1968

A polar space for document storage and retrieval

Joseph Paul Anthony Lombardo
Lehigh University

A POLAR SPACE

for

DOCUMENT STORAGE AND RETRIEVAL

by

Joseph Paul Anthony Lombardo

A Thesis

Presented to the Graduate Faculty

of Lehigh University

in Candidacy for the Degree of

Master of Science

in

Information Science

Lehigh University

1968

This Thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

September 13, 1968
(date)

Donald J. Hillman
Professor in Charge

Donald J. Hillman
Head of the Department

Acknowledgements

This Author wishes to thank Professor Donald Hillman for his suggestions and guidance in preparing this report.

Thanks are also extended to Ingersoll Rand Co. for their financial aid in the pursuance of the Master's Program.

Mr. Robert Pacific's typing efforts are also appreciated.

The Author's family,amd,friends,they all deserve a note of thanks for their patience and inspiration.

Table of Contents

	Page
Certificate of Approval	ii
Acknowledgements	iii
Table of Contents	iv
List of Figures	v
Abstract	1
Introduction	3
Theory	5
Description of Document Space in Polar Coordinates	7
Implementation	16
Utilization of System	17
Appendix	21
Vita	25

List of Figures

Figure		Page
1	List of Documents with Index Terms Including Document Density and Associated Spatial Locations	8
2	Documents Located in Polar Space	12
3	Computerized Output of Index Terms Obtained from Document Densities	13
4	Computerized Output of Densities Obtained from Specified Polar Spatial Locations	15
5	Typical Representations of Documents	18

Abstract

In a manual retrieval system, the process of associating a book or document with a specified term reflects the recall, associativity and flexibility of the human mind. Someone will enter a library and, by merely providing the librarian with a single word, will expect a book on a specific topic. To the librarian, this one word will bring to mind a specific topic which, in turn, will be associated with a unique field of endeavor. All that remains is for the librarian to go to the proper area of the library and retrieve the desired book. With the tremendous increase in the number of documents, it becomes necessary to design to automated mechanized retrieval systems.

It is hopeful that the mechanized systems will function in identically the same manner as the manual system.

The first step towards achieving this parallel between the manual system and mechanized system is to be able to uniquely define and catagorize all documents. In this report, the definition is achieved by calculating the density of a document. The density is a function of the index terms used to characterize the document. The specific index terms and the ordering of them relative to their importance in characterizing a document is preserved when calculating the density. The catagorization is accomplished by first defining a specific type of space and then assigning locations for the documents within it. The space selected is the polar space and the document locations are assigned by means of a radius and an angle which go towards making up the polar space. The locations reflect the document density values and as a result, form clusters which represent various topics common to all documents within the clusters. The size of the clusters can be determined so as to

fulfill the needs of the users of the system.

Included in this report are a number of computerized programming efforts which will enable the calculation of density values if given the spacial locations, and which will facilitate the identification of index terms when the density functions are known.

I. INTRODUCTION

If two requests R_1 & R_2 contain the same index terms i_1 & i_2, one might suspect that the same document should be retrieved for each request. Further negotiations do reveal that the person submitting R_1 is really interested in a document about the topic described by i_1 with just a very light treatment on topic covered by i_2 whereas the individual submitting R_2 is interested in the converse. There exists only a remote possibility that the same document would be adequate and suitable for both requests. In making a verbal request, one certainly could specify a preference of topics which, in turn, could instigate a search in a special area of a library. This last statement presupposes that a library does, in fact, exist and, in addition, the distribution and location of elements (books, documents, papers, etc.) are accomplished with utmost care. In a special type of system, one might find a library divided into many sectors each of which would supply the needs of a particular group (e.g. the mathematician would be directed to one sector, the horticulturist to another, the poet to still another, etc.). From a practical sense, this has a major disadvantage in that a great amount of duplication of elements would occur; another disadvantage one might forsee would result from the "gray" area or "overlapping" area where it would be difficult to decide on the sector for certain elements. A book which placed equal emphasis on "mathematics" and "physics" might provide a difficulty in deciding to which sector it belongs.

A more idealistic storage system might be one in which the formation of the sectors is a function of the elements stored therein. In this system, each element is defined in terms of its characteristics, called index terms, and the boundaries of the various sectors are established on the basis of the requirements of the users. An element common to mathematicians and to physicists would be accessible by both groups and, likewise, an element common to only one group would be accessible by only that one group. One decided advantage of this system would be the property of having flexible boundaries. These would be defined on the basis of the range of the subject material as specified by the user. One would approach this system not from the standpoint of saying, "Direct me to the mathematician's area of the library because I am a mathematician" but rather "Direct me to the area defined on the basis of index terms which suit my requirements and which I will specify". The index terms and the retrieved element from this library will reflect the unique aspects of both the professional endeavor of the person making the request and, also, the environment within which his request occurs.

