APPROACHES TO LIBRARY FILING BY COMPUTER

Jean M. Perreault

"The essence of library catalogue is arrangement of entries."[1]
(a) There is no intention, in this paper, of providing anything like a
new code of filing rules for use with the computer. Ted Hines and
Jessica Harris have made a valiant and largely successful try at this
task. It is recommended that you obtain it[2] and read it thoughtfully.
(b) Nor will this paper comment on the two classic American filing
codes[3] in such a way that the form subdivision for a subject-heading
on this paper would read " — Commentaries," but rather in such a
way as to give it " — Criticism, interpretation, etc." (c) Nor (as a
final disclaimer) is this paper thematically concerned with the hope
for code revision—though these pages come closer to such a treat-
ment than to (a) or (b).

Instead, there will be an attempt (d) to present some of the
intellectual or bibliographical problems involved in the notions of sort-
ing and filing, and then (e) an outline of some of the tools and tech-
niques which can be brought to bear upon their resolution. Together,
these ideas should (f) make possible a rational basis for the evalu-
ation of filing-code-revision suggestions.

The author is not entirely neutral in all this, as those who have
read my paper "The Computer and Catalog Filing Rules"[4] should
know; but the following discourse will subject that paper to consider-
able retractatio. Primary influences in moving toward this revalu-
ation were Lubetzky and Ranganathan. (To avoid confusion from the
outset, the author does not take "library filing" to include the ar-
rangement of documents, but only of their surrogates.)

Words—as anyone who has read, say, Finnegan's Wake, knows
—can be a lot of fun. But they can also be quite a burden, particular-
ly to whoever has the task of laying them end to end, or worse, one
atop the other, in what could be called a vertical rather than a hori-
zontal order. All discourse is the construction of "horizontal"

Jean M. Perreault is Lecturer, School of Library and Information
Services, University of Maryland.

strings of words; the problem of filing is the construction of "vertical" stacks of words <u>in discourse</u>.

This last phrase should be emphasized because the filing problem, that of organizing stacks of discursions, is far more complex than that of simple sorting. The ideal, the situation which could enable fully adequate filing by ritualizations such as computers or filing clerks, would be the reduction of filing rules to sorting rules.

Filing and sorting could be distinguished thus: to sort is to set in order, to file is to interpolate into an already existing order of sorted items. But, common and easy to understand as this may be, it is far too superficial; even in the sorting process as defined here, there must necessarily arise interpolation—unless the items being "sorted" are somehow already in order. And, then, what does "in order" mean? We shall try to see later.

To sort is to put like with like. The term was used from the beginning of printing shop practice to mean the replacing of the elements of the broken up form, each element of which belongs with its own <u>sort</u> (type) in the case. "A" is a sort, "a" is another, and all the exemplars of each are put back into their appropriate sector(s) of the upper- or the lower-case, and this putting of sort with sort is sorting.

Words too can be sorted, just as can complex letter-forms. There is an interesting historical example of this cited in Douglas McMurtrie's history of printing <u>The Book</u>, [5] where it is pointed out that the Chinese idea of moveable type (for printing in the Uigur language) was for the elementary pieces each to print a word, although Uigur is a language where words were built up from more elementary particles, namely letters. In our literal Western languages sorting is basically the ordering of the smallest elements, namely letters; but it is intriguing to try to imagine how an order of irreducible <u>words</u> could be set up. Let us not forget, however, that our own alphabetical (letter) order is every bit as arbitrary, although admittedly shorter than would be the (word) order of, say, the Mandarin vocabulary. Like the Chinese, a Western printer might well consider letter-complexes such as "á," "à", and "â" to constitute three additional sorts besides plain "a," but the breakdown into simple elements could go further, to "a," "´", "`", and "^", so that these accentual elements could be freely combined with other ordinary letters (as for instance "´" over "n" in Polish). By this technique, the total number of possible sorts is reduced, and this is only the logical extension of the reduction of the sorting of words themselves by means of the sorting of their elements, letters.

By "words", here, is meant each word—all the words of the dictionary taken individually. Discourse, where words are not taken individually, is quite another matter, because discourse cannot be standardized. It is this, in fact, that makes language fun, and it was

Joyce's refusal to accept the standardization of words, too, which makes Finnegan's Wake fun.

We can establish "sort-boxes" for everything from the elementary parts of an alphabet, or of all alphabets, up to the elementary parts (words) of a language. But how can the cross-over from this world of sorting to the quite different one of filing be accomplished?

First of all, since it has been postulated that it is non-standardizability that sets off the discursive from the sortably standardizable, we must try to see what it is, as far as order is concerned, that constitutes discourse. In answer to this I would suggest articulation, the same sort of phenomenon that lets a plurality of lines form various 2- or 3-dimensional figures by coming together at a variety of angles. Because of this variability, any two words can form as many patterns of discursive meaning as the language will allow—as is demonstrated, for instance, in the farce line "But you don't even know I'm alive" or in a great many puns.[6]

The usual way of handling these points of articulation between words is commonly expressed, among librarians, as "nothing before something," since the juncture between words is most commonly shown by a blank, a "nothing." This is true even in spoken language, where, although the blanks between words are not always clearly perceptible, at least those representing phrase-points normally are. These, indeed, are the historical antecedents in the development of our present-day rules of written-language punctuation.[7] The reduction of the principle "nothing before something" to a purely sorting process is accomplished by treating the articulating blank as a part of the collating sequence, so that the sequence reads "blank, a, b, . . ."

But there are other kinds of articulation in "horizontal" discourse than that constituted by the blanks between words. There are, as mentioned, in spoken discourse: intonation, tempo, accent, etc.; in written discourse: punctuation, paragraphing, abnormal capitalization, italicization, etc. Although it is clearly possible to take these into account in our filing codes—by the device explained in "The Computer and Catalog Filing Rules"[4] for instance—which device leads to phrase-by-phrase filing in addition to the word-by-word filing enabled by the use of the blank in the collating sequence—the real difficulty is the same as that encountered in mechanical translation and in classification for information retrieval with machines as the clerks. This central problem is ambiguity. It does not require much reflection to see that if each word in a language could denote one concept and one only, any word or words in that language could be easily enough translated into another similarly characterized language, provided that the conceptual underpinning and partitioning of the two languages was the same. That is, if there is no word in Eskimo for "to lie", then the English verb "to lie" cannot be adequately translated into Eskimo even if the structural similarity is there.

But in fact, languages are made up of words which are not univocal,[8] and whose equivocacy becomes even more marked in discursive situations,[9] since there the context sorts out the one sense among the several as appropriate. Take the phrases "a man of iron", "a feature of the land", "a man of Florence", "a book of John's." "Of" in these examples means successively "with the characteristic," "inherent in," "stemming from," and "belonging to." A precisely similar equivocacy can be seen in our marks of punctuation. (This is not, of course, to say that every preposition or mark of punctuation is equally burdened with equivocacy, but that some, such as "of" and the comma, are made to do a far greater share of the work than seems fair.)[10]

Now all this, in terms of the problem of filing, is incarnate in our desire to avoid purely mechanical rules and in our tendency to prefer an ordering of discursive elements in the same way that we articulate them in speech, in some way which transcends the purely temporal order of the sounds themselves. This tendency, or at least something faintly caricatural of it, is to be seen at work in the American Library Association and the Library of Congress filing rules. And lest we imagine that present-day Americans are alone in this tendency, bad or good, see the Prussian Instructions, part II.[11] In all these bodies of rules, we see a variety of solutions to the fact that in spoken discourse words can be variously emphasized, so that the usual order presumed as the basis of all filing is belied. This presumption is that the first element is always of higher importance than the second, and so forth.[12] (This point is given the name "Canon of Prepotence" by Ranganathan. Unfortunately, he forgets to indicate that, national as this position or unconscious policy may be, it is not the only one.)

This presumption should be examined, if only briefly. If there are two Dewey numbers, say 309.99 and 310, their filing order is determined by positional comparisons which accept the presumption mentioned above. The same is true as well for such a pair as 009 and 100, where the temptation would be strong for a naive person to put '1' earlier than '9' and not worry about the zeroes. But there is a different kind of number, one encountered even more frequently in real life: the integer as against the decimal number. In this case, two numbers are compared not on a first-to-first-element basis, and so forth, but in the exact reverse; here 30999 comes after, not before, 310.

Now although rules are needed for sorting of numbers, the verbal sorting/filing problem is far more difficult, or at least forms much the largest problem-sector in filing practice, except in the case of LC call numbers. There the presence of both decimal numbers and integers presents problems even to some librarians. And the interesting point is that sorting rules for words almost invariably operate according to the decimal sorting presumption. No one would

suggest that "mad" should come after "management" just because "-ent" is lower in the collating sequence than "mad" except, of course, in the organization of a rhyming dictionary. Yet there are cases where something close to the integer-sorting algorithm is at work with words, for instance where initial words such as articles are assumed to be of no value in the sorting order. (This tendency could, alternatively, be compared with the leading zeroes which get "9" in 009 to file ahead of "1" in 100. "The" at the head of a title is regarded as analogous to such ciphers.)

If sorting is the linear arrangement of simple or pseudo-simple inscriptions, and filing is the linear arrangement of semantic and discursive inscriptions, with all the difficulties implied, and if the goal is to make the problem of filing simple enough to be performed by clerks, human or mechanical, then we must seek for means to reduce these semantic-discursive aggregates to at least pseudo-simple ones. There are several tools or techniques for this task, ranging through quite a spectrum both in the complications of input required and in the output achieved.

If a clerk is given two entries to file relative to one another, one reading "Bible. English. 1964. Goodspeed." and the other reading "Bible. English. Knox.," the second files before the first not because letters, represented here by the "K" in "Knox" comes before the "1" in "1964", but because the Knox version, date unknown, would come among the first-order group of English (whole) Bibles with unspecified date-facets. (In such a case as this, a perhaps unexpected reflection of the rules for inverted file-order as derived from a citation-order can be seen quite clearly; see the Appendices for a fuller discussion.)

But for such an effect to be produced there must be instructions to the filing clerk that predict such possible variations, instructions such as "The facet after 'language' should be 'date'; if not, treat it as a blank date-facet, filed before all full date-facets." Essential in all this, if a clerk such as the computer is employed, is that each facet of such an entry be unambiguously labelled, so that the problem never occurs of the computer-clerk's not being able to "understand" the elements which it is called upon to manipulate.

Two instances of the same sort of discursive articulation can be seen in comparison (a) of such a pair of titles as The Man without Qualities and Man; Zephyr or Clod. Assuming that a dictionary catalog is being considered, the second of these will not be entered at all unless the explanatory sub-title is included, since otherwise it will be dropped because of a match with the subject heading apparatus. But if they are both entered and are to be filed relative to each other, the phrase-point introduced by the semi-colon cannot well be ignored lest the title be found only by someone who could imagine a title beginning (as if "man" were an adjective) Man Zephyr. . . . Thus, although

"w" in "without" precedes "Z" in "Zephyr", in the collating sequence, the phrase "Man", being shorter than "The Man without Qualities", should probably file first—just as a short word like "man" files before a long one like "management", and a short title like The Man before a long one like A Man and his Dog. It can also be seen in the order of (b) such a pair of entries as "Mass—Meditations" and "Mass (Canon law)." The parenthetical qualifier modifies the initial word in such a way as to make it represent a new concept, a longer one than the same word in its unqualified sense, even though the unqualified one is extended by a subdivision—and despite the priority of the "C" in "Canon" to the "M" in "Meditations." While something more than pure "alphabetic" or "symbolic" order is desired, there is an ineluctable dependence of all orders upon that which is "matter" to their "form"—in this case the ";" or "—" or "(."

