libraries to generate a shelf order for the material. However, searching in the shelves is generally regarded as myopic, except in the smallest libraries, even though it is very often effective. In a large library, books which are potentially useful to one reader may be widely separated spatially, and the separation of the short, but very important, documents published in the many periodicals devoted to any particular subject is much more pronounced. Hence the need for reference retrieval, *even for the browser.*

QUERY FORMULATION FOR REFERENCE RETRIEVAL

The crucial characteristic of a reference retrieval device is that it aims to help the user to make choices from among unseen documents. The searcher wants a document for the (subjective) information it contains, so we have the very difficult problem of finding a proxy for the information, which must be very much smaller and more manipulable than the document itself. We need a symbolic description of the document—there is no question of it being regarded as an alternative form of the information contained in the document, in the sense of information that we have in mind here. The most that we should aim for, at present, is a substitute which the user will interpret as meaning 'this document may contain information I want.' This is what class numbers (Dewey, UDC, Library of Congress, for information service, keywords will serve this purpose better than class numbers, because they are more meaningful. Moreover, numbers are arranged hierarchically, and therefore cannot have the infinite flexibility possible with keywords.

With a good descriptor language, documents which are relevant to a searcher's problem will have descriptions which he *recognizes* as being promising. The emphasis is on recognition: it is not necessarily true that a query can be formulated in advance by the searcher to match those same descriptions. It seems reasonable to assume that there will be some similarity between the descriptions of documents which are relevant to the same query.27 But the nature of the similarity may be very subtle and hard to recognize by anybody other than the enquirer. Even so, most approaches to query formulation attempt to *predict* the descriptions of the required documents. There are occasions when we should not force the searcher to make this prediction, and thus it is that we now recognize that search requirements should be formulated interactively. I would go a little further, and say that we should not, and need not, *oblige* the enquirer to formulate a question at all.

In the laboratory, Salton and his collaborators24 and also Bono5 have demonstrated that the user's relevance judgements concerning the documents retrieved in response to his query can be used to generate, automatically, a substantially more effective query. There are also instances of operational information services where emphasis is placed upon the value of relevance feedback during the formulation of queries and SDI profiles (see Barker *et al.*2 and Vernimb and Steven28 for examples). On-line information retrieval systems, designed for use by the enquirer himself, have been plentiful for some time now, particularly in North America. Lancaster and Fayen15 have treated this topic at length and describe several systems. A good example of a list of desirable capabilities of on-line systems is given by Williams.31 Much of the literature on interactive retrieval systems, whether specific to particular implementations, or containing more general discussions of the principles, is primarily concerned with man-machine interface engineering :4,11,13,15,18,19,26,29,31 for example, how to reduce the effort and training required of users and increase the computer's tolerance to their errors, and how to make users aware of the system's vocabulary or information structures. Software to achieve such ends acts as an interface between a conventional search program (Boolean usually, but sometimes employing matching functions)30 and the user of a terminal. This paper is not about interfaces in this sense. A user's browsing, or exploration of the system is normally controlled by the interface, whereas in the approach I shall describe it is an integral part of the reference retrieval process.

Lesk and Salton16 usefully summarize the structure of interactive systems by distinguishing two types of interaction in query formulation: 'presearch' and 'postsearch'. Presearch interaction consists of investigating the indexing vocabulary of the system in order to decide which terms should be included in the query. Postsearch interaction involves the use of relevance decisions about the result of a previous search, to modify the query. On-line systems generally provide facilities for exploring the vocabulary, but are rarely capable of making use of relevance feedback; it is left to the user to improve his formulation. (This is disturbing in view of the observation of Blanken and Stern4

that 'the ingenuity of the search formulator often plays an important role in determining the relevance and recall ratios obtained for even a relatively simple question.')

Interaction with an on-line system has two effects upon a user which should be distinguished. Firstly, the user learns how to express his need more effectively to the system. Both presearch and postsearch methods are clearly of value in this process, and automatic query modification based on relevance feedback24 may reduce the amount that he needs to learn. Secondly, the user's notion of what he requires becomes clearer to himself, and may shift in emphasis. Documents considered relevant at one stage in the browse/search may be discarded if seen again at a later stage, for instance. Therefore postsearch interaction is used by the enquirer to change the query to correspond to his *dynamic* understanding of his need. Laboratory work on information retrieval has, naturally, taken little account of this phenomenon: the question is assumed to be static; only the formulation is improved. The basis of the relevance feedback techniques explored on the SMART system is an expression for the query (a term-weights vector) which optimizes the retrieval of a given subset, designated 'relevant', of the test document collection.24 An approximation to the optimal query can be made using a previous attempt at a formulation together with partial knowledge of the relevant subset. The theoretical basis for this approach to relevance feedback appears to be at odds with the intellectual processes of the enquirer and their effect upon the composition of the 'relevant subset'.

