

Lib sc. 9; 1972; PAPER Y.

Computer Simulation of some Steps in Classifying : Built-in Features of a Theory-Based Scheme for Classification.
(Classification problems. 66). (Non-conventional methods in document retrieval. 15).

A Neelameghan, Documentation Research and Training Centre, Indian Statistical Institute, Bangalore 560003.

[The work done on a procedure for the assembly of the components of Class Number according to a freely-faceted analytico-synthetic scheme for classification with the aid of computer, given the kernel terms of a subject, is referred to. The particular built-in features of a scheme for classification designed on the basis of the General Theory of Library Classification which facilitates simulation on computer the assembly of the components of a subject—in kernel terms or as a Class Number—coextensively representing the given subject, are examined. The different variety of ideas recognised and the guiding principles for their arrangement in a helpful sequence provided by the General Theory of Library Classification are discussed with illustrative examples. An alternative procedure, using the alphabetical index to the schedules, for simulation of some of the steps in the classifying process, is briefly outlined.]

0 Introduction

01 EARLIER WORK

A procedure for the assembly of components of Class Number according to a freely-faceted analytico-synthetic scheme

for classification with the aid of computer, given the kernel terms of a subject, has been described in a series of papers, forming part of the work on the development of a computer-aided document finding system (3-5, 7, 11, 12, 19-22). In that work, the Colon Classification has been used as a representative of a near-approximation to a freely-faceted analytico-synthetic scheme for classification. The necessary computer programs have been developed and class numbers for different subjects have been constructed.

02 SCORE OF THE PAPER

This paper examines the particular features of the scheme for classification based on the freely-faceted analytico-synthetic model which facilitate simulation on computer the assembly of components of a subject, resulting in a class number or a structured subject heading coextensively representing the given subject.

1 Line of Approach

The basic line of approach in our work using computer in assembling the components of a Class Number with Colon Classification as a guinea pig representing a freely-faceted classification scheme, may be summarised as follows:

- 1 To specify as minutely as necessary, each of the successive steps involved in the process of classifying the subject of a document according to the Postulational Method (17);

- 2 To examine the work involved in each of the steps enumerated in order to identify such items of work which involve human judgment and those which do not; and

- 3 To develop the necessary programs for the computer to perform the items of work which do not involve human judgment.

2 Freely-faceted Classification

21 ANALYSIS AND REPRESENTATION OF SUBJECT

A freely-faceted analytico-synthetic classification, provides explicitly stated guiding principles for the

- 1 Analysis of subject into its respective component idea-units;

- 2 Recognising the varieties of interrelation and mutual filiation among the component idea-units obtained at 1; and

- 3 Assembling the idea-units in a preferred sequence on the basis of their mutual filiation recognised at 2.

The subject can be represented as a structured string of terms denoting the corresponding idea-units or as a class number made up of ordinal numbers representing the idea-units. The compo-

ment idea-units and their interrelation and mutual filiation recognised in the idea plane are to be preserved in the representation of the subject in either method.

3 Help from General Theory of Library Classification

31 GUIDING PRINCIPLES

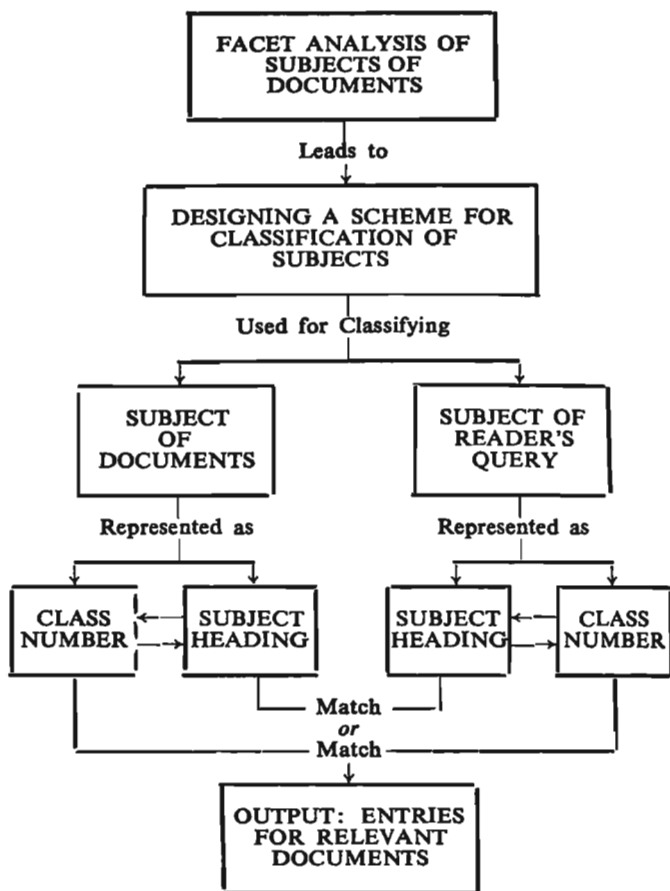
A set of guiding principles for the analysis and reassembly of idea-units of a subject is given in the General Theory of Library Classification (6, 15) and its refinements. It consists of sets of postulates, canons, and principles, and directions for their use in the above-mentioned analysis and assembly work. It results in a consistent pattern in the structuring of subjects; the sequence of the component ideas is deemed to be helpful to a majority of specialists in the respective subject-fields; and the subjects—macro or micro—falling in different asras of specialisation get arranged in a more or less helpful way. The overall result is an APUPA pattern (Alien-, Penumbral-, Umbral-, Penumbral-, Alien) of subjects. This pattern facilitates browsing and the selection of documents dealing with subjects most relevant to the reader's interest at the moment with minimum of noise, leakage, and search time (8, 10).

