I. THE SMART SYSTEM – TYPICAL PROCESSING SEQUENCES

Michael Lesk

In this section some output of the SMART retrieval system is presented to show the flexibility and potential of the system. One request is processed through seven of the processing methods available for retrieval, to illustrate the SMART text analysis and search techniques.

The request, named DIFFERNTL EQ, reads:

> "Give algorithms useful for the numerical solution of ordinary differential equations and partial differential equations on digital computers. Evaluate the various integration procedures (try Runge-Kutta, Milne's method) with respect to accuracy, stability, and speed."

The document collection being searched consists of the 405 abstracts published by the Professional Group on Electronic Computers in the <u>IRE Transactions on Electronic Computers</u> for March, June, and September of 1959. This collection covers all fields of the computer literature. Thirty-two of the 404 abstracts are classified by the IRE as: mathematics-logic; numerical analysis. Sixteen abstracts are believed to be relevant to the request DIFFERNTL EQ.

The first of the seven methods used to compare this request with the reference collection requires the use of the "null thesaurus"; this thesaurus maps each word stem into one distinct concept by a dictionary lookup operation. The request concepts are then matched against document concepts which are obtained in the same way. Thus, only directly matching word stems are taken into account with the null thesaurus procedure.

A more general method uses the regular "Harris" thesaurus, which defines a many-to-many mapping of words to concepts. Here, a request term can be made to match any document term which maps into any of the concepts corresponding to the original word in the request. Synonyms may be mapped into the same concept category, so that a match is obtained even if different word stems are involved. For example, in the regular thesaurus "console" and "panel" are both mapped into concept 43, since these words are considered synonymous with respect to the IRE abstracts. Ambiguous words can be mapped into several different concepts. A typical ambiguous word, "field," is included under concept 74 (topic), 89 (group, ring), 352 (land, terrain), and 370 (flux, intensity). The frequency of occurrence of each concept is counted in both the null and the regular thesaurus methods. In the regular thesaurus, the weighting of ambiguous concepts is lowered.

Other options may, respectively, ignore the concept weighting schemes, process only the titles of each text, or use an expansion through a hierarchy of concept numbers which is constructed to operate in conjunction with the regular dictionary. The hierarchy provides for each concept a parent concept, a set of sons and brothers, and a list of cross references. Two phrase searching methods are also available; one is a statistical method using co-occurrences of concepts within the sentences of a document to define a phrase. The other uses a complete syntactic analysis as part of a phrase searching device.

Each SMART processing run requires as basic input a set of processing instructions, a set of requests, and a reference document collection. The set of short form processing instructions used to specify a syntactic phrase searching run is shown in Fig. 1.† Figure 2 contains an expanded version of the same specifications, prepared by the computer. A complete list of specifications interpretable by SMART is included in Sec. II of this report as part of the operating instructions. The specifications shown in Fig. 1 have the following meanings:

ANSWER REQUESTS: A list of documents to which the computer assigns a high correlation with the submitted request is prepared, and the list is printed.

FORMAT JUMBO: The printout of the answers consists of a detailed list of the citations. (See Fig. 22.)

SCORES YES: A list of manually prepared relevance judgments is introduced and compared with the list of answers generated by the SMART system. (See Figs. 27-33.)

THESHR 2: The thesaurus used is the regular (Harris) thesaurus.

MAXCON 511: There are 511 concepts included in the thesaurus.

---

†The figures appear in the Appendix to this section.

CORMD3 COS: The cosine correlation is to be used to compare request vectors with document vectors (see Sec. II of this report).

CUTOF3 3500: The cutoff for determining what documents are to be printed as answers to a given request is a correlation greater than or equal to 0.35.

EXECUTE SYNTAX: The English input text is to be syntactically analyzed, and syntactic phrases are to be identified by means of the criterion tree procedure. See Information Storage and Retrieval, Report No. ISR-7, Sec. VI.

SYNTACTIC ANALYSIS: A syntactic analysis is printed for each sentence of the input text.

CRITERION TREES: The criterion trees which match phrases included in the input text are printed.

NODE CORRESPONDENCES: The correspondences between matching criterion tree nodes and sentence nodes are printed for all detected criterion trees.

CRITES 300: Criterion trees are weighted three times as heavily as word stems.

ENGLISH TEXTS: The texts submitted as input are printed.

WORDS NOT FOUND: Text words not found in the thesaurus during the dictionary lookup are printed.

REQUEST CORRELATIONS: The correlations of each request with each document are printed.

Two SMART instruction cards are also shown in Fig. 1. *TIME requests a printout of the current time; the *FIND card labels the document DIFFERNTL EQ as a search request. This *FIND card also stipulates that the request is being introduced as a BCD-punched English text.

The results of the dictionary lookup for this request are shown in Fig. 3. First the text of the request is printed; thereafter, the request words not found in the dictionary are shown. In this case, only the word "respect" is missing from the thesaurus.

Figure 4 shows the complete texts of two of the documents included in the reference collection.

Since in the example under discussion, the EXECUTE SYNTAX specification is used to process the sample request, a syntactic analysis is performed using the Kuno syntactic analyzer. The complete homograph codes for the words in the first sentence of the search request are shown in Fig. 5. These codes are obtained by combining the result of a stem lookup with that of a suffix lookup. For example, ALGORITHM is coded "NOU" (a partial homograph indicating a noun of indefinite number), and the S detected in the suffix list changes this to NOUP (plural noun). USEFUL, with the stem USE can be a noun, verb, adjective, or adverb; the output code correctly restricts the correspondences to an adjective code, using the suffix FUL.

The actual analysis produced by the analyzer is shown in Fig. 6. This analysis is not in fact the semantically accurate analysis for the sentence, since "partial differential equations" is identified as the object of "for" rather than "of." However, the phrase "differential equations" included in the sentence is analyzed correctly.

Figure 7 shows the results of the criterion tree search for this sentence. Two trees, "numerical solution" and "differential equations" are identified. The correspondences between nodes of the criterion trees and nodes in the sentence are shown for each tree.

Figure 8 shows a summary of the results of the syntactic phrase searching. The trees found are listed together with the included concept numbers.

The method used to introduce into the machine the preprocessed documents is illustrated in Fig. 9. A preprocessed request is preceded by a *LIST DIFFERNTL EQ card and followed by a *LIKE DIFFERNTL EQ card (to mark it as a request). The documents in the reference collection, previously looked up in the thesaurus, are also preceded by *LIST cards; no *LIKE cards are used for any of these documents since they do not function as requests. The *TAPE control card indicates that these documents are read from a separate tape.

Figure 10 shows the list of character identifiers for all documents in the reference collection, including the request.

Some typical concept-vectors resulting from the automatic analysis procedure are shown in Figs. 11, 12, and 13. The concept vector in Fig. 11 results from a null thesaurus (word stem) lookup; Fig. 12 is similarly produced by a lookup in the regular Harris 2 dictionary, and Fig. 13 shows a lookup using the statistical phrase dictionary. Each vector element is shown as a six-character alphabetic identifier followed by a numeric weight. In the null thesaurus concept vectors, the identifiers are represented by the first six characters of the word stem that maps into the concept. In the regular thesaurus, the numeric characters included in the identifier represent the concept number of the category; the alphabetic characters are mnemonic identifiers of the corresponding category. In the request on differential equations, for example, the word "numerical" is mapped into the concept "NUMER" through the null dictionary; the only processing used to generate the identifier is the suffix analysis. The regular thesaurus, however, maps the word "numerical" into concept 13. This concept represents words such as "numerical," "calculate" (whence the identifier 13CALC), "interpolate," and others. The statistical phrase search adds concept 375, among others, representing the phrase "numerical analysis" or "numerical solution." An example of the ambiguities which may be introduced through the thesaurus procedure is illustrated by the word "stability"; this word produces an entry in the concept vector representing "yaw," "pitch," and "roll."

Figure 14 shows the correlation list obtained by matching the sample request with the analyzed document collection; the correlations are ordered by increasing document number for both the Harris 2 and null thesaurus procedures. To save space, only nonzero correlations are printed. If the null thesaurus is used, the request can overlap only with documents that contain word stems identical with the word stems included in the request. For the Harris thesaurus, however, the requirements are not so strict; words may overlap with different words provided that they correspond to the same thesaurus category.

Thus, every document that exhibits a nonzero correlation with DIFFERNTL EQ in the null thesaurus will also have a nonzero correlation in the Harris thesaurus: some documents with nonzero correlations in the Harris thesaurus will however have zero correlations in the null thesaurus. In Fig. 14, for example, documents 3, 21, 22, 23, and 58 exhibit some overlap with the search request in the Harris thesaurus, but none in the null thesaurus. Documents with nonzero correlations using the null thesaurus may, of course, exhibit a different nonzero correlation coefficient using the regular thesaurus; note for example, the rise in the correlation of document 45 from 0.1525 in the null thesaurus to 0.2598 in the regular thesaurus (unlike the previously mentioned changes, this difference is large enough to be significant). Abstract 45, "A Calculator for Numerical Fourier Syntheses," although not directly relevant to the request, is in fact somewhat more relevant than most other documents included in Fig. 14.

Figure 15 shows the correlation list for the request DIFFERNTL EQ sorted in decreasing correlation order for both the null and Harris runs. Note that the average correlation near the top of the list is <u>higher</u> for the regular thesaurus than for the null thesaurus run. This reflects two influences: first, the null thesaurus includes more concepts than the regular thesaurus, so that two word stems are less likely to be mapped into the same concept; second, the Harris thesaurus permits overlap of nonidentical terms. Most documents near the top of the correlation list in the null thesaurus have higher correlations in the regular thesaurus, even though they may appear further down (with higher ranks) in the regular thesaurus than in the null thesaurus.

To simplify the study of the distribution of correlations, the SMART system prints a histogram, shown in Fig. 16. This list shows the number of documents with correlations greater than a number of specified argument values. It may be seen from the histogram, that there are more high correlation values for the regular thesaurus run than for the null thesaurus run. For example, the regular thesaurus exhibits six correlations above 0.5, whereas there are only three in the null thesaurus; similarly, eight correlations above 0.4 are included for the regular thesaurus, while the null thesaurus has only six correlations above 0.4.

All of the output included in Figs. 14, 15, 16 is controlled by the REQUEST CORRELATIONS specification.

The effects of using the hierarchy to alter a request are shown in Figs. 17-20. Figure 17 shows the request vector both before and after hierarchical expansion. In the example shown, the hierarchical parent of each concept in the request is added to the request vector. For example, concept 219 (program) is obtained from concept 8 (algorithms), and is given the weight originally assigned to concept 8. Similarly, concept 13 (calculation) is promoted to 375 (numerical analysis), and so on.

In Fig. 18, the correlation list obtained through hierarchical processing is shown in increasing document order. Since the hierarchical expansion adds concepts only to the concept vector of the search request, any document with a nonzero correlation in the regular thesaurus before expansion still exhibits a nonzero correlation after the expansion. Two documents (40 and 51), however, exhibit nonzero correlations after expansion, even though they had a zero correlation before the hierarchical process.

After all retrieval operations are effected, SMART can perform an automatic evaluation of the results by comparing the computer generated rank orders of relevant documents against human relevance judgments. The corresponding output appears in Figs. 27-33 for the seven runs used in conjunction with the request DIFFERNTL EQ. On the left side of these figures the top 15 documents are shown in decreasing correlation order. The documents judged relevant by a human relevance judgment are also listed, and their rank order in the correlation list is given. Various measures of retrieval effectiveness, discussed in Secs. III and IV of this report, are calculated to aid in evaluating the success of the retrieval procedure.

The computer methods illustrated are the following (all runs use the cosine correlation procedure):

Figure 27: Regular thesaurus lookup (Harris 2 dictionary).

Figure 28: Null thesaurus lookup.

Figure 29: Regular thesaurus lookup, only titles used.

Figure 30: Regular thesaurus lookup, logical vectors (no weighting).

Figure 31: Regular thesaurus lookup, hierarchy expansion by parents.

Figure 32: Regular thesaurus lookup, statistical phrase search with phrases weighted three times as much as word stems.

Figure 33: Regular thesaurus lookup, syntactic phrase searching with phrases weighted three times as much as word stems.

Figure 19 shows the correlation lists in correlation order both before and after expansion. Since the request vector contains more terms after expansion, it becomes more difficult for any document to match a large fraction of the request. The highest correlation coefficients shown after expansion are therefore lower than the highest correlations in the unexpanded output. The maximum correlation declined from 0.67 before expansion to 0.46 after expansion, and the histogram shown in Fig. 20 shows that the number of documents with correlations above 0.34 is halved by the hierarchical expansion (14 before, seven after).

At the same time, the larger size of the request vector after expansion increases the number of documents that have a small amount of overlap with the request. Thus there are more documents with low correlations after expansion than before expansion. For example, before hierarchical expansion 87 documents had a correlation with DIFFERNTL EQ equal to or greater than 0.16; after expansion 102 documents showed a correlation of 0.16 or better.

Figure 21 shows the processing specifications for the Harris 2 computer run. The answers obtained for the request DIFFERNTL EQ are shown for this method in Fig. 22. In Fig. 22, the documents are identified by full journal citation. A more compact output format is represented in Figs. 23-24. Figure 23 gives the specifications for the statistical phrase run. In Fig. 24, the answers are shown, identified only by the first line of their title. Finally, the most compact output format is illustrated in Fig. 25. The answers for seven processing methods are reproduced, identified only by the 12-character document abbreviations of the output documents.

Figure 26 displays the short form output both before and after the hierarchical expansion. Since the correlations are lower after the expansion and the cutoff is unchanged, fewer documents are retrieved. It may be seen that the hierarchical expansion effectively rearranges the order of the retrieved documents. Examination of the rank orders of the relevant documents (see Figs. 27 and 31) shows that the hierarchy has been successful in promoting several relevant documents.