Consideration of the unsatisfactory aspects of query formulation in information retrieval has resulted, in this work, in a computer program which offers no facilities for query formulation, in the usual sense, and proceeds on the basis of the user's reaction to the references and document descriptions which he is shown. The presearch-postsearch categorization will not be appropriate to this form of interaction. Structures such as the thesaurus are regarded as internal matters for the computer, with which a searcher should not be bothered.

A DIALOGUE STRUCTURE

Frequently, an enquirer can satisfy his need by talking to somebody with knowledge of an appropriate subject.22 Telling him the broad area in which the problem to be solved lies, is relatively easy. Their dialogue (in which, by definition, *each* participates) refines the region of enquiry, until the subject expert understands the other's problems in his own terms. He may then be able to offer information which may lead to a solution. The dialogue is not always a simple qucstion-andanswer interchange. The subject expert may miss the point and give a solution to the wrong problem; then the enquirer must bring him back on course—he must learn through conversation in what terms he should communicate his need. Each has his own world model, and each must try to construct a model of the other's interest, in relation to his own view of the world. Thus, each understands the other. (Quillian23 has advocated a similar notion for language understanding.) A librarian or information scientist will be limited in his ability to foresee possible solutions to many of the problems of his clients. Nevertheless, the structure and objective of their dialogues is much the same (see Taylor25). One component of the librarian's world model is his knowledge of the organization of literature in his library or subject. So, whereas the expert can solve problems in a restricted domain, the librarian can assist over a wide range of topics. The intellectual difficulty encountered in dialogues for information transfer is well expressed by Taylor:25 'one person tries to describe for another person not something he knows, but rather something he does not know.' (p. 180.)

It is upon this view of dialogue that I have modelled the interaction supported by the program to be described. It forms an image of the searcher's interest, derived from its world model, chooses references for display according to the state of the image, and modifies this image continuously in the light of his reactions to the displays. The program's world model consists solely of knowledge about the organization of literature. Specifically, it is a network of associations between documents, authors, and subject terms: any pair, of like or differing type, may be linked. In principle, any other type of entity of interest to a literature searcher may be represented in the network. No information is held on the nature of the associations: it is sufficient for this program to know simply that two entities are associated. Although the form of the structure is straightforward, the network corresponding to a typical indexed document collection will be extremely richly connected.

A program for understanding natural language has recently been described by Fisher,10 which constructs an image of the person with whom it is conversing. The image takes the form of a set of assertions, each tagged with the man's attitude to it (belief, approval, attribution of responsibility, etc.), but is not related to, or built upon a world model. The program is capable of detecting quite subtle inconsistencies in the man's utterances. Winograd's32 well-known program SHRDLU 'keeps track of when things have been mentioned, [so that] it can check a possible interpretation of a question to see whether the asker could answer it for himself from his previous sentences' (p. 185). Again, the

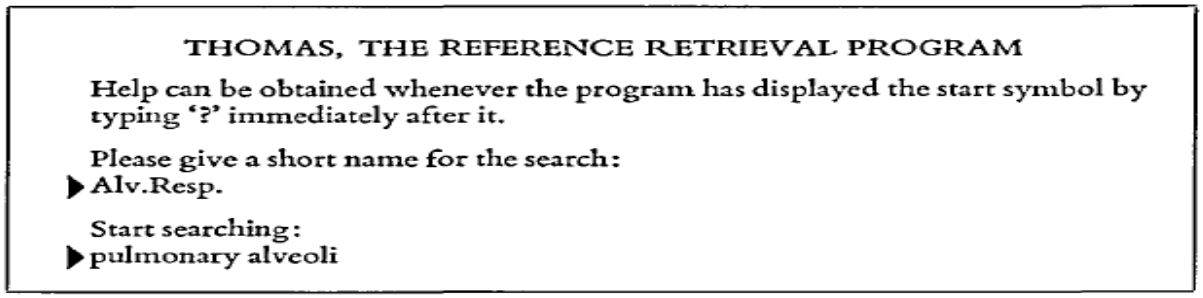
image is a set of assertions; this time, facts about the world as modelled by the program. With our present understanding of question negotiation and the browsing element in reference retrieval, we are not in a position to construct an image of the searcher consisting of (useful) assertions. The main component of the image used in this work is the context of the topic of enquiry, as represented by regions in the program's network knowledge base. Regions may shift, grow, shrink, split, and coalesce in the course of a dialogue, in response to what is learned of the searcher's preferences.

