



Neelameghan, A., Compiled by.
S.R. Ranganathan's Postulates and Normative Principles: Applications in Specialized Databases
Design, Indexing and Retrieval.
(Sarada Ranganathan Endowment for Library Science, Series 7).
Bangalore, Sarada Ranganathan Endowment for Library Science, 1997.

Copyright ©Sarada Ranganathan Endowment for Library Science (SRELS)

Digitization: Susan Ditch, SIRLS, University of Arizona

Quality Control: Megan Plesea, SIRLS, University of Arizona

Digitization Training: Han Yan, Information Systems, University of Arizona & dLIST Editor

Project Coordinator: Cheryl K. Malone, Assoc. Prof. SIRLS, University of Arizona & dLIST Editor

Digitized: Spring 2007

Acknowledgments: SRELS (A. Neelameghan, K.N. Prasad, K.S. Raghavan, DRTC) and dLIST
Advisory Board Member, S. Arunachalam (MS Swaminathan Research Foundation)

dLIST Classics (Book) Editor: Barbara Hutchinson, University of Arizona

dLIST Editor-in-chief: Anita Coleman, University of Arizona

dLIST Classics

Titles by S.R.Ranganathan

Five Laws of Library Science, Ed. 1 (1931)

Philosophy of Library Classification (1973)

Prolegomena to Library Classification, Ed. 3 (1967)

Classification and Communication (1951)

Documentation Genesis and Development (1973)

Documentation and its Facets (1963)

Library Book Selection, Ed. 2 (1966)

New education and school library: Experience of half a century (1973)

Reference Service, Ed. 2 (1961)

Other titles

S.R. Ranganathan's Postulates and Normative Principles: Applications in Specialized Databases
Design, Indexing and Retrieval, 1997. Compiled by A. Neelameghan.

Memorabilia Ranganathan: A compilation of useful quotations of S.R. Ranganathan from his various
works, 1994.

Putting Knowledge to Work: An American View of the Five Laws of Library Science, 1970. By
Pauline Atherton.

Read the dLIST Classics online!

<http://dlist.sir.arizona.edu/>



S.R. RANGANATHAN'S POSTULATES AND NORMATIVE PRINCIPLES

Applications in Specialized Databases Design, Indexing and Retrieval

Compiled by

A. NEELAMEGHAN

Hon. Visiting Professor

The Documentation Research & Training Centre

Indian Statistical Institute, Bangalore

Executive Director

Ranganathan Centre for Information Studies, Chennai

Foreword by

S. PARTHASARATHY

Hon. Director

Institute of Information Studies, Chennai



SARADA RANGANATHAN ENDOWMENT FOR LIBRARY SCIENCE

Bangalore, INDIA

1997

Ranganathan Centenary Series, 7

Ranganathan's FIVE LAWS of Library Science

Books are for Use
Every Reader his/her Book
Every Book its Reader
Save the Time of the Reader
Library is a Growing Organism

© Sarada Ranganathan Endowment for Library Science (1997)

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of the publisher.

Printed at Sharad Enterprises, Bangalore 560 003 and published by
Sarada Ranganathan Endowment for Library Science
c/o Institute of Information Studies,
702, 'Upstairs', 42nd Cross, III Block, Rajajinagar, Bangalore 560 010,
INDIA.

Dedicated to
Prof. S. Parthasarathy

For his significant contributions, in association with Dr. S.R. Ranganathan, to research in Classification, and the development of Science and Technology Information Systems and Services in India.

Contents

Foreword

S. Parthasarathy **i**

Preface

A. Neelameghan **iii**

Chapter 1

Hidden Roots of Classification

S.R. Ranganathan **1**

Chapter 2

Absolute Syntax and Structure of an Indexing and Switching
Language

A. Neelameghan **43**

Chapter 3

Design of Depth Classification: Methodology

S.R. Ranganathan **75**

Chapter 4

Subject Heading and Facet Analysis

S.R. Ranganathan **127**

Chapter 5

S.R. Ranganathan's General Theory of Knowledge
Classification in Designing, Indexing, and Retrieving from
Specialised Databases

A. Neelameghan **147**

Index **263**

RANGANATHAN CENTENARY SERIES

- 1 Information for Small Enterprises
Prepared by *A. Neelameghan*, 1992
- 2 Measuring Information
Some Problems and Direction for Information Science
By *Jean Tague-Sutcliffe*, 1991
- 3 Bibliometric Methods in Assessing Productivity and
Impact of Research
By *F.W. Lancaster*, 1991
- 4 Informetrics-91: Selected Papers from the Third International
Conference on Informetrics, 9-12 August, 1991, Bangalore
Edited by *I.K. Ravichandra Rao*, 1992
- 5 Memorabilia Ranganathan, 1994
- 6 Online Database Searching and Retrieval
Strategies, Procedures, Commands, and Problems:
A Brief Guide
Prepared by *A. Neelameghan*, 1995

FOREWORD

Nearly 70 years ago Dr. S.R. Ranganathan formulated his famous Five Laws of Library Science whose objective was to guide in maximizing the use of information and to meet the information needs of users efficiently, pinpointedly and expeditiously, in a fast changing environment. He recognised the user as the kingpin in designing all his innovative tools and techniques for organising information systems and services. His Colon Classification, Classified Catalogue Code, Reference Service, etc., are products of this fresh and vibrant approach, all designed toward the achievement of the goal of the Five Laws. In the process of designing tools for user-friendly systems, he developed many new concepts which are applicable in the newly emerging and fast developing area of digital information systems.

It is appropriate that Prof. A. Neelameghan has put together in this compilation some of the basic papers of Ranganathan on Subject Headings and Design of Depth Classification published in 1964. Ranganathan examines various possibilities for arranging the words in multi-worded subject headings in a helpful sequence which will match the approach of a majority of users. He states that the headings of subject entries taken together form an artificial language, where the syntax is one of position. Words in a multi-worded subject heading can be arranged in many sequences. So we cannot have a unique syntax of artificial language of subject headings. However, to arrive at a consistent and generally helpful syntax from the point of view of the users, the Postulates of Facet Analysis and Principles for Facet Sequence will be helpful. Ranganathan calls for an extensive study of the approach of users to arrive at a syntax acceptable to a large number of them. This would enable a deeper analysis of the problem for arriving at an Absolute Syntax which will give greatest satisfaction to the greatest number of users.

Ranganathan's papers on the Design of Depth Classification and the Hidden Roots of Classification consolidate his ideas developed over several years and fine-tuned from time to time. A major problem in the design of schemes for classification has been how to handle the two-fold infinity – the diversity in the diverse purposes of users and the infinity

in the dimensions of the universe of subjects. Another problem, which has a bearing on the notation of a scheme of classification, is the mapping of the multi-dimensional universe of subjects along a single dimension. For designing a scheme of classification, Ranganathan develops a series of postulates and normative principles. He spells out the various steps for classifying a document.

Prof Neelameghan was closely associated with Ranganathan during this period of development. He includes in this compilation two of his own papers. He shows how Absolute Syntax could function as a switching language – switching from one information System to another. Some parallel work done in the linguistic field by Chomsky, Fodor, Katz, Fillmore and others in search of universal linguistic forms are examined.

In presenting Ranganathan's General Theory of Knowledge Classification, Prof Neelameghan explains some of his experiences in the use of Normative Principles in the design, development and retrieval of information from machine readable databases, especially object-oriented specialised databases, and design of user-interfaces and hypertext links.

This publication brings together important contributions on Knowledge Classification and Subject Headings by Ranganathan and supplemented by Prof. Neelameghan. This will be useful for those engaged in and learning the designing of specialised schemes of classification, object-oriented databases, vocabulary control tools, various types of structured indexes, cognitive modelling and neural networks, etc.

It is most appropriate that this publication is released on Ranganathan's birthday in August this year and during the fiftieth year of India's Independence.

Chennai,
27 June 1997

S. Parthasarathy

PREFACE

The Five Laws of Library Science, Dr. S.R. Ranganathan's *magnum opus* were formulated in the 1920s and first published in 1931 under the same title. His observational visits to a number of libraries in Britain while he was completing his studies in library science at the University College London School of Librarianship and Archives in 1924, provided his fertile intellect necessary data to formulate these basic norms for library and information service. Subsequently, they formed, on the one hand the guidelines for further research in various aspects of information work and service, and on the other, they were used as the bench mark to test the value of any new development, be it theory, principle, technique or practice, in any aspect of library and information service. Over a period of close to half a century of research and service, there is hardly any aspect of library and information service that he had not touched and enriched with his worthy contributions. And the Five Laws served as guidelines and benchmark.

At the Cleveland Conference in 1959^{*}, Ranganathan indicated that the analytico-synthetic methodology and faceted classification model based on postulates and principles, which have formed the basis of his Colon Classification scheme, could give more helpful results in machine-based information storage and retrieval. In the late 1960s at the Documentation Research and Training Centre (DRTC) of the Indian Statistical Institute (ISI) in Bangalore, S. Venkataraman, a computer

International Conference for Standards on a Common Language for
Machine Searching and Translation, Cleveland, Ohio, USA

engineer associated with the teaching and research at the Centre, stressed that the machine should be used for more intelligent work than merely for printing catalogues, etc. and that the Colon Classification could provide a basis for such work. Semi-automatic synthesis of class numbers given a facet-analyzed co-terminus expression of a micro-subject and deriving feature headings and subject headings according to Ranganathan's Chain Procedure from a faceted class number, were achieved. Following this, by the beginning of the 1970s Ranganathan became increasingly interested in the use of the analytico-synthetic facet approach in computer-based information storage and retrieval, and in the design of special purpose machines. Prof Mitra of the ISI (Electronics Department), Calcutta, discussed with Ranganathan these aspects in Bangalore. Unfortunately, Ranganathan passed away in 1972. However, his ideas for the structuring, organization, and representation of knowledge/subjects in machine-based information systems have found productive applications and as areas for research in other countries.

Dr. Ranganathan was intensely patriotic and very proud of India's contributions to mathematics and the physical sciences, literature, and the moral, ethical, and philosophical disciplines. And we are proud of his contributions to the library and information field. Therefore, in celebrating the fiftieth year of India's Independence, it is only appropriate that we recall Ranganathan's contributions to knowledge organization and their relevance in the emerging environment of information technology.

In this publication of the Sarada Ranganathan Endowment for Library Science, established by Ranganathan in 1961, some of his basic

papers on classification, subject structuring and knowledge organization are reprinted along with two supporting papers of A. Neelameghan. It is hoped that teachers and students of library and information science will find this collection useful.

This book is dedicated to Prof S. Parthasarathy, former head of the Indian National Scientific Documentation Centre, New Delhi, one of the earliest students and a close associate of Dr. Ranganathan in many areas including research in classification. Prof Parthasarathy, who will be completing eighty years in 1997, is actively associated with the Institute of Information Studies, the Sarada Ranganathan Endowment for Library Science, and the Ranganathan Centre for Information Studies. We are grateful to Prof. Parthasarathy for writing the Foreword to this volume.

The sources of the papers included in the book are as follows:

Chapter	Title	Source
1	Hidden roots of classification	<i>Library Science</i> . 4(1); 1967 March; p.1-26
2	Ordering systems for global information networks	<i>Ordering Systems for Global Information Networks. Proceedings of the Third International Study Conference on Classification Research</i> . Bombay, 6-11 January 1975.
3	Design of depth classification: methodology	<i>Library Science</i> . 1(1); 1964; Paper A; p.1-42

Chapter	Title	Source
4	Subject heading and facet analysis	<i>Journal of Documentation</i> . 20(1); 1964, Sep.; p.109-119
5	S. R. Ranganthan's general theory of knowledge classification in designing, indexing, and retrieving from specialised databases	<i>Library Science</i> . 34(2); 1997, March; Paper A; p.3-53

We thank the respective copyright holders for permission to reproduce the papers in this collection.

Bangalore
1 July 1997

A.Neelameghan
Compiler

Chapter 1

HIDDEN ROOTS OF CLASSIFICATION*

S R RANGANATHAN

Points out that a twofold infinity characterises the design of classification — the infinity in the diverse purposes of the readers and that in the dimensions of the universe of subjects. Illustrates the risk in basing the design of classification on conjecture and suggests basing it on the results of a sound statistical study. With the aid of the analogy of roots of flames shows an alternative helpful method to be basing classification on postulates and principles inherent in the near-seminal level. Describes with illustrations a set of postulates and principles. Points out that classification is equivalent to mapping the multi-dimensional universe of subjects along a single dimension and that the crucial problem in such a mapping is, determining which of the many immediate neighbourhood relations among facets to the Basic Subject, should be kept invariant. The Wall-Picture Principle is shown to give the best result in this matter. This Principle fixes the syntactical relations among facets. Raises the issue whether there is an Absolute Syntax, and whether the syntax of facets implied in the Wall — Picture Principle is equivalent to it. Suggests the investigation of this problem by a team of Linguists, Psychologists, Classificationists, Reference Librarians and Statisticians.

0 INTRODUCTION

It has been kind of the organizers of this Symposium to have thought of me for this concluding talk in the series. On my part, I had been eagerly looking forward to be amidst you this week, observe your

* Based on an Address given on 11 June 1966 at the Symposium on Relational Factors in Classification, organised by the University of Maryland, College Park, Md., USA.

ways of thinking on the subject of the Symposium, and exchange ideas in this fascinating subject of the Foundations of Syntactic Relations in Classification. For, from the list of names of the participants furnished by Mr. Perreault, I knew that I would be in the midst of a galaxy of Classificationists.

0.1 Denial of Opportunity

But the old saw, "Man proposes; God disposes" asserted itself at the last moment. The health of the physical body reached a level which has made medical advisers to pass an "Order of Internment" on me. I have been, therefore, painfully obliged to deny myself the benefit of discussing the very problem in which I am now getting involved deeper and deeper, day by day, and you yourself have chosen for this Symposium.

0.2 First Approximation

However, I had the opportunity to get an inkling of your approach to the problem by a perusal of the introductory "Challenge" Paper sent by Perreault. This enables me to send you this script, as a crude first approximation to my sharing the thought with you in person.

1 PLIGHT OF CLASSIFICATIONIST

The problem to be faced by the Designer of a scheme for the depth classification of nascent micro-thought, gushing forth from the minds of the hundreds of research workers in each of thousands of narrow

Hidden Roots of Classification

subject fields and getting embodied in millions of articles published from year to year – the gravity of the problem of the Classificationist is not easily realized. Because, his design work concerns an invisible, intangible commodity called Thought or Idea. Therefore, I shall use the analogy of design work in a field where what is designed is visible to the eye.

1.1 Plight of Architect

Imagine the plight of an Architect designing a satellite township for a City Extension Board. Imagine one member after another of the Board suggesting changes day after day. Imagine the building material to be used having to be changed several times during the period of design and even during construction. Imagine also the sudden unexpected floods and earth tremors – the first within living memory in the locality – forcing a redesign of the foundation at an advanced stage. Imagine further the machinery and the tools for construction having to be changed frequently.

1.2 Greater Precariousness

The plight of Classificationist is even more precarious than that of the Architect. For, in the universe of ideas turbulent changes are becoming frequent, unpredictable, and violent. The disturbance caused in the universe of ideas by eruptions such as those of Electrons, Laser, Plastics, Statistical Analysis, and Breaking of the Atom are now straining the schemes for classification and even breaking some of them, The purpose of the consumers of ideas is infinitely more varied than those

in the case of township. Each reader wants it all his own way. In fact, the Classificationist has to reckon with a far greater number of factors than the Architect — almost tending to infinity.

2 TWO-FOLD INFINITY

As a matter of fact, a two-fold infinity characterises the domain of Depth Classification. Firstly, there is infinity inherent in the purposes of readers and consequently in their approaches to a collection of documents in the stack or their main entries in the catalogue. Secondly, there is infinity in the dimensions of the universe of ideas to be organized by the Classificationist. Let us leave infinity alone. Let us confine ourselves to large numbers.

3 LARGE NUMBER OF PURPOSES OF READERS

3.1 Problem of Classificationist

The problem of a Classificationist is as follows:

Different readers have different purposes; and even the same reader has different purposes at different times. But the Classificationist cannot provide a different scheme for classification to suit each of the large number of purposes. He cannot simulate the old man of Aesop's, going out with his son and the donkey. Classificationist is obliged to use the statistical idea "Mode". He has to design the scheme

to suit the most dominant purpose prevailing among readers. How to single out one out of many purposes as the most dominant one?

3.2 Method of Conjecture

We cannot deal with a situation involving large numbers in the way followed for small numbers. In dealing with situations involving large numbers, humanity depended, for a long long time, on mere conjecture by persons rich in experience and intelligence. That is what Classificationists are doing till now.

3.3 Method of Conjecture and Colon Classification

While starting the designing of Colon Classification in November 1924, I had little experience of the purposes of readers. I had to find them out, and also find out the dominant purpose. From 1924 to 1928 – that is, till I completed the “In-Service-Training” of my first staff – I was the only professional librarian in the Madras University Library. While studying in the University College of London in 1924, I had the opportunity to listen to a talk by E C Richardson on what he called, “Research Consultant”. His talk made me realise that Reference Service – helping the individual reader – was the *summum bonum* of library service. Therefore, in those early years I spent nearly six hours a day in floor duty – helping the readers. This gave me an insight into the varying purposes and approaches of readers. It also helped me to conjecture the sequence of the facets in a subject, that would serve the dominant purpose and approach of the readers. On this conjecture was based the sequence of facets in subjects going with each Main Subject. In other words, this conjecture helped me to lay down my Facet Formula for different Main Subjects. I began with a library of about 30,000

volumes. By the end of 1928 it grew to about 50,000 volumes. All these volumes were classified on the basis of the different Facet Formulae constructed according to the conjecture about the dominant purpose and approach. On the whole, the arrangement of books appeared to give satisfaction. I began to write out the First Edition of the *Colon classification* in 1929. I also began to teach the subject at that time. While developing the different Facet Formulae, I used to experiment upon the way in which the different students preferred the Facet Sequence. However, as a student of Statistics, I was aware of the limitation of conjecture. Therefore, before completing the press copy of the book, I made an intensive experiment with the senior readers.

4 RISK IN CONJECTURE

In compound subjects going with the Main Subject Law, we often have two facets. Our Law Collection was very meagre. Only a few students of the Law College used it. They had to read mostly Indian Law. They also read British Law to some extent. The Indian Law of those years was largely based on the British Law. This made the students always approach the subject from the angle of the Problem Facet instead of the Community Facet. In other words, the Problem Facet was the first that they would bring up. The Community Facet was quite secondary to them. This had led me to use the Facet Sequence:

Law [Problem] [Community]

as helpful to the dominant purpose and approach of the readers. To check up the validity of this conjecture, I invited some Members of the

Hidden Roots of Classification

Bar, Judges of the High Court, and experienced Jurists. I used to take them into the Law Gangway. I learnt from them that the Facet Sequence should be the very opposite *viz.*,

Law [Community] [Problem]

This was confirmed by most of the senior readers. They attributed the students' approach, resulting in the false conjecture, to the faulty method of teaching Law. Students were made to know only of the modern Indian and British Laws, which were alike; and they were not told that each country had its own system of Municipal law. I had also a similar experience with Chemistry. In these two Main Subjects all the books had to be reclassified and the class numbers had to be altered in all the places of their occurrence – in several places in the book itself, in the catalogue entries, shelf register, and accession register. This is the price to be paid when risk asserts itself in conjecture. Fortunately, in most of the other subjects the conjecture turned out to be correct. But it is the exception that highlights the “Risk in Conjecture” wherever large numbers are involved.

5 STATISTICAL METHOD

To minimize the risk of conjecture, the determination of the dominant purpose of readers must be based on a statistical study. This in its turn should be based on extensive observations in libraries of all standards – public, academic, and specialist – and in libraries of all countries. This will be far too expensive. At any rate, this has not been done. The result is that a certain amount of risk is taken by most of the

schemes for classification by basing themselves on conjecture. It is not easy to make a statistical study of the problem of dominant approach of readers, the most helpful sequence of facets, the most helpful succession of characteristics within each facet, and the most helpful sequence of isolates within each array. The library profession has yet to clarify its own ideas in formulating such problems needing statistical study. The precautions to be taken in statistical observation in these cases have yet to be formulated. There is every probability for the digestion of the observed statistical data demanding new statistical techniques. In spite of the immaturity and infancy of this line of statistical approach to classification, I hope that some rich Foundation will arrange for this statistical determination of the dominant purpose and approach of readers. Again, this may have to be repeated in different epochs to find out the change that might occur.

6 LARGE NUMBER OF DIMENSIONS IN THE UNIVERSE OF IDEAS

The statistical negotiation of the large number of purposes and approaches of readers with statistical methods will take long. Depth classification cannot wait till then. Let us therefore turn our attention to the large number of dimensions of the universe of ideas. Let us see whether we can find an *a priori* method for organizing the idea-masses – from Macro through Micro to Spot Idea – in a helpful way and thereby find methods for the design of depth classification. The aim and end of classification is to arrange ideas scattered in many dimensions, along a line – in a linear sequence. How are we to do it?

6.1 Analogy of Roots of Flame

The scatter of the tongues of flame shooting up from a pile of logs cannot be changed to our liking and brought into a line by catching the tongues of flame and rearranging them. The right method will be to manipulate the logs forming the root of the flames. So it is with classification of the universe of subjects – that is, the arrangement of subjects. The universe of subjects is dynamic. It throws forth subjects in a turbulent manner. It is doing so incessantly in our times. The scatter of the subjects so thrown forth is in many dimensions and is unhelpful. To anticipate the subjects and arrange them in advance in a helpful sequence would prove as futile and maddening as manipulating the tongues of flame. Because the subjects will not burn us, we are often tempted to secure a helpful classification of subjects by directly manipulating them. The result is frequent break-down and even the desperate declaration that classification is impossible and alphabetical arrangement by names of subjects is the only sensible way. The subjects are the tongues of flame. They belong to the phenomenal level. It is tiring and ineffective, if not futile, to manipulate the subjects directly and arrange them as desired. It is best to manipulate their roots at the near-seminal level at which all the subject-proliferations are traceable to a few roots. The roots of subjects, are hidden even at the near-seminal level. They are hidden in the sense that they cannot be reached by intellectual analysis. They will have to be apprehended with intuition. If intuition is functioning cent percent the roots can be unerringly and permanently located. Hardly anybody is found with cent

percent intuition. Further, in the scale of values of anybody worth cent percent intuition, classification may find itself very near the bottom. Therefore, we have to depend upon whatever can be got through the play of a momentary flash of intuition in some person or other — essentially intellectual. Postulates and Principles are usually disclosed by such momentary flashes. They may go a long way though not the full way. When they cease to be helpful, they may be replaced by another set of Postulates and Principles that may be disclosed at that time.

6.2 Postulates

We have now a set of postulates to guide the design of classification. The attributes “True” and “False” do not apply to postulates. The only attributes applicable to them are “Helpfulness” and “Unhelpfulness”. The following set of postulates have proved helpful.

6.2.1 *Postulate of Fundamental Categories*

There are five and only Five Fundamental Categories – *viz.*, Time, Space, Energy, Matter, and Personality.

These terms and the ideas denoted by them belong strictly to the context of classificatory discipline. They have nothing to do with their use in Metaphysics or Physics. In our context, their significance can be seen only in the statements about the facets of a subject — their separation and their sequence. This set of fundamental categories is, for brevity, denoted by the initionym PMEST.

Hidden Roots of Classification

6.2.1.1 Time

Perhaps the fundamental category “Time” gives the least difficulty in its identification. It is in accordance with what we commonly understand by that term. The usual Time Isolate Ideas — such as millennium, century, decade, year, and so on — are its manifestations. Time Isolate Ideas of another kind — such as day and night, seasons such as summer and winter, time with meteorological quality such as, wet, dry, and stormy — are also taken as manifestations of the fundamental category “Time”.

6.2.1.2 Space

The fundamental category “Space” comes next to “Time” in difficulty in its identification. It is in accordance with what we commonly understand by that term. The surface of the earth, the space inside it, and the space outside it, are manifestations of the fundamental category “Space”. The usual Geographical Isolate Ideas — such as continents, countries, and counties — and water formations — such as oceans and seas — are taken to be its manifestations. Physiographical Isolate Ideas — such as desert, prairie, rain-forest, plateau, mountain, river, and lake — are also taken to be manifestations of the fundamental category “Space”. So also an area occupied by a population-cluster — such as a city, a town, and a village — is taken to be a manifestation of the fundamental category “Space”.

6.2.1.3 Scope For Comparative Study

Both CC and UDC give schedules of Time and Space Isolates. Therefore, there is some material for comparative study. On the basis

of this study, it has been possible to cultivate the region of classificatory discipline falling within the purview of the fundamental categories “Time” and “Space”. The other schemes have not developed this region to a sufficient degree. CC is the only scheme with distinctive and consciously enumerated schedules of the three fundamental categories “Energy”, “Matter”, and “Personality”. The “Analytical Subdivisions” of UDC are of a casual nature. They are also mixtures of the manifestations of all the three fundamental categories. Therefore, there is no good scope for comparative study in respect of the manifestations of these three fundamental categories. In the circumstances, as the only meagre scope for comparative study, we have to use different editions of CC.

6.2.1.4 Energy

Even otherwise, the identification of the fundamental category “Energy” is a little more difficult than that of “Space” or “Time”. Generally speaking, its manifestation is an action of one kind or another. The action may be among and by all kinds of entities – inanimate, animate, conceptual, intellectual, and intuitive.

Till now, we have been taking Morphology, Physiology, Disease, Ecology, Hygiene, and some other isolate ideas also as manifestations of the fundamental category “Energy”. It was difficult to see any “Action”

Hidden Roots of Classification

in them. Therefore, we enumerated them in a schedule and labelled them as Energy Isolates. We were led into this position by a sheer accident. This was the use of the term 'Problem' to denote these isolate ideas – from Ed 1 (1933) onwards of the *Colon classification*. This gave rise to a blind tradition of thirty years' standing. Wrong traditions die hard. With the rush of preoccupation with many other ideas in classification and in other branches of library science, hardly any time or inclination was found to look at these 'Problems' critically. But the time has now come to give up this tradition. These isolate ideas are no longer taken to be manifestations of the fundamental category "Energy". Edition 7 of CC will incorporate this change.

6.2.1.5 Matter

The identification of the fundamental category "Matter" is more difficult than even of "Energy". Its manifestations are taken to be of two kinds – Material and Property. It may look strange that property should be taken along with material. But let us take a table as an example. The table is made of the material, timber or steel, as the case may be. The material is intrinsic to the table, but is not the table itself. Moreover, the same material can figure also in several other entities. So also, the table has the property of being two and a half ft. high and the property of having a soft top or a hard top. The property is intrinsic to the table, but not the table itself. Moreover, the same property can figure also in several other entities. Each of the isolate ideas Morphology, Physiology, Disease, etc., mentioned in the preceding section and now being excluded from the manifestations of the fundamental category

“Energy”, admit of being looked upon as Property. As such, they are now regarded as manifestations of the fundamental category “Matter”.

6.2.1.6 Personality

The fundamental category “Personality” presents the greatest difficulty in its identification. It is too elusive. It is ineffable.

6.2.1.7 Method of Residues

If a certain manifestation is easily determined not to be one of “Time” or “Space” or “Energy” or “Matter” it is taken to be a manifestation of the fundamental category “Personality”. This is the Method of Residues. For, according to the postulate, there are five and only five fundamental categories. Therefore, any entity, which is not a manifestation of “Time” nor of “Space” nor of “Energy” nor of “Matter”, should be a manifestation of “Personality”. The application of this Method of Residues may not be easy in certain cases. But experience will lead to the establishment of reflex action in recognizing the fundamental category manifesting itself in any isolate idea, even as experience leads to the establishment of a reflex action in recognizing Chinese, Indians, Egyptians, Italians, French, Germans, and Russians. This does not amount to saying that there is no difficulty at all. There are still some areas of doubt in distinguishing between manifestations of the fundamental categories “Energy”, “Matter”, and “Personality”. These difficult areas do not turn up very often. Therefore, we can get along, in spite of this difficulty, for the time being, and solve it in due

Hidden Roots of Classification

course as experience increases. We cannot give up the proven advantage of Classification Guided by Postulates and Principles, and say, "We shall begin to use them only when all the difficulties about them are finally solved".

6.2.2 Postulate of Basic Facet

Every Compound Subject has a Basic Facet.

This is implied in the very definition of the term 'Compound Subject'. A subject may have two or more basic facets. Then it will be a case of phase relation between the basic facets themselves or between the compound subjects of which they are the respective basic facets or a case of one of the subjects figuring as an isolate facet in a compound subject going with the other basic facet.

6.2.2.1 Recognition of the Basic Facet

To recognize the Basic Facet of a Compound Subject, a general knowledge of the schedules of Basic Subjects is necessary. Most of the Schemes for Classification give roughly similar schedules of them. The indication, by the title of a document of the Basic Facet of its subject may be either:

1 Explicit; or 2 Implicit in the context; or 3 Absent. Here are examples of the first two possibilities.

SN	Indication	Title	Basic Facet
1	Explicit	1.1 Treatise on coal mining	Mining
		1.2 Agricultural diseases	Agriculture
		1.3 Text-book of Indian History	History
2	Implicit	2.1 Structure of proteins	Chemistry
		2.2 Care of cows	Animal husbandry
		2.3 Income-tax	Economics

6.2.2.2 *Absence of Indication*

If the title does not express the subject at all but is oblique or fanciful, the contents-page and even the whole document may have to be perused to determine the Basic facet. Most of the Works in Literature and many Classics in diverse subjects come under this group.

6.2.3 *Postulate of Isolate Facet*

Each isolate facet of a compound subject can be deemed to be a manifestation of one and only one of the five Fundamental Categories.

It is generally easy to identify isolate ideas that are manifestations of the fundamental categories: Time, Space, Energy, and Matter. As already stated in Sec. 6.2.1.7 any isolate idea, not found to be a manifestation of any of these four categories, has a good chance to be a manifestation of the fundamental category "Personality". Its manifestation can also be directly sensed in some cases. Some examples are given in the succeeding sections.

Hidden Roots of Classification

6.2.3.1 Biological Sciences

1. In the Subjects in Botany — Plant Group. Plant.
2. In the Subjects in Agriculture — Cultivar Group. Cultivar.
3. In the Subjects in Zoology — Animal Group. Animal. Organ.
4. In the Subjects in Animal Husbandry — Animal. Organ.
5. In the Subjects in Medicine — Human Body. Organ.

6.2.3.2 Social Sciences

1. In the Subjects in Education — Child. Adolescent. Adult, Genius. Imbecile. Blind.
2. In the Subjects in History and Political Science — Head of the State. Executive. Legislature. Party. Public. Local Body. Judiciary. Civil Service.
3. In the Subjects in Sociology — Rural Folk. City Folk. Professional Group. Working Class. Royalty. Aristocracy. Middle Class. Military Class. Nomadics. Aryans. Semetics. Hindus. Christians. Muslims. Indians. Chinese. British.
4. In the Subjects in Law — Legal Personality. State. Association. Property. Contract. Treaty. Tort. Crime. Cause of Action. Court.

6.2.3.3 Humanities

1. In the Subjects in Linguistics — a Language. Phoneme. Syllable. Word. Phrase. Clause. Sentence. Piece of Composition. Punctuation.

2. In the Subjects in Religion — Vedic Religion. Hinduism. Vaishnavism. Saivism. Jainism. Buddhism. Christianity. Islam. Shintoism. Sikhism. Zoroastrianism. Sacred Book. Church. Sects.
3. In the Subjects in Psychology — Child. Adolescent. Adult. Genius. Imbecile. Blind.
4. In the Subjects in Literature — Poetry. Drama. Fiction. Author. Work.

6.2.3.4 Mathematics

1. In the Subjects in Arithmetic — Prime Number. Partition of Numbers. Arithmetical Function.
2. In the Subjects in Theory of Equation — Simple Equation. Quadratic Equation. Abelian Equation.
3. In the Subjects in Higher Algebra — Binary Form. Linear Form. Cremona Transformation.
4. In the Subjects in Differential Equation — Linear, Quadratic, Cubic, Quartic, Quintic, Sextic. First Order. Second Order. Third Order. Fourth Order. Fifth Order. Sixth Order.

6.2.3.5 Physical Sciences

1. In the Subjects in Properties of Matter — Solid. Glass. Crystal. Liquid-Surface. Liquid. Gas.

Hidden Roots of Classification

2. In the Subjects in Sound – Audible Sound. Infra Sound. Ultra Sound.
3. In the Subjects in Radiation – Light. Ultra-Violet Ray. X-Ray. Gamma Ray. Infra-Red Ray. Hertzian Wave.
4. In the Subjects in Electricity – Current. Direct Current. Alternating Current. Weak Current.
5. In the Subjects in Magnetism – Dia-Magnetism. Para-Magnetism. Terrestrial Magnetism.
6. In the Subjects in Nuclear Physics — Neutron. Nutrino. Proton. Beta Ray. Meson. Cosmic Ray.
7. In the Subjects in Chemistry — Inorganic Substance. Hydrogen. Calcium. aluminum. Carbon. Bismuth. Oxygen. Fluorine. Iron. Metal. Non-Metal. Alloy. Basic Oxide. Acid. Salt.
8. Organic Substance. methane. Phenophthalene. Carbohydrate. Starch. Aromatic Compound. Benzene. Heterocyclic Compound. Alkaloid. Amino Acid. Protein. Vitamin. Hormone. Chlorophyll. Enzyme.

6.2.3.6 Recognition of Isolate Ideas

The indication, by the title of a document, of the isolate facets of its subject may be either (1) Explicit, or (2) Implicit in the context, or (3) Hidden within a derived composite term, or (4) Absent. The basic

subject of the document will be of help in sensing the absence of the indication of a necessary facet of the compound subject. Experience will develop the capacity for sensing this. In that case, the contents page or even the whole document should be perused to find the absent isolate ideas, if any. Again, experience will develop the capacity to sense the derived composite terms in a title and to break it into its fundamental constituent terms. Some examples are given in the following table. In each example, against each isolate idea appropriate symbol is given to indicate the fundamental category of which it can be deemed to be a manifestation.

Symbols used:

(BF) = Basic Facet [S] = Space Facet
 [E] = Energy Facet [P] = Personality Facet
 [M] = Matter Facet [T] = Time Facet

SN	Indication	Title	Basic and Isolate Facets
1	Explicit	1.1 Coal washing	Mining (BF). Coal [P]. Washing [E].
		1.2 Control of virus diseases of the stem of rice plant in the winter of 1967 in Madras	Agriculture (BF). Control [E]. Virus disease [M]. Stem[P]. Rice plant [P]. Winter [T]. 1967 [T]. Madras [S].

Hidden Roots of Classification

SN	Indication	Title	Basic and Isolate Facets
	Explicit	1.3 Election of the President of the Congress Party in India in 1967	History (BF). Election [E]. President [P]. Congress Party [P]. India [S]. 1967 [T].
2	Implicit (The implied facets are in italics)	2.1 The structure of protein and electron microscope	<i>Chemistry</i> (BF). Structure [M]. Protein [P]. <i>De-termination</i> [E]. Electron microscope [M].
		2.2 X-ray diagnosis in cow farming	<i>Animal husbandry</i> (BF). X-Ray [M]. <i>Diagnosis</i> [E]. Disease [M]. Cow [P].
		2.3 Tape-record and protection of the folk songs of the Todas	<i>Sociology</i> (BF). Taperecord [M]. Dying out [M]. Protection [E]. Folk songs [M]. Todas [P].
3	Hidden within a Derived composite term (the hidden facets are in italics)	3.1 Phthisis	<i>Medicine</i> (BF). <i>Lungs</i> [P]. <i>Tubercular disease</i> [M].
		3.2 Indian franchise in 1967	<i>History</i> (BF). India [P]. <i>Citizens</i> [P]. Franchise [M]. 1967 [T].
		3.3 Birth control essential in India today (1966)	<i>Sociology</i> (BF). <i>Overpopulation</i> [M]. <i>Prevention</i> [E]. Birth control [M]. India [S]. Today [T].

6.2.4 Postulate of Rounds for Energy

The fundamental category “Energy” may manifest itself in one and the same subject more than once. The first manifestation is taken to

end Round 1 of the manifestation of the three fundamental categories “Personality”, “Matter”, and “Energy”. The second manifestation is taken to end Round 2. And so on.

We shall denote the manifestations of the fundamental category “Energy” in Rounds 1,2 etc., by the respective names Round 1 Energy Facet, Round 2 Energy Facet, etc. We shall represent them by the respective symbols [1E], [2E], etc. Which manifestation of the fundamental category “Energy” should be deemed to be the Round 1 which the Round 2 will be determined by the Wall-Picture Principle.

6.2.5 Postulate of Rounds for Personality and Matter

Each of the fundamental categories “Personality” and “Matter” may manifest itself in Round 1, Round 2, and so on.

We shall denote the manifestation of the fundamental category “Personality” in the Rounds 1, 2 etc., by the names Round 1 Personality Facet, Round 2 Personality Facet, etc. We shall represent them by the respective symbols [1P], [2P], etc.

So also we can have Round 1 Matter Facet, Round 2 Matter Facet, etc. We shall represent them by the respective symbols [1M], [2M], etc.

6.2.6 Postulate of Round for Space and Time

Ordinarily, any of the fundamental categories “Space” and “Time” may manifest itself only in the last Round in a subject.

Hidden Roots of Classification

We shall represent them by the respective symbols [S], [T].

6.2.7 Postulate of Level

Any of the fundamental categories “Personality” and “Matter” may manifest itself more than once in one and the same Round within a subject; and similarly with “Space” and “Time” in the Last Round. The first manifestation of a fundamental category within a Round will be said to be its Level 1 Facet in that Round. Its second manifestation within that Round will be said to be its Level 2 Facet in that Round. And so on.

6.2.7.1 Personality and Matter

We shall call the successive manifestations of the fundamental category “Personality” in the Round 1 by the respective names Round 1 Level 1 Personality Facet, Round 1 Level 2 Personality Facet, etc. We shall represent them by the respective symbols, [1P1], [1P2], [2P1], [2P2], etc. Similarly, [1M1], [1M2], [2M1], [2M2], etc.

6.2.7.2 Space and Time

Since Space and Time Facets can occur only in the last round of a subject, there is no need to indicate the Round in their names or their symbols. It is sufficient if we represent them by the respective symbols [S1], [S2], [T1], [T2], etc.

6.2.7.3 Energy

The fundamental category “Energy” can occur only once within a Round. Therefore, no Level in its case.

6.2.8 Postulate of First Facet

In a Compound Subject, the Basic Facet should be the first facet.

Every Compound Subject, should have a Basic Facet. Again, Isolate Facets can form a subject, if and only if they are attached to a Basic Facet. Helpfulness requires that all the Compound Subjects going with a Basic Facet should be arranged together. To secure this, the Basic Facet should be given the First Position among the facets of a Compound Subject.

6.2.9 Postulate of Concreteness

The five fundamental categories fall into the following sequence, when arranged according to their decreasing concreteness: P, M, E, S, T.

This Postulate conforms to what the majority of persons think in respect of the relative concreteness of the isolates which are manifestations in each of the five respective fundamental categories.

6.2.10 Postulate of Facet Sequence Within a Round

In any Round of Facets of a Compound Subject in which each of any of the first three fundamental categories occurs only once, their

Hidden Roots of Classification

sequence should be: Personality Facet, Matter Facet, and Energy Facet.

6.2.11 Postulate of Facet Sequence Within the Last Round

In the Last Round of facets of a Compound Subject, in which each of the fundamental categories other than Energy occurs only once, the sequence of the facets should be Personality Facet, Matter Facet, Space Facet, and Time Facet.

6.2.12 Postulate of a Level Cluster

Facets of different levels of the same fundamental category within a Round of facets in a Compound Subject should be kept together.

6.2.13 Alternative Sets of Postulates

It will be helping the cause of classification, if an alternative set of postulates is forthcoming. This will happen when the present set of postulates begins to prove unhelpful. This conjecture is based upon the history of science and upon the formulation of fundamental laws such as Postulates, Canons, Principles, and Hypotheses.

6.3 Wall-picture Principle

If two facets A and B of a subject are such that the concept behind B will not be operative unless the concept behind A is conceded, even as a mural picture is not possible unless the wall exists to draw upon, then the facet A should precede the facet B.

1. In “Cure of Disease” the concept behind the term ‘Cure’ is not operative unless the concept behind the term ‘Disease’ is conceded. Therefore, when expressed in transformed skeleton form, we shall have ‘Disease. Cure’.

In this case, the application of the Wall-Picture Principle has determined that the Round to which the concept ‘Disease’ should be assigned as the one preceding the Energy Facet ‘Cure’.

2. In “Prevention of Disease” also, the concept behind the term ‘Prevention’ is not operative unless the concept behind the term ‘Disease’ is conceded. Therefore, when expressed in transformed skeleton form, we shall have ‘Disease. Prevention’.

Thus, the application of the Wall-Picture Principle has determined that the Round to which the concept ‘Disease’ should be assigned as the one preceding the Energy Facet ‘Prevention’.

3. A comparison of examples 1 and 2 leads to an important warning in applying the Wall-Picture Principle. In example 1, ‘Disease’ actually comes in before ‘Cure’ begins. But in example 2, ‘Disease’ does not come in at all. Indeed, ‘Prevention’ is to secure that it does not come. In the former the concept as well as what is conceived are conceded before ‘Cure’ begins. In the latter, the concept ‘Disease’ alone is conceded, but not

Hidden Roots of Classification

'Disease' itself, before 'Prevention' begins. Thus, in applying the Wall Picture Principle it is only the concept that should be conceded, but not the correlate of the concept existing outside the mind.

4. In "President of India", the concept behind the term 'President' is not operative unless the concept behind the term 'India' is conceded. Therefore, when expressed in transformed skeleton form, we shall have 'India. President'.

In this case, the application of the Wall-picture Principle has determined the respective Levels to which the concepts 'India' and 'President' should be assigned.

5. Consider "Release of Contract in India". The concept behind the term 'Release' is not operative unless the concept behind the term 'Contract' is conceded. Further, the concept behind the term 'Contract' is not operative unless the concept behind the term 'India' is conceded. Therefore, when expressed in a transformed skeleton form, we shall have 'India. Contract Release'.

In this case, the application of the Wall-Picture Principle has determined the respective Levels to which the concepts 'India', 'Contract'. and 'Release' should be assigned.

6. Consider "Hamlet by Shakespeare, the English Dramatist". The concept behind the term 'Hamlet' is not operative unless the concept behind the term 'Shakespeare' is conceded. Again,

the concept behind the term 'Shakespeare' is not operative, unless the concept behind the term 'Drama' is conceded. So also, the concept behind the term 'Drama' is not operative unless the concept behind the term 'English' is conceded. Therefore, when expressed in transformed skeleton form, we shall have 'English. Drama. Shakespeare. Hamlet'.

In this case, the application of the Wall-Picture Principle has determined the respective Levels to which the concepts 'English', 'Drama', 'Shakespeare', and 'Hamlet' should be assigned.

7. Supplementary to Postulates. – The Wall-Picture Principle and the Postulates for Facet Sequence will produce the same result wherever they are both applicable. In those cases, we need not invoke the aid of the Wall-Picture Principle. But in the examples given above, the Postulates by themselves cannot determine the Round and the Levels indicated. Thus, the use of the Wall-Picture Principle is supplementary to the use of the Postulates. The former is more versatile than the latter.

6.3.1 Whole-Organ Principle

If, in a subject, facet B is an organ of facet A, then A should precede B.

Consider "The Public Accounts Committee of the Parliament of India". The Facet 'Public Accounts Committee' is an organ of the facet 'Parliament'. the facet 'Parliament' itself is an organ of 'India', When

Hidden Roots of Classification

expressed in a transformed skeleton form, we shall have 'India. Parliament. Public Accounts Committee'.

This sequence of Levels can also be inferred directly from the Wall-Picture Principle.

However, the Levels shown for a subject in Law, and the Levels shown for a subject in Literature, are not in the relation of 'Whole' and 'Organ'. Therefore, those Levels can be inferred only by directly invoking the Wall-Picture Principle.

6.3.2 Cow-Calf Principle

If a facet A and another facet B belonging to the same subject are not to be separated though they are distinct from each other and thus separable, A and B should be kept together in the same Round, even as a milch cow and its unweaned calf are not separately sold out though they are distinct entities and thus separable, but are kept together in possession of the same owner.