Thus there is a definite justification for the traditional attitude of librarians that absolute adherence to the letters (or symbols) alone of an entry is not enough to generate an intelligible and useful file. "Order" is that which, hopefully, is generated by a file; just what state of affairs it is that can be characterized by "in order" is not clear.

Fugmann defines order as "The meaningful contiguity (Beieinandersein) of the parts of a whole."[13] But this seems to imply perhaps too forcefully the naturalness of order; an organ in a body has such a part to play in the whole, but a word in a list of words is not related thus to the others. A bit more neutrally then, one might define order as "being where one ought to be." Thus "2" ought to be after "1", and "3" next; but the reason for "b" being after "a" is not quite the same. The order of letters is arbitrary, not as with numbers, natural. Yet even this last is not entirely true, since the number sequence could, if not radix-10, be quite other than is "naturally" expected. Thus it can be seen that it is finally the concepts referred to or represented by (1) the numbers not as mere digits, but as real numbers or (2) letters not as mere digits, but as elements of words, which are the basis for the order of a file—rather than that of a sorting, which depends entirely on "form" rather than "content." Or we could say, to follow the earlier terminology, "matter" rather than "form."

Even if there seems to be no real order except that of meaning, and if there remains a mixture of arbitrary and part-natural types, there is a common factor at work, namely that they both make prediction possible. The difference between sorting and filing comes sharply to the surface when the predictability to be found in each is sought. That of sorting is absolute, within the allowable limits of the decimal or integral presumption. But that of filing is only partial, (a) because of its dependence upon the meaning of the words (semantics), and (b) because of the meaning of the discourse constituted by these words as inter-related (syntax).

The essential ideas to keep in mind are:

(1a) Filing and sorting are not identical, nor do they refer to different stages of a single process. Sorting is putting like with like, whereas filing is the creation of a series of complexes by means of partial likenesses.

(1b) Discourse is what is to be filed, in libraries, and discourse is made possible by the articulation of semantemes (meaning-bearers) by syntax.

(1c) The result of this articulation, as well as of the character of many semantemes, is ambiguity. Even greater ambiguity characterizes a complex that is not articulated: thus our traditional and quite rational desire to retain as much as possible of the benefit of syntactic articulation.

(2a) There are two presumptions from which to choose in building a file: either the decimal or the integral. Numbers can be filed either way; words usually can be filed only decimally;

(2b) Means are needed to convert some words into "leading zeroes," as if the integral presumption were operative.

(3) The most general principles for filing are two: (a) "nothing before something," which can also be elaborated into the principle of inversion of the citation-order, and (b) order as predictability.

Some Tools and Techniques of Order Building

There can be two polar-antithetical positions taken with regard to the order of a file of complex verbalisms: the purely symbolic, and the purely eidetic. There are variations within each such pole, of course, so that within the adherents of eidetic order, there are (a) those who want nothing but classification, "Out with words altogether!," (b) those who want chain-indexing and the like, and (c) those who favor a strictly subjective or semantic interpretation of the words and punctuations, especially of subject headings. Also, within the adherents of symbolic order, there is the traditional disagreement over the recognition of the blank: (a) recognizing it yields word-by-word order, (b) ignoring it yields letter-by-letter order. Nor are these all the possible shades of divergence.

The scope of the present inquiry must accordingly be restricted, lest the luxuriance of detail make the presentation of the polarity unintelligible; the appendices, however, contain additional matter. Classification as an alternative to the subject heading catalog will not be considered even though I favor it myself. Nor will the chain-index. Nor will letter-by-letter filing be admitted as an acceptable device. There still remains a polarity between the symbolic and the

eidetic; and, as stated earlier, what should be sought is the reduction of filing rules (as called for by the eidetic nature of the items being manipulated) to sorting rules (as called for by the symbolic embodiment of these items).

The reason for this reduction is the need for univocity. The computer "knows" nothing; it cannot, as Fairthorne quips, find just the important words, because for a word to be important it must mean something. The computer can only manipulate words; meaning means nothing to it. So the equivocacy inherent in discourse must be avoided; and, if this can be accomplished, the computer is left with things to manipulate that it can handle (without needing to understand their meanings).

The two polar-antithetical positions then refer to the structure of the file that is to be generated. As to the means that are to lead to these positions, there are at least three, from which each position can choose at least two as its means of embodiment. The three means are (a) the collating sequence as such, (b) the numerical-interpolation device, and (c) the "programming" of the cataloger.

If there is a sample set of complexes to file, such as "WELLS (SOMERSET)" [a place as subject], "Wells (Somerset)" [a place as author], "Wells of despair" [a title], "WELLS, HERBERT GEORGE" [a person as subject], "Wells, Herbert George" [a person as author], "WELLS" [a topical subject], and "WELLS, OIL" [another topical subject], there can be two classes of order set up amongst them in accordance with the eidetic or the symbolic positions:

$EIDETIC_1$	$SYMBOLIC_1$
Wells, Herbert George [peA]	WELLS [tS]
WELLS, HERBERT GEORGE [peS]	Wells, Herbert George [peA]
Wells (Somerset) [plA]	WELLS, HERBERT GEORGE [peS]
WELLS (SOMERSET) [plS]	Wells of despair [T]
WELLS [tS]	WELLS, OIL [tS]
WELLS, OIL [tS]	Wells (Somerset) [plA]
Wells of despair [T]	WELLS (SOMERSET) [plS]

As mentioned, variations are possible within each of these polar positions; for instance, within the eidetic position all authors, subjects, and titles can be segregated (the divided catalog approach); or, within the symbolic, distinction between single- and double-blanks can be observed, giving:

$EIDETIC_2$	$SYMBOLIC_2$
Wells, Herbert George [A,pe]	WELLS [first word followed by many blanks]
Wells (Somerset) [A,pl]	Wells, Herbert George
WELLS, HERBERT GEORGE [S,pe]	WELLS, HERBERT GEORGE

WELLS (SOMERSET) [S,pl]	WELLS, OIL	[first word followed by double-blank]
WELLS [S,t]	Wells (Somerset)	
WELLS, OIL [S,t]	WELLS (SOMERSET)	
Wells of despair [T]	Wells of despair [first word followed by single-blank]	

Now, consider the means for achieving these various results. The collating sequence, (a) may have too much expected of it to manage the eidetic orders, at least without nominal assistance from (c), the "programming" of the cataloger. The computer, in other words, can set first in order "Wells, Herbert George" as author and then as subject, because there is a difference in the characters of which each is composed—just as the filing clerk can tell apart black entries for authors and red entries for subject. But it can scarcely be expected to put places ahead of topics unless it has available a ridiculously large table of place names for reference. And that, in any case, is dependence on the cataloger too, as is depending on the tracing-group number-families to distinguish authors, titles, and subjects.

The numerical-interpolation device, (b) on the other hand, can achieve any desired order at all, but while it may be considered reasonable for authors and subjects—which can be counted on to return with some degree of frequency[15]—it is assuredly not so with titles, which can be counted on to be almost invariably varied, thus requiring a new number for almost every document. The numerical-interpolation device seems to me to have no place in the symbolic position, which by definition should not call upon any outside assistance, at least not beyond that which the cataloger does anyway, such as assignment of tracing-group number-families. (It does, however, have a very definite place in several clerical operations such as the periodical check-in file.)[16]

A tiny matrix can be constructed to show the intersection of these two sets of ideas (positions and means), with checks indicating mutual appropriateness:

	EIDETIC	SYMBOLIC
COLLATING SEQUENCE ALONE		X
NUMERICAL INTERPOLATION	X	
'PROGRAMMED' CATALOGERS	X	X

The "programming" of the cataloger, (c) the only one of the three devices necessarily involved with both the eidetic and the symbolic positions, enters unrecognized even in the apparently pure collating-sequence technique as it is what distinguishes between otherwise identical entries. But it has a larger role to play in

situations of a higher complexity than those shown in the sample, such as the quasi-facet structure of Bible entries. (For greater detail, see the Appendices and the end of this section.)

It is well to detail the actual operations required by each of these three techniques before attempting to proceed to an evaluation of results and a recommendation of goals. The collating sequence might, if assumed to encompass a sufficiently wide gamut of discriminations, achieve even fairly good results in the eidetic as well as the symbolic ordering of entries. However, it has been my observation* that the letters of the alphabet can be called self-justificatory insofar as they either transcribe the given elements of the document or embody the controlled vocabulary decisions of the cataloger. This distinction between the a-scriptive and the de-scriptive is brought out more fully in my paper "Documentary Relevance and Structural Hierarchy."[17] Our cataloging rules, contrasted with the policies of analytical bibliographers who treat punctuation and the like as being as sacrosanct as the words of the transcription, take an unfortunately cavalier attitude about punctuation. The rules avoid a clear-cut decision between two attitudes, one of which would lead to leaving the punctuation just as it is found (descriptive), the other to transforming given punctuation into functional (controlled, ascriptive). Thus, what the computer is fed, in terms of guidance for the application of the collating sequence (which, other than the alphabet and the numeral digits, should not be considered as frozen forever by the hardware manufacturers), is far from the kind of thing that can help in the decision, for instance, that parenthetical expressions like "(Somerset)" create a place sub-group. To do this, the parenthesis would have to be used for such a meaning and no other. As mentioned before, codes of descriptive cataloging do not bind their practitioners to a thoroughgoing usage of any particular punctuation symbol for any particular function. Nor, indeed, do codes for the rendering of headings. To insist on this consistency of usage would be, to some extent, further "programming" of the cataloger.

The two alternatives in the use of the collating sequence alone as a filing device—and note that such a device does indeed reduce filing to sorting—are (a) the superposition of the collating sequence onto present descriptive and entry practice, or (b) its superposition onto descriptive and entry practice thoroughly re-worked for the sake of such results. The operative presumption here is decimal, so much so that even integral numbers would have to be fitted out with leading zeroes (for instance, in dates of the Christian era earlier than A.D. 1000). The computer can then operate upon these atomized elements

---

*In the course of an investigation of the permissible compression of book-titles for the Bro-Dart Foundation. See Appendix IV.