This report will explore a specific mathematical relationship and show that it exhibits the qualities of the system described above. It will also be shown that the elements of a document retrieval system can be associated with the elements of the mathematical relationship.

II. THEORY

The retrieval system which will be observed is made up of elements called index terms and documents. Given a collection called "C", it contains a set of documents "D" and a set of index terms I.

$$C=D=\sum_{n=1}^{N} d_n; \qquad I=\sum_{m=1}^{M} i_m$$

N= Total number of documents in the collection

M= Total number of index terms used to characterize all the documents in the collection.

For every document "d_n" in the collection D, there exists a unique set of index terms which describes or identifies the contents of that particular document. In order to properly identify the unique set of index terms, a measure of density "P_n" is assigned to each document d_n.

$$P_n = \sum_{j=1}^{K_n} S_{j_n} \left[(M) + 1\right]^{j-1} = \text{density of } d_n$$

K_n= Number of index terms assigned to d_n.

S_{j_n}= Subscripts of index terms assigned to d_n

e.g. given d_2 indexed by i_3, i_7, i_8, i_{12} there exists a K_2= 4 and $S_{1_2} = 3$; $S_{2_2} = 7$; $S_{3_2} = 8$; $S_{4_2} = 12$.

The qualifications for the density function are as follows:

1. Given K_n where $S_{g_n} = \ell$ for i_ℓ and also when $S_{f_n} = t$ for i_t then $S_{g_n} \neq S_{f_n} \longrightarrow i_\ell \neq i_t$
2. $i_a = i_b \longrightarrow S_{g_n} = S_{f_n}$ & a=b.
3. Given $P_1 = \sum_{j=1}^{K_1} S_{j_1}$ $(M)+1^{j-1}$ where $S_{1_1} = G$, $S_{1_2} = H$; & $P_2 = \sum_{j=1}^{K_2} S_{j_2} \left[(M)+1\right]^{j-1}$ where $S_{2_1} = H$; $S_{2_2} = G$ then $S_{1_1} \neq S_{2_1}$, $S_{1_2} \neq S_{2_1}$, $S_{1_1} = S_{2_2}$, $S_{1_2} = S_{2_1} \longrightarrow P_1 \neq P_2$.
4. Given $P_1 = \sum_{j=1}^{K_1} S_{j_1} \left[(M)+1\right]^{j-1}$ where $S_{1_1} = G$ & $P_2 = \sum_{j=1}^{K_2} S_{j_2} \left[(M)+1\right]^{j-1}$ where $S_{2_1} = H$; $K_1 = K_2 = 1$ then $G = H \longrightarrow P_1 = P_2$ also $G \neq H \longrightarrow P_1 \neq P_2$.
5. $P_L = P_G \longrightarrow L = G$ & $d_L = d_G$.
6. For $K_n = 1$ then $P_n = S_{1_n}$.

Items 1 and 2 reflect the unique assignments of subscripts to index terms.

From item 3, one recognizes that the ordering of index terms within the density function is important. Since ordering of index terms which characterize documents is a major consideration, and also, since uniformity in interpretation is desirable, it is imperative that an assignment of the convention for ordering be established. It shall be as follows:

For $K_n > 1$; S_{1_n} will have less to say about d_n than S_{2_n} which will in turn have less to say about d_n than S_{3_n} , likewise for each successive S_{j_n} until $S(K_n)_n$ is reached. In a multi-index term document, the right most index term reflects the most prominent characteristic of the document.

Item 4 discloses the property that each index term is associated with its own peculiar density.

To recap, given a set of index terms I such that each of its elements is assigned a unique position within that set, the numeric designation of that position functions as a subscript in the notation for the index term. There exist subsets of index terms which will serve as identifiers of documents d_n belonging to the collection D. Any member of the subset is also a member of the set I. A mathematical evaluation of the subset associated with d_n produces a numeric label for that specific document. This label called "density" is unique for each document in the collection; it also reflects the relative importance of each term in the subset.

III. DESCRIPTION OF A DOCUMENT SPACE IN POLAR COORDINATES

The next consideration of this report is to be able to prescribe a space which will encompass all document densities of the collection D and which will have properties such that each documents distinguishability will be maintained. At the same time, it is highly desirable to preserve the "topic-relationship" which may exist within the collection.