Consider "Enforcement of the Functions of the President of India". Here, the three facets 'India', 'President', and 'Functions' are not to be separated and put into different Rounds, although they are separable. They should all be put together in Round 1— that is, before the Energy Facet, 'Enforcement'— or after it. We cannot put any one of them in Round 1 and the other two in Round 2. The Cow-Calf Principle determines only that all the three facets should be put in one and the

same Round. To decide which Round it should be, we should invoke the direct aid of Wall-Picture Principle. This Principle would definitely assign them to Round 1. Therefore, when expressed in transformed skeleton form, we shall have 'India. President. Function. Enforcement'.

We can also get the same result by repeated application of the Wall-Picture Principle.

6.3.3 Actand-Action-Actor-Tool Principle

If, in a subject, facet B denotes action on facet A by facet C, with facet D as the tool, then the four facets should be arranged in the sequence A, B, C, D.

Consider "Charkha Cotton Spinning by Girls". (Charkha is a simple spinning instrument revived and brought into great prominence by Mahatma Gandhi during the days of Freedom Movement.) Here, the Action is 'Spinning'; the Actand is 'Cotton'; the Actor is 'Girls'; and the Tool is 'Charkha'. Therefore, when expressed in transformed skeleton form, we shall have 'Cotton. Spinning. Girls. Charkha'.

This result can also be got by the repeated application of the Wall-Picture Principle.

6.4 Linear Arrangement of Subjects and its Necessity

The human mind is, after all, at a very early stage in its evolution. Although we can speak of many dimensions, it usually works, more or

Hidden Roots of Classification

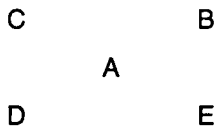
less, in one dimension. Even mathematicians have to work “bit by bit along the line”. There may be exceptions; but most serious thinkers have to think out one thing at a time in succession. In particular, the documents in the stack and their main entries in the catalogue have to be in linear sequence. The search for any one document or its entry has to be made by scanning along the line. But the universe of subjects has many dimensions. Let us say, n dimensions, where n is a large positive integer. The subjects in the Universe of Subjects have to be arranged in a line for the convenience of readers.

6.5 Mathematical Transformation and Mapping

To state this in mathematical terms, we have to transform the n -dimensional space into one-dimensional space. In other words, we have to map an n -dimensional space on a one-dimensional space. This is the problem in classification.

6.6 Invariant Among Immediate-Neighbourhood- Relations

Consider the five points spread out on a plane.



Here B, C, D, and E each claims Immediate-Neighbourhood-Relation with A. Let us arrange all the five points in one line. Let us put A at the left end. Then there can be only one Immediate-Neighbourhood-Position after A.

$$\frac{X}{A}$$

We can give that position *only to one* of B, C, D, and E and not to all. To which shall we give that position? In other words, which of the four Immediate-Neighbourhood-Relations should be kept invariant while arranging the five points along a line? Consider the points as subjects. This lays bare our inescapable problem in classification. It is a mischief created by the mathematics of transformation and mapping. If we begin to ask which of B, C, D, and E should be given the benefit of keeping invariant its Immediate-Neighbourhood-Relation with A, the chances will be equal to all the four elements. This tantalising problem attains colossal dimensions when we have to arrange millions of micro-subjects.

6.7 Descent Towards the Seminal Level

In the phenomenal world there are millions and millions of subjects. We do not know which of the Immediate-Neighbourhood-Relations should be kept invariant in arranging the subjects in a helpful way along a line. A suitable method would be to descend from the phenomenal level nearer to the seminal level. According to the Postulate of Fundamental Categories, we should descend down and down, and down and down, and allow the various subjects and ideas to become absorbed and reassembled, reabsorbed, and again reassembled, and so on, until we find only five ultimate generic ideas – seminal ideas, Fundamental Categories – standing out.

7 PRACTICAL CLASSIFICATION

7.1 Steps in Practical Classification

Classifying a document consists of the following steps in succession.

Step 0.— Raw Title (= Title as found in the document).

Step 1.— Full Title (= Title expressing each of the relevant basic and isolate ideas in the subject of the document, got by filling up all the ellipses in the Raw Title).

Step 2.— Kernel Title (= Full title minus all the auxiliary or apparatus words and with each composite term denoting a composite idea replaced by the fundamental constituent terms denoting its fundamental constituent ideas).

Step 3.— Analysed Title (= Kernel Title with each kernel term marked by a symbol, denoting the fundamental category of which the idea denoted by the term is a manifestation and also the round and the level to which it is assigned in conformity to the Postulates of Classification).

Step 4. — Transformed Title (= Analysed Title with the kernel terms rearranged according to the symbols of analysis attached to them).

Step 5. — Title in Standard Facet Terms (= Transformed Title with the kernel terms replaced, wherever necessary, by their respective equivalent terms as given in the appropriate schedules).

Step 6. — Title in Facet Numbers (= Title in Standard Terms with the kernel terms replaced by their equivalent numbers from the schedules).

Step 7. — Class Number (got by removing the symbols of analysis and inserting the appropriate Connecting Symbols between the facet numbers in accordance with the Rules).

The work in Steps 1 to 4 are done according to the principles enumerated in Sec 6.2 and 6.3 respectively. This work amounts to the analysis of the subject into facets and the determination of the sequence of the facets. This is done in the idea plane. The work in Step 7 amounts to synthesis of the facet numbers in the notational plane. The work in the idea plane deals in reality with the Syntax of the Facets; and this is reflected in the notational plane also.

7.2 Analytico-Synthetic Classification

By 'Analytico-Synthetic Classification' is meant a scheme for classification involving analysis and transformation in the idea plane and synthesis in the notational plane according to stated postulates and principles.

7.3 Advantage of Postulates and Principles

Practical classification based on the stated postulates and principles bypasses the work of thinking about all the subjects at one and the same time, analysing each of them in a helpful sequence in such a way that, in the finally resulting sequence of subjects, the intended

Immediate-Neighbourhood-Relation remains invariant. The *ad hoc* decision of these for each document is tantalising. Nightmare is often the result. For example, the number of the possible sequences of the facets in the diverse subjects, out of which one and only one is to be chosen consistently, is very large.

7.4 Example of Tantalisation

Let me illustrate with a small fringe of the problem. Consider the subject

“Agriculture of Wheat”

It has only two facets — Basic Facet “Agriculture” and Isolate Facet “Wheat”. The sequence of these facets does not give much trouble. Maintaining deeper consistency in such cases is quite easy. But, consider the micro-subject “Leaf virus of wheat and spraying of chemicals from aeroplane in the coastal areas of Florida during the wet summer months in the present decade”. This has 14 facets. The following 11 facets are explicitly mentioned in the title — Leaf. Virus. Wheat. Spraying. Chemicals. Aeroplane. Coastal area. Florida. Wet. Summer. Present decade. The following three facets are implied in the title — Agriculture. Disease. Cure.

There are 87,178,297,200 possible sequences in which these can be arranged. How are we to select the most helpful of these eighty-seven thousand million sequences? How are we to be consistent in their selection — consistent also at a deeper level, at which the same pattern of sequence of facets is followed in diverse subjects? An *ad hoc* decision of this in each case is tantalising.

7.5 Bypassing the Tantalisation

But step 5 in practical classification — described in Sec 7.1, and based on Postulates and Principles — secures consistency without undue strain to the mind — without being tantalised. Further, when all the subjects are given their respective class numbers as shown in Step 7 in Sec 7.1 and are arranged by their class numbers, they all fall automatically into a helpful sequence. This is the advantage of using an Analytico-Synthetic Classification guided by Postulates and Principles.

7.6 Helpfulness for the Majority of Readers

The sequence secured by the postulates and principles is found to be helpful to the majority of readers. This has now been tried out not only in the arrangement of books and of their main entries in libraries, but also in the arrangement of the main entries in documentation lists of current articles in about a hundred very specialised subjects, such as Production Engineering of Screw, Production of Diesel Engine, and Nuclear Medicine. The concerned expert specialists have expressed satisfaction with the sequence secured by this classification.

7.7 Minority Groups Among Readers

No minority group, however, should be left without help. The formation of special collections in the stack room is one form of help to a single minority in a particular library. But to meet the needs of several minorities, the catalogue can be pressed into service. A separate guide-card giving the interest of each minority may be inserted in the

alphabetical part of the catalogue. Behind the guide-card for a particular minority may be inserted a duplicate set of the main entry cards of all the documents of interest to that minority. These duplicate main entry cards should be arranged in the classified sequence. This is only an aside. We shall now resume our main line of thought.

8 SEARCH FOR THE HIDDEN ROOTS OF CLASSIFICATION

8.1 Syntax of Facets

My wish is that in our search for hidden roots of classification, we should not stop with the postulates and principles at the near-seminal level. These form only a first approximation. What pleased me most was the announcement that this Seminar intended to dive deeper still, to find the very tip of the tap-root in the intellect, so to speak. In analytico-synthetic classification, one of the vital steps is the determination of the helpful sequence of the facets of a subject. This means the Syntax of Facets. This in its turn means a harmonious sequence of the facets, that gives satisfaction to the human mind. Webster quotes the following passage to show the power of a good syntax:

“His mind moved in a rich erudite and complex *Syntax*
That turned all opposition into admiration”.

The syntax of facets in classification based on the postulates and principles gives satisfaction to the majority of readers.

8.2 Quest for the Reason for Satisfaction

What is the reason for the Syntax of Facets given by the postulates and principles being satisfactory to most people? The reason should be searched in the minds of the readers. This will lead us to the roots of classification hidden far deep in the intellect-in-action. In this behalf, I made a suggestion in my talk at the Washington Conference (1958) on "Retrieval of Scientific Information, that to help in the establishment of a fairly long-lived helpful scheme for classification, a team of epistemologists, psychologists, linguists, reference librarians, classificationists, and statisticians should investigate the way in which the human intellect works today – that is, the Syntax of Facets that will give the greatest satisfaction to the greatest number of readers. Probably, the time was not then ripe for it. For, I found scepticism in some. If I remember right, one even called it a "fall-order". I am glad that the theme has been seized in 1966 by the organisers of the Symposium on the Foundation of Syntactic Relations in Classification in the University of Maryland.

8.3 Absolute Syntax

By 'Absolute Syntax' is meant the sequence in which the facet ideas of a subject arrange themselves in the minds of the majority of persons. Linguistic Syntax is the Syntax of Words - that is, the sequence in which the words stand arranged in a sentence or in the name of a subject in a natural language. The Linguistic Syntax may vary with the language; often it does.

The result in Step 5 in the steps in classifying a document corresponds to Absolute Syntax – that is, Syntax of Facets. In it, the

Hidden Roots of Classification

kernel terms in the name of a subject stand rearranged according to the Syntactic Principles governing the sequence of the facets denoted by the respective kernel terms. On the other hand, the result in Step 3 corresponds to Linguistic Syntax. In it, the kernel terms stand arranged in the same sequence as the one in which these terms are found in the name of the subject in the natural language used. Here is an example.

Consider the Subject

“The Heart of the Frog”.

This is Step 0. the full title in Step 1 will be

“The Heart of the Frog (as studied in) Zoology”.

In Step 2, where only the kernel terms are retained, we shall have

“Heart. Frog. Zoology”.

This sequence is according to the Linguistic Syntax of the English language.

According to the Linguistic Syntax of the Tamil language, the kernel terms will stand arranged in Step 2 as follows: “

Zoology. Frog. Heart”.

In step 5, the kernel terms will stand arranged as:

“Zoology. Frog. Heart”.

This represents the Syntax of Facets.

It happens that the Linguistic Syntax of the English language differs from the Syntax of Facets; while the Linguistic Syntax of the Tamil language agrees with the Syntax of Facets.

There may be languages in which the Linguistic Syntax may give respectively

“Zoology. Heart. Frog
Frog. Heart. Zoology
Frog. Zoology. Heart
Heart. Zoology. Frog” – respectively.

The number of variations of Linguistic Syntax from the Syntax of Facets will increase with the number of the kernel terms in the name of the subject – which is the same as the number of the facets in it.

8.4 Absolute Syntax and Facet Syntax

In general, the number of Linguistic Syntaxes for the name of a subject, in the different natural languages all taken together, can become as great as factorial n , where n is the number of kernel terms in the name of the subject. But, there is only one Syntax of Facets for the subject. For this reason, it is conjectured that the Syntax of Facets is the same as the Absolute Syntax. This implies that the Absolute Syntax is the one conforming to the Postulates and Principles guiding the design of an Analytico-Synthetic Classification.

8.5 Problem for Investigation

The problem for investigation is, “Is there an Absolute Syntax governing the sequence of the facets of a subject, inherent in the human

intellect-in-action as it is today?" This investigation should be made by a team of specialists in Epistemology, Psychology, Linguistics, Reference Service, Design of Classification, and Statistical Analysis, mentioned in Sec 8.2.

8.6 Removal of Encrustation

It is not expected that the Absolute Syntax will be inherent in the minds of one and all the adults, without any exception. For, from childhood onwards the Linguistic Syntax of the mother tongue makes an incessant impact on the mind of a person. It is too much to expect that the encrustation formed by this incessant impact would not have become too hard and opaque for the inherent Absolute Syntax to become operative. On the other hand, I only expect the Absolute Syntax will be operative with the majority of persons.

8.7 Frequency Study

While investigating the problem, the Team of Specialists would use the same assortment of subjects. They should examine a reliable random sample of persons drawn from each of most of the natural languages. Probably, it will be helpful to have three sets of random samples – one for children, one for adolescents, and one for adults. The investigation may have to be done in five stages.

In the first instance, without any suggestion whatever being made by the Team, the sequence of facets naturally preferred in the various facets by the different persons should be found out and recorded.

Thereafter, an attempt should be made to break gently the encrustation of the Linguistic Syntax in the minds of the people. The degree of success in this work will throw the people into several groups.

At the third stage, the sequence of facets in the subjects, preferred by each of the groups, should then be found out and recorded.

The work in the fourth stage will fall largely to the share of the Statisticians. They will have to construct the necessary and possible frequency tables and curves, and the correlation tables and curves and surfaces; and they should also furnish all the necessary statistical constants emerging from the study of the problem.

In the fifth stage, it may be possible to find out, from the results tabulated by the Statisticians, whether there is an Absolute Syntax, and if so, what it is; and in particular, it can be verified whether the Syntax of Facets based on the Postulates and Principles for an Analytico-Synthetic Classification is the same as or at least a good approximation to the Absolute Syntax.

8.8 Help to Classificationist

The final findings of such a team of specialists will enable the classificationists to build schemes for classification on fairly firm foundations. At present, a good deal of professional energy and time get dissipated in discussing problems in classification guided solely by conjectures and conflicting opinions based on insufficient data. This wastage should be avoided. I wish that the deliberations of this Seminar leads to the solution of this problem along objective and scientific lines by Teams of Specialists, probing into the hidden roots of classification. The results will be of help in the study of absolute linguistics also. This piece of tiny research deserves to be provided for by a Foundation.

Chapter 2

ABSOLUTE SYNTAX AND STRUCTURE OF AN INDEXING AND SWITCHING LANGUAGE

A. NEELAMEGHAN

Switching from one information system to another would be convenient if the information languages — that is, the method of representation of subjects and other information content of discourse used in the systems are syntactically consistent, compatible with each other, and interconvertible at a reasonable cost. In this connection, the development of an intermediate language through which the switching from one information language to another is an important consideration. An idea is a pattern, a gestalt, a form, a structure that one perceives. A subject of a discourse of an information source or of a user's query is a combination of ideas, that is, of structures; therefore, the structure of a subject-representation that is, of a subject surrogate has a bearing on the user's 'perception' of the subject represented. Some characteristic features of an information structure helpful to users, the problems of transformation of information structures, the linear structuring of subject surrogates, and some criteria for the choice of a 'standard format' or framework or model for such structuring are considered. Absolute syntax is defined as the sequence of the component ideas in a subject helpful and acceptable to a majority of users. The helpfulness of structuring of subject parallel to the absolute syntax is indicated, together with supporting information based on postulations and research on deep structure of languages (Chomsky, Fodor, Katz, Fillmore, Birbbaum, and others), biocybernetics (Lazlo), syntax of knowledge (Meredith), common structure in preserving messages in a set of transformations (Rosenbleuth), etc. The generalised facet structure (model) of subject representation obtained on the basis of the general theory of classification and the guiding principles for helpful sequence formulated thereof (Ranganathan and the Bangalore School) is found to be helpful and acceptable to a large number of users of information systems, and therefore, conjectured to parallel the absolute syntax. Work done in this regard and in the development of specific schemes for classification and for the formulation of subject headings in different languages within the general framework (model), is mentioned.

1 TERMINOLOGY

The following are the operational definitions of some of the technical terms used in this paper[1]:

Idea: An idea is a product of thinking, imagining, etc., got by the intellect, by integrating with the aid of logic, a selection from the apperception mass, and/or what is directly apprehended by intuition, and deposited in memory. Alternative term: Concept.

Entity: An entity is any existent, concrete or conceptual that is, a thing or an idea.

Discourse: A discourse is an expression of ideas, especially systematic or orderly expression in speech or writing.

Subject: A subject is an organized or systematized account of a body of ideas, whose extension and intension are likely to fall coherently and comfortably within the intellectual competence and the field of inevitable specialization of a normal person.

Example: A systematized account of "Conduction of heat" is a subject; and so may be deemed a systematized account of "Thermodynamics", and of the ideas in "Physics". But, not all the discourses embodied in the McGraw-Hill *Encyclopedia of Science and Technology*, taken as a whole can be deemed to be a subject; for, the totality of the subjects embodied therein cannot form an inevitable and convenient field of specialization of a normal person.

2 STRUCTURE AND PATTERN

The kind of pattern one perceives in a representation of an entity lies in the perceived *structure*. For instance, in a pictorial representation, the idea of “triangle” can be conveyed by 300 dots, or 30 dots, or 3 dots, as shown in Fig 1.

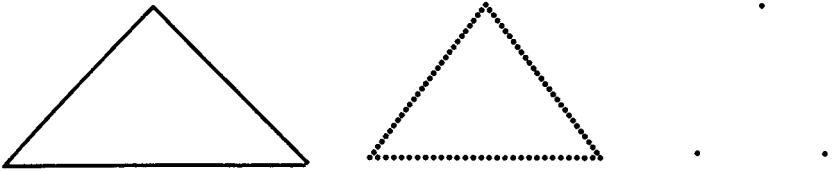


Fig. 1

Of these three representations, the last mentioned is deemed to be the most efficient, because it uses the fewest number of elements to convey the same amount of information as those using more number of elements. This indicates the important role of structure of a representation in relation to the perception of its “meaning”.

Structure is the way in which the components of an entity are put together. Lancelot Whyte defines the concept of structure as “effective patterns of relationships in any situation”[2]. In a general sense, structure denotes logical form. The content of a logical form may be physical, musical, psychological, temporal, or in some other way non-physical. Anything that has structure has parts, properties or aspects, which are in some manner related to each other. Thus, in every structure one can distinguish the relations and the items which are

related. The items may be qualities, values, or any conceptually distinguishable feature called elements of the structure[3].

An idea is a pattern, a structure, a gestalt, a form, a kind of picture that one perceives. A subject is constituted out of a combination of ideas that is, a combination of patterns. In understanding a complex structure, the human intellect finds it helpful to identify the substructures and categorize them. Such pattern recognition, pattern formulation, and categorization have been found to be involved in the human learning process and information handling[4]. Anderson and Bower point out that the representation of knowledge-structures is an important problem in cognitive psychology:

“..what are the primitive symbols or concepts, how are they related, how are they to be concatenated and constructed into larger knowledge-structures, and how is this ‘information file’ to be accessed, searched and utilized. The choice of a representation is central, since how one handles this issue causes widespread effects throughout the remainder of his theoretical efforts. As computer scientists working on problem solving have known for years, a good structural representation of the problem already constitutes half of its solution[5]”.

Therefore, the structure of *representation of subject* that is, surrogate of subject has a crucial role in conveying information about the subject denoted.

3 REPRESENTATION OF SUBJECT

An information system handles discourses. A discourse may be verbal as expressed in a query of a user of an information system. It may be in a recorded form as in a conventional document such as, a

book, an article in a periodical, and a technical report or on magnetic tape, film, etc., all of which may form information sources. Finding information and/or documents containing information co-extensively matching the subject of a user's query may depend, in a good measure, on the capacity of the system to identify and specify coextensively the subjects of discourses that is, subjects embodied in queries and those embodied in information sources. The representation of subjects expounded in discourses for example, subject headings, class numbers, data structures, algorithms or other kinds or surrogates may provide the first point of entry into an information retrieval system. An information system of this sort may form a node or component of a hierarchy of increasingly larger network of local, national, regional and global information systems. To facilitate the integration and collaborative functioning of the information systems developed in different contexts, it would be helpful if the "languages" used for representation of subjects in the different systems are syntactically consistent, compatible with each other, and inter-convertible at a reasonable cost. Thus, the representation of subject of discourses in the form of surrogates is central to the designing of information files for information storage and retrieval purposes.

4 PROBLEMS IN THE EFFICIENT USE OF THE LANGUAGE OF SURROGATE

The following are some of the factors which raise problems for the user of an information system in the efficient use of the language of the surrogate system.

4.1 For various reasons, it may be difficult for the user to perceive precisely and express coextensively the subject of his interest at the moment.

Therefore, the total semantic domain represented by the expression of the subject of his query may not be coextensive with the semantic domain of the subject of his interest as perceived by him.

4.2 The information scientist's perception of the semantic domain of the subject of interest of the user derived on the basis of the latter's expression of his interest at the moment, may not be coextensive with what the user purported to convey.

4.3 A user may not and, perhaps, cannot be concentrating attention or work on at one and the same time on all the component ideas potentially falling in the subject of his interest, even if it be a narrow one. He may study and concentrate attention only one or a few of the component ideas at a time. The recall value at the moment that is, the likelihood of being retained and recalled from memory for this component idea (or a few of them) would be relatively greater than that for the other component ideas in the subject of his interest. Therefore, he is more likely to bring up the name of this component idea (or of a few of them) in searching for information on the subject of his interest at the moment.

4.4 The information scientists' knowledge and understanding of the subject of interest to the user may be inadequate.

4.5 For various reasons, it would be difficult for the information scientist to perceive precisely and express coextensively the semantic domain of the subjects embodied in information sources. Therefore, the surrogate system prepared by him for representing the semantic domain of subjects embodied in information sources may not be coextensive with the semantic domain of the subject(s) purported to be described by the author of the work.

Fig. 2 illustrates the non-congruence of the different semantic domains.

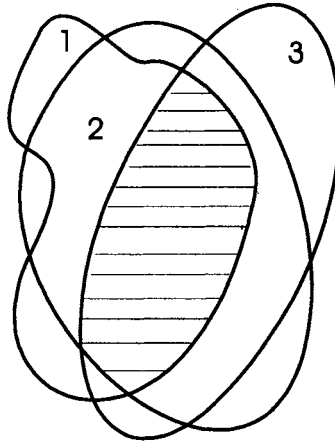


Fig. 2

- 1 User's perception of the semantic domain of his subject interest-at-the moment
 - 2 Actual semantic domain of the subject as expressed by user
 - 3 Librarian/Information Scientist's perception of the semantic domain of user's subject interest-at-the-moment
- ⊗ Common semantic domain

5 MINIMIZING THE CONSTRAINTS

5.1 Helpful Features of an Information System

Some of the features of an information system that may help in minimizing some of the constraints and difficulties mentioned in Section 4 and its subdivisions, are as follows:

5.1.1 Providing access to information on the subject of interest to the user by the name of the component idea(s) he may bring up in using the surrogate system.

5.1.2 Providing facility for browsing and selection of information to compensate for the dissimilarity and non-coextensiveness between the subject perceived and expressed by the user and the at perceived and understood by the information scientist at the time of query negotiation or user-system interfacing. This may involve providing access to:

- 1 Subjects greater in extension but subsuming the subject of the user's interest at the moment.
- 2 Subjects greater in intension but containing a substantial portion of it devoted to the specific subject of interest to the user at the moment.
- 3 Subjects in some other manner related to the specific subject of interest to the user at the moment (for example, collateral subjects, analogous subjects, and subjects studied in mutual relation to the subject of interest to the user).

5.2 Intersystem Connection and Compatibility

In order to facilitate switching over or movement from one information system to another with the longrange goal of establishing system interconnection on a global scale, there are at least two approaches. These are:

- 1 To use the same or very nearly the same information storage and retrieval language in all the information systems; and
- 2 To use an intermediate language or switching language through or by which one moves from one information system to another.

There can also be partial combination of 1 and 2.

6 FRAMEWORK FOR REPRESENTATION

6.1 Problems of Transformation

Of the two methods mentioned in the preceding section, the second one is the more practicable at present stage in the development of information systems throughout the world. However, in either of the methods, an important consideration relates to the framework elements, relations, and structure to be used for the analysis and representation of subjects of discourses that is, subjects embodied in information sources and in users' queries.

This paper mentions some of the suggestions about a common knowledge-structure and framework for representation of subjects and discusses one such framework.

As mentioned in Section 3, there are various methods of representing subjects, such as, class numbers, subject headings or strings of words, multi-dimensional arrays, tree-structures, etc. These arise from the process of analysis of subjects of discourses into constituent elements; recognition of the relevant relations among the elements as they obtain in the context of the subject concerned; and assembling the elements in a preferred pattern so as to represent as coextensively as possible the subjects. (See Fig. 3). Representation

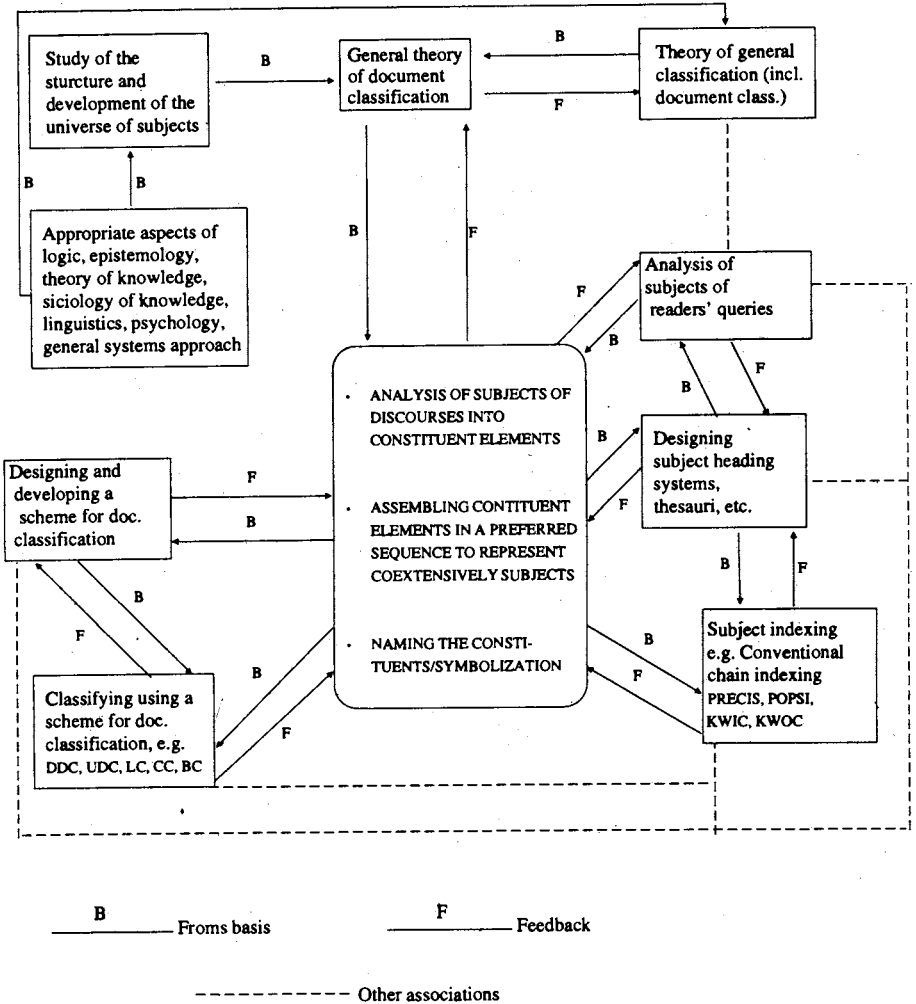


Fig. 3
Interrelation between discourse/subject analysis, classification, subject indexing, subject heading work, thesauri, etc.

of subject by a subject heading or a class number is equivalent to transforming the n -dimensional configuration of the subject into a linear configuration[6]. An arrangement of the component elements in each subject falling in a subject-field among themselves, in a sequence helpful to a majority of users requires keeping invariant every Immediate-Neighbourhood relation among all the subjects while transforming or mapping the n -dimensional configuration of subjects on to a line.

The number of subjects falling even in one subject-field is quite large and continues to increase rapidly such that it is difficult to arrange them in a helpful sequence consistently without the aid of guiding principles. In the transformation, only one of the many Immediate-Neighbourhood relations can be kept invariant. Determining which should this be, and which components should come respectively as remove 2, remove 3, etc., with respect to a reference component is a difficult decision. To depend, for this purpose, on the conjecture of different classificationists as to what is helpful to a majority of users may not yield a consistent pattern of arrangement of components of all subjects. But, such a consistency in pattern is helpful and necessary to the users, as well as the designers of the information system.

6.2 Criteria for Choice of Frame work

Meredith[7] summarizes the problem of transfer and transformation of knowledge-structures, as follows.

“The long-term task is not merely to analyse the problems but to design methodological instruments for carrying out practical researches into problems of communication. These may be treated as problems of mapping. Given an original territory of

factual phenomena how does this territory become mapped in the brain of the investigator? How does this map become transformed into a verbal or symbolic expression a linguistic map? How is this map transformed again into a language adopted to the needs of the ultimate recipient, the learner? Finally, how is this third map introjected into the learner's brain to form a pattern of knowledge? If we can establish a cartography for these maps, we can formulate "projective equations" leading from one map to the next. Each map will be a pattern of definable variables under appropriate controls and observing changes in the next map, the equations can be solved and laws of projection may be discovered."

Anderson and Bower[8] have listed the following considerations deemed helpful in the choice of a "standard format" (structure) for representing information:

- 1 The representation should be capable of expressing any conception which a human can formulate or understand.
- 2 The representation should allow for relatively efficient search for and retrieval of information. That is, specific information should remain relatively accessible even when the data-files grow to encyclopedic proportions.
- 3 The representation should saliently exhibit the substantive information extracted from a given input. It should not be influenced by the peculiarities of the particular natural language in which that information was communicated. This hope for language-invariance amounts to a wish for a universal *interlingua* in which any conception in any language could be expressed, but for which the format would not be specific to a particular language...

- 4 For reasons of parsimony, the representation should involve a minimum of formal categories. That is, it should make a minimum of *formal* (structural or syntactic) distinctions at the outset; more complex distinctions would be built up by the construction rules for concatenating primitive ideas.
- 5 The representation must allow for easy expression of concatenation operations, by which “duplex ideas” can be constructed out of “simple ideas”. This means, for example, that the representation should allow easy expression of conceptual hierarchies, or multiply embedded predications, or allow one to predicate new information on any old information-structure”.

7 ABSOLUTE SYNTAX

7.1 A Postulate

At the International Conference on Scientific Information (Washington DC) (1958), S. R. Ranganathan suggested that “to help in the establishment of a fairly long-lived helpful scheme for classification, a team of epistemologists, psychologists, linguists, reference librarians, classificationists and statisticians should investigate the way in which the human mind thinks that is, the Syntax of Facets that will give the greatest satisfaction to the greatest number of readers[9]”. In 1966, in his valedictory address to the Maryland Symposium on Relational Factors in Classification, Ranganathan postulated such a syntax of facets and named it as Absolute Syntax[10]. Absolute syntax in the sequence in which the component ideas of subjects falling in a subject-field arrange themselves in the minds of a majority of normal intellectuals, for instance when they think and communicate about the subject.

Ideas are largely products of intellection. Intellectual activity is known to be controlled by brain. There is considerable similarity in the structure and, therefore, in the functioning of the brain in a majority of normal human beings. Thus, a majority of normal human beings have more or less a similar mode of thinking and learning that is, in forming ideas and in combining them to build knowledge-structures. It is further stated that *biologically* man has not changed to any appreciable extent since the emergence of *Homo sapiens*; for, the structure of the genetic material has not appreciably changed since then that is, for some 500,000 years although we have changed *culturally*[11]. Therefore, the probability of a sudden change that is, a mutation in the mode of thinking and learning of a majority of normal persons in the immediate future is quite low. Hence, if the syntax of the representation of the component ideas of subjects is made to conform to, or parallel to, the Absolute Syntax, then the pattern of linking of the component ideas that is, the resulting knowledge structure is likely to be:

- 1 Helpful to majority of normal intellectuals;
- 2 Consistent in pattern in subjects falling in different subject-fields;
- 3 Relatively more stable and continue to be helpful to a majority of normal intellectuals so long as there is no mutation in their mode of thinking;
- 4 Free from the aberrations due to variations in linguistic syntax from the use of the verbal plane in naming subjects;
- 5 Capable of representing and indication of subjects co-extensively with a minimum number of variety of component elements;

- 6 Helpful in recognizing the less explored and unexplored regions in the universe of ideas; and
- 7 Helpful in probing deeper into the pattern of human thinking and modes of combination of ideas.

7.2 Analogy from Search for Linguistic Universals

In an earlier paper[12], it was pointed out that the formulation of a generic framework for structuring subjects has a parallel in the search for universal linguistic forms such as that expounded in the works of Chomsky, Fodor, Katz, and the generative grammarians. Birnbaum[13] suggests a multi-layered syntactic structure between the deepest of the deep structures and the surface structure. He points out: "As a result of the general trend toward a generative semantic framework, a new slightly modified model of generative grammar seems now to be taking shape. This model can be thought of as comprising three independent components:

- 1 A Semantic Component which will define the relations obtaining between semantic (including categorical) units or, rather hierarchically ordered clusters of semantic features (such as, (THING), (CONCRETE), (COUNTABLE), (ANIMATE), (HUMAN), (PERSONAL), (MALE), (ADULT); (PREDICATION), (AGENT), (DEFINITE), (ACTION), (PATIENT – ORIENTED), (TIME-DETERMINED), (ASPECT – DETERMINED), ETC.,
- 2 A Transformational Component which will convert the semantic deep structure representations into surface structure representations.
- 3 A Phonological (or Symbolization) Component.."

Fillmore[14] points out that “there may also be some psychological reasons that argue for the use of predication as a data-base language in a model of memory... Perhaps ‘thinking’ represents operations at the level of the semantic base structure, before it has been transformed into actual sentences through the application of syntactic rules”. The case categories suggested by Fillmore include the following:

“*Agentive* (A), the case of the typically animate perceived instigator of the action identified by the verb.

Instrumental (I), the case of the inanimate force or object causally involved in the action or state identified by the verb.

Dative (D), the case of the animate being affected by the state or action identified by the verb.

Factitive (F), the case of the object or being resulting from the action or state identified by the verb, or understood as a part of the meaning of the verb.

Locative (L), the case which identifies the location of spatial orientation of the state or action identified by the verb.

Objective (O), the semantically most neutral case, the case of anything representable by a noun whose role in the action or state identified by the verb is identified by the semantic interpretation of the verb itself..”

Vleduts and Stokolova also propose structures — standard phrases [15] at different levels for subject - representation in different disciplines.

Leibniz’s ideal language[16] and the Whorfian hypothesis[17] that “Every language contains terms that have come to attain cosmic scope of reference that crystallize in themselves the postulations of an

unformulated philosophy.. such are our words 'reality, substance, matter' and .. 'space. time, past, present, future", are worth noting here.

7.3 Biocybernetic View

In his book on Systems Philosophy, Ervin Lazlo[18] mentions about "basic modes of thinking".

"..It is also becoming evident that all men, regardless of the culture they happen to belong to, have basically similar nervous systems, are equipped with analogous sense receptors, command like patterns of response, and use patterns of thought (whether rationally or emotively motivated) which obey very similar laws or regularities. In other words, there appear to be some "universal" traits underlying cultural cognitive relativities: Chomsky could locate "linguistic universals" and Kluckhohn discovered a number of" universal categories of culture

"Finding such universals is rendered difficult if not impossible, by arguing out of one's own culturally or individually relativistic categories. In that light, every other world-model becomes but a special case of one's own, and is forced into the latter's structural scheme. But, in using the neutral frame work of a cybernetic mode, one is no more arguing out of his own culture-categories than out of that of a thermostat. Conceptualizing the cognitive process with such categories, we can reach universal structures, for we are not dealing with particular contents. Regardless of whether a person conceives a sensory pattern as *trees*, meaning "standing peoples, in whom winged ones built their lodges and reared their families" or interprets (presumably) the very same pattern as obstructions to be cut down and burnt; he is using a construct (or gestalt) which endows his perceptual input with meaning.

And the development of constructs and gestalts obeys some general regularities, already manifest in biological evolution and set forth in cultural development".

Lazlo further points out

"..Regardless of the genetically and empirically induced differences, however, basic modes of thinking characterize all human beings, and indeed all higher biological species. These are rooted in, and explained by, the fact that all such organisms are self-maintaining open systems using a specific mode of reproduction, and forming part of some similarly specific social structure. The mental capacities needed to maintain such systems in their environments are adaptive functions; they crystallize as cognition in the more evolved species, and culminate in man.

"..The most immediately pertinent to human cognition make up an ascending ordering of categories, universally human in principle but variously evolved in different real individuals. These categories may be listed as follows:

- (1) *Gestalt* (invariant patterns with established meanings to which the input patterns are assimilated);
- (2) Rational constructs (theoretic entities postulated through abstract reasoning and connected to the input patterns by means of some established rule of correspondence); and
- (3) Aesthetic construct (non-discursive meanings discovered in the input and illuminating some part of the knower's "felt experience"). These are the types of constructs which represent the limits of human cognition, given the kind of perceptions, cognitive organizations and effective output channels at our disposal. I argue that many forms of human experience do not

constitute disjunctive culture-conditioned categories, but a set of universal structures which transcends individual and cultural differences and relativities, and accommodates as subclasses, the many varieties of cognitive patterns as environment mappings and constructions of natural cognitive systems on the specially human level of nature's hierarchy"

7.4 Syntax of Knowledge and Epistemics

Meredith[19] suggests the existence of a "syntax of knowledge". The argument runs as follows:

"At a multi-lingual conference.. with a community of disciplines, experience and thought, the translators have no difficulty in transforming, virtually instantaneously, the most elaborate syntactic forms of one language into the quite different forms of another whilst reserving the essential structure of information and conceptualization in the speech. Thus, there is a 'syntax of knowledge' which, even if not entirely independent of the particular languages, can and does, in practice, follow its own course alongside the syntactic sequence of language. It may serve to sharpen the difference if, provisionally, we think of the latter as governed by temporal relations (by the sequence of words in the sentence) and the 'syntax of knowledge' as primarily a spatial structure only shredded into temporal filaments in order to conform to the sequential character of speech.

"This is a big step forward. Even though the syntax of language cannot be entirely divorced from the syntax of knowledge, we can pragmatically separate them by treating the one as a temporal sequence and the other as a spatial pattern. But, it may be objected, what about the temporal character of knowledge itself? Our knowledge of history, our understanding of sequential operations, of industrial processes, of astronomical events etc., all of which involve time.

Two points may be noted here: (1) Even though in a narrative the sequence of paragraphs normally (though by no means in every case) follows the time-sequence of the events narrated, this correspondence scarcely holds at all within the limits of a single sentence. And what is called linguistic syntax is largely based on the analysis of the single sentence. 'The assassin shot the President at the end of his speech'. In the actual *event*, the speech came before the shot; in the *sentence* after it. Thus, 'epistemic time' and 'linguistic time' are partially independent. (2) We speak a sentence sequentially, that is, at the beginning we have not yet spoken the end but what we are talking about even though it may be temporal event, is known to us throughout .. 'Epistemic time' is in fact 'dead' time, the completed past history, fossilised, and hence not "time" at all in the linguistic sense. It has a discernible sequence but no flow. Our knowledge of it is a geometric knowledge of evidence spread out in space or held in memory".

Constance Amsden[20] commenting on Vygotsky's ideas on "inner speech", also suggests a "syntax of thought".

7.5 Common Structure

Arturo Rosenbleuth postulates a "common structure" in preserving the message received through a set of transformations[21]:

"When a person hears a symphony, the messages sent by the orchestra reach the listener as air vibrations. These vibrations stimulate mechanically the receptors of the organ of corti, and these receptors set up nerve impulses along the fibres of the VIIIth nerve. It is clear that at this stage the physical events that are taking place are of an entirely different kind from those occurring in the instruments of the orchestra. Yet the message is preserved because there are similarities in certain features of the two series of events — sounds emitted by the orchestra

and nerve impulses travelling over the auditory nerves. The existence of these similarities of relations is precisely what is called a common structure. The mental decoding, which is the perception of the symphony, again preserves the corresponding relations. A common structure thus implies the quantitative preservation of the relations that exist between the independent constituents of an event or message through a set of transformations”.

7.6 Logic of Exposition and Linguistic Syntax

Rosenbleuth also comments on syntax of thought and linguistic syntax thus:

“As a further example of the fundamental difference between the mental events and the correlated neuro-physiological processes, let us consider the processes that would develop in my brain if I presented verbally a specific relatively complex, argument on three different occasions in Spanish, English, and French, respectively. Although the neurophysiological correlates corresponding to the logic of my exposition might be similar or identical in the three cases, clearly those corresponding to the selection of words and their syntactical organization, a very important aspect of the presentation of the argument would be absolutely dissimilar. If I should want to use dictionaries to translate from the language of the introspective data to that of the physical processes, I would need in this instance three different dictionaries, and more, if I were capable of using fluently other languages”

7.7 Concept and Conception

Suzanne Langer points out that the psychological context of our thoughts may be private and personal. Therefore, two persons talking about the same thing may perceive it in different ways [22]. They are

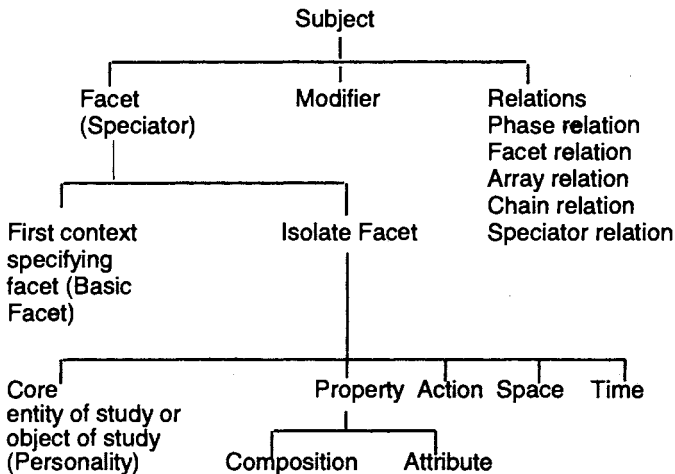
then said to have different *conceptions*. But, if they understand each other, then their respective conceptions embody the same *concept*. A concept is an abstracted form. Abstraction is the consideration of logical form (structure) apart from content [23].

8 GENERALIZED FACET STRUCTURE FOR SUBJECTS

Analysis into constituent ideas and structuring of several thousands of subjects in a variety of subject fields for the purpose of designing and developing of schemes for subject classification, preparation of feature headings and subject headings, and for indexing of subjects have helped in:

- 1 Categorizing the constituent elements in a subject into three types: Facet, Modifier (speciator), and Relations.
- 2 Sub-categorizing each of the three types of constituent elements into a few kinds.

For Example:



Absolute Syntax

- 3 Developing a typology of Basic Subjects, the modes of formation of Basic Subjects, and the arrangement of Basic Subjects[24].
- 4 Developing a typology of Modifiers (speciators) for basic facet and for isolate facet in different subject-fields.
- 5 Recognizing the relative strength of bond (relation) between the first context-specifying element (basic facet) and other types of facets in subjects[25].
- 6 Formulating principles for helpful sequence among
 - (a) Facets of a subject
 - (b) Speciators to a facet
 - (c) Compound subjects falling in a particular subject-field
 - (d) Subjects falling in different subject-fields (26).
- 7 Developing a Generalized Facet structure (Model) of subject, with specific models for different subject-fields.