The following interesting features of this output may be noted in particular: two of the methods employed disregard information otherwise retained. These are the "Titles Only" processing, which discards information obtained from the body of the text, and the logical vector procedure, which discards the weights of the elements of the concept vectors. Both of these methods, as expected, perform considerably less well than more complete processing systems. For example, relevant document 388 is completely missed by the "Titles Only" method, where it receives a correlation of 0.00. The title of document 388 ("A Method of

Normalized Block Iteration") contains no specific mention of differential equations, or of any of the other terms in the request, and as a result a correlation of zero with the request DIFFERNTL EQ is obtained.

The logical vectors processing also discards information; ignored, for example, is the fact that the request mentions "differential equations" twice, while "evaluate" is only mentioned once. This option also produces distinctly poorer recall and precision than the same run performed with numeric vectors. These conclusions, have, of course, been verified by using additional requests (see Secs. III and IV of this report).

The alteration of search requests by a hierarchical expansion provides generally poor system performance; there exist however, some exceptions: document 387, which is relevant, is promoted from eleventh place to ninth place, for example. The hierarchical expansion broadens the request, and loses some of the precision of the original search text. The accuracy of the retrieval is thereby often reduced. In some cases (for example, for poorly phrased requests where concepts needed for the search are omitted from the original text), the hierarchy does perform relatively well compared with a direct search.

The use of the null thesaurus has mixed effects. By allowing a match between concepts only in cases of precise agreement of English word stems, it becomes more difficult to retrieve unwanted documents. Thus, for the request DIFFERNTL EQ, the first eight, as well as thirteen of the first fifteen, documents found are relevant. The strict matching criteria also push two relevant documents far down the rank list (below, for example, the rank of any relevant document in the regular thesaurus processing). The precision of the null thesaurus methods is frequently comparable to that obtained with the regular thesaurus; but the recall produced by the null thesaurus is usually poorer than the recall obtainable with the regular thesaurus.

The most powerful methods are usually the statistical phrase and syntactic phrase searches. Phrase searching permits a more precise identification of document content than word counting, since phrases are inherently less ambiguous than individual words. The more context is included in the identification of a unit of information, the more precisely this unit is likely to be known. If the size of the phrase dictionary were of no importance, every possible sentence might be placed in this dictionary, and its exact meaning specified. A reasonable dictionary of phrases is, however, best restricted to pairs, triplets, etc., of concepts. This includes a considerable amount of context, and provides a substantial improvement in the accuracy of content identification.

The syntactic phrase searcher provides far more stringent matching criteria than the statistical phrase searcher. The statistical phrase searcher retrieves any phrase whose components occur in the same sentence, regardless of the distance between members, and regardless of their syntactic roles or relations. On the other hand, the syntactic phrase searching procedure, using the criterion tree dictionary, finds only phrases which match semantically, syntactically, and structurally with the dictionary entry. Contrary to what one might expect, the statistical phrase searcher usually works better than the syntactic phrase searching procedure. The reason for this seems to be that in the majority of cases use of the components of a phrase, although not necessarily implying the complete juxtaposition of the corresponding concepts, does nevertheless seem to result in discussion of the concept represented by the phrase. Consider, for example, the phrase "speech recognition," as it appears in the sentence "Segmentation of speech into discrete units suitable for recognition, including the possibility of overlapping elements, is considered." (from abstract #82). The statistical phrase searcher finds an occurrence of "speech recognition" while the syntactic phrase finder does not. Although the syntactic analyzer is technically correct, the sentence nevertheless deals with speech recognition, and the performance of a system admitting the phrase is better than that of a system which follows the strict rules of grammar, and ignores it. Sentences of the type "Some authors whose other work is solid state that information retrieval can be performed without accurate syntactic analysis," do require complete syntactic analysis and dependency tree matching, if a phrase such as "solid state," which does not reflect the meaning of the sentence, is to be eliminated. Such examples are rare in technical text.

The good over-all performance for the request under discussion is probably due to the precision of the terms contained in it. "Runge-Kutta" is highly specific to the topic requested; so is the phrase "differential equations" and "integration" (relative to the collection, of course). Any document using these words is almost certain to be relevant. It should be noted, however, that the request also contains "algorithms," "equation," "solution," "digital computer," "stability," and "speed," all of which are terms of somewhat smaller specificity. The computer nevertheless succeeds in extracting the relevant documents quite well for this request. A more detailed description of evaluation results is given in Secs. III, IV and X of this report.

```
INSTRUCTION CARDS TO SMART SUPERVISOR                         SEPTEMBER 28, 1964   PAGE 1

     ANSWER REQUESTS, FORMAT JUMBO, SCORES YES,  THESHR 2, MAXCON 511,            1
     CORMD3 COS, CUTOF3 3500, EXECUTE SYNTAX, SYNTACTIC ANALYSIS,                 1
     CRITERION TREES, NODE CORRESPONDENCES,  CRITES 300,                          1
     ENGLISH TEXTS, WORDS NOT FOUND, REQUEST CORRELATIONS                         1

     *TIME
THE CURRENT TIME IS   93.7 MINUTES. YOU WILL REMEMBER THAT START-OF-JOB
WAS AT   92.6 MINUTES, WHILE THE CLOCK READ   93.2 WHEN EXECUTION BEGAN.

     *FIND DIFFERNTL EQ  *** REQUEST ON DIFFERENTIAL EQUATIONS ***
```

SMART Specifications — Short Form

Figure 1

```
INSTRUCTION CARDS TO SMART SUPERVISOR                              SEPTEMBER 28, 1964   PAGE 375

            LIBRARY USED WAS VERSION NUMBER      2           OF THE HARRIS THESAURUS.
            THESAURUS DISCRIMINATES NOT MORE THAN      511    CONCEPTS.
            ENGLISH TEXTS WERE PRINTED DURING LOOKUP.
            WORDS NOT FOUND WERE PRINTED.


STATISTICAL INTRA-DOCUMENT PROCESSING --

            NONE.


STATISTICAL PHRASES --

            NONE.


CRITERION TREES --
            A SYNTACTIC ANALYSIS WAS PRINTED FOR EACH SENTENCE.
            CRITERION TREES DETECTED WERE PRINTED.
            NODE CORRESPONDENCES OF TREES TO SENTENCES WERE PRINTED.
            SYNTACTIC PHRASES HAD WEIGHT OF      3.0


THE ABOVE DATA WAS SUPPLIED BY THE PROGRAMMER AND MAY BE INCORRECT.

THE FOLLOWING DATA IS FROM INSTRUCTIONS FOR THIS RUN WHICH DEFINITELY WERE EXECUTED.


TITLES WERE GIVEN A WEIGHT OF       1.0


DOCUMENT CORRELATION --
            REQUEST CORRELATIONS WERE PRINTED.
            CORRELATION MODE USED WAS      COSINE.
            CUTOFF WAS  0.3500


HIERARCHY --

            NONE.


CONCEPT PROCESSING --

            NONE.


REQUESTS WERE ANSWERED.
AUTO-EVALUATION WAS REQUESTED AND WILL BE ATTEMPTED.
```

SMART Specifications – Long Form

Figure 2

```
ENGLISH TEXT PROVIDED FOR DOCUMENT  DIFFERNTL EQ                SEPTEMBER 28, 1964   PAGE 345

     GIVE ALGORITHMS USEFUL FOR THE NUMERICAL SOLUTION OF ORDINARY DIFFER-         1
     ENTIAL EQUATIONS AND PARTIAL DIFFERENTIAL EQUATIONS ON DIGITAL                1
     COMPUTERS .  EVALUATE THE VARIOUS INTEGRATION PROCEDURES (TRY                 1
     RUNGE-KUTTA, MILNE-S METHOD) WITH RESPECT TO ACCURACY, STABILITY, AND         2
     SPEED .                                                                       2


WORDS IN DOCUMENT DIFFERNTL EQ NOT FOUND IN THESAURUS           SEPTEMBER 28, 1964   PAGE 346

     WORD NOT FOUND       KIND    LOC  NUM       SENTENCE AND WORD NUMBERS

RESPECT                   STEM     8    1        2,   14
```

Dictionary Lookup of DIFFERNTL EQ

Figure 3

```
*TEXT 2MICRO-PROGRAMMING .

$MICRO-PROGRAMMING
$R. J. MERCER (UNIVERSITY OF CALIFORNIA)
$U.S. GOV. RES. REPTS. VOL 30 PP 71-72(A) (AUGUST 15, 1958) PB 126893

MICRO-PROGRAMMING . THE MICRO-PROGRAMMING TECHNIQUE OF DESIGNING THE
CONTROL CIRCUITS OF AN ELECTRONIC DIGITAL COMPUTER TO FORMALLY INTERPRET
AND EXECUTE A GIVEN SET OF MACHINE OPERATIONS AS AN EQUIVALENT SET
OF SEQUENCES OF ELEMENTARY OPERATIONS THAT CAN BE EXECUTED IN ONE
PULSE TIME IS DESCRIBED .


*TEXT 3THE ROLE OF LARGE MEMORIES IN SCIENTIFIC COMMUNICATIONS

$THE ROLE OF LARGE MEMORIES IN SCIENTIFIC COMMUNICATIONS
$M. M. ASTRAHAN (IBM CORP.)
$IBM J. RES. AND DEV. VOL 2 PP 310-313 (OCTOBER 1958)

THE ROLE OF LARGE MEMORIES IN SCIENTIFIC COMMUNICATIONS . THE ROLE
OF LARGE MEMORIES IN SCIENTIFIC COMMUNICATIONS IS DISCUSSED . LARGE
MEMORIES PROVIDE AUTOMATIC REFERENCE TO MILLIONS OF WORDS OF MACHINE-RE-
ADABLE CODED INFORMATION OR TO MILLIONS OF IMAGES OF DOCUMENT PAGES
. HIGHER DENSITIES OF STORAGE WILL MAKE POSSIBLE LOW-COST MEMORIES
OF BILLIONS OF WORDS WITH ACCESS TO ANY PART IN A FEW SECONDS OR COMPLE-
TE SEARCHES IN MINUTES . THESE MEMORIES WILL SERVE AS INDEXES TO THE
DELUGE OF TECHNICAL LITERATURE WHEN THE PROBLEMS OF INPUT AND OF THE
AUTOMATIC GENERATION OF CLASSIFICATION INFORMATION ARE SOLVED . DOCUMENT
FILES WILL MAKE THE INDEXED LITERATURE RAPIDLY AVAILABLE TO THE SEARCHER
. MACHINE TRANSLATION OF LANGUAGE AND RECOGNITION OF SPOKEN INFORMATION
ARE TWO OTHER AREAS WHICH WILL REQUIRE FAST, LARGE MEMORIES .
```

Sample Reference Documents

Figure 4

```
SYNTACTIC ANALYSIS OF SENTENCES FROM TEXT DIFFERNTL EQ         SEPTEMBER 28, 1964    PAGE 347

THE FOLLOWING ANALYSES WERE MADE USING ENGLISH GRAMMAR     00006C  PREDICTION POOL SIZE =   100  NESTER MAXIMUM =     8

ANALYSES OF SENTENCE NUMBER  000001

WORD                     HOMOGRAPHS

-----                    ------------

GIVE                     VT1P  IT1   VT2P  IT2   VI1P  II1

ALGORITHMS               NNNP  MMMP  NOUP

USEFUL                   AAA   ADJ

FOR                      YCC   PRE

THE                      AAA   ART

NUMERICAL                AAA   ADJ

SOLUTION                 NNNS  MMMS  NOUS

OF                       PRE

ORDINARY                 AAA   ADJ

DIFFERENTIAL             AAA   ADJ   NNNC  MMMC  NOVC

EQUATIONS                NNNP  MMMP  NOUP

AND                      XCC

PARTIAL                  AAA   ADJ   NNNS  MMMS  NOVS

DIFFERENTIAL             AAA   ADJ   NNNC  MMMC  NOVC

EQUATIONS                NNNP  MMMP  NOUP

ON                       PRE   AV2

DIGITAL                  AAA   ADJ

COMPUTERS                NNNP  MMMP  NOUP

.                        PRD
```

Homograph Codes Resulting from Lookup

Figure 5

SYNTACTIC ANALYSIS OF SENTENCES FROM TEXT DIFFERNTL EQ   SEPTEMBER 28, 1964   PAGE 348

***** ANALYSIS NUMBER 1   OF SENTENCE NUMBER 000001

ENGLISH	SENTENCE STRUCTURE	SWC	SWC CODE	SYNTACTIC ROLE	RL NUM	PREDICTION POOL
GIVE	3V	IT1	INFINITE VT1	IMPERATIVE VERB	SEIT10	SE
ALGORITHMS	30	NOUP	NOUN 1	OBJECT OF IMPERATIVE VERB	N2NNN1	PD N2B
USEFUL	3OA	ADJ	ADJECTIVE 1	POST-POSITIONAL ADJECTIVE	APADJ0	PD AP
FOR	3OAPR	PRE	PREPOSITION	PREPOSITION	PDPRE0	PD
THE	3OAPOA	ART	PRO-ADJECTIVE	OBJECT OF PREPOSITION	NQAAA0	PD NQG
NUMERICAL	3OAPOA	ADJ	ADJECTIVE 1	OBJECT OF PREPOSITION	N5ADJ0	PD N5G
SOLUTION	3OAPO	NOUS	NOUN 1	OBJECT OF PREPOSITION	N5MMM0	PD N5G
OF	3OAPOPR	PRE	PREPOSITION	PREPOSITION	XDPRE1	PD NQGZDB
ORDINARY	3OAPOPOA	ADJ	ADJECTIVE 1	OBJECT OF PREPOSITION	NQAAA0	PD NQGXDBNQG
DIFFERENTIAL	3OAPOPOA	ADJ	ADJECTIVE 1	OBJECT OF PREPOSITION	N5ADJ0	PD NQGXDBN5G
EQUATIONS	3OAPOPO	NOUP	NOUN 1	OBJECT OF PREPOSITION	N5MMM0	PD NQGXDBN5G
AND	3OAP+	XCO	COORDINATE CONJ1	COMPOUND OBJECT	XDXCO0	PD NQGXDB
PARTIAL	3OAPOA	ADJ	ADJECTIVE 1	OBJECT OF PREPOSITION	NQAAA0	PD NQG
DIFFERENTIAL	3OAPOA	ADJ	ADJECTIVE 1	OBJECT OF PREPOSITION	N5ADJ0	PD N5G
EQUATIONS	3OAPO	NOUP	NOUN 1	OBJECT OF PREPOSITION	N5MMM0	PD N5G
ON	3OAPOPR	PRE	PREPOSITION	PREPOSITION	PDPRE0	PD
DIGITAL	3OAPOPOA	ADJ	ADJECTIVE 1	OBJECT OF PREPOSITION	NQAAA0	PD NQG
COMPUTERS	3OAPOPO	NOUP	NOUN 1	OBJECT OF PREPOSITION	N5MMM0	PD N5G
.	3.	PRD	PERIOD	END OF SENTENCE	PDPRD0	PD

