L. A. RESEARCH PAMPHLET No. 2

# Classification for a General Index Language

# CLASSIFICATION FOR A GENERAL INDEX LANGUAGE

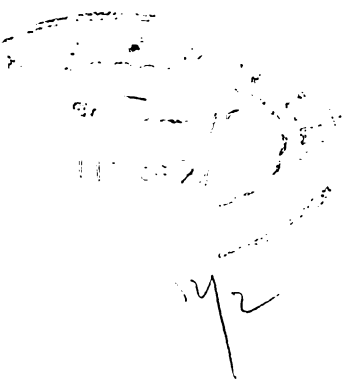
A review of recent research by the
Classification Research Group

by

D. J. FOSKETT, M.A., F.L.A.
University of London
Institute of Education

*Library Association Research Publication No. 2*

THE LIBRARY ASSOCIATION
1970



Published by The Library Association
7 Ridgmount Street, London, W.C.1.

SBN: 85365 032 2

*Printed by Wm. Pollard & Co. Ltd., Exeter*

## CONTENTS

# INTRODUCTION

The Classification Research Group has now been meeting regularly for more than 15 years, and the discussions and contributions of its members have been recorded periodically in the CRG Bulletin appearing in the *Journal of Documentation*. In 1963, the Library Association Research Committee negotiated a grant of £5,000 with the office of the Science Adviser to NATO, for research into the structure of a new general classification scheme for scientific documentation. Two research assistants, first Mrs. Helen Tomlinson and then Mr. D. W. Austin, produced a series of reports: their intention was to explore the theoretical foundations of general, as opposed to special, classification schemes, and to work out their applications in detail in a few selected disciplines. These reports are published by the Library Association in full, under the title *Classification and information control*.

The CRG thought that a short and simplified introduction to these reports might also be useful for students, and others with a general interest in the progress of subject analysis for information handling purposes. This pamphlet does not pretend to do full justice to the research, but I hope that it will serve the modest purpose of introducing it to a wider audience.

D.J.F.

## GENERAL AND SPECIAL CLASSIFICATIONS

The role of classification schemes in libraries and information services has probably caused more argument than any other professional activity. This would be surprising if classification were no more than a fairly convenient way of arranging books on shelves. Some librarians think it is, and support their view by heaping scorn on the heads of those who, like the Classification Research Group, actually spend years of their time in theoretical discussions that seem to result in more and more abstruse and difficult complications to what ought to be a straightforward exercise. Yet today we can see clearly two relatively new spectacles on the library scene: on the one hand, some librarians are criticising the Decimal Classification (especially as used in the *British National Bibliography*) for being too detailed and unwieldy; on the other hand, some librarians, and still more information officers, are busy revising the Universal Decimal Classification in order to make it more detailed. In the next field, as it were, computerised indexing and retrieval systems are pounding away at the ever-growing masses of literature, producing results that impress computer specialists but not information users, who are so deafened by the noise that they cannot hear what is new.

These are the problems that the CRG has been trying to solve, and they are intensely practical problems. If, as Professor John Ziman has recently suggested, science is public knowledge, then the organisation of the literature has to play a part exactly comparable to the organisation of knowledge in the human mind. If it is true, as I believe it is, that the workings of computers can teach us much about the workings of the brain, it is equally true that the study of the brain can teach us how to make the best use of computers. Classification plays a fundamental part in enabling the human mind to master its environment, and it is the contention of the CRG that it has just such a fundamental part to play in enabling a computer to organise its material.

In science, classification has been regarded as an essential tool since antiquity. For Plato, to know things actually meant to place them in their correct class; his theory of Forms was itself a scheme for classification. Aristotle came to reject the theory of Forms (with reluctance, he said, since those who propounded it were friends of his), but established a much more detailed classificatory method, that of division from genus to species, which has lasted to the present day. But this means neither that we are chained to the Aristotelian system alone, nor that we should cast it away completely. When a scientific "paradigm", or generally accepted pattern of ideas, becomes

unable to provide an explanation for newly-discovered phenomena, it is overthrown; but the new paradigm still incorporates all the unchallenged features of the old, where they apply. And, as T. S. Kuhn pointed out in advancing this notion of successive paradigms, the new one is accepted, not on its past performance, but on its future promise. Part of the acceptance derives therefore from intuitive and aesthetic considerations -- the feeling that the new pattern begins to make sense out of what increasingly seems a jumble of incompatible but incontrovertible observations.