32 USE OF THE GENERAL THEORY OF LIBRARY CLASSIFICATION

The General Theory of Library Classification mentioned above is used in several phases in designing the components of a document finding system. For example,

- 1 Analysis of the subjects of documents, needed for the
- 2 Design and development of a scheme for classification of subjects;
- 3 Classifying the subjects of documents using the Postulational Method (17) using the scheme for classification designed at phase 2, or independently by directly using the Postulates and Principles;
- 4 Classifying the subjects of readers query using the Postulational Method aided by the scheme for classification designed at Phase 2, or independently by directly using the Postulates and Principles; and
- 5 In formulating the structured representation of subjects for preparing subject headings.

The above mentioned phases can be represented schematically as shown in the diagram in P 502



33 BASIS FOR HELPFULNESS

In each of these phases of work, the same entity — namely, the universe of subjects, either the whole of it or a subset of it — is analysed and/or its parts assembled to represent a whole. A subject is a systematised body of ideas, the ideas being largely products of human intellect. These ideas and subjects made up of them are again to be used by the same or other intellectuals. A basic postulate in designing a document finding system based on the General Theory of Library Classification, is that there is an Absolute Syntax of Intellection — that is, the sequence in which a majority of intellectuals assemble the constituent idea-units in their mind when thinking about the subject concerned (14). Therefore, a pattern of assembling the constituent idea-units parallel to the Absolute Syntax would be helpful to a majority of the intellectuals (readers). The General Theory of Library Classification provides principles for Facet Sequence the application of which secures a sequence among the idea-units of subjects parallel to the Absolute Syntax (9).

34 CONFORMITY WITH SYSTEMS APPROACH

In Paper W in this issue, it has been shown that a subject can be considered as a system and that it is helpful to do so. A system can be viewed as a structure or organisation of an orderly whole, clearly showing interrelations of the elements to each other and to the whole itself (1, 2). Analysis is the process of examining a system and then breaking it down into its elements to disclose interrelationships to each other and to the whole. Such an analysis can help to produce a model or an analogy of an existing system. The General Theory of Library Classification provides for a process of analysis of subjects of the universe of subjects into idea-units. It has also given a linear model for representing a subject as a chain of idea-units of different varieties with different interrelations and mutual filiations between them (18).

4 Practicability of Use of Computer**41 APPLICATION OF PRINCIPLES INVOLVES JUDGEMENT**

The Postulational Method of Classifying is based on the use of the postulates, canons and principles of the General Theory of Library Classification. This naturally raises doubts such as to whether the computer can be easily programmed to

1 Select and apply the appropriate postulate or canon or principle at the different steps in the classifying process, and to structure the subject and arrange its component idea-units in a helpful sequence derivable on the basis of the appropriate Principle for Helpful Sequence; and

2 Select and prefix the appropriate indicator digit to the component digit or digit-group in the class number.

42 DIFFERENCE IN OBJECTIVES

However, it would be helpful to distinguish between the objectives of classifying subjects by a human classificationist/classifier and those of classifying aided by a computer. In the former case, in addition to constructing a class number for a given subject, the classificationist/classifier has the objective of knowing and learning about the structure, development and other attributes of diverse subjects, selection and application of the relevant postulates and principles, the structure and layout of the schedules in the scheme for classification, etc. This knowledge would help him in the design and development of the scheme for classification for subjects which cannot be classified with the existing schedules. On the other hand, in taking the aid of computer for classifying, our only objective was, given the kernel terms representing the component idea-units of the subject to be classified, to use the machine for assembling the component idea-units of the subject in a helpful sequence to coextensively represent the subject. The subject may be represented either as a structured subject heading or by a class number made up of ordinal numbers. We were not interested in the computer "learning" about the structure of subjects, the postulates and principles of classification, etc.

5 Classifying

51 MAIN STEPS

The steps in classifying the subject of a document essentially consist of the following:

1 Determining the kernel ideas forming components of the subject;

2 Determining the interrelation and mutual filiation among the component ideas;

3 Arranging (assembling) the component ideas in the sequence of their mutual filiation as determined by applying the appropriate Principle for Facet Sequence. At this stage, the subject may be represented as a

31 String of terms representing the corresponding idea-units; and/or

32 String of digits based on a preferred notational system of ordinal numbers — that is, the class number for the subject.

52 TWO METHODS: SAME RESULT

At present we have not computerized Step 1. That is, the kernel ideas are selected by the classifier by examining and

reading through the title, preface, introduction, abstract, summary, and text, if necessary, of the document.

Steps 2 to 31 can be done either using the General Theory of Library Classification — the postulates, and principles — directly, or by using a scheme for classification designed on the basis of the General Theory of Library Classification. Generally, in the Postulational Method of Classifying, all the steps, except Step 32 are expected to be done without using any scheme for classification. For, the work in these steps are not dependent on any scheme for classification; they are concerned with the analysis of the subject into idea-units and assembling the latter in a helpful sequence for which the Principles for Helpful Sequence provided by the General Theory of Library Classification are used.

