NASA Contractor Report 4569

A Comparison of Boolean-based Retrieval to the WAIS System for Retrieval of Aeronautical Information: Final Report

Gary Marchionini and Diane Barlow
University of Maryland at College Park
College Park, Maryland
Under contract to RMS Associates

Prepared for
National Aeronautics and Space Administration
Scientific and Technical Information Program
Code JTT
under Contract NASw-4584

National Aeronautics and
Space Administration

Scientific and Technical
Information Program

1994

A Comparison of Boolean-based Retrieval to the WAIS System for Retrieval of Aeronautical Information: Final Report

Gary Marchionini and Diane Barlow

College of Library and Information Services,
University of Maryland at College Park,
College Park, Maryland

Abstract

An evaluation of an information retrieval system using a Boolean-based retrieval engine and inverted file architecture and WAIS, which uses a vector-based engine, was conducted. Four research questions in aeronautical engineering were used to retrieve sets of citations from the NASA Aerospace Database which was mounted on a WAIS server and available through Dialog File 108 which served as the Boolean-based system (BBS). High recall and high precision searches were done in the BBS and terse and verbose queries were used in the WAIS condition. Precision values for the WAIS searches were consistently above the precision values for high recall BBS searches and consistently below the precision values for high precision BBS searches. Terse WAIS queries gave somewhat better precision performance than verbose WAIS queries. In every case, a small number of relevant documents retrieved by one system were not retrieved by the other, indicating the incomplete nature of the results from either retrieval system. Relevant documents in the WAIS searches were found to be randomly distributed in the retrieved sets rather than distributed by ranks. Advantages and limitations of both types of systems are discussed.

1. Introduction

Information retrieval has become an important issue in business and government due to the development of high-speed networking and finding aids such as Gopher, World-Wide-Web, Mosaic, and WAIS (see Obraczka, Danzig, & Li, 1993 for an overview of different systems). Such systems have provided broad ranges of computer users generally easy access to Internet resources, but access is not sufficient to assure information-seeking success. There is no evidence on how such systems perform either quantitatively or qualitatively and no guidance on what information seeking strategies are appropriate. Although the information retrieval (IR) community agrees that networked resources are important research and development domains, there have been no systematic studies of retrieval performance in networked environments. Practice rather than theory is driving development. This situation is problematic since networked end-user environments have enormous implications for IR researchers and information specialists. To explore how one such system performs, a comparative investigation was conducted for the WAIS system and a Boolean-based retrieval system. This report presents results from this study and makes suggestions for developers, evaluators, and users.

1.1. Boolean-based Systems

The development of systems to store and retrieve bibliographic records has been dominated for the past 30 years by the model that uses a Boolean-based retrieval engine and inverted file architecture to match query terms to document terms (these will be referred to as Boolean-based systems (BBS)). A search for relevant records is conducted in one of two modes: (1) using designated descriptors, typically including assigned index terms, authors names, and dates of publication; (2) using the free-text capability, which searches for individual words and phrases within citations. Among the prominent examples of such systems are Dialog, BRS, Medlars, and NASA RECON.

These systems support the bulk of IR practice because there is a large installed base of databases and a community of professional intermediaries who have learned to maximize their value. Over the years these systems have added new features and tools so that professional intermediaries can have even better control in mapping user needs to system capabilities. Today's BBSs allow skilled searchers to tightly control the scope and format of results (e.g., high recall or high precision; specific fields or full records), and new features such at Dialog's Target (Dialog, 1993) offer new search capabilities based on frequency of word occurrence as well as traditional exact match retrieval. The specialized features make these systems powerful but require that users be highly trained to use their rich command languages.

To attract end users, BBS vendors offer alternative systems that feature menu-driven interfaces, online thesauri, online help and abbreviated command sets (e.g., Knowledge Index, BRS AfterDark, Grateful Med). Thus, BBSs continue to evolve from the rigid, exact-match systems often criticized in the research literature to diverse collections of interfaces that aim to serve professional and novice searchers.

Although these systems continue to evolve, BBSs are often criticized in four regards (e.g., Larson, 1992): 1) users must use Boolean logic to formulate queries; 2) Boolean operators are either too coarse (OR is too inclusive) or too fine (AND is too restrictive); 3) query terms are treated equally regardless of their relative importance to the user's information need; and 4) all retrieved documents are treated as equally relevant (documents are not ranked).

1.2. Frequency of Occurrence-based Systems

An alternative approach to information retrieval uses a vector space model for documents and queries (Salton & McGill, 1983). The basic idea behind the vector model is to represent every document as an N-dimensional vector where N represents the number of words in the index language (this can be a small set of controlled terms or, in the extreme case, every word in the document collection) and each cell contains a value related to frequency of occurrence of that cell's index word in the document. Thus, a document that does contain the word "thermal" but does not contain the word "radiation" would be represented by a vector that contains a 0 in the cell for the index word "radiation" and some non-zero value, dependent on the weighting scheme used, in the cell for the index word "thermal." Queries are likewise represented as N-dimensional vectors, and each document vector is compared to the query vector according to a similarity metric (e.g., the dot product or the cosine of angle between the vectors).

Although computationally intensive, this type of representation allows all documents in the collection to be ranked in similarity to the query, thus mitigating the exact match limitations of Boolean-based approaches. Research related to this approach to information retrieval considers parameters such as types of term weighting schemes, types of similarity measures, size and composition of the indexing language (e.g., highly controlled or all words in the database), treatment of word phrases and proximity limits, and types of stemming.

In addition to the SMART studies conducted by Salton and his colleagues, Harman and her colleagues have demonstrated the usefulness of the vector approach (Harman & Candelea, 1990). Harman, McCoy, Toenese & Candelea (1991) argue that the frequency of a term's occurrence within the document taken in relation to its frequency of occurrence across the entire record file and in relation to the length of the document provides an accurate indication of the importance of the term in the documents and query. In this way, the system numerically determines that the term "AIDS" has great meaning for occurrences within a psychological counseling database, while being relatively unimportant within an AIDS database. Of course, this example also illustrates one of the dangers of both word matching and word counting approaches to full-text databases--homonyms can lead to irrelevant documents being retrieved in the BBS case and to misranked documents in the vector case.

Based upon the term weighting values and the relationships that document terms share with users' query terms, vector-based systems ordinally rank retrieved record sets in order of descending similarity to the query. The first citation presented to the user is the one with the highest similarity score. In theory, the system retrieves the entire set of documents that have any of the query terms (equivalent to a OR of all free-text terms in a BBS) ranked according to their similarity to the query. In practice, cutoff values are used to return only the top-ranked documents.

Another information retrieval technique is known as relevance feedback (Ide, 1971; Rocchio, 1971). This approach displays document citations to users and invites them to inform the system which citations are relevant. The relevant citations are then used to locate similar documents. Each iteration of the search is an opportunity to further refine the search by allowing the system to readjust the query vector by including and/or discarding selected terms. The user can proceed with as many iterations as necessary to retrieve those documents which best meet the perceived information need. Although either the Baleen or vector techniques may be used with relevance feedback, it is most often associated with the vector approach since term weights can be adjusted in the query[1].

The present study evaluated and compared the performance of a traditional system used for aeronautical engineering databases and a system based on vector retrieval and relevance feedback. Since these two systems take advantage of such divergent technical processes, representations of user queries, and methods for displaying documents, the selection of evaluation measures and the attributes used for comparison were carefully chosen and applied to capture the strengths and weakness of both types of systems.

1.3. NASA/RECON

NASA's implementation of RECON is one of the oldest of the large online bibliographic retrieval systems. In operation since 1969, it holds the special status of being the government's first large online retrieval system. A BBS, the system offers a wide range of commands for formulating queries and specifying the format of output. The system provides access to a body of citations to the scientific and technical literature for its users, who are primarily aerospace researchers and information specialists within the U.S. Government. In 1993, the database held about 3,000,000 citations organized within 29 files, which are grouped into 17 clusters. The largest of these are Scientific and Technical Aerospace Reports (STAR) and International Aerospace Abstracts (IAA).

1.4. Wide Area Information Server (WAIS)

WAIS is the name of a suite of programs using the client/server architecture and based on the NISO Z39.50 information retrieval protocol (Kahle & Medlar, 1991). This protocol allows searchers using a variety of interface "clients" to formulate queries which are translated into appropriate formats for particular "servers" that contain the bibliographic or primary information. A WAIS client allows users to formulate queries in "natural language" (terms or phrases can be typed in or cut and pasted from other sources in some clients) and to select as

---

[1] This automatic refinement technique is similar to the traditional manual citation pearl growing technique used in online searching wherein the user studies the characteristics of those valued retrieved documents ("pearls") and then attempts to retrieve more citations like them by adding the appropriate index terms to the query statement.

many databases (servers) as they wish for concurrent searching. A WAIS server is a retrieval engine that uses statistical ranking algorithms to return lists of information objects (typically documents, but possibly images, program code, etc.) that are ranked according to weighted frequency of occurrence. WAIS clients also invite users to provide relevance feedback to the servers by selecting relevant documents, phrases, or sections from the retrieved objects. Terms from these objects are fed back to the server and new ranked lists are sent to the client. A variety of public domain clients are available for different platforms (e.g., MS-DOS, Macintosh, X-windows), and there are public domain as well as commercial servers. Thus, WAIS searching requires high-speed connectivity to be effective and allows users to search a variety of servers without specialized training in query languages. Many WAIS servers are available through the Internet, and the client and server software packages are undergoing continued modification.

The proponents of WAIS system architecture make the following arguments: 1) The information search is driven by user perceptions of relevancy instead of the system's definition translated as an exact match between index and query terms. This distinction has been cast as the tension between relevancy, a match of document to query term, and pertinence, a match of document to actual information need. 2) The query representation automatically incorporates changes in the user's relevance judgments into the query vector and the order of the retrieved sets as the search proceeds. 3) The search process is made relatively transparent, allowing the user to concentrate fully on the intellectual demands of the information task instead of manual query refinement required by traditional systems. 4) Browsing by the user in the retrieved document sets can improve the final set's quality in terms of precision and recall and make it more reflective of the user's changing perception of the information need. Bates argues that this technique of systematic iterative retrieval, or "berrypicking," reflects how users actually approach the retrieval task and increases user understanding of their information need (Bates, 1989). 5) Searching different databases simultaneously insures a radical expansion of the retrieval engine's power and, probably, of the subsequent end value of the delivered document set (e.g., Stanfill & Kahle, 1986).

1.5. Study Perspective and Research Questions

As these brief descriptions imply, there are profound differences between BBS and WAIS in both surface and deep structure. Keen (1992) said, "Comparisons between Boolean and non-Boolean are not commonly attempted, and encounter difficulties with finding a control variable fair to both systems." Additionally, there are no accepted metrics for assessing the effectiveness of interactivity, an especially key aspect of WAIS. The approach that was taken for this work was to explore the differences between the two systems by controlling the database and search questions to ascertain strengths and differences for both approaches.

There is no overarching agreement within the information science field on the most appropriate measures to invoke and methods to use in evaluating an information retrieval system. Two measures that have been central to the debate about evaluation are recall and precision. Recall is the ability of the system to present all relevant items and is defined as the ratio of relevant documents retrieved to the number of relevant documents in the database.

Precision is the ability of the system to present only relevant items and is defined as the ratio of relevant documents retrieved to the total number of documents retrieved. It is generally recognized that recall and precision are inversely related and, thus, comparisons must be made at different levels of each.

The typical criticism of recall and precision as performance metrics relates to the problems of relevance (McCarn & Lewis, 1991). For example, 1) How does one determine the relevant items from those contained within a large information storage and retrieval system? 2) How can one be sure that users really care about relevant items not retrieved? 3) By what criteria should relevance be determined and to what degree are these judgements reliable? Beyond the individual interpretations of what is relevant to a question, even order effects have been found to influence judgments (Eisenberg & Barry, 1988). Nonetheless, recall and precision have become standard metrics for IR studies and were the first point of comparison for the work summarized here. Recall is the most troublesome value to compute in practice since it requires knowledge of the number of relevant documents in the entire database. Since this is impractical in working environments, this value is derived from extrapolation or estimates. For a system like WAIS that ranks documents and has quite severe cutoff levels (50 documents in the client used for these studies), recall is particularly problematic since not only must extrapolations be used to determine the number of relevant documents in the database, but adjustments for the cutoff limits must be invented. Additionally, recall has been questioned as a usable metric for large databases since there seems to be little overlap in relevant citations retrieved by different queries (Belkin, Cool, Croft, & Callan, 1993). Given these problems, recall was not used as a dependent measure in this study. The raw number of relevant documents is used for gross comparisons, and precision values are used for systematic comparisons.

Other criteria have been offered for IR evaluation. For example, Cleverdon (1964) not only used recall and precision but discussed effort (defined as the effort required by users to conduct the search), time (defined as the time necessary to generate the complete set of retrieved items), presentation, and coverage. Other possible criteria include cost, system learnability, system usability, and searcher satisfaction. Since the focus of this study was on retrieval performance, these criteria were not systematically assessed.

Two main questions guided this study and spawned more specific questions as data collection took place. These questions were:

- How do WAIS and the BBS perform when high recall, or the retrieval of all relevant documents, is the goal of a search?
- How do WAIS and the BBS perform when precision, or the retrieval of only relevant documents, is the goal of a search?

The end goal of the study was not to "prove" which is the "better" system. Rather, it was to provide a context wherein each system could display the features and capabilities which distinguish it from its counterpart. To this end, comparisons were used as the basis for discussions of the two different systems and as a springboard for gaining a better understanding of the particular values and limitations of both Boolean- and vector-based methods of information retrieval.

2. Methodology

2.1 Database and Search Systems

The database which served as the raw material for the two retrieval engines consisted of records in the Aerospace Database dated within the three year period 1990-1992. The Aerospace Database contains records for publicly available documents from NASA's STI Database. This set of documents consisted of approximately 205,000 document records which contained about 32 million words, about 540,000 of which were unique.

This database was mounted on a WAIS server (commercial version 1.0.10) on a Sun workstation at NASA's STI Program Office. The database used for the work reported here was not stemmed, so all results are based on exact matches of words in queries, a possible disadvantage from a recall perspective[2]. The WAIS database was accessed over the Internet via an ethernet connection using a public domain Macintosh WAIS client (version 1.1). This client offered a maximum cutoff value of 50 documents, but consistently returned a maximum of 45 documents.

As an example of a BBS, the Aerospace Database was accessed via the DIALOG network (Dialog File 108) using a MS-DOS platform and 2400 baud modem. All searches were date-limited to the same three year period as the records contained in the WAIS server.