(digits) of complex verbalisms, filing them by sorting them. But it must be noted too that unexpected difficulties can easily arise unless cataloging usage is modified, primarily in the direction of a far more stringent attitude about formats and the like.*

Use of the standard IBM collating sequence, for instance, in which the punctuations in our sample (comma, parentheses, blanks) are accepted as listed, gives (since the order built into the hardware is: blank, comma, opening parenthesis) $SYMBOLIC_2$. If policy were to dictate, instead, $SYMBOLIC_1$, it would be necessary to treat these various punctuations as blanks to be ignored in the creation of the "sort-tags." These then look (with "real" blanks indicated as "b") thus:

$$WELLS^{bbbbb}....$$
$$Wells^b Herbert^b George$$
$$WELLS^b HERBERT^b GEORGE$$
$$Wells^b of^b despair$$
$$WELLS^b OIL$$
$$Wells^b Somerset$$
$$WELLS^b SOMERSET$$

The numerical-interpolation device, (b) as mentioned is hardly applicable to symbolic-order filing, since it assumes decisions about order on the part of the cataloger. These decisions, once made, allow the computer to proceed to a kind of second-order symbolic-order sort. In the earlier example, for instance, the policy decision could be (instead) that $EIDETIC_1$ was the desirable order, and that a span of numbers, say 5000-6000, be allocated to all entries beginning "wells." The seven entries would thus generate an interval of $1000/7 = 142+$, which can be rounded out to 140. The seven entries are accordingly assigned numbers to "fossilize" their policy-determined order:

5070	Wells, Herbert George	5490	WELLS (SOMERSET)
5210	WELLS, HERBERT GEORGE	5630	WELLS
5350	Wells (Somerset)	5770	WELLS, OIL
	5910	Wells of despair	

Later additions to the file are assigned such numbers as will intercalate them in the place appropriate to the policy (for instance, "WELLS FAMILY," $5210 + 70 = 5280$). These numbers, once assigned, are either (a) automatically added to each entry by the computer from a look-up table (a program arrangement can provide for

*Daniel Gore's character, Melissa Spindrift, has a point in her fanatical insistence on obedience to trivial rules—at least in terms of the needs of computers. (See Reference 18.)

signalling the cataloger whenever an unassigned entry is input), or
(b) input by the cataloger on the basis of a reference table, thus sub-
stituting for the verbal entry-form entirely.

The device of "programming" the cataloger, (c), as mentioned,
is implied in every code of cataloging; but the needed specific con-
cern for the enabling of filing-manipulation is present only accident-
ally, if at all.

Substantial additions, dictated by the policy desired in terms of
variations within the eidetic position, would be required if this de-
vice were to be successful in reducing filing to sorting. For instance,
in the example of the two Bible entries, what is necessary to place
"Goodspeed" before "1964" would be the indication of facets: "^{6}Bible.
$[^{5,4}]$ ^{3}English. $[^2]$ ^{1}Goodspeed." as against "^{6}Bible. $[^{5,4}]$ ^{3}Eng-
lish. ^{2}1964. $[^1]$." (The blank facets "5" and "4" stand for section
and book, respectively.) An even more difficult example is the filing
of names such as those beginning with "John," which can be followed
by sobriquets, place-designators, given names, titles of nobility or
sovereignty, etc. Two lists of "John" entries are given in the Appen-
dices. Appendix I is a list of the subject-heading usage and ordering
of such entries, and Appendix II is a list of main entry usage and
ordering as drawn from LC printed catalogs. Analysis of these lists
shows that no principle is employed throughout which can be expected
to provide the predictability necessary to locate any single entry or
group of entries without scanning the whole list. Some seemingly
outlandish placements may be due to nothing more than individual
clerical errors rather than to the following of rules, but it should be
remembered that the more complex the rules (and the less they de-
pend on pure ritual) the more likely are cases of forgetfulness. How-
ever, in the filing of names of sovereigns, for any desired eidetic
order to be achieved, it is necessary not only to indicate empty facets
(as with the Bible entries), but to label each facet with a weighting
factor not determined by its position in the entry as written, in ac-
cordance with such a citation order as "Name. Rank. Area of
sovereignty. Number," as against the written order "Name. Number
[and sobriquet]. Rank. Area of sovereignty". This is not to say that
such relatively simple-minded means are capable of setting right
every problem of filing order. Note, for instance, the resumption of
a second sub-alphabet in the main entry list at "John Alcober." Sev-
eral entries that follow this one, although treated in accordance with
a logical rule similar to that mentioned in the earlier example of
"Mass (Canon law)"—namely that a longer first facet follows a short-
er, even where the shorter is followed by a subdivision-word of
higher alphabetic position—do not seem legitimately to be considered
as longer facets. They do not since the usual subdividing facet of the
first sub-alphabet is a sobriquet or a place-designator, such as
"John the Fearless, duke of Burgundy" or "John of Cappadocia".

The lengthening element of the supposedly longer first facet is similarly characterizable, for instance "John o'London, pseud." or "John Chrysostom, Saint" (Chrysostom is not a name, but a sobriquet meaning "golden-tongued.")

Implicit in these anomalies are the data that can lead to an important distinction between two aspects of the supreme desideratum, predictability. This distinction will be more fully discussed shortly.

An Attempt at Evaluation

Many solutions to these problems have been proposed, but if they do not recognize and control the actuality of the difficulties encountered in searching and in cataloging, they can be called "relatively simple-minded." Nor is it true that all such problems are necessarily real ones. A real problem in this area is one that not only presents us with an intricate tangle of circumstances, but one wherein a real need is prevented from being met. What are the real needs here? The most important, as can be seen from some of the foregoing, is predictability. To need to find a particular sovereign-person entry but to be unable to tell where to find it, even if the words used in it are known, is a need prevented from fulfillment. The same is true where it is necessary to find a particular subject heading made up of several phrases, even if one knows (or is told syndetically) the precise style even of the punctuation between the phrases; or where it is necessary to find a class of entries that form a group (among others) of divisions of some particular subject-heading or corporate entry. If it is impossible to predict where such entries or groups of entries are located in the file, there is a real problem.

Basically, the difficulty is with the lack of system in the cataloging policies in use among librarians. As a somewhat more protracted example, a discussion of the subject-heading "Art, Byzantine" follows. According to the LC authority-document Subject Headings,[19] the heading is among the fourth group of divisions of the heading "Art." The first of these sub-alphabets is that made up of (a) form-divisions such as "Art—Congresses" and "Art—Philosophy," as well as references from unused headings such as "Art—Negroes see Negro art." The second sub-alphabet is made up of (b) place names. If such a name were in use at present, we might expect to find "Art—Byzantium" here. The implication is, however, that this group is to comprise documents on the artifacts in a place, rather than the artistry characteristic of a place. The third sub-alphabet is somewhat of a hash, including inverted-phrase adjectival (c) styles-cum-periods like "Art, Abstract" and "Art, Baroque," (d) styles-cum-places like "Art, Cluniac" and "Art, Oriental," along with (e) miscellaneous

orientations like "Art, Commercial see Commercial art," "Art, Im-
moral," "Art, Muncipal," and "Art, Regional." (The application of
facet-analysis to the elements of this sub-alphabet reveals the full
extent to which lack of system can become enshrined in the hearts of
the unwary.) The fourth sub-alphabet, to which belongs "Art, Byzan-
tine," is made up of inverted-phrase adjectival (f) ethnic-cum-place
names like "Art, African," "Art, Celtic," and "Art, Greek" (to which
an xx-reference from "Classical antiquities" is prescribed, leaving
the user somewhat puzzled if he has in hand a document on modern
Greek art). Here is found the complement to the second sub-alphabet,
(b): such headings comprise the artistry characteristic of Byzantium,
rather than artifacts to be found in that city or empire. The fifth
sub-alphabet is composed of inverted-phrase adjectival (g) periods
like "Art, Ancient," "Art, Renaissance," and "Art, Modern—20th
century" arranged chronologically in defiance of their verbal form.
The sixth sub-alphabet comprises uninverted phrases beginning (h)
"Art and . . .," "Art in . . .," and the like.

Now let us assume a person in need of documentary information
about Byzantine art, a person to whom that turn of phrase appears
natural. The syndetic apparatus of LC subject headings will tell such
a person that a quasi-classification under "Art" is the proper place
to look for such information, rather than under the first-thought-of
expression. But it is not pointed out to him, except implicitly, where
such an entry can be predicted to be among all those beginning "Art".
What is explicitly (but still not thematically) pointed out is that the
form of the heading-to-be-sought is "inverted-phrase adjectival",
and that the adjective is formed from a proper name (for the sake of
argument, I am willing to grant that no one is ignorant of the general
meaning and formation of "Byzantine.")

Notice that this user cannot assign the heading "Art, Byzantine"
to the fourth (ethnic-cum-place) sub-alphabet under "Art" because
there is nothing in the syndetic link to tell him this. More funda-
mentally he cannot because even though he might be aware of the
groups generally to be expected under a broad heading like "Art" and
(c) even of the specific order of the ethnic-cum-place group among
the others, there remains a fundamental difficulty. He may very well
be thinking of "Art, Byzantine" as meaning something quite different
from that implied by its official place—as a style-cum-period or
style-cum-region heading (analogous to "Art, Gothic") in the third
sub-alphabet, or as a period as such in the fifth (analogous to "Art,
Renaissance"). This difficulty is similar to that noted earlier with
regard to some of the irregularities of order under "John": if a per-
son does not know the precise significance of part of a heading, an
eidetic order may well be such as to prevent him from finding it un-
less he takes the time and trouble to look through all the headings
beginning "John" or "Art". The difficulty arises because he does not

know the part of the heading's precise significance as used by the cataloger or filer, not that he does not know its precise form (which the syndetic apparatus reveals to him). The case of "Art" is analogous to that of "John" in the way in which it modifies the desideratum of predictability.

Notice also that in a case of sorting of meaningless and "atomic" signs (whether digit-particles like accents, digits, syllables—as in spoken Japanese—or even words—as in Uigur) the order is entirely arbitrary just because there is no meaning in their seriality. (The numerals "1" through "9" do not fall under this stricture except insofar as we presume the radix "10," and they are a paradigm then of both sorting and filing.) In an arbitrary sorting-order there is absolute predictability of the location of "q" between "p" and "r." The seriality of the whole alphabet (as governed by the Anglo-Saxon "arbitration") makes it possible to go from any entry-point in the series to the desired point: if one enters at "k", one knows that he must go "downstream" about half-way to the end of the series to hit "q."

In a file there is seriality too; but absolute predictability of the kind just mentioned requires an equally stringent application of arbitrary rules and an equally absolute awareness of the precise form of the sought sign-complex. It is the task of the syndetic apparatus to furnish the searcher with this absolute awareness because of the vagaries that seem inevitable in the cataloger's choice and rendering of entries. The seeker for "John Chrysostom, Saint" may well not be able to predict the location of such an entry precisely because he is aware of the fact that "Chrysostom" is a sobriquet, while the cataloger or filer is not.

A Conclusion of a Kind

It is easy enough to pick out flaws in any system, and far more so in any body of rules such as we are familiar with for filing, because none of them is really a system. I am not unsympathetic to the various attempts to provide the means for ordering a file of complex verbalisms. I am, however, convinced that only to those who look the difficulties full in the face and recognize the primordia of the concept of order in its various manifestations can there come hope of being able to devise or even to accept improvements.

The need for predictability is paramount, but there is not just one predictability to be sought for. In terms of the bifurcation "eidetic/symbolic" these insights appear: (a) symbolic order makes possible the predictability of entries known in precise detail, whereas (b) eidetic order makes possible the predictability of entries known in terms of their meaning. For instance, entries beginning

"M," "Mac," and "Mc," either in a separate alphabetic sector before or after all other "M. . ." entries or in a sub-sector of "M. . ." (as if all spelled "Mac"), make possible finding entries like "McClare," "MacClennan," and "M'Clore" in that order, without the searcher needing to advert to the precise form of the prefix. This is an example of the advantage of eidetic order. One advantage of the symbolic order is that the searcher for "Art, Byzantine" could afford to ignore the membership of this heading in one of the sub-alphabets. Instead, he could be assured that all comma-subdivisions of any heading like "Art" would be in a predictable place, distinct from double-dash-subdivisions, full-stop-subdivisions, parenthesis-subdivisions, or any other such.