From the diagram in Sec 32, it will be noted that classifying a document occurs as a process later to designing a scheme for classification. We have also mentioned in the same section, that the same set of postulates, canons and principles provided by the General Theory of Library Classification are used in designing a scheme for library classification of subjects as well as in classifying the subject of documents. Therefore, the result of classifying either by directly using the General Theory of Library Classification or by using the scheme for classification designed on the basis of the theory should be the same. That is, the component idea-units assembled in a sequence to represent a subject coextensively by either of the methods should be the same and, therefore, equally helpful, whether the component idea-units are represented in words of a natural language or by a notation. Therefore, in securing an assembly of the component idea-units of a subject in a preferred structural pattern using the scheme for classification, it would amount to simulating steps 2 to 32 of classifying by directly using the General Theory of Library Classification. This fact has been made use of in simulating the classifying process (Steps 2 to 32 in Sec 51).

6 Two Approaches for Simulation

In using a computer for assembling the component idea-units of a subject in a preferred sequence, two approaches are possible:

- 1 Use of the schedules of the scheme for classification; or
- 2 Use of an alphabetical index to the schedules specially constructed for the purpose.

61 USE OF SCHEDULES

611 *Matching*

In order to assemble the given kernel terms of a subject in a preferred sequence using the schedules of a scheme for classification for the subject,

1 The computer should be enabled to match each of the given kernel terms with the corresponding idea-term (isolate term or speciator term) enumerated in the schedule; and

2 This matching of the kernel terms should preferable be in the sequence in which they will be assembled one after the other in representing the subject — whether it be in words or in ordinal numbers.

612 *Implication*

This implies that

1 All the ideas of a subject should be enumerated in the schedules; and

2 The ideas should be enumerated in a sequence such that when picked up and assembled successively, they would be in the preferred sequence prescribed by the appropriate Principle for Helpful Sequence.

The first condition requires keeping the schedules continuously updated. We are not concerned with it in this paper. The second condition requires that the preferred sequence be built into the schedules — that is, every idea should be placed in a helpful sequence, the sequence being determined by the Principles for Helpful Sequence provided by the General Theory of Library Classification. In the succeeding sections it will be shown that in designing the schedules of a scheme for classification according to the General Theory of Library Classification, a helpful sequence among the different kinds of ideas is secured such that

1 The component ideas of a subject get arranged in a helpful sequence;

2 All the subjects falling in particular field — that is going with a particular Basic Subject get arranged among themselves in a helpful sequence; and

3 All the subjects going with different Basic Subjects get arranged among themselves in a helpful sequence.

7 Principles Used in Arrangement

70 KINDS OF COMPONENT IDEAS

The different kinds of ideas recognised by the General Theory of Library Classification that form component idea-units of a subject are as follows:

- 1 Main (Basic) subject;
- 2 Non-Main Basic Subject going with a Main (Basic) Subject;
- 3 Partial Comprehension;
- 4 Isolate (primary);
- 5 Speciator; and
- 6 Characteristic mentioned in the schedule (Quasi-isolate).

In the succeeding sections we shall consider only the Principles of Helpful Sequence generally used in Colon Classification (= CC) for arranging the different varieties of ideas in its schedules. These principles can, of course, be used in arranging ideas in the schedules of any other scheme for classification also.

71 MAIN SUBJECT

711 *Five Groups and their Sequence*

The Main (Basic) Subjects or Main Subjects enumerated in the schedule of Main Subjects have been arranged in the following sequences:

In the first instance the Main Subjects are arranged into five broad groups:

(a) Group of Main Subjects with which subjects of the following varieties may be respectively deemed to go.

Subjects dealing with the theory of a technique or of a practice-in-action;

For example:

Exhibition technique, Museology, and Conference technique;
Systems theory, Management science, Symbolism

It will be noted that the technique or practice-in-action can also occur as an isolate idea in subjects going with different Basic Subjects.

Subjects which can be associated with several subject-fields;

For example:

Library science, Book science, and Journalism.

(b) Group of Main Subjects with which subjects conventionally called the Natural Sciences, may be deemed to go. For example,

Physics	Biology	Botany
Chemistry	Geology	Zoology

(c) Useful Arts

(d) Group of Main Subjects with which subjects conventionally called the Humanities, may be deemed to go. For example:

Fine arts	Religion
Literature	Philosophy

(e) Group of Main Subjects with which subjects conventionally called the Social Sciences, may be deemed to go. For example:

Education	Political science	Sociology
History	Economics	Law

The five groups of Main Subjects are arranged in the sequence mentioned above. This sequence is a postulated one. However, the grouping into natural sciences, humanities, and

social sciences, is a generally accepted way of dividing the universe of subjects.