The CRG, which was founded as an informal discussion group after the Royal Society Scientific Information Conference of 1948, began with only their dissatisfaction with all existing schemes of classification linking its members. All had had considerable experience with general and special schemes, in general and special libraries. The membership, then and now, has been made up predominantly of special librarians and information officers who were faced with the problem of imposing order on collections of documents dealing with fairly specific areas of knowledge. They were not so much interested in arrangement of books on shelves as in providing quick and efficient access to subject information through indexes. Information Retrieval was their goal, in the eminently satisfactory definition of Calvin Mooers: "searching and retrieval of information from storage according to specification by subject".

Specification by subject may exist, however, in more than one mode: specification by an author, of what he takes to be the reality he is investigating; specification by an indexer, of what he in turn takes to be the intention of the author, which may be given a particular 'slant' for the purposes of the indexer's own organisation; specification by a user, who has become conscious of a gap in the knowledge he requires in order to deal with a current situation. The problem in an information library is to effect a match between the first and last of these three; and this match is effected by means of the second. The simplest case of this would be if an enquirer wanted to read about a single species – say, the Eagle – and knew the name of the bird, which would be the same as that used to identify and classify books about it. Here, all three specifications are identical, and there is no vagueness or ambiguity.

It is this kind of request that genus-species classification meets perfectly. It classifies by forming abstractions, ideas of groups or classes, and proceeds by separating out the sub-groups or individual members of the class, by enumerating the properties which differentiate one from another. Thus the class Birds produces the species Eagle, Sparrow, Starling, and so on. In their long series of tests on young children, Jean Piaget and his colleagues in Geneva have shown that this is the way in which the child combines sensorimotor activities (such as peeling, tasting, chewing) with perceptual data (such as colour and shape) to arrive at identification, and therefore knowledge,

of entities: "this is an orange". This can be demonstrated in two ways: the "descending" or analytical method of separating out one individual from a collection; and the "ascending" or synthetic method of forming a collection from a number of different individuals.

Readers searching for information in a library will also be relying on these two types of mental operation, but most library classification schemes rely only on analysis. The power of the analytical method in bringing order into science was perhaps best seen in the eighteenth century, when the work of the great Swedish botanist Carl Linnaeus laid the foundation for all subsequent biological nomenclature and taxonomy; even though his system has been superseded, it made an invaluable contribution to the great advances in the biological sciences during the next two centuries. Quite recently, Professor Donald MacRae has said that what Sociology now needs is not a Darwin or a Newton, but a Linnaeus to make a good classification of social institutions and their relationships.

The great achievement of Melvil Dewey in his classification for the Library of Amherst College in 1873 was to see how to apply "Scientific" classification to the arrangement of books on shelves. Instead of "fixed location", which was the assigning of a book to the same, constant, place (which could never mean more than a very broad arrangement by subject), he introduced "relative location" based on genus-species classification of subjects, giving a hierarchical sequence in which some classes can be seen to be subordinate to others. He made it practicable by abandoning the numbering of shelf-places, and numbering subjects with a decimal fraction notation. Just as, in principle, one can always make a further sub-division to a group of entities by separating out a lower-order species by one more characteristic feature, so one can always add another subject-number to any decimal number by adding one more digit at the right hand end.

| *Subject* | *Notation* |
|---|---|
| Mammals | 599 |
| Primates | 599.8 |
| Apes | 599.88 |
| Gorillas | 599.884 |

This meant that the shelf order of books became an up-to-date display of the structure – the internal relations – of subjects. In an open access library, shelf classification is indeed a very powerful tool for information retrieval; not only can one pin-point a subject in its hierarchical position, but by the placing together of related co-ordinate classes, one can browse; in this way the arrangement itself may well suggest new lines of thought.