THOMAS, THE PROGRAM

The program, THOMAS by name, was written for the IBM 360/67 at the University of Newcastle upon Tyne, and designed to communicate with a user at an IBM 2260 CRT character display terminal. The bibliographic data was obtained from the MEDUSA project1 being conducted in the Computing Laboratory of the University: it was reorganized into the network structure and accessed by the program from disk storage. The literature covered is in the fields of medicine and biochemistry, and records originated at the US National Library of Medicine as MEDLARS (Medical Literature Analysis and Retrieval System) records. The indexing vocabulary in MEDLARS is strictly controlled, and each subject term in our network either belongs to that vocabulary, or is a synonym added to the MEDUSA system by the Newcastle team.

There follows a sample dialogue with THOMAS conducted by a medical research worker—an anaesthetist. This searcher was of the opinion that very few articles had been written on his precise topic. However, we had ensured that the test file contained references in his broader field of interest. Lines typed by the user are preceded by the symbol ►. This 'start' symbol, as it is called, is used on the terminal to tell the user that he is required to type his next input, but it does not remain on the screen. A slight departure from the genuine computer displays is made in the interests of legibility in this printed form: I am using the lower case alphabet here, whereas the IBM 2260 terminals are without those characters.

The description which follows is necessarily brief and simplified. More detail is given in Oddy [21] and a full account of the program design can be found in Oddy [20].



The user has named the search, so that printed output will be identifiable. He has then typed the term for a subject related to his need. Even in a dialogue which is free of query formulation, the enquirer must take the initiative at least once and indicate his area of interest. It will be seen that throughout this particular search, he has introduced only one term. The first objective of the program is to find a point in the network knowledge base whose label bears a *textual* resemblance to the user's term. I shall not describe the methods in detail—they aim to cope with variants and errors, but do not invoke any syntactic or semantic analysis. At this early stage in the dialogue, the program's image of the user's interest is quite simple and is centred on the subject keyword 'pulmonary alveoli'. It includes documents, subject terms, and any other entity associated (in the network) with 'pulmonary alveoli'. Known associations between those items are also included. As in most circumstances in these dialogues, Thomas prefers to display a reference immediately, because references are presumed to be what most interest the searcher, within the ambit of the program. The choice of which reference to display is made on the basis of the notion of involvement in the image. The involvement of an item is defined to be the ratio of the number of associated items included in the image to the total number of associated items known. In this case, the most involved reference is chosen for display:

Influence of fasting on blood gas tension, pH, and related values in dogs.; Pickrell *et al, Am J Vet Res, 34*, 805–8, Jun 73
1. J A Pickrell, 2. J L Mauderly, 3. B A Muggenburg, 4. U C Luft, 5. animal experiments, 6. animal feed, 7. arteries, 8. blood, 9. body temperature, 10. carbon dioxide, 11. dogs, 12. fasting, 13. hemoglobin, 14. hydrogen-ion concentration, 15. irrigation, 16. lung, 17. oxygen, 18. pulmonary alveoli, 19. respiration, 20. time factors
▶?

The display contains the reference itself and a numbered list of all associated authors and subject terms (whether in the image or not). The searcher's request for assistance is answered by a display suited to this particular part of the dialogue:

There can be three parts to your statement (all optional):
1. Your reaction to the reference just shown (if any).
   This must come first:
   "Yes" or "No"
2. A selection from the names (authors) or terms shown, by number. A "not" in the statement signifies rejection of all numbers that follow it.
3. New names or terms (terms preferably in quotes). The elements of the statement should be separated by commas.

Examples: 'posture', 'circulatory system'
Yes, not 11,12
No, 7,13,4
'heart rate'
Yes

Press enter key when you are ready to proceed ▶

When he has finished with the instructions, the user returns to the dialogue:

. . . 12. fasting, 13. hemoglobin, 14. hydrogen-ion concentration, 15. irrigation, 16. lung, 17. oxygen, 18. pulmonary alveoli, 19. respiration, 20. time factors
?
▶No, 10,17,19,20

The reference is not what he wants, although it has interesting aspects—carbon dioxide, oxygen, respiration, and time factors. With the exception of these terms, and the previously requested term 'pulmonary alveoli', all numbered items in the display are assumed rejected by the searcher, along with the reference. The image changes to reflect this response: selected items and references linked to them are added to the image, and rejected items are removed. New associations consequently join the image, and others drop out. The program will remember which items have been rejected, and inhibit future growth of the image in those directions, thus forcing it into others. A reference is chosen from the updated image:

