Articulation in the Generation of Subject Indexes by Computer*

JANET E. ARMITAGE and MICHAEL F. LYNCH
Postgraduate School of Librarianship, University of Sheffield, Sheffield 10, England

Received May 19, 1967

A simple and logical model for the automatic generation of subject indexes from titlelike phrases is described, and its advantages and disadvantages are discussed. It is based on recent studies of the structure of articulated subject indexes, such as those to *Chemical Abstracts*. The model employs the prepositions and connectives of phrases of simple structure as articulating points, and selects from all possible forms of entries those which lead to optimal organization in an index. The technique is illustrated with part of an index to a recent abstracting journal. The wide variety of controls which can be exerted by the indexer and the program is discussed.

A wide variety of computer techniques to aid in the preparation of printed indexes has been developed since the keyword-in-context index (KWIC index) was first suggested by Luhn (*1*). These developments have been ably reviewed by Coblans (*2*) and Stevens (*3*). The articulated subject index, of which the most highly developed example is the subject index to *Chemical Abstracts*, has received very little attention other than in regard to mechanization of the sorting and printing routines involved in its preparation (*4*). Recently, however, one of the authors (*5*) has demonstrated that there is a simple and general logic which relates an entry, as it appears in an articulated subject index, to the normal, or titlelike form of the descriptive phrase from which the entry is derived, subject to easily determined conditions. This logic has since been embodied in a computer program which examines index entries, determines whether the conditions necessary for the transformation are present, and, if they are satisfied, converts the articulated entry into the corresponding titlelike form of the descriptive phrase.

It has now been found possible to deduce a simple model for the generation of entries in a subject index from the logic discovered for the transformation of entry to title. Furthermore, sorting procedures have been devised which enable a choice of entries to be made so that an optimum degree of organization in the display of information in a printed index can be attained.

It is thus possible to envisage a much higher degree of mechanization in the production of high quality subject indexes than has hitherto been possible, and it is likely that adoption of the procedures suggested here could result in relieving indexers of much routine work and in increased productivity on their part.

The problem is one that has already been considered by a number of workers. Thus, Herner (*6*) has outlined instructions to clerks for controlled manual permutation of the words in brief articulated phrases, in the context of the Permutext method of indexing. Farradane (*7*) has included rules for manipulating entries in the notation used in his relational indexing scheme to give subject index entries. Freeman (*8*) has studied the problem of generating index entries from titles, viewing it as a syntactic rather than a logical process.

A preliminary manual test of the logic of subject index generation has been carried out on the 479 abstracts in the first issue of *Documentation Abstracts*. It has been shown that all the steps which follow the compilation by an indexer of a phrase which describes a document can be carried out according to simple rules. The results of the preliminary test are encouraging, and the procedures are currently being programmed.

ARTICULATION OF INDEXING PHRASES

Examination of the form of an articulated subject index shows that it consists of a set of subject headings, in alphabetical order, under which are indented series of modifying phrases or modifications. An example for cesium is:

Cesium.
absorption by plants, fertilizer effect on, **60**:13833f
by plants, soil colloids and, **60**:11321h
by roots, Ca and, **60**:12620b
adenosine triphosphatase response to, **60**:4400b
adsorbed on Pt electrodes, hydroquinone-quinone system in relation to, **60**:3733g
adsorption of, by Hg electrodes, in presence of methylformamide, **60**:8668c
from radioactive waste water by clay, **60**:3865e
from Na soln. by clinoptilolite, heat-treatment effect on, **60**:15482h
from waste water by brown coal, clay and clinoptilolite, **60**:13009c
agaroid gel properties in presence of, **60**:6246e
argon elec. plasma contg., spectroscopic temp. detn. of, validity of, **60**:10077a
atomic scattering factor of, **60**:7528d,8655h
from barium-133 decay, γ-γ angular correlation in, **60**:12835h
base exchange of, in alcs. or aq. alcs., **60**:2359b
with NH_3 on faujasite-type zeolites, **60**:7490h
on $(NH_4)_3PMo_{12}O_{40}$, **60**:11405h
on Bio-Rex 70 and Dowex-50W, hydration in relation to, **60**:42e
with Ca and Li, solvents in relation to, **60**:7493c

* Presented before the Division of Chemical Literature, 153rd National Meeting of the American Chemical Society, Miami Beach, Fla., April 12, 1967.