But already Dewey had recognised that there is more than one way of

dividing a group of entities into smaller groups. He did not distinguish between "essential" characteristics such as the possession of a backbone, and "accidental" characteristics such as colour of hair or eyes; the first are characteristics essential to the definition (a "man" must have a backbone), the second are not (no "man" must have black hair, though he may). He did distinguish between sub-groups formed on the basis of characteristics possessed by the members, and sub-groups formed by a different method:

| | |
|---|---|
| Birds | Birds |
| Sparrows | Birds of California |
| Starlings | Birds of Massachusetts |

For the second group, he did not enumerate the sub-groups, with their separate notational symbols, but proposed to synthesise the sub-groups by adding a division from the geographical schedule, which brought with it its own notational symbol:

| | |
|---|---|
| 598.2 | Birds |
| 598.29 | Geographical treatment |
| 598.29793 | Birds of California |

This combination of analysis and synthesis was greatly developed in the Universal Decimal Classification, which was originally intended for the arrangement of a vast world bibliography of books and periodical articles, on cards. Paul Otlet and Henri La Fontaine, in 1895, found that the DC was neither detailed nor flexible enough for their specialised purposes, and over the years, through the work of the International Federation for Documentation (F.I.D.), several Auxiliary Signs and Tables have been added to make the UDC nowadays very much more able to deal even with highly specific topics, by synthesis. But many traces of the old basic structure remain. Consider the following sequence, from the latest revision of UDC Class 37, Education:

| | |
|---|---|
| 37 | Education |
| 37.018 | Fundamental forms of education |
| 37.018.2 | School education |
| 37.018.26 | Attitudes of parents to school |
| 37.018.263 | Parent-teacher relations |

It is obvious enough that this is not a generic hierarchy; neither of the last two terms is in any sense a species of its predecessor, yet they are presented in the schedules as if they were. This is doubly wrong: first, because we have a mixture of characteristics in one array of divisions; second, because compound subjects are presented as if they were simple sub-divisions of the

preceding subject. "Attitudes of parents to school" contains three elements, each differing radically in its nature from the others.

Only in the Colon Classification (CC) of S. R. Ranganathan was this difference recognised and represented. Ranganathan, in his many books and in particular in *Prolegomena to library classification*, made explicit the manifold nature of the features which may characterise any entity. Inhelder and Piaget, in *The early growth of logic in the child*, have shown that very early in life, the child learns to use both ascending and descending methods to form and divide groups on the basis of different characteristics; today it may be all oranges, all apples, all bananas, and tomorrow it may be all round fruit and all long, while the next day it may be all those that need peeling, and all those that don't. They call these "multiplicative" classifications, and have proved in extensive experiments that these cross-classificatory matrices develop in the child's mind at the same time as the notion of simple hierarchies. This is to be expected, because "they express one and the same general operational mode of organisation". Ranganathan used the term "facet analysis" for this process of separating out all the different strands of sub-divisions that could result from the analysis of a single class. The application of one characteristic of division produced a set or category of terms which he called a "facet" of the class. Each facet can be related to a fundamental abstract notion about the nature of the whole universe of knowledge.

The idea of fundamental categories is not new by any means. In addition to his analytic method of genera and species, Aristotle wrote a treatise on "The Categories" as part of his work on Logic, in which he defined a set of ten abstract notions in terms of which any single entity might be discussed: Substance, Quantity, Quality, Place, Time, and so on. Right down to the present, philosophers have gone on discussing the categories, differing among themselves to such an extent that while Professor P. F. Strawson believes that there are only two (bodies and persons), Professor Gilbert Ryle says that there neither is, nor can be, any limit to their number.

For library classification, Ranganathan nominates five. These are: Time, the chronological aspect of a subject; Space, its geographical aspect; Energy, the processes with which it is concerned; Matter, the substance of its entities; Personality, the thing-in-itself, the entity whose nature as a complete whole determines the Main Class of the subject. Ranganathan has not found it easy to define his category of Personality; it is itself, "ineffable", he says, and one can hear echoes of Plato's Theory of Forms. One can, however, see what he means. A Bicycle is an entity with a unique and recognisable identity as itself; it may be made of Steel, or Aluminium; it may be made by Forging, or Welding; it may be made in England, or India; it may be made in 1968, or 1969.