POOL OVERFLOWS= 0 NUMBER TEST FAILURES= 0 SHAPER OVERFLOWS= 229 NESTER OVERFLOWS= 59 TIME= 0.0 MINUTES 1

Syntactic Analysis of Sentence 1, DIFFERNTL EQ

Figure 6

```
THESE ARE THE TEXT,    NODE NUMBERS, AND STRINGS OF SENTENCE NO. 000001

GIVE                              2      3V
ALGORITHMS                        3      3O
USEFUL                            4      3OA
FOR                               6      3OAPR
THE                               8      3OAPOA
NUMERICAL                         9      3OAPOA
SOLUTION                          7      3OAPO
OF                               11      3OAPOPR
ORDINARY                         13      3OAPOPOA
DIFFERENTIAL                     14      3OAPOPOA
EQUATIONS                        12      3OAPOPO
AND                               5      3OAP+
PARTIAL                          16      3OAPOA
DIFFERENTIAL                     17      3OAPOA
EQUATIONS                        15      3OAPO
ON                               19      3OAPOPR
DIGITAL                          21      3OAPOPOA
COMPUTERS                        20      3OAPOPO

NODE CORRESPONDENCES OF TREE WITH INDEX =DIFEQU, SERIAL NO.     47, AND OUTPUT CONCEPT NOS
 TREE     SENTENCE
    1           12 - KEY
    2           14

NODE CORRESPONDENCES OF TREE WITH INDEX =NUMERI, SERIAL NO.     87, AND OUTPUT CONCEPT NOS
 TREE     SENTENCE
    2            9
    1            7 - KEY

THE CRITERION ROUTINE HAS PROCESSED      1 SENTENCES, HAVING      2 MATCHES OF       2 DISTINCT INDICES.
```

Trees Found in Sentence 1, DIFFERNTL EQ

Figure 7

TREES DETECTED SYNTACTICALLY IN DOCUMENT DIFFERNTL EQ SEPTEMBER 28, 1964 PAGE 350

TREE	CONCEPT	OCCURRED	COMPONENT CONCEPTS
DIFFQU	379DIF	1	274DIF 181QUA
NUMERI	375NUM	1	13CALC 11ALYS

Summary of Tree Dictionary Search

Figure 8

```
INSTRUCTION CARDS TO SMART SUPERVISOR                              SEPTEMBER 28, 1964   PAGE 1

     ANSWER REQUESTS, FORMAT JUMBO, PF, SCORES YES, THESHR 2, MAXCCN 511,             1
     REQUEST CORRELATIONS, CORMD3 COS, CUTOF3 3500, TEXTS PROCESSED                   1

     *LIST DIFFERNTL EQ NUMERICAL DIGITAL SOLN OF DIFFERENTIAL EQUATIONS

     *LIKE DIFFERNTL EQ

     *TIME
THE CURRENT TIME IS    72.1 MINUTES. YOU WILL REMEMBER THAT START-OF-JOB
WAS AT    69.9 MINUTES, WHILE THE CLOCK READ   71.8 WHEN EXECUTION BEGAN.

     *TAPE

     *NOTE THIS IS THE HARRIS THESAURUS VERSION TWO LOOKUP

     *LIST 1A COMPUTER ORIENTED TOWARD SPATIAL PROBLEMS .

     *LIST 2MICRO-PROGRAMMING .

     *LIST 3THE ROLE OF LARGE MEMORIES IN SCIENTIFIC COMMUNICATIONS

     *LIST 4A NEW CLASS OF DIGITAL DIVISION METHODS .

     *LIST 5ANALYSIS OF SHIFT REGISTER COUNTERS .

     *LIST 6GENERALIZED PARITY CHECKING .

     *LIST 7AN IMPROVED TECHNIQUE FOR FAST MULTIPLICATION ON SERIAL

     *LIST 8SHORT-CUT MULTIPLICATION AND DIVISION IN AUTOMATIC

     *LIST 9OPERATION AND SPECIFICATION OF TRANSISTORS FOR DIRECT-COUPLED

     *LIST 10ACCURATE TRANSISTORIZED MULTIPLIER-DIVIDER .

     *LIST 11A FULL BINARY ADDER EMPLOYING TWO NEGATIVE-RESISTANCE

     *LIST 12DIGITAL COMPUTER ADDING AND COMPLEMENTING CIRCUITS

     *LIST 13HALF-ADDERS DRIVE SIMULTANEOUS COMPUTER .

     *LIST 14THE SWITCHING CHARACTERISTICS OF 4-79 PERMALLOY CORES

     *LIST 15ELECTROLUMINESCENCE FROM BARIUM TITANATE .

     *LIST 16CONTROL APPARATUS FOR A SERIAL DRUM MEMORY .

     *LIST 17THE FUNCTION GENERATOR AS A BASIC TOOL FOR NON-LINEAR

     *LIST 18AN ACCURATE ANALOG FREQUENCY MEASURING CIRCUIT .

     *LIST 19RESISTANCE POTENTIOMETERS AS FUNCTION GENERATORS .

     *LIST 20DIFFERENTIATOR FOR A-C COMPUTERS .

     *LIST 21AN ERROR-CORRECTING ENCODER AND DECODER OF HIGH EFFICIENCY
```

Read-in of Preprocessed Documents

Figure 9

DOCUMENTS PROCESSED IN THIS EXECUTION   SEPTEMBER 28, 1964   PAGE 16

DIFF RNTL EQ	1A COMPUTER	2MICRO-PROGR	3THE ROLE OF	4A NEW CLASS	5ANALYSIS OF	6GENERALIZED	7AN IMPROVED	8SHORT-CUT M
9OPERATION A	10ACCURATE T	11A FULL BIN	12DIGITAL CO	13HALF-ADDER	14THE SWITCH	15ELECTROLUM	16CONTROL AP	17THE FUNCTI
18AN ACCURAT	19RESISTANCE	20DIFFERENTI	21AN ERROR-C	22LATCHING C	23MINIATURE	24SOME NOVEL	25A NEW TRAN	26SEMICONDUC
27TEN MEGAPU	28DESIGN OF	29INVESTIGAT	30A TRANSIST	31MAGNETIC C	32ANALOGUE I	33THE USE OF	34END-FIRED	35A LOAD-SHA
36FUNDAMENTA	37A HIGH-SPE	38AUTOMATIC	39AUTOMATIC	40COMMUNICAT	41COMMUNICAT	42A DIRECT R	43THE DATA C	44ACCURACY C
45A CALCULAT	46RAPID FIRE	47SPECIAL PU	48A BUSINESS	49A DUAL MAS	50ACCURACY C	51COMMUNICAT	52+ATHENA ,	53A COMPUTER
54AN AUTOMAT	55AUTOMATIC	56THE COMPUT	57CASE STUDY	58THE LARGES	59DATA PROCE	60INTELLIGEN	61AN INPUT R	62ON PROGRAM
63THE PROBLE	64PROPOSAL F	65FURTHER AU	66THE MATHEM	67AUTOMATIC	68FOUR YEARS	69COMPUTERS	70COMPUTERS	71AUTOMATIC
72HARMONIC A	73PROGRESS I	74DIGITAL CO	75METHOD FOR	76THE COMPUT	77WEAPON EVA	78DIGITAL SI	79MACHINE-MA	80RAPID RETR
81A TYPE OF	82ARTIFICIAL	83CHESS-PLAY	84ON THE LOD	85NOTE ON AN	86FOURIER AN	87AN ANALOG	88PROPOSED M	89CODING AND
90ON THE EQU	91ON SEQUENC	92SWITCHING	93THE APPLIC	94PSEUDO-RAN	95MINIMIZATI	96FORMAL ANA	97A NOTE ON	98SYNTHESIS
99TRANSLATIN	100ON THE IN	101METHOD FO	102ON THE SO	103RUNGE-KUT	104THE CALCU	105GENERATED	106COMPUTATI	107RECURSIVE
108 SECANT MO	109UNITARY T	110A METHOD	111A NOTE ON	112INTERLING	113A FRAMEWO	114SOME NEW	115INFORMATI	116ORDER OF
117COMPUTATI	118LINEAR LE	119CHANNELS	120COMPUTER	121LOGICAL M	122GLOSSARY	123THE SECON	124COMMUNICA	125PROBLEMS
126HOW MUCH	127LOGICAL D	128COMPUTER	129RELATION	130ON THE MI	131ANALYSIS	132LOGIC SYN	133MILLIMICR	134MILLIMICR
135THE LADDI	136A HIGH-SP	137COINCIDEN	138RESEARCH	139FERROELEC	140A BASIC T	141BISTABLE	142A MAGNETI	143TRANSISTO
144AN EMITTE	145A TRANSIS	146MULTIPLYI	147AN ELECTR	148MULTIPLIC	149DIODELESS	150MAGNETIC-	151A TRACK S	152TRANSISTO
153AN ANALOG	154MANUAL FO	155FEASIBILI	156FUNCTION	157DIGITAL A	158+BIDEC -	159DIGITAL R	160LOW COST	161ACHIEVING
162AN INPUT	163A COMPUTE	164LOGIC FOR	165CATALOG O	166COMPONENT	167AUTOMATIC	168GENERAL P	169THEORETIC	170COMPLEX P
171SMALL COM	172EFFECT OF	173AUTOMATIC	174DYNAMIC S	175THE DESIG	176FUNDAMENT	177TRANSAC C	178A LIBRARY	179PRELIMINA
180THE PROBL	181GENERALIZ	182ON MATRIX	183A PROGRAM	184THE PEGAS	185-- +DEUCE	186ALGORITHM	187COMPUTERS	188AUTOMATIO
189COMPUTERS	190DISTRIBUT	191DIGITAL C	192SOLVING E	193ARTIFICIA	194TECHNICAL	195TECHNICAL	196TECHNICAL	197TECHNICAL
198AIRCRAFT	199APPLYING	200SOLUTION	201ITERATIVE	202STABLE PR	203DESIGN OF	204DYES - MI	205AUTOMATIC	206CRYSTAL S
207AN APPLIC	208APPLICATI	209SOLUTION	210EVALUATIO	211MONTE CAR	212ABSTRACTS	213PREDICTIO	214COMPUTER	215DIGITAL C
216PREDISTOR	217SIMULATIO	218TEAM APPR	219THE USE O	220DATA PREP	221CHEMICAL	222RESEARCH	223THE USE O	224A METHOD
225DESCRIPTI	226+DEPI ..	227TEST ROUT	228AUTOMATIC	229MATRIX PR	230STUDY OF	231DIFFUSION	232STEAM-GEN	233ON THE SO
234ANALOGUE	235FREEZING	236OPERATING	237ON THE GE	238SIGNIFICA	239A METHOD	240A NEW PRO	241ANALYSIS	242THE MULTI
243ITERATIVE	244MINIMAL -	245SOME PROP	246FOLDING O	247SYMBOLIC	248METHOD FO	249MATHEMATI	250TWO SQUAR	251ERROR EST
252A CLASS O	253ROUND-OFF	254EXTRACTIO	255THREE LEV	256WHAT IS L	257SURVEY OF	258ELECTRONI	259TEN YEARS	260ORIGIN AN
261SYMBOLIC	262LOGICAL D	263MACHINE L	264THE DESIG	265TIME MULT	266UNIFYING	267MAGNETICS	268RECENT AD	269PRODUCING
270STUDY OF	271MILLIMICR	272EVALUATIO	273MILLIMICR	274HIGH-SPEE	275A NEW MEM	276SUPERCOND	277A CRYOGEN	278A NEW HIG
279SYMMETRIC	280DOUBLING	281A GENERAL	282DESIGN OF	283BINARY AR	284REVERSIBL	285QUADRATIC	286SYSTEM DE	287OPTICS AN
288BEAM-POSI	289STABLE FI	290COINCIDEN	291OPERATING	292REVERSIBL	293A FAST ME	294INHIBITED	295TRANSISTO	296A HIGH-SP
297A FIGURE	298A HIGH-SP	299A SENSING	300RANDOM NU	301THE X-1 C	302THE NEW U	303THE NATIO	304OPERATING	305MACHINE R
306ELECTRONI	307SIMULTANE	308ANALOG CO	309ORGANIZIN	310SIMULATIO	311A NONREAL	312CHARACTER	313THE REDUC	314THE MECHA
315COULD A M	316FINITE AU	317THE OPTIM	318FROM FORM	319ZURICH CO	320POSSIBLE	321SIGNAL CO	322USE OF CO	323RECURSIVE
324THE AIR F	325THE SHARE	326THE SHARE	327THE SHARE	328THE SHARE	329THE SHARE	330THE SHARE	331SESAME OP	332THE ARITH
333A CHECKLI	334BLUEPRINT	335BUSINESS	336UNIVERSIT	337PRELIMINA	338CALCULATI	339DIGITAL C	340THE IBM 6	341NODAL REP
342FORECASTI	343VERIFICAT	344SOME HELI	345WEATHER P	346MEMORY EF	347NEW METHO	348STRUCTURE	349THE STUDY	350PATTERN R
351A SYSTEM	352A CHESS P	353THE DESIG	354RADIX EXC	355A ROUTINE	356TECHNIQUE	357TEST PROG	358STABILIZA	359USE OF AN
360SIMULATIN	361SAINT COM	362GRID-CURR	363THE RESPO	364A RING MO	365A NOTE ON	366AN EFFICI	367ON EXPONE	368ABSOLUTE
369A NOTE ON	370A NEW COM	371A SYNTHES	372ON THE RF	373EXTENSION	374MATRIX SY	375A FUNCTIO	376ON CODES	377A NEW MET
378DECODING	379A MATHEMA	380TRANSPOSI	381A PROCEDU	382COMPUTATI	383A NOTE ON	384STABILITY	385NUMERICAL	386ELIMINATI
387BOUNDARY	388A METHOD	389AN ITERAT	390MONTE CAR	391NUMERICAL	392ON COMPUT	393SOME CYCL	394THE MORSE	395FULL DECO
396ON A PROP	3971958 PGEC	398POSITIONS	399AUTOMATIO	400IRE STAND	401A REVIEW	402PROGRAMMI	403A SURVEY	404SURVEY OF