The arterial-alveolar nitrous oxide difference: a measure of ventilation-perfusion mismatch during anesthesia.; Markello *et al*, *Anesth Analg* (*Cleve*), *52*, 347–50, May–June 73
1. R Markello, 2. R Schuder, 3. D Schrader, 4. E White, 5. aged, 6. anesthesia, inhalation, 7. blood, 8. carbon dioxide, 9. cardiac output, 10. human, 11. hydrogen-ion concentration, 12. methods, 13. middle age, 14. nitrous oxide, 15. oxygen, 16. partial pressure, 17. pulmonary alveoli, 18. respiration, 19. surgery, operative, 20. time factors, 21. ventilation-perfusion ratio
▶Yes

The enquirer was pleased to see this reference. It appeared to be an authoritative account of a different school of thinking on his own problem. When a searcher gives an unqualified 'yes', it is assumed that he has selected all items on display. The effect of his response is thus to enrich the image in the region of the displayed document. Documents close to the relevant one will then tend to have high measures of involvement in the image.

A Positive End-Expiratory Pressure—Nasal-Assist Device (PEEP–NAD) for treatment of respiratory distress syndrome.; Tummons, *Anesthesiology*, *38*, 592–5, June 73
1. J L Tummons, 2. blood, 3. carbon dioxide, 4. human, 5. hydrogen-ion concentration. 6. infant, newborn, 7. masks, 8. methods, 9. nose, 10. oxygen, 11. oxygen inhalation therapy, 12. positive-pressure respiration, 13. respiration, 14. respiratory distress syndrome
▶Yes, 13, not 6

The searcher judges this reference relevant; he is interested in 'respiration', but not in 'infant, newborn'. No assumptions are made by the program about the other numbered items.

Effects of continuous negative pressure on lung mechanics in idiopathic respiratory distress syndrome.; Beaucalari *et al*, *Pediatrics*, *57*, 485–93, Mar 73
1. E Beaucalari, 2. O L Garcia, 3. M. J. Jesse, 4. bicarbonates, 5. carbon dioxide, 6. esophagus, 7. female, 8. gestational age, 9. human, 10. hydrogen-ion concentration, 11. infant, newborn, 12. lung compliance, 13. male, 14. methods, 15. oxygen, 16. partial pressure, 17. pulmonary alveoli, 18. respiration, 19. respiration, artificial, 20. respiratory distress syndrome, 21. ventilation-perfusion ratio, 22. vital capacity
▶No, 19

This reference is judged not relevant, but the user makes it known that artificial respiration interests him. Apart from those he has previously chosen, the other aspects are rejected from the image. The term 'infant, newborn' (no. 11) is present in the display, implying that the reference has been chosen for display in spite of the fact that one of its descriptors has previously been explicitly rejected by the user. The assumption made about the user's intention when he says 'not X' is that he does not want references to be selected on the basis of association with X, rather than that he does not wish to see anything to do with X. A rejected item is inhibited from joining the image, and so cannot assist in the choice of a document. In fact, by being outside the image, it decreases the involvement measure of references associated with it, and thus weighs against those references.

Decrease in pulmonary capacity during lipid infusion in healthy men.; Sundstrom *et al, J Appl Physiol, 34,* 816–20, Jun 73
1. G Sundstrom, 2. C W Zauner, 3. M Arborelius, 4. adult, 5. carbon monoxide, 6. dietary fats, 7. human, 8. hyperlipemia, 9. male, 10. middle age, 11. oils, 12. oxygen, 13. parental feeding, 14. pulmonary alveoli, 15. pulmonary diffusing capacity, 16. respiration, 17. soy beans, 18. triglycerides, 19. ventilation-perfusion ratio
▶ [no reaction—user enters a null line]

The searcher prefers not to commit himself to a judgement on this reference. It is interesting, though not really pertinent to his present requirement. The image is not affected much by this response: the user is saying, in effect, 'no comment, give me another'. If he wishes, a user may mention topics or names without making known his decision regarding the document itself. It is preferable that he should be able to do this than that he should be forced to make a misleading relevance judgement.