712 *Pure Discipline-Applied Discipline Sequence*

A Main Subject with which the subjects in an applied discipline are deemed to go is placed after the Main Subject with which the subjects in the pure discipline on which the applied discipline is dominantly based, are deemed to go. For example:

D Engineering follows C Physics
 F Chemical technology follows E Chemistry
 J Agriculture follows I Botany
 K Animal husbandry follows K Zoology
 L Medicine follows KX Animal husbandry
 XV Applied economics follows X Economics
 YX Social work follows Y Sociology

713 *Principle of Increasing Concreteness*

The Main Subjects covering the pure disciplines in the Natural Sciences are arranged among themselves according to the Principle of Increasing Concreteness. Thus:

B Mathematics	H Geology
C Physics	I Botany
E Chemistry	K Zoology
G Biology	Δ Spiritual experience and Mysticism

714 *Principle of Increasing Artificiality*

The Main Subjects covering the pure disciplines in the Humanities and Social Sciences taken together are arranged among themselves according to the Principle of Increasing Artificiality or the Principle of Decreasing Naturalness. Thus

Δ Spiritual Experience and Mysticism	T Education
N Fine Arts	U Geography
O Literature	V History
P Linguistics	W Political science
Q Religion	X Economics
R Philosophy	Y Sociology
S Psychology	Z Law

It will be noted that, looked at from one angle, Spiritual Experience and Mysticism is the most concrete and looked at from another angle it is the most natural.

715 *Multiple Sequence*

The group of Main Subjects mentioned in Sec 711 (a) are arranged among themselves according to different principles which-

ever gives a helpful sequence among a subset of such Main Subjects. For example:

(a) The Principle of Decreasing Generality is used in arranging the Main Subjects

- 1 Universe of Subjects: Its structure and development.
- 2 Library science.
- 3 Book science
- 4 Journalism

(b) The Principle of Scheduled Mnemonics is used in arranging the Main Subjects

- | | |
|--------------------------------|---------------------------|
| 9b Career | 9p Conference methodology |
| 9c Metrology | 9t Commission methodology |
| 9d Standardisation methodology | 9P Communication theory |
| 9f Research methodology | 9Q Symbolism |
| 9g Evaluation methodology | |

72 NON-MAIN BASIC SUBJECTS

720 *Four Varieties and their Sequence*

Four varieties of Non-Main Basic Subjects going with a Main Subject have been recognised by the General Theory of Library Classification. They are:

- 1 Canonical Basic Subject;
- 2 Specials Basic Subject;
- 3 Environmented Basic Subject; and
- 4 Systems Basic Subject.

The four groups are arranged in the sequence mentioned above. It corresponds more or less to the Principle of Later in Time.

C Physics (Main Subject)

(a) *Canonical Basic Subjects* :

- | | |
|-------------------------|----------------|
| C1 Fundamentals | C5 Radiation |
| C2 Properties of matter | C6 Electricity |
| C3 Sound | C7 Magnetism |
| C4 Heat | |

(b) *Specials Basic Subjects* :

- | | |
|------------------------|----------------------------------|
| C-9B Molecular Physics | C-9E Elementary Particle Physics |
| C-9C Atomic Physics | |
| C-9D Nuclear physics | C-9F Cosmic ray Physics |

(c) *Environmented Basic Subjects* :

- | |
|---------------------|
| C-9H By Pressure |
| C-9J By Temperature |
| C-9L By Potential |

(d) Systems Basic Subjects :

C-K	Gravitation Theory	C-N	Quantum Theory
C-M65	Electro-magnetic Theory	C-N1	General Theory of Relativity

721 Canonical Basic Subjects

The Canonical Basic Subjects going with a Main Subject are arranged among themselves according to an appropriate Principle for Sequence of Array Isolates, or the Principle of Mnemonics — Scheduled, Systematic, and Seminal — whichever is helpful. For example:

The sequence

B1	Arithmetic	B5	Trigonometry
B2	Algebra	B6	Geometry
B3	Analysis		

conforms approximately to the Principle of Increasing Complexity.

The sequence

31	Author's work (1 = Starting Point, Creation)
32	Translator's work (2 = Two phases)
35	Illustrator's work (5 = Aesthetics)
36	Compiler's work (6 = Posterity, Time)
38	Editor's work (8 = Management, Organisation)

conforms to the Principle of Seminal Mnemonics.

722 Specials Basic Subjects

The Specials Basic Subjects going with a Main Subject are arranged among themselves according to an appropriate Principle for Sequence of Array Isolates, or the Principle of Mnemonics — Scheduled, Systematic, and Seminal — whichever is helpful.

For example,

The sequence

C-9B	Molecular Physics
C-9C	Atomic Physics
C-9D	Nuclear Physics
C-9E	Elementary Particle Physics

conforms to the Principle of Decreasing Quantity, and also the Principle of Periphery to Centre.

The sequence

L-9B	Medicine of Embryo
L-9C	Child medicine
L-9D	Adolescent medicine
L-9F	Old age medicine

conforms to the Principle of Increasing Quantity (Age), and also the Principle of Later in Time.

The sequence

- X-9B Small scale economics
- X-9C Medium scale economics
- X-9D Large scale economics

conforms to the Principle of Increasing Quantity (Size).

723 Environmented Basic Subject

The Environmented Basic Subjects going with a Main Subject are arranged among themselves according to the Principle of Scheduled Mnemonics giving a sequence parallel to the ideas in the schedule of Common Environment Isolates in CC, Ed 7. The isolates in the latter are themselves arranged according to the Principle of Spatial Contiguity and the Principle of Scheduled Mnemonics (parallel to the Schedule of Basic Subjects). Thus,

- | | | | |
|--------|---------------------------|--------|---------------------|
| L-9Un4 | High altitude
medicine | L-9UD7 | Aviation medicine |
| L-9Ux | Space medicine | L-9V | War medicine |
| L-9U3 | Tropical medicine | L-9XX | Industrial medicine |

wherein the component ideas

- | | | | |
|-----|---------------|-----|----------|
| Un4 | High altitude | UD7 | Aviation |
| Ux | Space | V | War |
| U3 | Tropical | XX | Industry |

are taken from the schedule of Common Environment Isolates.

