AN EXPERIMENT TO IDENTIFY DECISION POINTS IN ONLINE SEARCHING

Charles T. Meadow
Maria Loranger
Birgitta Olander

Faculty of Library and Information Science
University of Toronto
Toronto, Ontario M5S 1A1

ABSTRACT

This paper describes an experiment whose purpose was to identify and characterize major decision points in an online search. These are points at which a searcher might pause to assess progress and to decide upon the major course of action if improvement is necessary. The decision points will be used by a searcher assistance program to decide when to interrupt a search with advice to the user about how to proceed, and to provide him or her with a small number of broad options for how to proceed. Such an interruption should come at a time when the searcher, himself, senses the need for assessment. Hence, there is the need to determine at what points users are receptive to such advice.

DECISION POINTS IN ONLINE SEARCHING

PURPOSE OF THE STUDY

This study is part of a larger project whose objective is the development of a user training and assistance system for RECON, the U.S. Department of Energy's online information retrieval system. The user aid, called Online Access to Knowledge [Borgman et al, 1985] will be a "front end" program, resident in a personal computer and intended to help end users make direct use of RECON. A number of commercial systems for this purpose have recently been developed [Levy, 1984]. OAK is intended to display a higher degree of "intelligence" than front ends produced to date. By intelligence we mean understanding of users' problems and ability to offer assistance in the context of an actual user problem, rather than on searching in general.

Assistance can be meaningfully given at a number of levels, such as in the composition of a command or selection of a database. Our concern here is for help to the user in deciding what kind of action to take when he has completed some retrieval commands and needs to evaluate what has been done in order to plan his next commands, in other words, to assist in a decision having broad implications. There is also a need to know _when_ it is appropriate to offer help to the user. Help will be in the form of analysis of results and suggestion of options for the next action.

Our purposes in this experiment were: (a) to determine when users _say_ they have come to a decision point, (b) to determine _what decisions_ they come to, given a history of the search, and (c) to investigate whether a program can anticipate the decision points users identify and the decisions they make.

THE CONCEPT OF DECISION POINTS

Definition of Decision Points

The essential idea of a decision point is that it is a point at which a searcher makes a decision whose scope affects an ensuing sequence of commands. It is, in Bates' [1979a] terms, a tactic, but a high-order one. The decision to limit a set by date, or to terms in a title only, has limited scope of impact. The decision to introduce a new major search concept, or to drop one, or to start the search all over again with different terms, affects a number of commands that will follow the decision. We allow for one temporizing decision, that to seek more information before making another decision.

We believe, with varying degrees of intensity among us, that online searching might well be taught in terms of: (1) how to create a search plan (Bates' search strategy), (2) how to proceed from one major decision point to the next, and (3) how to carry out a decision with specific commands. The need for decisions and the making of changes as a result of them, does not negate the original plan, which should have been drawn up in recognition that some changing is necessary. Indeed, the plan is for how to anticipate decisions, not for how to execute the search perfectly on the first try.

In conducting this experiment, we did not try to change our subjects' views to correspond with the approach outlined above. We introduced the idea of decision points and asked them to identify the points as they occurred. We did not try to provide a list of possible decisions. They decided, and we interpreted what they said and did in our own terms.

Decision Alternatives

The decision to be made is the selection of one of the major actions listed below. Each will entail further decisions and in the assistance program there will be allowance for crossing over from one action path to another. An essential point in our approach is that none of these programs will be necessarily designed to complete the search, all at once. Our objective is only to achieve improvement and to count on iteration of the entire evaluation-enhancement process to achieve the overall goal of satisfying the user's information need. Hence, crossing from one path to another, or backing up, are acceptable procedures and this means we are not forced into selecting among mutually exclusive decision paths.

The use of a decision assistant does not preclude the user from performing his own evluatation and enhancements. Any of ours may be overridden at any time. Also, the user can invoke the evaluation process at any time, without the program detecting a decision point.

The decisions or next actions we set out with are listed below. Another, Combination of others, was added while analyzing the data:

BROWSE -- the user must browse some (more) of the records of the set, and evaluate each one seen, to provide more information. This decision is a temporizing one, a recognition that more information is needed in order to decide how to revise a search.

REDUCE SET -- the user is basically satisfied, but wishes to reduce the size of a set.

ENLARGE SET -- the user is basically satisfied but wishes to enlarge a set.

REFINE SET -- the user is basically satisfied, but wishes to consider some minor variations, without having a goal in terms of set size.