2.2. Search Questions

Volunteers from among graduate students in aeronautical engineering were solicited to participate in the study. Their role was to supply questions (one question from each participant) for search and to decide upon the relevance of retrieved citations. Four students volunteered. Each was at a different stage in his research project; their questions varied in subject, complexity, and specificity. Two reference interviews were held with each participant. Prior to the first interview, the students sent an initial statement of their questions to the research team. These questions are summarized in Figure 1. During the first interview, they were asked to state their questions again and to note any synonyms or related terms. They were also asked about the usefulness of sources already consulted.

2.3. Procedure

The first step in the search process was to retrieve as complete a set of relevant documents as possible from the database. This was done by conducting exhaustive high recall searches on the Aerospace Database accessed through the Dialog system. Search terms were taken from the participants' queries, the print version of the NASA Thesaurus (1988 edition with

---

[2] However, Harman (1991) has demonstrated that stemming offers few advantages when ranking algorithms are used.

updates), and the online thesaurus in Dialog, which supplied the newest indexing terms. Free text searching was used as needed. For each of the four questions, a search strategy that would maximize recall was developed. Experienced online searchers from outside the project reviewed the strategies and offered suggestions for improvement.

Two searches in the Aerospace Database accessed through Dialog were conducted for each question. The first search was exploratory, incorporating terms in the original query statements and from the NASA Thesaurus. The second search expanded the terms used to include as many applicable terms as possible. In one case (question 3), a set of about 2500 citations was limited by NOTing out the common phrase "computational fluid dynamics" so that the set provided to the requestor was more manageable. These searches are referred to as the BBS recall searches. See Appendix A for search strategies.

Figure 1. Original Search Questions

Subject 1. The forces of pressure used to predict what the effect would be on a wing using computer models.

Subject 2. Basic Objective: Design of Flight Control System for Flexible Aircraft using Modern Control Theory.

Some Necessary Keywords (or Key Sentences) are Application of H-Infinity & H-2 Control Theory to Flight Control System; Development of Flexible Aircraft; Dynamic Model for Flight Control Design; Modern Control Theory & Its Implication to Aircraft Flying/Handling Qualities; Application of Active Control Technology to Transport Aircraft; Stability and Control of Control Configured Vehicle.

Subject 3. Adaptive gridding on structured grids. Other keys words I am interested in are "redistribution" and "truncation error." Some key words which are not of interest are "unstructured grids" and "refinement."

Subject 4. Computational Fluid Dynamics. I am trying to model flow over a bluff body, (something like the automobile side view mirror). I am using a computer code which solves the equations governing fluid dynamics. This branch of study to simulate flow using computers is called Computational Fluid Dynamics as opposed to conventional wind tunnel testing. I usually look for following Key Words: Unsteady, Incompressible, Bluff Body flows; High Reynolds number Flows, Turbulence Modelling; Finite Analytic Method; Non-Stationary turbulence; Ground effect; Unsteady pressure in incompressible flow.

A member of the research team met with each of the subjects to present the retrieved citation set. The team member asked the participant to rank the relevance of each citation, using a five-point scale with 1 as "Not Relevant" and 5 as "Highly Relevant." The citations rated 4 "Relevant" and 5 "Highly Relevant" formed the first component of the set of relevant documents in the database for the three-year period.

Subsequently, BBS searches were conducted to maximize precision rather than recall. These searches were constructed and modified to focus on documents containing all facets and are termed the BBS precision searches. See Appendix B for high precision search strategies.

The same four questions formed the basis of searches in WAIS. Initially, three searches were run on each question. The first search was a keyword search using an abbreviated set of terms from the query; in the discussion that follows this search is called the Terse search. These queries ranged in length from five to eleven terms. The second search was a keyword search using a more inclusive set of terms; it is called the Verbose search. These queries ranged in length from 25 to 42 terms. See Appendix C for the WAIS search strategies. The third search was a Boolean search, formulated as close as possible to the Boolean searches performed in Dialog, although it was impossible to duplicate the searches exactly. Only one Boolean search was successful; the other three searches did not produce usable results (timed out) in spite of multiple efforts. Because there was data from only one Boolean search, the data were not included in the analysis.

The citation sets retrieved using WAIS were combined into a master set for each question. The participants were asked to rate the relevance of each citation in the set, using the same five-point scale. In each of the four cases, the WAIS searches produced citations rated as 4 "Relevant" or 5 "Highly Relevant" that were not retrieved by the BBS searches. Seventeen items retrieved by both systems and for which the relevance judgements differed enough to affect the interpretation were eliminated from the database. The resulting sets of citations are the basis for the data analysis.

Two additional WAIS searches were run on each question. These searches utilized the relevance feedback function of WAIS in order to refine the retrieved set of citations. In each case, up to five citations from the WAIS search that were rated "Very Relevant" were designated as example citations, and the search was run again. Since there are many variations on how relevance feedback may be used and it is most appropriate for end users themselves to make judgments about what documents to feed back, this procedure must be taken as exploratory at best.

3. Results

3.1. Comparison BBS and WAIS Results

3.1.1. Comparisons Based on Precision

The distribution of citations by relevance judgement for all searches is shown in Table 1. The data for relevance judgments were recoded to produce three categories, "Not Relevant" (1 or 2), "Somewhat Relevant" (3), and "Highly Relevant" (4 or 5). For each question, the number of citations and number judgments in each relevance category are given for the BBS recall search, BBS precision search, WAIS terse search, and WAIS verbose search. Note that percentage values in the "high" category represent precision values.

In all questions, the precision results for the WAIS searches fell between the high recall and high precision results for the BBS searches. Since in practice recall and precision are inversely related, it would be highly surprising if a system that ranks output and uses a cutoff limit did not offer better precision than a high recall BBS search.

Table 1: Relevance Judgments by Search

QUESTION 1	NOT		SOME		HIGH		TOTAL RECORDS
BBS Recall	193	(80%)	18	(7%)	32	(13%)	245
BBS Precision	2	(25%)	0		6	(75%)	8
WAIS Terse	15	(34%)	7	(16%)	22	(50%)	44
WAIS Verbose	20	(44%)	8	(18%)	17	(38%)	45
QUESTION 2							
BBS Recall	57	(12%)	257	(52%)	176	(36%)	490
BBS Precision	3	(18%)	0		14	(82%)	17
WAIS Terse	4	(10%)	21	(53%)	15	(38%)	40
WAIS Verbose	4	(10%)	17	(44%)	18	(46%)	39
QUESTION 3							
BBS Recall	735	(93%)	26	(3%)	43	(5%)	804
BBS Precision	9	(39%)	3	(13%)	11	(48%)	23
WAIS Terse	21	(48%)	12	(27%)	11	(25%)	44
WAIS Verbose	36	(80%)	4	(9%)	5	(11%)	45
QUESTION 4							
BBS Recall	196	(81%)	10	(4%)	35	(15%)	241
BBS Precision	26	(68%)	0		12	(32%)	38
WAIS Terse	30	(67%)	5	(11%)	10	(22%)	45
WAIS Verbose	18	(40%)	23	(51%)	4	(9%)	41

As expected, the precision values for the WAIS searches were above those of the high recall BBS search. The precision values for the WAIS searches were consistently below the results for the BBS high precision searches. To illustrate these differences, the distributions of judgments across the three categories for each of the four types of searches were compared. Since several cells in the matrix of values for each question contained zeroes or small values, the results for the four questions were aggregated for the Chi Square test. A Chi Square value of 1267.3 ($p < .000$, 6 DF) demonstrates the expected result that the distributions are independent. Thus, the proportions of relevant documents for the four types of search were different, i.e., these four types of search represent different search strategies when taken as a group (comparison of the two WAIS strategies are presented in a subsequent section and indicate that these strategies do not differ).

Several observations can be made from these data. First, the high recall searches done in the BBS retrieved a higher number of relevant documents than the WAIS searches for each of the four questions. This is not surprising since the recall searches retrieved 241, 245, 490, and 804 documents respectively for the four questions and the WAIS searches could return at most 45 documents. The consequence of this is that if users wish to obtain high recall and they are willing to spend the time examining large sets of citations, then a BBS should be strongly considered.

Second, the high precision searches done in the BBS provided better precision results, although in two of the four questions the WAIS searches returned a higher number of relevant items. The implications of these data are less clear but suggest that if users are unwilling to examine even moderately sized document sets, better results may be obtained with a BBS search.

3.1.2. Cases

There were citations retrieved by both systems that were treated differently by the other system. Two of these instances are examined below. They illustrate citations which were retrieved by one system and given high relevance ratings but which were not retrieved by the other system.

Case 1: Retrieved by Boolean-based System. The first case is that of a citation retrieved by the BBS but not by WAIS. Fourteen search terms that were in the WAIS verbose search strategy appear in the citation record. It seems likely, therefore, that the citation was identified by the WAIS system as a candidate for retrieval, but it was not included because it was not in the top 45 citations. The average frequency of occurrence of the terms is two. The terms in the search statement that occurred most frequently in the citation were "finite" and "element." It is possible that these two terms are pervasive enough in the NASA database that their weight is relatively low.

It is important to note that there was no truncation capability in the WAIS implementation used in this study. Only the singular forms of the terms "error" and "grid" were entered in the search statement, while the plural of each term appears in the citation. Therefore, the terms were not matched by the WAIS system. The lack of this functionality poses a major problem for users; in order to compensate, the user must include both singular and plural forms of all terms, as applicable.

Case 2: Retrieved by WAIS. The second case is for a citation retrieved by WAIS but not by the BBS. This citation has terms which match concepts used in the search statement used by the Boolean-based system. For example, one search statement was matched by the major descriptor "computational grids." Similar matches were found for four other search terms. However, the citation also contains a term that was eliminated from the search statement. The phrase "computational fluid dynamics" was removed from the final search set of the Boolean-based search by using the NOT operator because the set of retrieved citations was extremely large. The term was included in the WAIS search, where set size was not an issue. In retrospect, perhaps some other strategy to limit the set size should have been taken rather than eliminating a search term that retrieved a highly relevant document.

This case is a good example of the pitfalls that confront users and trained searchers in working with a traditional retrieval system in a very large database. From an information point of view, removing the phrase from the final search set was not a good strategy because it eliminated this and, perhaps, other relevant citations. On the other hand, from a practical point of view, a huge citation set creates problems for the user. Even the limited set which excluded this citation contained more than 800 items; without this limitation, the set would have contained approximately 2500 items. Not many users have enough time to consider each citation in such a large set.

3.2. Analysis of WAIS Searches

3.2.1. Effect of Citation Set Size

As described above, a WAIS system presents the user with a limited set of citations ordered by a score which is assumed to reflect relevance to the search statement, with the citation with the highest score at the top of the list. Therefore, the top part of a citation list should be richer in relevant citations than the bottom part.

To explore the distribution of relevant documents in the WAIS ranked document lists, the lists were subdivided into three parts: top, middle, and bottom third of the ranked list. Precision values for these three sets of documents were then computed to determine if there was a higher concentration of relevant documents in the top-ranked set of results. Table 2 presents the precision values for each question. Contrary to expectations that the ranking algorithm should distribute proportionally more relevant documents near the top of the rankings, no apparent pattern of distribution was found.

Table 2. Precision Values for Three Partitions of Ranked Results by Question.

QUESTION	TERSE QUERY	VERBOSE QUERY
Question 1		
Top Third	.53	.27
Middle Third	.38	.47
Bottom Third	.64	.43
Question 2		
Top Third	.55	.55
Middle Third	.36	.60
Bottom Third	.42	.46
Question 3		
Top Third	.27	.25
Middle Third	.27	.15
Bottom Third	.21	.00
Question 4		
Top Third	.33	.07
Middle Third	.07	.20
Bottom Third	.20	.00

A Chi-Square test was used to test whether the distribution of documents by relevance differed among the three subsets. The data from the four questions were combined by type of WAIS search then divided into three subsets, by position. The subsets are the top set (first 15 items), middle set (second 15 items), and bottom set (third 15 items). The results for the Terse and Verbose searches are shown in Table 3.

Table 3. Chi-Square Values for Distribution of Relevant Citations by Position in the Set

SEARCH	VALUE	DF	SIGNIFICANCE
Terse	1.88	4	.757
Verbose	2.59	4	.628

The results illustrated that for both types of query relevant documents were just as likely to occur in the middle and bottom thirds as in the top third of the retrieved document set.

Various WAIS clients offer different cutoff levels; the Macintosh client used here offered a range from 5 to 50 citations. These results suggest that increasing the cutoff level for a WAIS search increases proportionally the relevant citations retrieved at least up to a citation set of 45 items. Conversely, limiting the size of the retrieved set is likely to limit proportionally the number of relevant citations retrieved. This finding raises questions about the relationship between the WAIS score and the actual relevance of a citation as judged by the user. Further analysis of the relationship between relevance judgment and WAIS score is necessary.

In sum, the ranking of documents in the WAIS searches did not distribute relevant documents in decreasing order. There was just as good a chance to find a relevant citation in the bottom third of the 45 retrieved as in the top third. Since the concentrations of relevant documents were good throughout the sets of 45, it seems prudent to provide higher cutoff values for users to choose if they wish.

3.2.2. Effect of Differences in WAIS Search Strategies. Two types of search statements were used for the WAIS searches: terse and verbose (included synonyms and related terms). Chi-Square tests were run on each pair of results to determine whether or not the results of the searches differed in the distribution of citations by relevance. For Question 2, an expected value of less than five was obtained for one-third of the cells; therefore, the analysis for this question was discarded. The Chi-Square values for the other three questions are shown in Table 4.

Table 4. Chi-Square Values for Distribution of Relevance Judgments by Type of WAIS Query Statement: Terse vs. Verbose.

QUESTION	VALUE	DF	SIGNIFICANCE
1	1.41	2	.494
3	15.34	2	.004
4	17.14	2	.000

The Chi-Square value for Question 1 indicates that results of the two WAIS query strategies do not differ. That is, the results of the Terse search and the Verbose search were similar in the distribution of citations by relevance ratings. The results for Questions 3 and 4, however, indicate that there were differences between the results of the searches for each of these questions; however, they are in opposite directions. For question 3, the precision of the terse query was lower than the precision of the verbose query; and for question 4, the precision of the Terse search was higher than for the Verbose search. Thus, there are no clear indications favoring either of the two strategies for these data.

This result is somewhat surprising because the Verbose search included more terms and should, therefore, have been more representative of the full scope of concepts in the original query. However, it may be that expanding the search statement can dilute the precision of a WAIS search by introducing terms that are less closely related to the core of the query, i.e., more noise. In this implementation of WAIS, it is not possible to weight the search terms or to attach any measure of importance, such as order in the search statement--the WAIS algorithm weights all terms equally. A citation will be retrieved by the total score derived by the algorithm, even if a significant part of the score is contributed by terms that are less important to the query.

3.2.2.1. Effects of Adding Terms

To explore the length of query effects in more detail, WAIS searches were systematically conducted by adding terms one at a time. Figures 2-5 present graphical views of how the number of relevant documents changes as an additional term is added to a WAIS search. The number of relevant documents in the set of 45 retrieved is plotted as each term is added. Thus, for Figure 2, there were 8 relevant documents for the query consisting of only the term "unsteady," six relevant documents for the query "unsteady aerodynamics," and 12 relevant documents for the query "unsteady aerodynamics stall." Frequency of occurrence of the term in thousands is given below each term.