Now it is clear that a complex subject today may very well need more

than five facets for a complete analysis. We might have an article about "Presses for forging handlebars for bicycles". Here there are no less than three recognisable entities: presses, handlebars, bicycles. None of these will go into a Time, or Space, or Energy, or Matter, facet; and what remains after these are excluded, says Ranganathan, must be Personality. To cope with this type of situation, he has introduced the notions of Levels of a facet, and Rounds of facets. A handlebar is a sort of Personality, but not a whole in relation to Bicycles; it is a Part, a sub-level of the main Personality, the Bicycle. A press is another sort of Personality, a whole (separate and different from the Bicycle), but it does not appear until after the Energy facet, the process for which it is the Tool. This is therefore a Personality term which begins a new sequence, or Round, of facets.

In his latest work, Ranganathan has been considering in great detail the intellectual foundation of his theory of classification. To the refinements such as Rounds and Levels, he has added a number of other devices, in the Idea Plane, the Verbal Plane, and the Notational Plane. After the International Study Conference on Classification for Information Retrieval, held by Aslib and the CRG in Dorking in 1957, he concluded that no scheme could be maintained to keep up with the vast complexities of modern knowledge – "the unending formation of myriads of new micro-thought" – unless it were based on a set of Postulates for subject analysis, and of Principles for determining sequence among facets and isolate terms. This kind of scheme he now calls Analytico-Synthetic. "A faceted scheme is an analytico-synthetic scheme if and only if the design of its schedules and their augmentation are guided by a set of postulates and principles". No extensions or modifications can therefore be made to the scheme unless they can be shown to arise from, and be in conformity with, the postulates and principles, and Ranganathan has already made changes in the CC itself where he found that earlier decisions did not conform.

While the CRG has rejected allegiance to any existing scheme of classification, it has nevertheless recognised the enormous power and versatility of the method of facet analysis, and has always used it since the first CRG publication in the *Library Association Record* in 1955, "The need for a faceted classification as the basis of all methods of information retrieval". Members have contributed to many international conferences, and a second "Dorking" Conference on "Classification Research" was held under the auspices of the FID Committee on Classification Research, at Elsinore in 1964. For several years, the CRG avoided the issue of a general scheme covering the whole of knowledge. Members produced more than twenty schemes for special subjects, in connection with their own work or as consultants to other organisations. In the construction of these schemes, the "subject" was given; that is, the starting point for the analysis was determined by the special purpose of the exercise. We avoided asking the

question, where this subject might appear in the whole universe of knowledge, and indeed there was for some time a vague general agreement in the CRG that it might well be impossible to make a new and satisfactory general scheme of classification.

The procedure for making a special scheme is now fairly well understood. One has a subject, or discipline, or field of knowledge, which has several characteristic features: a collection of "core" or basic journals, professional groups which meet to discuss common problems, at least an approximation to an agreed terminology and methods of investigation. As a subject grows and reaches a recognised coherence, its literature begins to be organised by means of indexes, abstracts, reviews of progress. To make a special classification scheme, one can take a sample of the literature, obtain from it the terms in use, and sort them into categories. In the field of Education, for example, one finds subjects like these in the *British Education Index:*

A calculating device for teaching elementary arithmetic
Universities, politics and public opinion in Ceylon
Anxiety in the primary school child
Geographical field work in schools
Time-tabling in a junior comprehensive school
Local Education Authorities and teachers in England
Pakistanis in Bradford
Teaching cerebral palsied children in a partially hearing unit

From these, and others like them, one can easily deduce a set of facets:

Exceptional children (Pakistanis, cerebral palsied, partially hearing)
(Levels of) Educands (primary, junior)
Types of school etc. (universities, comprehensive)
Curriculum subjects (arithmetic, geography)
Teaching method and aids (calculating device, field work)
Educational psychology (anxiety)
School organisation (time-tabling)
Teaching profession (teachers)
General organisation etc. (Local Education Authorities, politics, public opinion)
Places, Countries (Ceylon, Bradford, England)

Once a facet has been identified by the appearance of only a few terms, one obviously does not need to go on searching the literature for actual instances of others. The faceted structure provides a framework into which one can insert terms from personal knowledge, glossaries, indexes to books, or any other appropriate source.