REDEFINE SET -- the user recognizes that results so far are not satisfactory and wants to make a basic change in his set definition.

ACCEPT -- the user decides the current set is an acceptable answer set, and ends the search

QUIT -- the user gives up on the search.

OTHER -- a user decision not in any of the categories listed above. A program cannot choose "other," but it can decline to make a choice, when asked.

Related Concepts

Conventional training exhorts searchers to work with structures, i.e. they build sets out of other sets [Meadow-Atherton, 1980]. They may do this, for example, by defining a series of concepts, each in turn consisting of one or more search terms, and then combining the concepts, or they may define a broad set, then gradually restrict its definition, decreasing its size as they do.

DECISION POINTS IN ONLINE SEARCHING

Bates and Fidel have discussed search strategy or tactics in terms of the actions to be taken at various times. Our decisions, while not expressed in the same ways, are fairly compatible with these two approaches. Bates [1979a, 1979b] proposes a set of *search tactics* and *idea tactics*. Because we want a program to be able to suggest the alternative to be taken, we wanted our list of options restricted to fewer choices than Bates' list. Fidel [1984a] uses the concept of *moves* approximately as we do decisions, the difference in terminology reflecting our current emphasis on deciding what to do, rather than taking the action. In a later paper, Fidel [1984b] describes some moves much like our decisions.

Penniman [1975] classified commands by type and noted that searches tend to progress through a series of command types, sometimes with one or more repetitions of these sequences. Typically, the sequence is:

BEGIN (select a file)
EXPAND (search a dictionary)
SELECT (retrieve records of the main file)
DISPLAY (browse through some records)
PRINT (print records off line)

Frequently, of course, users will find a need to move back to an earlier point in this sequence. We call each progression in the direction BEGIN ... PRINT a *cycle*, even if not all of the command types are executed. Hence, the sequence

BEGIN, SELECT, DISPLAY, SELECT, DISPLAY

contains two cycles, B, S, D and S, D.

The end of a sequence so defined seems, intuitively, a good point at which to pause for evaluation. Hence, we looked for coincidence between user-defined decision points and the ends of what we might call *Penniman cycles*. However, one weakness of this definition is that we can only detect the end of a cycle by either a logoff or the start of the next cycle. If a user is going to pause for reflection, it should be before starting the new cycle, not after.

The completion of a structure, a set defined by logical combination of other sets or terms, is an appropriate point to pause for evaluation, and likely corresponds to the end of a Penniman cycle. In our own definition, we tried to identify such points as following a set-defining command that showed a sharp increase in logical complexity over its predecessors, a simple example of which is:

SELECT TERMA, SELECT TERMB, SELECT TERMC,
SELECT S1 AND S2 AND S3.

By its increased complexity, the fourth SELECT may indicate a difference of purpose, not defining a basic component, but a structure.

In practice, users' decision points corresponded well with ends of Penniman cycles, but where they did not, they often differed by at most one DISPLAY command (See Table 2 below). This means the evaluation came after viewing one set of records then viewing some others to provide ammunition for the next approach.

Another factor in our thinking about the location of decision points has to do with intrusion. Our intent is to offer the user help in evaluating his search, or allow him to request it. We do not want to interrupt to offer help at inappropriate times. An inappropriate time is one at which the user would feel intruded upon by the offer. Particularly if we are communicating with users in line mode (each new line of text comes in at the bottom of the screen and pushes the display up), an interrupt message is intrusive unless wanted. It takes time and it breaks up the visual field on which the user is concentrating.

An interrupt to ask if help is wanted, when it was not, can only disrupt the user. Operating in screen mode, with windowing software available, would enable us to "pop" a short question into view in a corner of the screen, where it may more easily be ignored. We could, of course, leave the question permanently in view, obviating the need to determine decision points altogether. But, we feel it will help to give the user the sense that our OAK program is, indeed, intelligent to some extent. If it is, then its advice is more likely to be heeded. If not physically obtrusive, then even if it appears only approximately when it is wanted, its very appearance conveys a message to the user that a decision point is approaching.

METHOD

Subjects

26 subjects participated in the experiment. They were all University of Toronto students and with one exception enrolled in the course, Online Information Retrieval. This has a prerequisite of a course in reference work that includes about 1.5 hours of online search demonstration. With the exception of two, the subjects had no experience of online searching prior to taking the course.