It is not possible to discuss in detail these developments in this paper. The main developments are briefly outlined in a recent FID/CR report [27]. A condensed version of the Generalized Facet Structure is given in the Appendix.

Subject structuring obtained using the Generalized Facet Structure has been found to give a co-extensive representation of subjects and arrangement of subjects helpful to majority of users.

The chart in Fig. 3 shows the interrelation between subject-structuring, designing a classification scheme, generation of subject

indexes, etc. Depth classification schemes for over a hundred subject-fields have been designed and several hundreds of articles, technical reports etc., have been classified using these schemes in each subject-field [28]. The structuring of subjects and the sequence in which the subjects get arranged have been found to be acceptable to a large number of users.

In a small-scale experiment, subject-headings each with several components, structured in the above manner, were presented to about a hundred persons for indication by them of the subject that each of them perceived in the structuring. Although there were more than one way of representing each subject in the natural language, the subject perceived was the same in all the cases (results unpublished). That is, there was no homonym. Large scale experiments with other types of structuring of subjects has been planned.

Translation of the subject heading terms into different natural languages did not give rise to any difficulty in interpreting the subject represented by persons knowing the language concerned[28].

The facility of rearrangement of given terms into the preferred sequence and synthesis of class number given the descriptors, using computer, have been demonstrated[30].

These experiences indicate that the structuring of subjects conforming to the model developed according to the General Theory of Classification:

Absolute Syntax

- (1) Helps to secure a facet syntax parallel to that of the absolute syntax;
- (2) Gives a “standard format” for representing information considered helpful by Anderson and Bower (See Section 6.2); and
- (3) Provides a frame work for an intermediate or linking language that is helpful and consistent.

Appendix

A condensed Version of the Generalized Facet Structure for Subject Representation.

$$\begin{aligned}
 &B - b_1 - b_2 \dots - b_n \quad C' - c'_1 - c'_2 - \dots - c'_n \\
 &C'' - c''_1 - c''_2 - \dots - c''_n \dots, \\
 &C^x - c^x_1 \dots - c^x_n; \quad M' - m'_1 - m'_2 - \dots - m'_n; \\
 &M'' - m''_1 - m''_2 - \dots - m''_n; \\
 &M^x - m^x_1 - m^x_2 - \dots - m^x_n; \quad P' - p'_1 - p'_2 \dots - p'_n; \\
 &P'' - p''_1 - p''_2 - \dots - p''_n \dots; \\
 &P^x - p^x_1 - p^x_2 - \dots - p^x_n; \quad A' - a'_1 - a'_2 - \dots - a'_n *; \\
 &A'' - a''_1 - a''_2 - \dots - a''_n *; \dots *; \\
 &A^x - a^x_1 - a^x_2 - \dots - a^x_n \quad S' - s'_1 - s'_2 - \dots - s'_n \\
 &S'' - s''_1 - s''_2 - \dots - s''_n \dots \\
 &S^x - s^x_1 - s^x_2 - \dots - s^x_n \dots \quad T' - t'_1 - t'_2 - \dots - t'_n \\
 &'T'' - t''_1 - t''_2 - \dots - t''_n \dots \\
 &T^x - t^x_1 - t^x_2 - \dots - t^x_n
 \end{aligned}$$

*Repetitive occurrence of isolate ideas denoting Matter Material, and Property in that sequence.

where

- B represents the First Context Specifying Element, that is, Basic Facet.
- $C' C'' \dots C^x$ represent respectively the object of Study (Personality), its Part of Remove 1, Part of Remove 2, ... Part of Remove x.
- $c' c'' \dots c^x$ represent respectively Speciators to $C' C'' \dots C^x$.
- $M' M'' \dots M^x$ represent respectively Matter Material isolate of successive removes.
- $m' m'' \dots m^x$ represent respectively Speciators to $M' M'' \dots M^x$.

Absolute Syntax

- $P' P'' \dots P^x$ represent respectively Property Isolate of successive removes.
- $p' p'' \dots p^x$ represent respectively Speciators to $p' p'' \dots p^x$
- $A' A'' \dots A^x$ represent respectively Action Isolate of successive removes
- $a' a'' \dots a^x$ represent respectively Speciators to $A' A'' \dots A^x$
- $S' S'' \dots S^x$ represent respectively Space Isolate of successive removes.
- $s' s'' \dots s^x$ represent respectively Speciators to $S' S'' \dots S^x$
- $T' T'' \dots T^x$ represent respectively Time Isolate of successive removes.
- $t' t'' \dots t^x$ represent respectively Speciators to $T' T'' \dots T^x$
- $, ; : \dots$ represent respectively Indicator Digits C, M and P, A, S, and T.
- (hyphen) represents the Indicator Digit for a Speciator.
- > represents a chain (subordinate) relation.

Example

1 Document: Large-scale production of roller-dried infant food, by M R Chandrasekhara and others. (Infant food from buffalo milk series, 6). (*Food science*. 9;1960;1).

Chemical technology, Food-Infant-from Milk-of Buffalo-

B , C' - c'_1 - c'_2 - c'_3

with Fat 15 percent-with Protein 23 percent-Pasteurized:

- c'_4 - c'_5 - c'_6

Fortification-with Vitamin:Powdering:Homogenisation:

: A' - a'_1 : A'' : A'''

Drying-with Roller

: A'''' - a''''

2 Document: Cold storage studies on dusehri mango by K K Singh and ots. (*Journal of research (Ludhiana)*). 5,N4;1967; 516-22)

Chemical technology, Food Mango-Dusehri;

B , C' - c'_1

Storage life influenced by Mango:Preservation-by Coating-

; P' (phase relation) C' : A' - a'_1

Antimicrobial wax:Packaging-in Polyethelene bag: storing-

- a''_1 : A'' - a''_1 : A'''

at 2° C-for 3 Weeks

- a'''_1 - a'''_2

3 Document: Dowding & Doll accuratool type 6D programme controlled automatic cycle turret lathe (*Machine production engineering* 116;1970:694-5).

Commodity production engineering, Lathe-Turret type-

B , C' - c'_1

-Chandler Ford 6D-Automatic-with Speed 32-Range

- c'_2 - c'_3 - c'_4

7.5 to 4000 rpm-with collet and turret distance 460 mm

- c'_5 - c'_6

with Bedway-of Steel-with Carriage control-by Air cylinder-

- c'_7 - c'_8 - c'_9 - c'_{10}

with Saddle movement maximum 330 mm-with Turret side

- c'_{11}

travel maximum 75 mm-with Attachment for drilling-with

- c'_{12} - c'_{13}

Program control

- c'_{14}

Absolute Syntax

4 Document: Working capital advances to limited companies, by S P Sen Gupta. (*State Bank of India Monthly Review*, 7; 1968;276-90).

Banking, State Bank of India; Credit-by Lending

B , C' ; P' - p_1'

-to Joint stock companies-for Working capital-

- p_2' - p_3'

against collateral security-by Hypothecation-of

- p_4' - p_5'

Moveable goods; Risk: Calculation-Debt equity ratio method

- p_6' ; p'' : A' - a_1'

5 Document: On placing the blame for primary produce instability, by R C Porter. (*International Economic Review*, 11; 1970;175-78).

International trade, Commodity; Price; Variation; Trend

B , C' ; P' ; P'' ; P'''

-in Same direction as volume *in relation to* Price; Variation

- p_1''' (phaserelation) ; p' ; p''

:Evaluation-Elasticity of demand point of view

: A' - a_1'

6 Document: Maximization policies of less-developed exporting countries, by V K Smith (*Quarterly Review of Economics and Business*. 9;1969;84-6).

International trade-Developing country, Export-

B - b_1 , C'

Primary commodity; Grain-Long term: Determination-

- c_1' ; p' - p_1' : A'

by Theory of producers surplus-Johnson's method: Evaluation

- a_1' - a_2' : A''

7 Document: Undulant fever as an occupational disease, by R Rozansky. (*Lancet*, 2; 1964;416).

Medicine-Industrial, Human body; Disease-caused by

B $-b$, C' ; P'

Bacteria>Eubacterials>Brucella melitensis-Fever symptom-

$-P'_1$ $-P'_2$

Intermittent

$-P'_3$

8 Document: New India 1885; British official policy and the emergence of the Indian National Congress, by B Martin. 1972.

History, India; Independence movement-Indian National Congress;

B , C' ; P' $-p_1$

Emergence *influenced by* United Kingdom; Policy-Colonial

; P'' (Phaserelation) , C' ; P' $-p'_1$

It is possible to specify further the type of speciator by using a combination of indicator digits. For example:

- ' for a Time denoting speciator (e.g. John who came earlier than Smith)
- . for a Space of Location denoting Speciator (e.g. Box on the Conveyor belt)
- : for an Action denoting Speciator (e.g. Plastics produced by moulding)
- ; for a Property denoting Speciator (e.g. Paper coloured white).

REFERENCES

1. Ranganathan, S.R. *Prolegomena to Library Classification*. 3rd ed, Asst. by M.A. Gopinath, 1967.
Webster's Third New International Dictionary of the English Language, Unabridged, 1966.
2. Whyte, L.L. *Accent on form*. 1954, B 28.
3. Langer, S.K. *Introduction to Symbolic Logic*. 3d ed, rev, 1967, Chapter 1.
4. Bruner, J.S. and others. *Study of Thinking*. 1956, p. 12
5. Anderson, J.R. and G.H. Bower. *Human Associative Memory*, 1973, p. 151.
6. Ranganathan, S.R. *Classification and Communication*. 1961, Sec 14632.
7. Meredith, P. *Instruments of Communication: An Essay in Scientific Writing*. 1966, p. 64.
8. Anderson, J.R. and G.H. Bower. *op cit*. p. 152.
9. Ranganathan, S.R. *Quoted in ref 10. Sec 7.2.*
10. Ranganathan, S.R. "Hidden Roots of Classification" *Information Storage and Retrieval*. Vol. 3, 1967, Sec.7.
11. Russel, B. *Has Man a Future*. 191, Penguin Special, S P Kilpatrick, W.H. "Philosophy of Education". 1951, *Quoted in* J. Park ed, *Selected Readings in the Philosophy of Education*. 3d ed. 198, p.107-117.
12. Neelameghan, A "Sequence of Component Ideas in a Subject", *Library Science with a Slant to Documentation*. Vol. 8, 1971, Paper Q.
13. Birnbaum, H *Problems of Typological and Genetic Linguistics Viewed in a Generative Framework*. 1970, p. 9-18, 58- 70.
14. Filmore, C.J "The Case for Case", *In* E. Bach and R.T. Harns (eds.), *Universals in Linguistic Theory*. 1969, p. 1-88.
15. Vleduts, G.E. and N.A. Stokolova. "About a Method of Constructing Information Languages having Grammar", Tr. from the Russian by Joe Lineweaver *DID/CR Report* 13. 1974.
16. Neelameghan, A. *op cit*. Sec 431.

17. Whorf, B.L. *Language, Thought and Reality*. 195, p. 61.
18. Lazlo, E. *Introduction to Systems Philosophy: Toward a New Paradigm of Contemporary Thought*. 1972, p. 200-204.
19. Meredith, P. *op cit.*. Chapter, Section 2.
20. Amsden, C.E. "Syntax – more Sinned Against than Sinning" In, *Some Problems in Reading*. Claremont, Calif. Claremont Graduate School. 193, p. 90-93.
21. Rosenbleuth, A. "Mind and Brain", *A Philosophy of Science*. 1970
22. Langer, S.K. *op cit.* Chapter 3.
23. Langer, S.K. *op cit.* Chapter 1.
24. Neelameghan, A. "Basic Subject", *Library Science with a Slant to Documentation*. Vol. 10, 1973, Papers F to N.
25. Ranganathan, S.R. "Colon Classification", *Rutgers Series on Systems for Intellectual Organisation of Information*, 1965, Chapter V.
26. Parkhi, R.S. *Library Classification: Evolution of a Dynamic Theory*. 1972, Chapter R.
Neelameghan, A. *Presentation of Ideas in Technical Writing*, 1975.
27. Gopinath, M.A. "Classification Research (India): 1968-1973", *FID/CR Report 14*. 1974.
28. Gopinath, M.A. *op cit.* p. 37-44.
29. *See ref 12.*
30. Neelameghan, A. and Venkataraman, S. "Use of Computer for the Synthesis of Class Number: A case study with a Freely-Faceted Version of Colon Classification", *Library Science with a Slant to Documentation*, Vol. 5, 1968, Paper S.

Chapter 3

DESIGN OF DEPTH CLASSIFICATION : METHODOLOGY

S.R. RANGANATHAN

Begins with a rapid account of the evolution of the practice and theory of classification during the last one hundred years and of concepts and the terminology associated therewith. Then follows an account of the successive removal, during the last forty years, of six inhibitions in the Idea Plane in the development of CC caused by the Notational Plane and of the inhibitions in the Notational Plane itself caused by the DC tradition. Shows the greater practicability of the method of First-Link-Downwards Approach than that of the Last-Link-Upwards Approach. Recommends the blending of the speculative method and the pragmatic method as a corrective of each other, at convenient stages, in designing a scheme for classification. After introducing a new and more easily applicable use of the terms 'Sector' and 'Zone' in the Notational Plane of CC, shows the number of true (IN) available in an array of order 1 to be 100 if the number of digits in an (AIN) is not to exceed 1; 400 if it is not to exceed 2; and 1,000 if it is not to exceed 3. Gives a scheme of allocation of the sectors to (WI), (\overline{WI}), and ($\overline{\overline{WI}}$). Lays down a procedure for determining the necessary first characteristics — that is (QI) — and determining their sequence with the aid of the Wall-Picture Principle in the design of classification. States that the Telescoping in Array made possible in this procedure satisfies the finding of the physiology of the eye and of the psychology of memory. Describes seven steps in the design of depth classification. Traces the progressive elimination of Gap Device in notation. Shows the training in the methods of applied research in the designing of classification to be essential for a documentalist as distinct from a generalist librarian.

ABBREVIATIONS USED

- | | |
|---|---|
| (ACI) = Anteriorising Common Isolate | (M C) = Main Class |
| (AIN) = Array Isolate Number | (MCN) = Main Class Number |
| (BC) = Basic Class | [P] = Personality Facet |
| (BCN) = Basic Class Number | [1P1] = Level 1 in Round 1 of [P] |
| CC = Colon Classification | [1P2] = Level 2 in Round 1 of [P] |
| (CC) = Canonical Class | [P3] = Level 3 of [P] |
| (CCN) = Canonical Class Number | (QI) = Quasi Isolate |
| (CdC) = Compound Class | (QIN) = Quasi Isolate Number |
| (CEI) = Common Energy Isolate | [S] = Space Facet |
| (CI) = Common Isolate | SC = Subject Classification |
| (CIN) = Common Isolate Number | (SI) = Space Isolate |
| (CMI) = Common Matter Isolate | (SID) = Super-Imposition Device |
| (CN) = Class Number | (SII) = Super-Imposed Isolate |
| (CPI) = Common Personality Isolate | (SpEI) = Special Energy Isolate |
| (CS) = Connecting Symbol | (Spl) = Special Isolate |
| (CxC) = Complex Class | (SpMI) = Special Matter Isolate |
| DC = Decimal Classification | (SpPI) = Special Personality Isolate |
| [E] = Energy Facet | [T] = Time Facet |
| EC = Expansive Classification | (TI) = Time Isolate |
| FC = Fundamental Category | (W) = Whole |
| IAN = Indo-Arabic Numerals | (<u>W</u>) = Organ of Remove 1 |
| I = Isolate(s) | (<u>W</u>) = Organ of Remove 2 |
| (IF) = Isolate Facet | (WI) = Whole Isolate |
| (IN) = Isolate Number | (<u>WI</u>) = Organ of Remove 1 Isolate |
| LC = Library of Congress Classification | (<u>WI</u>) = Organ of Remove 2 Isolate |
| [M] = Matter Facet | |

1 INTRODUCTION

1.0 Classification

The term 'Classification' is a homonym. It can denote any one of three ideas. We shall denote them by the respective terms:

Design of Depth Classification

1. Formation of groups — classification in sense 1 — Grouping;
2. Formation of groups and arranging them in a helpful sequence — classification in sense 2 — Arranging; and
3. Formation of groups — arrangement of the groups in a helpful sequence, and representing the groups by ordinal numbers denoting their respective positions in the sequence — classification in sense 3 — Library Classification.

1.1 Sense 1 - Grouping

In its first sense, classification means dividing the existents of the universe of discourse — concrete or conceptual, things or ideas into different groups [10]. Each group is to contain only like existents but no unlike existents. The terms 'Like' and 'Unlike' are in respect of a single attribute or a complex of attributes. Classification in this primitive sense should have originated with the primitive man. It is practised very early in childhood. It may be done on the basis of a single characteristic to begin with. But, with the evolution and the development of the cortex of the brain, the single characteristic has been giving place to a characteristic-complex; the complex itself has been steadily increasing in complexity. A single characteristic gives place to a Train of Characteristics. DC represents this stage. A single train of characteristics has now given place to a sequence of Trains of Characteristics. CC represents this stage.

1.2 Sense 2 – Arranging

In its second sense, classification is arranging the groups formed by the classification in sense 1 [22]. The groups can be arranged in one or other of several sequences. In fact the number of

possible sequences equals the number of permutations of the groups. One of these sequences has to be preferred in classification in sense 2. The sequence preferred should be one as helpful as possible for the purpose to be served by the arrangement. We shall call it Helpful Sequence. There is an inherent urge in man to classify in sense 2. Probably it is a neural necessity. Arrangement in a sequence is in effect a Linear Arrangement. But the universe of knowledge is a multi-dimensional one. Therefore, classification in sense 2 – that is, arranging virtually amounts to mapping a multi-dimensional space on a uni-dimensional space – that is, a line. To change the analogy, it amounts to transforming a multi-dimensional space into a uni-dimensional one. It has been found in cartography and mathematics that such a mapping or transformation cannot keep invariant all the interrelations of the entities of the original, multi-dimensional universe.

1.2.1 *Invariant*

The problem is what should be preferred to be kept invariant. A fundamental approach to the answer to this question has yet to be made. The work of successive classificationists implies some of them having preferred different *invariants*. There was not any substantial difference in the *invariants* preferred in DC, EC, and LC. But the *invariant* preferred by the SC was different. The essence of this difference centres round the Categorical Tables of SC. The invariants preferred by CC are still more different. In fact they are of a complex nature. The essence of this difference centres round its phase-analysis and facet-analysis with its basis of five Fundamental Categories and its Round and its Levels. However helpful and even necessary such a complex of invariants may be in mapping the universe of knowledge of today on a uni-dimensional space, the inertia and the conservatism of some librarians make them allergic to it. They

Design of Depth Classification

oppose it. They denounce it. But this cannot hold back the work of the forward – looking librarian.

1.2.2 Qualities of the Universe of Knowledge

A theoretical determination of the helpful complex of *invariants* to be aimed at in classification in sense 2 is made even more difficult by some of the qualities of the universe of knowledge.

1.2.2.1 Many dimensions

In the first place it is a multi-dimensional universe as already stated. By this, we mean that many parameters are needed to describe the position of any existent or any group of existents in that universe. As the universe of knowledge is developed by research, the number of parameters steadily increases.

1.2.2.2 Infinite

Secondly, the universe of knowledge is infinite. In other words, the number of items of knowledge – that is existents – in that universe is unlimited. The mass or the extension of an item of knowledge may be very small at one extreme and very large at the other.

1.2.2.3 Dynamic

Thirdly, this infinite universe is ever – growing. There are new developments in it from time to time. It throws forth new items of knowledge. Some of these new items could not be anticipated almost

till they actually appeared. We do not know what comes round the corner at any moment. It may be a massive item of knowledge comparable to what we call (BC); or it may be tinear than any item known hitherto. Indeed the universe of knowledge is dynamic – turbulently dynamic – with capacity to throw forth, for ever, new items of knowledge, calling for their own respective positions among the ones already existing.

1.2.2.4 Continuum

Fourthly, it is conjectured that the universe of knowledge is a continuum, that is, it has no holes. Whatever be the holes currently present in the universe of knowledge, it is believed that, with the development of the universe of knowledge through research, the holes will be successively filled up. In other words, it is our faith that any spot in the universe of knowledge lying fallow up to any moment will be cultivated at a later time. All these factors call for a more powerful and penetrating analysis and methodology than available at present, to determine the most helpful complex of invariants to be preserved in the preferred linear sequence of the items of knowledge – past, present, and future.

1.2.3 Alternative Method

In this circumstance, an alternative method for meeting the problem of classification in sense 2 is to set up several alternative models of classification. The state of universe of knowledge at any one moment may indicate which of the alternative models would be most helpful [7,25]. For example, a hundred years ago a severely

Design of Depth Classification

enumerative model proved sufficient. DC is an approximation to it, though it deviates slightly from being severely enumerative. Several successive unconscious deviations were made later. In some the deviations were significant. A classification grafting an analytico-synthetic superficies to the old enumerative core was felt to be necessary at the turn of the present century. UDC was the result. Today, a classification going the whole hog along analytico-synthetic line is realized to be a necessity [3]. CC is the result – particularly in its current stage of development. Another alternative name for such a classification is faceted classification. Forward-looking librarians have realised that the necessity has already come for faceted models of classification. They are experimenting with diverse models of faceted classification [2].

1.2.4 Postulational Approach

The experiments show that the mere fact of a scheme being faceted does not go the whole way in determining the invariant-complex to be preserved. It is felt that the desirable invariant-complex cannot be seized with our present methodology and mode of thinking, by searching for them in the phenomenal, surface level of the universe of knowledge. At this level, the invariant-complex appears to play the will-o-the-wisp. To escape its tantalisation a break through the phenomenal surface level has been made and a dive taken towards the noumenal seminal level. This break-through has resulted in the avoidance of the tantalising situation experienced all along. Since 1957, this break-through has enabled us to base the design of classification on a set of Postulates for the identification and separation of the facets of a subject and a set of Principles for

determining the helpful sequence of all such facets [18], for work in the Idea Plane. These have reinforced the principles already formulated to arrive systematically at a helpful sequence of isolates in an array and of classes in general [19]. The advantage of postulates is that the question of their being right or wrong does not arise. The only reason for accepting them is that work based on them leads to a helpful sequence of the classes in the universe of knowledge. The helpfulness of the resulting sequence of classes is as if the invariant-complex were found out and preserved in the linear arrangement. This idea was first formulated in respect of design of classification at the Dorking Conference in 1957 [15]. During the last seven years it has been found that basing design of classification on the postulates and principles mentioned above has yielded a helpful sequence among classes presenting several rounds and levels of facets. This helpfulness implies that the right invariant-complex is preserved though we are not yet able to formulate it in explicit terms. With the confidence generated by this experience, schemes of depth classification for applied sciences are being designed at the DRTC, Bangalore. The results are encouraging.

1.2.5 Scheme of Classes

In relation to enumerative classification, a statement of the classes of knowledge in their helpful sequence is known as a *Scheme of Classes*.

1.2.6 Scheme for Classes.

Even in relation to faceted classification, a *Scheme of Classes* is possible. But it is not necessary to provide it except for the use of

Design of Depth Classification

beginners and for use in generalist libraries having to arrange only conventional books embodying macro thought and to arrange the main entries of whole books. Most classes in such a *Scheme of Classes* will be such composite ones as can be synthesised out of a (BC) and a few – say two or three or rarely five – facets, that is isolates. Whatever be the prevision of the classificationists, such a *Scheme of Classes* will become inadequate sooner or later. For, it would amount to attempting to foist an enumerative garb on an analytico-synthetic classification defying enumeration. Further, the length of a fairly exhaustive *Scheme of Classes* will be un-economical in the measure of the relative numerousness of $(m \times n \times p \times q \times r)$ over $(m+n+p+q+r\dots)$ Therefore in a faceted classification a *Scheme for Classes* is given instead of a *Scheme of Classes*. The *Scheme for Classes* consists of many lists of schemes instead of a single one. First comes the *Scheme of (BC)*. Then in association with each (BC), there are a number of *Schemes of Isolates*. Each *Scheme of Isolates* takes care of one of the many successive facets likely to be presented by one or other of the diverse subjects sharing that (BC). Of course, such of the *Schemes of Isolates* as are likely to recur in association with all or many (BC) may be given as *Schemes of (CI)* immediately after the *Scheme of (BC)*. The *Scheme of (BC)*, the *Scheme of (CI)*, and the *Scheme of (Spl)* set up in association with the respective (BC) taken together, constitute the *Scheme for Classes*. In the *Scheme for classes* of a faceted classification, enumeration is resorted to only in the constituent *Schemes of (BC)* and of the several kinds of isolates.

1.3 Sense 3 - Library Classification

In library classification, each (BC), each (I), and each class made of a (BC) and one or more isolates is represented by an ordinal

number called respectively (BCN), (IN), and (CN). Of course, if the Scheme is enumerative, there will be (CN) only.

1.3.1 Schemes for Classification

In an enumerative classification each class in the Scheme of Classes is fitted with its (CN). The result is a single Schedule of Classes. It forms the Scheme of Classification. On the other hand in a faceted classification we do not get a single schedule. Each (BC) in the Scheme of (BC) is fitted with its (BCN). The result is the Schedule of (BC). Each (I) in each Scheme for (CI) is fitted with a (CIN). These then become the respective Schedules of (CI). In each Scheme of (Spl) going with the respective (BC), each (I) is fitted with an (IN). The result is as many Schedules of (Spl) as there are Schemes of (Spl). The totality of the Schedule of (BC) and the Schedules of (I) – common as well as special – constitutes the Scheme for Classification in an analytico-synthetic or a faceted classification.

1.4 Chart of Equivalent Terms

For convenience in exposition a few other terms are introduced to denote generically or severally classes and isolates of different kinds. There are parallel terms for use in the Idea Plane, the Verbal Plane, and the Notational Plane respectively. The Documentation Committee of the Indian Standards Institution has established a Standard in respect of these terms. They are given below in tabular form [4]. This tabular context lays bare the meaning of each term.

Note : In any set of equivalent terms, the term in the Idea Plane has been defined. Some of the other terms have been defined when expediency required, but not all.

Design of Depth Classification

Generic	In the Idea Plane	In the Verbal Plane	In the Notational Plane
Focus	Class Isolate Idea Array Isolate Idea	Subject Isolate Term Array Isolate Term	Class Number Isolate Number Array Isolate Number
Basic Focus	Basic Class	Basic Subject	Basic Class Number
Main Focus	Main Class	Main Subject	Main Class Number
Canonical Focus	Canonical Class	Canonical Subject	Canonical Class Number
Isolate	Isolate Idea	Isolate Term	Isolate Number
Common Isolate	Common Isolate Idea	Common Isolate Term	Common Isolate Number
Anteriorising Common Isolate	Anteriorising Common Isolate Idea	Anteriorising Common Isolate Term	Anteriorising Common Isolate Number
Posteriorising Common Isolate	Posteriorising Common Isolate Idea	Posteriorising Common Isolate Term	Posteriorising Common Isolate Number
Array Isolate	Array Isolate Idea	Array Isolate Term	Array Isolate Number
Facet	Basic Class Isolate Idea	Basic Subject Isolate Term	Basic Class Isolate Number
Basic Facet	Basic Class	Basic Subject	Basic Class Number
Isolate Facet	Isolate Idea	Isolate Term	Isolate Number
Compound Focus	Compound Class	Compound Subject	Compound Class Number

Generic	In the Idea Plane	In the Verbal Plane	In the Notational Plane
Complex Focus	Complex Class	Complex Subject	Complex Class Number
Complex Isolate	Complex Isolate Idea	Complex Isolate Term	Complex Isolate Number
Complex Array Isolate	Complex Array Isolate Idea	Complex Array Isolate Term	Complex Array Isolate Number
Biasing Focus	Biasing Class	Biasing Subject	Biasing Class Number
Biasing Isolate	Biasing Isolate Idea	Biasing Isolate Term	Biasing Isolate Number
Biasing Array Isolate	Biasing Array Isolate Idea	Biasing Array Isolate Term	Biasing Array Isolate Number
Comparison Focus	Comparison Class	Comparison Subject	Comparison Class Number
Comparison Isolate	Comparison Isolate Idea	Comparison Isolate Term	Comparison Isolate Number
Comparison Array Isolate	Comparison Array Isolate Idea	Comparison Array Isolate Term	Comparison Array Isolate Number
Difference Focus	Difference Class	Difference Subject	Difference Class Number
Difference Isolate	Difference Isolate Idea	Difference Isolate Term	Difference Isolate Number
Difference Array Isolate	Difference Array Isolate Idea	Difference Array Isolate Term	Difference Array Isolate Number

Design of Depth Classification

Generic	In the Idea Plane	In the Verbal Plane	In the Notational Plane
Influencing Focus	Influencing Class	Influencing Subject	Influencing Class Number
Influencing Isolate	Influencing Isolate Idea	Influencing Isolate Term	Influencing Isolate Number
General Relation Focus	General Relation Class	General Relation Subject	General Relation Class Number
General Relation Isolate	General Relation Idea	General Relation Isolate Term	General Relation Isolate Number
General Relation Array Isolate	General Relation Array Isolate Idea	General Relation Array Isolate Term	General Relation Array Isolate Number

1.5 Basic Class and Isolate

All along we have been using the terms (BC) and (I) without definition. At present we have no formal definition for them. We have to define them only by enumeration.

E118	Chemistry of gold	ND,7;618	Numismatics of gold
F118	Metallurgy of gold	X61,2	Gold as a medium of currency
HX118	Mining of gold	Y:86,3	Sociology of gold
HX118.2	Mining of gold in India	Z26(E118)	Gold as a movable property

Note : For Gold currency and Gold ornament, the material characteristic is used to construct the schedule for [1P1]. The sector (s-1) is used for it. In this sector, digit 1 stands for the (QI) Material. The array of order 2 listing the individual materials is telescoped into array of order 1. Gold occupies the first place among the materials used for currency or ornament. therefore, it is represented by the digit 2. This is in accordance with the latest principles for design of classification being expounded in this series of articles.

1.6 Main Class and Canonical Class

(BC) is a generic term to denote a (MC) or a (CC). A formal definition of these two terms is also difficult. They have to be defined by enumeration only. We can only say that a (MC) may comprehend many (CC). For example, the (MC) B Mathematics comprehends the following (CC) among others:

B1	Arithmetic	B24	Determinant
B11	Lower Arithmetic	B245	Matrix
B13	Integer	B3	Analysis
B15	Ideal number	B32	Calculus
B16	Complex number	B321	Differential
B18	Transcendental number	B325	Integral
B2	Algebra	B33	Differential equation
B22	Continued fraction	B6	Geometry
B23	Theory of equation	B7	Mechanics

Similarly the (MC) D Engineering comprehends the following (CC) among others:

Design of Depth Classification

D2	Building	D554	Water pollution
D3	Irrigation	D556	Water distribution
D4	Track	D56	Sewage
D411	Highway	D6	Power(Production) engineering
D415	Rail road	D7	Service (Production) engineering
D416	Bridge	D7Z	Commodity (Production) engineering
D5	Sanitary	D92	Land vehicle
D55	Water supply	D9A	Special purpose machinery
D551	Head work	D9N	Electronic equipment
D8	Machine tool (Production)	D9R	Nuclear power equipment
D82	Wood cutting tool		
D85	Metal cutting tool		
D9a	Machine elements (Production)		
D91	Vehicle (Production)		

The above is a modification of the schedule given in CC ed 6. This modification has been made in the light of the experience gained during the last two years in discussing in DRTC the Principles for the Design of Classification and in getting a better understanding of the concept of (CC). This matter will be presented in a later paper after more experience is gained. At present it is difficult to say whether we begin with the ultimate (CC), and name the (MC) comprehending them or whether we begin with certain postulated (MC) and break down some of them into (CC). The only fact that we can now state is that each (MC) and each (CC) have a considerable literary warrant. There are also some partial comprehensions of (MC) with literary warrant. The following are some examples among others:

A	Natural science (General)	MZ	Humanities and Social sciences
AZ	Mathematical sciences		
BZ	Physical sciences	MZ	AHumanities
FZ	Mining and Metallurgy	NZ	Literature and Language
GZ	Earth sciences (Geology and Geography)	SZ	Social sciences

In CC, digit Z is used as the mnemonic digit for partial comprehension. It is taken to be an 'Emptying Digit' — that is it empties the preceding digit of its semantic richness, though allowing the retention of its ordinal value [8]. For example in SZ, the digit S does not represent Psychology. In fact in this context it does not by itself represent any subject whatever. But the digit-pair SZ represents Social Sciences. The result of retaining the ordinal value of S is that SZ is coordinate with S and T and comes between them.

1.7- Number of (BC)

Till now the number of (MC) was taken to be very limited. But in recent years several methodologies have gained literary warrant entitling them to the status of (MC). Here are some examples:

PX Communication theory XX Management (General theory)

Apart from this several classes, forcedly accommodated as sub-classes of the other (MC) in order to meet exigencies in notational construction, have been all along insisting on their status as co-ordinate (MC) being implemented in the Notational Plane. Here are some examples:

Design of Depth Classification

HX	Mining	LX	Pharmacognosy
KX	Animal husbandry		

In CC the digit X is taken to be an 'Emptying Digit'. The digit-pair made of an Emptied and an Emptying Digit is semantically rich. It represents a (MC). It is co-ordinate with the digit Emptied. Thus KX Animal Husbandry is co-ordinate with K Zoology and L Medicine. In CC the digits T to Y are given this property. This, taken with other concepts fits the Notational Plane to receive an infinity of (MC) – that is classes in the array of order 1 got by dividing the universe of knowledge. Further, in the clearer comprehension of the concept of (CC) now being got, the number of (CC) seems to be very large. Thus the number of (BC) is very large. Theoretically it tends to infinity.

1.7.1 Implication of Infinity

The above realization fits in with the concept of the universe of knowledge being infinite – that is the number of items of knowledge past, present, and future taken together tends to infinity. The elemental items of knowledge consists of (BC) and (I). The total number of all the (CI) – including (TI) of all levels, (SI) of all levels, the (CEI), the Common Property Isolates, the Common Value Isolates, the (CPI), the (ACI) – is very large. The number of Commodity Isolates is also very large almost of the order 10^{10} as they are today. The number of (SpI) likely to go with the respective (BC) – including the (SpEI) of all rounds, the (SpMI), (SpPI) of all levels in all rounds – is also very large. We may say that the number of (I) tends to infinity. Still more the number of (CdC) and (CxX) tends to infinity. By removing from the infinite universe of knowledge the infinite universe of (I), (CdC) and (CxX), we get the universe of (BC) as the residue.

That the residue should be also be infinite has been sensed from Vedic times. For, according to a Vedic statement if infinity is taken away from infinity, then infinity itself will be left as residue. This represents the intuitive grasp of a certain experience by the seers of yore. In recent years, the intellectual work of mathematicians has led to the statement of the same result as a Postulate about Infinity. Restricting the (BC) to a small number has all along been giving an uneasiness in mind. We now realize that it might have been due to the inhibition of the Idea Plane by the limitedness of the base in the Notational Plane. The concept of Emptying Digit has now fitted the Notational Plane to accept any number of (BC) interpolatable between any two existing (BC). The opening of Zone 4 by the Packet Notation has fitted the Notational Plane to accept any number of (BC) extrapolatable at the end of the already existing sequence of (BC).

1.7.2 Scope of this Series

In this series of papers on design of classification, the design of the sequence of the Schedule of (BC) may be taken up only sparingly by way of illustration. For, what has been said already gives sufficient indication as to how the problem can be dealt with. All that we have to do is to remember the Canon of Filiatory Sequence [5]. The papers in this series will be mostly on the design of schedules of (CI) and the schedules of (PI). Some work has already been done on the (CI) of [T] [24] and of [S] [23] of different levels needed for depth classification. The work on the other (CI) is in progress. The earlier papers in this series will be devoted to the working out of the schedules of (Spl) of all rounds and levels likely to be presented by subjects including those of deep intention, sharing some specific (BC). In any single paper we shall generally consider only a single (BC) and a schedule of (Spl) of the compound subjects derivable from it.

2 IDEA PLANE AND NOTATIONAL PLANE

2.1 Inhibition 1 and Connecting Symbol Colon

All along, work in the Idea Plane had been considerably inhibited by the limitations in the Notational Plane. For example, though a subject presented two or more unmistakably distinct facets, the limitations of the monolithic notation had inhibited DC from giving equal weight and equal facility for expansion of each of the facets. This made DC ignore the findings of the Idea Plane and forcedly prefer one of the facets alone for full representation in the (CN). In due course, practice with DC both during the process of classifying and while doing reference service had inhibited the mind from even recognizing the overlooked facets in a subject. But this inhibition was removed when the connecting symbol Colon lying in the lowest region of the ordinal scale was introduced by CC to serve at once both the functions of separating the facet numbers and as well as connecting them. This notation may be described as polyolithic.

2.2 Inhibition 2 and Five Connecting Symbols

The use of a single connecting symbol for all the facets worked fairly well in the classification of conventional books embodying macro thought. But when articles embodying micro thought had to be classified to serve the needs of documentation, to the full satisfaction of the Laws of Library Science, the Notational Plane was subjected to a terrible strain. As if in sympathy with this predicament of the Notational Plane, the Idea Plane was tempted to inhibit itself and suppress its full findings. The stage of the polyolithic notation of CC brought classification of micro thought to the same gross condition as

the monolithic notation of DC had brought classification of macro thought. In due course, practice with the first 1924 version of CC both during the process of classifying and while doing reference service, had inhibited the mind from even recognizing the difficulties. But this inhibition was removed when the single connecting symbol colon (:) was replaced by the five connecting symbols (single inverted comma, dot, colon, semicolon, and comma), all lying in the lowest region of the ordinal scale. The Notational Plane thus overcame its own limitations. This enabled the Idea Plane to overcome its inhibition 2 and put forth the Postulate of (FC) [17]. These results were implemented for the first time only in Ed. 4 of CC in 1952.

2.3 Inhibition 3 and Length of the Base

Even in the 1924 version of CC, the base of its notation had three zones:

1. The 23 Roman smalls excluding i, l, and o;
2. The 9 Indo-Arabic Numerals; and
3. The 26 Roman caps for the array of (MC) and 24 Roman caps, excluding I and O, for all other arrays.

The three zones taken together yielded 58 (MCN) and 56 (IN) for other arrays. However, the entire range was utilised only in the Array of (MC). But this was not done in the case of an array of order 1 in any (IF). In array of order 1 of [T] the nine (IN) of zone 2 were not used. And in array of order 1 of [S] the (IN) of zone 3 were not used. This was due to the same (CS) having been used for [T] and [S]. This self-imposed restriction was however removed the moment the single inverted comma was introduced as (CS) for [T]. However, the (IN) of zone 1 were not brought into use in the array of order 1 of any of the (FC). Further, the (IN) of zone 3 were not brought into use in any

Design of Depth Classification

array of any order of any facet except [T] used by itself or used to represent [P]. In [P] it was used to represent either persons or 'Systems'. The (IN) in zone 1 were not brought into use anywhere except to represent (ACI). This neglect of zones 1 and 3 due to the pressure – unconscious though it was – of the DC tradition of using (IAN) alone. This neglect of zones 1 and 3 inhibited the Idea Plane from asserting some of its requirements such as (CEI), Common Property and Value (I), (CPI), and the second and later levels of [T] and [S]. In due course the Notational Plane became aware of the neglect of zones 1 and 3. Once the Notational Plane gained this awareness, the Idea Plane overcame its inhibition, and demanded the implementation of the requirements mentioned above. The results were embodied for the first time in Ed 5 of CC in 1957.

2.4 Inhibition 4 and Super-imposition

Even in Ed 1 (1933) of CC, the concepts of 'Super-Imposition' in the Idea Plane and of the (SID) in the Notational Plane had been seized. The digit hyphen (-) was taken as the (CS) between the components of Super-imposed (IN). Till ed 6, it was called 'Auto-bias Device'. The use of the old term cannot explain an inhibition in the Idea Plane presenting all along. For some inexplicable reason, the Super-Imposition concept was brought into use only in the Schedule for Organs going with the (BC) L Medicine. In spite of its giving a considerable help in subjects in Medicine, it was not extended to other subjects. The subjects in Y Sociology often called for the application of the Super-Imposition concept. But the Idea Plane was inhibited from responding to it. It is only during the last few months that the extension of the Super-Imposition concept was made to all subjects. This has naturally removed the inhibition in the Idea Plane.

2.5 Inhibition 5 and (W) and (\overline{W})

One of the causes for failure to bring the concept of Super-Imposition into wide use was the experience with the Organ Schedule of L Medicine. In the human body, the Universe of Organs is made of sub-universes such as the Digestive System, the Circulatory System, and the Respiratory System. Each such sub-universe generally consists only of one (I). For example, there is only one mouth, one stomach, one liver, and so on. In certain cases, there are two (I). For example, there are two legs, two hands, two eyes, and so on. But there is not much occasion for distinguishing between two such (I) in the theoretical study of medicine – not taken to a great depth. Therefore, such a sub-universe also was treated as if it had only one (I). Thus, each sub-universe virtually turned out to be a unitary universe [26]. No further sub-division of it – that is a Unitary Universe – can yield a (W) Entity. - On the other hand, the sub-division of any such sub-universe yielded only (W) Entity. Further, in the Notational Plane, for the isolate in each such sub-universe of organs consisted only of one digit, while moving from (W) to (\overline{W}) belonging to different levels of facets of the same (FC) in the same round. This inhibition was removed only when the concepts of (W) and (\overline{W}) were explicitly stated in 1953 [27].

2.6 Inhibition 6 and Quasi-isolate

The array of order 1 in the schedule for [1P1] going with (MC) Y Sociology is made of (QI) – that is first characteristics – instead of regular (I). This was the result of the subdivision of the universe of Social Groups demanding more than one first characteristic. Each first characteristic started its own Train of Characteristics. This method of allowing several Trains of Characteristics – each based on its own first characteristic – in [1P1] going with a (BC) occurs in the

Design of Depth Classification

applied sciences. The occurrence of several Trains of Characteristics led to the inhibition of the Idea Plane from recognizing that a combination of (I) drawn from two or more such Trains of Characteristics still belong to [1P1] only. As a result, the (I) drawn from the different Trains of Characteristics were treated as if they belonged to different levels of [P] in spite of their combination still yielding only (W) Entities. In other words, (W) Entities and (\overline{W}) Entities became indistinguishable in the Notational Plane. The realization was that each Train of Characteristics though apparently different, together constitute only a complex Train of Characteristics yielding only (WI). This is of course due to the (I) in arrays of order 1 of all such Trains of Characteristics being only (QI). Once this was recognized, the concept of 'Super-Imposition' gave a ready made way for combining the (I) drawn from different Trains of Characteristics into a single 'Super-Imposed Isolate' in [1P1]. This has come to be of immense use in the design of depth classification, particularly for the applied sciences.

3 IDEA PLANE

3.0 Two Directions of Approach

There are two ways along which the schedules of (I) going with a particular (BC) can be approached. One way will be to start with listing all the ultimate (CdC) having that (BC) and having literary warrant. We may describe this as the 'Last-Link-Upwards Approach'. The other way will be to start with the (BC) as the universe. We may arrive successively at link after link in each chain in each possible facet likely to be presented by subjects sharing the (BC). We may describe this way as 'First-link-Downwards Approach'.

3.1 Last-Link-Upwards Approach

In the Last-Link-Upwards Approach, the last link will be a subject and not an (I). While moving upwards, each link will be a class and not an (I). In fact, we shall be moving up a chain of (CdC). At each step we should determine the constituent (I) or facets of the (CdC) concerned. The (I) of like kind occurring in all the links of the chain should be listed. These (I) should be arranged in a sequence of increasing extension. The sequences so got from all the last links should be established. The (IN) should be determined after the list of (I) is completed. This will prove to be maddening. Further, as a result of the ever growing nature of the universe of subjects sharing one and the same (BC), links of a lower order will be thrown forth by what was already taken as the last link in a chain. Thus what was taken to be the last link will thereafter cease to be the last link. Therefore, it looks as if the Last-Link-Upwards Approach will not be practicable unless the totality of the subjects sharing the given (BC) will be small even after full development. In fact, the last link either in the chain of subjects or in the chain of a particular facet is ever elusive. The term 'Inductive Scheme of Classification' has been suggested to denote the 'Last-Link-upwards Approach'. The impracticability of this approach was brought out in December 1952 [13].