on Dowex 50-X2, and Co in relation thereto, **60**:11404f
on Duolite C-65, **60**:15181g
with H ion on Zr phosphate, **60**:6244e
with H ion, thermodynamics of, **60**:1179e
on KU-2 resin, **60**:4837g
on micas, **60**:8670g
on mol. sieves of type A, **60**:4836a
with K and Na in zeolites, **60**:13024g
with Na-high porcelain membranes, in fused systems, **60**:12692h
with Na in two-temp. process, **60**:9951d
on zeolites, **60**:4836b,8913f

It is a characteristic of this index that the subject heading and the modification together form a full descriptive phrase. The modifications are listed under each subject heading in alphabetical order of their significant words—*i.e.*, function words such as prepositions and connectives are ignored in the alphabetization. The modifications themselves can be broken into components, the components being separated by commas. It is a further characteristic of the index that when the initial component is common to two or more modifications, it is printed once only, and the following modifications are again indented. In this way, a high degree of organization can be introduced into the display. Whaley (*9*) has given an excellent assessment of the advantages and the manner of use of this sort of subject index, and has shown how the form of organization is of value in scanning the display.

The logic for conversion of an entry into the titlelike form of the descriptive phrase depends upon the fact that the components of a modification usually consist of substantives—*i.e.*, noun phrases—plus a function word. The latter may either precede or follow the substantive. If each of the components of a modification fulfills the condition that it either begins or ends with a function word, it is possible to convert the entry into the normal form of the descriptive phrase. Such entries are termed regular entries. Irregular entries can, in many instances, be converted to normal phrases, but require a more extended analysis.

The procedure can be illustrated by the following examples, chosen from the index to *Chemical Abstracts*. In the entry

Energy levels
of zinc arsenide (Zn_3As_2) semiconductor,
gap in, and temp. coeff., **60**:3592a

the modification consists of three components. Each of these has a function word: "of" in the first component, "in" in the second, and "and" in the third. Each component thus fulfills the requirement for regularity—namely, that it should have a function word at either its beginning or end, but not at both. It is immaterial whether individual components contain further nonterminal function words. In converting the entry to the normal form, each component is considered in turn. If the function word of the component under consideration appears at the beginning of the component, then that component follows the subject heading, or any portion of the phrase already reconstituted. Thus, the first step in the conversion of the entry gives:

energy levels of zinc arsenide (Zn_3As_2) semiconductor

If, on the other hand, the function word of a component comes at its end, then the component must be placed to precede the subject heading, or any portion of the phrase already formed. Thus, taking the second component, the next step leads to:

gap in energy levels of zinc arsenide (Zn_3As_2) semiconductor

Finally, the third component must follow the part already formed, as the function word falls at its beginning. The full normal form of this entry is thus:

gap in energy levels of zinc arsenide (Zn_3As_2) semiconductor and temp. coeff.

An example with a different sequence of component types is as follows:

Isoleucine
formation of, by *Escherichia coli*, regulatory mechanisms in, **60**:9648e

Here the first component must be placed before the subject heading; the second follows these two; and the third must be placed before all three, giving rise in successive steps to the complete form of the phrase

formation of isoleucine
formation of isoleucine by *Escherichia coli*
regulatory mechanisms in formation of isoleucine by *Escherichia coli*

Over 1000 entries from the subject indexes to *Chemical Abstracts* have been examined. In all cases in which the conditions for regularity of an entry were satisfied, the normal form, derived according to the above rules, was identical with the intuitively correct form of the descriptive phrase. Furthermore, the incidence of regularity in entries is high. In a randomly selected sample of 1000 entries, 739 had modifications which did not consist solely of names of compounds. Of these entries, 60% were found to conform to the conditions necessary for regularity.

More recently, other major English language subject indexes have been examined from the same point of view. At least 15 of these made appreciable use of articulation in the structure of the entries, and about half of these had a degree of regularity comparable with, and often in excess of, *Chemical Abstracts* subject indexes. Again, the regular entries in these indexes conformed, without exception, to the general rule for transformation into the normal form.

About 40% of the entries from *Chemical Abstracts* subject index are irregular; in many of these cases, a somewhat deeper analysis can result in correct regeneration of the normal form. The commonest reasons for irregularity are the following:

(a) the presence of a comma in the normal form itself,
(b) elision of a preposition and comma in certain common constructions,
(c) extraction of the subject heading from a more complex substantive phrase.