Of the total of 26, 23 searches were considered valid, while 3 were discarded as invalid due to technical problems.

System and Database Used

The retrieval system used for the experiment was DOE/RECON, which is operated by the U.S. Office of Scientific and Technical Information (OSTI). DOE/RECON provides access to energy-related databases produced or obtained by OSTI. The data base searched in the experiment was the Energy Data Base. The system was accessed via TYMNET, through print terminals. The experiment went on for two consecutive days, and particularly the first day a number of system-related technical problems occurred. Response times were very slow most of the first day, but improved the second day. All truncation searches were very slow. The technical problems brought about search periods much longer than planned (20 minutes per search), and search time thus cannot be meaningfully related to the number of decision points, etc. The printed recordings of the searches were saved and subsequently analyzed together with other information collected by the observers during the searches.

Tasks

Before searching, the subjects were obliged to attend one of two instruction sessions offered. At the instruction sessions an oral presentation was made of the experiment and the tasks the subjects were asked to carry out, i.e. identify their decision points during the search. The concept of decision points was described in a way as uncomplicated as possible. Also during the oral presentation differences in the command languages between RECON and DIALOG were pointed out, since all the subjects were familiar with DIALOG. Written material supporting the oral presentation was handed out, including the search question.

All subjects were given the same search question, which read as follows:

> "The former premier of Ontario, Mr. William Davis, has been appointed to a special mission, with rank of ambassador, to deal with the problem of acid rain and the relations between Canada and the U.S. over this issue. Mr. Davis has asked for some background reading, starting with what has been done so far, i.e. what agreements exist already between the U.S. and Canada over acid rain."

During the searches an observer was always present. The subjects had been instructed to notify the observer when they considered themselves at a decision point. At such a notification, if the information was not volunteered, the subjects were asked by the observers why they identified a decision point at that particular point of the search, and what their decision was. This information was recorded and used later in the analysis of the searches.

STATISTICAL RESULTS

Illegal or erroneous commands were eliminated from the recordings of the searches. Since all commands were in the same database, logging on and the BEGIN command were done by the observer. BEGIN is not included in the statistical results.

Table 1 contains a summary of the commands given, by frequency. Since searchers were inexperienced and BEGIN was omitted, almost all commands were SELECT or COMBINE or DISPLAY. We have combined SELECT and COMBINE statistics, since there is virtually no logical difference between the two in RECON.

TABLE 1 Command frequencies for all subjects

Cmd type	No. of cmds	% of total
Select	128	49
Combine	56	21
Subtotal	184	70
Display	70	27
Other	8	3
Totals	262	100

Table 2 shows the degree of coincidence between u designated decision points and the ends of Penniman cy The first row shows number of coincident points. The second shows number of user decision points that diffe from a Penniman point by only one DISPLAY command. Row shows the number that differed by a single command othe than DISPLAY, and the remaining lines show number of decision points and the number of commands they were o from the end of Penniman cycles.

TABLE 2 Subject identified decision poin

Position of dp as compared to Penniman's	No of dp's	% of all dp's	Cum. No. / %
Coincident	47	46	47 / 46
1 cmd apart (DISPLAY)	26	26	73 / 72
1 cmd apart (other)	10	10	88 / 81
2 cmds apart	6	6	89 / 87
3 cmds apart	4	4	93 / 91
> 3 cmds apart	9	9	102 / 100
Totals	102	100	

Table 3 shows the decisions made by the subjects frequency and percent of total of each. This table includes the decision COMBINATION, not in our original uncomplicated as possible. Also during the oral presentation differences in the command languages between RECON and DIALOG were pointed out, since all the subjects were familiar with DIALOG. Written material supporting the oral presentation was handed out, including the search question.

All subjects were given the same search question, which read as follows:

> "The former premier of Ontario, Mr. William Davis, has been appointed to a special mission, with rank of ambassador, to deal with the problem of acid rain and the relations between Canada and the U.S. over this issue. Mr. Davis has asked for some background reading, starting with what has been done so far, i.e. what agreements exist already between the U.S. and Canada over acid rain."

During the searches an observer was always present. The subjects had been instructed to notify the observer when they considered themselves at a decision point. At such a notification, if the information was not volunteered, the subjects were asked by the observers why they identified a decision point at that particular point of the search, and what their decision was. This information was recorded and used later in the analysis of the searches.