Question 1 was motivated by a somewhat general topic, and the terms and question evolved over the two reference interviews. The dips in the graph (fewer relevant citations among the 45 retrieved) occurred as the frequently occurring and generic (for this database) terms "aerodynamics"(5500 occurrences in 4093 documents) and "air" (24,974 occurrences in 14,049 documents) were added to the query. The most common term "model" (63,547 occurrences in 34,027 documents) added some value since it is part of what may be considered a facet "stall delay model."

In question 2, the frequently occurring term "control" (76,573 occurrences in 27,170 documents) adds value to "flight" (37,273 occurrences in 18,684 documents), but the third term of that facet "system" (87,826 occurrences in 44,492 documents) is too common and increases noise. The next facet, "modern control theory," adds value word-by-word to reach a level almost as good as the simple "flight control" facet. Adding the new facet "flexible aircraft" adds surprisingly little value but does reinforce the need for using both terms of the facet.

Question 3 illustrates the importance of technical terms in this database. The term "grid" (7,513 occurrences in 3,605 documents) is the most salient term, adding most of the relevant documents for the entire search. The term "mesh" (2,310 occurrences in 1,344 documents), although relatively rare, adds considerable noise, perhaps due to its multiple technical usages.

Figure 2

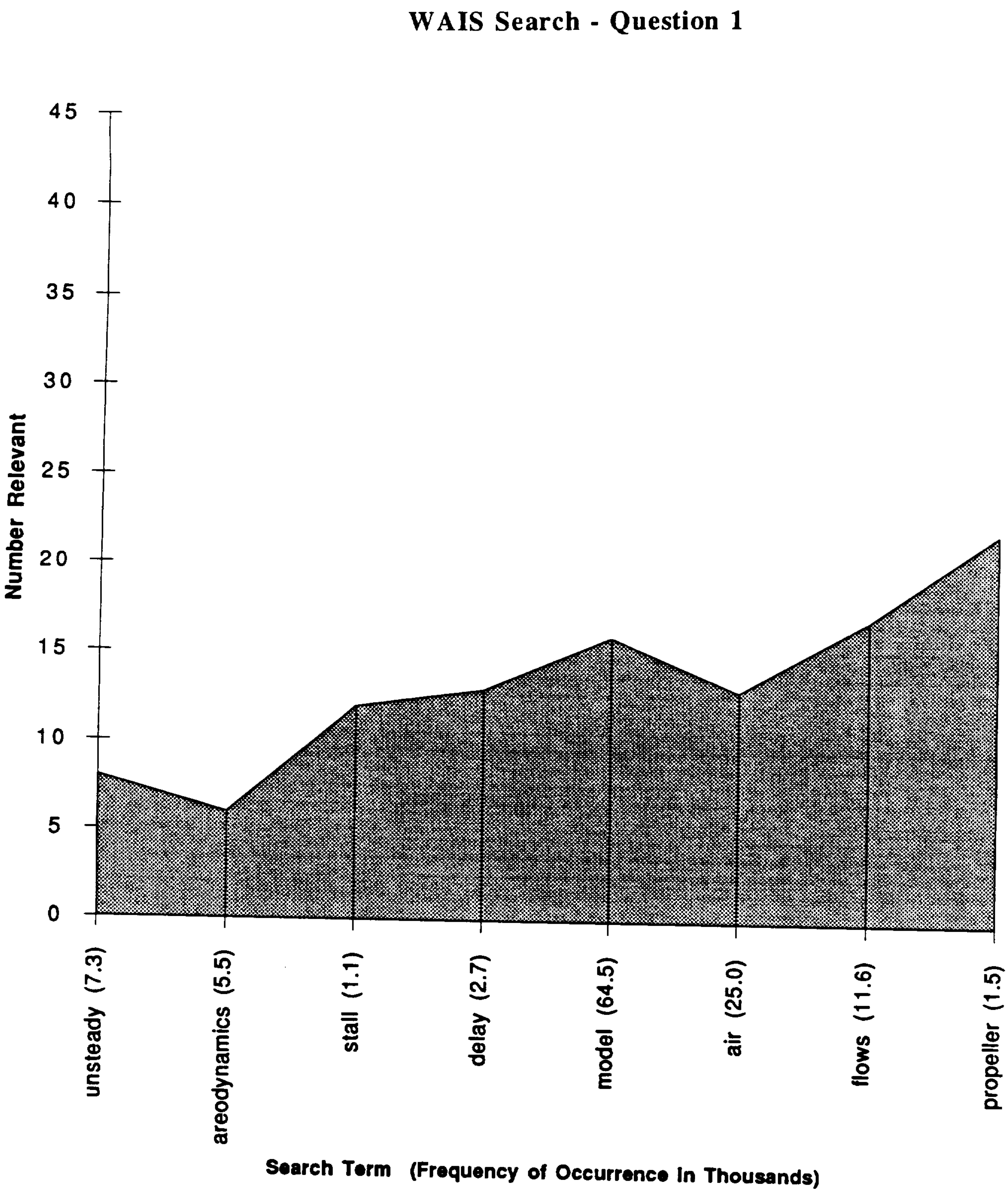

WAIS Search - Question 1
45
40
35
30
25
20
15
10
5
0
Number Relevant
unsteady (7.3)
areodynamics (5.5)
stall (1.1)
delay (2.7)
model (64.5)
air (25.0)
flows (11.6)
propeller (1.5)
Search Term (Frequency of Occurrence In Thousands)

Figure 3

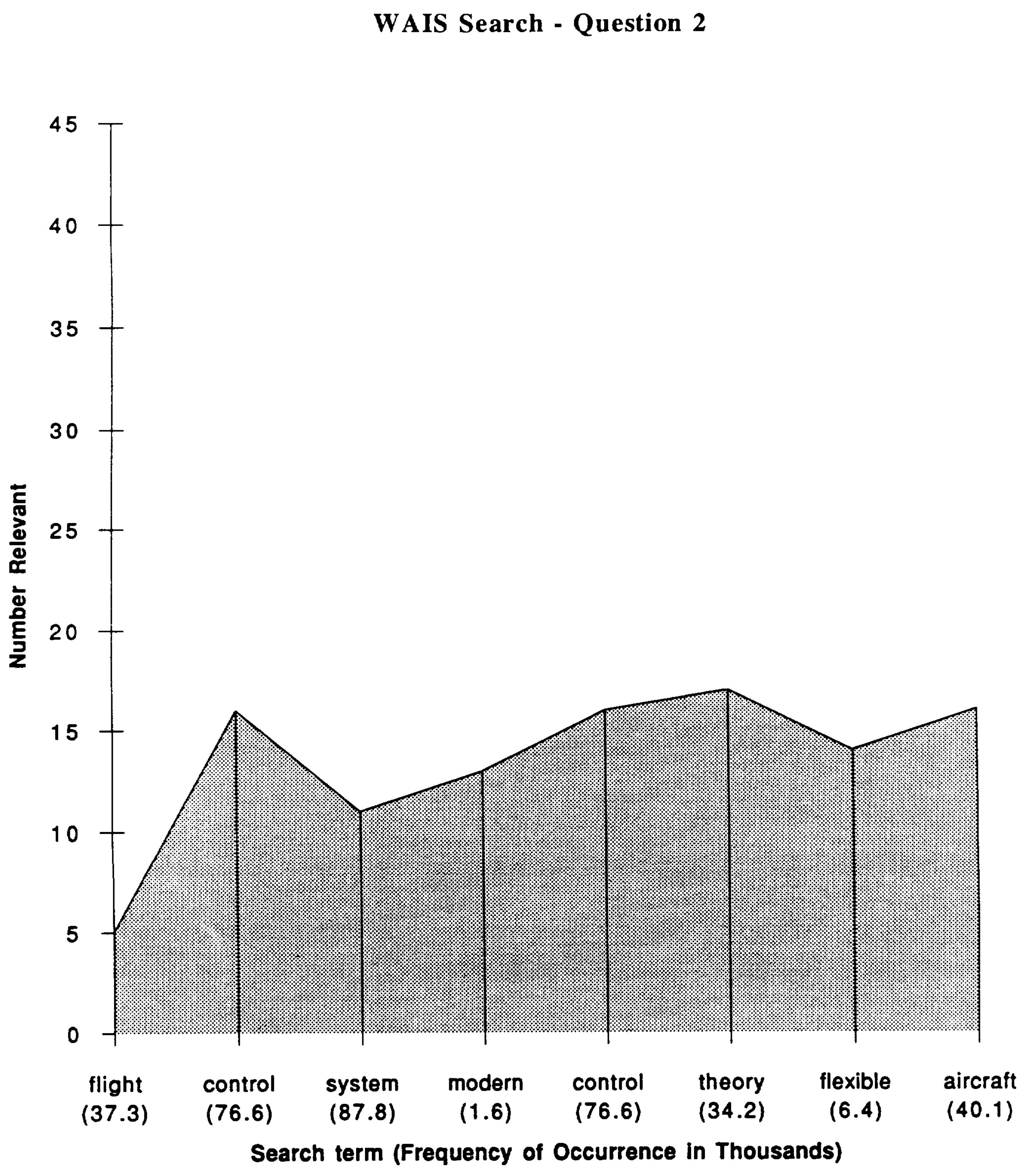

WAIS Search - Question 2
Number Relevant
45
40
35
30
25
20
15
10
5
0
flight (37.3)
control (76.6)
system (87.8)
modern (1.6)
control (76.6)
theory (34.2)
flexible (6.4)
aircraft (40.1)
Search term (Frequency of Occurrence in Thousands)

Figure 4

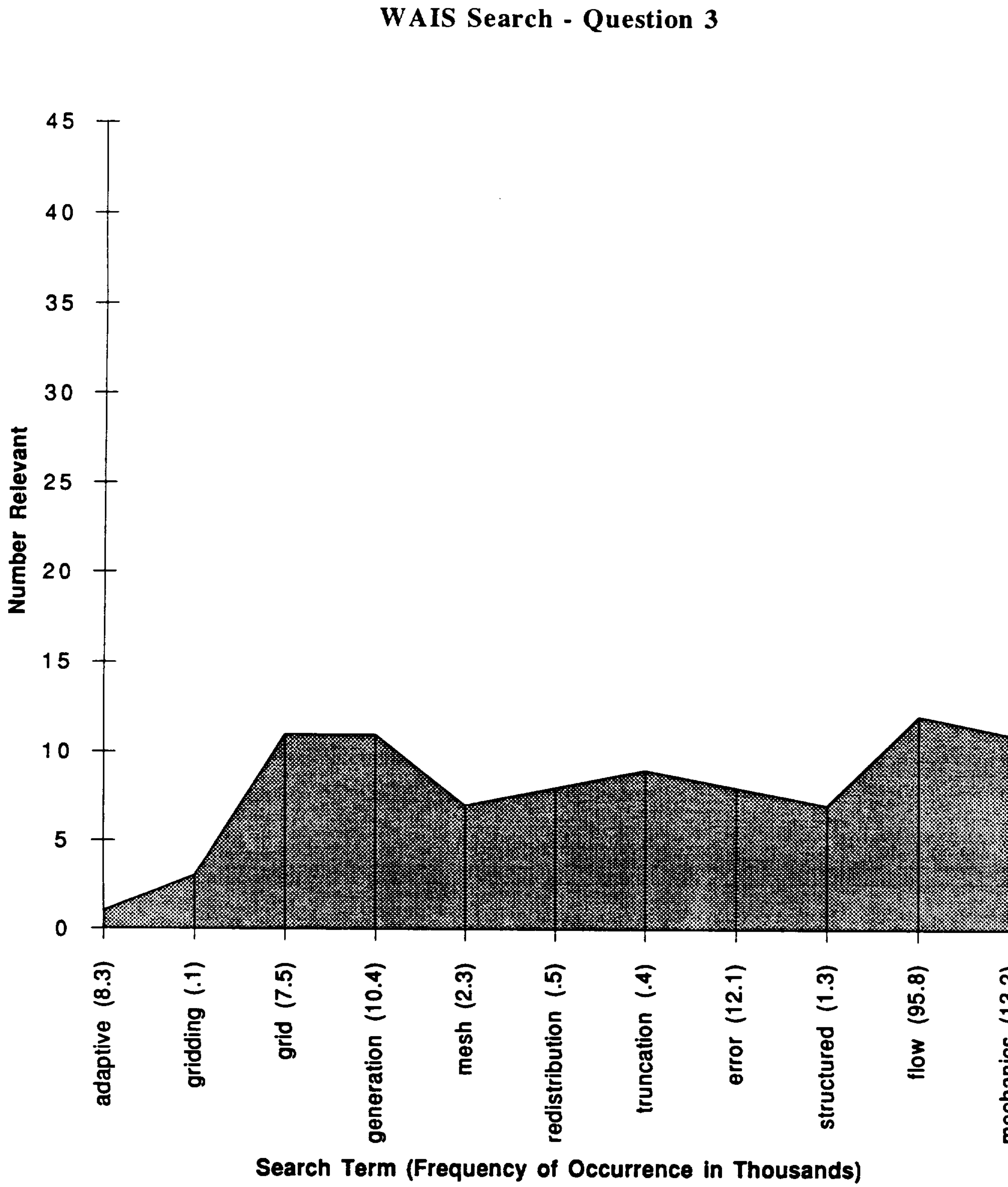
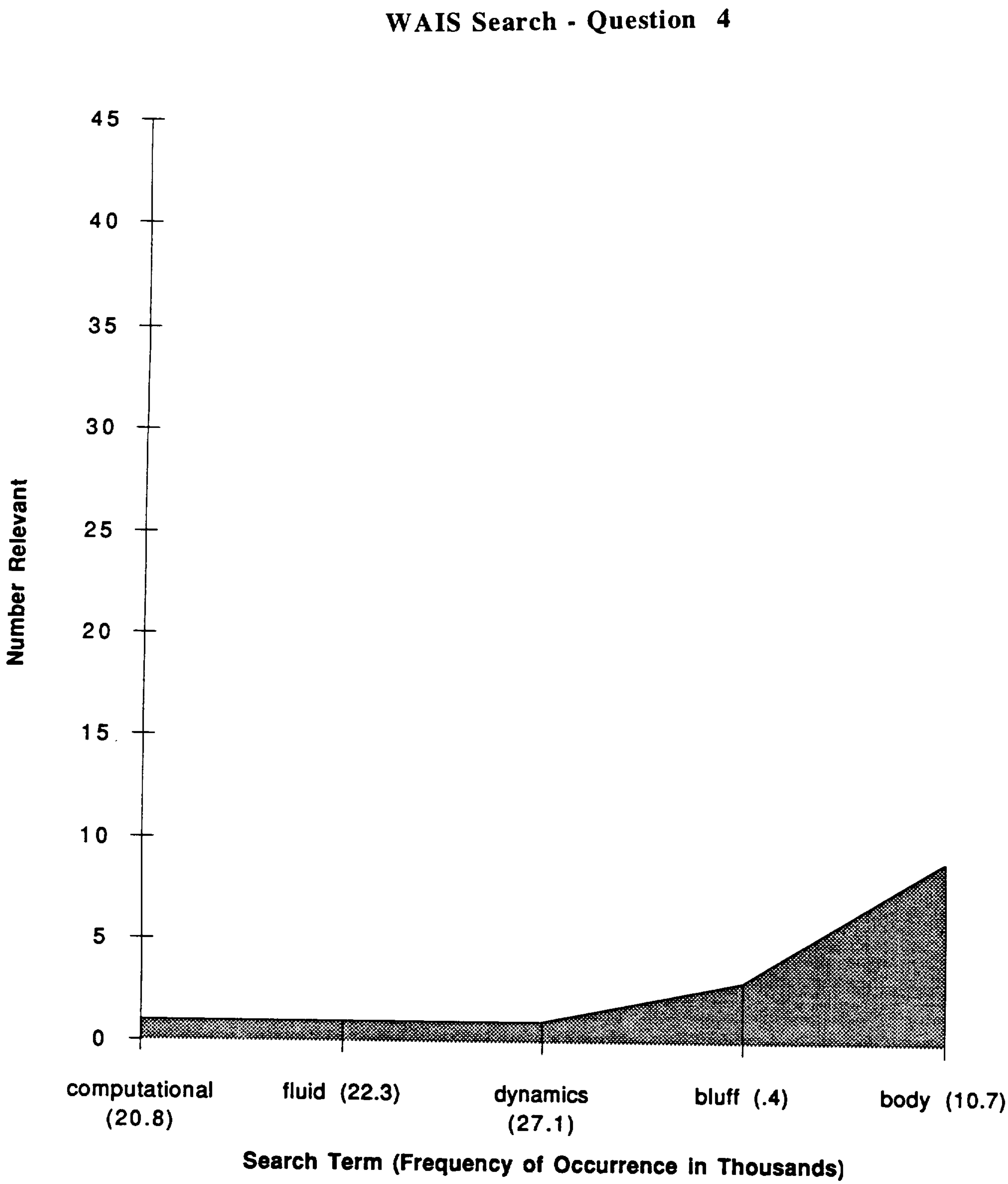