Document Collection Short Form Identifiers

Figure 10

SEPTEMBER 28, 1964 PAGE 103

OCCURRENCES OF CONCEPTS AND PHRASES IN DOCUMENTS

DOCUMENT	CONCEPT,OCCURS									
DIFFERNTL EQ	ACCUR	12	ALGORI	12	COMPUT	12	DIFFER	24	DIGIT	12
	EQU	24	EVALU	12	GIVE	12	INTEGR	12	METHOD	12
	NUMER	12	ORDIN	12	PARTI	12	PROCED	12	RUNGE-	12
	SOLUT	12	SPEED	12	STABIL	12	USE	12	VARIE	12
1A COMPUTER	BAS	12	CHARAC	12	COMPUT	36	DESCRI	12	DESIGN	12
	DIRECT	12	ENABLE	12	ESTIM	12	EXPLAI	12	FORM	12
	GIVE	12	HANDLE	12	ILLUST	12	INDEPE	12	INFORM	12
	MACHIN	24	OPER	12	ORD	12	ORIENT	12	PLANE	12
	POS	12	POSS	12	PROBLE	36	PROGRA	36	RECOGN	12
	SCANN	12	SIMPLE	12	SIZE	24	STORE	12	STRUCT	12
	TECHNI	12	TOWARD	12	TRANSF	12	USING	12	WRITT	12
2MICRO-PROGR	CIRCUI	12	COMPUT	12	CONTRO	12	DESCRI	12	DESIGN	12
	DIGIT	12	ELECTR	12	ELEMEN	12	EQUIVA	12	EXECUT	24
	FORM	12	GIVE	12	INTERP	12	MACHIN	12	OPERAT	24
	PULSE	12	SEQU	12	SET	24	TECHNI	12	TIME	12
3THE ROLE OF	ACCES	12	AUTOMA	24	AVAIL	12	CLASSI	12	CODE	12
	COMMUN	24	COMPLE	12	DENS	12	DISCUS	12	DOCU	24
	FAST	12	FILE	12	GENER	12	HIGH	12	INDEX	24
	INFORM	36	INPUT	12	LANGU	12	LARGE	48	LITERA	24
	MACHIN	12	MAKE	24	MEM	72	MILL	24	PART	12
	POSS	12	PROBLE	12	PROVID	12	RAPID	12	RECOGN	12
	REFER	12	REQUIR	12	ROLE	24	SCIENT	24	SEARCH	24
	SOLVE	12	STORE	12	TECHN	12	TRANSL	12	WORD	24
4A NEW CLASS	0	24	+	24	ARBITR	12	ARITHM	12	CLASS	24
	COMPUT	12	CONSID	12	CONVEN	24	CONVER	12	DESCRI	24
	DIGIT	84	DIVIS	96	DUR	12	EXAMPL	12	FACILI	12
	FLO	12	FORM	12	GENER	12	GIVE	12	INTEG	24
	-DASH	12	METHOD	60	NAT	12	POINT	12	PROCED	24
	PROCES	12	QUOTI	60	RADIX	36	SERI	12	SUIT	12
	USE	12	VALUE	36	XXX	60				
5ANALYSIS OF	ALGEBR	12	ANAL	12	APPROA	12	COUNT	24	CYCLE	12
	DERIV	12	DETERM	12	FLIP-F	12	GENERA	12	GIVE	12
	LENG	12	NUMB	12	POSS	12	PRESEN	12	PROBLE	12
	REGIST	24	SEVER	12	SHIFT	24	SOLVE	12	THEORE	12
6GENERALIZED	ANAL	12	ARITHM	24	BASI	12	CASE	24	CHECK	60
	CONCEP	24	CONSID	24	CONVEN	12	DEFIN	24	EQU	24
	GENERA	12	GENERA	12	GIVE	12	INCLUD	12	MATHEM	12
	MEAN	24	NOTE	12	NUMB	36	PAP	12	PAR	72
	PROPER	24	PROVID	12	SUIT	12	SYSTEM	12	USED	12
	USU	12								

Concept Vectors — Null Thesaurus

Figure 11

```
                                              SEPTEMBER 28, 1964   PAGE 17

OCCURRENCES OF CONCEPTS AND PHRASES IN DOCUMENTS

DOCUMENT       CONCEPT,OCCURS

DIFFERNTL EQ   4EXACT 12   8ALGOR 12   13CALC 18   71EVAL  6   92DIGI 12
               110AUT 12   143UTI 12   176SOL 12   179STD 12   181QUA 24
               269ELI  4   274DIF 36   356VEL 12   357YAW  4   384TEG 12
               428STB  4   505APP 24

1A COMPUTER    2INPUT  4   5LOCAT 12   10ALPH 12   15BASE  6   16BASC  6
               31BIT   3   32REQU  3   41MCHO  8   47CHNG  6   53DATA  6
               57DSCB 15   59AMNT 24   72EXEC  6   77LIST  4   83MAP   6
               87ENBL 12   930RDR 10   106NQU  6   107DGN 30   108LOO 12
               110AUT 36   112OPE  6   119AUT  8   121MEM  4   130MEA  4
               143UTI 12   146JOB 18   147SYS 12   149POG 36   158REL 12
               162ROF  6   163FAS 12   168ORD  4   176SOL 12   178SYM 18
               182SAV  4   187DIR 12   210OUT  4   212SIZ 12   216COM 12
               276GEM 18   327AST 12   332SEE 12   338MCH  8   340LET  3
               346JET  6   350IFO  6   419GEM  6   501ORD  4   508ACT  6

2MICRO-PROGR   16BASC 12   41MCHO  4   46TROL  6   72EXEC 24   81FRML 12
               90INIT 12   92DIGI 12   930RDR 36   94SAME 12   103SET 12
               107DGN 12   110AUT 12   119AUT  4   135OPE 12   176SOL  6
               198TIM 12   226CIR 12   227ELE 12   229DCO  6   290ROL  6
               313PUL 12   338MCH  4   503SEQ 12

3THE ROLE OF   2INPUT  6   12USES 12   17FILE  4   32REQU 12   34CLRK  4
               39PART  6   41MCHO  4   53DATA 24   59AMNT 12   77LIST  8
               86FAST 24   98TRSL  6   101TAG 24   102LNG 12   108LOO 36
               114TEX 72   118BIG 48   119AUT 28   121MEM 76   143UTI 12
               146JOB  6   168ORD 12   170PHR 12   176SOL 12   178SYM  6
               182SAV  4   198TIM 24   202GEN 12   208WRD 12   214ECO 12
               222HAV 12   229DCO 18   246HRD  6   279HI  12   281COD 12
               332SEE 12   334VIS 12   337TOO  6   338MCH  4   349SND 24
               350IFO 18   362VOI 12   412FLE  6   497SCI 36   509GEN 12

4A NEW CLASS   13CALC 24   15BASE 36   36MTCH 12   39PART 48   47CHNG  6
               40COOR  6   51SPEC 24   60DIV 144   64EFFI  6   68EQUP  6
               92DIGI 24   107DGN 24   110AUT 12   129NUM 72   143UTI 12
               147SYS 12   150PNT  6   158REL 24   204OPT 18   254VRT  6
               276GEM  6   376RRR  6   381FOA 12   387ROO 12   495HAZ 12
               503SEQ 12   508ACT  6   509GEN 12

5ANALYSIS OF   7ALGEB 12   11ALYS 12   52CYCL 12   59AMNT 12   66ORGN  6
               111REG 48   129NUM 12   136CHO 12   143UTI 12   146JOB  6
               171SHF 24   176SOL 12   178SYM  6   185SWI 12   196AXI 12
               202GEN 12   218SHO  6
```

Concept Vectors — Harris 2 Thesaurus

Figure 12

```
                                                  SEPTEMBER 28, 1964    PAGE 273
OCCURRENCES OF CONCEPTS AND PHRASES IN DOCUMENTS

DOCUMENT        CONCEPT,OCCURS

DIFFERNTL EQ    4EXACT 12   8ALGOR 12   13CALC 18   71EVAL  6   92DIGI 12
                110AUT 12   143UTI 12   176SOL 12   179STD 12   181QUA 24
                269ELI  4   274DIF 36   356VEL 12   357YAW  4   375NUM 36
                379DIF 72   384TEG 12   428STB  4   505APP 24

1A COMPUTER     2INPUT  4   5LOCAT 12   10ALPH 12   14CODR 72   15BASE  6
                16BASC  6   31BIT   3   32REQU  3   41MCHO  8   47CHNG  6
                53DATA  6   57DSCB 15   59AMNT 24   72EXEC  6   77LIST  4
                83MAP   6   87ENBL 12   93ORDR 10   106NQU  6   107DGN 30
                108LOO 12   110AUT 36   112OPE  6   119AUT  8   121MEM  4
                130MEA  4   143UTI 12   146JOB 18   147SYS 12   149PUG 36
                158REL 12   162RUF  6   163EAS 12   168ORD  4   176SOL 12
                178SYM 18   182SAV  4   187DIR 12   200DA- 72   210OUT  4
                212SIZ 12   216DUM 12   219POG 36   276GEM 18   292THK 36
                302LOO 72   327AST 12   332SEE 48   338MCH  8   340LET  3
                346JET  6   350IFU  6   419GEM  6   501ORD  4   508ACT  6

2MICRO-PROGR    16BASC 12   41MCHO  4   46TROL  6   72EXEC 24   81FRML 12
                90INIT 12   92DIGI 12   93ORDR 36   94SAME 12   103SET 12
                107DGN 12   110AUT 12   119AUT  4   135OPE 12   176SOL  6
                198TIM 12   226CIR 12   227ELE 12   229DCO  6   290ROL  6
                294AUT 36   313PUL 12   338MCH  4   503SEQ 12

3THE ROLE OF    2INPUT  6   12USES 12   14CODR 36   17FILE  4   32REQU 12
                34CLRK  4   39PART  6   41MCHO  4   53DATA 24   59AMNT 12
                77LIST  8   86FAST 24   98TRSL  6   101TAG 24   102LNG 12
                108LOO 36   114TEX 72   118BIG 48   119AUT 28   121MEM 76
                128SYB 36   143UTI 12   146JOB  6   168ORD 12   170PHR 12
                176SOL 12   178SYM  6   182SAV  4   198TIM 24   202GEN 12
                208WRD 12   214ECO 12   222HAV 12   229DCO 18   246HRD  6
                279HI  12   281COD 12   292THK 36   302LOO108   303MTR 72
                304COD 36   332SEE 48   334VIS 12   337TOO  6   338MCH  4
                349SND 24   350IFU 18   362VOI 12   412FLD  6   497SCI 36
                500DEX 36   509GEN 12

4A NEW CLASS    13CALC 24   15BASE 36   36MTCH 12   39PART 48   47CHNG  6
                49COOR  6   51SPEC 24   60DIV 144   64EFFI  6   68EQUP '6
                92DIGI 24   107DGN 24   110AUT 12   129NUM 72   143UTI 12
                147SYS 12   150PNT  6   158REL 24   204OPT 18   254VRT  6
                276GEM  6   376RRR  6   381FOA 12   387ROO 12   495HAZ 12
                503SEQ 12   508ACT  6   509GEN 12

5ANALYSIS OF    7ALGEB 12   11ALYS 12   52CYCL 12   59AMNT 12   66ORGN  6
                111REG 48   129NUM 12   136CHO 12   143UTI 12   146JOB  6
                171SHF 24   176SOL 12   173SYM  6   185SWI 12   196AXI 12
                202GEN 12   218SHU  6   292THK 36   295POV 36
```