Can we afford to sacrifice one such advantage to gain the other? I do not believe so; and I believe that a combination of two of the techniques outlined earlier, namely the collating sequence and the "programming" of the cataloger, can effect the solution desired without sacrifice of either advantage. These elements of a putatively adequate filing system are both indispensable anyway, since all filing implies (a) ritually sortable elements to constitute its filed complexes, and it is the collating sequence that makes such ritual sorting possible. It also implies (b) semantic content to these complexes, and the complementary control of them by the application of bibliographical control by the cataloger.[20] If these elements are combined with foresight, there can arise these two distinct and important results: (a) an eidetically determined filing order symbolically sortable, and (b) a flexible computer-contained surrogate capable of serving as a basis for a wide variety of output purposes,[21] both ritually manipulable. By such a combination of devices, which would imply the tagging of a carefully determined totality of the sub-elements of the whole complex of descriptive elements, there could be brought about any order desired, as long as such a desired order was based on ritually consistent principles, and any degree of abbreviation of the entry as a whole. The main point to grasp here is that the greater the number of uniquely tagged sectors, the greater the flexibility for both these purposes. These sectors will not be easy to enumerate, nor will it be a simple matter to devise unique codes for them. It is probably possible only with the use of mutually modifying codes.

For instance, the combination of eidetic and symbolic order, through the conscious cooperation of the cataloger and the collating sequence, could give this order to the sub-alphabets of "Art": "Art —Congresses" and other forms, "Art: London" and other places in which. . ., "Art 'Gothic'" and other styles etc., "Art; Byzantine" and other places of which. . ., "Art/1901- " (=20th century modern) and other periods, "Art and . . ., Art in . . ." and other uninverted phrases, and finally, though not given (nor needed here) in Subject Headings, "Art (Qualified)." (In such an arrangement the punctuation used must

not be allowed to be used for any other purpose, at least not in subject-headings. This is what I mean by mutually modifying codes.) Thus the heading "Art, Byzantine" would either become "Art; Byzantine," with see-references from "Art 'Byzantine'" and "Art/[Byzantine period]" and (perhaps) "Art: Byzantium"—to force all usages of the inverted phrase to be located at the putatively most helpful place, or there could be entries allowed under each of the three, depending upon the emphasis of the document being cataloged (ethnic/geographic, stylistic, or temporal), with see-also references connecting them.

The steps to such a solution must include (a) comparative analysis of the various eidetic and symbolic orders advocated, in the light of basic principles such as have been enunciated here, (b) tabulation of the sectorings necessary to achieve all these orders—insofar as they can be seen to be truly desirable in terms of the real problems of searching—from a single input-record, and (c) checking to see whether additional sectoring is necessary to enable selective abridgements. This ambitious program does not, of course, mean that we cannot achieve improvements right now. Already, in the design of various computerized cataloging systems (as at Florida Atlantic University,[22] the Ontario New Universities Library Project,[23] the University of California at Santa Cruz,[24] and the medical libraries' union catalog at Yale, Harvard, and Columbia Universities[25]), various systems of use of computers in filing have been worked out. These may have been less than ideal solutions, but if they were, it is primarily because of the ad hoc nature of the attempts. They were not worked out as consciously partial solutions of the problem as a whole. The combination of a careful re-examination of our prevalent attitudes on preferred eidetic order(s), of the introduction, where necessary, of flexibility in the collating sequence, of the development of the codifications which can make cataloging rules bear along with them all that is needed for the accomplishment of the two central purposes—this is what ultimately must be sought, even when we wish to make only small practical improvements.

REFERENCES

1. Ranganathan, Shiyali Ramamrita. Heading and Canons; Comparative Study of Five Catalogue Codes. Madras, S. Vigwanathan and Sons, 1955, p. 60.

2. Hines, Theodore C. and Harris, Jessica L. Computer Filing of Index, Bibliographic and Catalog Entries. Newark, Bro-Dart Foundation, 1966.

3. A.L.A. Rules for Filing Catalog Cards. Prepared by a Special Committee, Sophie K. Hiss, Chairman. Chicago, American Library Association, 1942.

U. S. Library of Congress—Processing Department. Filing Rules for the Dictionary Catalogs of the Library of Congress. Rev. ed. of the 1945 Filing Manual. Washington, Library of Congress, 1956.

4. Perreault, Jean M. "The Computer and Catalog Filing Rules," Library Resources & Technical Services, 9:325-331, Summer 1965.

5. McMurtrie, Douglas Crawford. The Book; The Story of Printing and Bookmaking. 3d rev. ed. New York, Oxford University Press, 1943, pp. 95-97.

6. Cf. e.g. Robinson, Jane J. The Transformation of Sentences for Information Retrieval. Santa Monica, California, The RAND corporation (P-3243), December 1965 (Paper presented at the 1965 Congress of International Federation for Documentation (FID), Washington, D. C., October 1965).

7. See e.g. MacLuhan, Herbert Marshall. The Gutenberg Galaxy; The Making of Typographic Man. Toronto, University of Toronto Press, 1962, pp. 159-161.
"Applied knowledge in the Renaissance had to take the form of translation of the auditory into visual terms, of the plastic into retinal form;" note in particular his quotation from Father Ong (p. 160): "The spatial processing of sound by means of the alphabet is not enough. Printed or written words themselves must be deployed in spatial relationships, and the resulting schemata thought of as a key to their meanings."

8. Hjelmslev, Louis. "Prolegomena to a Theory of Language," Translated by Francis J. Whitfield In Indiana University Publications in Anthropology and Linguistics, Vols. 1-10 (1948-54): Memoir 7, pp. 29-38.
He points out here the non-cotermineity of the individual members of two groups (in English and Welsh) of words—and again, of three (in Danish, German, and French)—which, collectively, are coterminous.

9. Ceccato, Silvio, et al. Linguistic Analysis and Programming for Mechanical Translation. New York, Gordon and Breach, 1961, p. 59 ff.

10. Lubetzky, Seymour. Letter dated Oct. 13, 1964.
This is forcefully pointed out (in answer to my paper cited in reference 4), in a passage whose import I have attempted to integrate with my own prevailing conviction of the supremacy of the collating sequence (i.e. of sorting as the ideal ritual to make possible

computer-filing); he argues that "the basic difficulty lies in our use of the same symbols—the dash or the comma—for various types of subdivisions, and I believe that the development and use of a system of symbols to designate and distinguish the different types of sub-division would add a new dimension to the organization of the catalog and contribute materially to its intelligibility and effectiveness." This hope stems from his conviction that "to theorize that filing is implicitly alphabetization, not classification, is quite erroneous and misleading." I am not willing to go that far, at least terminologically, but the same insight is stressed here that motivated Ranganathan's statement quoted at the beginning of this paper.

11. The Prussian Instructions; Rules for the Alphabetical Catalogs of the Prussian Libraries. Translated from the Second Edition, Authorized August 10, 1908, with an Introduction and Notes by Andrew D. Osborn, Ann Arbor, U. of Michigan Press, 1938, pp. 27-106 et passim.

12. Ranganathan, op. cit., pp. 59-63.

13. Fugmann, Robert. "Ordnung—oberstes Gebot in der Dokumentation," Nachrichten für Dokumentation, 13:121, Number 3, Sept. 1962.

14. Needham, Christopher D. Organizing Knowledge in Libraries. London, Andre Deutsch, 1964, pp. 181-183.

15. According to an unpublished study made by Barbara Markuson of the Information Systems Specialist's Office of the Library of Congress, 47 percent of the authors examined had only one entry, 61 percent one or two, 66 percent one, two, or three; thus the hope for consistent return of author entries may be disappointed as well.

16. Cf. e.g. Srygley, Ted F. "Serials Record Instructions for a Computerized Serial System," Library Resources & Technical Services, 8:248-256, Summer 1964.

Schultheiss, Louis A. "Two Serial Control Card Files Developed at the University of Illinois, Chicago," Library Resources & Technical Services, 9:271-287, Summer 1965.

Stewart, Bruce Warren. A Computerized Serials Record System for the Texas A & M University Library. College Station, Texas A & M University Library, 1965.

17. Perreault, Jean M. "Documentary Relevance and Structural Hierarchy," Information Storage and Retrieval, 3:13-18, Number 1, 1966.

18. Gore, Daniel. "Encounter with a Cataloger," Library Resources & Technical Services, 9:363-366, Summer 1965.

19. U. S. Library of Congress. Subject Headings Used in the Dictionary Catalogs of the Library of Congress. 6th ed. Washington, D. C., 1957.

20. For discussion of the nature and function of bibliographical control in its various senses, see my paper: Perreault, Jean M. "On Bibliography and Automation or How to Reinvent the Catalog," Libri, 15:287-339, Number 4, 1965.

21. Cf. e.g. Buckland, Lawrence F. The Recording of Library of Congress Bibliographical Data in Machine Form. A Report Prepared for the Council on Library Resources, Inc. Rev. Washington, Council on Library Resources, 1965.
        Note that Buckland's results by no means exhaust the possibilities, which can be made to extend from the full-standard analytical description to the briefest index entry.

22. See: Heiliger, Edward M., et al. "Florida Atlantic University," College & Research Libraries, 25:181-199, Number 3, 1964; and Perreault, Jean M. "Computerized Cataloging: The Computerized Catalog at Florida Atlantic University," Library Resources & Technical Services, 9:20-24, Number 1, 1965.

23. See: Bregzis, Ritvars. "Compatible Format for Machine Readable Record of Bibliographic Data," paper read at the 1965 Congress of the FID; Proceedings of the Congress to be published mid-1966; and also his "The Ontario New Universities Library Project-an Automated Bibliographic Data Control System," College & Research Libraries, 26:495-508, Number 6, 1965.

24. See: Computer Usage Company. Specification for an Automated Library System. Palo Alto, 1965.

25. See: Kilgour, Frederick G. "Development of Computerization of Card Catalogs in Medical and Scientific Libraries." In Herbert Goldhor, ed., Proceedings of the 1964 Clinic on Library Applications of Data Processing. Urbana, Illinois, University of Illinois Graduate School of Library Science, 1964, pp. 25-35.

ADDITIONAL REFERENCES

Graduate Research Center of the Southwest. Automation and Inter-Library Information Systems. A study sponsored by the Graduate Research Center of the Southwest, 1:24-105, Dallas, 1964.

Weinstein, Edward A., and George, Virginia. "Notes Toward a Code for Computer-Produced Printed Book Catalogs," Library Resources & Technical Services, 9:319-324, Number 3, 1965.

Appendix I

A list of subject heading uses of "John. . .," from: Library of Congress Catalog; (Books: Subjects) a Cumulative List of Works Represented by Library of Congress Printed Cards, 1955-1959, 11: 580-581. Patterson, New Jersey, Pageant Books, Inc., 1960.