Cardiovascular function after pulmonary surgery.; Wronne, *Int Anesthesiol Clin, 10,* 27–39, Winter 72
1. B Wronne, 2. adult, 3. aged, 4. arrhythmia, 5 blood pressure, 6. blood volume, 7. bronchial neoplasms, 8. cardiac output, 9. cardiovascular system, 10. human, 11. lung, 12. middle age, 13. postoperative complications
▶ No

Changes of venous admixture with inspired oxygen in hyaline membrane disease and foetal aspiration pneumonia.; Corbet *et al, Aust Paediatr J, 9,* 25–30 Feb 73
1. A J Corbet, 2. E D Burnard, 3. anoxemia, 4. fetal diseases, 5. human, 6. hyaline membrane disease, 7. infant, newborn, 8. oxygen, 9. pneumonia, aspiration, 10. pregnancy, 11. pulmonary alveoli, 12. pulmonary circulation, 13. respiration, 14. ventilation-perfusion ratio
▶ No

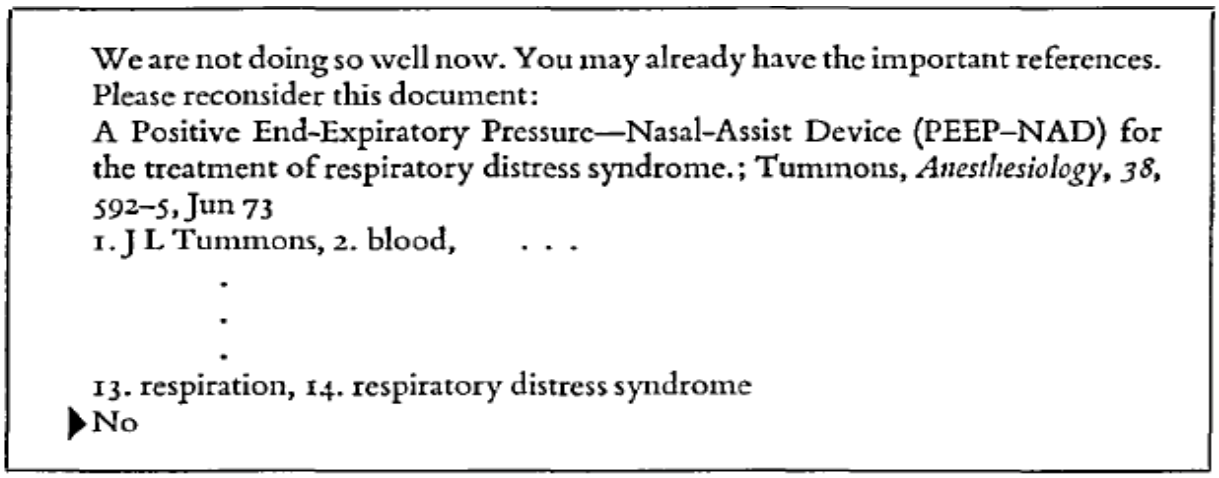
The anti-atelectasis factor of the lung. I; Lachmann *et al, Z Erkr Atmungsorgane, 137,* 267–87, Feb 73
1. B Lachmann, 2. K Winsel, 3. H Reutgen, 4. animal experiments, 5. carbon dioxide, 6. extra-corporeal circulation, 7. human, 8. lung, 9. lung compliance, 10. mice, 11. microscopy, electron, scanning, 12. models, theoretical, 13. pulmonary alveoli, 14. pulmonary embolism, 15. pulmonary surfactant, 16. rats, 17. respiration, 18. respiration, artificial, 19. review, 20, surface tension, 21. vagotomy, 22. ventilation-perfusion ratio, 23. work of breathing
▶ Yes, not 11

The dialogue continued until a further fifteen references had been displayed, as the user was obviously enjoying it, but no more relevant ones were found. We need not follow the search in detail through to the point at which the user felt that he had seen all that the program had to offer. Before leaving the example, let us jump forward a few steps in the dialogue. The situation is that the user has rejected several references in a row and the program, which measures