Concept Vectors After Statistical Phrase Search

Figure 13

CORRELATIONS WITH REGULAR THESAURUS		CORRELATIONS WITH NULL THESAURUS	
DIFFERNTL EQ 1A COMPUTER	0.1234	DIFFERNTL EQ 1A COMPUTER	0.0973
DIFFERNTL EQ 2MICRO-PROGR	0.0875	DIFFERNTL EQ 2MICRO-PROGR	0.1093
DIFFERNTL EQ 3THE ROLE OF	0.0293	DIFFERNTL EQ 4A NEW CLASS	0.2096
DIFFERNTL EQ 4A NEW CLASS	0.0844	DIFFERNTL EQ 5ANALYSIS OF	0.0364
DIFFERNTL EQ 5ANALYSIS OF	0.0658	DIFFERNTL EQ 6GENERALIZED	0.0907
DIFFERNTL EQ 6GENERALIZED	0.0741	DIFFERNTL EQ 7AN IMPROVED	0.1220
DIFFERNTL EQ 7AN IMPROVED	0.2090	DIFFERNTL EQ 8SHORT-CUT M	0.1298
DIFFERNTL EQ 8SHORT-CUT M	0.0861	DIFFERNTL EQ 9OPERATION A	0.1013
DIFFERNTL EQ 9OPERATION A	0.0611	DIFFERNTL EQ 10ACCURATE T	0.0716
DIFFERNTL EQ 10ACCURATE T	0.1100	DIFFERNTL EQ 12DIGITAL CO	0.1478
DIFFERNTL EQ 12DIGITAL CO	0.0883	DIFFERNTL EQ 13HALF-ADDER	0.1196
DIFFERNTL EQ 13HALF-ADDER	0.0548	DIFFERNTL EQ 16CONTROL AP	0.0535
DIFFERNTL EQ 16CONTROL AP	0.0336	DIFFERNTL EQ 17THE FUNCTI	0.0428
DIFFERNTL EQ 17THE FUNCTI	0.0580	DIFFERNTL EQ 18AN ACCURAT	0.1961
DIFFERNTL EQ 18AN ACCURAT	0.1397	DIFFERNTL EQ 19RESISTANCE	0.0851
DIFFERNTL EQ 19RESISTANCE	0.0177	DIFFERNTL EQ 20DIFFERENTI	0.2676
DIFFERNTL EQ 20DIFFERENTI	0.2123	DIFFERNTL EQ 24SOME NOVEL	0.0207
DIFFERNTL EQ 21AN ERROR-C	0.0105	DIFFERNTL EQ 25A NEW TRAN	0.0713
DIFFERNTL EQ 22LATCHING C	0.0057	DIFFERNTL EQ 26SEMICONDUC	0.0544
DIFFERNTL EQ 23MINIATURE	0.0307	DIFFERNTL EQ 27TEN MEGAPU	0.1233
DIFFERNTL EQ 24SOME NOVEL	0.0199	DIFFERNTL EQ 28DESIGN OF	0.0996
DIFFERNTL EQ 25A NEW TRAN	0.1068	DIFFERNTL EQ 29INVESTIGAT	0.0444
DIFFERNTL EQ 26SEMICONDUC	0.0653	DIFFERNTL EQ 30A TRANSIST	0.1216
DIFFERNTL EQ 27TEN MEGAPU	0.1004	DIFFERNTL EQ 31MAGNETIC C	0.1057
DIFFERNTL EQ 28DESIGN OF	0.1375	DIFFERNTL EQ 32ANALOGUE I	0.1538
DIFFERNTL EQ 29INVESTIGAT	0.0879	DIFFERNTL EQ 33THE USE OF	0.1591
DIFFERNTL EQ 30A TRANSIST	0.0736	DIFFERNTL EQ 34END-FIRED	0.0439
DIFFERNTL EQ 31MAGNETIC C	0.0575	DIFFERNTL EQ 35A LOAD-SHA	0.0345
DIFFERNTL EQ 32ANALOGUE I	0.2283	DIFFERNTL EQ 36FUNDAMENTA	0.0377
DIFFERNTL EQ 33THE USE OF	0.0802	DIFFERNTL EQ 37A HIGH-SPE	0.0186
DIFFERNTL EQ 34END-FIRED	0.0458	DIFFERNTL EQ 38AUTOMATIC	0.0981
DIFFERNTL EQ 35A LOAD-SHA	0.0331	DIFFERNTL EQ 41COMMUNICAT	0.1682
DIFFERNTL EQ 36FUNDAMENTA	0.0392	DIFFERNTL EQ 42A DIRECT R	0.0897
DIFFERNTL EQ 37A HIGH-SPE	0.0364	DIFFERNTL EQ 43THE DATA C	0.0314
DIFFERNTL EQ 38AUTOMATIC	0.1043	DIFFERNTL EQ 44ACCURACY C	0.1070
DIFFERNTL EQ 41COMMUNICAT	0.1185	DIFFERNTL EQ 45A CALCULAT	0.1525
DIFFERNTL EQ 42A DIRECT R	0.0439	DIFFERNTL EQ 46RADIO DIRE	0.1462
DIFFERNTL EQ 43THE DATA C	0.0333	DIFFERNTL EQ 47SPECIAL PU	0.0840
DIFFERNTL EQ 44ACCURACY C	0.1399	DIFFERNTL EQ 48A BUSINESS	0.0214
DIFFERNTL EQ 45A CALCULAT	0.2958	DIFFERNTL EQ 49A DUAL MAS	0.0335
DIFFERNTL EQ 46RADIO DIRE	0.0980	DIFFERNTL EQ 50ACCURACY C	0.0479
DIFFERNTL EQ 47SPECIAL PU	0.1268	DIFFERNTL EQ 52+ATHENA ,	0.1112
DIFFERNTL EQ 48A BUSINESS	0.0086	DIFFERNTL EQ 53A COMPUTER	0.1068
DIFFERNTL EQ 49A DUAL MAS	0.0575	DIFFERNTL EQ 54AN AUTOMAT	0.0777
DIFFERNTL EQ 50ACCURACY C	0.0668	DIFFERNTL EQ 55AUTOMATIC	0.0450
DIFFERNTL EQ 52+ATHENA ,	0.1030	DIFFERNTL EQ 56THE COMPUT	0.1634
DIFFERNTL EQ 53A COMPUTER	0.1327	DIFFERNTL EQ 57CASE STUDY	0.0306
DIFFERNTL EQ 54AN AUTOMAT	0.0763	DIFFERNTL EQ 59DATA PROCE	0.0347
DIFFERNTL EQ 55AUTOMATIC	0.0746	DIFFERNTL EQ 60INTELLIGEN	0.0572
DIFFERNTL EQ 56THE COMPUT	0.1513	DIFFERNTL EQ 61AN INPUT R	0.1641
DIFFERNTL EQ 57CASE STUDY	0.0950	DIFFERNTL EQ 62ON PROGRAM	0.0944
DIFFERNTL EQ 58THE LARGES	0.0256	DIFFERNTL EQ 63THE PROBLE	0.0176
DIFFERNTL EQ 59DATA PROCE	0.0302	DIFFERNTL EQ 64PROPOSAL F	0.0238
DIFFERNTL EQ 60INTELLIGEN	0.0091	DIFFERNTL EQ 65FURTHER AU	0.2011
DIFFERNTL EQ 61AN INPUT R	0.0404	DIFFERNTL EQ 66THE MATHEM	0.0769
DIFFERNTL EQ 62ON PROGRAM	0.1181	DIFFERNTL EQ 67AUTOMATIC	0.1207

Correlation Lists in Document Order

Figure 14

CORRELATIONS WITH REGULAR THESAURUS		CORRELATIONS WITH NULL THESAURUS	
DIFFERNTL EQ 384STABILITY	0.6675	DIFFERNTL EQ 103RUNGE-KUT	0.5508
DIFFERNTL EQ 360SIMULATIN	0.5758	DIFFERNTL EQ 384STABILITY	0.5462
DIFFERNTL EQ 200SOLUTION	0.5663	DIFFERNTL EQ 392ON COMPUT	0.5271
DIFFERNTL EQ 392ON COMPUT	0.5508	DIFFERNTL EQ 387BOUNDARY	0.4309
DIFFERNTL EQ 386ELIMINATI	0.5483	DIFFERNTL EQ 386ELIMINATI	0.4293
DIFFERNTL EQ 103RUNGE-KUT	0.5444	DIFFERNTL EQ 102ON THE SO	0.4073
DIFFERNTL EQ 85NOTE ON AN	0.4510	DIFFERNTL EQ 202STABLE PR	0.3981
DIFFERNTL EQ 192SOLVING E	0.4106	DIFFERNTL EQ 360SIMULATIN	0.3922
DIFFERNTL EQ 358STABILIZA	0.3986	DIFFERNTL EQ 192SOLVING E	0.3911
DIFFERNTL EQ 102ON THE SO	0.3986	DIFFERNTL EQ 200SOLUTION	0.3671
DIFFERNTL EQ 387BOUNDARY	0.3968	DIFFERNTL EQ 358STABILIZA	0.3441
DIFFERNTL EQ 202STABLE PR	0.3906	DIFFERNTL EQ 385NUMERICAL	0.3268
DIFFERNTL EQ 229MATRIX PR	0.3505	DIFFERNTL EQ 251ERROR EST	0.3202
DIFFERNTL EQ 88PROPOSED M	0.3451	DIFFERNTL EQ 85NOTE ON AN	0.3162
DIFFERNTL EQ 251ERROR EST	0.3329	DIFFERNTL EQ 229MATRIX PR	0.3044
DIFFERNTL EQ 234ANALOGUE	0.3176	DIFFERNTL EQ 359USE OF AN	0.3038
DIFFERNTL EQ 253ROUND-OFF	0.3152	DIFFERNTL EQ 169THEORETIC	0.3028
DIFFERNTL EQ 186ALGORITHM	0.3144	DIFFERNTL EQ 253ROUND-OFF	0.3013
DIFFERNTL EQ 169THEORETIC	0.3136	DIFFERNTL EQ 383A NOTE ON	0.2983

Correlation List Ordered by Size of Correlation

Figure 15

HISTOGRAM WITH REGULAR THESAURUS		HISTOGRAM WITH NULL THESAURUS	
0.9800	0	0.9800	0
0.9600	0	0.9600	0
0.9400	0	0.9400	0
0.9200	0	0.9200	0
0.9000	0	0.9000	0
0.8800	0	0.8800	0
0.8600	0	0.8600	0
0.8400	0	0.8400	0
0.8200	0	0.8200	0
0.8000	0	0.8000	0
0.7800	0	0.7800	0
0.7600	0	0.7600	0
0.7400	0	0.7400	0
0.7200	0	0.7200	0
0.7000	0	0.7000	0
0.6800	0	0.6800	0
0.6600	1	0.6600	0
0.6400	1	0.6400	0
0.6200	1	0.6200	0
0.6000	1	0.6000	0
0.5800	1	0.5800	0
0.5600	3	0.5600	0
0.5400	6	0.5400	2
0.5200	6	0.5200	3
0.5000	6	0.5000	3
0.4800	6	0.4800	3
0.4600	6	0.4600	3
0.4400	7	0.4400	3
0.4200	7	0.4200	5
0.4000	8	0.4000	6
0.3800	12	0.3800	9
0.3600	12	0.3600	10
0.3400	14	0.3400	11
0.3200	15	0.3200	13
0.3000	21	0.3000	18
0.2800	23	0.2800	26
0.2600	33	0.2600	30
0.2400	34	0.2400	36
0.2200	47	0.2200	39
0.2000	60	0.2000	44
0.1800	73	0.1800	55
0.1600	87	0.1600	77
0.1400	106	0.1400	111
0.1200	135	0.1200	136
0.1000	166	0.1000	177
0.0800	200	0.0800	215
0.0600	257	0.0600	257
0.0400	304	0.0400	300
0.0200	348	0.0200	357

Histogram of Correlation Distribution

Figure 16

```
CHANGES IN DOCUMENT-CONCEPT MATRIX THROUGH USE OF HIERARCHY    SEPTEMBER 28, 1964   PAGE 244

DIFFERNTL EQ  CONCEPT-PHRASE(WEIGHT)VECTOR BEFORE EXPANSION
4EXACT( 12) 8ALGOR( 12) 13CALC( 18) 71EVAL(  6) 92DIGI( 12) 110AUT( 12) 143UTI( 12) 176SOL( 12) 179STD( 12) 181QUA( 24) 269ELI(  4)
274DIF( 36) 356VEL( 12) 357YAW(  4) 384TEG( 12) 428STB(  4) 505APP( 24)

DIFFERNTL EQ  CONCEPT-PHRASE(WEIGHT)VECTOR  AFTER EXPANSION
4EXACT( 28) 219POG( 12) 375NUM( 42) 64EFFI(  6) 110AUT( 36) 145RNK( 12) 209AIM( 12) 163EAS( 12) 196AXI( 24) 379DIF( 48) 439MOV( 12)
346JET(  4) 428STB(  8) 8ALGOR( 12) 13CALC( 18) 71EVAL(  6) 92DIGI( 12) 143UTI( 12) 176SOL( 12) 179STD( 12) 181QUA( 24) 269ELI(  4)
274DIF( 36) 356VEL( 12) 357YAW(  4) 384TEG( 12) 505APP( 24)
```