JOHN BERCHMANS, SAINT
JOHN BOSCO, SAINT
JOHN CAPISTRAN, SAINT
JOHN CLIMACUS
JOHN FISHER, SAINT
JOHN FREE
JOHN, KING OF ENGLAND
JOHN OF GAUNT, DUKE OF LANCASTER
JOHN OF JANDUN
JOHN OF KRONSTADT, FATHER
JOHN OF LEYDEN
JOHN OF ROQUETAILLADE
JOHN OF SALISBURY, BP. OF CHARTRES
JOHN OF THE CROSS, SAINT
JOHN XXIII, POPE
JOHN, PRESTER
JOHN, SAINT, APOSTLE
JOHN THE BAPTIST
JOHN VIANNEY, SAINT
JOHN, OTTO

JOHN (BOOK OF THE NEW TESTAMENT)
JOHN, EPISTLES OF (BOOK OF THE NEW TESTAMENT)
JOHN BROWN RAID, 1859
JOHN CRERAR LIBRARY, CHICAGO
JOHN DICKINSON AND COMPANY, LTD.
JOHN GOFFE'S MILL, BEDFORD, N. H.
JOHN HANCOCK BUILDING, SAN FRANCISCO
JOHN LEWIS PARTNERSHIP, LTD.
JOHN MORE ASSOCIATION
JOHN MUIR TRAIL, SIERRA NEVADA MOUNTAINS
JOHN NEWBERY MEDAL BOOKS
JOHN RYLANDS LIBRARY, MANCHESTER
JOHN THE BAPTIST'S DAY
JOHN WOODMAN HIGGINS ARMORY, WORCESTER, MASS.

Appendix II

A list of author uses of "John. . .," from: A Catalog of Books Represented by Library of Congress Printed Cards, Issued to July 31, 1942, 76:616-621. Ann Arbor, Michigan, Edwards Brothers, Incorporated, 1944.

John III, King of Portugal
John, abbot of Ford
John (bar manager)
John, bp. of Lincoln
John, bp. of Nikiu
John, Brother

John de Bordeaux
John de Britto
John de Halton
John de Pontissara, bp. of Winchester
John de Sandale, bp. of Winchester

John, de Taxster
. . .
John of Burgundy
John of Cappadocia
John, of Damascus, Saint
John of Gaddesden
John, of Garland
John of Gaza
John of Genoa
John of Halton, bp. of Carlisle
John, of Hildesheim
John of Hoveden
John of Ire'land
John of Jandun
John of Lancaster, pseud.
John, of Nepomuk, Saint
John of Pian
John de Carpine
John of Ravenna
John of Salisbury, bp. of Chartres
John, of Tynemouth
John of Wales
John of Worcester
John, Prester
John, son of Sabanis
John, the Evangelist
John the Faster
John, the Fearless, duke of
   Burgundy
John, the Painter

John Wallensis
John Alcober
John Ambrose, father
John Amundesham
John Baptist, archduke of Austria
John Boston, of Bury
John Chrysostom, Saint
John de Burgh
John de Oxenedes
John Francis Regis, Saint
John Gabriel, sister
John Gualbert
John Henry, pseud.
John Joseph Laurence Charles
   Louis, archduke of Austria
John Mary of Reifenberg, father
John o'London, pseud.
John of God
John of Gaunt, duke of Lancaster
John of Saint Thomas
John Paul, pseud.
John Sackland
John III. Sobieski, King of
   Poland
John Zapolya, King of Hungary
John, Alfred

Appendix III

Inverted Citation Order

        In order to insure that whole comes before part, general before
specific, the principle of citation order in classification is subjected
to inversion. Citation order results in predictability among the indi-
vidual concepts ascribed to a single document; inversion of citation
order results in predictability among the complexes (i.e., the docu-
ments themselves or their surrogates).
        Two techniques are displayed below, either of which can assure
proper inversion. There must, however, be a decision ahead of time

as to the citation order upon which the inversions are to be based. Or, if there seems no clear preference as to citation order, filing order can be decided first, and the citation order can be derived as its inversion. In practice the two lines of attack may be exercised simultaneously.

TECHNIQUE A. With postulated facets "place," "period," and "form," of a subject Y, we can construct a matrix showing citation order, and then deduce all the patterns of occurrence:

	"place"	"period"	"form"	
Y			x	(document a)
Y		x		(b)
Y		x	x	(c)
Y	x			(d)
Y	x		x	(e)
Y	x	x		(f)
Y	x	x	x	(g)

The occurrences are in terms of foci from each facet, of course. For instance, if Y is "architecture," document (a) could be a bibliography of architecture, (b) a study of architecture in the 20th century, (c) a bibliography of architecture in the 20th century, (d) a study of the architecture of France, (e) a bibliography of architecture in France, (f) a study of 20th century architecture in France, and (g) a bibliography of 20th century architecture in France.

TECHNIQUE B. Alternatively, the citation order can be expressed as an (ad hoc) weighting device. If the facet "place" is to come first in the citation order, it should be accorded the greatest weight, "period" the next greatest, "form" the least, and empty facets zero. Thus "place" = 3, "period" = 2, and "form" = 1, so that document (a) receives the weight 001, (b) 020, (c) 021, (d) 300, (e) 301, (f) 320, and (g) 321, putting them into the same order as before.

These techniques are appropriate not only to subject classification, but to the cataloging code as well. For instance, facet analysis is implicitly present in the construction of anonymous-classic headings such as:

Bible. English. Goodspeed.
Bible. English. 1964.
Bible. N.T. English. 1964.
Bible. N.T. Galatians. English. Goodspeed.
Bible. N.T. Galatians. English. 1964.
Bible. N.T. Galatians. 1964. Goodspeed.

"part"	"book"	"language"	"date"	"version"
		x		x
		x	x	
x		x	x	
x	x	x		x
x	x	x	x	
x	x	x	x	x

The basic principle operating here is "nothing before something": the unspecified before the specified, the whole before the part, or the general before the specific. The desired effect is predictability—here, of <u>finding</u> the desired document or surrogate (while citation order makes the analysis of the document and the <u>creation</u> of the surrogate predictable).

Appendix IV

Compression of Titles

A sample of titles was chosen from the <u>Catalog of Books Represented by Library of Congress Printed Cards . . . 1942,</u> according to the following rule: the first English title in the third column or the next following page on which an English title appears, of pp. 100, 200, 300, etc. of vol. 1, the same from pp. 101, 201, 301, etc. of vol. 3, progressing similarly through each odd-numbered volume up to vol. 23; whereafter it was re-cycled from p. 100 of vol. 25, 101 of vol. 30, etc. This resulted in a random sample of 235 titles. English was the only language accepted, not because there is no problem with foreign titles, but because the probability of coming to a clear understanding of basic principles is higher under such restrictions.

Two simultaneous tabulations were made:

(A) Of types of main entry encountered, which yielded 11 title main entries, 10 corporate main entries, and 214 personal main entries. The smallness of the proportion of entries not of the "normal" pattern, <u>i.e.,</u> of a title distinctly subordinate to a personal responsibility, is significant, since the form of title in such "normal" cases is not as crucial in the finding operation as in corporate main entries or in title main entries. (It was not tabulated whether some of the "odd" 21 might, by a different decision, have been considered personal main entries, nor whether any of the "normal" 214 might have been considered "odd" by a different code for choice of entry.) It is perhaps interesting to note that none of the other main classes of main

entry occurred, namely anonymous classics or uniform headings.

(B) Of the structure of the titles themselves in terms of additional verbiage beyond the "core"-title, which yielded

      1 alternative title as expansion of acronym,

      1 sectional title,

      2 mentions of contents,

      4 initial words or phrases to be arbitrarily omitted as insignificant,

      5 mentions of redaction,

      7 initial possessives,

      9 alternative titles introduced by "or,"

      9 supplementary sub-titles,

      11 mentions of occasion,

      24 form sub-titles,

      30 explanatory sub-titles,

      39 mentions of scope, and

      91 initial non-filing articles.

And, where a new class can be defined by the simultaneous membership of one title in two or more of the above in the same sector of the title, the tabulation yielded

      1 explanatory/form/supplementary sub-title,

      1 mention of occasion/explanatory sub-title,

      2 alternative/explanatory sub-titles,

      2 mentions of occasion/form sub-titles,

      2 mentions of redaction/supplementary sub-titles,

      4 mentions of scope/explanatory sub-titles, and

      4 explanatory/form sub-titles.

The significant judgment to be made on these structural types is one of these three: (a) Is there a sector other than the "core"-title that should also be traced? (b) Is there a sector other than the "core"-title that should be omitted altogether at the time of compression? (c) Is there a sector other than the "core"-title that is necessary for the intelligibility of the "core"-title, but which should not be re-traced even when retained? Those classes that satisfy (a) are mentions of contents, sectional titles, and alternative sub-titles; those that satisfy (b) are explanatory and supplementary sub-titles, as well as some mentions of scope and some form sub-titles; those that satisfy (c) are mentions of redaction and the residual mentions of scope and form sub-titles. This, in terms of the sample, yields 12 sectors to be re-traced, ±71.5 to be dropped, and ±37.5 to be retained but not retraced. The remaining titles did not have such sectors, consisting of "core"-title alone.

Can the computer abbreviate titles according to these criteria? It might be that it could do so if standardized catalog-entry data were such as invariably to indicate these various functions—but the lack of

precision as to the various types of non-"core" sectors of titles, outside this investigation, makes it a priori impossible for such an indication to be expected. Something close to such an indication is implied by the practice of introducing alternative title in "or" by the invariable punctuation "; or,". But examination of the titles in the sample leads to near-total lack of faith as far as any expectation of such consistency from LC cataloging is concerned. (It can, of course, be argued that this inconsistency is traceable back to the title-pages themselves, but any cataloger can easily dispute such an allegation, there being all sorts of arbitrary re-punctuations to be found in LC cataloging—although not for the sake of the establishment of any function such as automatic filing/tracing/abbreviation decisions.)

It is therefore recommended that four punctuation symbols be used in titles for purposes of functional sectoring, and that they not be allowed to be used in their common senses. They are:

(a) the colon (:) to indicate the re-tracing of mentions of content and of sectional titles,

(b) parentheses (()) to indicate the re-tracing of alternative subtitles,

(c) the semi-colon (;) to indicate that the subsequent data are not to be retraced or even retained at the time of compression, and

(d) the double-dash (--) to indicate the retention of the subsequent data, without re-tracing it.

One further symbol is to be used in a restricted way, the virgule (/); see below. This leaves enough punctuation symbols with their common usages (. , ? " ' - * &) to avoid problems in the ordinary grammatical organization of the entry.

There is a large number of words which recur often enough to justify their invariable abbreviation, at least at the time of compression. The sample from which the following list is derived is certainly not large enough to lead us to think that they form an exhaustive tabulation, but they do form a sort of nucleus:

Amer(ican)	ess(ential)	p(ar)t
b(oo)k	extr(act) (acted)	pol(itical)
coll(ege)	gen(eral)	pr(inted) (inting)
collec(ted) (tion)	G(rea)t Br(itain)	publ(ished) (ication)
compl(ete)	h(an)db(oo)k	repr(inted) (inting)
conc(erning)	interpr(etation)	sect(ion)
cons(idered)	intro(duction) (duc-	sel(ected) (ection)
dev(elopment)	tory)	s(outh)
diss(ertation)	iss(ue)	tr(anslated) (anslator) (ans-
e(ast)	libr(ary)	lation)
ed(ited) (itor) (ition)	man(ual)	univ(ersity)
educ(ation)	mo(nth)	U(nited) S(tates)
Eng(lish)	n(orth)	w(est)
	org(anization)	y(ea)r

The standard abbreviations for the days of the week, the months of the year, the states of the Union should be used plus those for form-divisions given in <u>Subject Headings</u>.