This is not the whole of the story, of course. Major questions such as the sequence of facets, sequence of terms in facets, and notation, have to be

settled, and these, like the naming of the facets themselves, have always been decided by CRG members on a purely pragmatic basis, from the material in hand. It would be perfectly possible to relate the facets named above to Ranganathan's fundamental categories, but it would not be quite straightforward, and, in a restricted application, would not be particularly valuable either. Since all the CRG schemes were for restricted application, the idea of fundamental categories fell rather into the background, though they were undoubtedly useful as a check, and as part of the agreed specialised language used in discussion.

Some of the above facets, however, bring to light the major problem of the relations between various subject fields which Ranganathan calls "phase relations". B. C. Vickery had presented a paper on this subject to the Dorking Conference, showing how things can exist in more than one context, and how classification schemes had set out to cope with this. In some cases, such as "Alcohol", the same name may be applied to different things; in others, such as "Africa", "Aluminium" and "The Aged", the generic relations of the terms remain the same no matter what the context. In still others, such as "Animals", this may no longer hold: they may be pests, or pets, disease carriers or subjects for artistic representation. In Education, "Psychology", "Organisation" are facets that are not peculiar to it; certain specific terms, such as "Male", "Female", "Head", "Deputy", are necessary in many other fields.

In a special scheme, therefore, some terms are "core" terms, which are either peculiar to the subject, or so closely associated with it as to be usually thought of in this context; other terms are necessary for indexing the literature, but would be "core" terms in some other subject.

But this does not dispose of all the terms. Consider the "Curriculum" facet. We find terms such as "The basic subjects", "The three R's", "The Arts Sixth" which may be said to exist only in an educational context. But any subject could be taught in an educational establishment, which means that a "curriculum" facet ought to be comprehensive, to include a term denoting every subject over the whole field of knowledge. If these terms were listed in a systematic sequence, the facet would amount to a general scheme of classification, in outline at least. Exactly the same phenomenon occurs in other special schemes. All have to draw, to a greater or lesser extent, on a reservoir of terms that are not specific to their own field.

The result is that every special scheme has to cover more than its own share of knowledge; and if one were to try to make a general scheme by simply adding all these special schemes together, there would be enormous visible duplication and waste. This experience led the CRG to the conclusion that a general scheme is not only a possibility; it is a necessity for modern work in special fields of documentation. A new scheme might very well look different from traditional schemes (except CC), and it would probably have,

as its *primary* objective, the aim of serving as a starting point for specialist information control, rather than the arrangement of books on shelves.

## INFORMATION CONTROL

As knowledge advances, nations become more prosperous, and more people become involved in research and investigation of the world around us. This inevitably results in a steady increase, at an exponential rate, of the information available. As we know more, it becomes harder to find out more, and the focus of investigation shifts from "macro-thought" – the familiar objects, the large-scale, the all-embracing sweep – to "micro-thought" – the hidden, the small-scale, the concentration on detail. Some years ago, K. D. Puranik made a statistical analysis of the subjects of articles he found in periodicals between 1900 and 1950. He found that there had been a clear change from the simple, one facet or one class type of subject, to the complex, many faceted subject overlapping several classes. The subjects of research and publication, particularly in important documents such as pamphlets, reports and periodical articles, have become highly specific.

At the same time, because more people have become involved in research, communicating information has tended to become more difficult. Important work is no longer concentrated in a few centres where everybody can get to know everybody else. Yet, as we all know, our main method of controlling our own information is by personal contact. If we want to know something, our instinctive reaction is to ask someone who we think will know the answer. Even librarians and information officers do this, though we are usually surprised when research shows it to be the case. Person-to-person contact has always been the main mode of information control, and probably it always will be, even though, as Marshall McLuhan has over-emphasised, it need not necessarily always be face-to-face to be by word of mouth. The mass media of the electronic era enable us to overcome at least some of the obstacles created by time and distance.