The following illustrate these points:

Comma in normal form:

Coal
analysis for C, H, and N, **60**:15e

Acrylamide
polymers of
with acrylic acid, Na salt, hydraulic and lubricant emulsions
contg., **60**:P11827f

Elision in common constructions:

Beta rays
effect
on skin, **60**:13534b

Gas
detection in air, **60**:P2335f

Subject heading part of larger phrase:

Heat
balance
in open-hearth furnace, **60**:5120e

Catalysts
heterogeneous,
mechanism of, competitive reaction and, **60**:12698a

A computer program has now been developed, the purpose of which is to examine index entries to determine if the conditions for regularity are present, and if so, to regenerate the normal form of the phrase. It is written in SLIP (*10*), the symmetric list-processing language which has been incorporated in a version of FORTRAN implemented on the Science Research Council Atlas computer. The program examines entries, breaks them down into subject heading and components by detecting commas, and looks for a function word at either the beginning or end of each component by comparing terminal words against a short list of words. If the conditions are satisfied for each component, the program then regenerates and prints the normal form of the phrase. It has been successfully tested on sample entries, and its extension to deal with the more frequent instances of irregular entries is at present in hand.

It is clear that this rationalization of what has been largely an intuitive process reveals a general logic underlying the formation of subject index entries. This logic is based on a property of descriptive phrases in the English language—namely, that extended phrases frequently contain function words such as prepositions and connectives, about which the noun phrases can be articulated.

There are at least two reasons for the use of articulation in an index. In the first place, articulation permits the introduction of a high degree of organization into the display of an index. Thus, the form of the entries under a particular subject heading can be varied so that as many as possible of the first components are common to a number of entries. The use of indentation to indicate this results both in economy in space and in a reduction in the effort involved in scanning the display, due to the conjunction by subordination discussed by Whaley (*9*). A second reason, the importance of which is more difficult to assess, is that articulation permits the apposition of substantive phrases with the subject heading, in cases where the latter phrase may not follow the subject heading in the normal form. This point is illustrated by the following entry:

Cartilage
polysaccharide formation by, uridine nucleotide sugars in,

In logical terms, an alternative form of the entry would have been:

Cartilage
uridine nucleotide sugars in polysaccharide formation by,

A decision in favor of the first form of the entry was made by the indexer, and was presumably based on his knowledge of the subject field and his estimate of the more useful apposition. Subjective decisions are thus involved here, and it is scarcely necessary to point out that such decisions are less easily reflected in a computer program than decisions which have a purely logical basis.

INDEX ENTRY GENERATION

Whatever the relative importance of the reasons for the selection of a particular form of entry may be, it is worth-while considering an algorithm which will permit the generation of all the possible forms that an entry derived from an articulated phrase may take. This algorithm is based on the corollary of the logic which permits a regular entry to be transformed into the normal form. It operates on descriptive phrases in which noun phrases are separated by function words. It is presumed that the noun phrases, which may consist of one or more words, can act as subject headings in an index. Such a phrase can be represented as a string of n words or phrases, separated by $n-1$ function words:

Normal form:	—o—o—o—o—
Generalized as:	$—(o—)_n$ $n = 0,1,2,3.....$
Example:	articulation in subject indexes in English

This representation of an articulated phrase serves as the simplest model for the logical generation of subject index entries. The general rule then states that if one of the noun phrases be extracted from the string to serve as a subject heading, then all the possible modifications can be formed by choosing an adjacent function word and the noun phrase adjacent to it and continuing this selection provided an earlier choice has not reached one or the other end of the string. Thus at each stage, a binary choice is involved. Furthermore, the selection may include multiple sets of function words and noun phrases. The procedure is illustrated with the following example:

articulation in indexes for books on science

Books articulation in indexes for on science

At the first stage:

Books
indexes for,

Books
on science

At the second stage:

Books
indexes for, articulation in,

Books
on science, indexes for,

Books
indexes for, on science,

etc.

If multiple sets are chosen, some of the possible entries are as follows:

Books
articulation in indexes for, on science,

Books
on science, articulation in indexes for,

One implication of this model is the fact that when the subject heading chosen is the first of the string, only a single form of modification is possible—that in which the remainder of the string follows serially. Another is that any subject heading can be placed in conjunction with the noun which follows it, or with any which precedes it.

At this point it will also be quite clear that while the transformation:

index entry → normal form

gives a single, unequivocal result in the case of regular entries, the transformation

normal form — index entry

is multivalued, the number of possible forms depending on the number of nouns in the original phrase. Indeed, in the survey of *Chemical Abstracts* indexes, examples of all possible sequences of components were found for entries consisting of a subject heading and three or fewer components.