Such words are not as frequent of repetition as are certain prepositional-phrase-introducers (which, while they do not each take up so much space, do so in the aggregate), such as:

> as regards (a, an, the)
> at (a, an, the)
> by (a, an, the)
> for (a, an, the)
> from (a, an, the)
> in (a, an, the)
> of (a, an, the)
> to (a, an, the)
> with (a, an, the).

It is recommended that these phrase-introducers, together with the article following them, be programmed to be reduced to the virgule at the time of compression.

There are accordingly three steps in compression of titles, and they can be most satisfactorily accomplished in two stages in each of two possible ways:

(either)	(or)
(a,b) functional punctuation inserted by the cataloger; abbreviation of frequent words by the cataloger	(a) functional punctuation inserted by the cataloger
(c) programmed excision of phrase-introducers (='/')	(b,c) programmed abbreviation of frequent words; programmed excision of phrase-introducers (='/')

These techniques, when applied to the 235-title sample, result in compression of all titles to 40 or fewer characters except in 19 cases; the compressed titles thus truncated are somewhat less than intelligible in all but 8 cases. It seems unlikely that any further compression can be achieved except in terms of further truncation and of a direct increase in the number of unintelligibilities that result—except by the use of a longer list of frequent words to abbreviate.

The final problem is that of the elimination of words and phrases standing at the head of the title, rather than being sub-sectors coming after the "core"-title. There are three types of such initialisms to be eliminated: simple initial articles, initial possessives referring to the main entry, and insignificant phrases of introduction. Initial articles (in English) can be quite easily programmed out of the filing order in the same computer-pass which reduces phrase-introducers to virgules; however, the total elimination of such

articles in the compressed print-out is not recommended except where it is noticed (by program) that such retention will result in truncation of sectors neither re-traced nor to be eliminated, namely if the title is sectored by a double-dash or is not sectored at all. Of the 19 titles truncated at character 41, 12 began with articles, and the application of the ritual questions here recommended would have mollified the resultant unintelligibility by dropping the initial articles in 11 of those cases.

Initial possessives should always be eliminated from the very first if they refer to the main entry; there is only one case in the sample that does not reproduce the main entry, and must accordingly not be dropped. This "exception" occurs not so seldom in the real world, especially in humanistic and historical studies (e.g., "Aristotle's Contribution to . . .," by Barker, Mansion, or the like).

The 4 arbitrarily excised initialisms (type c) are characteristically equivalent to form sub-titles, but their position causes trouble in that the same symbol cannot be used to exclude them that was used for terminal exclusion of form sub-titles, namely the semi-colon. Accordingly, it is recommended that such initialisms, when recognized by the cataloger as accompanying a title that will result in being truncated by compression, should be excised from the very first, so as not to appear in the full title any more than in the compressed. Perhaps, as a concession, the brief lacuna-symbol (..) could be used to indicate such excision.

The titles that make up the sample from which these conclusions are drawn, first in raw form (with indication of structural type of sectoring) and then as compressed (with truncated parts underlined), are

History of King Richard the Second of England
HIST./KING RICHARD THE 2D/ENG.

Elementary mechanical drawing, for school and shop (explanatory/
    scope)
ELEMENTARY MECHANICAL DRAWING--SCHOOL + SHOP

The Joseph Leidy commemorative meeting, held in Philadelphia,
    Dec. 6, 1923 (occas.)
THE JOSEPH LEIDY COMMEMORATIVE MEETING; philadelphia,
    Dec. 6, 1923

A relation of the great river of Amazons in South-America (scope)
A RELATION/GREAT RIVER/AMAZONS; s. amer.

Provincial society, 1690-1763 (scope)

Address to the people of Pennsylvania (title main entry)
ADDRESS/PEOPLE/PA.

War against war; or, The joys of peace (alternative)
WAR AGAINST WAR (the joys of peace)
THE JOYS OF PEACE

A manual of instruction in Latin on the basis of a Latin method. . .
  (explanatory)
A MAN./INSTRUCTION/LATIN;/basis/latin method. . .

The Sumner assault. Speech of. . . (occasion, form)
THE SUMNER ASSAULT; speech of. . .

Hungary before, during and after the great war (scope)
HUNGARY--BEFORE, DURING + AFTER THE GREAT WAR

The Amaranth; a gift for all seasons (title main entry, supplementary)
THE AMARANTH; a gift for all seasons

Bulletin (corporate main entry)

The perfect tribute

Annable's grain calculator, showing at a glance the number. . . (init.,
  explan.)
GRAIN CALCULATOR; showing/glance the number. . .

Queen Elizabeth

The chace

Arbor day planting in eastern states (title main entry, scope)
ARBOR DAY PLANTING--EASTERN STATES

The Arena (title main entry)

Report of audit (corporate main entry)

Farm experiments with fertilizers
FARM EXPERIMENTS/FERTILIZERS

Guide to the material on Swedish history in the Augustana college
  library (scope)
GUIDE/MATERIAL/SWEDISH HIST.--AUGUSTANA COLL. LIBR.

Adelaide

Consular and other conventions between Austria-Hungary and Servia
  (initialism)
CONVENTIONS/AUSTRIA-HUNGARY + SERVIA

Translation of the Mining law of Mexico, effective August 1, 1926. . .
  (init., occas.)
MINING LAW/MEXICO; effective. . .

Progress in the science of education in the last twenty-five years
  (scope)
PROGRESS/SCIENCE/EDUC.--LAST 25 YRS.

A collection of rare and valuable books
A COLLECTION/RARE + VALUABLE BKS.   ('collection' not abbrev.,
   because first word)

Q 39

The Barbers' journal (title main entry)

An introduction to the Italian language
AN INTRO./ITALIAN LANGUAGE

Marti, a story of the Cuban war (explanatory/form)
MARTI--A STORY/CUBAN WAR

A brief sketch of the history of the Catholic church in the island. . .
   (init., scope)
THE CATHOLIC CHURCH IN THE ISLAND. . .

A sketch of the doctrine relative to commitments in bankruptcy
THE DOCTRINE RELATIVE/COMMITMENTS IN BANKRUPTCY

Sex, marriage and motherhood
SEX, MARRIAGE + MOTHERHOOD

Reconnaissance of the gold fields of southern Alaska with some
   notes. . . (suppl.)
RECONNAISSANCE/GOLD FIELDS/S. ALASKA;/some notes. . .

The eight crooked trenches

All in it

History of the Forty-eighth Ohio vet. vol. inf. giving a complete. . .
   (explan.)
HIST./48TH OHIO VET. VOL. INF.; giving a compl. . . .

Labor disputes and the President of the United States
LABOR DISPUTES + THE PRESIDENT/U. S.

The fundamentals of naval tactics
THE FUNDAMENTALS/NAVAL TACTICS

Joseph King Fenno Mansfield

Personal and literary memorials

Germany before the war
GERMANY--BEFORE THE WAR

Solar hydrodynamics

Commentaries on the law of England. . . the 19th ed. With the last
   corrections of the author, and copious notes. By J. E. Hoveden
   (redaction)
COMMENTARIES/LAW/ENGLAND--/NOTES/HOVENDEN

Co-operative movement in Russia (scope)
CO-OPERATIVE MOVEMENT--RUSSIA

The roads and road material of Indiana (scope)
THE ROADS + ROAD MATERIAL/INDIANA

Agricultural climatology of the United States compared with that of
  other parts of the globe (scope embedded in 'core')
AGRICULTURAL CLIMATOLOGY/U.S. COMPARED/OTHER PARTS/
  GLOBE

Effects of various dietary deficiencies upon the morphology of the
  suprarenal gland
EFFECTS/VARIOUS DIETARY DEFICIENCIES/MORPHOLOGY/
  SUPRARENAL GLAND (or, to force re-tracing of the second
  sector:)
EFFECTS/VARIOUS DIETARY DEFICIENCIES:MORPHOLOGY/
  SUPRARENAL GLAND
MORPHOLOGY/SUPRARENAL GLAND

Psychology; a factual textbook (form)
PSYCHOLOGY; a factual textbook

The history of educational legislation in Ohio from 1851 to 1925
  (scope)
THE HIST./EDUC. LEGISLATION--OHIO, 1851-1925

Boston investigator, devoted to the development and promotion of uni-
  versal mental liberty (explanatory, title main entry)
BOSTON INVESTIGATOR; devoted. . .

The art of dressmaking at home and in the workroom (scope)
THE ART/DRESSMAKING--/HOME +/WORKROOM

Canada under British rule, 1760-1905 (scope)
CANADA UNDER BRITISH RULE--1760-1905

A history of Philadelphia, with notice of villages, in the vicinity
  (supplementary)
A HIST./PHILADELPHIA; / notices/villages,/vicinity

The drama and music in New South Wales (scope)
THE DRAMA + MUSIC--NEW S. WALES

The layman's handbook to Daniel
The LAYMAN'S HDBK./DANIEL

The Jacobins; an essay in the new history (form)
THE JACOBINS; an essay. . .

New reading room and libraries
NEW READING ROOM + LIBRS.

British standard specification for the dimensions of grinding wheels
    and methods of attachment (initialism)
SPECIFICATION/DIMENSIONS/GRINDING WHEELS + <u>METHODS OF</u>
    <u>ATTACHMENT</u>

Lectures illustrative of various subjects in pathology and surgery
LECTURES ILLUS./VARIOUS SUBJECTS/PATHOLOG<u>Y + SURGERY</u>

The Senate finance bill

Sanctuary, Sunshine house sonnets (form)
SANCTUARY--SUNSHINE HOUSE SONNETS

Lands of the Andes and the desert
LANDS/ANDES + THE DESERT

Catalogue of the library and prints of. . . of. . .
CATALOGUE/LIBR. + PRINTS/. . .

The crimson cutlass

The Pan-American conferences and their significance
THE PAN-AMER. CONFERENCES + THEIR SIGNIFIC<u>ANCE</u>

Summer tours (corporate main entry)

The blessed bees

Moses Greeley Parker, M.D.

A tentative bibliography of the belles-lettres of the Argentine republic
    (scope)
A TENTATIVE BIBL./BELLES-LETTRES/ARGENT<u>INE REPUBLIC</u>

The third part of the Institutes of the laws of England: concerning
    high treason, and other pleas of the crown, and criminal causes
    (sectional)
THE 3RD PT./INSTITUTES/LAWS/ENG.: CONC. HIGH TREASON, +
    OTHER PLEAS/CROWN
CONCERNING HIGH TREASON, + OTHER PLEAS/C<u>ROWN</u>

English teaching in the Southwest; organization and materials for in-
    structing Spanish-speaking children (scope/explanatory)
ENG. TEACHING--S.W.; <u>org. + materials/instructing spanish-</u>
    <u>speaking children</u>

Christian salvation, a modern interpretation (explanatory)
CHRISTIAN SALVATION; <u>a modern interpr.</u>

Short texts from Coptic ostraca and papyri
SHORT TEXTS/COPTIC OSTRACA + PAPYRI

Meat purchasing a science

Suggestions for teaching primary reading and phonics
SUGGESTIONS/TEACHING PRIMARY READING + <u>PHONICS</u>

Saint Brigid of Ireland

North American ichneumon-flies, new and described, with taxonomic
   . . . (suppl.)
N. AMER. ICHNEUMON-FLIES--NEW + DESCRIBED; /taxonomic +
   <u>. . .</u>

The Greek herbal of Dioscorides; illustrated by a Byzantine, A.D. 512;
   Englished by John Goodyer, A.D. 1655; edited and first printed,
   A.D. 1933, by Robert T. Gunther (form, redaction, internal ref. to
   main entry)
THE GREEK HERBAL; illus./byzantine, a.d. 512; eng./goodyer, a.d.
   1655--ED. + . . .--ED. + 1ST PR., A.D. 1933/GUNTHER

Annelids from the Danmark expedition (scope)
ANNELIDS--DANMARK EXPEDITION.