The space which will be considered for this report is the two-dimensional space defined by the polar coordinates of "radius" and "angular displacement". The relationships between the document locations in the polar space and the polar coordinates are as follows:

r_n = radius relative to the center of the polar space $= \sum_{j=1}^{M} \delta_j$

$\delta_j = 1$ if i_j is used to characterize d_n.

$\delta_j = 0$ if i_j is not used to characterize d_n.

θ_n = angular displacement $= \left\{360/\left[(M+1)^{r_n-1}(M)\right]\right\}(P_n-(M+1)^{r_n-1})$.

Consider a collection of documents "C" with N=64 & M=4. In this collection, there are 4 index terms i_1, i_2, i_3, i_4, also, there are 64 documents characterized by a combination of the 4 index terms. For each document, the density function is evaluated.

See figure 1 for a complete listing of all associated values. The values for this table have been obtained by means of an IBM S/360 Model 30 computer.

The classification and evaluation of documents is an uncomplicated process. Each document of the collection is assigned a unique number from the set of numbers 1 to N where N represents the number of documents in the collection. The index terms which are used to characterize all documents in the collection are assigned numbers from the set of numbers 1 to M where M is equal to the total number of index terms used

// EXEC

FIGURE 1

List of Documents with Index Terms Including
Document Density and Associated Spatial Locations

n	K_n	S_{n_1}	S_{n_2}	S_{n_3}	S_{n_4}	P_n	r_n	θ_n (DEGREES)
1	1	1	0	0	0	1	1	0.0
2	1	2	0	0	0	2	1	90.0000
3	1	3	0	0	0	3	1	180.0000
4	1	4	0	0	0	4	1	270.0000
5	2	2	1	0	0	7	2	36.0000
6	2	3	1	0	0	8	2	54.0000
7	2	4	1	0	0	9	2	72.0000
8	2	1	2	0	0	11	2	108.0000
9	2	3	2	0	0	13	2	144.0000
10	2	4	2	0	0	14	2	162.0000
11	2	1	3	0	0	16	2	198.0000
12	2	2	3	0	0	17	2	216.0000
13	2	4	3	0	0	19	2	252.0000
14	2	1	4	0	0	21	2	288.0000
15	2	2	4	0	0	22	2	306.0000
16	2	3	4	0	0	23	2	324.0000
17	3	3	2	1	0	38	3	46.8000
18	3	4	2	1	0	39	3	50.4000
19	3	2	3	1	0	42	3	61.2000
20	3	4	3	1	0	44	3	68.4000
21	3	2	4	1	0	47	3	79.2000
22	3	3	4	1	0	48	3	82.8000
23	3	3	1	2	0	58	3	118.8000
24	3	4	1	2	0	59	3	122.4000
25	3	1	3	2	0	66	3	147.6000
26	3	4	3	2	0	69	3	158.4000
27	3	1	4	2	0	71	3	165.6000
28	3	3	4	2	0	73	3	172.8000
29	3	2	1	3	0	82	3	205.2000
30	3	4	1	3	0	84	3	212.4000
31	3	1	2	3	0	86	3	219.6000
32	3	4	2	3	0	89	3	230.4000
33	3	1	4	3	0	96	3	255.5999
34	3	2	4	3	0	97	3	259.2000
35	3	2	1	4	0	107	3	295.2000
36	3	3	1	4	0	108	3	298.7998
37	3	1	2	4	0	111	3	309.5999
38	3	3	2	4	0	113	3	316.7998
39	3	1	3	4	0	116	3	327.5999
40	3	2	3	4	0	117	3	331.1997
41	4	4	3	2	1	194	4	49.6800
42	4	3	4	2	1	198	4	52.5600
43	4	4	2	3	1	214	4	64.0800
44	4	2	4	3	1	222	4	69.8400
45	4	3	2	4	1	238	4	81.3600
46	4	2	3	4	1	242	4	84.2400
47	4	4	3	1	2	294	4	121.6800
48	4	3	4	1	2	298	4	124.5600
49	4	4	1	3	2	334	4	150.4800
50	4	1	4	3	2	346	4	159.1200
51	4	3	1	4	2	358	4	167.7600
52	4	1	3	4	2	366	4	173.5200

FIGURE 1 - (CONTINUED)