3.2 First-Link-Downwards Approach

In the First-Link-Downwards Approach, the design of an analytico-synthetic or a faceted scheme of classification is based on postulates for the Idea Plane and the Notational Plane respectively.

Design of Depth Classification

The design of CC, for example, is based on a definite set of postulates given in the *Elements* [16]. To apply the postulates we have to conjecture the possible rounds and levels of each of the five (FC) which may appear even in a subject of the highest possible order of (CdC) [6] having the given (BC). As already stated we have to think only of the (Spl) in [P], or in [M], or in [E] of any round or any level. According to the experience gained so far, the (SpEI) will be only a few in each (BC). The (SpMI) will also be equally few. What is more significant is that the rate of increase in the number (Spl) of either kind will be extremely small. Further, the (SpEI) will be largely made of quasi-common isolates determinable by their seminal equivalents. Moreover the later the round the easier will be the design of the schedule in [E] and of the schedule in any level of [M]. Therefore, it is not proposed to deal with the design of the schedule for the (Spl) of [E] or [M], in the earlier articles of this series.

3.3 Reduction of Work on Schedule to (Spl) in [P]

On the other hand the number of relevant characteristics to be used as the basis for designing the schedule of (Spl) of [1P1] will be fairly large. The number of levels of [P] in round 1 will also be many at least two and often many more. The (Spl) in each of these [P] of round 1 will also be large and may call for more than one relevant first characteristic. The rate of increase in the number of (I) in [1P1] will be great. The work of constructing the schedules for the levels of [P] in round 1 will be arduous in many of the (BC) and particularly in applied sciences. We shall therefore begin this series with articles on the design of classification of (Spl) in [1P1], and [1P2] for subjects in applied sciences.

3.4 Blending of Methods

In the design of the schedules mentioned above we shall have to depend upon speculative method that is to start with conjecture. But depending solely on it may throw the schedules out of focus. The number of the facets as well as the enumeration of (l) in each facet may be thrown out of focus. A corrective should be applied to this by frequently looking into the literary warrant that is by pragmatic method. In fact, a judicious blending of the speculative and pragmatic methods is necessary. It is a matter of compromise between the two methods. Compromise between conflicting claims is a well-known principle in practical life. So it is in the design of classification too. In other words each method should be used as a check on the other. At what stages and how often the check up should be made will vary with the context of the subject and the experience of the designer of classification.

4 NOTATIONAL PLANE

4.1 Mixed Notation

The methodology for the design work in the Idea Plane discussed in Sec 3 and its subdivisions is available for use in any Scheme of Classification. But the methodology for design work in the Notational Plane should necessarily vary with the Notational System used by a scheme. Therefore, the discussion about the notational Plane will be confined to the CC. The notation of CC is a mixed one. The advantages of having a distinctive species of digits for use as (CS) in an analytico-synthetic or a faceted scheme of classification has been already discussed in Sec 2.2. The semantically rich digits [21] of CC are drawn from three species. The use of packet notation

Design of Depth Classification

starter and arrester amounts virtually to the use of a fourth species of semantically rich digit; for, a packeted number can be treated as if it were a single digit of a new species. The number within the brackets may consist of any number of digits and of any species of digits, including (CS) and packet number itself.

4.2 Zone and Sector Formation in Array of Order 1

Let us remember that we are now interested only in the design of the schedule corresponding to the several (IF) going with a particular (BC). Further, let us also remember that in the earlier papers of this series we shall confine ourselves to the array of order 1 of a facet. Further, we shall take into account only such packeted (IN) as have a single semantically rich digit. Still further, we shall use the terms 'zone' and 'sector' to cover different parts of the array as shown in the following table:

Zone	Sector	(AIN)	Total of (AIN) for	
			Sector	Zone
1	(S - a)	a, b z	23	23
2	(S - 1)	1, 2 9	9	9
3	(S - A)	A, B Z	24	24
4	(S - (a))	(a) (b) (z)	23	23
5	(S - (1))	(1), (2) (9)	9	9
6	(S - (A))	(A), (B) (Z)	24	24
Total 6 sectors			112	

4.3 Table for Boundary Condition 1

Note: The few cases in which I and O can occur within brackets have been ignored in this census.

Annotation 1

The symbol used to denote a sector as shown in column 2 of the above table should be noted. It is a helpful descriptive symbol. It satisfies the Law of Parsimony in thinking, in expressing, and in writing.

Annotation 2

In what follows, when we speak of 'number of digits' in packeted (AIN) – that is, an (AIN) of zone 4 – we shall mean only the 'number of digits within the brackets'. From the table, it follows that the total number of one-digit (AIN) available in **any array of order 1** of a facet is apparently 112. But it is wise to **make the digits z, 9, and Z Empty Digits**, in order to have infinite hospitality within each zone. Further, we must be prepared to use the first digit or each zone that is *a*, 1 and A – to represent a (QI) – that is, a characteristic instead of a true (I). Therefore, we have the following results:

4.3.1 *Census for Boundary Condition 1*

With one digit as the boundary condition for the number of digits in an (AIN) of order 1, the number of sectors available is 6 and the number of (AIN) available is 100.

4.4 Table for Boundary Condition 2

We shall use the term sector to cover different parts of each of the different zones of the array as shown in the following table:

Design of Depth Classification

Zone	Sector	(AIN)	Total of (AIN) for		
			Sector	Zone	
1	(S-a)	a, b	.. y	22	75
	(S-za)	za, zb	.. zy	22	
	(S-z1)	z1, z2	.. z8	8	
	(S-zA)	zA, zB	.. zY	23	
2	(S-1)	1, 2	.. 8	8	61
	(S-9a)	9a, 9b	.. 9y	22	
	(S-91)	91, 92	.. 98	8	
	(S-9A)	9A, 9B	.. 9Y	23	
3	(S-A)	A, B	.. Y	23	76
	(S-Za)	Za, Zb	.. Zy	22	
	(S-Z1)	Z1, Z2	.. Z8	8	
	(S-ZA)	ZA, ZB	.. ZY	23	
4	(S-(a))	(a), (b)	.. (y)	22	
	(S-(za))	(za), (zb)	.. (zy)	22	
	(S-(z1))	(z1), (z2)	.. (z8)	8	
	(S-(zA))	(zA), (zB)	.. (zY)	23	
5	(S-(1))	(1), (2)	.. (8)	8	
	(S-(9a))	(9a), (9b)	.. (9y)	22	
	(S-(91))	(91), (92)	.. (98)	8	
	(S-(9A))	(9A), (9B)	.. (9Y)	23	
6	(S-(A))	(A), (B)	.. (Y)	23	212
	(S-(Za))	(Za), (Zb)	.. (Zy)	22	
	(S-(Z1))	(Z1), (Z2)	.. (Z8)	8	
	(S-(ZA))	(ZA), (ZB)	.. (ZY)	23	
Total Sectors		24			424

Here again we must be prepared to use the first digit of each sector to represent a (QI). Therefore we have the following result:

4.4.1 Census for Boundary Condition 2

With two digits as the boundary condition for the number of digits in an (AIN) of order 1, the number of sectors available is 24 and the number of (AIN) available is 400.

Annotation 1

In some of the sectors the (AIN) has apparently two digits. In all such cases the first digit is empty. Therefore, in all the (AIN) there is only one semantically rich digit. Therefore, all the 400 (I) are co-ordinate ones belonging to a single array as viewed from the Idea Plane.

4.5 Table for Boundary Condition 3

In passing from Sec 4.2.1 to 4.3.1 we increase the number of sectors from 6 to 24. We also increase the number of (AIN) from 100 to 400. This, we were able to secure by using the Empty Digits z, 9, and Z once and only once. This meant increasing the boundary condition for the number of digits in an (AIN) from 1 to 2. If necessary we can increase the boundary condition for the number of digits still further from 2 to 3. This can be done by using the following nine combination of Empty Digits:

zz	9z	Zz
z9	99	Z9
zZ	Zz	ZZ

The following table gives the result:

Design of Depth Classification

Zone	Sector	(AIN)	Total of (AIN) for		
			Sector	Zone	
1	(S-a)	a, b	.. y	22	
	(S-za)	za, zb	.. zy	22	
	(S-zza)	zza, zzb	.. zzy	22	
	(S-zz1)	zz1, zz2	.. zz8	8	
	(S-zzA)	zzA, zzB	.. zz8	23	
	(S-z1)	z1, z2	.. zzY	8	
	(S-z9a)	z9a, z9b	.. z8	22	
	(S-z91)	z91, z92	.. z98	8	
	(S-z9A)	z9A, z9B	.. z9Y	23	
	(S-zA)	zA, zB	.. zY	23	
	(S-zZa)	zZa, zZb	.. zZy	22	
	(S-zZ1)	zZ1, zZ2	.. zZ8	8	
	(S-zZA)	zZA, zZb	.. ZY	23	234
2	(S-1)	1,2	.. 8	8	
	(S-9a)	9a, 9b	.. 9y	22	
	(S-9za)	9za, 9zb	.. 9zy	22	
	(S-9z1)	9z1, 9z2	.. 9z8	8	
	(S-9zA)	9zA, 9zB	.. 9zY	23	
	(S-91)	91, 92	.. 98	8	
	(S-99a)	99a, 99b	.. 99y	22	
	(S-991)	991, 992	.. 998	8	
	(S-99A)	99A, 99B	.. 99Y	23	
	(S-9A)	9A, 9B	.. 9Y	23	
	(S-9Za)	9Za, 9Zb	.. 9Zy	22	
	(S-9ZA)	9ZA, 9ZB	.. 9ZY	23	220

Zone	Sector	(AIN)	Total of (AIN) for		
			Sector	Zone	
3	(S-A)	A, B	.. Y	23	
	(S-Za)	Za, Zb	.. Zy	22	
	(S-Zza)	Zza, Zzb	.. Zzy	22	
	(S-Zz1)	Zz1, Zz2	.. Zz8	8	
	(S-ZzA)	ZzA, ZzB	.. ZzY	23	
	(S-Z1)	Z1, Z2	.. Z8	8	
	(S-Z9a)	Z9a, Z9b	.. Z9y	22	
	(S-Z91)	Z91, Z92	.. Z98	8	
	(S-Z9A)	Z9A, Z9B	.. Z9Y	23	
	(S-ZA)	ZA, ZB	.. ZY	23	
	(S-ZZa)	ZZa, ZZb	.. ZZy	22	
	(S-ZZ1)	ZZ1, ZZ2	.. ZZ8	8	
	(S-ZZA)	ZZA, ZZB	.. ZZY	23	235
	4	(S-(a))	(a), (b)	.. (y)	22
(S-(za))		(za), (zb)	.. (zy)	22	
(S-(zza))		(zza), (zzb)	.. (zzy)	22	
(S-(zz1))		(zz1), (zz2)	.. (zz8)	8	
(S-(zzA))		(zzA), (zzB)	.. (zzY)	23	
(S-(z1))		(z1), (z2)	.. (z8)	8	
(S-(z9a))		(z9b), (z9b)	.. (z9y)	22	
(S-(z91))		(z91), (z92)	.. (z98)	8	
(S-(z9A))		(z9A), (z9B)	.. (z9Y)	23	
(S-(zA))		(zA), (zB)	.. (zY)	23	

Design of Depth Classification

Zone	Sector	(AIN)	Total of (AIN) for	
			Sector	Zone
	(S-(zZa))	(zZa), (zZb) .. (zZy)	22	
	(S-(zA1))	(zZ1), (zZ2) .. (zZ8)	8	
	(S-(zZA))	(zZA), (zZB) .. (zZY)	23	234
5	(S-(1))	(1), (2) .. (8)	8	
	(S-(9a))	(9a), (9b) .. (9y)	22	
	(S-(9za))	(9za), (9zb) .. (9zy)	22	
	(S-(9z1))	(9z1), (9z2) .. (9z9)	8	
	(S-(9zA))	(9zA), (9zB) .. (9zY)	23	
	(S-(91))	(91), (92) .. (98)	8	
	(S-(99a))	(99a), (99b) .. (99y)	22	
	(S-(991))	(991), (992) .. (998)	8	
	(S-(99A))	(99A), (99B) .. (99Y)	23	
	(S-(9A))	(9A), (9B) .. (9Y)	23	
	(S-(9Za))	(9Za), (9zb) .. (9ZY)	22	
	(S-9Z1))	(9Z1), (9Z2) .. (9Z8)	8	
	(S-(9ZA))	(9ZA), (9ZB) .. (9ZY)	23	220
6	(S-(A))	(A), (B) .. (Y)	23	
	(S-(Za))	(Za), (Zb) .. (Zy)	22	
	(S-(Zza))	(Zza), (Zzb) .. (Zzy)	22	
	(S-(Zz1))	(Zz1), (Zz2) .. (Zz8)	8	
	(S-(ZzA))	(ZzA), (ZzB) .. (ZzY)	23	
	(S-(Z1))	(Z1), (Z2) .. (Z8)	8	

Zone	Sector	(AIN)	Total of (AIN) for	
			Sector	Zone
	(S-(Z9a))	(Z9a), (Z9b) .. (Z9y)	22	
	(S-(Z91))	(Z91), (Z92) .. (Z98)	8	
	(S-(Z9A))	(Z9A), (Z9B) .. (Z9Y)	23	
	(S-(ZA))	(ZA), (ZB) .. (ZY)	23	
	(S-(ZZa))	(ZZa), (ZZb) .. (ZZy)	22	
	(S-(ZZ1))	(ZZ1), (ZZ2) .. (ZZ8)	8	
	(S-(ZZA))	(ZZA), (ZZB) .. (ZZY)	23	235
Total sectors	78			1,378

As in the earlier cases, we must be prepared to use the first digit of each sector to represent a (QI). Therefore we have the following result:

4.5.1 Census for Boundary Condition 3

With three digits as the boundary condition for the number of digits in an (AIN) of order 1, the number of sectors available is 78 and the number of (AIN) available is 1,300.

4.6 Consolidated Census for Array of Order 1

Max N of Digits	N of Sectors	N of (AIN)
1	6	100
2	24	400
3	78	1,300

Design of Depth Classification

4.7 Allocation of Sectors for Array or Order 1 In [1p1]

It can be seen from the tables in sections 4.2, 4.3, and 4.4 that each sector is denoted by a symbol such as (S-A). (S-A) denotes the sector with (IN) A as its first (AIN).

Similarly, the symbol (S-ZA) denotes a sector with (IN) ZA as its first (AIN).

Again, (S-ZZ1) denotes the sector whose first (AIN) is ZZ1.

The following table shows a convenient allocation of the sectors in any (P).

Sector	To accommodate in (P)
(S-a)	.. (CPI)
(S-za)	.. (WI) – that is (I) in a [P3]
(S-z1) to (S-zA)	$\overline{(WI)}$ – that is (I) in a [P2]
(S-1) to (S-ZA)	$\overline{\overline{(WI)}}$ – that is (I) in [P1]

It may be repeated that the first (AIN) in each sector should be normally reserved to represent the (QI)— that is, the first characteristic on the basis of which the (I) to be accommodated in the sector are to be derived. An illustration is given for the (BC) 'Highway Engineering'.

In sector (S-za), za represents the 'Organ of remove 2'.

In sector (S-z1) and sector (S-ZA)

z1 should represent 'Organ of remove 1';
z2 should represent the (QI) 'By Organ (Vertical)';
zA may represent the (QI) 'By Organ (Lateral)';
zM may represent the (QI) 'By Organ (Longitudinal)'; and
zT may represent the (QI) 'By Organ (as in Traffic Engineering)'.

In sector (S-A), the (IN) A may represent the (QI) 'By User'.

In sector (S-ZI), the (IN) Z1 may represent the (QI) 'By Use'.

In sector (S-ZA), the (IN) ZA may represent the (QI) 'By Terrain'.

In sector (S-(A)), the (IN) (A) may represent the (QI) 'By Brand — that is, by such a thing as 'Trade name'.

4.7.1 Overflow Quasi Isolate

If the number of first characteristics — that is (QI) — exceeds 24. The remaining 54 sectors given in the table of sectors and (AIN) in Sec 4.5.1 for the boundary condition of 3 digits in an (AIN) may be brought into use. The particular (QI) of such sectors brought into use will depend upon the position of the train of characteristics started by it among those already scheduled.

4.7.2 Combination of Trains of Characteristics.

If the subject of a document presents a (Cdl) made of (I) taken from two or more of the trains of characteristics, the component (IN) should be arranged in the sequence in which their respective trains of

Design of Depth Classification

characteristics stand arranged in the schedule. They should then be connected to one another with the (CS) hyphen (-) as prescribed by the (SID) for the Notational Plane. Unless the (BCN) is one digit, the (CS) comma (,) should be inserted after the (BCN). This means that the insertion of a comma (,) is necessary after all the (CCN) and in the case of (MCN) beginning with an Emptied Digit.

4.7.3 Levels of [P]

If the subject of a document involves facets of level 2 or level 3, the usual rules apply about the insertion of comma(,) between levels. If, however, level 2 alone occurs without level 1, it is enough if a single comma (,) is inserted between the (BCN) and [1P2]. The allotment of sectors to different levels of [P] in a round is so made that no homonym will be created while the contemplated helpful sequence of the resulting classes is secured.

4.8 Allocation of Sectors in Array of Order 2 or in Arrays of Higher Orders

Whether the facet is a [P], or a [M], or an [E] sector (S-a) in arrays of order 2 and above will be occupied by (ACI). Whether the other sectors of zone 1 can be used for (Spl) has not yet been decided. Even supposing they can be so used, it has not yet been decided how they should be utilised. But, all the sectors of zones 2, 3, 4, 5, and 6 can be used for the subdivision of the (AI) of array of order 1.

5 PROCEDURE FOR THE SCHEDULE FOR [1P1]

5.1 Number of First Characteristics

Find out by the blending of the speculative method and the pragmatic method mentioned in Sec 3.4, the various first characteristics to be used, for subjects likely to go with the (BC) under consideration. Let us denote them by the symbols (FP), (FQ), etc. (See Sec 6.1 for example).

5.2 Sequence of First Characteristics

Consider any two first characteristics at a time. By the application of the Principles for Facet Sequence, determine the sequence in which the first characteristic that is (QI) should occur in the array. It is found from experience that the Wall-Picture Principle proves sufficient in most cases. But, of course, one should cultivate the experience and sensitiveness necessary to apply this Principle. Suppose, we have fixed the sequence (FP) (FQ). Then fix the position of (FQ) by taking it in relation to (FP) and (FQ) successively. It can be easily seen there are three possible places for (FQ) that is, before (FP), between (FP) and (FQ), and after (FQ). If the decision is before (FP) and if the comparison of (FP) and (FR) is made first, there will be no need to compare (FQ) and (FR). But if by chance, we had begun comparing (FQ) and (FR) and we get the sequence (FR) (FQ) we should again compare (FR) (FP). We decide that (FR) should precede (FP). In other words, two operations will be necessary instead of one. With experience, one can develop this flair to arrive at the result with the least number of operations. The helpfulness of such a flair will be

realized when the number of first characteristics is large. (See Sec 6.2 for an example).

5.3 Allotment of First Characteristics to Sectors

After the sequence of first characteristics is established, each first characteristic should be allotted to a sector. The sequence of the sectors so allotted should be parallel to the sequence established for the first characteristics as prescribed in Sec 5.2.

5.4 Telecopying in Array

The (I) based on the first characteristic represented by the first (AIN) of the sector should be found out by blending the speculative method and the pragmatic method. Then the (I) should be arranged in a helpful sequence with the help of the appropriate principle found among the Principles for Helpful Sequence [19]. Then each (I) should be allotted an (AIN) in such a way that the sequence of (AIN) is parallel to the sequence of the (I) already determined. In relation to the first (I) of the sector, the other (I) belong to the array of the next order as viewed from the Idea Plane. In fact, they form array of order 2. Allotting these (I) to the (IN) in the first order array as suggested above amounts to telecopying the array of order 2 into the array of order 1. If some of the (IN) of the array of order 1 are left unused, an attempt should be made to use them by the telecopying array of order 3 into array of order 1. This requires judgment. This will be illustrated in the later articles of this series.

5.5 Economy in Notation

The examples in the later articles will show the great economy in notation secured by this way of utilising the mixed notational base and sector analysis in CC. The economy in notation pervades not only [1P1] but also [1P2] and [1P3]. This is the result of the allotment of sectors made in Sec 4.7.3.

5.6 Comfort to Eye and Memory

Experience shows that the need to further subdivide the (AIN) in a sector is only occasional. Even if it is necessary, one step of sub-division may prove sufficient. This will prevent the full (AIN) not to exceed three digits – that is, the optimum number of digits for the comfort of the physiology of the eye and the psychology of the memory mentioned at the International Conference of Libraries and Documentation Centres held in Brussels in 1955 [20]. When a (Cdl) is necessary, each component will also lie within this optimum limit and will get separated by its neighbouring components by a hyphen (-), thus giving relief to the eye and the memory.

5.7 Sum-up

To sum up, the design of classification, in respect of [1P1], for the subjects going with a specific (BC) consists of the following seven successive steps:

1. Determination of the possible first characteristics to be used as (QI) in the Array of Order 1;
2. Determination of a helpful sequence for the chosen first characteristics in the Array in the Idea Plane;

Design of Depth Classification

3. Allocation of the first characteristics to their respective sectors in the Array, in the Notational Plane;
4. For each of the first characteristics, construction of the provisional schedule for the isolates of all the necessary orders;
5. Classification of a variety of nascent micro documents by the provisional schedule, as a pilot project;
6. Finalisation of the schedule in the light of the experience gained in the pilot project; and
7. Additions and amendments to the schedule from time to time.

6 LONG PERSISTING CONFUSION

Till a few months ago, a confusion persisted in the Idea Plane. This was due to the spell caused by the new experience with the concept of (FC). Here is an example of such a spell. Consider the commodity 'Wooden Compasses' for use in drawing large figures on the black board. This term was taken to be a composite term. It was split into 'Compasses' and 'Wood'. The entity 'Compasses' was taken to belong to [P]; and the entity 'Wood' was taken to belong to [M]. Stated in words, the facet formula was taken to be:

[Compasses]; [Wood].

Consider a subject such as 'Breakage of Wooden Compasses'. Stated in words, the facet formula will then be:

[Compasses]; [Wood] : [Disorder] [Breakage].

The above may be represented in symbols as follows:

[1P1] ; [1M1] : [1E] [2P1].

This sequence would make 'Disorder' pertain to 'Wood' and not to 'Compasses'. Stated in general terms, the [1E] [Energy Isolate] would pertain to the [1M1] (Material Isolate) instead of to [1P1] (Commodity Isolate). This has been giving trouble for some years. During the last few months, the trouble has been traced to the initial error of looking upon 'Wooden Compasses' as a composite term and breaking it into a 'Personality Term' and a 'Matter Term.'

6.1 Removal of Confusion

It is now seen that the term 'Wooden Compasses' as a whole should be taken to be the Personality Term. To put it in other words, in classifying the universe of entities in [P] consisting of compasses, we can use the material of which the entity is made as a first characteristic – that is a (QI). On the basis of this first characteristic, we can get (I) such as:

- 1 Wooden Compasses;
- 2 Steel Compasses;
- 3 Aluminium Compasses; and so on.

Thus the material is not taken to be a manifestation in [M], but only as a means of identifying a commodity – that is, and (I) in [P]. Similarly, in classifying the universe of entities in [P] consisting of compasses, we may use as a first characteristic – that is a (QI):

Design of Depth Classification

- 1 Length of the arm;
- 2 Writing material attached at the end of the arm (such as inking compasses, penciling compasses, and chalking compasses);
- 3 Nature of the head joint of compasses; and so on.

Each such first characteristic will become a (QI) in the array of order 1 in [1P1]. This has already been mentioned in general terms in Sec 5.1.

6.2 Sequence of (QI)

As stated in Sec 5.2, the sequence of the (QI) in the array of order 1 can be determined by the Wall-Picture Principle. Let us assume for the time being that their sequence will be as follows:

- 1 Writing Material at the end of the arm (Purpose);
- 2 Material;
- 3 Length of the Arm; and
- 4 Nature of the Head Joint.

Let us further suppose the following to form part of the Schedule:

- 11 Chalking Compasses;
- 21 Wooden Compasses; and
- 37 Eighteen Inches Compasses.

Then the following (SII) are possible in [1P1]:

- 11-21 Wooden Chalking Compasses.
- 11-21-37 Eighteen Inches Wooden Chalking Compasses.
- 11-37 Eighteen Inches Chalking Compasses.

7 GAP DEVICE

In the design described in this series of articles, dependence on leaving gaps among (CN) or (BCN) or (IN) to provide for interpolation is reduced considerably.

7.1 Library of Congress Classification

The dependence on leaving gaps is greatest in LC. Its (CN) are all integers. Therefore, gap has to be left between two consecutive used integers to provide for 'Hospitality.' Any single gap has to provide for both Hospitality in Array [11] as well as Hospitality in Chain [12]. Sooner or later this leads to mix up of co-ordinate and subordinate classes within one gap. Further, whatever be the forethought with which the length of the gaps might have been decided, in course of time it happens that some gaps get choked up and make the scheme unfit for any further hospitality, while some other gaps get filled up very sparingly.

7.2 Decimal Classification

One lasting service of DC is the introduction of Decimal Fraction Notation. This takes care of subordinate classes and provides unlimited Hospitality in Chain [12]. But its limited base often makes an array of classes get completely filled up and unfit to give any more Hospitality in Array [11]. But when the number of classes in an array is smaller than the length of the base – less than 10 – DC notation can use the gap device to provide for Hospitality in Array[11]. But the gap is always small, if at all it exists and it gets choked up very soon.

7.3 CC Till Ed 6 (1960)

The notation of CC uses the Decimal Fraction Notation throughout and thus secures unlimited Hospitality in Chain [12] among (BC) and in each facet. Its Sector Device makes it possible to secure an unlimited Hospitality in Array [11] among (BC) and in each facet, in so far as extrapolation is concerned. But so far as interpolation is concerned, till ed 6 (1960), CC notation had to depend upon gap device with all its limitations. However, the very large base, with a length of 100 (if only one digit is allowed in an (AIN), of 400 (if the maximum number of digits is allowed in an (AIN) is 2, and of 1,300 if a maximum of 3 digits is allowed in an (AIN)) delays the choking up considerably. But, all the same, choking may appear and it does appear.

7.4 CC In 1963

After the postulation in 1963 of the concept of Emptying Digit and of the possibility of a digit being both Empty and Emptying, for example, U, W, Y, and Z [9], the sole dependence on Gap Device for interpolation of new (AIN) has now been obviated.

7.5 Threatened New Invasion of Gap Device and CC In 1964

In Sec 5.2 and 6.2 we saw the possibility of the array of order 1 of [1P1] having several (QIN) representing the respective first characteristics on which the universe of entities can be divided to yield the scheme for [1P1]. Suppose we have identified a certain number of first characteristics — that is (QI) — at the time of designing the Scheme *for* Classification. Suppose also that a few other first

characteristics – that is (QI) – are brought forth by newly emerging documents. These will ask for their respective places among the already existing (QI). In other words, need will arise for interpolating new (QIN) among the already existing ones in the array of order 1. All the economy and the elegance, got by allotting the different sectors and by telescoping in array the isolates based on them, would be lost if the Emptying Digits are used to carry out interpolation as prescribed in Sec 7.4. Therefore one is forced to have recourse to Gap Device in allotting the different (QI) to the different sectors in the array. In other words, some intermediate sectors will have to be left fallow at the time of designing. This danger is obviated and hospitality among the (QIN) can now be made as large as we like by bringing into use a three-digitated (AIN) for interpolation between two consecutive two-digitated (AIN). For example, if we have already brought into use the sectors (S-91) and (S-9A), we can interpolate between them the three sectors (S-99a), (S-991), and (S-99A). As and when need arises, we can provide any degree of Hospitality in Array [11] by increasing the maximum number of digits in (AIN) to 4, 5, and so on.

8 PLAN FOR FUTURE WORK

8.1 Versatility of Notation

We have thus got a very versatile notational system. It looks as if it can implement any finding whatever in the Idea Plane. It has taken nearly forty years to sense and harness the potential versatility of the mixed notation of CC and bring it into active use. It is the failure to have realized this that has been responsible for the delay in working

Design of Depth Classification

out Schemes *for Classification* – that is the schedules for all possible facets likely to go with a (BC) in ‘Engineering’, ‘Chemical Technology’, and ‘Useful Arts’ in particular and in any (BC) in general.

8.2 Frozen vs Growing Depth Classification

The present pace of industrial development all the world over caused by the current population pressure calls for the maximum conservation of the research potential of the world. This requires documentation work and service in respect of nascent micro-thought necessary as a Supporting Set-Up of Remove 1. In its turn, this requires the design of depth classification as a Supporting Set-Up of Remove 2. In documentation work, the depth classification has to organize not the universe of the frozen micro-thought of the past but with an extremely turbulent and fast developing universe of nascent micro-thought.

8.3 Differentia Between Training of Documentalist and of Generalist Librarian

Therefore, it is not sufficient if a documentalist is trained merely to use frozen, published Schemes *for Classification*. But he must be trained to design new Schemes *for Classification* and to extend them from time to time in a consistent way. Thus, the training of a documentalist has to differ essentially from the training of a generalist librarian in the public or academic sector. The training of the documentalist should include training in applied research in depth classification.

8.4 Training in Applied Research in Depth Classification

Continuing applied research in the design of depth classification by documentalists working in specialist libraries is now possible. This is so because the necessary fundamental research in the subject has now reached a stage when it can become the basis for applied research. As usual the quantity of applied research and the number of persons required for it will be far greater than those for fundamental research. Moreover, applied research is best done by persons working at the point of service. The training being given in the DRTC is turned on this fact. Documentalists with experience are also welcome to work in DRTC for short periods to get oriented to applied research of the kind mentioned above.

8.5 A Passing Mood of Despair

The Schemes *for* Classification of applied subjects available till now have been improvised rather superficially. Their roots do not run deep enough to make them stand the challenge of the new developments in the universe of nascent micro-thought. They had been put up *ad hoc* to meet the exigencies of the moment. They therefore break-down frequently. This frequent break-down has even thrown some documentalists into despair, made them question even the very utility and applicability of classification to the retrieval of micro-thought, and driven them almost to the point of abandoning classification. Some seek refuge in the possibility of subject heading giving the necessary service. This leads to failure sooner or later as the depth of the micro-thought increases. Some others try to escape

Design of Depth Classification

into machine retrieval. These would hand over the retrieval work to electronic engineers, and rest on their oars saying as it were “No more bother for us; the engineers will do all that.”

8.6 Sight of Recovery

However, a few are beginning to realize that the machine can do the work efficiently only if a considerable prior classificatory work is done by the library profession. Further, machine retrieval is not likely to become viable for many years in an individual library or even in a group of libraries. Retrieval of micro-thought and of the organization of micro-thought needed for this purpose will have to be done by the documentalist working in specialist libraries attached to research centres and industrial enterprises. They will have to depend on an efficient depth classification kept continuously up to-date by applied research. This means that applied research in depth classification by the documentalists in each specialist library is a necessity. This means that applied research in depth classification by the documentalists in each specialist library is a necessity. This is the reason for the DRTC training documentalists not only in the conventional methods of documentation work and service but also in the design of depth classification for diverse subjects.

8.7 Estimate of Work to be done

It is estimated that modern industrial and research requirements call for the design of a Scheme *for* Depth Classification for each of about 5,000 (BC). To carry out the necessary applied research in this

field, intimate knowledge is required not only of the latest methods in the design of classification but also of the developments in the wave-front of knowledge in the respective (BC). Thus, this applied research in design is best done by the documentalist working in libraries specializing in the respective (BC). This will be a continuing process. But the initial setting up of the basic scheme for each (BC) may take about a year. Thus 5,000 man years will be needed to cover all the 5,000 (BC). DRTC solicits the co-operation of the documentalists all the world over in getting the initial basic scheme for the diverse (BC) established as quickly as possible.

Bibliographical References

- Note:-
- 1 The following is the list of the documents used.
 - 2 Column 1 gives the S N of the respective documents.
 - 3 Column 2 gives the number of the Sec in the text, containing the reference
 - 4 Unless indicated otherwise, the author of the document is S R Ranganathan.
-
- 1 Sec1.2.1 BROWN (James Duff). Subject classification. 1906.
 - 2 Sec1.2.3 FOSKETT (D J). Current trends. (Foskett (D J). Modern trends in classification. P 99. *In* International Congress of Libraries and Documentation Centres (Brussels) (1955). V 2B. Communications. 1955. P 96-102).
 - 3 Sec1.2.3 Modern trends in classification. (International Congress of Libraries and Documentation Centres (Brussels) (1955). V 2B Communications. 1955. P 97).
 - 4 Sec1.4 INDIAN STANDARDS INSTITUTION, DOCUMENTATION (Committee). (Indian standard glossary of classification terms. IS: 2550-1963).

Design of Depth Classification

- 5 Sec 1.7.2 Canon of filiation sequence. (Ranganathan (S R). Prolegomena to library classification. Ed 2. 1957. Chap 16).
- 6 Sec 3.2 Compound focus. (Ranganathan (S R). Colon classification. Ed 6. 1960. Sec 053).
- 7 Sec 1.2.3 Documentation and abstract classification. (International Congress of Libraries and Documentation Centres (Brussels) (1955). V 2B. Communications 1955. P 108-13).
- 8 Sec 1.6 Emptying digit. (Ranganathan (S R). Notational plane: Extrapolation and interpolation. Sec A22. *In An lib sc.* 10; 1963; Paper A).
- 9 Sec 7.4 Empty and emptying digits. (*ibid.* Sec A24)
- 10 Sec 1.1 First sense – primitive use. (Ranganathan (S R). Classification and communication. 1951. Chap 11).
- 11 Sec 7.1, 7.2, 7.3, 7.4 Hospitality in array. (Ranganathan (S R). Prolegomena to library classification. Ed 2. 1957. Chap 22).
- 12 Sec 7.1, 7.2, 7.3 Hospitality in chain. (*ibid.* Chap 23).
- 13 Sec 3.1 Inductive vs classificatory approach. (*Annals*, Ind Lib Assoc. 2; 1952; 233-45).
- 14 Sec 1.3 Invariants in classification. (Ranganathan (S R). Classification and communication. 1961. Sec 14632).
- 15 Sec 1.2.4 Library classification as a discipline. (Proceedings, International Study Conference on Classification and Information Retrieval 1 (Dorking) (1957). P 3-12).
- 16 Sec 3.2 Postulates for classification. (Ranganathan (S R) Elements of library classification. Ed 3. 1962. Chap U).
- 17 Sec 2.2 Postulate of fundamental categories. (*ibid.* Sec H1)
- 18 Sec 1.2.4 Principles for facet sequence. (*ibid.* Chap V).
- 19 Sec 1.2.4 Principles of helpful sequence of isolates within an array. (*ibid.* Chap W).
- 20 Sec 5.7 Psychology and notational structure. (Ranganathan (S R). Documentation and abstract classification. P 111. *In International*

Congress of Libraries and Documentation Centres (Brussels) (1955).
V 2B. Communications. P 108-13).

- 21 Sec 4.1 Rich digit. (Ranganathan (S R). Notational plane: Extrapolation and interpolation. Sec A26. *In An lib sc.* 10; 1963 Paper A).
- 22 Sec 1.2 Second sense – common use. (Ranganathan (S R) Classification and communication. 1951. Chap 12).
- 23 Sec1.7.2 Space isolate. (Common isolates in documentation work. 4). (Rev doc. 24; 1957; 18-28).
- 24 Sec1.7.2 Time isolate. (Common isolates in documentation work. 3). (Rev doc. 23; 1956; 70-9).
- 25 Sec1.2.3 Tools and models. (Ranganathan (S R). Prolegomena to library classification. Ed 2. 1957. Chap 85).
- 26 Sec 2.5 Unitary class. (Ranganathan (S R). Prolegomena. to library classification. Ed 2. 1957. Sec 112).
- 27 Sec 2.5 W Universe: Portion, constituent, organ. (*Annals*, Ind Lib Assoc. 3; 1953; 1-6).

Chapter 4

SUBJECT HEADING AND FACET ANALYSIS*

S.R. RANGANATHAN

After establishing the terminology, shows how the choice of the name of the subject of a document and the rendering of the name in the heading of the specific subject entry can be got by facet analysis based on postulates and principles. After showing that subject headings constitute an artificial language, points out that using facet analysis for subject heading does not amount to using class number. Marks out the area for an objective statistical survey of sought heading for subject entry. Calls on Council for Library Resources Incorporated to provide for this project.

1 TERMINOLOGY

1.1 Subject

Thought-content of a document.

1.2 Basic Subject

Assumed term. It is usually of a large extension. It is capable of being sub-divided on the basis of one or more trains of characteristics. It admits of a book on it being written by a single author and being sought

* Abbreviations used: BTI=British Technology Index (1962); CCC=Classified Catalogue Code with additional rules for dictionary catalogue code (1958) by S.R. Ranganathan; Cutter=Rules for a dictionary catalogue (1904) by C.A. Cutter; Elements=Elements of library classification (1962) by S.R. Ranganathan; Prolegomena=Prolegomena to library classification (1957) by S.R. Ranganathan.

for its entire contents by an appreciable number of general readers or even of specialist ones. There is no unique list of Basic Subjects. Each School of Cataloguing may put up its own list – in other words, *postulate* the list. It may include a few hundreds of Basic Subjects. It is desirable not to include in the list a subject capable of separation into a Basic Subject and an Isolate.

1.3 Isolate

Assumed term. In a sense, it is not a subject by itself. But if it is attached to an appropriate basic subject, the result is a subject of smaller extension. For example, 'Gold' is an isolate.

Chemistry of Gold,	Gold Smithy,
Metallurgy of Gold,	Gold Sculpture, and
Economic Geology of Gold,	Gold Currency
Mining of Gold,	

are different subjects derived by attaching the isolate 'Gold' to different basic subjects. A number of isolates also may be attached to one and the same subject.

Prospecting for Gold,
Prospecting for Gold in India, and
Prospecting for Gold in India in 1964,

are all subjects derived from the basic subject 'Economic Geology' by attaching the isolate 'Gold' and some other isolates to it. A book embodying every possible subject derivable by attaching the isolate 'Gold' to every possible basic subject cannot obviously be written by a

Subject Heading and Facet Analysis

specialist in a single subject. It is likely to be only a Composite Book, consisting of several contributions by different specialists. And readers for the whole of such a book will be few. It is in this sense that an Isolate cannot be a Subject by itself.

1.4 Facet

A generic term for 'Basic Subject' and 'Isolate'.

1.4.1 Focus

A generic term for 'Subject', 'Basic Subject', and 'Isolate'. It may be distinguished as Overall Focus, Basic Focus, and Isolate Focus. In the Verbal Plane, the two latter may be denoted by the terms 'Basic Term' and 'Isolate Term' respectively. We may also speak of the Constituent Facets of a Subject.

1.4.2 Kernel

A constituent facet – that is, the Basic Subject or an Isolate of a subject.

1.4.3 Kernel Term

The term denoting a kernel in a subject.

1.5 Facet Analysis

Recognizing the Basic Subject and each of the Isolates of a subject, and arranging them in a preferred sequence in accordance with prescribed rules.



Neelameghan, A., Compiled by.
S.R. Ranganathan's Postulates and Normative Principles: Applications in Specialized Databases
Design, Indexing and Retrieval.
(Sarada Ranganathan Endowment for Library Science, Series 7).
Bangalore, Sarada Ranganathan Endowment for Library Science, 1997.

Copyright ©Sarada Ranganathan Endowment for Library Science (SRELS)

Digitization: Susan Ditch, SIRLS, University of Arizona

Quality Control: Megan Plesea, SIRLS, University of Arizona

Digitization Training: Han Yan, Information Systems, University of Arizona & dLIST Editor

Project Coordinator: Cheryl K. Malone, Assoc. Prof. SIRLS, University of Arizona & dLIST Editor

Digitized: Spring 2007

Acknowledgments: SRELS (A. Neelameghan, K.N. Prasad, K.S. Raghavan, DRTC) and dLIST
Advisory Board Member, S. Arunachalam (MS Swaminathan Research Foundation)

dLIST Classics (Book) Editor: Barbara Hutchinson, University of Arizona

dLIST Editor-in-chief: Anita Coleman, University of Arizona

dLIST Classics

Titles by S.R.Ranganathan

Five Laws of Library Science, Ed. 1 (1931)

Philosophy of Library Classification (1973)

Prolegomena to Library Classification, Ed. 3 (1967)

Classification and Communication (1951)

Documentation Genesis and Development (1973)

Documentation and its Facets (1963)

Library Book Selection, Ed. 2 (1966)

New education and school library: Experience of half a century (1973)

Reference Service, Ed. 2 (1961)

Other titles

S.R. Ranganathan's Postulates and Normative Principles: Applications in Specialized Databases
Design, Indexing and Retrieval, 1997. Compiled by A. Neelameghan.

Memorabilia Ranganathan: A compilation of useful quotations of S.R. Ranganathan from his various
works, 1994.

Putting Knowledge to Work: An American View of the Five Laws of Library Science, 1970. By
Pauline Atherton.

Read the dLIST Classics online!

<http://dlist.sir.arizona.edu/>



1.6 Scheme of Classes

A scheme for the arrangement of subjects in a helpful sequence.

Chapter 7.1 of the Prolegomena mentions nine well-known Schemes of Classes.

1.6.1 Scheme of Classification.

The classes in a scheme of classes — each provided with its class number — arranged according to the class numbers representing the respective classes.

Chapter 7.2 to 7.8 of the Prolegomena mention seven Schemes of Classification.

1.6.2 Classifying

Assigning a subject to its class in a preferred Scheme of Classification and determining its class Number.

1.7 Main Entry

The one entry of a document in a catalogue, giving maximum information about it.

1.7.1 Book Index Entry

An entry, other than the Main Entry, mentioning a document in a brief way.

Subject Heading and Facet Analysis

1.7.2 Specific Subject Entry

A Book Index Entry with the name of the specific subject of a document in its Leading Section.

1.7.3 See also Subject Entry

A general Index entry referring from the name of one subject to that of another.

1.7.4 Heading

The occupant of the Leading Section of an entry.

1.8 Annotation on Facet Analysis

1.8.1 Idea and Verbal Planes Only

Facet analysis is work done in the Idea Plane and in the Verbal Plane. It does not involve any work in the Notational Plane. In fact it does not require any notational system. It does not depend on any Scheme of Classification.

1.8.2 A Misconception about Facet Analysis

Some current writings disclose a misconception about Facet Analysis. They seem to imply the belief that Facet Analysis is either by itself Classification or that it is a technique designed exclusively for Classification. At any rate, Facet Analysis and Classification are taken to be inseparable. This is not correct.

1.8.3 Three Uses of Facet Analysis

On the other hand, Facet Analysis can be used for three different purposes, *viz.*,

1. Determination of Class Number;
2. Reference Service; and
3. Determination of Subject Heading.

Basing classifying work on Facet Analysis is now well-known. In reference service, the requirement of a reader — particularly of a specialist reader — at the moment is formulated pin-pointedly and expeditiously by Facet Analysis — that is, by ascertaining, through a careful conversation, the Facets of his requirement and the exact Focus in the Basic and each of the Isolate Facets respectively. The determination of the subject of a document and its subject heading has been a matter of flair all along. But it can be made in a systematic and standard way with the aid of Facet Analysis.

2 SPECIFIC SUBJECT

2.1 Choice of Heading

It is now accepted without question that any catalogue or documentation list should give a Specific Subject Entry to each document. Whether the name of the subject is made of words in a natural language (as in Dictionary Catalogue), or of the digits of an ordinal language (as in Classified Catalogue), the Subject has first to be ascertained in precise terms. In other words, the Choice of the Name

of the Specific Subject is the first step. It is desirable to base this on a Standard Procedure.

2.2 Rendering of the Heading

If the name of the subject is a single-worded term, there is no further problem. By single-worded term is meant a word comprehending each and every facet of the subject. Hardly any natural language has such a single worded comprehensive term to denote the majority of the myriads of multi-faceted subjects attracting documents today. In general, a subject has to be denoted by a multi-worded term. When put on the Leading Section of the Subject Entry, this multi-worded term becomes a Multiple Heading.