Old Kensington palace and other papers
OLD KENSINGTON PALACE + OTHER PAPERS

Doderlein's Hand-book of Latin synonyms (initialism)
HAND-BOOK/LATIN SYNONYMS (first word not abbreviated)

Statement of the minister of the Dominican Republic, . . ., concerning
   the incident of October last, in one section of the Haitian-
   Dominican border (scope)
STATEMENT CONC. THE INCIDENT/OCT. LAST; <u>in 1 sect./haitian-
   dominican border</u>

Historical sketch of the Congregational Church in Belchertown, Mass.,
   from its reorganization, 114 years, with. . . (scope)
HISTORICAL SKETCH/CONGREGATIONAL CHURCH/<u>BELCHER-
   TOWN, MASS.</u>; <u>from. . .</u>

Latin grammar

Madame Therese; or, The volunteers of '92 (alternative)
MADAME THERESE (THE VOLUNTEERS/'92)
THE VOLUNTEERS/'92

Jane Clegg; a play in three acts (form)
JANE CLEGG--A PLAY; /<u>3</u> acts

Nicholas II, prisoner of the purple (explanatory)
NICHOLAS 2D--PRISONER/PURPLE

"The North Pole", a tale of the Arctic sea; a 100% tone and speaking
   picture with songs, choruses and dances (explanatory/form/sup-
   plementary)
THE NORTH POLE; <u>a tale. . .</u>

The margin of music

Two Tudor books of arms
TWO TUDOR BKS./ARMS (Numeral word not transliterated because
   initial)

An elementary geography; being also a key to the new series of out-
   line maps (suppl.)
AN ELEMENTARY GEOGRAPHY; being. . .

Text-book of bacteriology
TEXT-BK./BACTERIOLOGY

Inventory of Memorials, presented, by the deputies of the Council of
   Trade in France, to the Royal Council, in 1701 (initialism, oc-
   casion, repet. of main ent.)
MEMORIALS, PRESENTED/ROYAL COUNCIL; /1701

Paris following the terror (scope)
PARIS--FOLLOWING THE TERROR

Romances of the law
ROMANCES/LAW

Map of southern Idaho and the adjacent regions (supplementary)
MAP/S. IDAHO; + the adjacent regions

Simplified spelling from the scholarly standpoint and with reference
   to the masses: the argument for simplified spelling (explanatory,
   alternative)
SIMPLIFIED SPELLING; /scholarly standpoint + /masses (THE
   ARGUMENT/SIMPLIFIED. . .
THE ARGUMENT/SIMPLIFIED SPELLING

A manual of procedure for the clinical laboratory
A MANUAL/PROCEDURE/CLINICAL LABORATORY

How I became a preacher

The cementation of iron and steel
THE CEMENTATION/IRON + STEEL

Methods of recording retail prices and measuring the cost of living
   in Italy (scope)
METHODS/RECORDING RETAIL PRICES + MEASURING THE COST/
   LIVING/ITALY

Principles of American forestry
PRINCIPLES/AMER. FORESTRY

Charter rights of the Greenville. . . railroad company (ref. to main
   entry)
CHARTER RIGHTS

Man-midwifery exposed and corrected; or, The employment. . .
   (alternative)
MAN-MIDWIFERY EXPOSED + CORRECTED (THE EMPLOYMENT...
THE EMPLOYMENT. . .

Indians of yesterday
INDIANS/YESTERDAY

The ancestress, tragedy in five acts (form)
THE ANCESTRESS; tragedy. . .

Children's toys of bygone days; a history of playthings of all peoples
   from prehistoric times to the XIXth century (explanatory/scope)
CHILDREN's TOYS/BYGONE DAYS; a hist./. . .

May dust

The story of mankind; science of anthropology--man's evolution--his
   physical, mental and moral development (explanatory, alternative)
THE STORY/MANKIND (SCIENCE/ANTHROPOLOGY); man's . . .
SCIENCE/ANTHROPOLOGY

The book of electrical wonders
THE BOOK/ELECTRICAL WONDERS

The little schoolmistress

The first (-second) book of word and sentence work; or, Easy steps
   in spelling (inserted content, alternative)
THE 1ST-2ND BK./WORD + SENTENCE WORK (EASY STEPS/SPELL-
ING)
EASY STEPS/SPELLING

Map and aerial photo reading simplified
MAP + AERIAL PHOTO READING SIMPLIFIED

How not to treat Illinois soils

Case studies in the development of social attitudes
CASE STUDIES/DEV./SOCIAL ATTITUDES

The absorption of resonance neutrons by boron, chlorine, cobalt and
   manganese
THE ABSORPTION/RESONANCE NEUTRONS/BORON, CHLORINE,
   COBALT + MANGANESE

Our debt to the red man; the French-Indians in the development of the
   United States (explanatory/alternative)
OUR DEBT/RED MAN (THE FRENCH-INDIANS/DEV./U. S.)
THE FRENCH-INDIANS/DEV./U. S.

A manual of the timbers of the world; their characteristics and uses
   (explan.)
A MANUAL/TIMBERS/WORLD; their. . .

Assessment and taxation in Kansas (scope)
ASSESSMENT + TAXATION--KANSAS

The Italian navy in the world war 1915-1918 (scope)
THE ITALIAN NAVY--WORLD WAR 1915-1918

Body form in growing chickens
BODY FORM/GROWING CHICKENS

The present status of the certification of teachers in the United
    States (scope)
THE PRESENT STATUS/CERTIFICATION/TEACHERS/U. S.

The temperaments; or, The varieties of physical constitution in man,
    considered in their. . . (alternative/explanatory)
THE TEMPERAMENTS (THE VARIETIES/PHYSICAL CONSTITU-
    TION/MAN; cons./mental. . .)
THE VARIETIES/PHYSICAL CONSTITUTION/MAN; cons./mental. . .

Report of the Central board. . . (corporate main entry)

Shakespeare's heroines; characteristics of women, moral, poetical,
    and historical (explan.)
SHAKESPEARE'S HEROINES; characteristics. . .

The memoirs of the Marquise de Keroubec (1785-1858) being ex-
    tracts. . . (explanatory/form, reference to main entry)
THE MEMOIRS; being. . .

The Peacock farm

Technology, employment, and output per man in petroleum and
    natural-gas production (scope)
TECHNOLOGY, EMPLOYMENT, + OUTPUT/MAN--PETROLEUM +
    NATURAL-GAS PRODUCTION

The way of the West
THE WAY/W.

Poems; including The saint's tragedy, Andromeda, songs, ballads, etc.
    (contents)
POEMS; including: THE SAINT'S TRAGEDY: ANDROMEDA; songs
    . . .
THE SAINT'S TRAGEDY
ANDROMEDA

A digest of the statutes of Arkansas, embracing. . . (explanatory/
    scope)
A DIGEST/STATUTES/ARK.; embracing. . .

The adventurous life of Count Lavalette, Bonaparte's aide-de-camp
    and Napoleon's postmaster-general (explanatory)
THE ADVENTUROUS LIFE/COUNT LAVALETTE--BONAPARTE'S. . .

Guy Livingstone; or, "Through" (alternative)
GUY LIVINGSTONE (THROUGH)
THROUGH

The book that gave to Iowa its name, a reprint (formal)
THE BOOK THAT GAVE/IOWA ITS NAME; a repr.

Nutrition

Driven to bay
DRIVEN/BAY

The perception of light and color
THE PERCEPTION/LIGHT + COLOR

Morphological studies of the head and mouth-parts of the mature
   codling-moth larva, Carpocapsa pomonella (Linn.) (explanatory)
MORPHOLOGICAL STUDIES/HEAD + MOUTH-PTS./MATURE
   CODLING-MOTH LARVA; Carpocapsa. . .

Madame L'Orient's art of beauty (initialism)
ART/BEAUTY

The American vine dresser's guide
THE AMER. VINE DRESSER'S GUIDE

A manual of modern scholastic philosophy
A MANUAL/MODERN SCHOLASTIC PHILOSOPHY

An oration, pronounced July 4, 1799, at the request of the inhabitants
   of the town of Boston, in commemoration of the anniversary of
   American independence (occasion)
AN ORATION; pronounced. . . Boston--IN COMMEMORATION. . .
         --IN COMMEMORATION/ANNIVERSARY/AMER.
   INDEPENDENCE

The book of prescriptions, with notes on the pharmacology. . . (sup-
   plementary)
THE BOOK/PRESCRIPTIONS; /notes. . . .

Adventures in interviewing
ADVENTURE/INTERVIEWING

Catalogue of Marietta library (corporate main entry referred to)
CATALOGUE

The horn-fly (Haematobia serrata Rob.-Desv.) (explanatory)
THE HORN-FLY; Haematobia. . .

The law of insolvency in South Africa (scope)
THE LAW/INSOLVENCY--S. AFRICA

The vegetation of twin island
THE VEGETATION/TWIN ISLAND

Martin's natural history (initialism)
NATURAL HIST.

Upper night

Churches in lower Nubia (scope)
CHURCHES--LOWER NUBIA

The new State department

What is wrong?

Oil recovery investigations of the Petroleum experiment station of
the U. S. Bureau of mines
OIL RECOVERY INVESTIGATIONS/PETROLEUM EXPERIMENT
STATION/U. S. BUREAU/MINES

The olive branch; or, The evil and the remedy (alternative)
THE OLIVE BRANCH (THE EVIL + THE REMEDY)
THE EVIL + THE REMEDY

The doctrine of fascism
THE DOCTRINE/FASCISM

The law of suffrage and elections
THE LAW/SUFFRAGE + ELECTIONS

The archaeological collections (corporate main entry)
THE ARCHAEOLOGICAL COLLECS.

Rules and practice of the Supreme Court of South Africa, Transvaal
provincial division, Witwatersrand local division, and Appellate
division (contents)
RULES + PRACTICE/SUPREME COURT--S. AFRICA: TRANSVAAL
PROVINCIAL DIVISION: WITWATERSRAND LOCAL DIVISION:
APPELLATE DIVISION
TRANSVAAL PROVINCIAL DIVISION (a more desirable result would
WITWATERSRAND LOCAL DIVISION be to get each of these re-
APPELLATE DIVISION tracings as a continuation of
the opening phrase)

Suggested forms for internal financial reports of colleges and uni-
versities (scope)
SUGGESTED FORMS/INTERNAL FINANCIAL REPORTS/COLLS. +
UNIVS.