Hierarchical Expansion of Request DIFFERNTL EQ

Figure 17

CORRELATIONS
BEFORE EXPANSION

DIFFERNTL EQ	1A COMPUTER	0.1234
DIFFERNTL EQ	2MICRO-PROGR	0.0875
DIFFERNTL EQ	3THE ROLE OF	0.0293
DIFFERNTL EQ	4A NEW CLASS	0.0844
DIFFERNTL EQ	5ANALYSIS OF	0.0658
DIFFERNTL EQ	6GENERALIZED	0.0741
DIFFERNTL EQ	7AN IMPROVED	0.2090
DIFFERNTL EQ	8SHORT-CUT M	0.0861
DIFFERNTL EQ	9OPERATION A	0.0611
DIFFERNTL EQ	10ACCURATE T	0.1100
DIFFERNTL EQ	12DIGITAL CO	0.0883
DIFFERNTL EQ	13HALF-ADDER	0.0548
DIFFERNTL EQ	16CONTROL AP	0.0336
DIFFERNTL EQ	17THE FUNCTI	0.0580
DIFFERNTL EQ	18AN ACCURAT	0.1397
DIFFERNTL EQ	19RESISTANCE	0.0177
DIFFERNTL EQ	20DIFFERENTI	0.2123
DIFFERNTL EQ	21AN ERROR-C	0.0105
DIFFERNTL EQ	22LATCHING C	0.0057
DIFFERNTL EQ	23MINIATURE	0.0307
DIFFERNTL EQ	24SOME NOVEL	0.0199
DIFFERNTL EQ	25A NEW TRAN	0.1068
DIFFERNTL EQ	26SEMICONDUC	0.0653
DIFFERNTL EQ	27TEN MEGAPU	0.1004
DIFFERNTL EQ	28DESIGN OF	0.1375
DIFFERNTL EQ	29INVESTIGAT	0.0879
DIFFERNTL EQ	30A TRANSIST	0.0736
DIFFERNTL EQ	31MAGNETIC C	0.0575
DIFFERNTL EQ	32ANALOGUE I	0.2283
DIFFERNTL EQ	33THE USE OF	0.0802
DIFFERNTL EQ	34END-FIRED	0.0458
DIFFERNTL EQ	35A LOAD-SHA	0.0331
DIFFERNTL EQ	36FUNDAMENTA	0.0392
DIFFERNTL EQ	37A HIGH-SPE	0.0364
DIFFERNTL EQ	38AUTOMATIC	0.1043
DIFFERNTL EQ	41COMMUNICAT	0.1185
DIFFERNTL EQ	42A DIRECT R	0.0439
DIFFERNTL EQ	43THE DATA C	0.0333
DIFFERNTL EQ	44ACCURACY C	0.1399
DIFFERNTL EQ	45A CALCULAT	0.2958
DIFFERNTL EQ	46RADIO DIRE	0.0980
DIFFERNTL EQ	47SPECIAL PU	0.1268
DIFFERNTL EQ	48A BUSINESS	0.0086
DIFFERNTL EQ	49A DUAL MAS	0.0575
DIFFERNTL EQ	50ACCURACY C	0.0668
DIFFERNTL EQ	52+ATHENA ,	0.1030
DIFFERNTL EQ	53A COMPUTER	0.1327
DIFFERNTL EQ	54AN AUTOMAT	0.0763
DIFFERNTL EQ	55AUTOMATIC	0.0746
DIFFERNTL EQ	56THE COMPUT	0.1513
DIFFERNTL EQ	57CASE STUDY	0.0950
DIFFERNTL EQ	58THE LARGES	0.0256
DIFFERNTL EQ	59DATA PROCE	0.0302
DIFFERNTL EQ	60INTELLIGEN	0.0091
DIFFERNTL EQ	61AN INPUT R	0.0404
DIFFERNTL EQ	62ON PROGRAM	0.1181

CORRELATIONS
AFTER EXPANSION

DIFFERNTL EQ	1A COMPUTER	0.1812	
DIFFERNTL EQ	2MICRO-PROGR	0.0950	
DIFFERNTL EQ	3THE ROLE OF	0.0177	
DIFFERNTL EQ	4A NEW CLASS	0.0673	
DIFFERNTL EQ	5ANALYSIS OF	0.0794	
DIFFERNTL EQ	6GENERALIZED	0.0447	
DIFFERNTL EQ	7AN IMPROVED	0.1792	
DIFFERNTL EQ	8SHORT-CUT M	0.0919	
DIFFERNTL EQ	9OPERATION A	0.0922	
DIFFERNTL EQ	10ACCURATE T	0.1659	
DIFFERNTL EQ	12DIGITAL CO	0.0888	
DIFFERNTL EQ	13HALF-ADDER	0.0614	
DIFFERNTL EQ	16CONTROL AP	0.0405	
DIFFERNTL EQ	17THE FUNCTI	0.0909	
DIFFERNTL EQ	18AN ACCURAT	0.1721	
DIFFERNTL EQ	19RESISTANCE	0.0428	
DIFFERNTL EQ	20DIFFERENTI	0.1565	
DIFFERNTL EQ	21AN ERROR-C	0.0318	
DIFFERNTL EQ	22LATCHING C	0.0137	
DIFFERNTL EQ	23MINIATURE	0.0185	
DIFFERNTL EQ	24SOME NOVEL	0.0169	
DIFFERNTL EQ	25A NEW TRAN	0.1289	
DIFFERNTL EQ	26SEMICONDUC	0.0911	
DIFFERNTL EQ	27TEN MEGAPU	0.1596	
DIFFERNTL EQ	28DESIGN OF	0.1131	
DIFFERNTL EQ	29INVESTIGAT	0.0530	
DIFFERNTL EQ	30A TRANSIST	0.0444	
DIFFERNTL EQ	31MAGNETIC C	0.0347	
DIFFERNTL EQ	32ANALOGUE I	0.1377	
DIFFERNTL EQ	33THE USE OF	0.1129	
DIFFERNTL EQ	34END-FIRED	0.0277	
DIFFERNTL EQ	35A LOAD-SHA	0.0542	
DIFFERNTL EQ	36FUNDAMENTA	0.0433	
DIFFERNTL EQ	37A HIGH-SPE	0.0487	
DIFFERNTL EQ	38AUTOMATIC	0.0676	
DIFFERNTL EQ	40COMMUNICAT	0.0202	←
DIFFERNTL EQ	41COMMUNICAT	0.1207	
DIFFERNTL EQ	42A DIRECT R	0.0265	
DIFFERNTL EQ	43THE DATA C	0.0603	
DIFFERNTL EQ	44ACCURACY C	0.2016	
DIFFERNTL EQ	45A CALCULAT	0.3127	
DIFFERNTL EQ	46RADIO DIRE	0.1477	
DIFFERNTL EQ	47SPECIAL PU	0.2125	
DIFFERNTL EQ	48A BUSINESS	0.0133	
DIFFERNTL EQ	49A DUAL MAS	0.0760	
DIFFERNTL EQ	50ACCURACY C	0.0878	
DIFFERNTL EQ	51COMMUNICAT	0.0233	←
DIFFERNTL EQ	52+ATHENA ,	0.2175	
DIFFERNTL EQ	53A COMPUTER	0.2033	
DIFFERNTL EQ	54AN AUTOMAT	0.1041	
DIFFERNTL EQ	55AUTOMATIC	0.0930	
DIFFERNTL EQ	56THE COMPUT	0.1662	
DIFFERNTL EQ	57CASE STUDY	0.1347	
DIFFERNTL EQ	58THE LARGES	0.0154	
DIFFERNTL EQ	59DATA PROCE	0.0547	
DIFFERNTL EQ	60INTELLIGEN	0.041	

Correlation Lists Before and After Hierarchical
Expansion – Document Order

Figure 18

CORRELATIONS BEFORE EXPANSION		CORRELATIONS AFTER EXPANSION	
DIFFERNTL EQ 384STABILITY	0.6675	DIFFERNTL EQ 103RUNGE-KUT	0.4566
DIFFERNTL EQ 360SIMULATIN	0.5758	DIFFERNTL EQ 392ON COMPUT	0.4365
DIFFERNTL EQ 200SOLUTION	0.5663	DIFFERNTL EQ 384STABILITY	0.4261
DIFFERNTL EQ 392ON COMPUT	0.5508	DIFFERNTL EQ 360SIMULATIN	0.3970
DIFFERNTL EQ 386ELIMINATI	0.5483	DIFFERNTL EQ 386ELIMINATI	0.3827
DIFFERNTL EQ 103RUNGE-KUT	0.5444	DIFFERNTL EQ 200SOLUTION	0.3677
DIFFERNTL EQ 85NOTE ON AN	0.4510	DIFFERNTL EQ 234ANALOGUE	0.3455
DIFFERNTL EQ 192SOLVING E	0.4106	DIFFERNTL EQ 128COMPUTER	0.3238
DIFFERNTL EQ 358STABILIZA	0.3986	DIFFERNTL EQ 387BOUNDARY	0.3224
DIFFERNTL EQ 102ON THE SO	0.3986	DIFFERNTL EQ 226+CEPI ..	0.3171
DIFFERNTL EQ 387BOUNDARY	0.3968	DIFFERNTL EQ 45A CALCULAT	0.3126
DIFFERNTL EQ 202STABLE PR	0.3906	DIFFERNTL EQ 202STABLE PR	0.3123
DIFFERNTL EQ 229MATRIX PR	0.3505	DIFFERNTL EQ 358STABILIZA	0.3100
DIFFERNTL EQ 88PROPOSED M	0.3451	DIFFERNTL EQ 306ELECTRONI	0.3042
DIFFERNTL EQ 251ERROR EST	0.3329	DIFFERNTL EQ 192SOLVING E	0.3019
DIFFERNTL EQ 234ANALOGUE	0.3176	DIFFERNTL EQ 85NOTE ON AN	0.3015
DIFFERNTL EQ 253ROUND-OFF	0.3152	DIFFERNTL EQ 171SMALL COM	0.2962
DIFFERNTL EQ 186ALGORITHM	0.3144	DIFFERNTL EQ 318FROM FORM	0.2925
DIFFERNTL EQ 169THEORETIC	0.3136	DIFFERNTL EQ 173AUTOMATIC	0.2853

Correlation Lists Before and After Hierarchical
Expansion — Correlation Order

Figure 19

HISTOGRAM BEFORE EXPANSION		HISTOGRAM AFTER EXPANSION	
0.9800	0	0.9800	0
0.9600	0	0.9600	0
0.9400	0	0.9400	0
0.9200	0	0.9200	0
0.9000	0	0.9000	0
0.8800	0	0.8800	0
0.8600	0	0.8600	0
0.8400	0	0.8400	0
0.8200	0	0.8200	0
0.8000	0	0.8000	0
0.7800	0	0.7800	0
0.7600	0	0.7600	0
0.7400	0	0.7400	0
0.7200	0	0.7200	0
0.7000	0	0.7000	0
0.6800	0	0.6800	0
0.6600	1	0.6600	0
0.6400	1	0.6400	0
0.6200	1	0.6200	0
0.6000	1	0.6000	0
0.5800	1	0.5800	0
0.5600	3	0.5600	0
0.5400	6	0.5400	0
0.5200	6	0.5200	0
0.5000	6	0.5000	0
0.4800	6	0.4800	0
0.4600	6	0.4600	0
0.4400	7	0.4400	1
0.4200	7	0.4200	3
0.4000	8	0.4000	3
0.3800	12	0.3800	5
0.3600	12	0.3600	6
0.3400	14	0.3400	7
0.3200	15	0.3200	9
0.3000	21	0.3000	16
0.2800	23	0.2800	19
0.2600	33	0.2600	30
0.2400	34	0.2400	39
0.2200	47	0.2200	46
0.2000	60	0.2000	64
0.1800	73	0.1800	80
0.1600	87	0.1600	102
0.1400	106	0.1400	122
0.1200	135	0.1200	144
0.1000	166	0.1000	179
0.0800	200	0.0800	217
0.0600	257	0.0600	259
0.0400	304	0.0400	312
0.0200	348	0.0200	348

Histograms Before and After Hierarchical Expansion

Figure 20

```
INSTRUCTION CARDS TO SMART SUPERVISOR                                 SEPTEMBER 28, 1964   PAGE 82

          LIBRARY USED WAS VERSION NUMBER      2           OF THE HARRIS THESAURUS.
          THESAURUS DISCRIMINATES NOT MORE THAN      511   CONCEPTS.


STATISTICAL INTRA-DOCUMENT PROCESSING --

          NONE.


STATISTICAL PHRASES --

          NONE.


CRITERION TREES --

          NONE.


THE ABOVE DATA WAS SUPPLIED BY THE PROGRAMMER AND MAY BE INCORRECT.

THE FOLLOWING DATA IS FROM INSTRUCTIONS FOR THIS RUN WHICH DEFINITELY WERE EXECUTED.


TITLES WERE GIVEN A WEIGHT OF       1.0


TEXTS PROCESSED WERE PRINTED.
CONCEPT OCCURRENCES BY DOCUMENT WERE PRINTED.


DOCUMENT CORRELATION --
          REQUEST CORRELATIONS WERE PRINTED.
          CORRELATION MODE USED WAS      COSINE.
          CUTOFF WAS  0.3500


HIERARCHY --

          NONE.


CONCEPT PROCESSING --

          NONE.


REQUESTS WERE ANSWERED.
AUTO-EVALUATION WAS REQUESTED AND WILL BE ATTEMPTED.
```