724 Systems Basic Subjects

The Systems Basic Subjects going with a Main Subject are arranged among themselves in the chronological sequence, that is, according to the Principle of Later in Time (of origin). Thus,

- S-M Experimental Psychology (1800s)
- S-M9 Psycho-analytical Psychology (1890s)
- S-N Gestalt Psychology (1900s)
- S-N1 Behaviorist Psychology (1910s)
- S-N14 Individualistic Psychology (1914)
- S-N17 Reflexology Psychology (1917)
- S-N2 Eidetic Psychology (1920)
- S-N3 Field Psychology (1930s)
- S-N3b 'We' Psychology (1936)

73 PARTIAL COMPREHENSION

A Partial Comprehension is interpolated immediately earlier to the first of the set of ideas — Main Subject Idea, Basic Subject Idea, or Isolate Idea — enumerated in the schedules,

- which it comprehends or covers. Thus,
- A *Natural sciences* (Partial Comprehension)
 - B Mathematics
 - C Physics
 - D Engineering
 - E Chemistry
 -
 - L Medicine
 - T*Z *Social sciences* (Partial Comprehension)
 - T Education V History
 - U Geography W Political science
 - V*Z *History and Political Science* (Partial Comprehension)
 - X Economics
 - Y Sociology
 - Z Law

74 ISOLATE

740 *Varieties*

The following varieties of isolate ideas have been recognised by the General Theory of Library Classification:

- Personality Isolate
- Matter Material Isolate
- Matter Property Isolate
- Matter Method Isolate
- Energy Isolate (Action Isolate)
- Space Isolate
- Time Isolate

741 *Sequence*

The isolate derived on the basis of a characteristic in the schedules of the different varieties of isolates are arranged among themselves using an appropriate principle from the set of Principles for Array Isolate Sequence.

75 SPECIATOR

The speciators derived on the basis of a characteristic are arranged among themselves using an appropriate principle from the set of Principles for Array Isolate Sequence. Thus, the speciators derived on the basis of the characteristic "By Colour" may be

Violet	Yellow
Indigo	Orange
Blue	Red
Green	

When they are arranged in the sequence given above, it conforms to the Principle of Decreasing Quantity (Wave-length).

The attachment of a speciator to an isolate gives rise to a Compound Isolate. Thus, if "Coat" is the isolate, we get a variety of coloured coats (Compound Isolates) by attaching each of the speciators mentioned above — Blue coat, Green coat Yellow coat, Red coat. These compound isolates get arranged in the sequence in which the speciators are arranged; in the present case conforming to the Principle of Decreasing Quantity.

76 CHARACTERISTIC

Isolates can be grouped on the basis of a characteristic—that is, a schedule of isolates can be formed on the basis of a characteristic. Similarly, a schedule of speciators can be formed on the basis of a characteristic. When a large number of isolates are to be arranged in a helpful sequence, it may be helpful to group them first on the basis of different characteristics. Similarly, in the case of speciators. Then, to get a helpful sequence among the isolates or speciators, as the case may be, the groups of isolates (or speciators) may be first arranged in a helpful sequence—that is, first the characteristics used for grouping, should be arranged in a helpful sequence. The following principles are usually applied for this purpose:

- 1 Wall-Picture Principle and its associated Principles for Facet Sequence (13)
- 2 Principle of Decreasing Concreteness in facet structure, by correlating the idea denoted by each of the characteristics with one and only one of the five Fundamental Categories (16).
- 3 Principle of Scheduled Mnemonics.
- 4 Principles for Array Isolate Sequence.

For example, consider the following characteristics—that is, Quasi Isolates (QI)—used to derive speciators for the Personality Isolate "Fountain Pen" occurring in subjects going with the Basic Subject "M0J Stationery material production".

(QI) of Order 1

By Attributes associated with Fountain Pen as a Whole

By Attributes associated with Organs of Fountain Pen

The sequence between the above (QI) is determined using the Whole-Organ Principle, from among the Principles for Facet Sequence.

(QI) of Order 2 for the (QI) of Order 1

By Attributes associated with Fountain Pen as a Whole

By Standard

By Purpose

By Brand

By Overall size

By Country of make

By Overall shape

Each of the first four (QI) correlate with Personality and each of the last two (QI) correlate with Property. Therefore, the two groups can be arranged in the sequence of Decreasing Concreteness.

ness in Facet Structure — that is, those correlating with Personality followed by those correlating with Property.

Among the first four (Q1) correlating with Personality, the sequence is determined using the Principle of Decreasing Concreteness.

Among the two (Q1) correlating with Property the sequence is determined using the Principle of Scheduled Mnemonics, that is, to parallel the enumeration of isolates in the schedule of Common Property Isolates. The isolates in the schedule of Common Property Isolates are themselves arranged in the sequence of the Basic Subjects with which the idea normally goes.