STATISTICAL RESULTS

Illegal or erroneous commands were eliminated from the recordings of the searches. Since all commands were in the same database, logging on and the BEGIN command were done by the observer. BEGIN is not included in the statistical results.

Table 1 contains a summary of the commands given, by frequency. Since searchers were inexperienced and BEGIN was omitted, almost all commands were SELECT or COMBINE or DISPLAY. We have combined SELECT and COMBINE statistics, since there is virtually no logical difference between the two in RECON.

TABLE 1 Command frequencies for all subjects

Cmd type	No. of cmds	% of total
Select	128	49
Combine	56	21
Subtotal	184	70
Display	70	27
Other	8	3
Totals	262	100

Table 2 shows the degree of coincidence between user-designated decision points and the ends of Penniman cycles. The first row shows number of coincident points. The second shows number of user decision points that differed from a Penniman point by only one DISPLAY command. Row 3 shows the number that differed by a single command other than DISPLAY, and the remaining lines show number of decision points and the number of commands they were offset from the end of Penniman cycles.

TABLE 2 Subject identified decision points

Position of dp as compared to Penniman's	No of dp's	% of all dp's	Cum. No. / %
Coincident	47	46	47 / 46
1 cmd apart (DISPLAY)	26	26	73 / 72
1 cmd apart (other)	10	10	83 / 81
2 cmds apart	6	6	89 / 87
3 cmds apart	4	4	93 / 91
> 3 cmds apart	9	9	102 /100
Totals	102	100	

Table 3 shows the decisions made by the subjects, with frequency and percent of total of each. This table includes the decision COMBINATION, not in our original list, but added during the evaluation phase, to cover those cases in which even stretching the definitions of the other types, we could not adequately illustrate the event in the experiment by using either of the labels. We identified 9 "combination" decisions out of the 102 total.

TABLE 3 Decisions by category for all subjects

Decision type	No. of dec.	% of dec's
Browse	28	27
Accept	22	22
Redefine	21	21
Combination	9	9
Other	8	8
Refine	7	7
Reduce	6	6
Quit	1	1
Enlarge	-	-
Totals	102	100

FINDINGS

We consider this to have been a preliminary experiment, to be repeated later with more subjects and more variation in their backgrounds, and possibly with an active assistance program to compare results with.

We feel that this experiment did show that searchers do make major decisions in patterns which follow closely the cycles defined by Penniman. The major differences were in intervening display commands. Often, our observers felt, a decision was announced and only then were more records examined to find out how to perform the next series of commands. The observers also felt that the test subjects did not differentiate adequately between high order and low order decisions, i.e. that they often announced a decision that was nothing more than to execute some particular command, not to change the course of the search.

We feel, although the experiment was not designed to produce evidence for this, that searching ought to be taught in terms of strategic planning of the major steps to be taken in implementation of the strategy. This is "top down" searching, following a technique found quite successful in computer programming.

Although users hewed to a pattern, they were often unclear about their own decisions. That they were not dealing with clear cut, mutually exclusive decisions may confirm that we do not have to either, and that we can content ourselves with offering help at approximately the time most users would begin considering that a decision point had been reached, and that the decision alternatives we offer need not present mutually exclusive paths for the user to follow.

REFERENCES

Bates, Marcia J. (a) "Information Search Tactics," in Journal of the American Society for Information Science, Vol. 30, No. 4 (July 1979), pp 205-14.

Bates, Marcia J. (b) "Idea Tactics," in Journal of the American Society for Information Science, Vol. 30, No.5 (Sept 1979), pp 280-89.

Borgman, Christine L., Case, Donald O. and Meadow, Charles T. "Designing an Information Retrieval Interface Based on User Characteristics," ACM SIGIR Annual Meeting, Montreal, June 4-6, 1985.

Fidel, Raya. (a) "Online Searching Styles: A Case-Study-Based Model of Searching Behavior," in Journal of the American Society for Information Science, Vol. 35, No. 4 (July 1984), pp 211-21.

Fidel, Raya. (b) "Moves in Online Searching", unpublished manuscript, 1984.

Levy, Louise R. "Gateway Software: Is It for You?" in ONLINE, Vol. 8, No. 6 (November 1984), pp 67-79.

Meadow, Charles T. and Atherton, Pauline. Basics of Online Searching, New York, John Wiley & Sons, Inc., 1980, pp 133-42.

Penniman, W. David. Rhythms of Dialogue in Human-Computer Conversation, Ph D dissertation, Ohio State University, 1975.