2.3 Problem of Multiple Heading

The core of the problem of multiple-heading is the sequence in which the heading — made of each of the different words or word groups is the multi-worded name of the subject — should be written in the Leading Section of the Specific Subject Entry. To take a particular example, a multiple-heading — of, say, five constituents — would admit of 120 different sequences. To give a Book Index Entry for the document under each of these 120 sequences would make the catalogue an octopus, apart from the prohibitive cost. We have therefore to select a single sequence as the Standard Rendering. At best, we may allow a small number of *See also* Subject Entries to direct attention to the entry with the Standard Rendering out of the many possible renderings and the selection of the necessary and sufficient minimum number of *See also* Subject Entries forms the core of the problem of Multiple Subject

Heading. This has always been left to the flair of the cataloguer. Occasionally, he has been helped by the flair of the author of a Dictionary of Subject Headings. We say 'occasionally' because documents embodying many new multi-faceted subjects, not covered by such dictionaries, are coming out today at a rapid rate. Therefore, flair and dictionary should be replaced by a Standard Procedure. It will be convenient if one and the same Standard Procedure is sufficient for the Choice as well as the Rendering of the Specific Subject Heading.

3 PROCEDURE OF FACET ANALYSIS BASED ON POSTULATES AND PRINCIPLES

The protean, unpredictable ways of formation of multi-faceted subjects would indicate the helpfulness of basing the determination of all the relevant kernels of a subject, and of the standard rendering of the Kernel Terms, on a set of Postulates for Facet Analysis, and Principles for Facet Sequence. By varying the postulates and the principles, we may get different procedures. One set of postulates and principles has been hit upon by the Indian School of Thought. Perhaps this is the first of its kind. More powerful or sharper standard procedures may be hit upon in future. Till then, the standard procedure based on Facet Analysis and illustrated in the next section and its subsections can liberate us to some extent from the present predicament of leaving decisions about Subject Headings of micro-documents to the individual cataloguer in a 'do your own best' kind of situation, without sufficient guidance being given to him in the form of a standard procedure.

4 STANDARD PROCEDURE FOR THE CHOICE OF THE SPECIFIC SUBJECT

The standard procedure evolved in chapters H, J, K, and N of the *Elements* will now be demonstrated by determining the kernel terms needed in choosing and specifying the name of the subject of the following document and the rendering of its name in the Heading.

Jones (D M) and Noone (T M). Some approaches to the permanent flame-proofing of cotton: Systems containing phosphorus (J. ap. Chem. 12; 1962; 349-405).

This entry is based on *BTI*, 1962;275.

The standard procedure has four steps yielding successively the Full Title, the Kernel Title, the Analysed Title, and the Transformed Title respectively of a document.

4.1 Step 1: Full Title

The Full Title is derived from the Raw Title found in the document by filling up all the ellipses in it and breaking down each composite term into its fundamental constituent elements. Usually the Basic Facet would be absent and will have to be filled up. Here it is 'Fabrics'. For, according to the Postulate of Basic Facet, every subject has a Basic Facet (See Sec H1.1 of the *Elements*). The concept of 'fundamental constituent terms' is still left to flair based on experience. We have yet to find the semantic depth to which the breaking down should be carried. Sometimes, the whole document may have to be glanced through in this process. In the document under consideration, there is no composite word to be broken down. But the following terms are to be supplied to

fill up ellipses: Fabrics and Finishing. The Full Title of the document is then

Some approaches to the permanent flame-proofing in the finishing of cotton fabrics: Systems containing phosphorus.

4.2 Step 2: Kernel Title

The Kernel Title is derived from the Full Title by removing all the auxiliary terms and puffs and replacing all the surviving essential terms in their respective nominative forms. The Kernel Title of the document is then Flame-Proof. Finishing, Cotton, Fabrics & Phosphorus.

4.3 Step 3 : Analysed Title

This step is convenient, but not essential. Step 3 and Step 4 may be merged into a single step. Step 3 is illustrated in what follows. The Kernel Term 'Finishing' is first spotted as denoting the Energy Facet of the subject. Then, the position of each facet other than the Basic Facet, is determined one by one, relative to the Energy Facet. This is done with the aid of the Wall-Picture Principle, which is, 'if two facets A and B of a subject are such that the concept behind B will not be operative unless the concept behind A is conceded, even as a mural picture is not possible unless the wall exists to draw upon, then the facet A should precede the facet B'. (See Sec N 3.2 of the *Elements*.)

As between the concepts of 'Cotton' and 'Finishing', the latter concept does not become operative (will not be thought of), unless the

Subject Heading and Facet Analysis

latter concept is conceded (had been already thought of). Therefore 'Phosphorus' succeeds 'Finishing'.

As between the concepts of 'Flame-Proof' and 'Finishing', the former concept does not become operative (will not arise) unless the latter concept is conceded. Therefore, 'Phosphorus' succeeds 'Flame-Proof'.

4.4 Step 4: Transformed Title

Thus the following Transformed Title of the Document is derived from the Kernel Title:

Fabrics. Cotton. Finishing. Flame-Proof. Phosphorus.

We have now completed the Choice of the Name of the Specific Subject of the document. This name is a term with five words should be used to satisfy the Canon of Individualization. (See Sec. 0.2.3 of CCC.)

5 RENDERING OF NAME OF THE SPECIFIC SUBJECT IN HEADING

Then comes the rendering of the multi-worded name of the subject chosen for the Heading. As already stated in Sec 2.2, there are 120 possible arrangements of the five words in the name. Therefore, there are 120 ways of rendering the name of the specific subject. We have to choose one and only one of them. For convenience of reference, let

us call the sequence arrived at in Sec 4.4 the 'Forward Rendering'. An obvious alternative is the reverse one:

Phosphorus. Flame-Proof. Finishing. Cotton. Fabrics.

Let us call this 'Reverse Rendering'.

5.1 Choice of First Heading

Let us first concentrate on the first heading. Only five first headings are possible. In the Forward Rendering, the first heading is 'Fabrics'. In the Reverse Rendering, the first heading is 'Phosphorus'. We can also have three other first headings, viz. 'Cotton', 'Finishing', 'Flame-Proof'. The first problem is the choice of one of these words for the first heading. This we can determine on the basis of one or other of two of the Canons of Cataloguing given in CCC.

5.1.1 Canon of Prepotence

The Canon of Prepotence is 'The Principle:

1. that the Potency to decide the position of an entry among the various entries in a catalogue should, if possible, be concentrated totally in the Leading Section;
2. that it should be concentrated, as much as possible, in the entry element; and further
3. that, if total concentration in the Leading Section is not possible, the minimum possible potency should be allowed to overflow beyond it to later sections'. (See Sec. 0.2.20 of CCC.)

Subject Heading and Facet Analysis

The following table gives the potency of each of the five first headings, as based on the *BTI*.

First Heading	Number of Entries	Potency
Fabrics	200	1/200
Cotton	37	1/37
Finishing	71	1/71
Flame-Proof	2	1/2
Phosphorus	8	1/8

Left to itself the Canon of Prepotence will therefore choose 'Flame-Proof' as the first heading. If the reader should be helped to have the least number of entries to be looked into in order to spot out his documents, 'Flame-Proof' will have the greatest claim to be the first heading of the specific subject entry of the document; and 'Fabrics' will have the least claim.

5.1.2 Canon of Sought-Heading

But we should consult also the Canon of Sought Heading. (See Sec 0.2.40 of *CCC*). This Canon will determine the first heading in the light of the answer to the question 'Are the majority of the readers likely to look for the document under "Fabrics", or under "Cotton", or under "Finishing", or under "Flame-Proof", or under "Phosphorus" as the first heading? What is the relative frequency of each of these terms being sought as the first heading'.

5.1.3 Statistical Survey or Opinion

The answer, to these questions, can be based on an objective statistical survey or on conjecture based on experience. The former basis is not available. Conjecture is the only recourse. In reality it degenerates into opinion. Here comes the apple of discord.

5.1.4 Only Two Schools of Thought

At present there are only two dominant Schools of Thought – one choosing the first heading as in the Forward Rendering, and the other choosing the first heading as in the Reverse Rendering. *BTI* prefers the Forward Rendering. The School of Specific Subject Heading prefers the Reverse Rendering. The dispute between these two Schools of Thought should be decided either by a statistical study or by casting lots. Otherwise, to carry on the discussion on this issue will amount only to an indulgence in polemic. The polemic may have really to range not only over these two Schools of Thought; but it can also range over any number from five to 120 Schools of Thought if the name of the subject has five kernel terms, and if the choice of the second and the later headings also are to be open to all possible variations.

5.1.5 Omission of First Term

Can we omit the first or the first few kernel terms? In the case of the Reverse Rendering, the First Term should never be omitted. In the case of the Forward Rendering, the answer to the question will depend upon whether the library or the documentation list is for the generalist reader or for the specialist reader in the subject field 'Fabrics' or in the subject field 'Cotton Fabrics'. If it be a documentation list prepared on

'Cotton Fabrics', the two headings 'Fabrics' and 'Cotton' need not be used at all. If it be a documentation list on 'Fabrics', the heading 'Fabrics' need not be used at all. On the other hand, if it is a general documentation list such as the BTI, the first heading should be 'Fabrics'.

5.2 Omission of Last Term

In the case of the Forward Rendering, the last term should never be omitted. In the case of the Reverse Rendering, it may be possible to omit the last one or more terms in certain contexts. Strict rules for this have not yet been provided.

5.3 Omission of Intermediate Term

With only a limited number of documents, it may look as if some of the intermediate terms might be dropped. For example, in the subject under consideration, it may look as if the heading 'Finishing' might be dropped whether the Forward Rendering or the Reverse Rendering is adopted. But one has to beware of the Fifth Law of Library Science. For in that case, any day two different documents may appear with the same Specific Subject Heading but embodying two different subjects. The only way of distinguishing them will then be the restoration of the missing link 'Finishing' in one case and some other term in the other.

6 CHAIN PROCEDURE

Once the choice and rendering of the Specific Subject Heading of a document is established on the basis of the preferred set of Postulates for Facet Analysis and of Principles for Facet Sequence, the necessary

and sufficient minimum number of *See also* Subject Headings are derived from the Specific Subject Heading mechanically with the aid of Chain Procedure. (See Part 3 of *CCC*).

6.1 Headings for *See also* Subject Entries

Whether we use the Forward Rendering or the Reverse Rendering, the necessary and sufficient minimum number of *See also* entries is four for the document considered. In the case of the Reverse Rendering, these headings will be:

1. Flame-Proof, Finishing, Cotton, Fabrics;
2. Finishing, Cotton, Fabrics;
3. Cotton, Fabrics; and
4. Fabrics.

These entries will help the minorities of readers seeking the document in question through a first heading other than 'Phosphorus'. In the case of the Forward Rendering, the headings for the four *See also* entries will be:

1. Cotton. Fabrics;
2. Finishing. Cotton. Fabrics;
3. Flame-proof. Finishing. Cotton. Fabrics; and
4. Phosphorus. Flame-proof. Finishing. Cotton. Fabrics.

Subject Heading and Facet Analysis

These entries will help the minorities of readers seeking the document in question through a first heading other than 'Fabrics'.

6.2 Versatility of Chain Procedure

The Rules of Chain Procedure can be so framed as to implement any kind of decision about Sought First Heading and the other successive headings in conformity with the Principle of Local Variation (See Sec. 0.3.5 of CCC).

6.3 Common Practice

For example, in some catalogues and documentation lists, the Common Isolate Terms — denoting either Approach Material, or Common Personality or Common Energy Isolates, or Property, or Geographical Area, or Year — occurring last in the Forward Rendering are usually disqualified from eligibility to be First Heading either in a Specific Subject Entry or in a *See also* Subject Entry. For example, consider a document which embodies the Bibliography of the document we have considered so far. Even in the Reverse Rendering, the Heading of the Specific Subject Entry will be only Phosphorus. Flame-Proof. Finishing. Cotton. Fabrics. Bibliography; and not Bibliography. Phosphorus. Flame-Proof. Finishing. Cotton. Fabrics.

There will not be even a *See also* Subject Entry with the latter heading. *See also* Subject Entry with Bibliography as the First Heading will not be given, even if the Forward Rendering is used for the Specific Subject Entry.

6.4 Special Practice

On the other hand, it may be useful to make each of the common isolate terms eligible in certain kinds of documentation list and even in the catalogue of a library such as the one specializing in regional studies. *My Union catalogue of learned periodicals in South-East Asia* follows this practice for special reasons.

6.5 Decision in the Idea Plane

Whatever it be, the decision should first be made in the Idea Plane. Then, the Rules of Chain Procedure can be framed so as to implement in the Verbal Plane, the policy adopted in the Idea Plane.

7 ARTIFICIAL LANGUAGE

7.1 Artificial Syntax

The headings of subject entries taken together form an artificial language. No doubt the words used are taken from a natural language. But the syntax is one of position and is different from the syntax of the natural language from which the words are taken. It is well known that syntax varies from one natural language to another. Formal linguistics and psychology do not appear to have as yet discovered the natural syntax – ‘Patterns of linguistic thought and expression’ – independent of the natural language used. Therefore, the sequence of the words in a multiple subject heading – that is, the syntax of the artificial language of subject headings – is not yet known to be unique. But all the same we have to evolve a consistent syntax for it. The Postulates of Facet

Subject Heading and Facet Analysis

Analysis and the Principles for Facet Sequence yield one possible syntax – that is, for the arrangement of the words in a heading and the meaning depending on it.

7.2 Punctuation Mark

The symbols such as “,” and “.” inserted between the successive headings in a multiple subject heading belong to the artificial language and their function and purpose are determined by the rules of that language. It is for this reason that CCC has some special rules on punctuation marks. To pick up the thought embodied in a language – natural or artificial – it is not always necessary to be thoroughly well-versed in every detail in the grammar of the language. The very subject-context helps one to follow the expression and the thought conveyed by it.

8 REFLECTION

8.1 Wrong Assumption

The Postulates for Facet Analysis and the Principles for Facet Sequence are of as much help in finding out the names of the Specific Subject of a document and in rendering it in the Subject Heading, as they are in establishing its Class Number. The tasks of cataloguing and of classifying are equal beneficiaries of these postulates and principles. The use of one and the same procedure in cataloguing and in classifying does not warrant the assumption of subject heading being derived from class number or of class number being derived from subject heading.

Of two different branches A and B of one and the same tree, we do not say either that A stems from B or that B stems from A.

8.2 The Outstanding Problem and its Difficulties

The School of Specific Subject Heading would trace its origin to Cutter. The commentaries on Sec 1.6.1, 1.7.4, 1.7.5 of Cutter are as penetrating and forward looking as the commentaries on most of its other sections. Cutter was finalized in 1891. In that distant age, Cutter contemplates only books of that day and neither the multifaceted books nor the micro-documents of today. Even so, the choice of the name of the subject of a book was nebulous. Its rendering was equally so. No definite procedure was possible. Now, we are in a better position. The only task to be done is the Statistical Survey of the reader's approach to the First Heading as well as the sequence of the other headings in a multiple subject heading. This survey is not easy. For it will have to reckon, not with a freely thinking population, but with one mentally conditioned by the cataloguing practices – good or bad – to which they had been all along exposed. The Survey Project should devise methods to get behind such mental conditioning.

8.3 An Appeal

My appeal goes to the Council for Library Resources Incorporated. It has monetary resources. Its Head is keen to pursue such difficult statistical projects. I know that Verner Clapp has already dealt with several problems. Some of them had arisen out of our mutual contact. I commend this Statistical Project on Multiple Subject Headings to his fostering care.

Chapter 5

S.R. RANGANATHAN'S GENERAL THEORY OF KNOWLEDGE CLASSIFICATION IN DESIGNING, INDEXING, AND RETRIEVING FROM SPECIALIZED DATABASES

A. NEELAMEGHAN*

The principles and postulates of S.R.Ranganathan's General theory of Knowledge Classification, and the implied freely-faceted analytico-synthetic methodology provided a firm foundation for the design, development and use of subject classification, indexing, preparation of vocabulary control tools, and search and retrieval of information in conformity with his Five Laws. This paper summarizes some of our experiences of the use of these normative principles in the design, development, and retrieval of information from machine-readable databases, specially object-oriented specialized databases, including the design of user-interfaces and hypertext links. Enumerates some of the earlier instances of the facet method in machine-based systems beginning with Hollerith's punched card system for the U.S. Census data processing. Elaborates on Ranganathan's holistic approach to information systems and services provided by his normative principles. Mentions several of the similarities in the concepts and approach to database design and those of faceted classification system design. Examples from working systems are given to demonstrate the usefulness of selected canons and principles of classification and the analytico-synthetic methodology in database design.

KEYWORDS/DESCRIPTORS : Information systems; Database design; Ranganathan S.R.; Normative principles; Knowledge organization; Analytico synthetic methodology

* Hon. Visiting Professor, Documentation Research and Training Centre, ISI, Mysore Road, 8th mile, Bangalore 560 059
Contact address: 216, 16th Cross Road, 4th Main Road, Bangalore 560 055, India. e-mail: nndrtc@isibang.ernet.in

1 SCOPE OF THE PAPER

This paper summarizes experiences of the application of the postulates and principles of the General Theory of Knowledge Classification, the freely-faceted approach, and the analytico-synthetic methodology, the foundations of which have been laid by S.R. Ranganathan, to the design and development of object-oriented specialized databases, indexing, user-interfaces and retrieval from such databases [1]. Some of the earlier instances of use of the faceted method in machine-based systems, for example, Hollerith's punched card system for the U.S. census data processing are mentioned. Elaborates on Ranganathan's holistic approach to information systems and services, design and development guided by his normative principles.

The database systems given as examples are mostly operational systems developed using Unesco's Micro CDS-ISIS software.

2 RANGANATHAN'S HOLISTIC INTEGRATIVE APPROACH

2.1 Normative Principles and Postulates

In discussing Ranganathan's approach to the design, development and use of information systems and services, it is useful to bear in mind his holistic approach. With a view to securing consistency and integrity in designing and developing information systems, he proposed a range of normative principles [51,64], a typology as well as a hierarchy among them, and formulated guidelines for their application. The typology of Ranganathan's normative principles may be summarized as follows:

General Theory of Knowledge Classification

Level 1 : Basic Laws (Law of Symmetry, Law of Impartiality, Law of Context/Local Variation, Laws of Interpretation and Law of Parsimony), applicable to intellectual work in all areas of the universe of knowledge;

Level 2 : Fundamental Laws, (e.g., Laws of Physics, Laws of Biology, and the Five Laws of Library Science postulated by Ranganathan) applicable to a discipline as a whole;

Level 3 : Canons, (e.g., Canons of Classification in Library Science) applicable to a branch of a discipline; and

Level 4 : Principles (e.g., Principles of Facet Sequence in Classification), applicable to a subdivision of a branch of a discipline.

Additionally, Ranganathan provided several postulates for use at one or more of the above levels.

The principal functions of the normative principles is to guide research and development from the micro to the macro level, to serve as criteria for assessing the helpfulness of findings or innovations in library and information systems and services. Any problem or conflict that may arise, for example when more than one principle appears applicable at one level (e.g., the Canons of Classification), it may be resolved by applying the principles of higher levels (e.g., Laws of the subject and Basic Laws).

2.2 Focus on Information Use and Users

Ranganathan's *Five Laws of Library Science* [48] are information use and user-oriented and imply that any new principle, technique or strategy adopted should enable users obtain more pertinent, a wider range and timely information, presented in convenient forms so as to

save users' time and effort, and ensure a more efficient management of the information systems and services in question, than was the case prior to the application of the new technique, strategy, theory, or principle.

3 NEW INFORMATION TECHNOLOGIES

3.1 Library is a Growing Organism

“Library is a growing organism” is Ranganathan’s Fifth Law of Library Science. Here the term “Library” denotes the species of institutional forms created by society to manage information resources to provide information to people, in other words for the institutionalisation of information services. Its internal system is composed of information sources, personnel, information processing methods and techniques, facilities and equipment, and its users, outputs such as information products and services, together with its organizational structure and management, in other words, the hardware, software, brainware and firmware.

The term “growing” is to be interpreted not only as change in physical dimensions but also in the sense of *biological growth* signifying a living organism, renewing constantly its internal system cells, evolving and adapting itself to the environmental changes. Information systems exhibit attributes of physical and biological systems [1].

Like any structure created by society, an information system, such as a library, is affected by changes in the institutional, local, national and international environment in which it operates. The environmental

General Theory of Knowledge Classification

elements include political, social, cultural, technological, regulatory and other influences. Hence, information institutions must have their antenna constantly fine-tuned so as to anticipate the emerging trends and changes in the environmental elements and their likely impact. They will then be able to (a) help people cope with the societal changes and shape the future; and (b) adapt themselves to those changes conducive to the optimal performance of their functions *vis-a-vis* the individual and the society. New information technologies (NITs) enhance the capacity of information systems to access a wider range, quantity, and forms of information and enable provision of a variety of information services and products to meet the needs of people to cope with change and thus satisfy Ranganathan's Laws of Library Science.

Ranganathan defined 'book' as a trinity consisting of a 'soul', that is, the ideas embodied; a 'subtle body', that is, the language and mode of exposition and presentation of the ideas; and a 'gross body', that is, the medium in which the ideas are recorded and/or transmitted. The NITs have not only brought about many changes in the gross body, and in the form of presentation (text, graphics, images, audio, video, or combinations of these, and animation), in the methods of transmission of the messages, but also enable manipulation of the presentations with greater ease. But remember what people need and seek are the ideas and messages embodied (soul). Thus, Ranganathan's normative principles are valid in the changing information and communication technologies environment as well.

3.2 Value-Added Information

In the familiar information hierarchy triangle (Fig. 1), as one moves up the hierarchy closer to end-users — decision makers, planners,

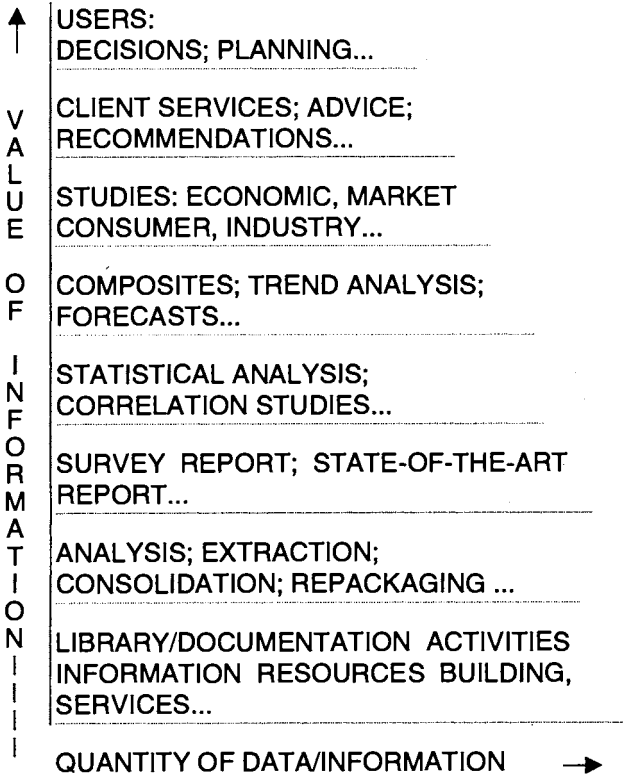


Fig. 1

executives, researchers — the type of information and information products needed are of the value-added type, that is, analyzed, synthesized, evaluated and repackaged information, and there is a growing demand for these. Bibliographies in many subject areas are now more widely available on national, regional and global networks, online database services and CD-ROMs, but value-added products of local or institutional interest on specific objects need to be prepared by information specialists serving specific user groups at the institutional or local level.

General Theory of Knowledge Classification

NITs, among other things, can relieve library and information professionals of a good proportion of the routine and repetitive library tasks, such as, cataloguing and circulation, and also enable them perform other management functions more efficiently. Therefore, information professionals can devote more time and effort at providing value-added products and services. Provision of these calls for the design and development of specialized databases of the object-oriented type forming the basis for preparing value-added products, and knowledge bases for use in expert systems, intelligent information systems, decision support systems, etc. Such databases and value-added products can be developed with some of the widely available low-cost software, and the application of the analytico-synthetic methodology and the principles and postulates of the General Theory of Knowledge Classification. Until recently these postulates and principles, and the techniques of Facet Analysis and Analytico-Synthetic Approach have mainly been applied to bibliographical work and there again largely in the preparation of manual catalogues and bibliographies, indexes, subject headings, etc. Their application to machine-readable data-bases, particularly to object databases, knowledge bases and expert systems are more recent [37, 38].

With the availability of low-cost powerful micro-computers and DBMS packages in the 1980's and the need to develop specialized object databases and information products thereof at the institutional level is becoming more widespread.

The present report is largely based on the following earlier reports:

“Design and development of object-oriented databases: application of the principles and postulates of the general theory of knowledge classification”. A. Neelameghan, assisted

by Rahel Bekele. Addis Ababa: School of Information Studies for Africa, Addis Ababa University, 1994 April. [unpublished]

“Application of Ranganathan’s general theory of knowledge classification in designing specialized databases”. A. Neelameghan. *Libri*. 42 (1992): 202-226.

“Home-grown object-oriented specialized databases”. A. Neelameghan. In *NIT’92 5th International Conference New Information Technology ... Hong Kong, November 30 — December 2, 1992. Proceedings* / edited by Ching-Chih Chen. Boston: MicroUse, 1992. p. 252-284.

“Concept categorization and knowledge organization in specialized databases: a case study”. A. Neelameghan *International Classification* 18(1991): 92-97

4 WORKING DEFINITIONS

Working definitions of selected terms used in this report are given below [35,51]:

4.1 Entity

An existent, for example, an object, an event, a process, a phenomenon, etc., in the real world or what is imagined in the abstract.

4.2 Idea

An idea is a product of thinking, reflecting, imagining, etc. about an entity, formulated by the intellect by integrating with the aid of logic, a selection from the apperception mass, and/or what is directly apprehended by intuition. In the framework of this definition, the

General Theory of Knowledge Classification

meaning of the terms 'concept', 'thought', and 'notion' is each taken to be near-synonymous with the term 'idea'.

4.3 Knowledge

Knowledge is the totality of ideas conserved by human beings at any point in time, and in this sense, the term 'knowledge' is synonymous with the term 'universe of ideas'.

4.4 Subject

A subject is an organized or systematic account of an entity or group of entities (an idea or group of ideas) whose extension and intension are likely to fall coherently within the field of interest and comfortably within the intellectual competence and the field of inevitable specialization of a normal person. For example, 'Housing needs of metropolitan areas in transition', 'Planning for metropolitan areas', 'Town planning' and 'Planning' is each a subject. For, the extension and intension of the totality of ideas comprehended by each of the combination of terms can form a coherent, comfortable and convenient field of specialization of a normal person. On the other hand, not all that is embodied in the *Encyclopaedia Britannica* taken as a whole can be deemed to be a subject as the totality of those ideas cannot be deemed to form a coherent, comfortable and convenient field of specialization of a normal person. It is a collection of descriptive accounts of a large number of subjects.

4.5 Discipline

A discipline is a branch of learning; or a subject that is taught; or a field of study. For example, Mathematics, Physics, Philosophy, Economics, and Sociology [73].

4.6 Data Entity

A data entity is a set of objects that share in common among themselves one or more attributes and about which a person or an organization may be interested in collecting descriptive data for use. The descriptive data thus collected is usually data on some attributes of the entity or 'attribute data'. For example, an urban planner collects descriptive data about the attributes of the spatial unit for which the plan is to be prepared, such as, attributes of the population, housing facilities, transportation facilities, educational facilities, etc. A company may collect descriptive data about market entity, such as, location, composition, competing products, distance from warehouse, etc., to enable proper planning of product promotion, marketing, pricing policy, and so on. Again, a hospital and the medical doctors record data about patient entity, such as, name, age, sex, family history, social history, patient's complaints, disease symptoms, clinical investigation results, etc., so as to facilitate diagnosis and treatment of the patient's illness, or other management action. Data entities may be related in one-to-one, one-to-many or many-to-many relationship.

4.7 Data Model

A data model is a schema to represent some real world entity using information concepts and structures. For example, in a database on Schools, the entity "school" may be represented by data attribute concepts, such as 'Level' (nursery, primary, secondary, etc.), 'Size' (less than 500 pupils, over 500 but less than 1500 pupils, etc.), 'Teacher-Pupil ratio' and so on. In a database on computers, the entity "computer" may

be represented by such data attributes as Type, Model, RAM size, Access time/Speed, Number of drives, Number of ports, Mass storage device, Monitor etc.

Hierarchic, network and relational models represent stages in the evolution of data models with increasing capability for representing or handling the different kinds of relations among data entities. A database may conform to one or the other of the data models, although in an operational database there may be a mix of elements of two or more models.

5 ANALYTICO-SYNTHETIC FACET TECHNIQUES

5.1 In Mechanized Information Systems

In normal persons, the facet approach is natural to the processes of learning and knowledge gathering about an entity, at the conceptual as well as practical planes. Application of facet analysis and structured representation of data/information about an entity in mechanized information storage and retrieval systems is not new. In the 1890s Herman Hollerith used holes punched in a card to represent or code respective attributes (e.g., age, sex) of each person (the object of the census) and used the same codes to retrieve the required data, for example, number of females, or number of persons above 60 years of age. In this manner Hollerith was using facet techniques though not explicitly stated so. The state-of-the-art of such techniques in mechanized information systems in the 1950's and early 1960's was presented in the papers contributed to the International Conference for

Standards on a Common Language for Machine Searching and Translation, Cleveland, Ohio, 1959 [31], the FID/IFIP Joint Conference, Rome, 1967 [56], and other documents. Three papers by Ranganathan and one by S. Parthasarathy contributed to the Cleveland Conference dealt with faceted classification and analytico-synthetic methodology. In the paper *Natural, Classificatory and Machine Languages*, Ranganathan discussed analytico-synthetic method, and the postulational approach, pointing out the advantages of using classificatory language as a bridge language, it being "more economical to make the machine translate this by-product of isolate numbers into the code symbol of the machine language than to make facet analysis independently and translate directly from natural language." In *Classification and Retrieval- Problems for Pursuit*, Ranganathan identified several problems relating to facet analysis that "are of vital interest to work in machine search." In their papers *Classifying, Indexing, Coding and Faceted Classification as an Approach to Machine Coding*, by Ranganathan and Parthasarathy respectively, suggest the use of faceted class number for machine coding and in search.

Information storage and retrieval systems in the engineering fields, especially on equipment and machines, involved design of object-oriented information bases, exemplifying in some cases the facet approach. A good example of an object-database design using a facet approach is the report on an information retrieval system on Internal Combustion Engine, by Kinzo Tanabe, presented to the above Cleveland Conference [65]. Researches on designing information systems, indexing, concept representation and related areas, carried out at the Center for Documentation and Communication, School of

General Theory of Knowledge Classification

Library Science, Western Reserve University, Cleveland, during this period are noteworthy contributions to the analytico-synthetic approach in mechanized systems. During this period, the use of faceted schemes of classification, such as the UDC (Universal Decimal Classification), in computerized systems is exemplified in the work of Freeman and Atherton [20,21] and K. Schneider [59] and in some of the papers contributed to the International Symposium on UDC in Relation to Other Indexing Languages, 1970 [27]. Schneider's design of information systems to support cancer research, exemplified the analytico-synthetic approach in computerized systems.

In the chemical sciences, information systems describing chemical substances by their physical, chemical, and structural attributes and searches in the related factual databases using a combination of these attribute names also exemplify the faceted methodology. The GREMAS system described by Fugmann (1983) [22] is an example. Also, information systems in pharmacology and medicine have used similar techniques and tools. Databases of case records of hospital patients, land records, etc. use facet methodologies in developing object databases.

Object databases, on a variety of objects – gene banks, plants, national parks, museum objects, rocket engines, and rocket launch vehicles and other machinery – have been developed. The data structures in most of these systems manifest the faceted approach.

Ranganathan's work, beginning with the design of the Colon Classification in the 1920's (and its successive editions) [46], the formulation of a philosophy of classification [50], the analytico-synthetic

methodology based on the postulates and principles of his General Theory of Knowledge Classification [51] and elaboration of the Chain Indexing [54], provided for a sound theoretical basis for the organization of subjects/knowledge and for analytico-synthetic facet methodology [64].

In the late sixties and the seventies, significant developments have been reported on indexing systems (pre-coordinate and post-coordinate indexes) and vocabulary control methods and tools such as, classification schemes and thesauri some of which are based on facet analysis. Work of the CRG (Classification Research Group) in U.K., and more particularly researches relating to the classification and indexing aspects of the *British National Bibliography*, the *British Technology Index*, development of PRECIS [04], the Thesaurofacet, etc. are noteworthy contributions [02]. Ranganathan's facet methodology and the associated analytico-synthetic approach had an influence on the work leading to these developments [45].

In 1963, Ranganathan formulated a methodology for the in-depth classification of micro-subjects (specialized and complex subjects with many facets and component isolates) based on the systematic application of the postulates and principles of the General Theory of Classification and analytico-synthetic method [47]. His associates at the Documentation Research and Training Centre in Bangalore, India, designed and developed schemes for in-depth classification and indexing of micro-subjects in a number of fields [57].

In the late sixties, a computer-based method for synthesis of faceted class numbers given the facet-analyzed terms of a subject representation and for translation of a faceted class number to faceted

feature-headings (subject headings) from which KWOC and pre-coordinated indexes could be generated, was developed [44,71].

5.2 Indexes

Information storage and retrieval systems find it useful to provide alphabetical subject indexes as a means of fast access to the information contained in databases. The subject index may be a part of a single inverted file (author, title, subject, etc.) or a separate file. The helpfulness of the subject index entries to users depends in a large measure on how adequately the subject content (ideas) of each document (if a bibliographical database) or the data about the entity described (if an issue or object-oriented database) has been analyzed and represented precisely and comprehensively in the index structures/expressions. In this work, the analytico-synthetic procedure can be useful as already indicated above. Chain Procedure is one such technique based on the analytico-synthetic methodology [54].

In the first formulation of Chain Procedure in 1937, the subject index entries were derived from the corresponding Colon Class Number. As the latter was based on facet analysis of the subject of the document, each component idea and to some extent the nature of relationship among the ideas, could be represented in the class number. Hence, the subject index entries could be made to represent, as precisely as desired, the specific subject(s) of a document. But when Chain Procedure is applied to a class number that does not represent all component ideas of the subject, it leads to subject index entries that are less precise and therefore less helpful in retrieval.

In a paper, Ranganathan [53] clarified that the work of facet analysis of the subject of a document and expressing it precisely by analysis and synthesis of the component ideas using relevant postulates and principles for mapping the specific subject onto a data model for linear representation, must be done before a class number is assigned according to a scheme for classification and that the expression of the specific subject could be used to derive subject index entries. Thus, derivation of subject index entries was clearly separated from the class number. POPSI or Postulate-based Permuted Subject Indexing, explicitly uses facet analysed subject strings based on the General Theory of Knowledge Classification, to derive pre-coordinated index entries [06-10].

Using precise expression(s) of the subject(s) or information on the object or issue discussed, a variety of index entries of the KWOC, PRECIS, and POPSI type can be computer-generated. Work on POPSI led to techniques for the computer-generation of various types of indexes and vocabulary control tools, such as, thesaurus and classaurus from the faceted strings [15-18,32,33, 62,63].

The facet analyzed expression of subject can be used in the database records instead of, or in addition to, descriptors. In query analysis and information retrieval the helpfulness of the analytico-synthetic approach and facet analysis has been demonstrated [24,26,28-30,72]. Experiments on knowledge-based systems, expert systems, and user interfaces have shown the use-fulness of facet analysis and the analytico-synthetic methodology in such systems [26,70, 71]

6 DATA MODELS AND CLASSIFICATION MODELS

6.1 Similarity

It is noteworthy that there is a parallel in the evolution of classification/index language models (e.g., hierarchic, faceted, relational, freely-faceted) with the evolution of data models (hierarchic, network and relational). The main reasons for their development are more or less similar. Rapid and new developments in the universe of knowledge not only give rise to new ideas and subjects but also change the relationships and give rise to new relationships among existing ideas and subjects. These call for interpolation of new intermediate divisions in existing hierarchy of concepts, without restructuring the whole system of representation of ideas in a scheme for classification or data structure. Further, a concept may be related to other concepts in different hierarchies; and a variety of non-hierarchic associative relationships among concepts should also to be provided for [51]. In this regard, the postulates and principles of the General Theory of Knowledge Classification used in the semantic categorization and linking of categories according to the type and strength of the relationship or strength of bond, are helpful in designing databases for specific knowledge areas (i.e. specialized databases) [37,38,43].

6.2 Relational Database Model

There are features common to the relational data model and the structured representation of subject/information about entity according to the analytico-synthetic facet model.

The relational data model concept developed by Edgar F. Codd [14, 23] has its basis in set theory. In the relational model (RM) entities are represented in the form of tables, the table being named a 'relation'

in mathematics and hence the name of the model. Generally, a relational database consists of three parts: the structural part defines relations of data and their interrelations. The integrity part assures that each occurrence of a relation is unique. The manipulative part provides operators for processing relations [11]. Structurally, the relation consists of columns and rows. The column part of the relation refers to the entity attributes. Each of the attributes has values selected from its permissible domain (for example, for the attribute gender the permissible domain may be 'female' and 'male'). The value of the attributes form a number of rows in the table. A single row, usually known as a 'tuple' (or n-tuple) in the RM, represents a distinct entity as shown in the following table for Population:

Name#	RHHH*	Sex	Age	Religion	Sect
T. Raghavendra	Head	Male	63	Hindu	Brahmin
T. Rohini	Wife	Female	57	Hindu	Brahmin
T. Srinath	Son	Male	22	Hindu	Brahmin
Abdul Ahmed	Head	Male	52	Moslem	Sunni
Nuriya	Wife	Female	48	Moslem	Sunni
Halim Abdulla	Son	Male	21	Moslem	Sunni

*RHHH : Relation to head of household

Similarly, for relational table for Farm For esty:

Forest div.#	Species	Area (hec)	Yield (ton)	Value (000\$)	Expend (000\$)	Year
Palaghat	Teak	36	250	5000	235	1984
Trichur	Bamboo	27	150	325	125	1984
Idukki	Eucalyp.	28	145	1234	345	1984

The relation is a collection of similar tuples or a subset of the cartesian product of the underlying domains.

General Theory of Knowledge Classification

A relation scheme is a list of relation name and attributes. Thus, the relation scheme for Population can be:

Population scheme (name:string, RHHH:string,sex:string,
age:integer,religion:string, sect:string)

The relation scheme for Farm Forestry can be:

Farm forestry scheme (Forest divn. : string, species:string,
area: integer, yield : integer, value:
integer, expenditure: integer, year:
integer)

Each occurrence of entity type in the relation should be uniquely identified. An attribute that has no repeated value on different tuples in the relation can be considered as an entity identifier. Such an attribute used for uniquely identifying each entity in the relation is denoted as a 'Primary key', usually positioned in the first column of the relation (marked with an #) and has a unique value. On the other hand, a non-key attribute can have repeated value.

In the above table the attribute 'name' representing a person is deemed a primary key in the population relation. The primary key has a dominant role in the relation in that if the value of the key attribute is changed/deleted, all of its other values should be changed or deleted from the relation as the case may be.

In a faceted structure representing the information about an entity, especially according to the postulational approach, the Primary key will be equivalent to the Personality isolate or the Object of Study, and the data attributes will be equivalent to Property isolates, giving rise to Entity-Property relation. The generalized subject structure model (See

Chapter 2, page 52) recognizes also other types of relations between the Primary Key (Personality isolate/Object of study) and the associated concepts.

6.3 Object-Oriented Analysis and Design

The approach to conceptualizing reality in Object-oriented Analysis and Design (OOAD) of databases has several features similar to that of the process of conceptualizing reality for designing a scheme for classification for a specific subject or information about an object [05, 58].

In OOAD the computer is adapted to the characteristics of the problem unlike in the procedural approach. Such analysis is independent of programming languages. This is also the basis of the postulate-based analytico-synthetic approach.

The basic construct in OOAD is objects. An 'object' represents or corresponds to a physical or conceptual entity in the real world considered in the problem. For example, in simulating queuing at a manually operated circulation desk in a library, the users, the documents borrowed, the documents returned, the personnel at the circulation desk, the book card, the borrower identification, etc. can be the objects, some of whose attributes may be relevant to the study. Also, the library's policy and rules regarding circulation of different kinds of materials, over-due charges and other aspects may also be objects of the problem. The entities share one or more attributes in common among themselves. The attributes, including behaviour and functions, are modelled and represented by an object in the programme. Which attributes are to be taken into consideration depends upon the problem (purpose) in question. Appropriate attributes of objects are used or evoked by sending a message to the objects. Some basic features of

objects in OOAD are *identity, classification, inheritance* and *polymorphism*. These are also basic concepts in classification theory and practice.

'Identity' refers to the delimitation of boundaries and distinguishing each object from all others considered in the problem while each may share one or more attributes in common with other objects. (see Sec. 6.4) For example, consider duplicate copies of a book in a library. Attributes, such as name of author, title, subject, date of publication, etc. will be common to all the copies of the book (object). But physically they are distinct objects and should be so identified by their different accession numbers, copy numbers, etc. Such a distinction may not be essential to serve a user who is interested in the subject of the book, but is necessary in circulation transactions, in acquisition and payment for purchase of the copies.

'Classification' is used in the conventional sense of grouping entities (objects) to form classes. The objects in a class share properties in common. For each problem relevant attributes are selected to form classes. Each object then represents an instance of the class in which it is grouped.

'Inheritance' is the formation of hierarchic classes of successive levels, that is superordinate and subordinate or sub-classes. A subclass is of greater intension and lesser extension than its immediate superordinate class. Hierarchic relations such as master-servant, genus-species, father-son, whole-part/portion are recognized.

An object-oriented operation is 'polymorphic' if it evokes different behaviour when applied to different classes of subjects. Therefore, there are different procedures/methods of implementing an operation on different classes of objects. As an instance, in graphics application, the 'draw' operation will have different procedures for different shapes.

Applied to the class line the draw operation will draw a line whereas when applied to the class circle it will draw a circular shape.

In the General Theory of Knowledge Classification there is a parallel to polymorphism. For example, the concept 'construct' has the seminal meaning of "to build", but applied to different classes of entities, for example, "house", "triangle", "sentence" at the phenomenal level, the result differs.

The above brief description indicates that the OOAD approach has features common to the process of identification, categorization and organization of information concepts for representing reality in designing a scheme for classification conforming to a model (hierarchical or faceted). Hence, several of the postulates and principles and techniques of the General Theory of Knowledge Classification, for example, the analytico-synthetic methodology, the canons for characteristics, postulate of fundamental categories, etc. can be helpful in OOAD (see below).

6.4 Demarcation of Subject Fields

Thinking about and communicating an idea or a body of ideas about an entity presupposes cognition of that entity. It involves demarcating or discriminating that entity from other entities. This means recognizing a boundary about that entity. In the physical world, the sensory experiences, e.g., visual, auditory, and tactile, provide for such discrimination. In intellection, the boundary may be created, that is, discrimination may be done by recognizing the dissimilarities in attributes (that is, by using differentiating attributes or characteristics) of the entities — the basic principle of classification. The definition of 'subject' (Sec. 4.4) conforms to such requirements of the cognitive

process. The creation of boundaries among component ideas of a subject can lead to the formation facets [01,03,12].

The universe of knowledge is ever-changing; it is developing at an increasingly faster rate. The nature of inter-relationships among entities may change and new kinds of relationships may be recognized from time to time by subject specialists. The grouping and demarcation of ideas into fields of study/subjects acceptable to specialists a few decades earlier lose their sharpness while new fields and/or relationships over-lapping and criss-crossing older boundaries arise from time to time. Nevertheless, the basic principle that any scheme for organization of subjects or component ideas within a subject, that is, mapping of the ideas should be compatible with the perception/mapping of ideas and subjects by specialists in those fields, holds good. This implies that database design should be based on a conceptual data model that parallels human information seeking behaviour (cognitive model).

Another aspect of the proliferations in the universe of ideas is that the range of search among information sources, for example, documents or their surrogates in manual catalogues or machine-readable databases, for information on a specific subject should be narrowed down to a convenient and viable range. This has been discussed elsewhere in the framework of studies on modes of formation of subjects and the concept of Basic Subjects [25,40].