53	4	4	2	1	3	414	4	208.0800
54	4	2	4	1	3	422	4	213.8400
55	4	4	1	2	3	434	4	222.4800
56	4	1	4	2	3	446	4	231.1200
57	4	2	1	4	3	482	4	257.0398
58	4	1	2	4	3	486	4	259.9199
59	4	3	2	1	4	538	4	297.3599
60	4	2	3	1	4	542	4	300.2397
61	4	3	1	2	4	558	4	311.7598
62	4	1	3	2	4	566	4	317.5198
63	4	2	1	3	4	582	4	329.0398
64	4	1	2	3	4	586	4	331.9199

in the collection. Each document number, with its associated index-term numbers, is recorded; utilizing these numbers, the density function and the polar coordinates can be calculated. There need be no prior relationship between the document and the document numbering system, however, once the numbers have been assigned them, they must be retained. Referring to figure 1, it can be said that there exists a document referenced by the number 1. Associated with this document is an index term with subscript 1. No other document in the collection will have an assignment which is identical to this one. Because of this unique assignment of index-term numbers, document densities can be calculated which will possess the property of distinguishability. Again referring to figure 1, it can be observed that all the density values are distinguishable one from the other.

With the addition of new documents, it will not always be necessary to re-evaluate the entire collection. If the new document is characterized by index terms which are already included in the collection, the only calculation which need be performed is that of evaluating the density of the added document. This is quite obvious when one observes that the density is a function of the total number of index terms in the collection. When new index terms are added, then it becomes necessary to recalculate all the density values. This may appear to be a cumbersome chore if the collection is quite large and the addition of index terms is frequent; however, with computerized programming, this becomes less burdensome than one would imagine. One way of minimizing this recalculation effort would be to anticipate new additions by introducing pseudo-terms into the system. Then, as new terms are to be added, they can replace the pseudo-terms. In the sample problem, it appears as if 4 index terms are used to characterize the documents in the collection; however, it is conceivable that there could exist only two terms i_1, and i_2 with i_3 and i_4 serving as pseudo-terms. All the calculations for a 4 term system are performed even though it contains only documents of a two term system. With the advent of 2 additional terms, i_3 & i_4 can then be reassigned to represent the two new terms, thus forming a four term system whose numeric calculations have already been performed.

Figure 2 is a representation of the polar space with the locations of each document in the sample problem. It can be observed that each document of the collection has a unique location in this polar space. The locations, just as the densities, are distinguishable. The range of the space coordinates are as follows: the radius ranges from values of 1 to K_{max} where K_{max} is equal to the maximum number of index terms used to characterize a document. All values of radii are integer numbers. The values of "θ" range between 0 and 2 .