At present, however, it is undeniable that some control has been lost as more people, spread over wider distances, have become involved. The invention of printing with movable type, making large editions possible, was a huge step towards overcoming some obstacles, and so was the introduction of the scientific periodical in the seventeenth century. Both helped to make systematic the communication of knowledge, and to widen the range of people a scholar could expect to reach. But, as McLuhan points out, the printed page imposes its own obstacles and restrictions on the flow of thought. As I write, I am acutely aware of how slow my pen is compared to

the flow of my thoughts or even of my speech. I have to make a conscious check on what I am thinking in order to set it down in some sort of coherent sentence. The reader likewise has to discipline and concentrate his mind in order to understand what I am saying, and this is much more difficult in the absence of another human personality. This is why the "lecture" still survives, and is even enjoyed when it is given by a good lecturer, even though it can only reach a relatively few people (unless it is on radio or television!).

More people write more, more publications appear, and the proportion of the total that one can manage tends to decrease. There is a serious danger nowadays that "information" has become a commodity, and the idea is spreading that the more we produce, the better we are. But let us not forget the dangers of over-production. We do not want to build up great systems of document control if many of them contain nothing but rubbish. On the other hand, we must remember the research worker, particularly the young worker. If the people in his field are scattered all over the world, how is he to make contact with them? How can he make known his own interest, his line of work? He is just as concerned that he should know them, as that they should know him. He wants to build up his own lines of communication.

He does this by means of documentation. He publishes on his own account; he reads the current journals to get to know the others. As journals have multiplied, techniques like Selective Dissemination of Information with computers are invented. The computer becomes an intermediate screen between the current journals and the individual user, channelling all that is relevant to his work towards him, regardless of where it may have been published. The SDI system matches his specification of his interests with the specification it will apply to articles currently being published. This is to take care of his current awareness.

When he comes up against a problem, he also wants to know what has already been done to solve it. He makes a retrospective search. He looks up what has already been published, by consulting subject indexes in libraries, books, periodicals, encyclopedias, and so on. He is now trying to match his specification with those already made by other authors. The results, in retrospective searching and in current awareness, depend on the accuracy of the match between the two specifications.

Subject indexing is, of course, a very old and well-established library technique. In fact, classification itself is a form of subject indexing put into practice through notational symbols. Subject indexing with words, "subject headings", either in books or in card indexes in libraries, has not always been carried out with the actual needs and practices of users in mind. We are still haunted by the ghost of Charles Ammi Cutter, who made a tremendous contribution to cataloguing and subject indexing in the nineteenth century, but who claimed that when a reader expressed an interest in "the movement

of fluid in plants", his *subject* was "Physiological Botany". What he should have said was that the main class in which he would expect to find this subject was Physiological Botany; the subject itself is "the movement of fluid in plants". But Cutter was actually trying to match the specification of the reader to what he thought was the most likely specification to have been used by an indexer when faced with a document on this subject. Cutter was not wildly wrong, either; given a book on Physiological Botany, you would certainly find in it information on the movement of fluid in plants. But the subject would still not require the whole book; the part is not co-extensive with the whole.

Once a classifier or an information officer has tried to translate a reader's terms into some other terms, we immediately run up against the danger of a mis-match, however slight. As subjects have grown more complex, this danger has increased, especially in library classification and thus in retrospective searching. If it is true, as Mortimer Taube once said, that "standard library practice assumes that, in general, it is possible to express any specific topic in a single word or phrase", then the chances of compiling an efficient subject index are rather small. This belief has led to what might at first appear to be the obvious answer: in subject indexing, use the same terms as the author of the document to be indexed. And there is no doubt that this technique will lead to much better indexes than the former. Some research workers, such as Cyril Cleverdon in Aeronautics, and Colin Tapper in Law, have shown that good results can be achieved by indexing in the natural language of authors, that is, using authors' terms.