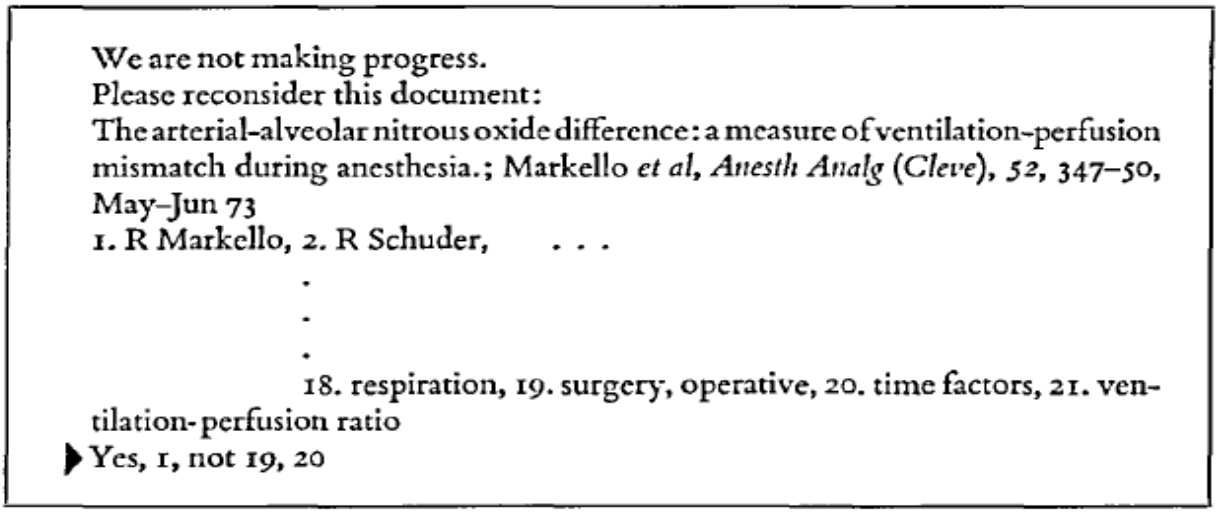
its own performance in the task of extracting favourable reactions from the user, now makes an attempt to get back on course. The strategy is as follows:

(i) Look for a reference which the user has already seen and not rejected, and display it again, asking him to reconsider it. Preference is given to references which no longer have high involvement in the image.
(ii) If the search for a suitable reference fails, show the user a topic that he has explicitly requested together with associated subject terms.
(iii) If no such topic can be found, ask the user to take the initiative, i.e. to think of a new term or name.

In this case, the program shows him again a reference that he has previously judged relevant:



Now, this judgement is a complete reversal of the earlier one (the searcher's privilege!), and does not improve the performance measure. The next display is:



This was still the most important reference seen. The user had noted that the term 'time factors' was attached to several of the references, and had a wide variety of meanings, so he now stated that he was no longer interested in it. (Here is explicit evidence that the searcher is also building a model, or image: to represent the program's view of the world.) The response enabled the program to display a few more new references on topics in anaesthetics.

There are many aspects of the program which are not illustrated in the dialogue above (see references 20 and 21). Nevertheless, it should give the reader an impression of the simplicity of dialogues with program THOMAS. A searcher may take as much, or as little, initiative as he wishes while using the program. He may mention new subjects, authors or titles as they come to mind, or he may confine himself to reacting to what he is shown. There are no commands for him to learn to use, such as one usually finds in 'conversational' systems.

The syntax of the user's statements is very simple. No attempt has been made to parse natural English sentences. However, the statement structure and the interpretation of responses is meant to correspond with the searcher's natural reactions to references which crop up in a dialogue. He may think *'Yes,* that reference looks promising', or *'No,* it is not likely to be useful', or 'I don't know about that one', and he may qualify his judgement by pointing out the descriptive attributes of the document which appeal, or do not appeal, to him. The assumptions that the program makes regarding the attributes that he does *not* mention depend upon his general assessment of the reference, whether he has mentioned any of the descriptive items, and what has been learned from his past responses. This is illustrated in the sample dialogue above.

It can be seen that the program is not suitable in itself for large-scale exhaustive literature searches. Even for such requirements, however, it may be a useful tool for getting a search under way. Finding a few references will help the searcher to decide what he is looking for, and should also provide a lead-in to the literature through chains of citations. The program is intended to be illustrative of, and a test for, a new approach to man-machine interaction in bibliographic search. There is plenty of potential for refinement of the method and it will be necessary to consider the ways in which a browsing facility of this type might be integrated with the well-established information retrieval tools. The user who begins his search, glad that he need not say anything definite, suddenly finds that he knows what he needs, and wants to pose precise questions.

EXPERIMENTAL RESULT

The evaluation of an on-line information retrieval system is difficult. The high level of interaction between man and computer and the use of machinery to aid browsing tempts one to regard the user as part of the system, and take his aims and performance into account in evaluating the system. In the work reported here it has not been possible to observe a significant number of genuine searches, conducted by real users. The aim of the initial tests was to determine whether the retrieval effectiveness of a technique which avoided query formulation could be comparable to that of a more conventional on-line system. Thus, the test data was chosen in such a way that a direct comparison with non-delegated searches with the MEDUSA system was possible. The MEDUSA system is an on-line reference retrieval system which gives access to MEDLARS data and provides a very effective set of facilities for incremental query formulation. It has been used for experiments to compare the performance of scientific enquirers conducting their own searches on-line with that of professional information workers.1