6.5 Generalized Subject Structure Model

Mapping of component ideas of a subject, whether for designing a scheme for classification, for classifying subjects, for query analysis for search and retrieval, or in database design is based on subject representation. Essentially representation of a subject or of information

about an object involves analysis of the subject or the information about the object into component ideas, recognizing the nature of relationships among them and representing or mapping these as data structures, subject index entries, subject headings, class numbers, and other surrogates, that is, assembling the constituent elements in a pattern/model compatible with the perception of users/specialists in the subject concerned. This process is equivalent to transforming the multi-dimensional configuration of the subject for uni-dimensional representation.

In the process of such mapping every immediate-neighbourhood relation identified among the component ideas of every subject and among all subjects is to be kept invariant. This is not possible in practice in a linear representation. However, Ranganathan's General Theory of Knowledge Classification presents a cognitive model of subject structure that is economical, providing more information about the subject with less cognitive effort [01,03,12].

The generalized model for structuring of subjects based on the postulate of Absolute Syntax [41,34,49] (See Chapter 2, page 67)

- (a) conforms to the definition of subject (Sec. 4.4);
- (b) helps to minimize the range of search/ browsing necessary in a set of records so as to zero-in on the specific subject(s) of interest to the user at the moment, because of the heuristics provided (cybernetic hunting);
- (c) the power to discriminate/distinguish between specific subjects falling in different subject areas rests with the first component in

the linear representation of subject, namely, the 'Object of Study' or Personality facet (cf. relational data model); and

- (d) is object or entity-oriented (cf. Object-oriented Analysis and Design), implying that information seekers in general are interested in information about the attributes (Properties) of an object or entity (concrete or abstract) with a view to using, or manipulating, or changing the attributes; or in finding out which object(s)/entity(ies) possess given property(ies), or particular values for the selected attributes. Examples:

- (1) From a database on Medical Syndromes, doctors may wish to know which 'disease(s)' manifests the following combination of attributes or symptoms:

congenital; cyanosis; clubbing of fingers and of toes

- (2) From a database of hospital Patients records, doctors may wish to get information as to which 'patient(s)/disease(s)' show a particular symptom or a combination of given symptoms.
- (3) From a database of socio-economic development indicators for the constituent states of a country, planners or researchers may wish to know which Constituent States have particular values for a given set of indicators.
- (4) From a database of profiles of small enterprises, users may wish to know alternative means of financing small enterprises in a particular country or state or may want comparative data

on small enterprises in a particular area (e.g., State in India) receiving following types of funding and/or subsidies:

Term loan, State subsidy, Promoter's share, Soft loan, Seed capital, Development loan, Deferred payment loan, IFCI subsidy, IDBI subsidy, EDP subsidy, DIC subsidy etc.

7 ANALYTICO-SYNTHETIC METHODOLOGY APPLICATION

The postulates and principles of the General Theory of Knowledge Classification, especially facet analysis and the analytico-synthetic approach to it, constitute a basis for and have several applications in the design and development of information storage and retrieval systems (see Fig. 2 in Annex 1; and Fig. 3 in Chapter 2, page 52).

7.1 Organizing Ideas in Specialized Databases

Concept categorization and knowledge organization are used, in one form or another, in knowledge bases (e.g., specialized and object databases). Such organization of concepts also helps in recognizing the types of interrelationship that may exist among concepts in the particular subject domain, and in linking them accordingly with a view to assisting users browse and navigate more conveniently in the knowledge base; and in presenting the retrieved records or ideas in a sequence helpful to users.

7.2 Help to the Heuristic Process

Thinking about and communicating ideas about an entity will be more effective if the account about the entity is systematized or

General Theory of Knowledge Classification

structured. A similar communication process takes place when an information seeker is perusing/examining records in a database, that is, between the user and the records (purported to map real world entities using information concepts and structures). Therefore, the organization and structuring of information in each record and those of the records retrieved in response to a query taken as a whole should preferably be compatible with the perception or mapping of the real world by specialists on the entity (subject specialists) as already mentioned. Such a presentation of information can contribute to enhancing the user-friendliness of the system. The analytico-synthetic method based on postulates and principles is helpful in deriving a consistent concept categorization and knowledge organization in databases and in the outputs thereof. The resulting representation of entity/subject has a higher heuristic value in browsing and to zero-in on the specific entity/subject the user is interested in at the moment. The facet syntax so arrived at is deemed to parallel more closely the way in which a majority of normal persons arrange the component ideas of a subject when thinking, communicating, and/or seeking information about the subject. Hence, users are likely to find the system more user-friendly than is the case if the descriptors are placed any which way like 'peas in a bag', other aspects of the system remaining the same [19]. In the database consisting of case histories of hospital patients, developed by doctors who are also users of the system and who have not had any prior exposure to the postulates and principles of the General Theory of Classification or to analytico-synthetic method, the grouping and organization of the database fields and their preferred sequence conformed to what would have been arrived at by applying those

principles and methods. (*see* Sec. 10.2) A similar preference by specialists in forest resources management database has been noted [35,38]. (*see* Sec. 10.4)

The analytico-synthetic approach has been used in conceptualizing, mapping and visualizing in a convenient framework/ matrix the issues/concepts/data entities of interest to specialists such as energy resources planners. The frame-work could be translated into fields of a database or schedules in a scheme for classification. Examples are given in Sec. 10.3. These and other examples of organizing concepts in knowledge bases of expert systems confirm that the principles and postulates of the General Theory of Knowledge Classification make explicit the sequencing of component ideas in the thinking process of specialists. Hence they can serve as general guidance for arriving at a consistent and helpful organization of ideas in specialized databases.

7.3 Designing a Scheme for Classification and a Database

7.3.1 Parallel Steps

Designing a database is the process of arranging the data fields needed in one or more applications into an organized structure. The two phases of database design are: Logical database design, and Physical database design.

Logical database design is an implementation-independent exercise performed on the fields and relationships needed for one or more applications. It is independent of any particular database model, such as, hierarchical, network, or relational model. It describes the users' view of the data.

General Theory of Knowledge Classification

Physical database design is an implementation-dependent exercise using the results of the logical database design. It usually conforms to a database model and may be constrained by the capacity and capabilities of the software, hardware, tools and techniques available at a given point in time.

The designing of a scheme for knowledge classification also has a logical phase, designated by Ranganathan as work in the Idea Plane. This is application-independent, whether a scheme for macro-thought or one for micro-thought, the notation systems, etc.). It is also independent of any classification model such as, hierarchical, faceted or freely-faceted model. It presents an ideal view of knowledge organization required by users.

The implementation or physical design phase of designing a scheme for classification uses the results of the work in the Idea Plane (logical phase), conforms to some model, and may be constrained by the notational system used in achieving the findings of the Idea Plane (the logical phase). Ranganathan discusses such constraints and inhibitions imposed by the notational system on the Idea Plane in his paper on the methodology for the design of depth classification schemes for micro subjects [47].

The steps in identifying and delimiting data entities, their attributes and relationships among them (in relation to the needs of potential end-users) and mapping these on to a data model are similar to those for representing information about reality in designing schemes for classification of subjects or a knowledge domain.

Ranganathan's General Theory of Knowledge Classification recommends [51] demarcating the work of designing a scheme for classification into three planes – Idea Plane, Verbal Plane, and Notational Plane. Such a demarcation can also be usefully applied to the work involved in designing databases. The major steps in the three planes of work are similar in both cases as shown in the table below. The work preceding each of the steps is also similar in the two cases. Hence, the applicability of the postulates, principles and the related analytico-synthetic method of the General Theory of Knowledge Classification to the design and development of specialized databases not only in theory but also in practice.

Designing Database	Designing Scheme for Classification
<i>Idea Plane</i>	
1. Identifying data entities about which attribute data are to be collected for use	1. Identifying/defining the subject domain for which the scheme for classification is to be designed.
2. Selecting attributes of data entities of interest to potential users.	2. Selecting attributes of the entities constituting the subject in relation to users' interests.
3. Selecting data model, a scheme to map the entities, their attributes, their relationships (hierarchical, network, relational, or a mix of these).	3. Selecting classification model (hierarchical, faceted, freely faceted) for mapping information about the entities (concepts/isolates).

General Theory of Knowledge Classification

Designing Database	Designing Scheme for Classification
4. Grouping/dividing the data entities by their common attributes and differentiating attributes (characteristics).	4. Grouping/dividing the concepts/ isolates by their common attributes and differentiating attributes (characteristics).
5. Organizing, arranging the groups, subgroups and units derived at step 4.	5. Organizing, arranging the groups sub-groups and isolates derived at step 4.
<i>Verbal Plane</i>	
6. Naming fields and data elements	6. [Work in the verbal plane].
<i>Notational Plane</i>	
7. Assigning tags, coding...	7. [Work in the notational plane]

Specification of indexing parameters, of search strategies, of data display etc., is essentially dependent on data entities selected and data collected on them. A data entity and related attributes can constitute a field, each field with its contents or value thus representing a part of the database. The information content of a database may be divided into fields/subfields by relevant common/differentiating attributes or characteristics of the entities. In this process, the Canons for Characteristics and the Canons for Succession of Characteristics are helpful in preparing Field/Data Definition Table (see Sec. 7.4).

7.3.2 Application of Canons of Classification

The Canons for Characteristics are:

- Canon of Differentiation
- Canon of Relevance
- Canon of Ascertainability
- Canon of Permanence

In a database on local lands for use by the land records office, selecting characteristics such as

Registration number	Use of land
Name of owner	Area of land
Type of land	Location of land

conforms to the Canon of Differentiation and to the Canon of Relevance.

In a database on chemicals for use by chemists, selecting such characteristics of division as:

Molecular weight	Solubility in water
Melting point	Solubility in alcohol
Optical rotation	Molecular formula
UV maxima	Structure
NMR	Derivatives

conforms to the Canon of Differentiation and Canon of Relevance.

If the same database is also to be used by pharmacologists and medical personnel, additional characteristics, such as:

Medical uses	Side reactions
Mean lethal dose	Antidotes
Pharmacology	

will also be relevant.

General Theory of Knowledge Classification

The Canon of Relevant Succession requires that the application of two or more characteristics in succession should give rise to a sequence of sub-divisions of the universe of entities helpful to a majority of the target users.

In a database on on-going projects, the characteristics

Date of proposal	Date of starting
Date of revision	Expected date of completion
Date of approval	Actual date of completion

are relevant and their application in the sequence given above organizes the data in the chronological sequence of the events and is therefore deemed to be helpful to a majority of the users.

In a database on coal supply system, the characteristics

Source	Transport
Extraction method	Distribution
Refining method	End use

are relevant and the sequence of their application as mentioned above parallel the sequence of operations and the resulting organization of the information segments is deemed to be helpful to a majority of the users of the database.

The Principles for Helpful Sequence of Facets and of Isolates are helpful in the organization of information retrieved from a database [39].

In 1963, as mentioned earlier, based on analytico-synthetic procedures and the generalized freely-faceted model of subject, Ranganathan formulated a practical procedure for designing scheme for classification and for classifying micro subjects in great depth and

detail. The technique, further refined principally at the Documentation Research and Training Centre, Bangalore, India, proved to be a powerful tool for the analysis, synthesis and representation of even very minute subject areas, thus providing a map of the subject acceptable to specialists in the entity/subject. At that time, the technique appeared to be beyond the needs of document classification and indexing in libraries and complex to handle manually. With the availability of powerful micro-computers and DBMS packages the methodology is proving to be valuable in creating object data bases, and knowledge bases for expert systems. For example:

A scheme for depth classification of subjects in Motor Vehicle Production Engineering was designed in 1967 [42] applying the analytico-synthetic methodology. Using the characteristics as fields and sub-fields in a database on the subject, the data in each record and in the output could be organized to conform to the principles of helpful sequence. The output is a structured abstract, the information on the constituent ideas being arranged in a sequence acceptable to motor vehicle design and production engineers. Detailed indexes could be generated. (*see* Sec. 10.1) A query, such as,

Which Motor Car uses Synchromesh Gearbox and Rear Suspension composed of Cylindrical Helical Spring and Telescopic Damper?

can be responded to quickly.

7.4 Decisions on Object, Fields, Data Elements

A database is made up of one or more files; a file is made up of one or more records; a record may be made up of one or more fields

General Theory of Knowledge Classification

and a field may be made up of one or more data elements. A designer of an information system may begin with deciding on or defining what data elements should make up which field using the findings of an analysis of user needs, organizational analysis, the type of data model adopted and other parameters. Which set of fields should make up a record is another decision to be taken. The Canons for Characteristics and the postulate-based analytico-synthetic methodology are helpful in these decisions.

In a specialized database, generally each record consists of one or more fields each providing information/data about one or more attributes relating to an object (Personality facet) and on methods of studying or manipulating them, as illustrated in the following Object databases:

Database on/for	Object of Study	Master Record Devoted to	Fields: Attributes of
Forest resources management	Forest division	Forest division	Forest division
Market for nonconventional energy	Nonconventional energy source	Energy source	Market for energy source
Biogas equipment	Biogas equipment	Biogas equipment	Biogas equipment
Syndromes	Syndrome	Syndrome	Syndromes
Patients records	Patient	Patient	Patient
Toxicity of chemicals	Chemicals	Chemicals	Chemicals
Small industries	Small industry	Small industry	Small industries

This amounts to selecting the 'principal component' in a database design using the relational model discussed above.

7.5 Organizing Fields/Data Elements in a Field Definition Table

For a machine-readable database arranging the fields/data elements in a particular sequence in the Field Definition Table (FDT) or Data Dictionary may not appear to be important as any desired sequence of the fields and data elements can be obtained in the output. Even predefined output formats can be used for the purpose even if the fields are not arranged in any particular sequence in the FDT. However, organizing the fields in the FDT in a helpful sequence has its advantages. First, it can provide a structured overall view of the information about the entities described in the database such that the relationships among them can be more easily perceived; second, the interpolation of new fields/sub-fields or characteristics of division or sub-division will be easier when the structure of the information base is 'visible' and the relationship among the segments clearer; third, gaps in the information content of the database can be more easily identified and the exhaustiveness or completeness of coverage can be assessed; and fourth, if the fields are arranged in a helpful sequence, it will help decisions on splitting the information into two or more records or databases, if necessary, and linking of whole records or specific fields thereof as may be necessary to serve the needs of different groups of user and at the same time be able to make a more optimal use of computer resources.

A number of Principles for Helpful Sequence of Facets and of isolates (unit concepts) in an array formulated by Ranganathan are available [39, 51]. These can be used in the arrangement of fields in the database and its FDT. When the number of fields to be organized is large,

the application of a Principle of Helpful Sequence may be tedious. It will be helpful first to group the fields into broad categories and arrange these groups in a helpful sequence applying one of relevant principles of helpful sequence; then the fields in each sub-group may be arranged using relevant principles. This Group Strategy can be applied to any sub-group if the number of fields within it is large. For example:

In the database of hospital Patients records, each attribute of the patient given in the case history sheet and enumerated in the FDT was correlated with one or other of the Five Fundamental Categories postulated by Ranganathan, namely, Personality (P) or the object of study taken as a whole, Property (M) or attribute of the object of study, Energy (E) or external action, Space (S), and Time (T), and the fields arranged applying the principle of decreasing concreteness or PMEST sequence. Thus,

1. Patient's identity and administrative information
2. Patient's general background
3. Patient's complaints
4. Signs/symptoms
5. Interrogation
6. Investigations/Examination findings
7. Treatment
8. Follow-up.

Attribute groups 1 and 2 correlate with the whole object, that is, patient as a whole; groups 3 and 4 correlate with properties of the particular state of the patient, namely, illness; groups 5 to 8 correlate with external action. A similar sequence of the groups can be arrived at by using the Principle of Later in Time instead of the Principle of Decreasing Concreteness. As mentioned earlier, such a sequence was the preferred sequence of fields in the case history sheet used in hospitals.

Where the number of attributes within a group is large, a similar strategy can be followed for that group. In the FDT for a database on Forest Divisions of a specific geographical area, the first grouping of the fields is by district; the attributes of the forest division of each district are grouped using a relevant principle of helpful sequence; and within some of these sub-groups there may be further divisions and groupings. (see below)

Level 1	Level 2	Level 3
Division: Districts [arranged by Spatial Contiguity principle]	Attributes of Division [arranged by using Correlation technique and Decreasing Concreteness principle] Headquarters Location Area Elevation Peaks Rivers Reservoirs Hydro-electric proj. Irrigation proj. Temperature Rainfall Forest types Farm forestry Population Number of schools Number of hospitals etc.	Ranges [arranged by using Spatial Contiguity principle; attributes of each range arranged by using Correlation technique and Decreasing Concreteness principle] Name Headquarters Ranger Population Number of sections Number of beats etc.

Another method of grouping is to correlate the attributes to the steps in the study of a system and to organize the groups in the sequence of the steps [39].

7.6 Mnemonics

Ranganathan proposed four kinds of mnemonic devices for use in constructing isolate numbers and whole class numbers. They are: Seminal mnemonics, Systematic mnemonics, Scheduled mnemonics, and Verbal mnemonics [52]. These concepts can be usefully applied in database design, especially integrated databases and vocabulary control tools.

An integrated database may hold records describing different kinds of entities, some of whose attributes may be common to two or more entities. Users often wish to retrieve information on related entities quickly or concurrently in response to a query without having to change databases and reexecuting the query again and again. Integration of records into a single database facilitates such concurrent search and retrieval. Devices such as hypertext links and multiple window displays are also used for similar purposes.

Cognate and similar ideas occur in different contexts and may have different verbal designations, that is, in the idea plane there is equivalence of ideas in the different contexts, but in the verbal plane the names are different even in the same language. This equivalence of ideas at the near-seminal level formed the basis of Ranganathan's system of seminal mnemonics. The concept of polymorphism in object-oriented analysis and design mentioned in Sec.6.3 is similar to the concept of seminal mnemonics. Examples: the concept 'Correction' occurs in the context of proof-reading, navigation, and surgery; the concept 'Construction' occurs in the context of building construction, sentence construction in language studies, in geometry, etc. If the

seminally equivalent concept is assigned the same isolate or class number, the same number can be assigned if an equivalent idea arises in future in a different context. If one uses the number (for the concept) already enumerated (say, in classification schedule) then the principle of Scheduled mnemonics will also be conformed to. This will ensure consistency even when different persons assign numbers or codes.

Let us consider fields for databases of documents (bibliographic database), profiles of research projects, of institutions, of specialists, and of events. 'Name of Person' field occurs in all of these although with different designations: Personal Author (bibliographic database), Project Leader (research project database), Head of Institution (institution database), Name of Person (database of specialists), and Organizer of Event (events database). Similarly, 'Name of Corporate Body' as Corporate Author (Bibliographic data-base), Performing Institution (research project database and events database), Affiliation (data-base of specialists); 'Date of starting' as Starting Date (research project and events databases), Date of Establishment (institution database). 'Descriptors' as Descriptors or Subject Key-words (bibliographic, research projects and events databases), Fields of Specialization (data-bases of institutions and specialists). And other equivalences.

In an integrated database of bibliographic records, profiles of corporate bodies, experts, research projects, information systems, research projects, and events all the fields required to describe the different entities can be listed in a single FDT in such a way that seminally equivalent idea fields will be listed only once and given a single Tag number (e.g., Name of person, Name of corporate body, starting date, subject descriptor, etc.) If desired, separate online input worksheets can be prepared for each of the different entities); a single specification of the indexing parameters; and one or more display formats. In the

respective input worksheets an appropriate field name can be used even though the tag may be the same. For example, the field tag 300 of the Common Communication Format can be used for 'Name of Person' field, with the name of field as Personal Author for the bibliographical records, Name of Person in the records for specialists biodata, Name of Head for the institution records, Name of Leader for the research projects database, and Name of Organizer for the events records. Similarly, for Name of Corporate Body, Starting Date, etc. On the one hand this will economize on computer resources and on the other enable the retrieval of relevant records (bibliographical, institution, expert, research project, event) in response to a single query by Name of Person or Name of Corporate Body or Descriptors.

In other factual databases too one can find seminally equivalent fields. For example, the fields Signs, Symptoms, Etiology, Pathology, Diagnostic Procedures, Treatment, Prognosis, Disease Classification, etc. are common to the Syndromes database, the Hospital Case Records database, etc.

In such integration using same field tag for equivalent ideas occurring in different contexts, the canons for scheduled and systematic mnemonics will be automatically, though indirectly, conformed to.

8 SEARCHING AND USER-SYSTEM INTERACTION

8.1 Humanization of Information Service

Ranganathan emphasized the humanization of reference/information service, so as to make the system more user-friendly and to enhance its capacity to meet the specific requirements of users, which, in turn, would satisfy the Five Laws of Library Science.

He recommended the provision of a human interface (intermediary)

- to receive the query of the end-user,
- to dialogue with the end-user to understand and interpret the query in terms of his/her precise needs, and if necessary, to use subject structuring tools such as a scheme for classification as an interface, in the process,
- to analyze the query into appropriate elements based on a knowledge of or using a relevant scheme for classification,
- to facet structure and express the analyzed query in the form of an appropriate search statement,
- to select appropriate domain of information materials based on a knowledge of classification or organization of the materials,
- to perform or help the end-user in performing search and retrieval operations,
- to facilitate user examination of the intermediate search results,
- to provide for user feedback,
- to modify the search expression, if necessary to obtain a more precise and/or wider range and/or better quality of retrieved information/information source records,
- to organize or assist the user in organizing and presenting the retrieved information/information source records in a form and format convenient to the user,
- to help, if necessary, in carrying out further processing of the retrieved information (e.g., tabulation, statistical processing, modelling, etc.), and

General Theory of Knowledge Classification

- to procure and deliver the original source materials selected by end-user, if necessary.

The application of information technology to information handling does not do away or preclude the humanization of information service; on the other hand more sophisticated interfaces are required to help the user cope with and derive optimum benefits from the hardware/software capabilities and limitations.

Structured Search request forms (manual or online) are designed to facilitate the more precise expression of user's information need, consulting a structured thesaurus, classification scheme or other vocabulary control tool in constructing the search profile, etc. Some forms, presenting a structured view of the subject assist users in refining the query statement in a systematic way (see Sec. 10.5).

End-user studies indicate that the application of information technology could change the information seeking behaviour of users as, for instance, the accessibility to a wide variety and large amounts of information relatively quickly. Such facilities change their perception of the capabilities of information systems, information seeking, etc. On the other hand, quick response is expected of users, more particularly when using online services.

It is generally known that in many cases a user may not be able to state precisely his/her information requirement to begin with. System offerings, browsing of retrieved items and heuristics provide for a learning base during the search and retrieval session. A well-designed and structured user-interface can enhance such learning feature of the system through effective interactive query refinement. Online use of classification schemes, faceted thesaurus, etc., can assist users in selecting appropriate search terms.

Users may wish to navigate between different databases during a search session, to assemble data (even different types) from different sources (interleaved access). Again, a structured interface can help in this regard (see below).

8.2 Search Strategy

Search strategies include: briefsearch, building block method, successive fractions method, citation pearl growing method, beginning a search with the most discriminating facet first, and starting the search with the most specific term/element as measured against the file. In any of these strategies, it is advisable to start the search with the term(s) denoting the Object of Study (Personality Isolate), as the discriminatory power among different subjects rests with the Object of Study (Personality isolate) in the Generalized Subject Structure model.

In the data format in a database devoted to a specific field, the leader/header/data directory for each record will register, among other things, the field tag and its attributes. If the search is started with the Object of Study term, which is often relatively unique compared to other field values (e.g., attribute), the search will be faster and more economical. For example, the attribute 'blue colour' can apply to many different objects; so also an action denoting term e.g., 'construction' would apply to geometrical figure, sentence, building, etc.

From the Syndrome database, the combination of symptoms alone or with treatment methods retrieves several records. If the Object of Study (Personality facet) e.g., Heart is combined with symptoms (attributes), then a fewer number but more precise set of records, is retrieved. (see Sec. 10).

General Theory of Knowledge Classification

Experience shows that user-intermediary-information system interaction or user-information system direct interaction helps in the more precise formulation of user's information need and in the search process so as to improve the retrieval of information or records pertinent to user's needs. Usually, terms representing concepts of interest to user are selected and combined using boolean and other operators to formulate search expressions. This process of facet analysis and synthesis should preferably match user's perception of reality or knowledge gap. Ranganathan's analytico-synthetic methodology formalizes the process leading to a more or less consistent search formulation. Experience also indicates that a properly facet-analyzed query can give better retrieval performance [24]. Vocabulary control tools, such as, thesaurus, subject headings list, classification scheme, etc., are helpful in the analysis and processing of queries.

For a Patients' case records database, a thesaurus-like tool was created using terms derived from the database records to facilitate user-system interaction in information retrieval. The user is helped to start from broad divisions of patient's case history and move in steps through the hierarchy of sub-divisions. At the same time, where appropriate the user can also view related terms. (*see Sec. 10*). Alternatively, or in addition, one may use a conventional thesaurus, a faceted scheme for classification, a thesaurofacet, or a classaurus.

User interaction and search formulation can be facilitated, made more productive by linking the facet-analyzed query terms to appropriate thesaurus terms so as to provide for browsing and a wider choice of concept terms. This can be done manually by looking up relevant terms in a thesaurus or by automatically using a machine-readable thesaurus and database of queries. A record for a query can have fields for the Statement of Query (by user), for Facet Analyzed Terms, for Name and

Address of User, for Dates (receipt of and response to), for Notes on Response, etc.

When the query is displayed, each of the facet terms can be linked to corresponding terms in the thesaurus, for example, by using the REF(L) function of MicroISIS in the display format of the Query record. We used the machine-readable Macrothesaurus of OECD and the corresponding CDS/ISIS Pascal programmes. As the OECD Macrothesaurus is multi-lingual (English, French, Spanish), each query term can be linked to terms in the different languages. Search using terms from one language will automatically search also with the corresponding terms in the other languages. Here is an example:

QUERY: Women's role in development education in Africa

FACETS: Development education. Women. Africa

DEVELOPMENT EDUCATION (3947)

- SN Education on issues of concern in economic and social development for youth and adults
- BT Education
- RT Development personnel
- RT Economic and social development

WOMEN (2150)

- NT Married women
- NT Mother
- NT Rural women
- NT Women worker
- NT Education of women
- RT Gynaecology
- RT Women's organizations
- RT Women's participation
- RT Women's rights
- RT Women's role
- RT Women's status

General Theory of Knowledge Classification

What has been demonstrated does not preclude direct search in the database using query terms or facet-analyzed terms; or search first in the thesaurus to select appropriate search terms followed by search in the database. Free text search is also possible.

A database of queries, recording facet-analyzed search terms, has other uses as well, e.g., research on user needs, informetric analyses, updating thesaurus/classification scheme, search in the Query database first when a new query is received so as to avoid duplication of effort, and preparation and updating of user profiles for SDI service.

A similar facet analysis of the subject of a document (or information relating to an entity in a specialized database) can be done while inputting a record. The terms can be linked to the corresponding terms in a thesaurus and displayed when the record is retrieved [13]. For example, consider the query

“Breeding of pyrethrum in Africa, particularly Kenya”

Using a Pascal interface, each of the terms

Pyrethrum. Breeding. Africa. Kenya

can be linked online to a relevant thesaurus (we used FAO's AGROVOC).

Looking up the term “Breeding” led to “Breeding Methods” and the following thesaurus links:

- BREEDING METHODS
- UF genetic improvement
- NT cross breeding
- NT 2 backcrossing
- NT 3 convergent improvement
- NT 2 reciprocal crossing

- NT hybridization
- NT 2 intergeneric hybridization
- NT 2 interspecific hybridization
- NT 2 somatic hybridization
- NT inbreeding
- NT outbreeding
- NT progeny testing
- NT 2 diallel crossing
- NT pure breeding

[Next] B[ack] F[irst] P[age] S[elect] T[erm select] Q[ue]ry] M[ark term]
G[Mark code] ?[Disp query] A[dd rel] D[el] C[reate] W[rite in Fld] X[eXit]

Looking up “pyrethrum” in AGROVOC led to “Chrysanthemum
Cinerariaefolium” and the following links:

CHRYSANTHEMUM CINERARIAEFOLIUM

- UF pyrethrum cinerariaefolium
- BT chrysanthemum
- BT 2 compositae
- RT pesticide crops
- RT pyrethrum extracts

[Next] B[ack] F[irst] P[age] S[elect] T[erm select] Q[ue]ry] M[ark term]
G[Mark code] ?[Disp query] A[dd rel] D[el] C[reate] W[rite in Fld] X[eXit]

Ranganathan’s analytico-synthetic methodology is also helpful in the analysis and structuring of data, in grouping them into fields and sub-fields describing, for example, various attributes of resources, roads, landforms, soil, rock type, vegetation, etc., of a given spatial area as may be needed in designing a geo-coded or geographical information system. These then can be converted and represented in the form of tables, graphic patterns, maps and text using appropriate software.

8.3 Hypertext Link for Multiple Database Search

At the institutional level, it may be necessary to develop several databases providing different types of information and to be able to search some or all of the databases concurrently either using one and the same set of search terms or different sets of search terms. It may not be expedient to enter all of the different types of records in a single database. For example, a medical information system may develop separate databases respectively on patients' case records, syndromes, pharmacological, toxicological and therapeutic properties of drugs, profiles of specialists and of institutions, documentary materials, etc. It may be necessary to search some or all of the databases concurrently to retrieve all related information to meet users' needs (and of the Five Laws). Especially when several databases carrying different types of information relating to an entity (factual, tabulated data, bibliographical, biographical etc.) are available, the user needs to be provided information about each database to enable the selection of appropriate database(s) with respect to a specific information requirement.

With this in view, a user interface (SELECT.PAS) was developed for Micro CDS-ISIS databases (using the CDS Pascal language) to provide for:

- onscreen listing of the names of the databases;
- online information about the content, number of records, updating frequency etc. about each of the databases;
- selecting online one or more databases for searching;
- online guidance (help) in formulating search expressions in the search language acceptable to the system, e.g., use of boolean and other operators, truncation, specific field search, etc.

- index-based as well as free-text search in the database(s) selected;
- online selection of available pre-defined display/print formats for each of the databases selected;
- online use of vocabulary control tool, such as, thesaurus;
- browse through the retrieved items, and selectively save records and print them out subsequently;
- *selection of term(s) from one or more displayed fields of a record by moving the cursor on the desired terms to formulate new search expression;*
- *Searching in the same or another database using the new search expression, and thus surf and move back and forth from one database to another with the same or new search expressions (by selecting onscreen terms from the records displayed); and*
- Deletion of a saved file, if desired. [60]

Users find this hypertext facility quite useful in assembling a wide range of information for analytical and comparative study, to understand concepts in different contexts, etc. [38a]

8.4 Data Mining and Data Mapping

Information and communication technologies, networks and information highways enable accessing and retrieving interactively a wide range and vast quantities of information from data sources located at global distances even on very specialized topics. A large number of databases can be searched (mined) to extract the required information. Often an entity or concept of apparent interest to a user may occur in

different contexts (with the same, near-synonymous, or different connotations in the different contexts). The amount of data and information retrieved can and do create difficulties to the user, for he/she has to browse through the records and identify the data sources really pertinent to his or her interest at the moment.

The entity/concept may be related to other entities/concepts in different ways in different contexts. A facility is needed to organize the retrieved items into convenient classes indicating or according to the inter-relationship between the entity/concept of interest and other entities/concepts. This will create different clusters of data (around the entity/concept of interest). Another capability required is to be able to reorganize the clusters in different patterns with different foci of interest. This is like rotating a kaleidoscope to view different patterns with the same set of coloured beads in each group. Such mapping of the structures in different patterns is tedious to obtain manually. Information technology should be of help in this regard.

8.5 Helpfulness of the Principles and Postulates: Summary

In summary, the postulates and principles of Ranganathan's General Theory of Knowledge Organization, postulational approach to facet analysis and synthesis, and the Principles of Helpful Sequence are helpful in:

- Organizing concepts in structured knowledge bases and in specialized databases.
- Designing schemes for classification.
- Generating various types of structured indexes.
- Designing and developing vocabulary control tools, such as thesaurus, classaurus, subject headings, etc.

- Preparing field definition table or data dictionary for databases.
- Recognizing inter-relationships among concepts and linking them accordingly.
- Assisting user to navigate more conveniently in the databases or among the records to zero-in on the specific subjects of interest to him/her at the moment.
- Facet-analyzing user's query and structuring search expressions for better results in online retrieval.
- Presenting retrieved records/information in a sequence helpful to user.
- Object-oriented analysis and design.
- Cognitive modelling and neural networks.

9 SPECIALIZED OBJECT DATABASES

9.1 General Features

As mentioned earlier, object databases (ODBs) are becoming increasingly the basis for knowledge-based systems, expert systems, etc. and for generating value-added products and services, especially at the institutional level. We may now examine some features differentiating ODBs from conventional bibliographic databases.

An ODB usually provides information about an 'object', such as market, equipment, hospital patient, rocket-launch vehicle, forest resources, chemicals, etc. At a given time an end-user is usually interested in selected attributes of the object, with a view to manipulating, modifying, or using the object for a particular purpose. Such databases can be home-grown on PCs to meet the needs of specialized user groups and can supplement or be integrated with,

General Theory of Knowledge Classification

conventional bibliographic and referral type databases and services thereof.

The designing of such ODBs and the provision of value-added services, however, require:

- the involvement of end-users at various stages of the design and development of the system;
- close interaction between information professionals and end-users;
- that the subject background and work experience of the information specialist(s) be generally compatible with those of the end-users served;
- that the education and training of information professionals should cover indepth information extraction, analysis, synthesis and repackaging methods, the theories and principles of knowledge organization, natural language interface with information systems, as well as information technology applications in designing ODBs and preparation of value-added products.

9.2 ODB and End-User Involvement

The need to involve end-users in the design and development of ODBs arises for the following reasons among others:

- In designing a bibliographical database as a component of an information storage and retrieval system, involvement of end-users is usually minimal, apart from the occasional assistance in the choice of subject descriptors and/or in preparing abstracts of specialized materials. Other bibliographic data elements are mostly normalized (through cataloguing rules, authority lists, or vocabulary control tools, etc.) and these are

generally applied to documents in all subjects. But in a specialized ODB the data elements (usually attributes of the object and methods of manipulating them) differ from one type of object to another. For example, the data elements of a record in a database on biogas equipment will be different from those of a record in a database of hospital patients' records; and the latter will differ from those of a record in a database for energy resources planning.

- For a bibliographical record the source of information for input to most of the fields – author, title, edition, publisher, series, subject, etc. – is mainly the document being catalogued. On the other hand, for a record in a non-bibliographic ODB the input data may be taken from different sources including remarks of individuals, even for a single field, and the data may already be in an analyzed, evaluated and reorganized form.
- An ODB (e.g., database of records of hospital patients or of forest resources) may be required to respond to questions, such as, whether a particular attribute/field is present or not, or whether a particular attribute is positive or negative, or whether a particular action (e.g., a specific surgical procedure) was carried out or not and, if yes, then the details.
- An ODB usually attempts to provide information/data asked for rather than just references to other sources.
- The output record from a bibliographic database is largely standardized e.g., AACR2 format; on the other hand, end-users of a specialized ODB may demand outputs in different formats or for a facility to create their own formats.
- The information/data recorded in an ODB is often of local interest/use (e.g., hospital patients records, forest resources) and

may not be sourced in international or national abstracting services. Therefore such ODBs are usually to be developed at the institutional or small specialized group level.

These and other factors necessitate close involvement of end-users at various stages in the design of non-bibliographic specialized ODBs, more so than in the design of bibliographic databases.

9.3 Roles of End-User

Involvement and assistance of end-user in the design and/or development of an ODB is needed in some or all of the following aspects:

- to define the specialized objects or subject areas in which databases are required at the institutional or individual or small group level;
- to assist the designer in a more precise understanding of the structure of the field – the object, its attributes of interest to users, the interrelation among the concepts, etc.
- to develop a reference frame-work(s) to analyze and synthesize the information in a pattern acceptable to end-users;
- to identify and advice on relevant sources of information, including views of experts, and sometimes for providing the data/information itself;
- to assist in the analysis, evaluation, filtering and organization of the data for input into the system;
- to advise on the fields/data elements of a record, that is essentially the selection of attributes and other concepts relating to the object(s) of the database;
- for data rating;

- to provide the types of queries that the system should be designed to respond to;
- to advise on the specialized terminology to be used in the system;
- to specify the types of outputs and their formats that may be frequently required with a view to formulating predefined output formats, while at the same time providing for the formulation of other formats as well; and
- to provide frequent feedback to the system.

9.4 Modes of End-User Participation

In the examples discussed in this paper subject specialist/end-user involvement in system design was secured in the following ways:

- Information specialist(s) designing the system working or interacting closely with subject specialists to get assistance/advice on the aspects mentioned above. (Example: Forest Resources database; Rocket Launch Vehicle database).
- Subject specialist(s) himself/herself designing and developing the system after receiving short training and demonstration of the use of the software. (Example: Database of Hospital Patients);
- Information specialist having the subject background and experience in the field being in a position to understand the information needs of end-users, develops appropriate databases and on the basis of user-feedback modifies them (Example: Databases for Technology Evaluation; Database of Biogas Equipment).

General Theory of Knowledge Classification

As already mentioned the theories, principles and techniques of concept categorization, classification and knowledge organization, and of the analytico-synthetic facet approach, as well as those of methods of vocabulary control and natural language indexing developed for organizing and retrieving subject information from bibliographic databases are helpful in designing non-bibliographic ODBs [37, 38].

Several of the examples presented here are taken from working or operating information systems. Each database forms a component of a group of databases serving the needs of an institution or user group. Such a group of databases may consist of bibliographic databases, referral data-bases (e.g. profiles of institutions, experts, projects, etc., often integrated into a single data-base) and other non-bibliographic databases. All the databases were developed using Micro Unesco's CDS-ISIS software.

The specialized ODBs developed include:

- hospital patients records databases (3 databases);
- database of medical syndromes;
- forest resources database;
- national parks description;
- databases for technology evaluation;
- databases on the chemistry and toxicology of selected chemicals;
- biogas equipment database;
- databases for energy resources planning;
- database of land records;

- urban planning information support;
- automobile production engineering;
- gene data bank;
- museum objects;
- software packages;

and others

Examples of some of the databases are presented in Sec. 10. In each case some particular application of the analytico-synthetic methodology and of the postulates and principles of Ranganathan's General Theory of Knowledge Classification/Organization discussed earlier is explained. Sample outputs from the databases are also presented.

10 EXAMPLES

10.1 Motor Vehicle Production Engineering

10.1.1 Fields and Sub-fields

The *fields* and *sub-fields* for this database were formed using the quasi-isolates (characteristics) originally formulated for designing a freely-faceted scheme for classification for subjects in the field of Motor Vehicle Production Engineering [42]. The choice of characteristics or quasi-isolates for the division and application of normative principles, such as the Canons for Characteristics, are discussed in detail in the paper cited. The characteristics of division are enumerated below in the sequence organized in the earlier work. The formation of fields and sub fields and structured outputs are illustrated.

General Theory of Knowledge Classification

Serial No.	Characteristic
001-003	By Brand
001	By Company
002	By Model
003	By Year
004	By Make (country)
005-021	By Purpose/use
005	By Service
006	By Service area
007	By Portability
008	By Mass communication use
009	By Funeral use
010	By Medical/Hospital use
011	By Recreational use
012-013	By Cargo transport
012	By Type of cargo
013	By Weight
014	By Animal transport
015-016	By Passenger transport
015	By Type of passenger
016	By Number
017-021	By Environment adapted to
017	By Weather
018	By Physiography
019	By Altitude
020	By Latitude
021	By Gradient climb
022-031	By Design parameters (whole vehicle)
022-023	By Weight
022	By Vehicle weight
023	By Weight distribution
024-027	By Dimension
024	By Overall length
025	By Overhang
026	By Overall width
027	By Ground clearance
028-029	By Speed
028	By Mean maximum speed
029	By Maximum speed in each gear
030-031	By Acceleration

Serial No.	Characteristic
030	By Acceleration from standing to 0.25 mile
031	By Acceleration through gears
032-137	By Design parameters of organs
032-034	By Power system
032	By Equipment
033	By Voltage
034	By Current
035-065	By Engine
035	By Kind
036	By Brand
037	By Mounting
038	By Supercharging
039	By Power
040	By Thrust
041	By Maximum torque
42	By Compression ratio
043	By Cycle
044-046	By Cylinder
044	By Number
045	By Bore
046	By Arrangement
047	By Displacement
048	By Stroke
049	By Fuel
050-061	By Fuel injection system
050-056	By Carburator
050	By Number
051	By Brand
052	By Draught
053	By Number of barrels
054	By Number of venturi tubes
055-056	By Choke
055	By Brand
056	By Operation
057	By Method of injection
058-059	By Fuel pump
058	By Brand
059	By Kind
060-061	By Oil filter

General Theory of Knowledge Classification

Serial No.	Characteristic
060	By Brand
061	By Kind
062	By Combustion chamber
063	By Valve gear
064	By Ignition
065	By Cooling
066-070	By Frame
066	By Construction
067	By Profile
068	By Kind
069	By Articulation
070	By Material
071-091	By Cargo body
071-073	By Truck trailer
071	By Kind
072	By Axle assembly
073	By Number of axles
074	By Construction type
075-084	By Purpose
075	By Body for speed
076	By Body for mixed freight
077	By Body for passenger vehicle
078	By Body for food transport
079	By Body for liquids transport.
080	By Body for livestock transport.
081	By Body for garbage
082	By Body for vehicle transport.
083	By Body for construction and repair equipment
084	By Body for building equipment
085	By Material
086	By Length
087	By Width
088	By Height
089	By Floor area
090	By Capacity
091	By Number of decks
092-096	By Seat
092	By Number

Serial No.	Characteristic
093	By Position
094	By Kind
095	By Cover
096	By Cover of headlining
097-099	By Door
097	By Name
098	By Position
099	By Op
100-104	By Window
100	By Number
101	By Position
102	By Material
103	By Curvature
104	By Operature
105-106	By Roof
105	By Kind
106	By Material
107-108	By Suspension
107	By Position
108	By Kind
109-114	By Spring
109	By Material
110-112	By Shape
110	By Axial load
111	By torsion
112	By Bending
113	By Length
114	By Rate of spring
115	By Control
116-137	By Wheel
116	By Position
117	By Number
118	By Number of driving wheels
119	By Wheel base
120	By Alignment
121	By Track
122	By Number of revolutions
123-125	By Rim
123	By Material

General Theory of Knowledge Classification

Serial No.	Characteristic
124	By Width
125	By Fixing
126-137	By Tyre
126	By Brand
127	By Kind
128	By Material
129	By Fabrication
130	By Maximum load
131	By Overall diameter
132	By Static load radius
133	By Size
134	By Weight of casting
	By Inflation pressure (alternative to 219-221)
135-136	By Traction characteristic
135	By Air resistance
136	By Total air-and-rolling resistance
137-216	By Operation associated characteristics
137-152	By Transmission
137-144	By Gear box
137-138	By Drive
137	By Type
138	By Number of speeds
139	By Number of speeds
140	By Number of auxiliary speeds
141	By Gear ratio
142-144	By Final drive
142	By Kind of gearing
143	By Arm
144	By Gear ratio
145	By Rear axle
146	By Kind of shift
147	By Method of transmission
148-152	By Clutch
148	By Brand
149-152	By Disc
149	By Number
150	By State
151	By Diameter

Serial No.	Characteristic
152	By Engagement
153-160	By Steering system
153By	Brand
154	By System
155	By Steering gear
156	By Linkage
157	By Gear ratio
158	By Rotation of steering wheel
159	By Number of turning circles, lock-to-lock
160	By Diameter of turning circles
161-172	By Brake system
161	By Brand
162	By Purpose
163	By Assistance
164	By Relation to wheel
165	By Construction
166-168	By Shoe
166	By Number
167	By Kind
168	By Width
169	By Brake-pedal pressure
170	By Swept area
171	By Lining
172	By Cooling
173-188	By Safety
173-178	By Indicator/measuring device
173	By Oil
174	By Electrical power
175	By Fuel
176	By Temperature
177	By Speed
178	By Time
179-187	By Light
179	By Number
180	By Brand
181	By Purpose
182	By Position
183	By Power
184	By Shape

General Theory of Knowledge Classification

Serial No.	Characteristic
185	By Contour of reflector
186	By Filament
187	By Operation
188	By Safety device
189-205	By Facility and comfort accessory
189-194	By Baggage accommodation
189	By Position
190	By Fixity
191	By Length
192	By Width
193	By Height
194	By Capacity
195-199	By Ventilation and heating
195-197	By Purpose
195	By Climatic control
196	By Heating
197	By Ventilation
198	By Equipment
199	By Operation
200	By Telecom facility
201	By Audiovisual facility
202	By Interior trimming
203	By Sanitation
204	By Housekeeping
205	By Executive facility
206-218	By Maintenance factors
206	By Overall fuel consumption
207	By Fuel consumption at constant speed
208	By Fuel tank capacity
209	By Distance per tank full
210-213	By Engine maintenance
210	By Cooling system
211	By Engine sump oil
212	By Change of oil
213	By Change of filter element
214	By Final drive oil
215	By Slip differential change of oil
216	By Overdrive oil
217	By Gearbox oil

Serial No.	Characteristic
218	By Grease application
219-221	By Tyre pressure
219	By Normal driving
220	By Fast driving
221	By Full load driving

10.1.2 Arrangement of the Characteristics

The quasi-isolates/characteristics of division were arranged in a helpful sequence using Group Strategy and the Principles for Helpful Sequence. Some examples are given below:

Quasi-isolates of order 1 arranged by Correlation Technique:

Quasi-isolate group	Correlated with	Serial number
By Brand	Personality	1-3
By Make	Personality	4
By Purpose	Personality	5-21
By Properties of Motor vehicle as a whole	Matter (Property)	22-31
By Properties of Organs of vehicle	Matter (Property)	32-137
By Operations associated characteristics	Energy	138-221

Quasi-isolates of order 2 associated with quasi-isolate of order 1 "By Brand" arranged by Correlation technique:

Quasi-isolate group	Correlated with	Serial number
By Company	Personality	1
By Model (Design)	Matter (Property)	2
Year	Time	3

General Theory of Knowledge Classification

Quasi-isolates of order 3 associated with quasi-isolate of order 2
By Body arranged by Correlation technique:

Quasi-isolate group	Correlated with	Serial number
By Purpose	Personality	75-84
By Material	Matter ((Material)	85
By Length	Matter (Property)	86
By Width	Matter (Property)	87
By Height	Matter (Property)	88
By Floor area	Matter (Property)	89
By Capacity	Matter (Property)	90

Quasi-isolates of order 4 associated with quasi-isolate of order 3
"By Tyre" arranged by Correlation technique:

Quasi-isolate group	Correlated with	Serial number
By Brand	Personality	126
By Kind	Personality	127
By Material	Matter (Material)	128
By Fabrication (Structure)	Matter (Property)	129
By Maximum load	Matter (Property)	130
By Overall diameter	Matter (Property)	131
By Static load radius	Matter (Property)	132
By Size	Matter (Property)	133
By Weight of casing	Matter (Property)	134
By Inflation pressure	Matter (Property)	135
By Traction characteristics	Matter (Property)	136

The following quasi-isolates are arranged according to the Principle of Spatial Contiguity, left-to-right:

- By Clutch
- By Transmission
- By Rear axle
- By Final drive

The following data elements are arranged using the Principle of Increasing Quantity, the attribute concerned being the increase in the Lumens per watt (lwp) of the filament of the lamp:

- By Light
 - By Filament
 - Carbon
 - Tantalum
 - Osmium
 - Tungsten

The following data elements are arranged using the Principle of Decreasing Quantity, the attribute concerned being the decrease in the tension of the material of the Spring.