WAIS Search - Question 3
Number Relevant
45
40
35
30
25
20
15
10
5
0
adaptive (8.3)
gridding (.1)
grid (7.5)
generation (10.4)
mesh (2.3)
redistribution (.5)
truncation (.4)
error (12.1)
structured (1.3)
flow (95.8)
mechanics (13.2)
Search Term (Frequency of Occurrence in Thousands)

Figure 5


WAIS Search - Question 4
45
40
35
30
25
20
15
10
5
0
Number Relevant
computational (20.8)
fluid (22.3)
dynamics (27.1)
bluff (.4)
body (10.7)
Search Term (Frequency of Occurrence in Thousands)

Little value is added as additional generic terms are added until the enormously frequently occurring but technically important term "flow" (95,802 occurrences in 28,340 documents) is added to the query.

Question 4 illustrates how a common facet, "computational fluid dynamics," (computational occurs 20,767 times in 14,461 documents; fluid occurs 22,250 times in 13,977 documents; and dynamics occurs 27,051 times in 19,588 documents) is not helpful but the specific facet, "bluff body," (bluff occurs 373 times in 191 documents; body occurs 10,652 times in 6, 412 documents) does yield some positive results. Examining ratios of total occurrences to the number of documents having occurrences is suggestive. Recall may be improved when these ratios are relatively high (e.g., model, control, grid, and flow improve recall and have ratios of 2 or higher) and the occurrences are fairly high. It may be useful to examine whether such ratios are useful for query formulation (since these could be provided to the user during query formulation automatically) and whether some tolerance levels for usefulness can be developed.

This discussion illustrates the importance of mapping user need to facets and then considering which terms of a facet should actually be included in the query. This is certainly the case in BBS search strategy, but the use of controlled vocabularies makes these mapping decisions more tractable. For a system like WAIS, where the interplay between terms, facets, and the content of the entire database is so complex, the end-users special domain knowledge and immediately available intelligence seem necessary to give appropriate results. The end user's understanding of the information problem, the vocabulary of the field, and the scope of the database may substitute for query formulation tools such as thesauri.

3.2.3. Effect of Relevance Feedback. WAIS allows a user to use citations retrieved in one search as feedback for subsequent searches. The user marks citations and asks for other citations that are similar. Relevance feedback should re-sort the citations to bring more relevant items to the top of the ranked list, thus increasing both the recall and the precision within the cutoff value.

Relevance feedback was used with each of the eight WAIS searches. In each case, the documents designated as examples for the new iteration were the five documents with the highest WAIS scores that had been rated as "Very Relevant" by the requestor. (In the WAIS search which returned only four documents rated as "Very Relevant," only four documents were designated.) The results of the new search were compared with the results of the original WAIS search. It is important to iterate that this was a simplistic approach to relevance feedback. First, several different approaches should be taken (e.g., how many documents to feedback, feedback of document sections or words as well as citations, etc.). Second, relevance feedback is best done by end users, a condition made impossible by the artificial constraints of this study.

In six of eight searches, the results using relevance feedback were the same as the results from the original search. There were no differences in the specific citations retrieved nor in the order of presentation of the citations. However, the results for two searches, both for Question 4, were different. The distribution of relevant citations for these two searches is shown in Table 5.

Table 5. Relevance Judgments for Searches with Relevance Feedback for Question 4.

SEARCH	NOT RELEVANT		SOMEWHAT RELEVANT		VERY RELEVANT	
	NO.	%	NO.	%	NO.	%
Terse with Relevance Feedback	29	(64%)	7	(16%)	9	(20%)
Verbose with Relevance Feedback	28	(62%)	15	(33%)	2	(4%)

A Chi-Square test was used to test the expectation that the results of the searches using relevance feedback differed from the original WAIS searches. For the Verbose search, an expected cell frequency of less than five was found for two of the six cells; therefore, the results are not reported. The results for the Terse search yield a value of .40 [p=.82]. Thus, there are no significant differences between the original WAIS search and the search using relevance feedback.

However, other comparisons are of interest. The search with relevance feedback for the Terse search yielded six new citations out of a total retrieved set of 45 citations. None of the new citations was judged as "Very Relevant." Therefore, relevance feedback did not increase recall. The precision value for the Terse search with relevance feedback is .20, which is approximately the same as the precision value for the original Terse search, .22. The Verbose search with relevance feedback yielded 20 new citations, none of which was judged "Very Relevant." The precision value for the Verbose search with relevance feedback is .04, which is much lower than the precision value for the original Verbose search, .09.

The use of relevance feedback did not increase recall or precision in either of the two searches, although this result must be qualified by the small number of questions tested and the use of four or five citations as a whole as feedback. More extensive analyses of relevance feedback are needed to determine when and how it is most useful.

4. Summary and Discussion

4.1. Limitations of This Work

Although simple Boolean-based retrieval systems (BBS) have been the standard in operational systems for many years, today's systems have been enhanced to provide capabilities such as field limits, proximity qualification, and term weighting in the form of major or minor index terms. Term weighting and the new rank or target features of Dialog were not used; therefore, the Dialog BBS system was not fully exploited. Likewise, the WAIS system was not fully exploited. One of the most basic features of WAIS is its ability to distribute queries to multiple servers; this feature was not used at all. Likewise, the highly promoted "natural language" feature was not used; strings of terms and phrases (structured queries) were used rather than full-sentence queries or English clauses. Since both systems are evolving, the results from this study must be considered in the context of the current state of the art. This is particularly crucial for the WAIS system that has a relatively small experience base--a fact that makes explorations such as this one more important since findings can be used to influence subsequent iterations.

In addition to these limits on how the systems were used, this study was limited by the small number of questions, the fact that relevance judgments were made only by the requestor rather than a panel of experts, and the often-noted problems of making relevance judgments in general.

As stated earlier, the objective of this investigation was to compare the capabilities of Boolean and vector-based retrieval systems, emphasizing the strengths and appropriate uses of each. The preliminary results of this investigation and the experience gained in working with both systems on the same questions provide information useful in reaching this goal and point out questions needing further study. Certain strengths and weaknesses of each system can be discussed within the context of these findings, with particular emphasis given to WAIS.

4.2. Pros and Cons of BBS.

BBSs offer powerful search capabilities for professional searchers. The main strengths are related to high levels of control over how search is conducted. The system can yield high recall values if high recall strategies are used, and high precision if high precision strategies are used. Thus, BBSs provide a good range of features for constructing searches customized to specific user needs. Features such as proximity operators and file limits give the searcher greater control in constructing a search query and allow results to be interpreted directly. Additionally, there is a large base of experience in using this type of system and extensive research on various search techniques and strategies.

Limitations of BBS include the "flip side" of multiple features and good control--high levels of expertise are required. Searches require a considerable investment of proximity resources--time, effort, and money. The searchers in this project spent an average of three hours preparing for each search. Each search consisted of as many as 56 search statements and required at least an hour in an online session. Additionally, the size of the retrieved set can be very large, particularly when trying to maximize recall. The case cited earlier points out a trade-off encountered in trying to limit set size.

4.3. Problems with WAIS

This study illustrated several problems with the WAIS system. The lack of tools, such as proximity operators other than immediate adjoining to the right, means that the user has less control over a search statement. Use of relevance feedback should be a compensating feature, but the results from relevance feedback were not encouraging. Secondly, the established size of the citation set artificially limits the number of relevant citations retrieved. Thirdly, the experience base, documentation, and training for vector-based systems are undeveloped, at this time. Some of these problems are likely due to the limited nature of the investigation and problems with server or client implementations, but others require longer-term consideration. First, WAIS provides a "black box" effect for the user. Users put in queries and get results based upon many hidden computations. For end users this may be just what is needed if performance can be demonstrated to be good. For professional intermediaries, however, using WAIS can be quite frustrating since it is impossible to deduce exactly how results were obtained. Furthermore, searchers have good control over how they conduct searches in a BBS; e.g., they can design queries to maximize either precision or recall. Strategies for constructing queries, setting cutoffs and using relevance feedback with WAIS are yet to be identified and tested for different types of needs.

Second, the ranking algorithm used in WAIS yielded disparate results in this study. Substantial additions and modifications may be needed if WAIS is to be used in high-performance environments where results are mission-critical. This has been recognized by WAIS Inc. since the commercial version of the server provides for Boolean operators and a query report that gives posting data. Continued movement toward a more hybrid approach to information retrieval seems prudent if WAIS is to be useful for end users and professional intermediaries alike. More problematic is the way that relevance feedback is handled. At present, users have no control over how weights are assigned--a double-edged sword from learnability and user control perspectives. The low weights given to terms in the feedback query compared to terms in the original query reduced the effectiveness of relevance feedback in query refinement. Likewise, there is very poor control for setting cutoff values which affects both relevance judgments for ranked sets and feedback decisions. Also, distinctions between different iterations of a search are difficult to make unless each iteration is saved and then compared separately with a word processing or text management package. Adding new features for setting feedback parameters and for manipulating intermediate result sets will occupy the research and development community for some time to come.

4.4. Potentials with WAIS

Several positive aspects of using WAIS were illustrated by this work. First, the system yields generally acceptable precision values for little user effort. Second, less time and effort are required to conduct a search in WAIS than in the BBS. The time required to conduct a WAIS search ranged from approximately 40 seconds to slightly over 60 seconds. It must be noted, of course, that the preliminary work of structuring the search had been done for the BBS searches and was not counted as part of the WAIS search process. Also, WAIS utilizes natural language expressions as search queries, which should minimize the time needed to structure search statements. Third, the established size of the citation set provides a means of controlling the number of document records that the requestor must review, which is valuable in producing results that are immediately useful.

Although it is sometimes the case that no documents are returned in a WAIS search, it is much more likely that something is found that can serve as a starting point for subsequent search or browsing. This is in contrast to BBS systems such as OPACs and Grateful Med that typically yield "no hits" for about one-third of all end user searches. Thus, WAIS offers users starting points for interactive sessions. Although the purpose of this study centered on retrieval performance rather than the interface, usability effects were clearly apparent and represent significant advantages over BBS. First, there is no query language to learn--a clear advantage for novice and occasional users. Second, WAIS is easy to use, offering direct manipulation interaction styles in the Macintosh and X-Windows clients. These clients invite exploration and interaction and allow editing/cutting and pasting operations to be leveraged as reinforcement to continuous interactivity rather than as simple short-cuts for discrete, linear search.

4.5. Implications

Highly interactive systems invite end users to do their own searching and offer massive amounts of information. Although much time can be spent exploring with WAIS (other systems such as Mosaic depends almost entirely on exploration), questions about efficiency and cost-effectiveness will become more critical as these systems are used in commercial settings and as the costs of network access and use become less transparent.

There are two sets of problems that require research. First, it seems clear that highly interactive front-ends such as WAIS demand new information-seeking strategies. Just as it took time and experience for the online community to develop and teach strategies for BBS such as "building block," "successive fraction," "pearl growing," and "iterative scanning" (Hawkins & Wager, 1982), it will take time to discover strategies that serve different search purposes and users in WAIS and other emerging systems. For example:

- What effects will query length have in such environments, and how should users control relevance feedback?
- Will domain experts use relevance feedback differently than intermediaries?
- Should term weighting be used, and if so, how?
- How will multiple servers affect results?

Just as intermediaries have developed divide and conquer strategies to handle problems with many facets using a BBS, there may be a need for staged strategies for different databases or customization of queries for specific databases using WAIS. Although WAIS clients provide options to display results by server, this is a simple beginning to distinguish and develop more carefully targeted revisions/feedback to specific servers. Finally, better ways are needed to keep the user informed of how the search is progressing and why results have been obtained. Intermediate posting data and sophisticated query reports may be helpful here.

Second, there is a need for new evaluation metrics that address the interactive nature of systems like WAIS. Variables such as time, number of queries, number of iterations, number of screens/windows displayed, number or distribution of records displayed or saved, cost, and satisfaction may take on new meanings in such environments. New variables such as number of servers queried, number of servers returning results, and filtering values for different servers may be needed. As electronic environments become more seamlessly integrated, it may become feasible to trace how information actually gets integrated into user work (e.g., text or citations cut and pasted into documents, figures imported into spreadsheets or local databases, etc.) to gain a fuller understanding of how information seeking supports user work. It will become more feasible to assess where information objects came from as electronic documents are constructed as a series of links between existing and original text, images, and other objects. Some attention to assessing the overall process rather than simply the products of searching must also be considered (e.g., search patterns from symbolic or graphical perspectives, Lin, Liebscher, & Marchionini, 1991).