(Q1) of Order 2 for the (Q1) of Order 1

By Attributes associated with Organs of Fountain Pen
 By Attributes of nib By Attributes of barrel
 By Attributes of tongue By Attributes of cap
 By Attributes of neck

The sequence among these (Q1) is determined using the Principle of Spatial Contiguity (Horizontal line, with pen open, cap on, and with the nib as starting point).

77 FACET

A compound subject consists of a Basic Subject (Facet) and one or more isolate facets as components. Seven varieties of isolates have been mentioned in Sec 740. The sequence of the component facets in a compound subject is determined by the Generalised Facet Structure for a subject prescribed by the General Theory of Library Classification (18):

(BS)=Basic Subject. L=Level. [M]=Matter Facet. [P] Personality Facet. R=Round. [S] = Space Facet. [T] =Time Facet. (BS),

[1P1], [1P2], ..., [1PLp];
 [1M1]; [1M2]; ...; [1MLm];
 [1E],
 [2P1], [2P2], ..., [2P1q];
 [2M1]; [2M2]; ...; [2MLn];
 [2E],

 [RP1], [RP2], ..., [RPLr];
 [RM1]; [RM2]; ...; [RMLr];
 [RE].
 [S1]. [S2]. [SLs]
 [T1][T2] ... [TLt]

where Lp, Lm, Lq, Ln, Lr, Ls, Lt may have any integral value.

The above generalised facet structure is for Rounds started by Energy Facet.

An Isolate in (P), (M), (S) and (T) can be a Simple Isolate or a Compound Isolate.

Implied and used in this model are the postulates about Basic Facet, the Postulate of Fundamental Categories, Postulate of Round, Postulate of Level, and the Principles for Helpful Sequence in Facet Structure (the Postulate of Absolute Syntax of Intellection, the Principle of Decreasing Concreteness in Facet Structure etc).

78 RESULTING ADVANTAGE

Thus, the ideas in each variety of ideas that form components of subject get arranged among themselves in a helpful sequence in the respective schedules according to an appropriate guiding principle. The notation implements the findings of the idea plane by fixing the relative position of the ideas in the respective schedules — that is, the schedule of Basic Subjects, schedule of Special Isolates for subjects going with particular Basic Subject, schedule of Speciators for attachment to isolates in a schedule. This built-in feature of a scheme based on the General Theory of Library Classification ensures that:

- 1 The component idea-units of a subject going with a particular Basic Subject get arranged in a helpful sequence and consistent pattern;
- 2 The compound subjects going with a particular Basic Subject get arranged among themselves in a helpful sequence and consistent pattern; and
- 3 The subjects going with different Basic Subjects get arranged more or less in a helpful sequence and consistent pattern.

Therefore, given the kernel ideas of a subject, with the schedules of a scheme designed according to the General Theory of Library Classification, a helpful sequence of the component ideas in the representation of the subject can be secured "automatically" — that is, without exercise of judgment. It merely involves the presentation of different schedules of ideas in the sequence in which the different component ideas of the subject is to be assembled in the representation of the subject as a whole. In our earlier method the procedure consisted in matching the kernel terms of the subject with the terms in the schedules of the scheme, picking out the digit or digit group (together with the appropriate indicator digit already prefixed to it in the schedules) given against the term in the schedule matching with a kernel term, and assembling these to form the Class Number for the subject. Simultaneously, the kernel terms also get arranged in the sequence of component ideas as represented in the Class Number. This gives rise to a structured subject heading for use as Feature Heading and for preparing various types of alphabetical

subject indexes, using computer. The steps of scanning the schedules in a preferred sequence, matching the kernel term with the terms in the schedule, picking up the digit or digit group and assembling the component numbers and/or the kernel terms in a sequence can be programmed to be done by computer.

7Z METHODS USED

8 Method 1

81 SCHEDULE-ON-TAPE

The schedules of Colon Classification used in the construction of Class Number were put on magnetic tape. It consisted mainly of three parts:

- 1 Schedule of Basic Subjects;
- 2 Separate schedules of Special Isolates and associated speciators for subjects going with each of the different Basic Subjects; and
- 3 Schedule of Common Isolates — such as, schedule of Environment Isolate, schedule of Space Isolates, and schedule of Time Isolates.

The appropriate indicator digit is prefixed to each of the Isolate Ideas and Speciators enumerated in the schedules.

82 SCANNING THE SCHEDULES

The subject of the document to be classified is facet analysed. That is, the fundamental constituent ideas and Basic Subject are determined and each idea expressed in a standard term used in the document finding system. The standard term may be picked up using a thesaurus. This last step can be computerized.

Given these kernel terms for each of the kernel ideas in the subject, the computer is programmed to scan the schedules on the Schedule-on-Tape in the following sequences.

1 Schedule of Basic Subjects.

Each kernel term is matched with each of the terms in the schedule of Basic Subjects. When a match is established, the idea denoted by the term concerned would be the Basic Subject. The number against this term in the schedule is picked up.

2 The schedules of Special Isolates for subjects going with the Basic Subject established in the preceding step.

21 Each of the remaining kernel terms is matched first with the isolate terms in the Schedule of (IPI) isolates. When a match between a kernel term and an isolate term is established, the corresponding number given against the isolate term is picked up along with the indicator digit (that is, a comma) prefixed to it. This digit group is suffixed to the number for the Basic Subject picked up in the previous step.