Specifications for Regular Thesaurus Processing

Figure 21

```
ANSWERS TO REQUESTS FOR DOCUMENTS ON SPECIFIED TOPICS          SEPTEMBER 28, 1964    PAGE 83

                 CURRENT REQUEST - *LIST DIFFERNTL EQ NUMERICAL DIGITAL SOLN OF DIFFERENTIAL EQUATIONS

REQUEST    *LIST DIFFERNTL EQ NUMERICAL DIGITAL SOLN OF DIFFERENTIAL EQUATIONS
-------
            GIVE ALGORITHMS USEFUL FOR THE NUMERICAL SOLUTION OF ORDINARY
            DIFFERENTIAL EQUATIONS AND PARTIAL DIFFERENTIAL EQUATIONS ON DIGITAL
            COMPUTERS . EVALUATE THE VARIOUS INTEGRATION PROCEDURES (E.G. RUNGE--
            KUTTA, MILNE-S METHOD) WITH RESPECT TO ACCURACY, STABILITY, AND SPEED .


ANSWER              CORRELATION                        IDENTIFICATION
------              -----------                        --------------
3?4STABILITY         0.6675                  STABILITY OF NUMERICAL SOLUTION OF DIFFERENTIAL EQUATIONS
                                             W. E. MILNE AND R. R. REYNOLDS (OREGON STATE COLLEGE)
                                             J. ASSOC. FOR COMPUTING MACH. VOL 6 PP 196-203 (APRIL, 1959)


ANSWER              CORRELATION                        IDENTIFICATION
------              -----------                        --------------
3?CSIMULATI?         0.5758                  SIMULATING SECOND-ORDER EQUATIONS
                                             D. G. CHADWICK (UTAH STATE UNIV.)
                                             ELECTRONICS VOL 32 P 64 (MARCH 6, 1959)


ANSWER              CORRELATION                        IDENTIFICATION
------              -----------                        --------------
2??SOLUTION          0.5663                  SOLUTION OF ALGEBRAIC AND TRANSCENDENTAL EQUATIONS ON AN AUTOMATIC
                                             DIGITAL COMPUTER
                                             G.N. LANCE (UNIV. OF SOUTHAMPTON)
                                             J. ASSOC. FOR COMPUTING MACH., VOL 6, PP 97-101, JAN., 1959


ANSWER              CORRELATION                        IDENTIFICATION
------              -----------                        --------------
392ON COMPUT         0.5558                  ON COMPUTING RADIATION INTEGRALS
                                             R. C. HANSEN (HUGHES AIRCRAFT CO.), L. L. BAILIN (UNIV. OF SOUTHERN
                                             CALIFORNIA, AND R. W. RUTISHAUSER (LITTON INDUSTRIES, INC.)
                                             COMMMUN. ASSOC) FOR COMPUTING MACH. VOL 2 PP 28-31 (FEBRUARY, 1959)


ANSWER              CORRELATION                        IDENTIFICATION
------              -----------                        --------------
366ELIMINATI         0.5483                  ELIMINATION OF SPECIAL FUNCTIONS FROM DIFFERENTIAL EQUATIONS
                                             J. E. POWERS (UNIV. OF OKLAHOMA)
                                             COMMUN. ASSOC. FOR COMPTING MACH. VOL 2 PP 3-4 (MARCH, 1959)
```

Answers to Request DIFFERNTL EQ — JUMBO Format

Figure 22

ANSWERS TO REQUESTS FOR DOCUMENTS ON SPECIFIED TOPICS   SEPTEMBER 28, 1964   PAGE 84

CURRENT REQUEST - *LIST DIFFERNTL EQ NUMERICAL DIGITAL SOLN OF DIFFERENTIAL EQUATIONS

ANSWER	CORRELATION	IDENTIFICATION
1 3RUNGE-KUT	0.5444	RUNGE-KUTTA METHODS FOR INTEGRATING DIFFERENTIAL EQUATIONS ON HIGH SPEED DIGITAL COMPUTERS D.W. MARTIN (NAT. PHYS. LAB., TEDDINGTON) THE COMPUTER J., VOL 1, PP 118-123, OCT., 1958
[illegible]NOTE ON AN	0.451[illegible]	NOTE ON ANALOG TECHNIQUES FOR RESOLVING TWO-POINT BOUNDARY VALUE PROBLEMS D. GREENSPAN (PURDUE UNIVERSITY) REV. SCI. INSTR. VOL 29 PP 787-788 (SEPTEMBER, 1958)
1 2SOLVING	0.41[illegible]6	SOLVING EQUATIONS OF STATE IN TELEPHONE TRAFFIC THEORY WITH DIGITAL COMPUTERS S.G. CARLSSON AND A. ELLDIN (TELEFONAKTIEBOLAGET LM ERICSSON) ERICSSON TECHNICS, VOL 14, NO. 2, PP 221-224, 1958
3 8STABILIZA	0.3986	STABILIZATION OF COMPUTER CIRCUITS ED. E. FOCHFELD (UNIV. OF CHICAGO) U. S. GOV. RES. REPTS. VOL 31 P 97(A) (FEBRUARY 13, 1959) PB 151 2[illegible]
1 2ON THE SO	[illegible].3986	ON THE SOLUTION OF POISSON-S DIFFERENCE EQUATION P. LAASONEN (UNIVERSITY OF CALIF.) J. ASSOC. COMPUTING MACH., VOL 5 PP 370 - 382, OCT., 1958
3 7BOUNDARY	[illegible].3968	BOUNDARY CONTRACTION SOLUTION OF LAPLACE-S DIFFERENTIAL EQUATION H. W. MILNES (G. M. CIRP. AND R. B. POTTS (UNIV. OF TORONTO) J. ASSOC. FOR COMPUTING MACH. VOL 6 PP 226-235 (APRIL, 1959)

Figure 22 (continued)

```
ANSWERS TO REQUESTS FOR DOCUMENTS ON SPECIFIED TOPICS          SEPTEMBER 28, 1964    PAGE 85

           CURRENT REQUEST - *LIST DIFFERNTL EQ NUMERICAL DIGITAL SOLN OF DIFFERENTIAL EQUATIONS


ANSWER                CORRELATION                        IDENTIFICATION
------                -----------                        --------------
202STABLE PR           0.3906                      STABLE PREDICTOR-CORRECTOR METHODS FOR ORDINARY DIFFERENTIAL
                                                   EQUATIONS
                                                   R.W. HAMMING (BELL TEL. LABS.)
                                                   J. ASSOC. FOR COMPUTING MACH., VOL 6, PP 37-47, JAN., 1959


ANSWER                CORRELATION                        IDENTIFICATION
------                -----------                        --------------
220MATRIX PR           0.3505                      MATRIX PROGRAMMING OF ELECTRONIC ANALOG COMPUTERS
                                                   R. E. HORN (WESTINGHOUSE ELEC. CORP.) AND P. M. HONNEL (WASHINGTON
                                                   UNIV.)
                                                   COMMUN. AND ELECTRONICS NO. 39 (AIEE TRANS. PT 1 VOL 77) PP 420-428
                                                   (SEPTEMBER, 1958)
```

Figure 22 (continued)

```
INSTRUCTION CARDS TO SMART SUPERVISOR                  SEPTEMBER 28, 1964    PAGE 340

           LIBRARY USED WAS VERSION NUMBER      2          OF THE HARRIS THESAURUS.
           THESAURUS DISCRIMINATES NOT MORE THAN      511   CONCEPTS.


STATISTICAL INTRA-DOCUMENT PROCESSING --

           NONE.


STATISTICAL PHRASES --
           PHRASE SEARCH WAS DONE.
           STATISTICAL PHRASES HAD WEIGHT OF    3.0


CRITERION TREES --

           NONE.


THE ABOVE DATA WAS SUPPLIED BY THE PROGRAMMER AND MAY BE INCORRECT.

THE FOLLOWING DATA IS FROM INSTRUCTIONS FOR THIS RUN WHICH DEFINITELY WERE EXECUTED.


TITLES WERE GIVEN A WEIGHT OF       1.0


CONCEPT OCCURRENCES BY DOCUMENT WERE PRINTED.


DOCUMENT CORRELATION --
           REQUEST CORRELATIONS WERE PRINTED.
           CORRELATION MODE USED WAS     COSINE.
           CUTOFF WAS  0.3500


HIERARCHY --

           NONE.


CONCEPT PROCESSING --

           NONE.


REQUESTS WERE ANSWERED.
AUTO-EVALUATION WAS REQUESTED AND WILL BE ATTEMPTED.
```

Specifications for Statistical Phrase Search

Figure 23

ANSWERS TO REQUESTS FOR DOCUMENTS ON SPECIFIED TOPICS   SEPTEMBER 28, 1964   PAGE 341

CURRENT REQUEST - *LIST DIFFERNTL EQ NUMERICAL DIGITAL SOLN OF DIFFERENTIAL EQUATIONS

REQUEST   *LIST DIFFERNTL EQ NUMERICAL DIGITAL SOLN OF DIFFERENTIAL EQUATIONS

GIVE ALGORITHMS USEFUL FOR THE NUMERICAL SOLUTION OF ORDINARY DIFFERENTIAL EQUATIONS AND PARTIAL DIFFERENTIAL EQUATIONS ON DIGITAL COMPUTERS . EVALUATE THE VARIOUS INTEGRATION PROCEDURES (E.G. RUNGE--KUTTA, MILNE-S METHOD) WITH RESPECT TO ACCURACY, STABILITY, AND SPEED .

ANSWER	CORRELATION	IDENTIFICATION
3[illegible]4STABILITY	0.8575	STABILITY OF NUMERICAL SOLUTION OF DIFFERENTIAL EQUATIONS
36[illegible]SIMULATIN	0.7741	SIMULATING SECOND-ORDER EQUATIONS
3[illegible]6ELIMINATI	0.7407	ELIMINATION OF SPECIAL FUNCTIONS FROM DIFFERENTIAL EQUATIONS
3920N COMPUT	0.6571	ON COMPUTING RADIATION INTEGRALS
2[illegible]SOLUTION	0.6443	SOLUTION OF ALGEBRAIC AND TRANSCENDENTAL EQUATIONS ON AN AUTOMATIC
3[illegible]NOTE ON AN	0.6372	NOTE ON ANALOG TECHNIQUES FOR RESOLVING TWO-POINT BOUNDARY VALUE
3[illegible]7BOUNDARY	0.6071	BOUNDARY CONTRACTION SOLUTION OF LAPLACE-S DIFFERENTIAL EQUATION
1[illegible]3RUNGE-KUT	0.5874	RUNGE-KUTTA METHODS FOR INTEGRATING DIFFERENTIAL EQUATIONS ON HIGH
1[illegible]70N THE SO	0.5648	ON THE SOLUTION OF POISSON-S DIFFERENCE EQUATION
3[illegible] MONTE CAR	0.5448	MONTE CARLO SOLUTIONS OF BOUNDARY VALUE PROBLEMS INVLVING THE
358STABILIZA	0.5436	STABILIZATION OF COMPUTER CIRCUITS
366A METHOD	0.5318	A METHOD OF NORMALIZED BLOCK ITERATION
2[illegible]2STABLE PR	0.5162	STABLE PREDICTOR-CORRECTOR METHODS FOR ORDINARY DIFFERENTIAL
365NUMERICAL	0.4941	NUMERICAL SOLUTION OF LAPLACE-S EQUATION - GIVEN CAUCHY CONDITIONS
1[illegible]9THEORETIC	0.4793	THEORETICAL CONSIDERATION OF COMPUTING ERRORS OF A SLOW TYPE
229MATRIX PR	0.4085	MATRIX PROGRAMMING OF ELECTRONIC ANALOG COMPUTERS
1[illegible]6SECANT MO	0.3937	SECANT MODIFICATION OF NEWTON-S METHOD
363A NOTE ON	0.3743	A NOTE ON THE DOWNHILL METHOD
4[illegible]A CALCULAT	0.3555	A CALCULATOR FOR NUMERICAL FOURIER SYNTHESIS

Answers After Statistical Phrase Search — BIG Format

Figure 24

ANSWERS TO REQUESTS FOR DOCUMENTS ON SPECIFIED TOPICS SEPTEMBER 28, 1964 PAGE 81

DIFFERNTL EQ	384STABILITY 360SIMULATIN 200SOLUTION 392ON COMPUT 386ELIMINATI 103RUNGE-KUT 85NOTE ON AN 192SOLVING E 358STABILIZA 102ON THE SO 387BOUNDARY 202STABLE PR 229MATRIX PR	REGULAR THESAURUS
DIFFERNTL EQ	103RUNGE-KUT 384STABILITY 392ON COMPUT 387BOUNDARY 386ELIMINATI 102ON THE SO 202STABLE PR 360SIMULATIN 192SOLVING E 200SOLUTION	NULL THESAURUS
DIFFERNTL EQ	103RUNGE-KUT 384STABILITY 202STABLE PR 360SIMULATIN 387BOUNDARY 386ELIMINATI 215DIGITAL C 88PROPOSED M 200SOLUTION 253ROUND-OFF 128COMPUTER	TITLES ONLY REGULAR THESAURUS
DIFFERNTL EQ	384STABILITY 217SIMULATIO 392ON COMPUT 103RUNGE-KUT 192SOLVING E 387BOUNDARY	LOGICAL VECTORS REGULAR THESAURUS
DIFFERNTL EQ	103RUNGE-KUT 392ON COMPUT 384STABILITY 360SIMULATIN 386ELIMINATI 200SOLUTION	REGULAR THESAURUS & HIERARCHICAL EXPANSION (UP & ADD)
DIFFERNTL EQ	384STABILITY 360SIMULATIN 386ELIMINATI 392ON COMPUT 200SOLUTION 85NOTE ON AN 387BOUNDARY 103RUNGE-KUT 102ON THE SO 39OMONTE CAR 358STABILIZA 388A METHOD 202STABLE PR 385NUMERICAL 169THEORETIC 229MATRIX PR 108SECANT MO 383A NOTE ON 45A CALCULAT	STATISTICAL PHASE SEARCH
DIFFERNTL EQ	384STABILITY 360SIMULATIN 386ELIMINATI 85NOTE ON AN 392ON COMPUT 200SOLUTION 103RUNGE-KUT 387BOUNDARY 385NUMERICAL 390MONTE CAR 248METHOD FO 252A CLASS O 358STABILIZA 202STABLE PR 45A CALCULAT 102ON THE SO 388A METHOD 169THEORETIC 229MATRIX PR 251ERROR EST	SYNTACTIC TREE MATCHING

Short Form Answers From Seven Processing Methods

Figure 25

```
ANSWERS TO REQUESTS FOR DOCUMENTS ON SPECIFIED TOPICS          SEPTEMBER 28, 1964   PAGE 243

   DIFFERNTL EQ     384STABILITY 360SIMULATIN 200SOLUTION  392ON COMPUT 386ELIMINATI
                    103RUNGE-KUT 85NOTE ON AN 192SOLVING E 358STABILIZA 102ON THE SC
                    387BOUNDARY  202STABLE PR 229MATRIX PR
```