The results summarized here demonstrate some of the possibilities as well as problems that systems such as WAIS provide for the IR community. Rather than aiming to compare systems on simplistic metrics rooted in tradition, we are challenged to develop systems that apply different techniques and interfaces and to test those systems from formative as well as comparative perspectives. This effort used comparison as the basis for better understanding a specific innovative system. The results temper the "hype" associated with new technologies and raise questions about assumed performance and features. It suggests directions for improving this particular system in subsequent iterations, but more importantly, raises issues about what types of questions seem promising for research, development, and evaluation.

Acknowledgements

The authors wish to acknowledge and thank Dr. Linda Hill for her contributions to this study; Dr. Brewster Kahle for making the commercial version of the WAIS server software available for use; and Dialog, Inc. for supporting the searches of the Aerospace Database on Dialog. Thanks are also due to Kristi Mashon, Vince Boisselle, and Jill Strass who conducted searches and assisted in the data analysis.

References

Bates, M. (1989). The design of browsing and berrypicking techniques for the online search interface. Online Review, 13(5), 407-424.

Belkin, N., Cool, C., Croft, B., Callan, J. (1993). The effect of multiple query representations on information retrieval system performance. Proceedings of the 16th SIGIR Conference (Pittsburgh, June 27-July 1, 1993), ACM Press. 339-346.

Buchan, R. L. (1991). Technology transfer at NASA--A librarian's view. Science and Technology Libraries Winter.

Cleverdon, C.W. (1964). Evaluation of operational information retrieval systems. Part 1: Identification of criteria. Cranfield, England: College of Aeronautics.

Dialog Inc. (1993). TARGET--A new way to search on Dialog. Dialog Chronicle, Nov/Dec. 319-321.

Eisenberg, M. & Barry, C. (1988). Order effects: A study of the possible influence of presentation order on user judgments of document relevance. Journal of the American Society for Information Science, 39(5), 293-300.

Harman, D. (1991). How effective is suffixing? Journal of the American Society for Information Science, 42(1), 7-15.

Harman D. & Candelea, G. (1990). Retrieving records from a Gigabyte of text on a minicomputer using statistical ranking. Journal of the American Society for Information Science 41: 581-89.

Harman, D., McCoy, W., Toenese, R., & Candelea, G. (1991). Prototyping a distributed information system that uses statistical ranking. Information Processing and Management 27: 449-60.

Hawkins, D.T. & Wagers, R. (1982). Online bibliographic search strategy development. Online, 6(3), 12-19.

Ide, E, (1971). New experiments in relevance feedback. In G. Salton (Ed.) The Smart System-Experiments in Automatic Document Processing. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Kahle, B. & Medlar, A. (1991). An information system for corporate users: Wide Area Information Servers. Online, September, 55-60.

Keen, E.M. (1992). Presenting results of experimental retrieval comparisons. Information Processing and Management, 28(4), 491-502.

Larson, R. (1992). Evaluation of advanced retrieval techniques in an experimental online catalog. Journal of the American Society for Information Science, 43(1), 34-53.

Lin, X., Liebscher, P., & Marchionini, G. (1991). Graphic representation of electronic search patterns. Journal of the American Society for Information Science, 42(7), 469-478.

McCarn, D. & Lewis, C. (1990). A mathematical model of retrieval system performance. Journal of the American Society for Information Science, 41(7), 495-500.

Obraczka, K., Danzig, P., & Li, S. (1993). Internet resource discovery services. IEEE Computer, 26(9), 8-22.

Rocchio, J. J. (1971). Relevance feedback in information retrieval. In G. Salton (Ed.) The Smart System-Experiments in Automatic Document Processing. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Salton, G. and McGill, M.J. (1983) Introduction to Modern Information Retrieval. New York: McGraw-Hill.

Stanfill, C. & Kahle, B. (1986). Parallel free-text search on the Connection Machine System. Communications of the ACM, 29(12), 1229-1239.

Appendix A: BBS Recall Search Strategies

Subject One
DIALOG Recall Search Statement:

```
File 108:AEROSPACE   62-93/9306B2
    Set  Items  Description
    ---  -----  -----------

?ss unstead?
                                        Search the unsteady state concept
   S1  20326  UNSTEAD?

?ss s1()(state? or flow? or aerodynam? or pressure?)
          20326  S1
   S2   153877  STATE?
   S3   247128  FLOW?
   S4    58099  AERODYNAM?
   S5   150226  PRESSUR?
   S6    14334  S1()(STATE? OR FLOW? OR AERODYNAM? OR PRESSUR?)

?ss turbulen?
   S7    70594  TURBULEN?

?ss s7()(flow? or wake? or model? or effect?)
          70594  S7
         247128  S3
   S8    12835  WAKE?
   S9   335200  MODEL?
   S10  455590  EFFECT?
   S11   29748  S7()(S3 OR WAKE? OR MODEL? OR EFFECT?)

?ss s7()boundar?()layer?
          70594  S7
   S12  135730  BOUNDAR?
   S13  121839  LAYER?
   S14   11509  S7()BOUNDAR?()LAYER?

?ss nonstation?
   S15    4758  NONSTATION?

?ss non()station?
   S16   37066  NON
   S17   66489  STATION?
   S18     525  NON()STATION?

?ss s15 or s18
           4758  S15
            525  S18
   S19    5108  S15 OR S18
```

```
?ss  s19()s7
       5108  S19
      70594  S7
   S20     34 S19()S7

?ss s1 or s5 or s6 or s7 or s11 or s14 or s19 or s20
```

Combine synonyms for the unsteady state concept

```
        20326  S1
       150226  S5
        14334  S6
        70594  S7
        29748  S11
        11509  S14
         5108  S19
           34  S20
  S21 223959  S1 OR S5 OR S6 OR S7 OR S11 OR S14 OR S19 OR
              S20

?ss  s21  or  (model?()air()flow?)  or  (air()flow?)
        223959  S21
   S22  335200  MODEL?
   S23  115109  AIR
   S24  247128  FLOW?
   S25        1  MODEL?(W)AIR(W)FLOW?
   S26  115109  AIR
   S27  247128  FLOW?
   S28    7465  AIR(W)FLOW?
    S29  227922  S21 OR (MODEL?()AIR()FLOW?) OR (AIR()FLOW?)

?ss  s29/de*
```

Sift to see how many citations were pulled from major descriptors

```
   S30   83616  S29/DE*

?ss s29/de
```

Sift to see how many citations were pulled from all descriptors

```
   S31  164207  S29/DE

?ss model?
```

Search the mathematical/computer models concept

```
  S32  335200  MODEL?

?ss  math?()s32
  S33  166591  MATH?
       335200  S32
  S34    87641  MATH?()S32
```

```
?ss computer?()s32
   S35 199247 COMPUTER?
       335200 S32
   S36   4145 COMPUTER?()S32

?ss (numeric?(2n)s35)()s32
   S37 131256 NUMERIC?
       199247 S35
       335200 S32
   S38     34 (NUMERIC?(2N)S35)()S32

?ss digit?()simulat?
   S39  61786 DIGIT?
   S40 126899 SIMULAT?
   S41   5918 DIGIT?()SIMULAT?

?ss s32 or s34 or s36 or s38 or s40 or s41
```

Combine synonyms for the mathematical/ computer models concept

```
       335200 S32
        87641 S34
         4145 S36
           34 S38
       126899 S40
         5918 S41
   S42 408079 S32 OR S34 OR S36 OR S38 OR S40 OR S41

?ss s42/de*
```

Sift to see how many citations were pulled from major descriptors

```
   S43 109315 S42/DE*

?ss s42/de
```

Sift to see how many citations were pulled from all descriptors

```
   S44 237530 S42/DE

?ss propell?
```

Search the propeller state concept

```
   S45  29531 PROPELL?

?ss high()solid?()s45
   S46 289326 HIGH
   S47  86661 SOLID?
        29531 S45
   S48      2 HIGH()SOLID?()S45

?ss s45()blade?
```

```
       29531  S45
  S49  23717  BLADE?
  S50   1027  S45()BLADE?

?ss (compress? or fan? or rotor? or stator or turbine?)()s49
  S51  54002  COMPRESS?
  S52   6742  FAN?
  S53  22141  ROTOR?
  S54   1985  STATOR
  S55  30977  TURBINE?
       23717  S49
  S56  15931  (COMPRESS? OR FAN? OR ROTOR? OR STATOR OR
              TURBINE?)()S49

?ss wind()s55
  S57  85351  WIND
       30977  S55
  S58   2248  WIND()S55

?ss (s57()power?)()(utiliz? or generat?)
       85351  S57
  S59 146391  POWER?
  S60  63826  UTILIZ?
  S61 121687  GENERAT?
  S62     89  (S57()POWER?)()(UTILIZ? OR GENERAT?)

?ss streamline?()bod?
  S63   4040  STREAMLIN?
  S64  93722  BOD?
  S65    275  STREAMLIN?()BOD?

?ss engine()part?
  S66  47311  ENGINE
  S67 296392  PART?
  S68   3034  ENGINE()PART?

?ss windmill?
  S69    886  WINDMILL?

?ss s45 or s48 or s49 or s50 or s56 or s58 or s62 or s65 or s68
    or s69 or s55
```

Combine synonyms for the propeller concept

```
       29531  S45
           2  S48
       23717  S49
        1027  S50
       15931  S56
        2248  S58
          89  S62
```

```
          275  S65
         3034  S68
          886  S69
        30977  S55
   S70  72042  S45 OR S48 OR S49 OR S50 OR S56 OR S58 OR S62
               OR S65 OR S68 OR S69 OR S55
```

```
?ss s70/de*
```

Sift to see how many citations were pulled from major descriptors

```
   S71   43374  S70/DE*
```

```
?ss s70/de
```

Sift to see how many citations were pulled from all descriptors

```
   S72   59245  S70/DE
```

```
?ss delay?
```

Search the stall delay concept

```
   S73   18147  DELAY?

?ss stall?()s73
   S74    4056  STALL?
         18147  S73
   S75      25  STALL?()S73

?ss aerodynam?()(s74 or stabil?)
   S76   58099  AERODYNAM?
          4056  S74
   S77  135757  STABIL?
   S78    4726  AERODYNAM?()(S74 OR STABIL?)

?ss engine?()fail?
   S79  155807  ENGINE?
   S80   45414  FAIL?
   S81    1209  ENGINE?()FAIL?

?ss (aircraft? or airplane?)()perform?
   S82 121142 AIRCRAFT?
   S83   6021 AIRPLANE?
   S84 228951 PERFORM
   S85   5671 (AIRCRAFT? OR AIRPLANE??)()PERFORM?

?ss stop???? ?
   S86    5945  STOP???? ?

?ss lateness
   S87      22  LATENESS
```

```
?ss  time()lag?
    S88  251319  TIME?
    S89   20890  LAG?
    S90    7038  TIME?()LAG?

?ss s73 or s74 or s75 or s77 or s78 or s81 or s85 or s86 or s87
    or s90
```

Combine synonyms for the stall delay concept

```
         18147  S73
          4056  S74
            25  S75
        135757  S77
          4726  S78
          1209  S81
          5671  S85
          5945  S86
            22  S87
          7038  S90
   S91  168912  S73 OR S74 OR S75 OR S77 OR S78 OR S81 OR S85
                OR S86 OR S87 OR S90

?ss  s91/de*
```

Sift to see how many citations were pulled from major descriptors

```
    S92   73227  S91/DE*

?ss  s91/de
```

Sift to see how many citations were pulled from all descriptors

```
    S93  121954  S91/DE

?ss s30 and s43 and s71 and s92
```

Combine the four separate concepts together (unsteady state, mathematical/computer models, propeller, and stall delay) for the sets with major descriptor terms

```
         83616  S30
        109315  S43
         43374  S71
         73227  S92
   S94      11  S30 AND S43 AND S71 AND S92

?ss s31 and s44 and s72 and s93
```

Combine the four separate concepts together (unsteady state, mathematical/computer models, propeller, and stall delay) for the sets with all descriptor terms

```
        164207  S31
        237530  S44
```

```
       59245  S72
      121954  S93
  S95    243  S31 AND S44 AND S72 AND S93
```

?ss s29 and s42 and s70 and s91

Combine the four separate concepts together (unsteady state, mathematical/computer models, propeller, and stall delay) for the sets with natural language terms

```
      227922  S29
      408079  S42
       72042  S70
      168912  S91
  S96   1334  S29 AND S42 AND S70 AND S91
```

?ss ud=9001:9212b2

Set the update restriction of 1990-1992

```
   S97  205080  UD=9001:9212B2
```

?ss s94 and s97

Limit the major descriptor citations by date

```
          11  S94
      205080  S97
  S98      3  S94 AND S97
```

?ss s95 and s97

Limit all descriptor citations by date

```
         243  S95
      205080  S97
  S99     48  S95 AND S97
```

?ss s96 and s97

Limit the natural language citations by date

```
        1334  S96
      205080  S97
  S100   244  S96 AND S97
```

?ss s99 not s98

Compared sets by NOTing to compare citations pulled in through major descriptors, all descriptors, and natural language

```
          48  S99
           3  S98
  S101    45  S99 NOT S98
```

?ss s100 not (s98 or s101)

```
         244  S100
```

```
              3  S98
             45  S101
   S104     196  S100 NOT (S98 OR S101)

?logoff
```

Subject Two

DIALOG Recall Search Statement:
File 108:Aerospace Database 1962-1993/Jul B1

```
    Set  Items  Description
    ---  -----  -----------
```

Expand on select authors' names and titles to look at indexing

```
S1       10  AU="MYERS, THOMAS T."
S2       20  AU=MCRUER, D.
S3       47  AU=MCRUER, D. T.
S4        5  AU=MCRUER, DUANE
S5       16  AU=MCRUER, DUANE T.
S6       88  E4-E7
S7        2  AU=THOMSON, P. F.
S8        2  AU=THOMSON, P. J.
S9        4  E4-E5
S10       2  AU=ASHKENAS, I.
S11      42  AU=ASHKENAS, I. L.
S12       4  AU=ASHKENAS, IRVING L.
S13      48  E6-E8
S14      16  AU=DOYLE, J.
S15      32  AU=DOYLE, J. C.
S16      48  E4-E5
S17 15649  ADVANCE?/TI
S18  6969  PILOT?/TI
S19 40127  AIRCRAFT/TI
S20 30490  FLIGHT/TI
S21 61210  CONTROL/TI
S22181326  SYSTEM?/TI
S23       2  ADVANCE?()PILOT?()AIRCRAFT()FLIGHT()CONTROL()
           SYSTEM?/TI
S24       2  RN=STI-TR-1228-1
```

?ss modern()control?()theor?