22 Each of the remaining kernel terms is matched against each of the speciator terms in the schedule of Speciators asso-

ciated with the primary isolate identified in the preceding step. When a match between a kernel term and a Speciator term is established, the number given against the Speciator term in the schedule is picked up together with the indicator digit (that is, a hyphen) prefixed to it. This digit group is suffixed to the class number consisting of the Basic Subject Number and Primary Isolate Number assembled in the preceding steps.

23 The schedules for the later levels of (P) isolates — (1P2), (1P3), etc — are then successively scanned. And whenever a match between a given kernel term and a schedule term is established, the corresponding number together with the indicator digit prefixed to it is picked up and assembled in the Class Number in the sequence they are picked up.

24 The procedure is repeated with the schedules of isolates in (1MMt1), (1MMt2) etc, (1MPI), (1MP2) etc and their respective schedules of speciators; then the schedule of isolates in (1E), followed by the schedules for (2MM1), (2MM2) etc, their respective schedules of Speciators; then the schedules for (2P1), (2P2) etc, and the respective schedules of Speciators; then the schedule for (2E); and the cycle is repeated until the scanning of the entire set of schedules of special isolates going with the Basic Subject concerned is completed, or until all kernel terms have been matched.

25 Any kernel term left over is matched with the isolate terms in the schedules of Common Isolates for Space and Time. And the procedure is the same as for the Special Isolates and Speciators mentioned above.

26 Any kernel term remaining after this would be for a new idea not included in the existing schedules.

Thus, by a proper presentation of the different schedules, the components of the Class Number get assembled automatically in the facet structure sequence. Further, the appropriate indicator digit gets prefixed to the component number, even as each of them is picked up from the schedules.

In our earlier paper the method of presentation of schedules, the standardisation of kernel terms, etc have been fully discussed (11). The assembly of Class Number involving the various devices prescribed by Colon Classification—Alphabetical Device, Chronological Device, Geographical Device, Numerical Device, and Environment Device—has also been dealt with (21).

In the succeeding sections we shall discuss briefly an alternative approach to the assembly of the components of a Colon Class Number using computer.

89Z Method 2**8A USE OF ALPHABETICAL INDEX TO SCHEDULES**

In Method 1 the schedules of CC were used for assembling the Class Number with the aid of computer. In the alternative method outlined below, it is shown that an alphabetical index to the schedules of the scheme could be used for assembling the components of the Class Number in the preferred sequence. The schedules are not scanned by the computer. However, for preparing the alphabetical index to the terms in the schedules, it would normally be necessary to prepare the schedules first.

8B STRUCTURE OF ENTRY IN ALPHABETICAL INDEX

The structure of an entry in the alphabetical index to the schedules is similar to the one adopted in the alphabetical index to the terms in the schedules in CC. Here are some examples:

EXECUTIVE
 V (1P2) ,2
 W (1P2) ,2
 Z (1P2) ,8
 EXPERIMENTATION
 9f
 EXPLOSIVE
 E (1P1) -f54
 F (1P1) ,554
 (CPI) ;f54
 EXTRUDE
 F (2MM1) ;52

Under each term, the first digit or digit-group is the Basic Subject Number of the subjects with which the idea denoted by the term is deemed to go: the last digit or digit-group is the Isolate Number or Speciator Number, as the case may be, for the idea denoted by the term with the appropriate Indicator Digit prefixed to the substantive digits; and this is preceded by the symbol to indicate the manifestation of the idea denoted by the term, in the context of the subject-field concerned. Thus, in the first entry under 'EXECUTIVE'

V is the Basic Subject Number
 (1P2) denotes Personality Isolate, Round 1, Level 2
 ,2 is the Isolate Number for "Executive" with the Indicator Digit for Personality (comma) prefixed to the number.

8C SCANNING THE INDEX

Given the kernel terms for the subject to be classified, the computer first rearranges the terms in the alphabetical sequence. Each of the kernel terms is matched with each of the terms in the alphabetical index. The Isolate Number or Speciator Num-

ber given against the matching term is picked up, and the Class Number assembled with the Basic Subject Number first followed by the component numbers in the PMEST sequence. An example is given below.

8D EXAMPLE

Consider the subject: Circulation of periodicals to university post graduate departments.

The kernel terms are:

CIRCULATION
LIBRARY-SERVICE
PERIODICALS
POSTGRADUATE-DEPARTMENT
UNIVERSITY

The term "Library Service" denoting the Basic Subject with which the subject of the document is deemed to go, has been supplied as required in Step 1 of the Postulational Method of Classifying. However, this is not necessary; it can be "determined" by the computer. This will be discussed in a later paper.

The parts of the alphabetical index to the schedules where matching of the kernel terms with the index terms may occur are indicated below:

CEILING	
NA (1P3) ;2	
CELL	
G (1P1) ;11	
I (1P1) ;11	
K (1P1) ;11	
L (1P1) ;11	
CIRCULATION	PERIODICALS
2 (1MP1) ;6	2 (1P2) ;46
2V (1MP1) ;6	29D(1P2) ;46
D2(1MP1) ;93	2V (1P4) ;46
L (1MP1) ;36	PERIODICITY
	CPI ;c93
CIRCUS	POSTGRADUATE DEPARTMENT
MMR	2V (1P3) ;95
... ..	7 (1P2) ;95
LIBRARY-SCIENCE	UNIVERSITY
2	29D (1P2) ;J
LIBRARY-SERVICE	2V (1P2) ;J
2V	7 (1P1) ;4
...