Answers Before Hierarchical Expansion

```
ANSWERS TO REQUESTS FOR DOCUMENTS ON SPECIFIED TOPICS          SEPTEMBER 28, 1964   PAGE 253

   DIFFERNTL EQ     103RUNGE-KUT 392ON COMPUT 384STABILITY 360SIMULATIN 386ELIMINATI
                    200SOLUTION
```

Answers After Hierarchical Expansion

Answers Before and After Hierarchical Expansion

Figure 26

SEPTEMBER 28, 1964 PAGE 86

EVALUATION OF REQUEST DIFFERNTL EQ WITH 16 RELEVANT DOCUMENTS

THE TOP FIFTEEN DOCUMENTS				RELEVANT DOCUMENT RANKS		
1	X	384STABILITY	0.6676	1	384STABILITY	0.6676
2	X	360SIMULATIN	0.5758	2	360SIMULATIN	0.5758
3	X	200SOLUTION	0.5664	3	200SOLUTION	0.5664
4	X	392ON COMPUT	0.5508	4	392ON COMPUT	0.5508
5	X	386ELIMINATI	0.5484	5	386ELIMINATI	0.5484
6	X	103RUNGE-KUT	0.5445	6	103RUNGE-KUT	0.5445
7	X	85NOTE ON AN	0.4511	7	85NOTE ON AN	0.4511
8		192SOLVING E	0.4106	9	102ON THE SO	0.3987
9	X	102ON THE SO	0.3987	10	358STABILIZA	0.3986
10	X	358STABILIZA	0.3986	11	387BOUNDARY	0.3968
11	X	387BOUNDARY	0.3968	12	202STABLE PR	0.3907
12	X	202STABLE PR	0.3907	15	251ERROR EST	0.3329
13		229MATRIX PR	0.3506	17	253ROUND-OFF	0.3152
14		88PROPOSED M	0.3452	23	390MONTE CAR	0.2866
15	X	251ERROR EST	0.3329	24	388A METHOD	0.2788
				40	385NUMERICAL	0.2301

RANK RECALL= 0.7196 LOG PRECISION= 0.9169
NORMALIZED RECALL=0.9914626 NORMALIZED PRECISION=0.9572
RNK REC + LOG PRE=1.6365 WEIGHTED NORMED RECALL + NORMED PREC=1.9146

Evaluation – Regular Thesaurus

Figure 27

SEPTEMBER 28, 1964 PAGE 165

EVALUATION OF REQUEST DIFFERNTL EQ WITH 16 RELEVANT DOCUMENTS

THE TOP FIFTEEN DOCUMENTS				RELEVANT DOCUMENT RANKS		
1	X	103RUNGE-KUT	0.5509	1	103RUNGE-KUT	0.5509
2	X	384STABILITY	0.5462	2	384STABILITY	0.5462
3	X	392ON COMPUT	0.5271	3	392ON COMPUT	0.5271
4	X	387BOUNDARY	0.4309	4	387BOUNDARY	0.4309
5	X	386ELIMINATI	0.4294	5	386ELIMINATI	0.4294
6	X	102ON THE SO	0.4073	6	102ON THE SO	0.4073
7	X	202STABLE PR	0.3982	7	202STABLE PR	0.3982
8	X	360SIMULATIN	0.3922	8	360SIMULATIN	0.3922
9		192SOLVING E	0.3911	10	200SOLUTION	0.3672
10	X	200SOLUTION	0.3672	11	358STABILIZA	0.3441
11	X	358STABILIZA	0.3441	12	385NUMERICAL	0.3269
12	X	385NUMERICAL	0.3269	13	251ERROR EST	0.3203
13	X	251ERROR EST	0.3203	14	85NOTE ON AN	0.3162
14	X	85NOTE ON AN	0.3162	18	253ROUND-OFF	0.3014
15		229MATRIX PR	0.3045	68	388A METHOD	0.1682
				113	390MONTE CAR	0.1373

RANK RECALL= 0.4610 LOG PRECISION= 0.8806
NORMALIZED RECALL=0.9743879 NORMALIZED PRECISION=0.9360
RNK REC + LOG PRE=1.3416 WEIGHTED NORMED RECALL + NORMED PREC=1.8080

Evaluation — Null Thesaurus

Figure 28

```
                                          SEPTEMBER 28, 1964   PAGE 190

EVALUATION OF REQUEST DIFFERNTL EQ WITH  16 RELEVANT DOCUMENTS

    THE TOP FIFTEEN DOCUMENTS               RELEVANT DOCUMENT RANKS

      1 X 103RUNGE-KUT 0.7243                  1 103RUNGE-KUT 0.7243
      2 X 384STABILITY 0.7008                  2 384STABILITY 0.7008
      3 X 232STABLE PR 0.5560                  3 232STABLE PR 0.5560
      4 X 360SIMULATIN 0.5376                  4 360SIMULATIN 0.5376
      5 X 387BOUNDARY  0.4765                  5 387BOUNDARY  0.4765
      6 X 386ELIMINATI 0.4390                  6 386ELIMINATI 0.4390
      7   215DIGITAL C 0.3748                  9 200SOLUTION  0.3652
      8   88PROPOSED M 0.3725                 10 253ROUND-OFF 0.3573
      9 X 200SOLUTION  0.3652                 13 390MONTE CAR 0.3429
     10 X 253ROUND-OFF 0.3573                 14 385NUMERICAL 0.3421
     11   128COMPUTER  0.3519                 18 102ON THE SO 0.2986
     12   192SOLVING E 0.3439                 21 392ON COMPUT 0.2688
     13 X 390MONTE CAR 0.3429                 23 251ERROR EST 0.2634
     14 X 385NUMERICAL 0.3421                 61 358STABILIZA 0.1626
     15   306ELECTRONI 0.3226                 75 85NOTE ON AN 0.1458
                                             303 388A METHOD  0.0000

RANK RECALL= 0.2394 LOG PRECISION= 0.7766
NORMALIZED RECALL=0.9304124 NORMALIZED PRECISION=0.8643
RNK REC + LOG PRF=1.0160 WEIGHTED NORMED RECALL + NORMED PREC=1.5163
```

Evaluation — Titles Only

Figure 29

```
                                             SEPTEMBER 28, 1964   PAGE 218

EVALUATION OF REQUEST DIFFERNTL EQ WITH  16 RELEVANT DOCUMENTS

    THE TOP FIFTEEN DOCUMENTS                RELEVANT DOCUMENT RANKS

     1 X 384STABILITY 0.6121                  1 384STABILITY 0.6121
     2   217SIMULATIO 0.3800                  4 392ON COMPUT 0.3690
     4 X 392ON COMPUT 0.3690                  4 103RUNGE-KUT 0.3690
     4 X 103RUNGE-KUT 0.3690                  6 387BOUNDARY  0.3603
     5   192SOLVING E 0.3656                  7 202STABLE PR 0.3430
     6 X 387BOUNDARY  0.3603                 14 386ELIMINATI 0.3034
     7 X 202STABLE PR 0.3430                 15 358STABILIZA 0.2943
     8   226+DEPI ..  0.3378                 17 85NOTE ON AN 0.2854
     9   383A NOTE ON 0.3241                 20 253ROUND-OFF 0.2750
    10   74DIGITAL CO 0.3153                 21 200SOLUTION  0.2684
    11   18AN ACCURAT 0.3100                 28 360SIMULATIN 0.2572
    12   128COMPUTER  0.3049                 40 390MONTE CAR 0.2334
    14 X 386ELIMINATI 0.3034                 49 251ERROR EST 0.2214
    14   234ANALOGUE  0.3034                 59 388A METHOD  0.2068
    15 X 358STABILIZA 0.2943                 71 385NUMERICAL 0.1980
                                             78 102ON THE SO 0.1917

RANK RECALL= 0.3134 LOG PRECISION= 0.6917
NORMALIZED RECALL=0.9519974 NORMALIZED PRECISION=0.7897
RNK REC + LOG PRE=1.0051 WEIGHTED NORMED RECALL + NORMED PREC=1.5497
```

Evaluation – Logical Vectors

Figure 30

```
                                           SEPTEMBER 28, 1964   PAGE 256

EVALUATION OF REQUEST DIFFERNTL EQ WITH  16 RELEVANT DOCUMENTS

    THE TOP FIFTEEN DOCUMENTS               RELEVANT DOCUMENT RANKS

      1 X 103RUNGE-KUT 0.4567                1 103RUNGE-KUT 0.4567
      2 X 392ON COMPUT 0.4365                2 392ON COMPUT 0.4365
      3 X 384STABILITY 0.4261                3 384STABILITY 0.4261
      4 X 360SIMULATIN 0.3970                4 360SIMULATIN 0.3970
      5 X 386ELIMINATI 0.3827                5 386ELIMINATI 0.3827
      6 X 200SOLUTION  0.3677                6 200SOLUTION  0.3677
      7   234ANALOGUE  0.3455                9 387BOUNDARY  0.3224
      8   128COMPUTER  0.3239               12 202STABLE PR 0.3123
      9 X 387BOUNDARY  0.3224               13 358STABILIZA 0.3101
     10   226+DEPI ..  0.3171               16 85NOTE ON AN 0.3016
     11   45A CALCULAT 0.3127               21 385NUMERICAL 0.2776
     12 X 202STABLE PR 0.3123               31 102ON THE SO 0.2580
     13 X 358STABILIZA 0.3101               53 251ERROR EST 0.2123
     14   306ELECTRONI 0.3042               55 253ROUND-OFF 0.2082
     15   192SOLVING & 0.3020               88 390MONTE CAR 0.1729
                                            94 388A METHOD  0.1682

RANK RECALL= 0.3293 LOG PRECISION= 0.7653
NORMALIZED RECALL=0.9553801 NORMALIZED PRECISION=0.8553
RNK REC + LOG PRE=1.0946 WEIGHTED NORMED RECALL + NORMED PREC=1.6322
```

Evaluation — Hierarchy

Figure 31

SEPTEMBER 28, 1964 PAGE 342

EVALUATION OF REQUEST DIFFERNIL EQ WITH 16 RELEVANT DOCUMENTS

THE TOP FIFTEEN DOCUMENTS	RELEVANT DOCUMENT RANKS
1 X 384STABILITY 0.8576	1 384STABILITY 0.8576
2 X 360SIMULATIN 0.7741	2 360SIMULATIN 0.7741
3 X 386ELIMINATI 0.7408	3 386ELIMINATI 0.7408
4 X 392ON COMPUT 0.6571	4 392ON COMPUT 0.6571
5 X 200SOLUTION 0.6444	5 200SOLUTION 0.6444
6 X 85NOTE ON AN 0.6372	6 85NOTE ON AN 0.6372
7 X 387BOUNDARY 0.6072	7 387BOUNDARY 0.6072
8 X 103RUNGE-KUT 0.5875	8 103RUNGE-KUT 0.5875
9 X 102ON THE SO 0.5648	9 102ON THE SO 0.5648
10 X 390MONTE CAR 0.5448	10 390MONTE CAR 0.5448
11 X 358STABILIZA 0.5437	11 358STABILIZA 0.5437
12 X 388A METHOD 0.5318	12 388A METHOD 0.5318
13 X 202STABLE PR 0.5163	13 202STABLE PR 0.5163
14 X 385NUMERICAL 0.4942	14 385NUMERICAL 0.4942
15 169THEORETIC 0.4794	21 251ERROR EST 0.3444
	25 253ROUND-OFF 0.3157

RANK RECALL= 0.9007 LOG PRECISION= 0.9751
NORMALIZED RECALL=0.9975900 NORMALIZED PRECISION=0.9880
RNK REC + LOG PRE=1.8758 WEIGHTED NORMED RECALL + NORMED PREC=1.9759

Evaluation — Statistical Phrases

Figure 32

SEPTEMBER 28, 1964 PAGE 380

EVALUATION OF REQUEST DIFFERNTL EQ WITH 16 RELEVANT DOCUMENTS

THE TOP FIFTEEN DOCUMENTS				RELEVANT DOCUMENT RANKS		
1	X	384STABILITY	0.7792	1	384STABILITY	0.7792
2	X	360SIMULATIN	0.6155	2	360SIMULATIN	0.6155
3	X	386ELIMINATI	0.6078	3	386ELIMINATI	0.6078
4	X	85NOTE ON AN	0.5951	4	85NOTE ON AN	0.5951
5	X	392ON COMPUT	0.5737	5	392ON COMPUT	0.5737
6	X	200SOLUTION	0.5736	6	200SOLUTION	0.5736
7	X	103RUNGE-KUT	0.5593	7	103RUNGE-KUT	0.5593
8	X	387BOUNDARY	0.5547	8	387BOUNDARY	0.5547
9	X	385NUMERICAL	0.5179	9	385NUMERICAL	0.5179
10	X	390MONTE CAR	0.5053	10	390MONTE CAR	0.5053
11		248METHOD FO	0.4620	13	358STABILIZA	0.4587
12		252A CLASS O	0.4596	14	202STABLE PR	0.4582
13	X	358STABILIZA	0.4587	16	102ON THE SO	0.4273
14	X	202STABLE PR	0.4582	17	388A METHOD	0.4205
15		45A CALCULAT	0.4417	20	251ERROR EST	0.3698
				26	253ROUND-OFF	0.2866

RANK RECALL= 0.8447 LOG PRECISION= 0.9535
NORMALIZED RECALL=0.9959833 NORMALIZED PRECISION=0.9770
RNK REC + LOG PRE=1.7982 WEIGHTED NORMED RECALL + NORMED PREC=1.9569

Evaluation — Syntactic Phrases

Figure 33