Search the modern control theory concept

```
   S25    9189 MODERN
   S26  235735 CONTROL?
   S27  304734 THEOR?
   S28     239 MODERN()CONTROL?()THEOR?
```

?ss mct
```
   S29    77 MCT
```

?ss control()theor?

```
S30 235735  CONTROL?
S31 304734  THEOR?
S32  13852  CONTROL?()THEOR?
```

```
?e h-in
```

Expand on h-infinity control in the basic index

```
Ref   Items  Index-term
E1       8   H-BETA LINE
E2       4   H-GAMMA LINE
E3       0   *H-IN
E4      53   H-INFINITY CONTROL
E5      59   H-LINE
E6      53   H-WAVE
E7      10   H-1 ENGINE
E8       7   H-1 ROCKET ENGINE
E9       3   H-126 AIRCRAFT
E10      1   H-126 JET FLAP AIRCRAFT
E11      1   H-13 AIRCRAFT
E12      6   H-17 HELICOPTER
```

```
?ss e4
```

Select the appropriate term

```
   S33     53  "H-INFINITY CONTROL"
```

```
?ss  h()infinit?()control?
   S34   40267  H
   S35   23343  INFINIT?
   S36  235735  CONTROL?
   S37     247  H()INFINIT?()CONTROL?
```

```
?ss  h()infinit?
   S38   40267  H
   S39   23343  INFINIT?
   S40     635  H()INFINIT?
```

```
?e h-2
```

Expand on h-2 control in the basic index

```
Ref   Items  Index-term
E1       6   H-17 HELICOPTER
E2       1   H-19 HELICOPTER
E3       0   *H-2
E4       1   H-25 HELICOPTER
E5       1   H-34 HELICOPTER
E6       1   H-43 HELICOPTER
E7      82   H-53 HELICOPTER
E8       7   H-54 HELICOPTER
E9      14   H-56 HELICOPTER
E10     43   H-60 HELICOPTER
```

```
E11   304  HA
E12    11  HAA

?e h 2
Ref   Items  Index-term
E1      13  H REGION
E2     161  H WAVES
E3       0 *H 2
E4      63  H-ALPHA LINE
E5       8  H-BETA LINE
E6       4  H-GAMMA LINE
E7      53  H-INFINITY CONTROL
E8      59  H-LINE
E9      53  H-WAVE
E10     10  H-1 ENGINE
E11      7  H-1 ROCKET ENGINE
E12      3  H-126 AIRCRAFT

?ss  h()2()control?
   S41   40267  H
   S42  195123  2
   S43  235735  CONTROL?
   S44       3  H()2()CONTROL?

?ss s28 or s29 or s32 or s33 or s37 or s40 or s44
```

Combine synonyms for the modern control theory concept

```
       239  S28
        77  S29
     13852  S32
        53  S33
       247  S37
       635  S40
         3  S44
   S45  14284  S28 OR S29 OR S32 OR S33 OR S37 OR S40 OR S44

?ss  s45/de*
```

Sift through citations in which these concepts were major descriptors

```
   S46   7280  S45/DE*

?ss s45/de
```

Sift through those citations in which these concepts were major or minor descriptors

```
   S47  12827  S45/DE

?ss  flight?()simulat?
```

Searching the flight control concept

```
   S48  110968  FLIGHT?
```

```
    S49  126899  SIMULAT?
    S50    8968  FLIGHT?()SIMULAT?

?ss  flight?()control?()design?
    S51  110968  FLIGHT?
    S52  235735  CONTROL?
    S53  260987  DESIGN?
    S54     125  FLIGHT?()CONTROL?()DESIGN?


?ss  flight?()control?()theor?
    S55  110968  FLIGHT?
    S56  235735  CONTROL?
    S57  304734  THEOR?
    S58       5  FLIGHT?()CONTROL?()THEOR?

?ss  flight?()control?
    S59  110968  FLIGHT?
    S60  235735  CONTROL?
    S61   10621  FLIGHT?()CONTROL?

?ss  active?()control?
    S62  104513  ACTIV?
    S63  235735  CONTROL?
    S64    3531  ACTIV?()CONTROL?

?ss act
    S65    4525  ACT

?ss  activ?()control?()technol?
    S66  104513  ACTIV?
    S67  235735  CONTROL?
    S68  108774  TECHNOL?
    S69     231  ACTIV?()CONTROL?()TECHNOL?

?ss  activ?()control?()tech?
    S70  104513  ACTIV?
    S71  235735  CONTROL?
    S72  344509  TECH?
    S73     270  ACTIV?()CONTROL?()TECH?

?ss  flight?()control?()system?
    S74  110968  FLIGHT?
    S75  235735  CONTROL?
    S76  509309  SYSTEM?
    S77    2896  FLIGHT?()CONTROL?()SYSTEM?

?ss  control?()system?()design?
    S78  235735  CONTROL?
```

```
   S79 509309 SYSTEM?
   S80 260987 DESIGN?
   S81  10793 CONTROL?()SYSTEM?()DESIGN?

?ss aircraft?()control?
   S82 121142 AIRCRAFT?
   S83 235735 CONTROL?
   S84   6658 AIRCRAFT?()CONTROL?

?ss advanc?(2w)s82(advanc?(2w)s82)(5n)s54
   S85  47210 ADVANC?
       121142 S82
          125 S54
   S86      0 (ADVANC?(2W)S82)(5N)S54

?ss (robust(5n)multivari?)(7n)(control?()method?)
   S87   5109 ROBUST
   S88   5523 MULTIVARI?
   S89 235735 CONTROL?
   S90 351482 METHOD?
   S91      2 (ROBUST(5N)MULTIVARI?)(7N)(CONTROL?()METHOD?)

?ss (single or singular?)(w)(value?(w)base?)(7n)s77
   S92  86999 SINGLE
   S93  14469 SINGULAR?
   S94 145770 VALUE?
   S95 216862 BASE?
         2896 S77
   S96      2 (SINGLE OR SINGULAR?)(W)(VALUE?(W)BASE?)(7N)S77

?ss computer?()simulat?
   S97 199247 COMPUTER?
   S98 126899 SIMULAT?
   S99  43205 COMPUTER?()SIMULAT?

?ss aircraft?()guid?
   S100 121142 AIRCRAFT?
   S101  34575 GUID?
   S102   1478 AIRCRAFT?()GUID?

?ss stabiliz?
   S103  25301 STABILIZ?

?ss s50 or s54 or s58 or s61 or s64 or s65 or s73 or s77 or s81
    or s84
```

Combine synonyms for the flight control theory concept

```
          8968 S50
           125 S54
```

```
            5  S58
        10621  S61
         3531  S64
         4525  S65
          270  S73
         2896  S77
        10793  S81
         6658  S84
   S104  38491  S50 OR S54 OR S58 OR S61 OR S64 OR S65 OR S73
                OR S77 OR S81 OR S84

?ss s104 or s82 or s91 or s96 or s99 or s102 or s103

         38491  S104
        121142  S82
             2  S91
             2  S96
         43205  S99
          1478  S102
         25301  S103
   S105 201581  S104 OR S82 OR S91 OR S96 OR S99 OR S102 OR
                S103

?ss s105/de*
```

Sift through those citations in which these concepts were major descriptors

```
   S106  110943  S105/DE*

?ss s105/de
```

Sift through those citations in which these concepts were major or minor descriptors

```
   S108  161600  S105/DE

?ss (flex? or elast?)()(aircraft? or bod? or model? or airplane?)
```

Search on the flexible aircraft concept

```
   S109   27359  FLEX?
   S110   71257  ELAST?
   S111  121142  AIRCRAFT?
   S112   93722  BOD?
   S113  335200  MODEL?
   S114    6021  AIRPLANE?
   S115    8589  (FLEX? OR ELAST?)()(AIRCRAFT? OR BOD? OR MODEL?
                    OR AIRPLANE?)

?ss s115()(dynamic? or design? or handl? or model?)
          8589  S115
   S116 168075  DYNAMIC?
   S117 260987  DESIGN?
   S118  18085  HANDL?
```

```
  S119 335200 MODEL?
  S120     405 S115()(DYNAMIC? OR DESIGN? OR HANDL? OR MODEL?)

?ss s120()model?
           405 S120
  S121 335200 MODEL?
  S122     354 S120()MODEL?

?ss s116()s121
        168075 S116
        335200 S121
  S123   10541 S116()S121

?ss s116()system?
        168075 S116
  S124 509309 SYSTEM?
  S125    7219 S116()SYSTEM?

?ss transport?()aircraft?
  S126  66796 TRANSPORT?
  S127 121142 AIRCRAFT?
  S128    6061 TRANSPORT?()AIRCRAFT?

?ss s127()s121
        121142 S127
        335200 S121
  S129    2726 S127()S121

?ss control?()configure?()vehicle?
  S130 235735 CONTROL?
  S131  78429 CONFIGUR?
  S132  53481 VEHICLE?
  S133     600 CONTROL?()CONFIGUR?()VEHICLE?

?ss activ?()control?()tech?()s127
  S134 104513 ACTIV?
  S135 235735 CONTROL?
  S136 344509 TECH?
        121142 S127
  S137        2 ACTIV?()CONTROL?()TECH?()S127

?ss aeroelastic?
  S138   5395 AEROELASTIC?

?ss (structur? or s127 or s130)()stabili?
  S139 282772 STRUCTUR?
        121142 S127
        235735 S130
  S140 135626 STABILI?
```

```
S141   15126  (STRUCTUR? OR S127 OR S130)()STABILI?

?ss rigid?()structur?s139

S142  18975  RIGID?
     282772  S139
S143   4296  RIGID?()S139

?ss s127()configur?
     121142  S127
S144  78429  CONFIGUR?
S145   4300  S127()CONFIGUR?

?ss s115 or s120 or s123 or s125 or s128 or s129 or s133 or s137
    or s138
```

Combine synonyms for the flexible aircraft concept

```
       8589  S115
        405  S120
      10541  S123
       7219  S125
       6061  S128
       2726  S129
        600  S133
          2  S137
       5395  S138
S146  38615  S115 OR S120 OR S123 OR S125 OR S128 OR S129 OR
             S133 OR S137 OR S138

?ss s146 or s141 or s143 or s145
      38615  S146
      15126  S141
       4296  S143
       4300  S145
S147  58305  S146 OR S141 OR S143 OR S145

?ss s147/de*
```

Sift through those citations in which these concepts were major descriptors

```
S148   30559  S147/DE*

?ss s147/de
```

Sift through those citations in which these concepts were major or minor descriptors

```
S149   49829  S147/DE

?ss s46 and s106 and s148
```

Combine the separate concepts (modern control theory, flight control, and flexible *aircraft) for those citations in which the search terms were major descriptors*

```
        7280 S46
      110943 S106
       30559 S148
  S153   201 S46 AND S106 AND S148
```

?ss s47 and s108 and s149

Combine the separate concepts (modern control theory, flight control, and flexible aircraft) for those citations in which the search terms were minor descriptors

```
       12827 S47
      161600 S108
       49829 S149
  S154  1135 S47 AND S108 AND S149
```

?ss s45 and s105 and s147

Combine the separate concepts (modern control theory, flight control, and flexible aircraft) for those citations in which the search terms were in natural language

```
       14284 S45
      201581 S105
       58305 S147
  S155  1819 S45 AND S105 AND S147
```

?ss ud=09001:9212b2

Set the update for 1990-1992

```
  S156 205080 UD=9001:9212B2
```

?ss s153 and s156

Restrict the citations with major descriptors by update

```
         201 S153
      205080 S156
  S157    71 S153 AND S156
```

?ss s154 and s156

Restrict the citations with major or minor descriptors by update

```
        1135 S154
      205080 S156
  S158   357 S154 AND S156
```

?ss s155 and s156

Restrict the citations with natural language by update

```
        1819  S155
      205080  S156
   S159    498  S155 AND S156

?ss s158 not s157
```

Compared sets by NOTing to compare citations pulled in through major descriptors, minor descriptors, and natural language

```
         357  S158
          71  S157
   S160    286  S158 NOT S157

?ss s159 not (s157 or s160)
         498  S159
          71  S157
         286  S160
   S161    141 S159 NOT (S157 OR S160)

?logoff
```

Subject Three

DIALOG Recall Search Statement:
File 108:AEROSPACE 62-93/9306B2

```
    Set  Items  Description
    ---  -----  -----------
?ss computation?()grid?
                                        The computational gridding concept
    S1   79544  COMPUTATION?
    S2   21657  GRID?
    S3    8475  COMPUTATION?()GRID?

?ss adapt?()grid?
    S4   33301  ADAPT?
    S5   21657  GRID?
    S6     470  ADAPT?()GRID?

?ss mesh()redistrib?
    S7    6208  MESH
    S8    3424  REDISTRIB?
    S9       2  MESH()REDISTRIB?

?ss structur?()grid?
    S10 281815  STRUCTUR?
    S11  21657  GRID?
    S12    154  STRUCTUR?()GRID?

?ss grid?()generat?
    S13  21657  GRID?
    S14 121348  GENERAT?
    S15   2688  GRID?()GENERAT?

?ss computat?()fluid?()dynamic?
    S16  79574  COMPUTAT?
    S17  95669  FLUID?
    S18 167348  DYNAMIC?
    S19  25040  COMPUTAT?()FLUID?()DYNAMIC?

?ss adapt?()control?
    S20  33301  ADAPT?
    S21 235043  CONTROL?
    S22   8383  ADAPT?()CONTROL?

?ss flow()distrib?
    S23 235042  FLOW
    S24 272625  DISTRIB?
    S25  22463  FLOW()DISTRIB?
```

```
?ss s3 or s6 or s9 or s12 or s15 or s19 or s22 or s25
```
Combine similar terms for the computational gridding concept

```
        8475  S3
         470  S6
           2  S9
         154  S12
        2688  S15
       25040  S19
        8383  S22
       22463  S25
  S26  56686  S3 OR S6 OR S9 OR S12 OR S15 OR S19 OR S22 OR S25

?ss  finite()differen?()theor?
```
The finite difference and math theory concept

```
   S27   76274  FINITE
   S28  230475  DIFFEREN?
   S29  304131  THEOR?
   S30   15945  FINITE()DIFFEREN?()THEOR?

?ss  truncat?()error?
   S31    4976  TRUNCAT?
   S32   83056  ERROR?
   S33    1658  TRUNCAT?()ERROR?

?ss  numeric?()analy?
   S34  130784  NUMERIC?
   S35  546777  ANALY?
   S36   31774  NUMERIC?()ANALY?

?ss  math?()model?
   S37  166198  MATH?
   S38  333985  MODEL?
   S39   87417  MATH?()MODEL?

?ss coordinate? ?
   S40   33623  COORDINATE? ?

?ss coordinat?
   S41   37270  COORDINAT?

?ss  problem?()solv?
   S42  218104  PROBLEM?
   S43  391479  SOL?
   S44   10792  PROBLEM?()SOL?

?ss  precision?
```

```
   S45   13757 PRECISION?

?ss approximat?
   S46 106466 APPROXIMAT?

?ss finite()element?()method?
   S47   76274 FINITE
   S48 117011 ELEMENT?
   S49 350376 METHOD?
   S50   29009 FINITE()ELEMENT?()METHOD?

?ss s30 or s33 or s36 or s39 or s41 or s44 or s45 or s46 or s50
                                   Combine synonyms for the finite difference
                                   and math theory concept
        15945 S30
         1658 S33
        31774 S36
        87417 S39
        37270 S41
        10792 S44
        13757 S45
       106466 S46
        29009 S50
  S51 289158 S30 OR S33 OR S36 OR S39 OR S41 OR S44 OR S45 OR
                  S46 OR S50

?ss s26 and s51
                        Combine the two main concepts
        56686 S26
       289158 S51
  S52  20390 S26 AND S51

?ss ud=9001:9212b2
                        Set the update for 1900-1992
  S53 205080 UD=9001:9212B2

?ss s52 and s53
                        Limit the final set by the update
        20390 S52
       205080 S53
  S54   4447 S52 AND S53

?ss s54/de*
                                   Sift through the final set to see how many
                                   citations were pulled from major descriptors
  S57   3679 S54/DE*

?ss s54/de

                                   Sift through the final set to see how many
                                   citations were pulled from major or minor descriptors
```

```
   S58   4433  S54/DE

?ss s54 not s58
```