FIGURE 2 Documents Located in Polar Space

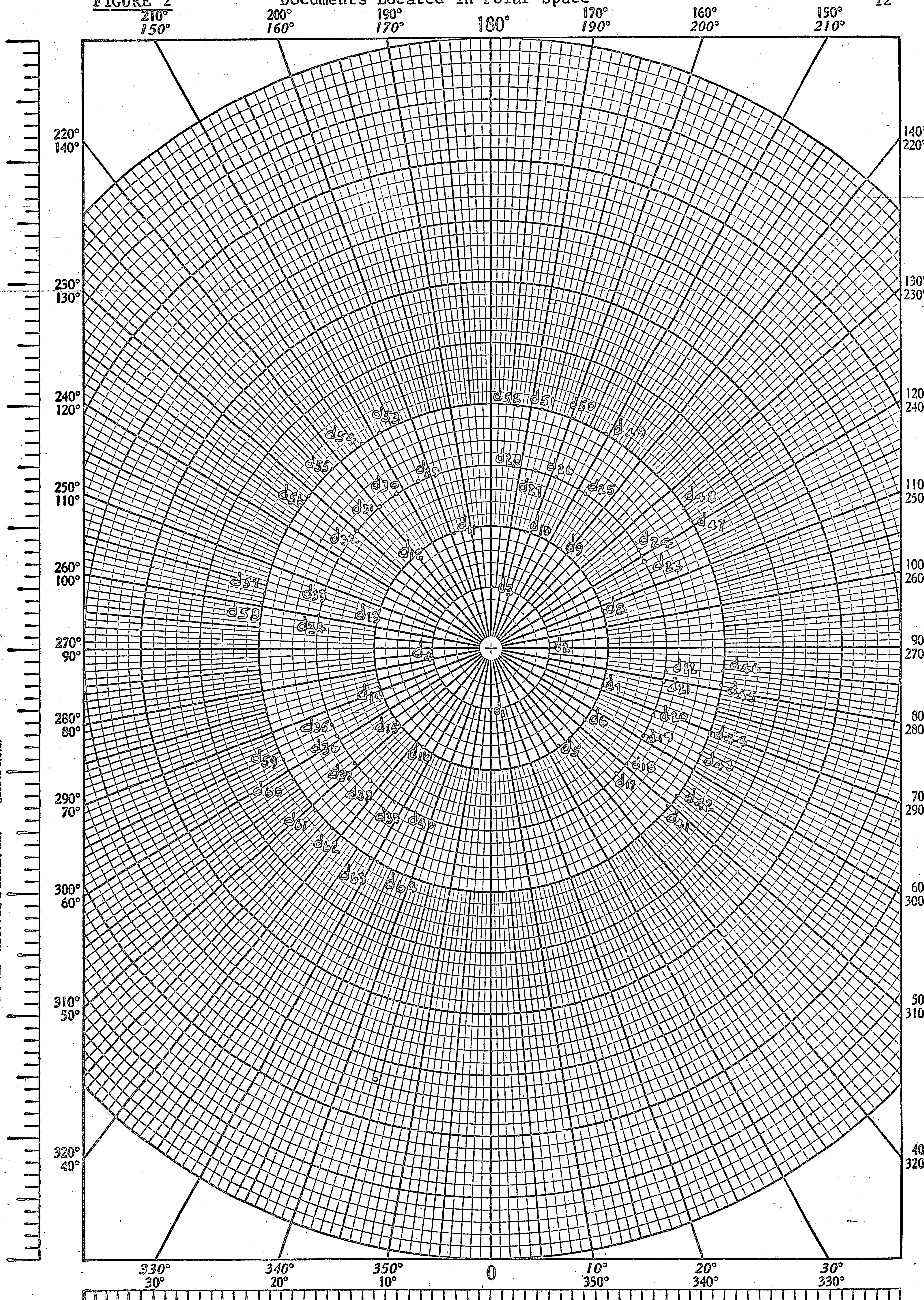


FIGURE 3
Computerized Output of Index Terms
Obtained from Document Densities

```
// EXEC
NUMBER OF INDEX TERMS IN COLLECTION =     4     DOCUMENT DENSITY = 111.

SN 1 =     1
SN 2 =     2
SN 3 =     4
NUMBER OF INDEX TERMS IN COLLECTION =     4     DOCUMENT DENSITY = 222.

SN 1 =     2
SN 2 =     4
SN 3 =     3
SN 4 =     1
NUMBER OF INDEX TERMS IN COLLECTION =     4     DOCUMENT DENSITY = 366.

SN 1 =     1
SN 2 =     3
SN 3 =     4
SN 4 =     2
NUMBER OF INDEX TERMS IN COLLECTION =     4     DOCUMENT DENSITY = 482.

SN 1 =     2
SN 2 =     1
SN 3 =     4
SN 4 =     3
NUMBER OF INDEX TERMS IN COLLECTION =     4     DOCUMENT DENSITY = 194.

SN 1 =     4
SN 2 =     3
SN 3 =     2
SN 4 =     1
NUMBER OF INDEX TERMS IN COLLECTION =     4     DOCUMENT DENSITY = 358.

SN 1 =     3
SN 2 =     1
SN 3 =     4
SN 4 =     2
NUMBER OF INDEX TERMS IN COLLECTION =     4     DOCUMENT DENSITY = 538.

SN 1 =     3
SN 2 =     2
SN 3 =     1
SN 4 =     4
NUMBER OF INDEX TERMS IN COLLECTION =     4     DOCUMENT DENSITY = 414.

SN 1 =     4
SN 2 =     2
SN 3 =     1
SN 4 =     3
NUMBER OF INDEX TERMS IN COLLECTION =     4     DOCUMENT DENSITY = 406.
```

```
SN 1 =      1
SN 2 =      1
SN 3 =      1
SN 4 =      3
NUMBER OF INDEX TERMS IN COLLECTION =      4      DOCUMENT DENSITY = 486.

SN 1 =      1
SN 2 =      2
SN 3 =      4
SN 4 =      3
NUMBER OF INDEX TERMS IN COLLECTION =      4      DOCUMENT DENSITY =   1.

SN 1 =      1
NUMBER OF INDEX TERMS IN COLLECTION =      4      DOCUMENT DENSITY =   2.

SN 1 =      2
NUMBER OF INDEX TERMS IN COLLECTION =      4      DOCUMENT DENSITY =   3.

SN 1 =      3
NUMBER OF INDEX TERMS IN COLLECTION =      4      DOCUMENT DENSITY =  44.

SN 1 =      4
SN 2 =      3
SN 3 =      1
NUMBER OF INDEX TERMS IN COLLECTION =      4      DOCUMENT DENSITY =  97.

SN 1 =      2
SN 2 =      4
SN 3 =      3
NUMBER OF INDEX TERMS IN COLLECTION =      4      DOCUMENT DENSITY =  99.

SN 1 =      4
SN 2 =      4
SN 3 =      3

  IJT219I
```