It is impractical to consider an index in which each of many potential combinations can appear. A means of selection of one form of modification to appear with the appropriate subject heading is necessary. Clearly, it is possible to direct the selection so that an important characteristic of the articulated subject index, the high degree of organization in the display, is achieved. Moreover, it is likely that the second objective, significant apposition of important words, will also be achieved to some extent, in that it may depend on the frequency of cooccurrence of pairs of words or phrases. The simple rules for sorting all possible forms of modifications under each subject heading are as follows:

(a) Select all invariant forms of entry and alphabetize by the significant words of the modifications. Examples of invariant entries are those in which there are only two components, or those in which the subject heading chosen is the first word of the phrase.

(b) In cases where variable forms of modifications are possible, select all the words which are possible first components. For the i^{th} word in a phrase, there are usually i-possible first components, these being the word which follows the particular subject heading and any which precede it. The exception occurs when the subject heading is the last word in the phrase, when there are only $i-1$ possible first components. These possible first components are sorted against the initial components of the invariant entries. Those forms which have possible first components which are also found in the invariant entries are selected and other potential forms are discarded. This process can be continued, if necessary, to second and subsequent components.

(c) When no selection has yet been made, the possible first components of variable entries are sorted against one another, and those which lead to the greatest numbers of common first components are selected. All other potential forms of these phrases are discarded.

(d) Any remaining variable entries are inserted in the least rearranged form.

The first two steps are illustrated with the following phrases:

classification by computers and analysis of English language.
information retrieval and classification by computers.

The only form of entry possible for the first phrase under the subject heading "classification" is

Classification
by computers and analysis of English language

This enables a choice to be made from the variable forms of entry of the second phrase—*i.e.*, "classification," followed by "information retrieval" or "computers," since, of these, "computers" is common to both this and the previous entry, and is therefore selected. The two entries are displayed as follows:

Classification
by computers and analysis of English language, 66/1-428
information retrieval and, 66/1-439

The third step is illustrated by the following examples:

education toward research in information retrieval by computer.
information retrieval by computer in metallurgy.

Under the heading "computer," modifications beginning with "education," "research," or "information retrieval" are possible for the first phrase, and for the second, "information retrieval" or "metallurgy." Since "information retrieval" gives a common initial component, these forms are chosen, resulting in:

Computer
information retrieval by, education toward research in, 66/1-314
in metallurgy, 66/1-248

It is necessary to invoke the fourth rule in the case of an entry such as:

cooperation in cataloging in Czechoslovakia,

which, under "cataloging," becomes

Cataloging
cooperation in, in Czechoslovakia, 66/1-109

since no other entry under this heading has an initial component containing either "cooperation" or "Czechoslovakia."

These, then, are the basic rules for manipulating and sorting the components of articulated phrases to form a well-organized index.

To these must be added a further rule which relates to a specific function word—namely "of." In applying the logical rules to phrases in which the noun or gerund form of an active verb is followed by "of," it was noted that certain ill-formed entries were produced. They were correct in the logical sense, but did not read well. Examples of ill-formed entries are:

Catalogs
by computer, production of,

and,

Computer
catalogs by, production of,

One further simple rule excludes these ill-formed entries:

(e) When the preposition "of" occurs between two words or phrases and either of these acts as a subject heading, the other must come at one end of the first component, which may consist of multiple sets. Furthermore, the function word "of" does not act as an articulating point when indexing the phrase at other headings.

The permissible entries for this phrase are then:

Catalogs
production of, by computer,

and

Computer
production of catalogs by,

These rules have been found to give acceptable forms of entries and to result in a high degree of organization in sample indexes when applied to phrases in which certain parts of speech are excluded. These parts of speech are articles, infinitives, participles (except when qualifying nouns), and gerunds (except when used as nouns, as in "Abstracting"). These exclusions are necessary to ensure that the substantives which will be used as subject headings are separated by function words about which articulation can take place. Otherwise, for instance, the "to" of an infinitive would be falsely treated as an articulating "to," and gerunds and participles would be treated as parts of subject headings. However, it may be possible in time to develop this simple model yet further to accommodate more complex forms of structure.