The users were carefully simulated in thirty-two searches corresponding to actual MEDUSA searches. Relevance judgements obtained from the users and terms occurring in the MEDUSA sessions were used to determine reactions to the displays generated by the program THOMAS. The reactions of the simulated users were always simple and exhibited rather less initiative than one might expect of real users. The quantitative experiments have been restricted to making sure that the program finds the relevant references in its data base, quickly and without a great deal of effort expended by the enquirer. Search length has been proposed by several people7,8,26 as a basis for evaluation, particularly of on-line systems, and effortless interaction on the user's part is often considered desirable.15 Using a small test collection, dialogues conducted with THOMAS were continued until all the relevant references retrieved by the searches formulated with MEDUSA's aid had been displayed. THOMAS showed the user about as many references as MEDUSA (the small difference was not statistically significant), but the demand on user effort by THOMAS was much less than that demanded by MEDUSA— about a quarter, in fact, in terms of the number of subject phrases and other symbols typed. The actual average precision (which, in such experiments, could be regarded as an inverse measure of search length) was 58% in these dialogues. THOMAS achieved this performance without giving the user the opportunity to formulate a query.

CONCLUDING REMARKS

The art of information system design (which, I am certain, has a long future) is to find the form and timing of information presentation which will best aid the system user in whatever task he has in hand. The problem is his and the 'search' is a subordinate problem—for him to solve, because automated information systems cannot yet solve problems for their users. The job of the mechanical part of the system is to create a helpful framework within which

the user can make problem-solving decisions. The decisions a literature searcher wishes to make concern documents, and the program which I have described briefly in the foregoing paragraphs shows him references to documents, one at a time, and invites him to assess their relevance.

From the user's point of view, a dialogue with THOMAS is a browse through a collection of document surrogates, in which he may take as much initiative as he wishes. The subject terms occurring in the collection are seen through their use in describing individual documents. The searcher has the opportunity to pass judgement upon each reference shown to him, and the program uses his comment to try to form a better image of his region of interest before showing him more. The man may react to the reference itself and to individual items in its description. He may also name topics of interest to him; however, after the first few interactions with THOMAS, he will normally find it much easier to acknowledge what interests him and dismiss or ignore what does not. In this way the program design induces the user to express his need in the language of the information available to the program.

I do not wish to claim that the computer program THOMAS could provide a comprehensive information facility for the research worker. It is actually for this reason that I have tried to avoid terms like 'system' when referring to THOMAS, and used 'program' instead. In many circumstances, users would currently find other information tools more appropriate and satisfying. I am reluctant, however, to label THOMAS simply as another information retrieval tool which must find its place in the 'total information system', because this suggests that a straightforward separation of the various functions of an information facility is possible. I do not believe that browsing is separable from literature searches involving the use of, say, citation indexes and abstracting journals. On-line access can enhance the utility of any computer-stored information bank (bibliographic or otherwise) whenever there is uncertainty that the searcher's need can be accurately conveyed to the machine. The reason for this is usually that the degree of interleaving of intellectual and mechanical effort is high when a searcher uses an interactive program, whatever its design.9 However, if one is not quite sure what one wants, issuing orders is inappropriate and wasteful of the resources of a slave. The more balanced form of man-machine *dialogue* in THOMAS is meant to resemble the personal communication between human minds through conversation which is so effective in the transference of complex ideas.

ACKNOWLEDGMENTS

This work was supported originally by the Office for Scientific and Technical Information and then by the British Library Research and Development Department, to whom I am extremely grateful. Elizabeth Barraclough of the University of Newcastle upon Tyne was continually ready with ideas and encouragement, and I am therefore very much in debt to her.

REFERENCES

1. BARBER, A. S., BARRACLOUGH, E. D. *and* GRAY, W. A. On-line Information retrieval as a scientist's tool. *Information Storage and Retrieval,* 9 (8), 1973, 429-40.

2. BARKER, F. H., VEAL. D. C. *and* WYATT, B. K. Towards automatic profile construction. *Journal of Documentation, 28*(1), 1972, 44-55.

3. BELKIN, N . J . Towards a definition of information for informatics. *Informatics 2, Proceedings of a conference held by the Aslib Co-ordinate Indexing Group on 25-27 March 1974 at New College Oxford* (*edited by Verina Horsnell*), Aslib, 1975, pp. 50-6.

4. BLANKEN, R. R. *and* STERN, B. T. Planning and design of on-line systems for the ultimate user of biomedical information. *Information Processing & Management,* 11 (8/12), 1975, 207-27.