FIGURE 4

Computerized Output of Densities Obtained from Specified Polar Spatial Locations


```
// EXEC

      1.0       0.0       1.000
  DENSITY     ANGLE      RADIUS

      8.0    54.000       2.000
  DENSITY     ANGLE      RADIUS

     23.0   324.000       2.000
  DENSITY     ANGLE      RADIUS

     58.0   118.800       3.000
  DENSITY     ANGLE      RADIUS

     86.0   219.600       3.000
  DENSITY     ANGLE      RADIUS

     97.0   259.200       3.000
  DENSITY     ANGLE      RADIUS

    116.0   327.600       3.000
  DENSITY     ANGLE      RADIUS

    242.0    84.240       4.000
  DENSITY     ANGLE      RADIUS

    334.0   150.480       4.000
  DENSITY     ANGLE      RADIUS

    414.0   208.080       4.000
  DENSITY     ANGLE      RADIUS

    446.0   231.120       4.000
  DENSITY     ANGLE      RADIUS

    542.0   300.240       4.000
  DENSITY     ANGLE      RADIUS

    586.0   331.920       4.000
  DENSITY     ANGLE      RADIUS

  IJT219I
```

IV. IMPLEMENTATION

Now that the numbering system and associated space locations have been defined, it would be highly desirable to investigate the possibility of programming for a computer, the means of a transformation of values from the numbering system to the spatial locations. If given a specific document density, one would wish to be able to arrive at the subscripts of the index terms used to characterize the document. It would be desirable to obtain the density of a document if its polar space location was known. Two programs were written to accomplish these goals. Figure 3 is a representation of a listing obtained from one of the programs, the one which determines the subscripts of the index terms for a document whose density is specified. In addition to providing the identity of the index terms, the program also preserves the ordering of the terms.

The second program was designed so as to be capable of calculating the density of a document whose location in polar space was known. See figure 4 for a sample of the output from this program. One merely need specify the radius and angle of the document location; this information will lead to the determination of the density value; Both programs were written in "FORTRAN IV"; they were assembled and executed on an IBM S/360 Model 30.

V. UTILIZATION OF SYSTEM

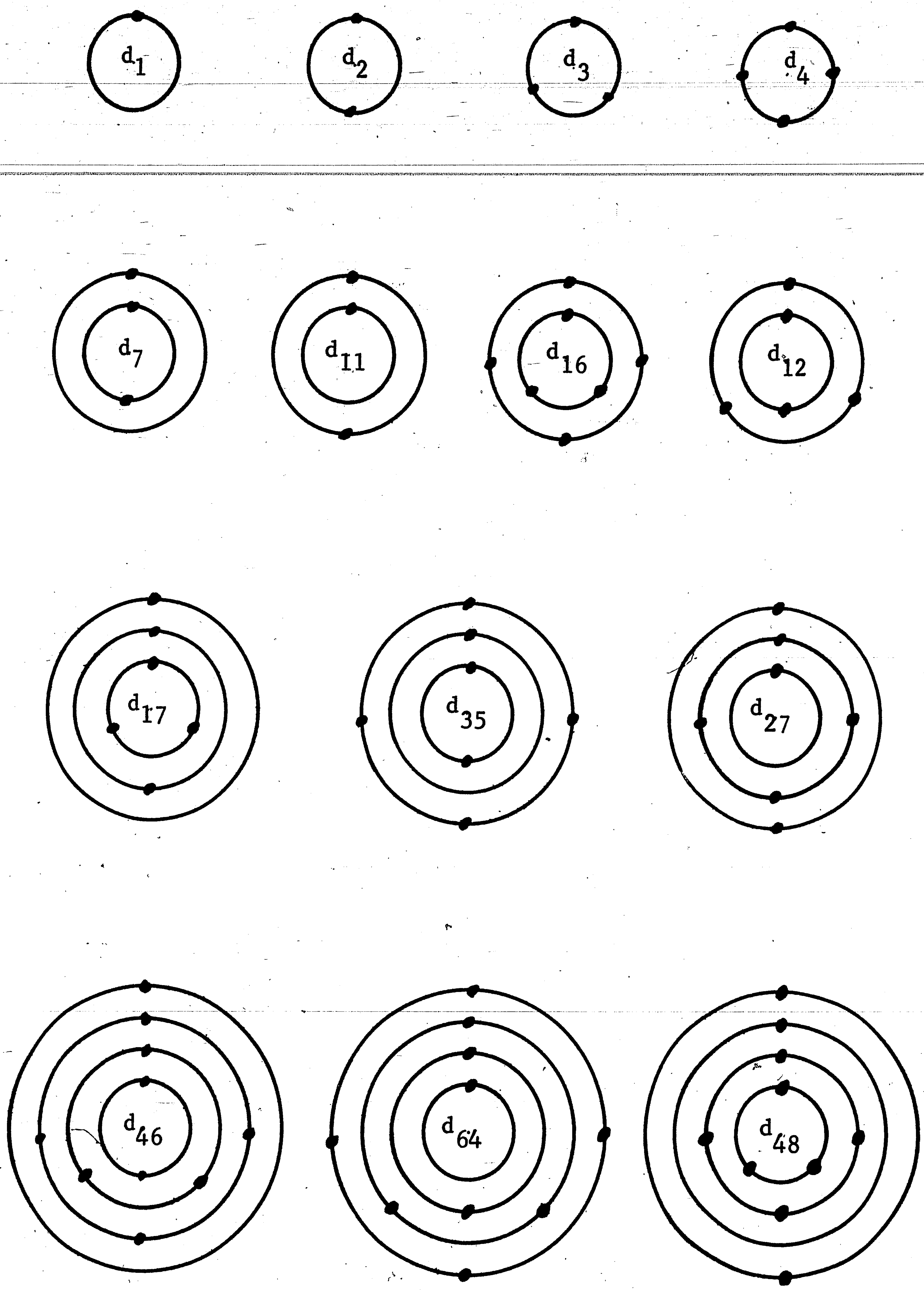
The balance of this report will concern itself with the utility and potential of the density and polar space used to define and locate documents. The distinguishability of both density and spacial locations transmits its characteristics to the document. This eliminates the occurance of any ambiguity in the identification of documents.