Even with the simple model, a number of elaborations is possible, some of which have been incorporated in the manual test of this indexing procedure on the first issue of *Documentation Abstracts*. Thus, it is often necessary to refer to multiple and coordinate concepts, as in the phrase:

cooperation and standardization in cataloging

If the "and" is replaced by an ampersand, a means of differentiating between the articulating "and" and the coordinating "and" is available:

cooperation & standardization in cataloging

Separate entries under "cooperation" and "standardization" can then be provided, giving the following entries:

Cataloging
cooperation & standardization in,

Cooperation
in cataloging,

Standardization
in cataloging,

Various controls can be exerted, both by the indexer at the time of compilation of the descriptive phrases and by program. Thus, a standard stop-word list, as used in KWIC indexing, can be included. Alternatively, the indexer can bracket words of low indexing value to prevent their being used as subject headings. Again, a dictionary can be compiled, which may include instructions on the treatment of specific compound phrases. Thus it may be desirable to index "Systems analysis" both at:

Systems
—analysis

and at

Analysis
systems —

or to treat some articulated phrases, such as "Library of Congress," as single entities. Again, especially with frequently occurring headings such as "Information science" and "Library science," a similar splitting could lead to overloading at "Science." Instead, the splitting would be inhibited and cross references would be inserted automatically.

A further elaboration which increases the power of the indexing language is the use of generic posting. It is often necessary to use highly specific terms to describe the content of a document accurately, but these terms may not fit into the vocabulary. If the specific term is followed by the generic term, the entry can be included under the latter, while a cross reference can be entered at the specific term. Thus:

organization of crystallographic data (chemistry) by computer

would be represented by:

Chemistry
organization of crystallographic data by computer,

Computer
organization of crystallographic data by,

Crystallographic data
see *Chemistry*

A similar device can be used to index books by title.

These procedures have now been tested by indexing the 479 abstracts in the first issue of *Documentation Abstracts*. Sample pages from the index are given in Appendix A. In preparing the index, an indexing phrase was compiled for each abstract. A vocabulary was developed and refined during the process. No constraints were imposed on the structure used in the phrases other than exclusion of the parts of speech mentioned previously, nor was any subsequent attempt made to modify structure to bring about greater uniformity.

Each subject heading was treated in turn. All phrases in which the heading appeared were brought together, and all possible first components identified for each occurrence of the subject heading in a phrase. The sorting procedures described earlier were then applied so that finally only one of the many possible forms in which each phrase might appear at a particular heading was chosen, all alternative forms being discarded.

The index which resulted from the application of these procedures, all of which are logical and programmable, shows a high degree of organization in spite of the fact that the number of abstracts included was relatively small. It can be expected that an even greater degree of organization can be attained when larger numbers of entries are used.

It is useful to note at this point that the scheme is well-suited to the production of periodic indexes, followed at longer intervals by cumulations. In producing cumulations, the original indexing phrases rather than the entries produced for a particular index would be reprocessed. The form in which a particular entry appeared in a cumulative index might well differ from its form in an index covering a shorter time span, as it would depend on the nature of the other entries occurring with it under each heading.

The degree of success attained in this preliminary trial is, we believe, considerable; we are now engaged in programming the procedures to provide a comprehensive indexing scheme which we will then make available for

APPENDIX A

Draft of Index to *Documentation Abstracts*, Vol. 1, No. 1. (Mockup of computer output)

testing in a variety of contexts by interested organizations. The scheme will in no sense be tied to any specific discipline or vocabulary; rather, we see it as a widely adaptable procedure which can be applied in any of a wide range of disciplines.

We believe that it offers the prospect of considerable increase in the output of indexers, and of substantial improvements in the final indexes. Thus the scheme ensures that an optimal degree of organization is attained in the display of information in the index. It also ensures that all entries which relate to the same document but appear at different headings carry essentially the same information. This is in marked contrast to many articulated subject indexes, in which cognate entries frequently differ widely in content.

Looking further ahead, we envisage the scheme's eventually being coupled with an on-line editing facility, which would enable an index-editor to control the final form of the display. Using a console, he could view the entries under each subject heading, assess the usefulness of the form of organization which the logical procedures had resulted in, and modify them as seemed appropriate.

MATHEMATICAL ASPECTS

Finally, it seems appropriate to provide a brief review of the mathematical basis of subject index generation. Considering again the simplest model of the articulated indexing phrase, which consists of a string of n subject headings and $n - 1$ function words, it is apparent that the total number of logically permissible index entries is less than $n!$, since the combinations are subject to certain constraints.