Number of records in the set retrieved with only title and abstract words, i.e. not descriptors. See also set 68.

```
        4447  S54
        4433  S58
   S59    14  S54 NOT S58

?ss s59 not s57
          14  S59
        3679  S57
   S60    14  S59 NOT S57

?ss s54 not (s41 or s44 or s45 or s46)
```

Remove terms that seem irrelevant to the query and caused major jumps in retrieval

```
        4447  S54
       37270  S41
       10792  S44
       13757  S45
      106466  S46
   S61  2549  S54 NOT (S41 OR S44 OR S45 OR S46)

?ss s61 not (s19 or s25)
        2549  S61
       25040  S19
       22463  S25
   S62   805  S61 NOT (S19 OR S25)

?ss s61/de*

   S63  2146  S61/DE*

?ss s61/de

   S64  2546  S61/DE

?ss s62/de*
```

Sift through the final set to see how many citations were pulled from major descriptors

```
   S65   643  S62/DE*
```

```
?ss s62/de
                                  Sift through the final set to see how many
                                  citations were pulled from major or minor descriptors
   S66     803  S62/DE

?ss s66 not s65

        803  S66                  Number of records pulled from minor descriptors only
        643  S65
   S67  160  S66 NOT S65

?ss s62 not (s65 or s66)
       805  S62                   Number of records retrieved with
       643  S65                   only title and abstract words, i.e. not descriptors
       803  S66
   S68   2  S62 NOT (S65 OR S66)

?ss logoff
```

Subject Four

DIALOG Recall Search Statement:
File 108:AEROSPACE 62-93/9306B2

```
    Set  Items  Description
    ---  -----  -----------

?ss unstead?
                                   Search the unsteady/turbulent concept
    S1   20243  UNSTEAD?

?ss unstead?()state?
    S2   20243  UNSTEAD?
    S3  153467  STATE?
    S4    1478  UNSTEAD?()STATE?

?ss computat?()fluid?()dynamic?
    S5   79574  COMPUTAT?
    S6   95669  FLUID?
    S7  167348  DYNAMIC?
    S8   25040  COMPUTAT?()FLUID?()DYNAMIC?

?ss turbulen?()flow?
    S9   70384  TURBULEN?
    S10 246544  FLOW?
    S11  21177  TURBULEN?()FLOW?

?ss unstead?()flow?
    S12  20243  UNSTEAD?
    S13 246544  FLOW?
    S14  10785  UNSTEAD?()FLOW?

?ss unstead?()aerodynamic?
    S15  20243  UNSTEAD?
    S16  57908  AERODYNAMIC?
    S17   2784  UNSTEAD?()AERODYNAMIC?

?ss turbulen?()wake?
    S18  70384  TURBULEN?
    S19  12792  WAKE?
    S20   2451  TURBULEN?()WAKE?

?ss turbulen?()model?
    S21  70384  TURBULEN?
    S22 333985  MODEL?
    S23   4584  TURBULEN?()MODEL?

?ss turbulen?
```

```
   S24   70384 TURBULEN?

?ss pressur?
   S25 149936 PRESSUR?

?ss unstead?()pressur?
   S26   20243 UNSTEAD?
   S27 149936 PRESSUR?
   S28      734 UNSTEAD?()PRESSUR?

?ss turbulen?()effect?
   S29   70384 TURBULEN?
   S30 454543 EFFECT?
   S31     4812 TURBULEN?()EFFECT?

?ss nonstation?()turbulen?
   S32    4751 NONSTATION?
   S33   70384 TURBULEN?
   S34        29 NONSTATION?()TURBULEN?

?ss nonstation?
   S35    4751 NONSTATION?

?ss turbulen?()boundar?()layer?
   S36   70384 TURBULEN?
   S37 135404 BOUNDAR?
   S38 121533 LAYER?
   S39   11472 TURBULEN?()BOUNDAR?()LAYER?

?ss s1 or s4 or s8 or s11 or s14 or s17 or s20 or s23 or s24 or
    s25 or s28
```

Combine terms for the unsteady/turbulent concept

```
          20243 S1
           1478 S4
          25040 S8
          21177 S11
          10785 S14
           2784 S17
           2451 S20
           4584 S23
          70384 S24
         149936 S25
            734 S28
   S40 232016 S1 OR S4 OR S8 OR S11 OR S14 OR S17 OR S20 OR
              S23 OR S24 OR S25 OR S28

?ss s40 or s31 or s34 or s35 or s39
         232016 S40
```

```
       4812  S31
         29  S34
       4751  S35
      11472  S39
S41  235453  S40 OR S31 OR S34 OR S35 OR S39
```

```
?ss incompressib?
```

Search the incompressible concept

```
S42   18388  INCOMPRESSIB?

?ss  compressib?()effect?
S43   15458  COMPRESSIB?
S44  454543  EFFECT?
S45    1437  COMPRESSIB?()EFFECT?

?ss  incompressib?()effect?
S46   18388  INCOMPRESSIB?
S47  454543  EFFECT?
S48       3  INCOMPRESSIB?()EFFECT?

?ss  incompressib?()flow?
S49   18388  INCOMPRESSIB?
S50  246544  FLOW?
S51    9317  INCOMPRESSIB?()FLOW?

?ss  high()reynold?()number?
S52  288546  HIGH
S53   27196  REYNOLD?
S54  138099  NUMBER?
S55    3039  HIGH()REYNOLD?()NUMBER?

?ss  incompressib?()boundar?()layer?
S56   18388  INCOMPRESSIB?
S57  135404  BOUNDAR?
S58  121533  LAYER?
S59     748  INCOMPRESSIB?()BOUNDAR?()LAYER?

?ss  incompressib?()fluid?
S60   18388  INCOMPRESSIB?
S61   95669  FLUID?
S62    5832  INCOMPRESSIB?()FLUID?

?ss s42 or s45 or s48 or s51 or s55 or s59 or s62
```

Combine synonyms for the incompressible concept

```
      18388  S42
       1437  S45
          3  S48
       9317  S51
```

```
          3039  S55
           748  S59
          5832  S62
   S63   22118  S42 OR S45 OR S48 OR S51 OR S55 OR S59 OR S62

?ss  bluff()bod?
```

Search the bluff bodies concept

```
   S64    1148  BLUFF
   S65   93453  BOD?
   S66    1039  BLUFF()BOD?

?e nonli
```

Expanding on the term nonlifting in the basic index

```
Ref   Items  Index-term
E1        1  NONLEVELING
E2        1  NONLEVITATED
E3        1  *NONLI
E4        1  NONLIDAR
E5        1  NONLIEAR
E6        1  NONLIENAR
E7        1  NONLIFE
E8        2  NONLIFT
E9      244  NONLIFTING
E10       8  NONLIFTING VEHICLE
E11       3  NONLIGHT
E12       1  NONLIGHTLY

?ss  nonlift?
   S67     246  NONLIFT?

?ss  non()lift?
   S68   36901  NON
   S69   19086  LIFT?
   S70      59  NON()LIFT?

?ss  bluff()bod?()flow?
   S71    1148  BLUFF
   S72   93453  BOD?
   S73  246544  FLOW?
   S74      36  BLUFF()BOD?()FLOW?

?ss  ground?()effect?
   S75   70051  GROUND?
   S76  454543  EFFECT?
   S77    3156  GROUND?()EFFECT?

?ss s66 or s67 or s70 or s74 or s77
```

```
          1039  S66
           246  S67
            59  S70
            36  S74
          3156  S77
   S78    4486  S66 OR S67 OR S70 OR S74 OR S77
```

Combine synonyms for the bluff body/ground effect concept

```
?ss s41 and s63 and s78

        235453  S41
         22118  S63
          4486  S78
   S79     165  S41 AND S63 AND S78
```

Combine the three concepts

```
?ss  ud=9001:9212b2

   S80  205080  UD=9001:9212B2
```

Set the update for 1990-1992

```
?ss s79 and s80

           165  S79
        205080  S80
   S81      26  S79 AND S80
```

Limit by the update

Due to small search set, decided to reexamine the relation of the concepts to each other. Current use of ANDing might be inappropriate due to the meaning of the terms and their relation to one another in the indexing language hierarchy.

```
?ss  finite  analy()analy?()method?
```

Unsure of the meaning of these terms and where they fit in the structure of the query so they were excluded (s85 and s89) from final sets

```
   S82   76274  FINITE
   S83  546777  ANALY?
   S84  350376  METHOD?
   S85      32  FINITE()ANALY?()METHOD?

?ss  finite()differen?()theor?
   S86   76274  FINITE
   S87  230475  DIFFEREN?
   S88  304131  THEOR?
   S89   15945  FINITE()DIFFEREN?()THEOR?

?ss s41 or s63
        235453  S41
```

```
        22118  S63
  S90  245019  S41 OR S63

?ss s90 and s78
```

Combine the unsteady concept and the bluff body concept

```
      245019  S90
        4486  S78
  S91   1665  S90 AND S78

?ss s90 or (s85 or s89 or (finite()element?)
```

Include finite element in the unsteady concept

```
       245019  S90
           32  S85
        15945  S89
  S92   76274  FINITE
  S93  117011  ELEMENT?
  S94   31200  FINITE(W)ELEMENT?
  S95  278359  S90 OR S85 OR S89 OR (FINITE()ELEMENT?)

?ss s95 and s78
```

Combine the unsteady concept and bluff body concept

```
      278359  S95
        4486  S78
  S96   1697  S95 AND S78

?ss s80 and s96
```

Limit by the update

```
      205080  S80
        1697  S96
  S97    241  S80 AND S96

?ss s85 or s89 or s94
```

Combine finite difference, finite element and finite analysis

```
           32  S85
        15945  S89
        31200  S94
  S98   45673  S85 OR S89 OR S94

?ss s90 and s98 and s78
```

Combine the unsteady concept, the finite difference/element/analysis concept, and the bluff body concept

```
      245019  S90
       45673  S98
        4486  S78
```

```
S99      95  S90 AND S98 AND S78

?ss s99 and s80
                                              Limit by the update
         95  S99
     205080  S80
S100     15  S99 AND S80

?ss s97/de*
                                   Sift for major descriptors
S101    190  S97/DE*

?ss s97/de
                                   Sift for major or minor descriptors
S102    234  S97/DE

?ss s97 not s101
                                             The NOTed sets are checking for duplicates.
                                             Used set 97 as the final set.
        241  S97
        190  S101
S103     51  S97 NOT S101

?ss s101 not s102
        190  S101
        234  S102
S104      0  S101 NOT S102

?ss s102 not s101
        234  S102
        190  S101
S105     44  S102 NOT S101

?ss s97 not (s101 or s102)
        241  S97
        190  S101
        234  S102

S106      7  S97 NOT (S101 OR S102)

?logoff
```

Appendix B. BBS Precision Search Strategies

QUESTION 1

```
   Set  Items  Description
   ---  -----  -----------
?ss aerodynamic?()stall?/de
   S1  51164  AERODYNAMIC?/DE
   S2   2527  STALL?/DE
   S3   1907  AERODYNAMIC?()STALL?/DE

?ss stall?/de
   S4   2527  STALL?/DE

?ss s3 or s4
        1907  S3
        2527  S4
   S5   2527  S3 OR S4

?ss computer()program?de
   S6 202425  COMPUTER?
   S7      0  PROGRAM?DE
   S8      0  COMPUTER?()PROGRAM?DE

?ss computer?/()program?/de
   S9 177936  COMPUTER?/DE
   S10 125934  PROGRAM?/DE
   S11  66375  COMPUTER?()PROGRAM?/DE

?ss math()?()model?/de
   S12 154874  MATH?/DE
   S13 177623  MODEL?/DE
   S14  85604  MATH?()MODEL?/DE

?ss s11 or s14
        66375  S11
        85604  S14
   S15 144914  S11 OR S14

?ss turbulen?/de
   S16  62291  TURBULEN?/DE

?ss boundar?()layer?/de
   S17 102138  BOUNDAR?/DE
   S18  66592  LAYER?/DE
   S19  46540  BOUNDAR?()LAYER?/DE

?ss unstead?/de
   S20  13212  UNSTEAD?/DE

?ss s16 or s19 or s20
```

```
      62291  S16
      46540  S19
      13212  S20
  S21  101505  S16 OR S19 OR S20

?ss wind?/de
  S22   83479  WIND?/DE

?ss turbine?/de
  S23   26313  TURBINE?/DE

?ss rotor?/de
  S24   14745  ROTOR?/DE

?ss s22 or s23 or s24
      83479  S22
      26313  S23
      14745  S24
  S25  117748  S22 OR S23 OR S24

?c 5 and 15 and 21 and 25
      2527  5
     144914  15
     101505  21
     117748  25
  S26      24  5 AND 15 AND 21 AND 25

?ss ud=9001:9212b2
  S27  205080  UD=9001:9212B2

?c 26 and 27
        24  26
     205080  27
  S28       3  26 AND 27

?ss s21 or flow?
     101505  S21
  S29  251436  FLOW?
  S30  282542  S21 OR FLOW?

?ss s21 or flow?/de
     101505  S21
  S31  205276  FLOW?/DE
  S32  239863  S21 OR FLOW?/DE

?c 5 and 15 and 32 and 25
      2527  5
     144914  15
     239863  32
```

```
        117748  25
    S33      52  5 AND 15 AND 32 AND 25

 ?c 27 and 33
        205080  27
            52  33
    S34       8  27 AND 33

 ?ss  s25  or  wind?/ti
        117748  S25
    S35   32419  WIND?/TI
    S36  119527  S25 OR WIND?/TI

 ?ss  s32  or  (turbulen?/ti  or  unstead?/ti)
        239863  S32
    S37   33702  TURBULEN?/TI
    S38    8260  UNSTEAD?/TI
    S39  241305  S32 OR TURBULEN?/TI OR UNSTEAD?/TI

 ?c 5 and 15 and 39 and 36
          2527  5
        144914  15
        241305  39
        119527  36
    S40      54  5 AND 15 AND 39 AND 36

 ?ss  s3  or  stall?()delay?
          1907  S3
    S41    4156  STALL?
    S42   18433  DELAY?
    S43      26  STALL?(W)DELAY?
    S44    1913  S3 OR STALL?()DELAY?

 ?c 44 and 15 and 39 and 36
          1913  44
        144914  15
        241305  39
        119527  36
    S45      42  44 AND 15 AND 39 AND 36

 ?c 27 and 45
        205080  27
            42  45
    S46       6  27 AND 45

 ?ss  aerodynamic?/de*  or  stall?/de
    S47   31838  AERODYNAMIC?/DE*
    S48    2527  STALL?/DE
    S49   32857  AERODYNAMIC?/DE* OR STALL?/DE
```

```
?c 49 or 15 or 39 or 36
       32857  49
      144914  15
      241305  39
      119527  36
   S50 444822  49 OR 15 OR 39 OR 36

?ss s5 or stall?()delay?
        2527  S5
   S51   4156  STALL?
   S52  18433  DELAY?
   S53     26  STALL?(W)DELAY?
   S54   2531  S5 OR STALL?()DELAY?

?c 54 and 15 and 39 and 36
        2531  54
      144914  15
      241305  39
      119527  36
   S55     54  54 AND 15 AND 39 AND 36

?c 27 and 55
      205080  27
          54  55
   S56      8  27 AND 55
```