The density of any document is a physical characteristic which is atomic in structure. Given a document, it can be physically represented by a series of concentric circles. Each circle contains elements which represent one of the index terms used to characterize that document. A document with density equal to 3, which means it is characterized by one index term whose subscript is equal to 3, will have one ring on which are located 3 electron type elements equally spaced. The maximum number of elements per ring is equal to the maximum number of index terms in the collection. See figure 5 for some examples of document density representations. This type of structure can open the door for future research relative to combinations of documents to determine similarity of topics, dissimilarity of documents, etc. The principle concern of this paper is the definition of the system and an investigation of the potential which exists in the spatial representations of documents.

After reviewing the results which appear in figures 1 and 2, certain interesting properties are observed. On the circle with radius equal to 1, the four documents characterized by a single index term are spaced ½ radians apart. It can then say that index term 1 occupies the first quadrant of the polar space. After examining all

FIGURE 5

TYPICAL REPRESENTATION OF DOCUMENTS



documents which exist within this first quadrant, one notices that all of these documents have more to say about index term 1 than about any of the other index terms. A similar observation can be made about index term 2 and the documents in the 2nd quadrant; likewise for each of the remaining index terms and their associated quadrants. To generalize, the following can be stated: for M index terms, the polar space can be divided into M sectors. Each sector will contain documents which emphasize the index term on ring of radius 1 in that sector. As one travels from ring to ring, sub-divisions of sectors exhibit similar properties. The document locations can be considered as clusters where each cluster deals with a specific topic. The establishment of flexible boundaries, which can be adjusted as the needs occur, becomes quite feasible. Let us say that the test collection referred to previously, represents documents written about 4 distinct index terms each of which has associated with it, a special field of endeavor. That is to say that if i_1 is a mathematical term and i_2 is a "physics" term, then documents dealing with i_1 and i_2 will be of interest to the mathematician if the document is more concerned about i_1 than i_2 ; however, if the document has more to say about i_2 , then it will be of greater value to the "physics" minded individual. On the basis of this, the boundaries of the system can be established by the needs of the user.

It is highly conceivable that the ring of radius 1 can be used as a browsing or exploratory region. Since all documents on this ring are indexed by a single term, one could define, very easily, specific topics and their associated sectors. The appealing aspect of this procedure is that the documents characterized by one index term need not even exist. The entire space can be generated on the basis of known index terms regardless of whether documents exist or not. As the documents become available, they can be inserted into their pre-defined locations. A by-product of this system could be in the form of a service which informs the users or managers of the collection that a deficiency of available documents exists. The most important property of this system is that a natural grouping of documents does occur and that this grouping reflects the ordering of index terms. The size of each group can be tailored to the needs of the individual.