The number of possible entries for a string of n headings is easily determined, as the rules for selection state that when the i^{th} heading is selected from a phrase of n headings, the permissible entries can be formed by following it either by the $(i-1)^{th}$ or by the $(i+1)^{th}$ heading, or by sets of adjacent headings which either end with the $(i-1)^{th}$ heading or begin with the $(i+1)^{th}$. This procedure is repeated until the end of the string is reached on either side. When the right-hand end of the string is reached, those remaining to the left can be chosen singly or in multiple sets. One implication of this rule is that if the heading is the first of the string, only a single form of modification is possible, because single selections lead to the same result as selections of multiple sets. On the other hand, if the heading is the last of the string, 2^{n-2} combinations are possible, because in this case single selections and selections of multiple sets lead to different modifications. While the number of possible entries for the case in which the heading is the first of the string is always 1, the number for the second heading is $n - 1$, or the natural numbers. The numbers for subsequent headings are again series, the differences between individual sums being related to the triangular numbers.

The logically permissible combinations for phrases with up to four headings are:

No. of headings in title, n	Phrase	Possible Index Entries				
1	A	A				
2	A·B	A_B	B_A			
3	A·B·C	A_{BC}	B_{AC} CA	C_{AB} BA		
4	A·B·C·D	A_{BCD}	B_{ACD} CAD CDA	C_{ABD} BAD BDA DAB DBA	D_{ABC} BCA CAB CBA	

The maximum number of permissible entries for each heading in a series of n, and the total number of possible entries for all headings, is:

n	No. of index entries under t^{th} heading						Total
	1	2	3	4	5	6	
1	1						1
2	1	1					2
3	1	2	2				5
4	1	3	5	4			13
5	1	4	9	12	8		34
6	1	5	14	25	28	16	89

These are the alternate numbers in the Fibonacci series.

The general expression for the total number of possible entries for a phrase with n headings takes the form:

$$a_n = \frac{1}{(5)^{1/2}} \left[\left(\frac{(5)^{1/2}+1}{2} \right)^{2n-1} + \left(\frac{(5)^{1/2}-1}{2} \right)^{2n-1} \right]$$

SUMMARY

A theoretical basis has been laid for the transformations involved in deriving articulated subject index entries from phrases of simple structure. Sorting procedures have been devised which lead to the selection of those forms of entry which result in a highly organized subject index.

The resulting automatic index generation scheme promises to be a powerful aid toward the production of subject indexes of high descriptive quality, in that it combines human skills with the logical processing capability of the computer in an optimal fashion.

ACKNOWLEDGMENT

The authors gratefully acknowledge a grant from the Office of Scientific and Technical Information, London, and thank Dr. G. M. Dyson for valuable discussions, Dr. I. J. Good for assistance with the mathematical aspects of the study, and the Director and staff of the Atlas Computing Laboratory for the provision of computing facilities.

LITERATURE CITED

(1) Luhn, H. P., "Keyword-in-Context Indexing for Technical Literature (KWIC Index)," *Am. Doc.* **11**, 288–95 (1960).
(2) Coblans, H., "Use of Mechanized Methods in Documentation Work," ASLIB, London, 1966.
(3) Stevens, M. E., "Automatic Indexing: A State-of-the Art Report." NBS Monograph 91, U.S. Department of Commerce, National Bureau of Standards, Washington, D.C., 1965.
(4) Tate, F. A., "Progress Toward a Computer-Based Chemical Information System," *Chem. Eng. News* **45**, No. 4, 78–90 (1967).
(5) Lynch, M. F., "Subject Indexes and Automatic Document Retrieval: The Structure of Entries in *Chemical Abstracts* Subject Indexes," *J. Doc.* **22**, 167–85 (1966).
(6) Herner, S., "Deep Subject Indexing by Manual Permutation Methods," *Automation and Scientific Communication*, American Documentation Institute, Washington, D. C., 1965, pp. 437–9.
(7) Farradane, J., Scientific Theory of Classification, *J. Doc.* **6**, 83–89 (1950).
(8) Freeman, R. R., unpublished paper.
(9) Whaley, F. R., "The Use of Subordinate and Coordinate Indexes *vs.* the Scanning of Their Outputs," J. CHEM. DOC. **5**, 102–7 (1965).
(10) Weizenbaum, J., "Symmetric List Processor," *Commun. Assoc. Comput. Mach.* **6**, No. 9, 524–44 (1963).