- By Material of spring
 - Steel
 - Hard drawn steel
 - Oil-tempered steel
 - Annealed high carbon steel
 - Chromium-vanadium steel
 - Silicon-manganese steel
 - Flat CR strip
 - Stainless steel 18-8
 - Chromium stainless

The following data elements are arranged using the Principle of Later-in-Time:

- By Climatic control
 - Thermostatic control
 - Air-conditioning
 - Pressurization
- By Curvature of window glass
 - Flat
 - Curved
 - Compound curved
 - Wrap over

General Theory of Knowledge Classification

10.1.3 Fields and Sub-fields

The characteristics at various levels were used to derive fields and sub-fields in the Field Definition Table (FDT). Some examples are given below.

Field Tag	Field/Subfield Name	Repeatability
1	Brand	Yes
A	Company	
B	Model	
C	Year	
	Design parameters (Whole vehicle)	
22	Weight	Yes
A	Vehicle weight	
B	Weight on rear axle	
C	Weight on front axle	
D	Laden weight	
24	Dimension	Yes
A	Length	
B	Width	
C	Ground clearance	
D	Height	
71	Cargo body	Yes
A	Truck trailer	
B	Construction type	
C	Purpose	
D	Material	
E	Length	
F	Width	
G	Height	
H	Floor area	
I	Capacity	
J	Number of decks	

The display format has been designed to provide a structured abstract of each record. An example follows:

10.1.4 A Record Output

TITLE :FORD V-six model: analysis of the design of the largest of the British Ford range.

SOURCE : *Automobile engineering* 1966 May, 164-172

ABSTRACT :

BRAND : Ford Zephyr 1966

MADE IN : USA

PURPOSE : Hospital Car; Ambulance; Recreational

WEIGHT:

KERB WT : 26 cwt

WT DISTRIBUTION ON REAR AXLE : 42.5 cwt

WT DISTRIBUTION ON FRONT AXLE : 57.5 cwt

DIMENSIONS:

OVERALL LENGTH : 15'5"

OVERALL WIDTH : 5'1"

GROUND CLEARANCE : 6.5"

OVERALL HEIGHT : 4'1"

ENGINE

TYPE: Petrol engine

MAX POWER : 112 bhp at 4,800 rpm

MAX TORQUE : 137 lb-ft per sec at 3,000 rpm

COMPRESSION RATIO : 9:1

CYLINDERS : 6

CYLINDER ARRANGEMENT : V-type 60 degree

DISPLACEMENT : 211 cu in

STROKE DISTANCE : 2.33" per sec

CARBURATOR BRAND : Zenith 381 VT brand

CHOKE DIAMETER : 1'1"

COMBUSTION CHAMBER : Toroidal

VALVE GEAR : Overhead pushrod and rocker

General Theory of Knowledge Classification

SUSPENSION :

POSITION : Rear: independent, cylindrical helical spring; Telescopic damper; Front: Cylindrical helical spring, Anti-roll bar, Coaxial telescopic damper

WHEEL

WHEEL BASE : 9'7"

TRACK : Rear 4'1"; Front 4'9";

RIM (MATERIAL) : Pressed steel disc;

RIM (WIDTH) : 4'5"

RIM (FIXING) : 5-stud fixing

TRANSMISSION GEAR BOX

TYPE : Synchromesh

NO. OF SPEEDS : Reverse speed, Forward speed 4

GEAR RATIO : Reverse 3.346 to 1, Top 1 to 1, Third 1.412 to 1, Second 2.144 to 1, First 3.163 to 1

AXLE

REAR AXLE : Sprung

SHIFT : Manual hydrostatic

TRANSMISSION CLUTCH

BRAND : Borg and Beck

DISC DIAMETER : 9"

DISC ENGAGEMENT : Diaphragm

STEERING SYSTEM

GEAR RATIO : 20.6 to 1

NO. OF TURNING CIRCLES (LOCK TO LOCK) : 4.75

DIAMETER OF TURNING CIRCLES : 36'

BRAKE SYSTEM

ASSISTANCE : Hand assistance

CONSTRUCTION : Hand assistance

SWEPT AREA : Rear Brake: Disc diameter 9.91", Swept area 139 sq in;

Front brake: Disc diameter 9.63", Swept area 214 sq in

MAINTENANCE FACTORS

FUEL TANK CAPACITY : 15 Imp gallons

TYRE PRESSURE :

NORMAL DRIVING : Rear 24 psi, Front 24 psi

10.2 Hospital Patients Records Database

In December 1986, the Head of a large hospital in Bombay, India, feeling that case history records of hospital patients would be better managed and utilized if the information contained in them could be easily accessed and retrieved in response to specific queries of doctors and hospital administrators. Demonstration of the possibilities of computerizing the records was requested through the librarian of the institution.

Copies of patients' records of the Cardiac Department were made available and a database PREMA was developed using Micro CDS-ISIS ver. 1. Each record in the database was a mirror image of the corresponding case sheet, each field holding information on a particular attribute of the patient and/or the disease as recorded in the case sheets. Retrieval of information and records responding to specific queries was demonstrated to some fifty doctors.

Two years later a similar demonstration was organized in Madras, India, for heads of medical departments and specialized clinics. The system demonstrated included in addition to patients records database, a bibliographical database, a database of medical syndromes and an integrated database of profiles of experts and institutions. The retrieval of various types of information, such as bibliographic references, records of patients, syndromes data, profiles of experts and of institutions, responding to a single query was demonstrated.

General Theory of Knowledge Classification

The doctors realized that a system of the kind demonstrated could provide them access to valuable information recorded in the case sheets thus facilitating clinical diagnosis and treatment of specific diseases, comparative study of cases, education, research, writing of research papers and administrative actions.

Three doctors of different hospitals were subsequently given training at their request in the use of Micro CDS-ISIS as well as assistance in designing databases in specialized areas:

- surgical cases of pituitary tumours (PITU database)
- tumours of the CNS (TMR1 database)
- heart diseases in pregnancy.

On comparison it was noted that several fields were common to the databases PREMA, PITU and TMR1, and that it would be helpful to integrate them into a single database by enumerating all the fields in one Field Definition Table, creating a single input worksheet, and a single Field Select Table (FST). The fields of this data base are listed at the end of Sec. 10.2.

Examples of queries to this database:

In which cases, in how many cases and in what percentage of the total cases

- of optic atrophy and bilateral field defect in males above 25 years?
- of astrocytoma patients complained of temporal headache, visual and hearing disturbances?

- of prolactinoma with high prolactin level and negative immunochemistry for prolactin secreting tumours?
- had second transphenoidal?
- of glioma or glioblastoma had previous CNS surgery?
- of acoustic nerve tumours showed 8th bilateral nerve deafness, abnormal spinal motor, abnormal gait, third ventricle dilated, periventricular lucency absent and single tumour in CT scan of head?

Each record in the database was more or less a mirror image of the corresponding patient's case record as the fields were formulated and data entered by doctors, the end-users. The fields of a patient's record may be grouped into the following broad categories:

- **Patient's identity and administrative information** (e.g., name, identification numbers, file references, address, contacts, date of birth, age group, sex, blood group, dates of admission and discharge, admitted for, hospital ward/bed number, duration of stay, etc.)
- **Patient's general background** (e.g., religion, community, education, economic class, residence environment, occupation, work environment, etc.)
- **Complaints** of the patient
- **Signs and symptoms** observed
- **Interrogation findings** (e.g., personal history, family history, social history, social worker's data, house officer's notes, etc.)
- **Examinations/Investigations**

General Theory of Knowledge Classification

- **Treatment**
- **Follow-up**, progress, complications, prognosis
- **Recurrence**, re-admission, treatment, progress.

This organization of the content of the patient's record follows the organization of the details in the case records in broad categories and is in conformity with the Principle of Later in Time. The organization of the fields is reflected in the FDT (of the database), in the online worksheet and in the display format(s) as desired by end-users. In each of the above broad categories, the recorded information is divided into subunits and of the latter again divided further as decided upon by the specialist doctors. For example:

EXAMINATIONS/INVESTIGATIONS

General	Homones and levels
NMR	Growth hormone
Head CT scan	T3
Skull x-ray	T4
Cardiovascular system	TSH
ECG	ACTH
Four vessel angiogram	LH
Carotid angiogram	FSH
Ventriculogram	Cortisol
Respiratory system	Testosterone
Chest x-ray	Nervous system
Pneumoencephalogram	CT scan spine
	Muscular system
	Myelogram

Several display formats showing all fields or only selected fields were prepared to meet the needs of users.

Micro CDS-ISIS permits index-based search, free text search as well as a combination of these. The query:

In which cases, in how many cases and in what percentage of the cases x-ray skull abnormal, CT scan intrasellar, prolactin level greater than or equal to 50, consciousness level not less than 15 (Glasgow sc.) and transthemoidal surgery and transphenoidal, were done?

The free text search expression will be:

? val(v210)>=15 and val(v260)>=50 and v320:'abnorm' and v330:'intrasel' and v350:' transthemmo' and v352:'y'

Response:

Set 1: 7 hits out of a sample of 44 records (15.91%)

Further specification of the query:

Above query parameters plus Diabetes Insipidus as a Post-operative Complication?

Search expression:

? #1 v370:'diabetes insip'

Response:

Set 2: 1 hit out of 44 records (2.27%)

[In the display of the record below, names of the hospital, the department and of the patient and a few other details of the case are not displayed for reasons of confidentiality]

General Theory of Knowledge Classification

XXXXXXXXX MEDICAL CENTRE
DEPARTMENT OF XXXXXXXXXXXXXXXXX

Name of Patient : xxxxxxxxxxxxxxxx
Age : 28 ys
Sex : Female
Case no. : 588/88
Address : 34A Metcalfe Street, Calcutta-13
Admission date : 19.08.88
Discharge date : 04.09.88
Surgery date : 24.08,88
Endocrine symptoms : Yes
Endocrine hyperfunc : Hyperprolactin
Visual disturbances : No
Headache : Yes. Temporal
Vomiting : No
Visual acuity right : 6/6
Visual acuity left : 6/6
Cranial nerve palsy : No
Consciousness (GS) : 15
Higher functions : Normal
Hemiplegia : No
Cerebellar signs : No
Apoplexy : No
Prolactin : 71
Growth hormone : 1.4
X-ray skull : Abnormal; Sella Gr. II
CT scan : Intrasellar; Suprasellar
Surgery : Transethmoidal
Transphenoidal : Yes; Septum soft; Normal gland seen; subtotal
Postoperative compl. : Diabetes insipidus
Treat. recur./med : Bromocriptine
Treat. recur./DXRT : No

This database and the associated bibliographic database can be linked to the Syndrome database such that if the descriptor in the bibliographic database and/or the final diagnosis field in the patient's record specifies a syndrome name, the description of the syndrome will be retrieved from the Syndrome database. This is done using REF(L) function in the display format specification for the bibliographic and patients records databases or by using the CDS-ISIS Pascal interface SELECT.PAS.

With a view to facilitating navigation through the fields of a patient's record, especially for a new user, a thesaurus-like facility can be provided. This Pascal interface to the Patient Records database (PREMA or MEDIS) suggests in the first screen that the user may enter the term CASE and the system will display the broad categories of fields mentioned above. The cues in the submenu at bottom of the screen will enable the user to select a specific concept of interest in a series of successive screen displays.

***** ASSISTED SEARCH FACILITY *****

When you press any key, a screen will be displayed and you will be asked to keyin a search term. Type the term CASE, in upper or lower case. A list of the areas into which a case history has been divided will be displayed. Press the Space Bar as many times as necessary to move the cursor against the desired term e.g. INVESTIGATIONS. Press S key to select the term. Sub-divisions of and concepts related to INVESTIGATIONS will be displayed. As before move the cursor to the desired term and press S key to select. Use the options at bottom of the screen for formulating search queries.

Press any key _

General Theory of Knowledge Classification

CASE

CASE

- _ SN Patient Case History is divided into following areas
- _ NT Administrative information
- _ NT Complaints
- _ NT Symptoms/Signs
- _ NT Interrogation
- S N TInvestigation**
- _ NT Pathology
- _ NT Treatment
- _ NT Postoperative
- _ NT Follow-up
- _ NT Recurrence
- _ NT Readmission

[SpB] Next [B]ack [F]irst [P]age [S]elect [T]erm select [M]ake file add
te[R]m(file) [C]reate term [A]dd relation [D]elete [Q]uery sa[V]e [?]display
Query/save search[/]add e[X]it

INVESTIGATION

- BT CASE
 - NT ANGIOGRAM
 - NT BLOOD
 - NT CARDIOVASCULAR SYSTEM
 - NT CT SCAN
 - NT HORMONES
 - NT MYELOGRAM
 - NT NMR
 - S NT X-RAY**
- X-RAY
- BT INVESTIGATION
 - NT X-RAY CHEST

S NT X-RAY SKULL
NT X-RAY SPINE

X-RAY SKULL
BT X-RAY
QN ABNORMAL
NT SELLA GR I
QNT SELLA GR II

A similar interface was found useful with a database for supporting urban planning (see Sec. 10.5)

10.2.1 Integrated Database

In addition to the database of patient's records, databases CHEM and TOXIS consisting of records describing the physical and chemical properties, and pharmacological properties respectively of selected drugs, a database SYNDRO describing selected medical syndromes, and an integrated database (IIS) of bibliographical references, and of profiles of institutions, information systems, experts, and ongoing research projects, were developed for use of the doctors. IIS and SYNDRO have a common FDT, Field Select Table but separate data entry worksheets. When a descriptor in a record retrieved in response to a query is the name of a syndrome, it will also retrieve the related record from the SYNDRO database. Again, this is achieved either by using the REF(L) function of MicroIis or by using the Pascal Program SELECT.PAS [60].

SAMPLE RECORD FROM IIS BIBLIOGRAPHIC DATABASE LINKED TO SYNDROME DATABASE

TITLE: Moyamoya-like disease associated with a lenticulostriate region aneurysm : Case report

AUTHOR(S): Grabel Jordan C.; Levine Mitchell; Hollis Peter; Ragland, Ronald

ABSTRACT : A case of moyamoya-like disease associated with an intracerebral hemorrhage and an unusual aneurysm is reported. The patient's

General Theory of Knowledge Classification

clinical status and the initial and follow-up angiographic appearance of the aneurysm are presented. The etiology of the moyamoya phenomenon and of associated aneurysms are discussed.

KEYWORDS: Cerebral aneurysm; Intracerebral hemorrhage; Moyamoya disease; Pseudoaneurysm

IN: J Neurosurg 70(1989):802-803

[As the term Moyamoya disease occurs as one of the descriptors in the above bibliographical record, the system retrieved also the corresponding record describing the Moyamoya syndrome from the SYNDRO database]

SYNDROME

NAME: MOYA-MOYA syndrome

SYNONYM: Progressive arterial intracranial occlusions; Intracranial arteries; progressive occlusions; Kawakita's syndrome; Leed's syndrome; Maki's syndrome; Multiple progressive intracranial arterial occlusions; Taveras syndrome.

SYMPTOMS: Onset from infancy to adulthood, usually following some non-specific infectious process or cold.

INCIDENCE: Most cases in Japanese; a few Black races and Caucasian patients.

INCIDENCE-SEX: M, F.

INCIDENCE-AGE: Infancy to youth.

SIGNS: Paralysis and focal epileptic attacks, alternating between both sides, together with twitching, speech disturbances, unsteady gait, hemianopia, and headache. In some cases especially adults, psychiatric manifestations which may be followed by signs of intracranial hemorrhage.

ETIOLOGY: Unknown; considered a non-specific inflammation due to autoimmune reaction.

PATHOLOGY: Occlusion of distal internal carotid artery and middle cerebral arteries, and sometimes basilar arteries and proximal posterior cerebral arteries. Absence of atherosclerotic changes.

DIAGNOSTIC PROCEDURES: X-ray: of skull usually normal. Angiography: Occlusion usually situated in internal carotid artery at its bifurcation; develop-

ment of a large network of vessels in basal ganglia and upper brain stem areas from basilar artery and trunk of anterior and middle cerebral arteries. Marked degree of vascularization and visualization of rete mirabile. Spinal tap: Occasionally hemorrhagic.

THERAPY: Corticosteroids apparently not effective. Surgery for removal of obstruction attempted.

PROGNOSIS: Progress to complete obstruction of major arterial network along base of brain that eventually stops spontaneously. Majority of patients survive with moderate or no disability. Mental retardation in one-third of children.

REFERENCES : Kawakita, Y.; Abe, K.; Miyata, Y. et al. Spontaneous thrombosis of internal carotid artery in children. (*Folia Psychiatr. Neurol. Japan.* 19;1965; 245-255); Leeds, N.E.; Abbott, K.H. Collateral circulation in cerebrovascular disease in childhood via rete mirabile and perforating branches of anterior choroidal and posterior cerebral arteries. (*Radiology* 85;1965;628-634); Kudo, T. Spontaneous occlusion of circle of Willis; disease apparently confined to Japanese. (*Neurology* 18;1968; 485-496); Taveras, J.M. Multiple progressive intracranial arterial occlusion... (*Amer. J. Roentgenol. Rad. Ther. Nucl. Med.* 106; 1969;235-268).

SOURCE: DMS.

RECORD ON ASPIRIN FROM TOXIS DATABASE

ACETYSALICYLIC ACID

CATEGORY: Analgesic.

ALTERNATIVE NAMES: Aspirin.

OCCURRENCE/USE: Analgesic. Antipyretic. Anti-rheumatic.

PROPERTIES: [For more details use CHEM data base]. MP 137°C. Odourless. Bitter taste. White crystalline soluble in ethanol, ether, or ethylene dichloride.

SIGNS/SYMPTOMS: Acidosis, metabolic acidosis as a later effect. Nausea and vomiting may be early or delayed. Large doses first stimulate and then depress the CNS; initial excitement and convulsions are followed by stupor and depression. Central stimulation may also produce hyperventilation (loss of carbon dioxide) and early transient alkalemia; followed by severe metabolic acidosis. Other possible symptoms: headache, thirst, dizziness, irritability, cyanosis, p-aminophenol in urine (brown black urine) diaphoresis, sugar in

General Theory of Knowledge Classification

urine, gum bleeding, gastritis with or without gastric hemorrhage, prothrombin level lowering, hearing disturbance, vision disturbance, tinnitus, vasodilatation, peripheral, hyperpnea with dyspnea, renal damage, brain damage, weakness, fatigue, hypotension, delirium, confusion, respiratory failure, collapse, coma, death.

EXPOSURE: Oral, Tablet.

MLD: Approx. 15 g for a 150 pound person. Sodium salicylate: c.15 g. Phenyl salicylate: c.5 g. Methyl salicylate: 10 ml. Salicylic acid: 8 g.

METABOLISM/TOXICITY: First completely hydrolyzed in plasma to salicylic acid, most of which is eliminated unchanged. Part is conjugated and a very small amount oxidized to gentisic acid.

INFLUENCING FACTORS: [For a checklist of influencing factors use ACTORS data base].

TREATMENT: Gastric lavage with warm water; or 15 ml of syrup of ipecac to produce vomiting in all cases. Patient may appear deceptively well when first seen; signs and symptoms may be delayed; vomiting may be delayed. Fluids (oral or parenteral) to increase urinary output and to counteract de-hydration due to vomiting and sweating. Maintain body heat, fluids, and electrolyte balance. Keep patient warm and quiet. Give alkaline drinks. Give sodium bicarbonate or sodium lactate (IV, M/6) or citrate to combat acidosis and sodium loss as necessary. Periodic blood determinations for alkaline reserve to estimate severity of poisoning and progress of treatment. Recovery may depend on renal elimination and controlled acidosis. Elimination is more rapid if urine is kept alkaline. Vitamin K derivatives and ascorbic acid for low prothrombin and capillary fragility. In severe cases exchange transfusion for children, or hemodialysis.

[The CHEM and TOXIS databases can be linked in the same way as IIS and SYNDRO are linked so as to retrieve corresponding records from the two databases]

RECORD ON ASPIRIN FROM CHEM DATABASE

***** ASPIRIN *****

2-(Acetyloxy)benzoic acid; Salicylic acid acetate.

ALTERNATIVE NAME(S): Acetylsalicylic acid; 2-Acetoxybenzoic acid; Acidum acetylsalicylicum; Acetilum acidulatum; Acenterine; Aceticyl; Acetophen; Acetosal; Acetosalic acid; Acetosalin; Acetylin; Acetyl-SAL; Acemetten; Acylpyrin; A.S.A; Asatard; Aspro; Asteric; Caprin; Colfarit; Contrheuma retard; Delgesic; Duramax; Ecotrin; Empirin; ECM; Endydol; Entrophen; Enterosarine; Helicon; Neuronika; Rhodin; Salcetin; Salcetogen; Salcetin.

CATEGORY: Analgesic. Antipyretic. Antirheumatic. Anti-inflammatory.

MOLECULAR FORMULA: C₉H₈O₄.

MOL. WT.: 180.5.

STABILITY: Stable in dry air. Long half-life, about 24 hours.

PHYS/CHEM.PROPERTIES: C 60.00 p.c., H 4.48 p.c., O 35.53 p.c. Manufactured from salicylic acid and acetic anhydride. Odourless, but in moist air gradually hydrolyses into salicylic acid and acetic acid and acquires the odour of acetic acid. Bitter taste. White crystalline monoclinic tablets or needle-like. MP 135°C (rapid heating); the melt solidifies at 118°C. Soluble in ethanol, chloroform, ether or ethylene chloride. 1 g dissolves in 300 ml water at 25°C, in 100 ml water at 37°C, in 5 ml alcohol, 17 ml chloroform, 10-15 ml ether. Less soluble in anhydrous ether. Decomposed by boiling water or when dissolved in solutions of alkali hydroxides and carbonates. Pharmaceutical incomp.; Forms a damp to pasty mass when triturated with acetalide, phenacetin, antipyrine, aminopyrine, methanamine, phenol or phenyl salicylate. Aspirin powder with an alkali salt becomes gummy on contact with atmospheric moisture. Hydrolysis occurs in admixture with salts containing water of crystallization. Solutions of the alkali acetates and citrates, as well as alkalis themselves, dissolve aspirin but the resulting solutions hydrolyse rapidly to form salts of acetic and salicylic acids. Sugar and glycerol hinder this decomposition. Liberates very slowly hydrochloric acid from potassium or sodium iodide. Subsequent oxidation by air produces free iodine.

IDENTIFICATION: Salicylates are rapidly absorbed and may be easily detected in the urine within 30 min. As little as 0.3 g. tablet gives a positive test in the urine. Separated by acid-ether extraction. Specimen must be strongly acid to affect a good extraction. Aspirin (salicylates): (1) To 5 ml of urine, add 1 ml of 10 p.c. ferric chloride. Heat the urine gently to rule out possible false reaction due to acetone bodies. Salicylate gives purple colour. This test also positive for sodium, phenyl, or methyl salicylates, and phenol derivatives. (2a)

General Theory of Knowledge Classification

Acidify 5 ml of serum by adding 0.5 ml of conc.HCl. Extract with 10 ml of ethylene dichloride. Allow layers to settle and then separate. Discard upper blood layer and add 2 ml of 0.2 p.c. ferric nitrate to the ethylene dichloride. Salicylates give purple colour. The colour can be compared with standards similarly prepared or may be read in a spectrometer. (2b) (Nelson): Place 0.01 ml of serum or urine on a white porcelain dish; add one drop of 1 p.c. ferric nitrate (in 0.07N HNO₃). Purple colour for salicylates. Report as negative, faint, moderate, or large amounts. (2c) Quantitative: Measure 0.2 ml of serum or urine plus 0.8 ml of water into each of two small tts. To one tt (which will be the blank), add 1 ml of 0.07N HNO₃. To the other (which will be the unknown), add 1 ml of 1 p.c. Fe(NO₃)₃. Mix and allow to stand for about 5 min; read in a spectrophotometer at 540 mf. For standard: Use 1 ml of water in a tt (blank) and add 1 ml of 0.07N HNO₃. To another tt, add 0.8 ml of water and 0.2 ml of standard and then 1 ml of 1 p.c. Fe(NO₃)₃. Absorbance unknown/Absorbance standard x 25 = mg of salicylate per 100 ml. Reagents: Ferric nitrate: 1 p.c. in 0.07N nitric acid. Nitric acid: 0;07N (4.69 ml of HNO₃ [1.42] and 70.5 p.c. made up to 1 litre. Salicylate standard: 25 mg per 100 ml; 29 mg of sodium salicylate or 25 mg of salicylic acid per 100 ml of water. Thiocyanates interfere if taken in large amounts. (3) Benedict's reagent positive for sugar in urine. (4) Pure crystals extracted from urine can be differentiated with 1 ml of ferric chloride (10 p.c.). Antipyrine + ferric chloride gives red colour. Salicylate + ferric chloride gives deep violet colour. Amynopyrine (Pyramidon) + ferric chloride gives fading light violet colour. Methyl salicylate + ferric chloride gives violet colour. (Must add alcohol to dissolve methyl salicylate). Also by NIR spectrophotometry. (5) UV spectrophotometry: Acidify specimens. Use 5 ml of serum or tissue plus 0.5 ml of conc.HCl. Gently extract with 25 ml of chloroform; reextract with 25 ml of chloroform; reextract with 0.9N NaOH and scan between 230 and 340 mf. Max density, 300 mf; min density, 265 mf. (6) TLC may be used. Salicylates can be detected in urine within 1 hr after ingestion and may be found upto 24 hrs or more depending upon the quantities taken.

DERIVATIVES: Sodium salicylate. Phenyl salicylate. Salol. Methyl salicylate. Oil of winter green. Salicylic acid.

REMARKS: All tests described are for the salicylate radical. Compound radicals such as methyl salicylate, salol must first be hydrolyzed if tests are done on tablets or original compounds. Levels above 20 mg p.c. may be considered toxic.

REFERENCES: Kaye, Sidney / Handbook of emergency toxicology / 3rd ed / 1971; Merck index / 10th ed / 1983.

Fields of PITU and TMR1 databases combined (Subfields are indicated in parenthesis)

Institution	PC on duty
Institution acronym/code	Attending officer (name, designation)
Department/Service	Admitted on (date, time)
Head of Department (name, designation)	Admitted for
Name of patient	Provisional diagnosis
Patient type	Final diagnosis
Patient ID	Secondary diagnosis
OPD number	Discharged (status, transfer, hospital, date)
Casualty number	Expired (date, time)
Period of stay	Cause of death
Ward/bed number	Patient's complaints
Previous admission nos.	House officer's notes
File references	Personal history
Address of patient (address, phone..)	Family history
Contacts (name, address, phone...)	Social history
Associated persons/instns (name, relation, address)	Other history
Recommended by (name,date)	Social worker's data
Date of birth	Symptoms/general
Age at admission (age, range)	Headache (y/n, site)
Sex	Vomiting
Blood group	Convulsions
Marital status	Gait
Religion/community	Visual acuity (left, right)
Occupational class	Colour vision
Economic class	Visual fields
Residence environment	Lobar function
Educational level	Diplopia
Employment (function, environment)	Ptosis
Nationality	Nystagms
Race	Ophthalmoplegia
Brought by	Examination/general

General Theory of Knowledge Classification

Cardiovascular system

ECG

X-ray chest

X-ray skull

Pupil changes

Ocular nerve

Cranial nerve

Fundus

X-ray spine

Four vessel angiogram

Carotid angiogram

CT scan head

NMR

Pneumoencephalogram

Consciousness/Glasgow sc.

Higher function

Hemiplegia

Paraplegia

Quadriplegia

Monoplegia

Cerebellar signs

Apoplexy

Memory

Speech

Anosmia

Phonation

Swallowing

5th nerve

5th corneal

5th sensory

5th motor

6th nerve

7th nerve

8th nerve

9th nerve

10th nerve

11th nerve

12th nerve

Spinal Motor

Reflexes

Sensation

Extensor plantar

Sphincters

Urinary continence

Rectal continence

Associated illnesses

Endocrine symptoms

Endocrine hypofunction

Endocrine hyperfunction

Prolactin

Growth hormone

T3

T4

TSH

ACTH

LH

FSH

Myelogram

Ventriculogram

CNS surgery

Pathology

Complications

Treatment/medical

Reactions

Surgery

Attending doctors

Transphenoidal

Postoperative complication

Prognosis

Follow-up visual

Follow-up hormonal

Follow-up CT scan

Pathology altered

Operated more than twice

Progress

Advice

Readmissions (date, time)	ACTH
Recurrence	GH
Immunochemistry	Recurrence treatment/medical
TSH	Recurrence treatment/surgery
LH	Recurrence treatment DXRT
FSH	International disease classif.
PL	Keywords/Descriptors

10.3 Databases to Assist Energy Resources Planning

Information specialists interacting with State energy planners (end-users) helped in identifying and visualizing the broad issues and aspects of energy resources planning. These were then organized into a matrix/framework as shown below:

Framework for Energy Planning Information Analysis

Issue relating to energy resource	Aspect			
	Current status	Forecast	Policy	Impact
End uses				
Energy demand		B		
Energy supply	A			
Energy balance				
Energy flow				
Physical plants		D		
Energy prices			C	
Social data				
Economic data				
Energy research				
Environmental issues				

General Theory of Knowledge Classification

A box formed by the cross-section of an issue and an aspect represents a specific area of interest of end-users. For example:

- Box A: Current status of energy supply
- Box B: Forecast on energy demand
- Box C: Current policy on energy pricing
- Box D: Future developments in energy plant technology

This analytical framework may be applied either to a particular form/source of energy (e.g., coal, solar, electricity, wind) or to a group of energy sources (e.g., non-conventional energy).

Each box can be further divided by geographical area or by local, district, state, national, and international levels. Each domain so formed can be used for providing SDI service as well as analyzed, consolidated, repackaged information. It may also be noted that a particular interest area, such as:

Electricity consumption in the industrial sector in
Tamil Nadu in India during 1990-1992

can be structured using the postulational approach to facet analysis and synthesis as follows:

Electricity [P]: Consumption [E]-Domestic sector (Qualifier).
Tamil Nadu (India) [S] '1990-1992 [T]

The framework helped to determine the fields and sub-fields, for an FDT, the fields and concepts to be indexed and arrangement of data

in the output from the database. The issues were used as fields and the aspects and spatial divisions were used as descriptors. Sample of output from the ENERGY database is given below.

Other foci of interest were also identified relating to specific energy forms, as flow models, from source to end uses (*see also* [38]). From these, fields and subdivisions for a database as well as a scheme for classification were formulated.

For example:

Coal

By Source/Extraction

Mining

Deep mine

Strip mine

Auger mine

Refuse mine

Import

By Refining

Cleaning

Refuse

By Transportation means

Rail

Truck

River

Overseas

Other

By End use

Electric utilities

Industry

Coke

Gas

Byproduct

Commercial use

Domestic use

SAMPLE RECORD FROM ENERGY DATABASE

FILE : NONCON - 00007

TOPIC : Nonconventional energy equipment.

DETAILS ON : Demand and market for non-conventional energy equipment.

COVERAGE: World. Developing countries.

CURRENT SITUATION

Considerable potential exists for expanding trade among developing countries in nonconventional energy producing equipment. The demand for such equipment as wind pumps, wind generators, water turbines and solar panels is expected to grow rapidly during the remainder of this century and

General Theory of Knowledge Classification

into the next. International trade in such equipment has not yet developed, but there is a rapidly growing and attractive market which should be explored by developing countries.

FORECAST

Demand already exists in some developing countries in which non-conventional energy-generating technologies are being applied, for example, Argentina, Brazil, China, India, Mexico and the Republic of Korea. Future sales prospects are much greater especially if there be change in oil price situation along with the expected overall expansion of energy needs in the long term. Demand in industrialized countries is also expected to follow an upward trend. Several of them are already well into nonconventional energy production. In USA, for example, a national programme has led to widespread use of solar energy systems in some parts of the country. This market as well as others in industrialized countries, could offer sales opportunities for developing countries. Although the overall market for nonconventional energy equipment is difficult to define because of the scarcity of data, general indicators point to attractive future sales prospects for developing countries for several types of equipment, more particularly wind energy converters, photovoltaic cells, small hydropower machinery and solar cookers.

POLICY

The viability and extent of application of energy systems depend greatly on government policy. One of the reasons for the development of the market for non-conventional energy technologies in the USA was the very favourable policy of the government: a national goal was set to obtain 20 percent of national energy supplies from solar energy by 2000 A.D. This was supported by tax reductions for solar energy installations. The policy also led to the establishment of large "wind farms" on the east and west coasts. Other countries that have adopted national policies on energy systems development including nonconventional energy systems, have also achieved good results. The energy substitution project in Cape Verde is a good example.

OPERATIONAL EXPERIENCE

Developed countries, particularly in Europe, are relatively more active in the trade of nonconventional energy equipment. Almost all exports of wind turbines and water turbines and about two-thirds of such imports over the past several years have been accounted for by developed countries. Developing countries have been importing about one-third of the world total, and they export

about 2 per cent of the total traded. This export share is quite small given the relatively simple technology required to produce low-powered non-conventional energy equipment.

SOURCE OF INFORMATION

1. Vorobiev, Andrei (International Trade Centre, UNCTAD/GATT, Geneva). Power-generating and transforming equipment for nonconventional energy: an overview of market prospects. *International Trade Forum*, 1987 April-June.
2. Attractive market prospects for nonconventional energy equipment. *Asia-Pacific Tech Monitor* (APCTT), 1987 September-October, p. 20-22.

FILE : WIND - 00001

TOPIC : Wind Energy Equipment.

DETAILS ON : Demand and market for equipment and machinery — Wind pumps and wind mills.

COVERAGE : World. Developing countries.

CURRENT SITUATION

It is estimated that 10 per cent of the energy needs of most countries could be met by wind energy conversion, using contemporary technology. The areas with the best prospects are seashores, small islands, flat plains and mountainous regions. The number of wind pumps or wind mills in operation throughout the world is estimated at about 1 million. Most of these are installed in the plains of USA, Australia and Argentina.

FORECAST

The major prospective markets for wind pumps are countries frequently stricken by droughts and floods, those seeking to improve their water supply or those already widely using wind power. The latter are replacement markets or markets requiring spare parts. About 25,000 new wind mills will be required globally during the next 10-15 years, amounting to a market value of US\$ 50 million. The demand may go up to 40,000 wind mills, that is, a market value of US\$ 80 million. As wind pumps are simple in design and can be/are being manufactured in many developing countries, the latter could supply a sizeable share of the market, including sales to industrialized countries, such as, the USA. The development of the market for replacement windmills and spare parts

General Theory of Knowledge Classification

is another possibility. The market demand between the years 1986 and 2000 is estimated at US\$ 50 million to 80 million for replacement windmills and US\$ 10 million to 15 million for spare parts.

OPERATIONAL EXPERIENCE

In Cape Verde a programme and strategy for wide utilization of nonconventional energy equipment including wind pumps and wind generators is underway. The call is for the installation of about 400 wind machines, of which 320 have already been imported and set up and 18 have been locally made. About 52 of the machines are for generating electricity. A number of other developing countries also have the potential to produce energy from wind, for example, Chile, Syria, Lesotho, Cyprus and Argentina.

SOURCE OF INFORMATION

1. Vorobiev, Andrei (International Trade Centre, UNCTAD/GATT, Geneva). Power-generating and transforming equipment for nonconventional energy: an overview of market prospects. *International Trade Forum*, 1987 April-June.
2. Attractive market prospects for nonconventional energy equipment. *Asia-Pacific Tech Monitor* (APCTT), 1987 September-October, p.20-22.

An associated database is on equipment, specifically for biogas generation. In this case the users required data on the characteristic features of such equipment. A record output from the EQUIP database is presented below:

SAMPLE RECORD FROM EQUIP DATABASE

FILE: EQUIPMENT

FIELD : Biomass

EQUIPMENT: Gasifier

BRAND & MODEL: Charung.

CHARACTERISTICS

Rated capacity	90 KVA
Gas flow rate	3-5 cu.m per min. (approx)
Average calorific value	750-1000 K Cal/cu.m

Gas temp. at gasifier outlet	60°C to 80°C
Rice husk consumption	2 to 2.5 kg/KVA
N of hrs. in operation	Over 11,000 hrs
Permissible contin. operation	125 to 200 hrs.
Full load operation rate	About 240 days in a year
Expected diesel replacement	100 per cent
Rate of tar deposition	5 l tar + water every 8 hrs.
Construction materials	Gasifier - stainless steel; Other parts:mild steel
Ash output	50 to 100 kg/hr. automatic removal

BIOMASS

Biomass type	Rice husk, uncharred
Acceptable moisture	Below 18 per cent

MACH INERY & EQUIPMENT

Blower & water pump	Blower - centrifugal 2 HP rating; Water pump - centrifugal 2-3 HP rating
Diesel engine & alternator	Diesel engine - 280 HP at 2300 rpm Nissan engine, Japan (modified to gas engine spark ignition; Alternator - 100 KVA, 3-phase
Control/protections provided	Starter control, phase, overload, reverse current and auto off protections
Auxiliary power req.	10 per cent of rated capacity to start with, from the grid, and the reafter 100 p.c. self-sufficient and no auxiliary power required
Site requirements	6m x 12m x 6m. Foundation for engine and motor only

General Theory of Knowledge Classification

TOTAL INVESTMENT

Total	50-60 per cent of total cost
Husk	25-30 per cent
Labour & maintenance	7 percent

ADVANTAGES

The rice husk produced electricity is cheaper than grid electricity - US\$ 0.038 to 0.049 as compared to US\$ 0.060 per unit. If there is no expenditure on husk, the cost comes down even further. Company claims that the plants are economical and most of the factories where the units are installed save enough energy that the gasifier will repay itself within 1 to 1.5 years.

LIMITATIONS

World Bank Expert Study Report points out that the company's figures regarding economics of the system have been arrived at taking into account only saved electricity charges, equipment and maintenance costs. If all other costs, such as of stand-by equipment, building, installation and training are included, the payback time would be more than two years.

SOURCE/VENDOR: Charung Engineering Co. Ltd. Thailand

DISTRIBUTOR-FOREIGN: Charung Engineering Co. Ltd. Thailand

REMARKS

The Charung gasifier is being promoted by APCTT. The Charung Engineering Co. has installed more than 20 units, mostly in rice mills and ice and jute factories indifferent parts of Thailand.

SOURCE OF INFORMATION

Rice husk gasifier power plants. *Asia-Pacific Tech Monitor*, 1987 September-October, p. 17-19.

10.4 Database on Forest Resources

The database forms a component of an information system developed at a forest research institute, and relates to forest divisions in a specific geographical area in India. Categorization of the geographical area into districts, and the physical features, meteorological

data, features of the mountains in each district, farm products, etc. were identified by the librarian of the institute through interactions with the end-users and using appropriate information source materials. The interaction also helped in determining the fields and sub-fields, the sources of data and the type of outputs required frequently. For example, information/data relating to:

- a district or forest division as a whole;
- the products of a forest division in a district;
- temperature range (maximum and minimum) in a district;
- farm products or plantation crops output for the different districts over a period of years in a comparative tabular form.

The interaction between users and librarian helped to decide:

- that a record should provide data on a specific forest division (geographical area);
- that each field of a record should provide information on a specific characteristic of the forest;
- appropriate indexing parameters; and
- on predefined output formats, etc.

The grouping of the main fields using principles of helpful sequence has been discussed earlier. For each data entity constituting a field, certain aspects were specified. Unlike in the Energy database discussed above, these aspects were defined as subfields. Some examples follow:

General Theory of Knowledge Classification

Field	Sub-fields
Peaks	Name, height, location, forest range
Sanctuaries	Name, animals, birds, location, forest range, area, year
Farm forestry	Species, yield, expenditure, revenue, year
Tribal settlements	Name of settlement, location, tribe, designation, number of settlements
Forest fires	Locality, range, area affected, year

Interests of users about each entity enumerated in the main field correspond to Property, Space and Time facets enumerated in the sub-fields.

Sample records using different output formats are presented below:

SAMPLE RECORDS FROM 'FORD' DATABASE (DISPLAY FORMAT FORD)

PALAGHAT DIVISION

HEADQUARTERS AREA	Mannarkadu, 680 694
POPULATION	5,96,600
NUMBER OF SCHOOLS	15
HOSPITALS	6
RANGES	4
RANGERS	4
FORESTERS	15
GUARDS	32
VEHICLES	6

DIVISIONAL FOREST OFFICER: Ramamurthy, K.N.