The first kernel term to match with a term in the index will be CIRCULATION. All the entries given below the term are picked up and stored. The next term to match will be LIBRARY-SERVICE. The first digit or digit group under the term (that is, a Basic Subject Number) is matched with the first digit or digit group in the entries picked up under CIRCULATION. Here, the only match will be for the entry:

2V (1MP1) ;6

The other entries stored earlier are deleted and only the above entry is retained.

Scanning further, there will be a match for the term PERIODICALS. On the basis of further matching of the digit or digit groups in the entries under this term, the entry.

2V (1P4) ,46

will be picked up and stored.

In a similar way, the following entries will be picked up from under POSTGRADUATE-DEPARTMENT and UNIVERSITY respectively:

2V (1P3) ,95

2V (1P2), J

The computer is programmed to assemble digits or digit-groups of the components in the PMEST sequence after the Basic Subject Number. Thus the assembled class number will be

2V,J,95,46;6

In actual work, the labels (1P1), (1P2), (2MM1) etc are not used. Instead, numbers such as (C) for Basic Subject number, 1 for P, 2 for MMt, 3 for MP, 4 for E, 5 for MM, 6 for S and 7 for T are used.

Thus, (111) = (1P1) (132) = (1MP2)
 (112) = (1P2) (211) = (2P1)
 (121) = (1MMT1) (212) = (2P2)
 etc etc

Thus, the entries picked up will be represented as

2V (0) (131);6 2V (0) (113) ,95
 2V (0) 2V (0) (112) ,J
 2V (0)(114),46

This facilitates the assembly of the components in the preferred sequence.

8E DEvised ISOLATE

The programs developed earlier for construction of Class

Number involving devices of CC could be used in Method 2 also with necessary modifications.

91 Bibliographical References

- 1 Sec 34 ACKOFF (R L) and EMERY (F E). On purposeful systems. 1972.
- 2 Sec 34 DE GREENE (K B). Systems psychology. 1970.
- 3 Sec 01 GUPTA (B S S). Application of two-way merge technique for arranging variable length class numbers consisting of alpha-numeric and special characters. (Lib sc. 8; 1971; Paper T).
- 4 Sec 01 —. Preparation of feature heading and class index heading using computer. (Lib sc. 7; 1970; Paper R).
- 5 Sec 01 —. Program package for a system for document finding. (Lib sc. 7; 1970; Paper H).
- 6 Sec 31 NEELAMEGHAN (A). Classification, Theory of (*In Encyclopedia of library and information. Ed by Harold Lancour and Allen Kent. 1971. V 5).*
- 7 Sec 01 —. Design of the document finding system: General features. (Lib sc. 5; 1968; Paper P).
- 8 Sec 31 —. Integrated approach of India to the design and development of document retrieval systems. (*In International forum on informatics (Moscow) (1969). Volume of papers. 1969. V2; P 114-47).*
- 9 Sec 33 —. Sequence of component ideas in a subject. (Lib sc. 8; 1971; Paper)
- 10 Sec 31 — and GOPINATH (M A). User-based and user-oriented classification system: Sequence of subjects and of components in a subject. (*In Islic International Conference on Information Science (Tel Aviv) (1967). Proceedings. Ed by Lydia Vilentchuk. 1972. V2; P 253-66).*

- 11 Sec 01 NEELAMEGHAN (A) and VENKATARAMAN (S). Formu-
 Sec 82(26) lation of kernel terms for a subject and isolate
 terms for a classification schedule for use in
 the synthesis of class number by computer.
 (Lib sc. 6; 1969; Paper D).
- 12 Sec 01 — and —. Use of computer for the systhesis
 of class number: A case study with a freely
 faceted version of Colon Classification. (Lib
 sc. 5; 1968; Paper S).
- 13 Sec 76(1) RANGANATHAN (S R). Elements of library classifi-
 cation. Ed 3. 1962. Chapter N, Sec 32.
- 14 Sec 33 —. Hidden roots of classification. (Infor
 stor retr. 3; 1967; Sec 7).
- 15 Sec 31 —. Prolegomena to library classification.
 Ed 3. Assist by M A Gopinath. 1967.
- 16 Sec 76(2) —. —. Chapter RB.
- 17 Sec 1 —. —. Chapter SB.
- Sec 32 —. —. Chapter SB, Sec 5.
- 18 Sec 34 —. —. Chapter SR, Sec 5.
- Sec 77
- 19 Sec 01 VENKATARAMAN (S) and NEELAMEGHAN (A).
 Catalogue-on-tape. (Lib sc. 5; 1968; Paper
 Q).
- 20 Sec 01 — and —. Document finding. (Lib sc.
 5; 1968; Paper R).
- 21 Sec 01 — and —. Formation of isolate number by
 Sec 82(26) computer using the devices of Colon Classifi-
 cation. (Lib sc. 6; 1969; Paper K).
- 22 Sec 01 — and —. Preparation of schedule-on-tape
 for synthesis of class number by computer.
 (Lib sc. 6; 1969; Paper J).