The war and South American trade, issued. . . by. . . (occasion, refer-
ence to m. e.)
THE WAR + S. AMER. TRADE; iss. . . .

Report of the secretary of the. . . on the census. . . 1861 (ref. to corp.
main ent.)
REPORT/CENSUS

Relation of. . ., tr. from the Spanish by Buckingham Smith (reference to main entry, redaction)
RELATION--TR./SPANISH/SMITH

Oaths and obligations of free masonry, as published to the world, by a convention of seceding masons, held. . . (title main entry, explanatory/occasion)
OATHS + OBLIGATIONS/FREE MASONRY; as publ./world. . .

A Gael's guide to Waterford and the Deise country
A GAEL's GUIDE/WATERFORD + THE DEISE COUNTRY

The church of Rome: a view of the peculiar doctrine, religious worship, ecclesiastical polity, and ceremonial observances, of the Roman Catholic Church (expl.)
THE CHURCH/ROME; a view/(THE PECULIAR DOCTRINE, RELIGIOUS WORSHIP, ECCLESIAST. . .
THE PECULIAR DOCTRINE, RELIGIOUS WORSHIP, ECCLESIASTICAL. . .

Poems. Dedicated to Thomas Moore, esq. (form)
POEMS--DEDICATED/THOMAS MOORE, ESQ.

The man with the lantern
THE MAN/LANTERN

Love in its extasie; or, The large prerogative (alternative)
LOVE/EXTASIE (THE LARGE PREROGATIVE)
THE LARGE PREROGATIVE

Export opportunities in South and Central America (scope)
EXPORT OPPORTUNITIES/S. + CENTRAL AMER.

An answer to the remarks of the Plymouth company, or (as they call themselves) the Proprietors of the Kennebeck purchase. . . on the plan and extracts of deeds
AN ANSWER/REMARKS/PLYMOUTH COMPANY; or. . . (LC has a tracing "Plymouth Company, 1749-1816. Remarks on the plan and extracts of deeds" which probably cannot be automatically extracted)

A handbook to the palace of Minos at Knossos, with its dependencies (scope, suppl.)
A HANDBOOK/PALACE/MINOS/KNOSSOS; /its. . .

Nutrition and the school lunch
NUTRITION + THE SCHOOL LUNCH

Some decisions on the Property acts
SOME DECISIONS/PROPERTY ACTS

Taxation in New York, 1924; the complete New York tax law. . .
(scope, explanatory)
TAXATION--N. Y., 1924; the. . .

Essays on the political history of the fifteenth, sixteenth, and seven-
teenth centuries
ESSAYS/POL. HIST./15TH, 16TH, + 17TH CENTURIES

Church music in the light of the motu proprio; a guide for the Catholic
choir-master and organist (explanatory)
CHURCH MUSIC/LIGHT/MOTU PROPRIO; a guide. . .

D. Francisco Manuel de Mello

The dollar crop, and its effect upon labor, agriculture, commerce
and manufacture
THE DOLLAR CROP, + ITS EFFECT/LABOR, AGRICULTURE, . . .

Reynold's code. Polyglot nautical telegraph for the use. . . (initial-
ism, explan.)
CODE (POLYGLOT NAUTICAL TELEGRAPH; /use. . .
POLYGLOT NAUTICAL TELEGRAPH; /use. . .

Acts relating to the public schools of Rhode Island, with remarks +
forms. (scope, supplementary)
ACTS/PUBLIC SCHOOLS/R. I.; with. . .

Rustic architecture

Fair exchange; a novel in the first person (form)
FAIR EXCHANGE; a novel. . .

Memoirs of Duke de Richelieu (reference to main entry)
MEMOIRS

Racing bits; a story of the oil-fields of Texas (form/explanatory)
RACING BITS--A STORY/OIL-FIELDS/TEX.

Floats for the so-called fattening of oysters
FLOATS/SO-CALLED FATTENING/OYSTERS

An address to the flocks of the reverend approvers of Blanco White's
internal evidences against Catholicism
AN ADDRESS/FLOCKS/REVEREND APPROVERS/BLANCO WHITE'S:
INTERNAL EVIDENCES AGAINST CATHOLICISM
INTERNAL EVIDENCES AGAINST CATHOLICISM

Obstetrical notes for juniors (occasion)
OBSTETRICAL NOTES/JUNIORS

Canary breeding for beginners; a practical up-to-date guide (oc-
casion/form)
CANARY BREEDING/BEGINNERS; A PRACTICAL UP-TO-DATE
GUIDE

St. Philip's church. Memorial poem (title main entry, form)
ST. PHILIP's CHURCH. MEMORIAL POEM (retention of original
    punctuation leaves whole as single field)

Annual report of the town (corporate main entry)
ANNUAL REPORT/TOWN

The wayfarer; a novel (form)
THE WAYFARER; <u>A NOVEL</u>

A bibliographical manual for the student of criminology (occasion)
A BIBLIOGRAPHICAL MAN.--/STUDENT/CRIMINO<u>LOGY</u> (first word
    not abbrev.)

Syllabus of Davis' International law
SYLLABUS/DAVIS': INTERNATIONAL LAW
INTERNATIONAL LAW

On the wing through Europe

The second Sexton cook book for the Sexton market (repetition of
    main entry, but not given as an initialism; best to arbitrarily
    abridge as:)
THE 2D COOK BK./SEXTON MARKET

All's well that ends well, by. . .; ed., with notes, introduction, glos-
    sary, list of various readings, and selected criticism, by. . .
    (redaction/form)
ALL's WELL THAT ENDS WELL--ED.;<u>/NOTES, INTRO., GLOS-
    SARY, LIST/VARIOUS READINGS, + SEL. CRITICISM</u>; BY. . .
        (traced as:)
ALL'S WELL THAT ENDS WELL--ED. BY. . .

Thanksgiving plays for boys and girls, a collection of six. . . (form,
    occasion)
THANKSGIVING PLAYS/GIRLS + BOYS; <u>A COLLEC./6. . .</u>

Final report of the. . . (corporate main entry repeated)
FINAL REPORT

The young citizens league; helps and suggestions for organizing
    (form)
THE YOUNG CITIZENS LEAGUE; <u>HELPS. . .</u>

Not a sparrow falls, a novel (form)
NOT A SPARROW FALLS; <u>A NOVEL</u>

Manifesto of the Spanish nation to Europe
MANIFESTO/SPANISH NATION/EUROPE

O C D, outline of lectures on civilian defense (alternative to abbreviation)
O C D, OUTLINE/LECTURES/CIVILIAN DEFENSE (Note need for spaces in abbrev. entries)

Sketching methods

The Swiss cross, a monthly magazine of popular science (title main entry, form/explanatory)
THE SWISS CROSS; A MONTHLY. . .

Shelley

The Germania of Tacitus, with ethnological dissertations and notes. By R. G. Latham (supplementary/redaction)
THE GERMANIA;/ETHNOLOGICAL DISSERTATIONS + NOTES--BY R. G. LATHAM
   (traced as:)
THE GERMANIA--BY R. G. LATHAM

Carolina humor; sketches by. . . (form)
CAROLINA HUMOR; SKETCHES. . .

Brief memorials of an only daughter
BRIEF MEMORIALS/ONLY DAUGHTER

The distant princess

Tumminello theory of numerical roots (initialism)
THEORY/NUMERICAL ROOTS

The minority of Henry III.
THE MINORITY/HENRY 3D

The twin sisters; or, The effects of education: a novel; in a series of letters (title main entry, alternative, form)
THE TWIN SISTERS (THE EFFECTS/EDUC); A NOVEL. . .
THE EFFECTS/EDUC.; A NOVEL. . .

A history of philosophy, from Thales to the present time (scope)
A HISTORY/PHIL.;/THALES. . .

Under the crown, American history; selections from the Youth's companion (title main entry, explanatory, form)
UNDER THE CROWN--AMER. HIST.; SEL./. . .

The Unitarian; conducted by. . . (title main entry)
THE UNITARIAN; . . .

The application of statistical methods to the problems of psychophysics
THE APPLICATION/STATISTICAL METHODS/PROBLEMS/PSYCHOPHYSICS

The geology of the country around Newton Abbey (scope)
THE GEOLOGY/COUNTRY AROUND NEWTON ABBEY

Studies in the Asclepiadaceae
STUDIES/ASCLEPIADACEAE

An essay on the antiquity of the Irish language, with a preface, proving
    Ireland to be the Thule of the ancients (supplementary)
AN ESSAY/ANTIQ./IRISH LANGUAGE; WITH. . .

Hand book of the United States tariff, containing the tariff act of 1897
    rev. to July 1, 1902, with complete schedules of. . . (explanatory)
HAND BK./ U. S. TARIFF; cont. . .

Deterioration of steels in the synthesis of ammonia
DETERIORATION/STEELS/SYNTHESIS/AMMONIA

An integrated course of study in speech
AN INTEGRATED COURSE/STUDY/SPEECH

The story of milk and how it came about
THE STORY/MILK + HOW IT CAME ABOUT

Paintings of the sea
PAINTINGS/SEA

An introduction to English grammar, on an analytical plan, adapted
    to the use of students in colleges. . . (form/occasion)
AN INTRODUCTION/ENG. GRAMMAR,/ANALYTICAL PLAN;
    ADAPTED. . .

Laird & Lee's vest-pocket Webster pronouncing dictionary. . .
    (initialism for publ.)
LAIRD + LEE'S VEST-POCKET WEBSTER PRONOUNCING DICT.

Anticlinal structure in parts of Cotton and Jefferson counties, Okla-
    homa (scope)
ANTICLINAL STRUCTURE/PTS./COTTON + JEFFERSON COUN-
    TIES, OKLA.

The analytical chemist's assistant: a manual of chemical analysis,
    both qualitative and quantitative of natural. . . (explanatory)
THE ANALYTICAL CHEMIST'S ASSISTANT: A MANUAL. . .
A MANUAL/CHEMICAL ANALYSIS, BOTH QUALITATIVE + QUANTI-
    TATIVE. . .

The Greek romances in Elizabethan prose fiction
THE GREEK ROMANCES/ELIZABETHAN PROSE FICTION

Memorials of James Wood, LL.D., J.P. of Grove House, Southport
MEMORIALS/JAMES WOOD; LL.D. . .

The silence, thoughts on the silence, and on various other subjects
(explanatory)
THE SILENCE; <u>THOUGHTS. . .</u>

Punches, dies and tools for manufacturing in presses (scope)
PUNCHES, DIES + TOOLS--/MANUFACTURING IN <u>PRESSES</u>

Social life at the English universities in the eighteenth century (scope)
SOCIAL LIFE/ENG. UNIV./18TH CENTURY

 The filing of such a sample of entries has not yet been under-
taken, though its satisfactory use is obviously necessary before such
a compression can be judged acceptable. However, the principles
and desiderata enunciated in the preceding text, together with the
Hines/Harris code, should produce the basis on which to attain such
an acceptability. The two major problems, of course, are the fil-
ing of abbreviations with other similar-stemmed words (a problem
encountered in the filing of L.C. subject headings, with their com-
pulsory abbreviations, already), and of the filing of the virgule
(=prepositional-phrase-introducers)—this last particularly in that
the same symbol equivalated (in the recommendations) to a multipli-
city of such prepositions, from <u>a</u> ("at". . .) almost to <u>z</u> ("with". . .).