QUESTION 2

```
     Set  Items  Description
     ---  -----  -----------
?ss ud=9001:9212b2
      S1  205080  UD=9001:9212B2

?ss modern()control()theor?
    S2    9259 MODERN
    S3  206837 CONTROL
    S4  306568 THEOR?
    S5     239 MODERN()CONTROL()THEOR?

?ss h()2
    S6   40803 H
    S7  197069 2
    S8     449 H()2

?ss h()infinit?
    S9   40803 H
    S10  23500 INFINIT?
    S11    653 H()INFINIT?

?ss s5 or s8 or s11
           239 S5
           449 S8
           653 S11
   S12    1325 S5 OR S8 OR S11

?ss robustness
   S13    5387 ROBUSTNESS

?ss robust?(3w)control
   S14    8098 ROBUST?
   S15  206837 CONTROL
   S16    1112 ROBUST?(3W)CONTROL

?ss flight()control()system?
   S17  107537 FLIGHT
   S18  206837 CONTROL
   S19  513422 SYSTEM?
   S20    2912 FLIGHT()CONTROL()SYSTEM?

?ss multivariable control
   S21     169 MULTIVARIABLE CONTROL

?ss (control()synthesis)/ti
   S22   61577 CONTROL/TI
```

```
    S23  10881  SYNTHESIS/TI
    S24    212  (CONTROL()SYNTHESIS)/TI

?ss (feedback control)/de*
    S25   8826  (FEEDBACK CONTROL)/DE*

?ss feedback control
    S26  15751  FEEDBACK CONTROL

?ss multivariable()feedback()control()system
    S27   3069  MULTIVARIABLE
    S28  26930  FEEDBACK
    S29 206837  CONTROL
    S30 340935  SYSTEM
    S31     12  MULTIVARIABLE()FEEDBACK()CONTROL()SYSTEM

?ss active()modal()suppression()system
    S32  34732  ACTIVE
    S33   9402  MODAL
    S34   6541  SUPPRESSION
    S35 340935  SYSTEM
    S36      1  ACTIVE()MODAL()SUPPRESSION()SYSTEM

?ss s20 or s13 or s21 or s16 or s24
          2912  S20
          5387  S13
           169  S21
          1112  S16
           212  S24
    S37   8648  S20 OR S13 OR S21 OR S16 OR S24

?ss s37 or s31 or s36
          8648  S37
            12  S31
             1  S36
    S38   8657  S37 OR S31 OR S36

?ss transport aircraft
    S39   4754  TRANSPORT AIRCRAFT

?e flexible aircraft
Ref   Items  Index-term
E1        2  FLEXIBITY
E2    14693  FLEXIBLE
E3        0  *FLEXIBLE AIRCRAFT
E4     3318  FLEXIBLE BODIES
E5       98  FLEXIBLE BODY
E6     1777  FLEXIBLE SPACECRAFT
E7       45  FLEXIBLE WING
```

```
E8    297 FLEXIBLE WINGS
E9     27 FLEXIBLEN
E10     4 FLEXIBLER
E11    20 FLEXIBLES
E12     2 FLEXIBLILITY

?ss (flexible()bod? or wing?)/de
    S40   5622 FLEXIBLE/DE
    S41  61501 BOD?/DE
    S42   3416 FLEXIBLE/DE(W)BOD?/DE
    S43  25162 WING?/DE
    S44  28500 (FLEXIBLE()BOD? OR WING?)/DE

?ss (flexible()(bod? or wing?))/de
    S45   5622 FLEXIBLE/DE
    S46  61501 BOD?/DE
    S47  25162 WING?/DE
    S48   3754 (FLEXIBLE()(BOD? OR WING?))/DE

?ss x()29
    S49  95226 X
    S50   6525 29
    S51    296 X()29

?ss lynx()helicopter?
   S52    146 LYNX
   S53  16531 HELICOPTER?
   S54     67 LYNX()HELICOPTER?

?ss maneuverable()aircraft()design?
   S55    692 MANEUVERABLE
   S56 121850 AIRCRAFT
   S57 263297 DESIGN?
   S58      2 MANEUVERABLE()AIRCRAFT()DESIGN?

?ss s39 or s48 or s51 or s54 or s58
         4754 S39
         3754 S48
          296 S51
           67 S54
            2 S58
   S59   8853 S39 OR S48 OR S51 OR S54 OR S58

?c 12 and 38 and 59
         1325 12
         8657 38
         8853 59
   S60     31 12 AND 38 AND 59
```

```
?c 60 and 1
         31  60
       205080  1
   S61      25  60 AND 1

?ss  s12  or  (control()theory)/de*
          1325  S12
   S62  105897  CONTROL/DE*
   S63   68393  THEORY/DE*
   S64    7289  CONTROL/DE*(W)THEORY/DE*
   S65    8339  S12 OR (CONTROL()THEORY)/DE*

?c 65 and 38 and 59
          8339  65
          8657  38
          8853  59
   S66      65  65 AND 38 AND 59

?c 1 and 66
       205080  1
           65  66
   S67      42  1 AND 66

?ss s67 not s61
            42  S67
            25  S61
   S68      17  S67 NOT S61
```

QUESTION 3

```
     Set Items Description
         ---  -----  -----------
 ?ss  grid?()generat?/de*
     S1    4795  GRID?/DE*
     S2   22734  GENERAT?/DE*
     S3    1243  GRID?()GENERAT?/DE*

 ?ss  finit?()different?()theor?/de
     S4   47451  FINIT?/DE
     S5   67381  DIFFEREN?/DE
     S6  154251  THEOR?/DE
     S7   16218  FINIT?()DIFFEREN?()THEOR?/DE

 ?ss  finit?()element?()method?/de
     S8   47451  FINIT?/DE
     S9   54517  ELEMENT?/DE
    S10  105570  METHOD?/DE
    S11   29106  FINIT?()ELEMENT?()METHOD?/DE

 ?ss s7 or s11
        16218  S7
        29106  S11
    S12   43955  S7 OR S11

 ?ss s3 and s12
        1243  S3
        43955  S12
    S13    360  S3 AND S12

 ?ss  ud=9001:9212b2
     S14   205080  UD=9001:9212B2

 ?ss s14 and s13
      205080  S14
         360 S13
    S15  236  S14 AND S13

 ?ss  s7/de*
    S16    6363  S7/DE*

 ?ss  s11/de*
 S17    16425  S11/DE*


 ?ss s16 or s17
       6363 S16
      16425 S17
 S18 22413 S16 OR S17
```

```
?ss s3 and s16
        6363 S16
        1243 S3
     S20  52 S3 AND S16


?ss s14 and s20
      205080 S14
          52 S20
     S21   37  S20 AND S14

?ss  s21  not  computat?()fluid?()dynamic?
        37  S21
  S22  82336  COMPUTAT?
  S23   97938  FLUID?
  S24  171698  DYNAMIC?
  S25   26184  COMPUTAT?(W)FLUID?(W)DYNAMIC?
  S26      23  S21 NOT COMPUTAT?()FLUID?()DYNAMIC?
```

QUESTION 4

```
?ss flow? or wake? or turbulen? or stabilit?
     S2  248235  FLOW?
     S3   12907  WAKE?
     S4   70918  TURBULEN?
     S5  136185  STABILI?
     S6  383912  FLOW? OR WAKE? OR TURBULEN? OR STABILI?

?ss  s6/ti
     S7  146622  S6/TI

?ss  bluff()bod?/ti
     S8     453  BLUFF/TI
     S9   24657  BOD?/TI
     S10    387  BLUFF()BOD?/TI

?ss s7 and s10
         146622  S7
            387  S10
     S11    286  S7 AND S10

?ss  ud=9001:9212b2
     S13  205080  UD=9001:9212B2

?ss s11 and s13
            286  S11
         205080  S13
     S14     38  S11 AND S13

?ss computat? or computational () fluid ()
    dynamic? or reynold's () number
  S15   80575  COMPUTAT?
  S16   52140  COMPUTATIONAL
  S17   81766  FLUID
  S18  169087  DYNAMIC?
  S19   25391  COMPUTATIONAL(W)FLUID(W)DYNAMIC?
  S20       0  REYNOLD'S
  S21  127151  NUMBER
  S22       0  REYNOLD'S(W)NUMBER
  S23   80575  COMPUTAT? OR COMPUTATIONAL()FLUID()DYNAMIC? OR
               REYNOLD'S()NUMBER

? e reynold
E1        1  REYNODS
E2        4  REYNOL
E3      174  *REYNOLD
E4    27314  REYNOLDS
E5      786  REYNOLDS EQUATION
E6    15690  REYNOLDS NUMBER
E7       75  REYNOLDS NUMBER EFFECT
```

```
E8   2540 REYNOLDS STRESS
E9      2 REYNOLDSA
E10     1 REYNOLDSD
E11     2 REYNOLDSOVA
E12     1 REYNOLDSOVO

?ss e4,e7
   S24  27314 REYNOLDS
   S25     75 REYNOLDS NUMBER EFFECT
   S26  27314 E4,E7

?ss s15 or (s15 or s19 or s26)/ti
   S27  12189 S15/TI
   S28    529 S19/TI
   S29   3926 S26/TI
   S30  16054 (S15 OR S19 OR S26)/TI

?ss s30 and s13
        16054 S30
       205080 S13
   S31   2449 S30 AND S13

?ss s7 and s31
       146622 S7
         2449 S31
   S32    824 S7 AND S31

?ss s32 and s13
          824 S32
       205080 S13
   S33    824 S32 AND S13
```

Appendix C. WAIS Search Strategies

WAIS Terse Strategies

QUESTION 1

unsteady aerodynamics stall delay model air flows propeller

QUESTION 2

flight control system modern control theory flexible aircraft

QUESTION 3

adaptive gridding grid generation mesh redistribution truncation error structured flow mechanics

QUESTION 4

computational fluid dynamics bluff body

WAIS Verbose Strategies

QUESTION 1

unsteady state flow aerodynamics pressure turbulence flow wake model effects boundary layers mathematical models computer models digital simulation high solidity propellers compressed rotor stator blades wind power generator

QUESTION 2

modern control theory flight control design flight control active control h-infinity h-2 control system design flexible aircraft body model dynamic model transport aircraft control configured vehicle"

QUESTION 3

grid adaptive gridding mesh redistribution structural gridding grid generation computational fluid dynamics adaptive control flow distribution finite difference theory truncation error numeric analysis numerical analysis mathematical models math model mathematical model coordinate coordinates coordination coordinating problem solving approximate approximation approximating finite elemental method finite element method

QUESTION 4

unsteady state computational fluid dynamics turbulence flow turbulent flow unsteady flow unsteady aerodynamics turbulent wake turbulent wakes turbulent model turbulence pressure unsteady pressure turbulence effect nonstationary pressure nonstationary turbulence turbulence boundary layer incompressible compressible effect incompressible flow high reynolds number incompressible boundary layer incompressible fluid bluff bodies nonlifting non-lifting bluff body flow ground effect finite analysis method finite difference theory

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (leave blank)	2. REPORT DATE March 1994	3. REPORT TYPE AND DATES COVERED Contractor Report
4. TITLE AND SUBTITLE A Comparison of Boolean-based Retrieval to the WAIS System for Retrieval of Aeronautical Information: Final Report		5. FUNDING NUMBERS NASw-4584
6. AUTHOR(S) Gary Marchionini and Diane Barlow		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Maryland College Park, MD *Under contract to RMS Associates*		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546		10. SPONSORING/MONITORING AGENCY REPORT NUMBER NASA CR-4569
11. SUPPLEMENTARY NOTES		
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified - Unlimited Subject Category 82		12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

An evaluation of an information retrieval system using a Boolean-based retrieval engine and inverted file architecture and WAIS, which uses a vector-based engine, was conducted. Four research questions in aeronautical engineering were used to retrieve sets of citations from the NASA Aerospace Database which was mounted on a WAIS server and available through Dialog File 108 which served as the Boolean-based system (BBS). High recall and high precision searches were done in the BBS and terse and verbose queries were used in the WAIS condition. Precision values for the WAIS searches were consistently above the precision values for high recall BBS searches and consistently below the precision values for high precision BBS searches. Terse WAIS queries gave somewhat better precision performance than verbose WAIS queries. In every case, a small number of relevant documents retrieved by one system were not retrieved by the other, indicating the incomplete nature of the results from either retrieval system. Relevant documents in the WAIS searches were found to be randomly distributed in the retrieved sets rather than distributed by ranks. Advantages and limitations of both types of systems are discussed.

14. SUBJECT TERMS	Information retrieval, architecture (computers), aeronautical engineering, search profiles, information systems, precision		15. NUMBER OF PAGES 74
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclass	18. SECURITY CLASSIFICATION OF THIS PAGE Unclass	19.SECURITY CLASSIFICATION OF ABSTRACT Unclass	20. LIMITATION OF ABSTRACT Unlimited

Available from NASA Center for AeroSpace Information
800 Elkridge Landing Road
Linthicum Heights, MD 21090-2934
(301) 621-0390