5. BONO, P. R. *Adaptive procedures for automatic document retrieval.* Ph.D. Thesis, University of Michigan, 1972.

6. BROOKES, B. C. Robert Fairthorne and the scope of information science. *Journal of Documentation, 30* (2), 1974, 139-52.

7. CLEVERDON, C. W. User evaluation of information retrieval systems. *Journal of Documentation, 30* (2), 1974, 170-80.

8. COOPER, W. S. Expected search length: a single measure of retrieval effectiveness based on the weak ordering action of retrieval systems. *American Documentation,* 19(1), 1968, 30-41.

9. DOYLE, L. B. Expanding the editing function in language data processing. *Communications of the A.C.M., 8,* 1965, 238-43.

10. FISHER, J. N. D. *Computational semantics: a study of a class of verbs.* Ph.D. Thesis, University of Aston in Birmingham, 1974.

11. HALL, J. L., NEGUS, A. E. *and* DANCY, D. J. On-line information retrieval: a method of query formulation using a video terminal. *Program,* 6(3), 1972, 175-86.

12. HERNER, S. Browsing. *Encyclopedia of library and information science,* Marcel Dekker, 1970.

13. KATTER, R. V. On the on-line user of remote-access citation retrieval services. SDC Report TM-(L)-4494/000/00, 1970.

14. LANCASTER, F. W. Information retrieval systems: characteristics, testing, and evaluation. Wiley, 1968.

15. LANCASTER, F. w. *and* FAYEN, E. G. Information retrieval on-line. Wiley, 1973.

16. LESK, M. E. *and* SALTON, G. Interactive search and retrieval methods using automatic information displays (Chapter 25). *The SMART retrieval system* (*edited by G. Salton*), Prentice-Hall, 1971.

17. LORD TODD. Introduction: the problem stated. *Communication in science: documentation and automation* (*edited by A. de Reuck and J. Knight*), London, Churchill, 1967, pp. 4-15.

18. MARTIN, T. H., CARLISLE, J. *and* TREU, A. The user interface for interactive bibliographic searching: an analysis of the attitudes of nineteen information scientists. *Journal of the A.S.I.S., 24, 1973,*142-7.

19. MEADOW, C. T. Man-machine communication. Wiley, 1970.

20. ODDY, R. N. *Reference retrieval based on user induced dynamic clustering.* Ph.D. Thesis, University of Newcastle upon Tyne, 1975.

21. ODDY, R. N. Retrieving references by dialogue rather than by query formulation. (To be published in the proceedings of Informatics 4 conference held at Lancaster, April 5-7, 1976.)

22. OLNEY, J. C. Building a concept network for retrieving information from large libraries. Part 1. SDC Report TM-634/001/00, Jan. 1962.

23. QUILLIAN, M. R. The teachable language comprehender: a simulation program and theory of language. *Communications of the A.C.M., 12* (8), 1969, 459-76.

24. ROCCHIO, J. J., Jr. Relevance feed back in information retrieval (Chapter 14), *and* SALTON, G. Relevance feedback and the optimization of retrieval effectiveness (Chapter 15). *In: The SMART retrieval system* (*edited by G. Salton*), Prentice-Hall, 1971.

25. TAYLOR, R. S. Question-negotiation and information seeking in libraries. *College & Research Libraries, 29,* 1968, 178-94.

26. THOMPSON, D. A. Interface design for an interactive information retrieval system: a literature survey and a research system description. *Journal of the A.S.I.S.,* 22, 1971, 361-73.

27. VAN RIJSBERGEN, C. J. *and* SPARCK JONES, K. A test for the separation of relevant and nonrelevant documents in experimental retrieval collections. *Journal of Documentation, 29* (3), 1973. 251-7.

28. VERNIMB, C. O. *and* STEVEN, G. 'ENDS'—European Nuclear Documentation Service. *Nuclear Engineering and Design,* 25, 1973, 325-33.

29. WALKER, D. E. (Ed.). Interactive bibliographic search: the user/computer interface. AFIPS Press, 1971.

30. WILLIAMS, J . H., Jr. BROWSER: an automatic indexing on-line text retrieval system. Annual report. IBM Federal Systems Division, 1969. AD 693 143.
31. WILLIAMS, J. H., Jr. Functions of a man-machine interactive information retrieval system. *Journal of the A.S.I.S., 22,* 1971, 311-17.
32. WINOGRAD, T. A procedural model of language understanding (Chapter 4). *In: Computer models of thought and language* (*edited by R. C. Schank and K. M. Colby*), San Francisco, W. H. Freeman, 1973.

(Received final version 6 October 1976)