Appendix

The following three program listings represent the programs written in FORTRAN IV for the IBM S/360 Model 30 computer.

The first program's objective is to prepare a complete listing of all possible combinations of M index terms belonging to a collection of documents. In addition, all corresponding densities and polar coordinates are calculated.

The second program has as its objective, the evaluation of all index terms associated with a document whose density is specified. The order of terms as prescribed for the document is preserved through the programming effort.

The third program will permit the calculation of the density of a document when its spatial location is known.

```
      READ (1,100) N
  100 FORMAT(I5)
      WRITE(3,113)
  113 FORMAT(///)
  102 FORMAT(13X,I5,3X,I2,3X,4(I2,2X),3X,I5,3X,I2,3X,F10.4)
      K1=0
      N1=0
      N2=1
      N3=5
      NP1=N+1
      DO 10 I5=1,NP1
      IA5=I5-1
      DO 10 I4=1,NP1
      IA4=I4-1
      DO 10 I3=1,NP1
      IA3=I3-1
      DO 10 I2=1,NP1
      IA2=I2-1
      DO 10 I1=1,N
      IA1=I1
      N1=N1+1
      II=N1/N3+N2
      GO TO (9,4,5,6,21),II
    4 N3=20
      N2=2
      IF (IA2) 8,10,8
    8 IF (IA1-IA2)9,10,9
    5 N3=100
      N2=3
      IF (IA2) 15,10,15
   15 IF ((IA1-IA2)*(IA1-IA3)*(IA2-IA3)) 9,10,9
    6 N3=500
      N2=4
      IF (IA2)  11,10,11
   11 IF (IA3) 12,10,12
   12 IF((IA1-IA2)*(IA1-IA3)*(IA1-IA4)*(IA2-IA3)*(IA2-IA4)*(IA3-IA4)) 9,
     110,9
    9 IP=IA1+IA2*NP1+IA3*NP1**2+IA4*NP1**3
      IR=(IA1+3)/4+(IA2+4)/5+(IA3+4)/5+(IA4+4)/5
      THETA=(N+1)**(IR-1)*N
      TETA=IP-(N+1)**(IR-1)
      THETA=360./THETA*TETA
      K1=K1+1
      WRITE(3,102)K1,N2,IA1,IA2,IA3,IA4,IP,IR,THETA
   10 CONTINUE
   21 CONTINUE
      END
```

```
      DIMENSION INDEX(15)
  100 FORMAT(2I5)
   25 READ (1,100) M,IP
      AM=M
      P=IP
      K=ALOG(P)/ALOG(AM+1.)
      IF (K) 3,4,3
    3 DO 1 I=1,K
      INDEX(I)=IP/((M+1)**(K+1-I))
      IP=IP-((M+1)**(K+1-I))*INDEX(I)
    1 CONTINUE
    4 INDEX(K+1)=IP
      K=K+1
      WRITE(3,101) M,P
      DO 2 I=1,K
      N=K+1-I
    2 WRITE(3,102) I,INDEX(N)
      GO TO 25
  101 FORMAT(' NUMBER OF INDEX TERMS IN COLLECTION = 'I5,5X,'DOCUMENT DE
     1NSITY =',F5.0//)
  102 FORMAT(' SN',I2,' =',I5)
      END
```

```
 40 READ (1,100) DEGREE,AM,R
100 FORMAT(3F10.3)
    DEN=(AM+1.)**(R-1.)/360.*(AM*DEGREE+360.)+.0009
    WRITE(3,101) DEN,DEGREE,R
101 FORMAT(F10.1,2F10.3)
    WRITE(3,110)
110 FORMAT('  DENSITY    ANGLE      RADIUS'/)
    GO TO 40
    END
```

VITA

Joseph Paul Anthony Lombardo, son of Salvatore Lombardo and Mary Baka Lombardo, was born in Phillipsburg, N. J. on December 19, 1927.

He is married to the former Mary Edna Stranzl of Northampton, Penna. They are the parents of three children, one boy, Daniel, and two girls, Joyce Ann and Marian.

He was graduated from Sts. Peter and Paul's Grammar School, Sts. Philip and James' High School and received a Bachelor's Degree in Electrical Engineering from Lafayette College in 1964.

After serving in the U. S. Navy in 1945 and 1946, Mr. Lombardo went to work for the Ingersoll-Rand Company, where he is currently employed as a Systems Analyst.