RANGES

Name : Mannarkadu
Headquarter : Mannarkadu PIN. 680 654
Ranger : Purushotham,P.K

Population : 1,25,313
No. of Sections : 15
No. of Beats : 75

Name : Panallur
Headquarter : Puthukotta PIN. 680 313
Ranger : Meghanathan,P.M
Population : 85,460
No. of Sections : 13
No. of Beats : 55
Name : Eriyad
Headquarter : Azhikode PIN. 680 304
Ranger : Krishnankutty,R
Population : 1,75,510
No. of Sections : 14
No. of Beats : 53

Name : Porkulam
Headquarter : Perinchantha PIN. 680 215
Ranger : (vacant)
Population : 2,10,217
No. of Sections : 20
No. of Beats : 83

* * * * *

LOCATION	Latitude 12.01; Longitude 10.38
ELEVATION	min. 235 max. 878
PEAKS	Paramudi (878); Thekkumudi (856)
RIVERS	Bharathapuzha; Palanadhi
RESERVOIRS	Malampuzha

General Theory of Knowledge Classification

IRRIGATION PROJ TEMPERATURE	Malampuzha	
	Year	Celsius min max
	1984	20 31
	1985	19 32
	1986	18 37
	1987	23 29

RAINFALL	Year	Cm
	1984	107
	1985	108
	1986	98
	1987	123

FOREST TYPES Evergreen (code III)
Deciduous (code IV)
Moist Deciduous (code VII)

FARM FORESTRY	Species	Year	Area
	Teak	1984	36
		1985	40
		1986	40
		1987	45
	Eucalypt	1984	28
		1985	36
		1986	30
		1987	33
	Bamboo	1984	25
		1985	30
		1986	22

xxxxxxx RECORD INCOMPLETE xxxxxxxx

(Display format FARMPR)

PALAGHAT DIVISION - FARM PRODUCTS		
Species	Year	Area
Teak	1984	36
	1985	40
	1986	40
	1987	45
Eucalypt	1984	28
	1985	36
	1986	30
	1987	33
Bamboo	1984	25
	1985	30
	1986	22

(Display format PLANT1)

FARM PRODUCTS: COMPARATIVE DATA

Division	Teak				Eucalyptus				Bamboo			
	1984	1985	1986	1987	1984	1985	1986	1987	1984	1985	1986	
Palaghat	36	40	40	28	45	36	30	33	25	30	22	
Trichur	26	46	42	35	38	45	32	23	27	29	32	
Idukki	46	40	39	46	38	32	40	23	34	24	32	
Trivandrum	43	45	23	35	38	23	51	42	35	32	21	
Kottayam	45	40	40	32	38	36	29	42	35	30	21	

(Display format TEMPER)

TEMPERATURE Division:Palaghat		
Year	Celsius	
	min	max
1984	20	31
1985	19	32
1986	18	37
1987	23	29
1988	22	29

Other out put formats such as TRIBE (for data on tribes), FIRE (for Forest fire data), INJURY (for forest damages data) etc. have been prepared. In arranging The Fields and Data in each field (Row-wise and Column-wise) appropriate Principles of Array Isolates sequence have been used.

10.5 Information Support for Urban Planning

10.5.1 Information Request Form

For assisting users in expressing their information requirement as precisely as possible, a structured query worksheet was developed (printed and online format) [67]. The query elements or facets were derived from studies on data requirements of users and the fields of the related databases discussed above. CDS-ISIS Pascal programmes facilitate matching of the terms in the query worksheet with those in machine-readable vocabulary control tools, such as, a thesaurus, scheme for classification, etc. and/or the index terms of the database(s), thus enhancing the retrieval performance of the system. An example of the Search Request worksheet is given below:

NATIONAL URBAN PLANNING INSTITUTE INFORMATION
SERVICES CENTRE
Structured Information Search Request Form

Note: Tick mark against the search topic of your interest

Page 1 of 3

Information Category	Search Topic
1. Maps and Plans	<input type="checkbox"/> Maps: <input type="checkbox"/> Original <input type="checkbox"/> Semi-original
	<input type="checkbox"/> Subject _____
	<input type="checkbox"/> Scale _____
	<input type="checkbox"/> PLans
	<input type="checkbox"/> Subject _____
	<input type="checkbox"/> Scale _____
	<input type="checkbox"/> Type _____
	Other _____
2. Physical characteristics	<input type="checkbox"/> Location
	<input type="checkbox"/> Topology
	<input type="checkbox"/> Topography
	<input type="checkbox"/> Meteorology
	<input type="checkbox"/> Temperature
	<input type="checkbox"/> Humidity
	<input type="checkbox"/> Rainfall
	<input type="checkbox"/> Natural Resources
	<input type="checkbox"/> Water <input type="checkbox"/> Mineral <input type="checkbox"/> Soil <input type="checkbox"/> Forest
	Other natural resources _____

<input type="checkbox"/> Other _____	

General Theory of Knowledge Classification

Information Category	Search Topic
3. Population (Demography)	<ul style="list-style-type: none">_ Population size_ Population distribution<ul style="list-style-type: none">_ By age_ By sex_ Population density_ Population dynamics<ul style="list-style-type: none">_ Fertility_ Birth rate_ Death rate_ Population projection_ Population structure_ Migration<ul style="list-style-type: none">_ Rural-Urban_ Urban-Urban_ Attraction (Push-pull factors)_ Birth place<ul style="list-style-type: none">_ Rural _ Urban _ Region_ Length of residence data <p>Other _____</p> <p>_____</p>

Information Category	Search Topic
4. Housing	<ul style="list-style-type: none">_ Housing size_ Age of housing unit_ Housing tenure<ul style="list-style-type: none">_ Owned_ Rented_ Rent distribution_ Purpose of housing_ Residence _ Business_ Residence and business

Information Category	Search Topic
	<ul style="list-style-type: none">_ Typology<ul style="list-style-type: none">_ Unit type_ No. of rooms_ Building type_ Service quarters_ Housing structure/material<ul style="list-style-type: none">_ Foundation_ Walling_ Roofing_ Flooring_ Kitchen_ Ceiling_ Housing services<ul style="list-style-type: none">_ Water_ Toilet_ Telephone_ Garbage collection_ Bathing_ Lighting_ Fuel/Energy <p>Other _____</p>
	<ul style="list-style-type: none">_ Housing demand_ Construction plan_ Household size_ Total size_ Occupancy rate <p>Other _____</p>
5. Transport and Communications	<ul style="list-style-type: none">_ Origin-destination_ Private transport_ Public transport_ Road network_ Pedestrian path_ Route capacity_ Mode of transport <p>Other _____</p>

General Theory of Knowledge Classification

Information Category	Search Topic
6. Social infrastructure	<ul style="list-style-type: none"> _ Education facilities <ul style="list-style-type: none"> _ School type _ Student class ratio _ Eligible population size _ Demand for education facilities _ Libraries _ Health facilities <ul style="list-style-type: none"> _ Hospitals _ Clinics _ Health stations _ Pharmacies _ Religious facilities _ Entertainment facilities _ Sports facilities/centres _ Cinema/theatre halls <p>Other _____</p>

Page 3 of 3

Information Category	Search Topic
7. Economic activity	<ul style="list-style-type: none"> _ Occupation type _ Occupational status <ul style="list-style-type: none"> _ Government employment _ Self employed _ Private employment <p>Other _____</p> <ul style="list-style-type: none"> _ Economic sector <ul style="list-style-type: none"> _ Agriculture _ Industry/Manufacture <ul style="list-style-type: none"> _ Modern _ Cottage _ Building/construction _ Installation/maintenance

Information Category	Search Topic
	<ul style="list-style-type: none">_ Service_ Trade_ Transport/communication_ Finance; Insurance_ Public administration_ Labour force participation_ Dependency ratio_ Economically active<ul style="list-style-type: none">_ Female_ Male
	Other _____
8. Finance and management	<ul style="list-style-type: none">_ Income distribution_ Expenditure distribution_ Management and administration<ul style="list-style-type: none">_ Responsible persons_ Accountability_ Power hierarchy
	Other _____
9. Legal	<ul style="list-style-type: none">_ Planning policy_ Building and housing standards_ Housing policy_ Landuse policy
	Other _____
10. Historical	<ul style="list-style-type: none">_ Year founded_ Growth trend_ Founder_ Previous plans
	Other _____
11. Other information not specified above	

General Theory of Knowledge Classification

10.5.2 Thesaurus-like Search Interface

An assisted search facility, similar to the one mentioned in Sec. 10.2 was developed. The categorization or broad structuring of the database contents parallels the categorization in the Search Request form given in the preceding section (10.5.1)[68].

**ASSISTED SEARCH FACILITY USER INTERFACE FOR
URBAN PLANNING DATABASE**

1. When you press any key a blank box will be displayed and you will be asked to enter a search term in the rectangular blank box. Type in the term DATA in lower or upper case. A list of the areas into which the Urban Planning Database has been divided will be displayed. Use the Space Bar to move the cursor to the desired term, e.g. "POPULATION" and press S key to select. Sub-divisions and concepts related to POPULATION will be displayed. Move the cursor, as before, to the desired term and press S key to select. Use options at the bottom of the screen for formulating search queries.
2. You may also key-in the term (or a stem of it) representing your interest and press <ENTER> key. A list of the terms beginning with the term or stem string you have keyed-in will be displayed. Move the cursor, using Space Bar to the desired term and press S key to select. Then proceed as mentioned at 1.

Press any key _

DATA

DATA

- SN Urban data is divided into the following main areas
- NT Physical characteristics

S NT Population

- NT Housing
- NT Transport and Communication
- NT Social infrastructure
- NT Maps and plans
- NT Economic activity
- NT History
- NT Finance and management
- NT Legal
- NT Other

[SpB] Next [B]ack [F]irst [P]age [S]elect [T]erm select [M]ake file add
te[R]m(file) [C]reate term [A]dd relation [D]elete [Q]uery sa[V]e [?]
[?]display Query/save search[/]add e[X]it

POPULATION

POPULATION

- UF Demography
 - NT Population size
 - NT Population distribution
 - NT Population density
- S NT Population dynamics**
- NT Population projection
 - NT Migration
 - NT Birth place
 - NT Length of residence

[SpB] Next [B]ack [F]irst [P]age [S]elect [T]erm select [M]ake file add
te[R]m(file) [C]reate term [A]dd relation [D]elete [Q]uery sa[V]e [?]
[?]display Query/save search[/]add e[X]it

POPULATION DYNAMICS

POPULATION DYNAMICS

- NT Fertility
- NT Birth rate
- NT Death rate

[SpB] Next [B]ack [F]irst [P]age [S]elect [T]erm select [M]ake file add
te[R]m(file) [C]reate term [A]dd relation [D]elete [Q]uery sa[V]e [?]display
Query/save search[/]add e[X]it

REFERENCES

1. ACKOFF (E). *On Purposeful Systems*. Chicago: Aldine-Atherton. 1972.
2. AITCHISON (J). *et al. Thesaurofacet; A Thesaurus and Faceted Classification for Engineering and Related Subjects*. Whetstone, Leicester: English Electric Co.; 1979.
3. ANDERSON (J R); BOWER (G N). *Human Associative Memory*. New York: Halstead Press; 1973.
4. AUSTIN (D). "The Development of PRECIS: A Theoretical and Technical History." *Journal of Documentation*. Vol. 30; 1974; p47-102.
5. BARKAKATI (N). *The Wailer Group's Turbo C++ Bible*. Indiana: SAMS: 1990. Chap. 2.
6. BHATTACHARYYA (G). "Fundamentals of Subject Indexing". In *Ordering Systems for Global Information Networks: Proceedings of the Third International Study Conference on Classification Research*, Bombay, 1975. Edited by A. Neelameghan. Bangalore, Sarada Ranganathan Endowment for Library Science; 1978, p83-93.
7. — . "POPSI: Its Fundamentals and Procedure Based on a General Theory of Subject Indexing Languages." *Library Science with a Slant to Documentation*. Vol. 16; 1979; p1-34.

8. BISWAS (S C). "Two Decades of POPSI, 1969-1988: Literature Review." *Libri*. Vol. 38; 1988; p279-96.
9. —; SMITH (F). "Computerized Deep Structure Indexing System - A Critical Appraisal." *International Classification*. Vol. 15; 1988; p2-12.
10. —; —. "Term and Cross-Reference Structure in Computerized Deep Structure Indexing Systems." *International Classification*. Vol. 15; 1988; p139-44.
11. BRACKETT (Michael H). *Developing Data Structured Databases*. Englewood Cliffs: Prentice-Hall; 1987.
12. BRUNER (Jerome). *Study of Thinking*. New York: Wiley; 1956.
13. CHOWDHURY (Govinda G); NEELAMEGHAN (A); SUDATTA CHOWDHURY. "VOCON: Vocabulary Control Online in Microslis Databases". *Knowledge Organization*. Vol. 22(1); 1995; p18-22.
14. CODD (E F). "A Relational Model of Data for Large Shared Data Banks. *Communications of the Association of Computing Machinery*. Vol. 13 (6); 1970; p377-87.
15. DEVADASON (F J). "Postulate-based Permuted Subject Indexing Language as a Metalanguage for Computer-Aided Generation of Information Retrieval Thesaurus." *International Forum on Information and Documentation*. Vol. 8; 1983; p22-29.
16. —. "Computerization of Deep Structure Based Indexes". *International Classification*. Vol. 12; 1985; p87-94.
17. —. *Computerized Deep Structure Indexing System*. Frankfurt: INDEKS verlag; 1986. (FID/CR Report No. 21).
18. —; BALASUBRAMANIAN (V). "Computer Generation of Thesaurus from Structured Subject- Propositions." *Information Processing and Management*. Vol. 17; 1981; p1-11.
19. FOSKETT (D J). "Ranganathan and 'User-friendliness'" *Libri*. Vol. 42; 1992; p227-34.

General Theory of Knowledge Classification

20. FREEMAN (R R); ATHERTON (P). *AUDACIOUS - An Experiment with an Online Interactive Reference Retrieval System Using the Universal Decimal Classification as the Indexing Language*. 1967. Report AIP/UDC-7 under NSF Grant Gn- 433.
21. —; —; “File Organization and Search Strategy Using UDC in Mechanized Reference Retrieval Systems.” In (1968) *Mechanized Information ... Edited by Kjell (Samuelson)*. *op. cit.* p. 122-152.
22. FUGMANN (Robert). *The Analytico-synthetic Foundation for Large Indexing and Information Retrieval Systems*. Bangalore: Sarada Ranganathan Endowment for Library Science; 1983.
23. GILLENSON (Mark L). *Database: Step-by-Step*. New York: John Wiley & Sons; 1985.
24. GODERT (Winfred). “Facet Classification in Online Retrieval”. *International Classification*. Vol. 18; 1991; p98-103.
25. GOPINATH (M A); SEETHARAMA (S). “Interdisciplinary Subjects and Their Classification.” In *Ordering Systems ... op.cit* p121-36.
26. INGWERSEN (Peter); WORMELL (Irene). “Ranganathan in the Perspective of Advanced Information Retrieval.” *Libri*. Vol. 42; 1992; p184- 201.
27. International Symposium on Relational Classification, College Park, Md., 1967. *Proceedings*. Vol. 3 (1967).
28. IYER (Hemalata). “Facet Structure of Subjects: An Empirical Study of Concreteness and Predictability of Categories.” *Library Science with a Slant to Documentation*. Vol. 19; 1982; Paper L.
29. —. “Online Searching: Use of Classificatory Structures”. In *Tools for Knowledge Representation and the Human Interface. Proceedings of the 1st International ISKO Conference, Darmstadt, 14-18 August 1990*. Edited by R. Fugmann. p159-67. Frankfurt: INDEKS Verlag; 1991.
30. -. “Subject Representation and Entropy.” *International Classification*. Vol. 19; 1992; p15-18.

31. KENT (Allen) (ed.). *Information Retrieval and Machine Translation Based on the International Conference for Standards for Machine Searching and Translation*, Sponsored by Western Reserve University and Rand Development Corporation, held in Cleveland, Ohio, September 6-12, 1959. New York: Interscience Publishers Inc.; 1960. 2 parts.
32. KURBAKOV (K J). "[Problems of Unification of Classification Systems: One Possible Approach]." *Nauchno-Tekhnicheskaya Informatsiya. Seriya 2. Informatsionnye Protsessy I Systemy*. Vol. 7; 1968; p24-6.
33. LEONOV (R A). "[Chain Method of Building Alphabetical Indexes]." *Nauchno-Tekhnicheskaya Informatsiya. Seriya 2. Informatsionnye Protsessy I Systemy*. Vol. 1; 1970; p15-20.
34. NEELAMEGHAN (A). "Absolute Syntax and Structure of an Indexing and Switching Language". In *Ordering Systems ... op. cit.*, p165-76.
35. — . "Application of Ranganathan's General Theory of Knowledge Classification in Designing Specialized Databases". *Libri*. Vol. 42; 1992; p202-26.
36. — . "Application of Ranganathan's Postulates and Principles of the General Theory of Knowledge Classification to Database Design and Information Retrieval." *International Cataloguing and Bibliographical Control*. Vol. 22(3); 1993; p46-50.
37. — . "Concept Categorization and Knowledge Organization in Specialized Databases: A Case Study". *International Classification*. Vol. 18; 1991; p92-7.
38. — . "Home-grown Object-oriented Specialized Databases." In *Ching-Chih Chen (ed.) NIT'92: 5th International Conference New Information Technology...*, Hong Kong, Nov. 30-Dec. 2. *Proceedings*. 1992, p251-84.
- 38a. — . A Knowledge Base and Hypertext Facility to Support Spiritual, Moral and Ethical Studies. paper contributed to the *6th International Study Conference on Classification Research*, Londown, 16-19 June 1997.

General Theory of Knowledge Classification

39. — . *Presentation of Ideas in Technical Writing*. Delhi: Vikas; 1975.
40. — . "Ranganathan's Generalized Model of Subject Structure and Modes of Formation of Subjects." In *Cognitive Paradigms in Knowledge Organisation. Second International ISKO Conference, Madras, 26-28 August 1992*. Edited by A. Neelameghan et al. Bangalore: Sarada Ranganathan Endowment for Library Science; 1992, p259-85.
41. — . "Sequence of Component Ideas in a Subject." *Library Science with a Slant to Documentation*, 8 (1971) Paper Q.
42. — ; GOPINATH (M A); DENTON PATRICIA (H). "Motor Vehicle Production: Depth Classification: A Demonstration." *Library Science with a Slant to Documentation*. Vol. 4; 1967; p97-178.
43. — ; RAVICHANDRA RAO (I K). "Nonhierarchical Associative Relationship: Their Types and Computer Generation of RT Links." *Library Science with a Slant to Documentation*. Vol. 3; 1976; p24-42.
44. — ; VENKATARAMAN (S). "Use of Computer for the Synthesis of Class Number: A Case Study with a Freely Faceted Version of Colon Classification." *Library Science with a Slant to Documentation*. Vol. 5; 1968; Paper S.
45. RAJAGOPALAN (T S). (ed.) *Ranganathan's Philosophy: Assessment, Impact and Relevance*. New Delhi: Vikas; 1984.
46. RANGANATHAN (S R). *Colon Classification*. Madras; 1933. Ed. 7 revised and edited by M.A. Gopinath. Bangalore; Sarada Ranganathan Endowment for Library science; 1987.
47. — . "Design of Depth Classification: Methodology." *Library Science with a Slant to Documentation*. Vol. 1; 1964; Paper A.
48. — . *Five Laws of Library Science*. Bombay: Asia Publishing House; 1963.
49. — . "Hidden Roots of Classification. " *Information Storage and Retrieval*. Vol. 3; 1967; section 7.

50. — . *Philosophy of Classification*. Copenhagen: Munksgaard 1951.
51. — . *Prolegomena to Library Classification*. Ed 3. Assisted by M.A. Gopinath. Bombay: Asia Publishing House; 1967.
52. — . — . Part H.
53. — . "Subject Heading and Facet Analysis". *Journal of Documentation*. Vol. 20; 1964; p109-19.
54. — . *Theory of Library Catalogue*. Madras; 1937.
55. ROSCH (Eleanor). "Principles of Categorization." In *Cognition and Categorization*. Hillside, N.J.: L.Erlbaum Assoc; 1978; p28-9.
56. SAMUELSON (Kjell) (ed.). *Mechanized Information Storage, Retrieval and Dissemination. Proceedings of the FID/IFIP Joint Conference, Rome, June 14-17, 1967*. Organized by FID and IFIP. Amsterdam: North-Holland Publishing Co.; 1968.
57. SATIJA (M P). "Colon Classification: Literature Analysis, 1970-1984." *International Classification*. Vol. 12; 1985; p133-42.
58. SAXENA (D). "Object-Oriented Programming." *Computers Today*. 1991 July; p61-3.
59. SCHNEIDER (K). "UDC in Mechanized Indexing and Information Retrieval." In *Mechanized Information... op cit*. p153-59.
60. SRILATHA (G); NEELAMEGHAN (A). "A Microisis Pascal Interface for Concurrent Multiple Databases Search and Retrieval. *Information Studies*. Vol. 1; 1995; p114-29.
61. STILES (W G). "Ranganathan, Cognition and Expert Systems." *Canadian Journal of Information Science*. Vol. 10; 1985; p16-24.
62. SUKHMANEVA (E G). "[Questions of Indexing and Faceted Classification.]" *Nauchno-Tekhnicheskaya Informatsiya. Seriya 2. Informatsionnye Protsessy I Systemy*. Vol. 9; 1967; p10-12.

General Theory of Knowledge Classification

63. — . “[Some Questions of Contemporary Development of Faceted Classification].” *Nauchno-Tekhnicheskaya Informatsiya. Seriya 2. Informatsionnye Protsessy i Systemy*. Vol. 11; 1967; p5-8.
64. SVENONIUS (Elaine). “Ranganathan and Classification Science”. *Libri*. Vol. 42; 1992; p176-83.
65. TANABE (Kinzo). “Functional Operators in Engineering Language.” In *Information Retrieval... op. cit.* Part 1, p295-336.
67. TEMTIM ASSEFA. *Information Support System for Urban Planning With Special Reference to the National Urban Planning Institute*. Thesis. M.Sc (I.S.), Addis Ababa University. Addis Ababa: Addis Ababa University; 1992. [unpublished]
68. TESHOME BULA. *Computer-generation of Vocabulary Control Tools and Their Application to Data-bases in urban planning*. Thesis. M.Sc (I.S.), Addis Ababa University. Addis Ababa: Addis Ababa University; 1993. [unpublished]
69. VICKERY (B C). “Knowledge Representation: A Brief Overview.” *Journal of Documentation*. Vol. 42; 1986; p145- 159.
71. VENKATARAMAN (S); NEELAMEGHAN A. “Formulation of Isolate Number by Computer Using the Devices of Colon Classification. *Library Science with a Slant to Documentation*. Vol.6; 1969; Paper K.
72. VICKERY (A); BROOKS (H M). “PLEXUS - the Expert System for Referral.” *Information Processing and Management*. Vol. 23; 1987; p99-117.
72. WATTERS (C R); SHEPHERD (M A). “A Logical Basis for Information Retrieval.” *Information Processing and Management*. Vol. 23; 1987; p433-45.
73. Webster's *Third New International Dictionary of the English Language*. 1961.

Annexure 1

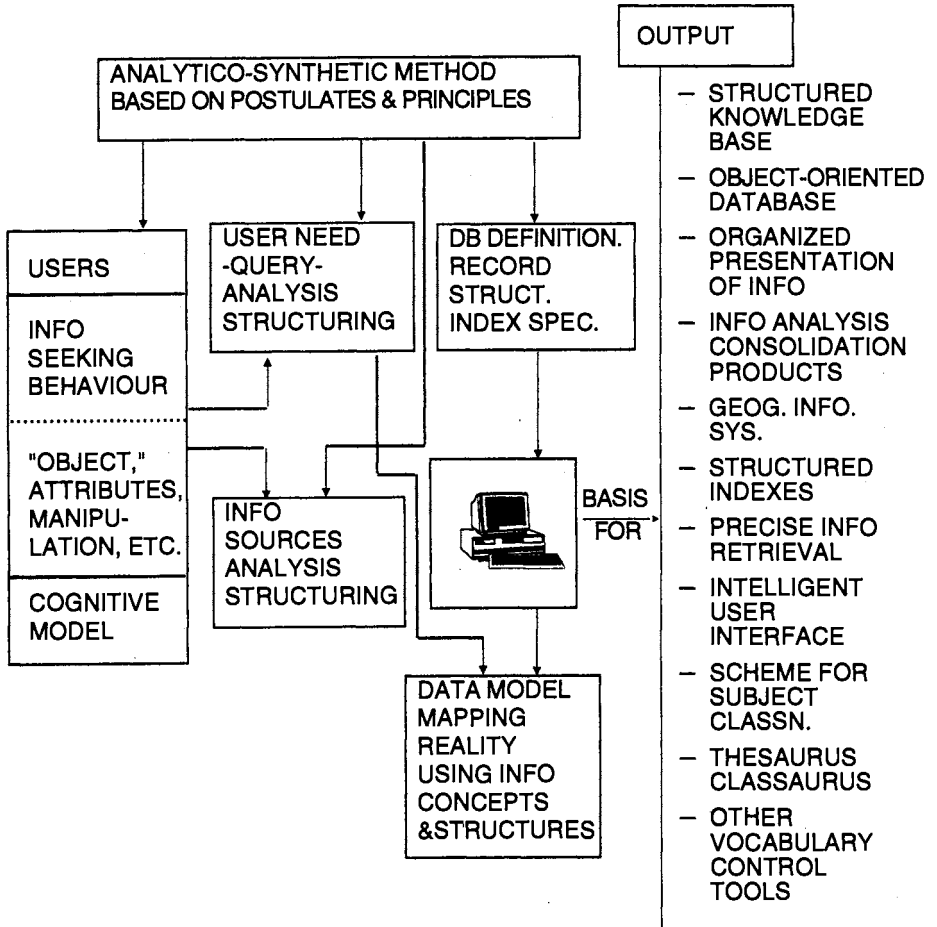


Fig. 2 Analytico-Synthetic Methods in Database Design, Retrieval, and outputs.

INDEX

Note: *cf.*: compare with; *def.*: defined; *irt.*: in relation to; *qirt.*: quoted in relation to; *riirt.*: referred in relation to

- Absence of Basic Facet**, 16
- Absolute syntax**
- irt* Facet sequence, 38-42
 - Indexing language, 43-71
 - Linguistic syntax, 38-42
 - Linguistic universals, 57-59
 - Roots of classification, 38-42
 - Subject structure model, 170
 - Switching language, 43-71
 - Usefulness of, 56-57
- Actand-Action-Actor-Tool principle**
- irt* Helpful sequence, 30
- Amsden, Constance**
- qirt* Syntax of thought, 62
- Analytico-synthetic classification**, 34
- irt* Scheme of classes, 82
- Analytico-synthetic technique**
- irt* Classification, 135-137
 - Database design, 172
 - Mechanized info systems, 157
- Analysed title**
- irt* Analytico-synthetic technique, 136-137
- Architect, Plight of**, 3
- Arrangement of subjects, Linear**, 30-32
- Array of Order 1**
- Allocation of sectors in , 109-110
- Array, Telescoping in**, 113
- Artificial language**
- irt* Artificial syntax, 144-145
- Artificial syntax**
- irt* Headings of subject entries, 144-145
- Basic Class, Enumerative**
- definition of, 87
- Basic Classes, Number of**, 90-91
- Basic Facet Absence of**, 16
- irt* Compound subject, 15
 - Postulate of, 15
 - Recognition of, 16
- Basic Subject** *def.*, 127
- Biocybernetic view**
- irt* Modes of thinking, 58-60
- Birnbaum, H.**
- qirt* Universal linguistic forms, 57-58
- Book index entry** *def.*, 130
- Canon of prepotence**
- irt* Heading, Rendering of, 138-139
- Canon of Sought Heading**
- irt* Heading, Rendering of, 139
- Canonical Class** *def.*, 88-90
- Canons of classification**
- irt* Database design, 177-179
- CCC**
- qirt* Chain procedure, 141
 - Heading, Rendering of , 134
- Chain procedure**, 141
- irt* Indexing, 161-162
 - Versatility of, 143-144
- Characteristic**
- see also* Quasi Isolate
- Characteristics**
- Combination of, 110-111
- Classification**
- Analytico-synthetic, 34
 - def* Arranging , 77-78
 - Grouping , 77
 - Design of, 2
 - cf* Database design, 174
 - First-Link-Downwards approach, 98-99

Index

- Classification. . .
 Last-Link-Upwards approach, 98
 irt Facet analysis , 131-132
 Knowledge organization in
 databases , 173-174
 Purposes of users , 4
Library, 83-84
Micro-subjects, 3
models
 irt Universe of knowledge, 80-81
 irt Data model, 162
Practical, 33-37
Research in, 120-121
Roots of, 1
see also Analytico-synthetic classification
 Enumerative classification
 Faceted classification
Terminology of, 84-87
Classificationist
 Plight of, 2
 Problems of, 4
Classified Catalogue Code
 see CCC
Classifying *def*, 130
Cleveland Conference (1959)
 qirt Facet technique , 158
Colon Classification *see* CC
Comfort to eye and memory, 114
Compound subject
 irt Basic Facet, 15
Concept categorization
 Analytico-synthetic technique in, 173
Concreteness Postulate of, 24
Conjecture,
 Method of *irt* CC, 5-6
 Risk of, 6-7
Continuum
 irt Universe of knowledge, 80
Council for Library Resources
 qirt Funding, 146
Cow-Calf principle
 irt Helpful sequence, 29-30
Cutter, Charles Amni
 Specific subject heading, 146
D
Data entity *def*, 156
Data fields arrangement,
 Principles of helpful for, 182-184
Data mapping, 196-197
Data mining, 196-197
Data model
 def, 156-157
 irt Classification model, 163
Database design
 Application of
 Principles and postulates, 204
 Classification design, 174
 Examples of, 204
Database elements
 Determination of, 180-181
Databases
 Knowledge organization in, 173
 Organization of ideas in, 172
Databases search
 Hypertext links in, 195-96
Depth classification
 Design of, 75-126
Dewey Decimal Classification
 see DDC
Discipline *def*, 155
Dorking Conference
 qirt Postulational approach, 82
Dynamic nature of
 Universe of knowledge, 79-80
E
Empty digit, 101
Empty-Emptying digit, 101
Emptying digit, 101
End-user involvement
 irt Object database design, 199
Energy
 irt Fundamental Categories, 12
 Level in, 24
 Rounds of, 21-22

Index

- Energy isolate
 - irt* CC Matter isolate , 13
- Entity *def*, 154
- Enumerative classification
 - irt* Scheme of classes , 82
- Examples of
 - database design , 204
- Facet** *def*, 129
- Facet analysis
 - def*, 129
 - irt* Artificial syntax, 144-145
 - Classification, 131-132
 - Postulates and principles, 134
 - Search strategy, 191
 - Subject heading, 127-146
 - Procedure for, 135-137
- Facet, First
 - Postulate of, 24
- Facet sequence
 - Absolute syntax in, 38-42
 - Principles for
 - irt* Artificial syntax, 144-145
 - Statistical study of, 41-42
 - within a Round
 - Postulate of, 24-25
 - within Last Round
 - Postulate of, 25
- Facet structure
 - Generalized, 64-71
- Facet syntax
 - Hidden roots of, 37-42
- Facet technique
 - irt* Mechanized info systems , 157
- Faceted classification
 - irt* Scheme for classes , 82-83
- Facets
 - Syntax of , 37-42
- FID/FIP Joint Conference (1967)
 - qirt* Facet technique , 158
- Fields in database
 - Organization of , 182
- Filmore, C.J.
 - qirt* Case categories , 58
- First term omission
 - irt* Heading, Rendering of, 140-141
- First-Link-Downwards Approach
 - irt* Design of classification, 98-99
- Five Laws of Library Science
 - qirt* Users, 149
- Flame analogy, Roots of, 9
- Focus *def*, 129
- Forward rendering
 - irt* Heading, 140
- Framework
 - irt* Transformation of subject, 53-54
- Full title
 - irt* Analytico-synthetic procedure, 135
- Fundamental Categories
 - at Seminal level, 32
 - Energy, 12
 - Manifestation of
 - Examples, 17-19
 - Matter, 13
 - Personality, 14
 - PMEST, 10
 - Space, 11
 - Time, 11
- Gap Device**
 - irt* CC, 119-120
 - DDC, 118
 - LC , 118
- Generalized facet structure, 64-71
 - irt* Absolute syntax, 66
- Heading**
 - Choice of, 132
 - def*, 131
 - First
 - Choice of, 138-141
 - irt* Specific subject, 132
 - Multiple, 133-134
 - Rendering of, 133

Index

- Helpful sequence
 - Determination by
 - Actand-Action-Actor-Tool principle, 30
 - Cow-Calf principle, 29-30
 - Wall-Picture principle, 25-28
 - Whole-Organ principle, 28-29
 - Intuition about, 9
 - irt* Information presentation, 179-180
 - Users, 36
- Heuristics
 - irt* Database browsing, 173
- Hidden roots
 - irt* Classification, 1, 37-42
- Holistic approach, 148-149
- Hypertext link
 - irt* Multiple databases search, 195-196
- Idea *def*, 154-155
- Idea plane
- Ideas
 - Intangible nature of, 3
 - see also* Universe of ideas
 - Universe of
 - see* Universe of ideas
- Immediate neighbourhood relations
 - irt* Transformation of
 - n-dimensional space, 31-32
 - Universe of subjects
 - Transformation of, 31-32
- Index
 - Computer generation of, 162
- Indexing
 - Chain procedure in, 161-162
 - Facet technique in, 161
- Indexing language
 - Absolute syntax of, 43-71
- Infinite nature
 - irt* Universe of knowledge, 79
- Infinity
 - irt* Purposes of users, 4
 - Universe of ideas
 - Dimensions of, 4
- Infinity . . .
 - Universe of knowledge, 91-92
 - Two-fold, 4
- Information
 - service
 - Humanization of, 187
 - systems
 - Compatibility among, 50
 - technology
 - Impact of, 150-154
- Inhibition in *irt*
 - Base of notation, 94-95
 - Colon connecting symbol, 93
 - Conecting symbols, 93-94
 - Quasi-isolate, 96-97
 - Sub-universe, 95
 - Super-imposition, 95
 - Whole universe, 96
 - irt* Design of Classification
 - Approaches to, 97-100
 - Notational plane, 93
- Intermediate term omission
 - irt* Heading, Rendering of, 141
- Intuition
 - irt* Helpful sequence, 9
- Invariant
 - irt* Linear arrangement, 78
 - Neighbourhood relations, 31-32
- Isolate
 - def*, 128-129
 - Enumerative definition of, 87-88
 - facet, Postulate of, 16
 - ideas, Recognition of, 20-21
- K**ernel
 - def*, 129
 - term *def*, 129
- Kernel title
 - irt* Analytico-synthetic procedure, 138
- Knowledge
 - def*, 155
 - organization
 - irt* Databases, 173

Index

- Knowledge. . .
 see also Universe of knowledge
 Syntax of, 60-61
- L**anger, S.K.
 qirt Logical form of thought, 64
- Last-Link-Downwards approach
 irt Classification design, 98
- Last term omission
 irt Heading, Rendering of, 141
- Lazlo, Ervin
 qirt Modes of thinking, 59-60
- Leibniz
 qirt Ideal language, 59
- Level, Postulate of , 23-24
 irt Matter, 23
 Personality, 23, 111
 Space, 23
 Time, 23
- Level cluster, Postulate of, 25
- Library classification, 83-84
- Library of Congress Classification
 see LC
- Linear arrangement
 irt Mathematical transformation, 31
 Subjects, 30-32
- Linguistic syntax
 irt Absolute syntax , 38-42
 Syntax of thought, 63
- Linguistic universals
 irt Absolute syntax, 57-59
- M**acro thought, 8
- Main class
 def, 88-90
- Main entry *def*, 130
- Mapping
 irt n-Dimensional space, 31
 One-dimensional space, 31
 Subjects, 51-53
 Universe of subjects, 31
- Mathematical transformation
 irt Linear arrangement , 31
- Matter
 as Quasi Isolate, 115-117
 irt Fundamental Categories, 13
 isolate *irt* Energy isolate, 13
 Material, 13
 Postulate of Level, 23
 Property, 13
 Rounds of, 22
- Meredith, P.
 qirt Syntax of knowledge, 61-62
- Method of residues
 irt Personality recognition , 14
- Micro subjects
 Classification of, 3, 75-126
- Micro thought, 8
- Mixed notation, 100
- Mnemonics
 irt Polymorphism, 185-187
- Multi-dimensions
 irt Universe of knowledge, 79
- Multiple heading, 133-134
- N**-dimensional space
 Mapping to
 One-dimensional space, 31
- Near-seminal level
 irt Universe of ideas , 9
- Neelameghan, A
 Absolute Syntax... , 43
 S.R.Ranganathan's General Theory.. , 147
- Neighbourhood relations
 Invariance in, 31-32
 irt Mapping of
 Universe of subjects , 31-32
- Normative principles
 irt Holistic approach, 148-149
- Notation
 Economy in, 114
 Mixed, 100
 Sector formation in, 101
 Zqne formation, 101

Index

- Notational Plane, 100-115
irt Idea plane, 93
- O**bject database
Features of, 198-199
- Object databases design
End-user involvement in, 199
Examples of, 204
irt Classification, 166-168
- P**attern
irt Perception of entity, 45-46
- Perception
Pattern in, 45-46
Structure in, 45-46
- Personality
Facet
Levels of, 111
Schedules for, 112-114
irt Fundamental Categories, 14
Manifestation of, 14
Postulate of Level, 23
Recognition
Method of residues in, 14
Rounds of, 22
- Phenomenal level
irt Universe of ideas, 9
- PMEST, 10
- Polymorphism
irt Mnemonics, 185-187
- POPSI, 162
- Postulate of
First Facet, 24
Fundamental Categories, 10
Level *irt*
Space, 23
Time, 23
Rounds of Energy, 21-22
Rounds of Personality, 22
Rounds of Space, 22
Rounds of Time, 22
- Postulate(s)
Advantages of, 34-37
- Postulate(s) . . .
Alternative, 25
irt Classification, 10
Concreteness, 24
Facet analysis, 134
Facet sequence
within a Round, 24-25
within Last Round, 25
Holistic approach, 148-149
Isolate facet, 16
Level, 23-24
Level cluster, 25
Level of Energy, 24
Rounds of Matter , 22
- Postulational approach
irt Invariant complex,
Determination of, 81
- Practical classification, 33-37
- Principles and postulates
Advantages of, 34-37
irt Database design
Examples of, 204
- Principles of helpful sequence
Actand-Action-Actor-Tool principle, 30
Cow-Calf principle, 29-30
irt Data fields arrangement, 182-184
Wall-Picture principle, 25-28
Whole-Organ principle, 28-29
- Punctuation mark
irt Subject heading, 145 -
- Q**uasi Isolates
Over flow of, 110
Sequence of, 112-113, 117
- R**anganathan, S.R.
Design of Depth Classification, 75-126
Hidden Roots of Classification, 1
*Subject Heading and
Facet Analysis*, 127-146
- Relational database
irt Faceted classification, 163-165

Index

- Relational Factors in Classification,
Symposium on (1966) *irt*
Syntax of facets, 1
- Research in classification, 120-121
- Reverse rendering
irt Heading, 140
- Roots
irt Classification, 1, 37-42
of flame analogy, 9
- Rosenbleuth, Arturo
qirt Linguistic syntax, 63
Syntax of thought, 63
- Rounds of Energy
Postulate of, 21-22
- S**chedules for
Personality Facet, 112-114
- Schedules of Special Isolates
irt Personality Facet
Design of, 100
- Scheme for classes, 82-83
- Scheme of classes, 82
def, 130
- Scheme of classification
def, 130
- Scientific Information,
International Conference on (1958)
irt Syntax of facets, 55
- Search strategy
Facet analysis in, 191
for databases, 190
- Sector allocation
irt Array of Order 1, 109-110
irt Array of Order 2 and above, 111
- Sectors *irt* Notation, 101-114
- See also subject entry
def, 131
irt Chain procedure, 142
- Seminal level
Descent to, 32
irt Fundamental Categories, 32
- Space
irt Fundamental Categories
isolates in CC and UDC
Rounds of, 22
- Specific subject
Choice of, Procedure for, 135
def, 131
heading, Rendering of, 137-138
irt Choice of heading, 132
- Statistical study
irt Heading, Rendering of, 140
Risk in conjecture, 7-8
Sequence of facets, 41-42
Specific subject heading, 146
- Step-by-step procedure
irt Classifying, 136-137
- Structure
Common
irt Message transformation, 63
irt Perception of entity, 45-46
Subject representation, 46-47
- Subject
def, 127, 155
fields, Demarcation of, 168
heading
irt Facet analysis, 127-146
irt Facet structure, 64-71
representation
irt Information systems, 47
User, 48-49
Problems in, 48-49
Resolution of, 49-50
Structure in, 46-47
Surrogates for, 47-48
Transformation in, 51-53
Framework for, 53-54
structure, Generalized
model of, 169
- Subjects
Arrangement of, 30-32

Index

- Surrogate
 - irt* Representation of subject, 47-48
- Switching language
 - Absolute syntax of, 43-71
- Syntax,
 - Facet
 - Hidden roots of, 37-42
 - irt* Linguistic syntax, 38-42
 - Linguistic
 - irt* Syntax of thought, 63
- Telescoping in Array, 113
- Terminology of
 - Classification, 84-87
- Thinking, Modes of
 - Biocybernetic view of, 58-60
- Thought *see* Ideas
- Time
 - irt* Fundamental Categories, 11
 - isolates in CC and UDC, 12
 - Rounds of, 22
- Training of
 - documentalist in
 - design of depth classification, 121-122
 - generalist librarian in
 - design of depth classification, 121
- Transformation *irt* Subjects, 51-53
- Transformed title
 - irt* Analytico-synthetic technique, 137
- U**niverse of ideas
 - Dimensions of, 8
 - Multi-dimensional, 9
 - Turbulent nature of, 3
- Universe of knowledge
 - Qualities of, 79-80
- Universe of subjects
 - Transformation of, 31-32
- User assistance
 - irt* Database search, 191
- User satisfaction
 - irt* Syntax of facets, 38
- User-system interface
 - irt* Searching, 187
- Users, Majority of
 - Helpful sequence for, 36
- Users, Minority of
 - Helpful sequence for, 36-37
- V**alue-added information, 151-154
- W**all-Picture Principle
 - irt* Helpful sequence, 25-28
- Whole-Organ principle
 - irt* Helpful sequence, 28-29
- Z**ones *irt* Notation, 101-114



Neelameghan, A., Compiled by.
S.R. Ranganathan's Postulates and Normative Principles: Applications in Specialized Databases
Design, Indexing and Retrieval.
(Sarada Ranganathan Endowment for Library Science, Series 7).
Bangalore, Sarada Ranganathan Endowment for Library Science, 1997.

Copyright ©Sarada Ranganathan Endowment for Library Science (SRELS)

Digitization: Susan Ditch, SIRLS, University of Arizona
Quality Control: Megan Plesea, SIRLS, University of Arizona
Digitization Training: Han Yan, Information Systems, University of Arizona & dLIST Editor
Project Coordinator: Cheryl K. Malone, Assoc. Prof. SIRLS, University of Arizona & dLIST Editor
Digitized: Spring 2007

Acknowledgments: SRELS (A. Neelameghan, K.N. Prasad, K.S. Raghavan, DRTC) and dLIST
Advisory Board Member, S. Arunachalam (MS Swaminathan Research Foundation)
dLIST Classics (Book) Editor: Barbara Hutchinson, University of Arizona
dLIST Editor-in-chief: Anita Coleman, University of Arizona

dLIST Classics

Titles by S.R.Ranganathan

Five Laws of Library Science, Ed. 1 (1931)
Philosophy of Library Classification (1973)
Prolegomena to Library Classification, Ed. 3 (1967)
Classification and Communication (1951)
Documentation Genesis and Development (1973)
Documentation and its Facets (1963)
Library Book Selection, Ed. 2 (1966)
New education and school library: Experience of half a century (1973)
Reference Service, Ed. 2 (1961)

Other titles

S.R. Ranganathan's Postulates and Normative Principles: Applications in Specialized Databases
Design, Indexing and Retrieval, 1997. Compiled by A. Neelameghan.
Memorabilia Ranganathan: A compilation of useful quotations of S.R. Ranganathan from his various
works, 1994.
Putting Knowledge to Work: An American View of the Five Laws of Library Science, 1970. By
Pauline Atherton.

Read the dLIST Classics online!

<http://dlist.sir.arizona.edu/>