This movement away from classification as a tool in indexing has roughly coincided with the development of mechanised indexing techniques. The first of these were based on punched cards, which also enabled a new kind of indexing to be used. In conventional library catalogues, each entry represented a single document, or item in the collection. With punched cards, however, one could also make use of an alternative form of indexing, and allot a card to each term, and punch on the card the numbers of all documents that dealt with that term. The idea of term cards, or "aspect" cards, was presented by W. E. Batten to the Aslib Conference of 1947, using classification as a basis for selecting his terms. Taube introduced this technique in the U.S.A., calling it Co-ordinate Indexing, but without at first using punched cards, simply writing the numbers on the cards. Since searching for multi-term subjects involved a visual scanning of the cards to find which numbers appeared on all the cards, it proved rather cumbersome and liable to error. With punched cards of the "optical coincidence" kind, holding the cards together in front of a light source quickly revealed the common numbers.

Similarly, with machine-sorted cards, subject classification was first discarded in favour of random numbering; but as this meant that all the

cards had to be searched for every enquiry, some form of grouping into smaller batches was needed. The idea of a controlled vocabulary grouped according to "families of notions", as an aid to both indexing and search strategy was discussed in 1951 by H. P. Luhn, and ever since then most documentation systems based on machines have relied to a greater or lesser extent on a "thesaurus".

The original *Thesaurus of English words and phrases*, of Peter Mark Roget, is a compilation in which words are grouped together according to certain abstract notions such as Existence, Quantity, Time, Change, Causation, Space, Matter, and so on. There are five classes of Categories, and more than 20 categories, some with many sub-divisions. In other words, it is a classification scheme based on Fundamental Categories; the similarity, in general and in detail, to the Colon Classification needs no emphasis. Attached to Roget's *Thesaurus* is a detailed and specific alphabetical index, and many words are shown to appear in several different categories according to their different meanings or contexts of use (including metaphorical):

> 183. *Situation* – *N.* situation, position, locality, locale, status, latitude and longitude; standpoint, post; aspect, attitude, posture, pose.
>
> place, site, base, station, seat, venue, whereabouts, environment, neighbourhood; bearings &c. 278; spot &c. 182.
>
> top– ge-ography; map &c. 554.
>
> *V.* be -situated, -situate; lie; have its seat in.
>
> *Adj.* situ-ate, -ated; local, topical, topographical &c. *n.*
>
> *Adv.* in situ; here and there; here-, there-, where-abouts; in place, here, there.

After the categories comes an alphabetical list of all these words, referring to the categories in which each word appears:

> position
> *circumstances* 8
> *term* 71
> *situation* 183
> *proposition* 514
> *assertion* 535
>
> situation
> *circumstances* 8
> *place* 183
> *location* 184
> *business* 625
> *out of a* – 185

When one turns from the index to the actual categories, one finds the word one is looking up grouped together with many others of related meaning. As every writer knows, this grouping is a powerful stimulus to thought, because it displays a structured pattern of relationships which corresponds to the way we think. An alphabetical index does not, though it is usually (not always) required as a point through which to gain access to the categories.

Few modern thesauruses resemble Roget, however. Some are merely alphabetical lists of subject headings with references to synonyms and related terms under each. Some are alphabetical lists, with cross-references, but with sets of terms in categories added as a secondary, instead of the primary, listing. An example of the first kind was ERIC, the first Educational Resources Information Centre thesaurus produced by the U.S. Office of Education. An example of the second is MeSH, the Medical Subject Headings List used by the U.S. National Library of Medicine in MEDLARS, The Medical Literature Analysis and Retrieval System, which comprises both a computerised store of bibliographical data on tapes, which are distributed to co-operating libraries in other countries, and also the *Index Medicus*, a printed index which appears twice a month. A very detailed and valuable report on the operation of MEDLARS was published in 1968: *Evaluation of the MEDLARS demand service*, by F. W. Lancaster. This was a review of the success or failure in 300 actual searches carried out in 1966 and 1967. Lancaster notes that "the categories of MeSH do not attempt to be authoritative classification, but empirically-derived from users' requests". He then goes on to assert the need for an "entry vocabulary"; "an adequate entry vocabulary is essential to ensure that indexers and searchers consistently use the same terms, or term combinations, to describe identical items of subject matter .... The quality of an entry vocabulary can substantially affect the recall performance of an information retrieval system". He finds fault with MeSH for its lack of specificity, its inadequate use of hierarchical relationships, and its inconsistency in the use of terms and entering of new terms. Experience is similar at the University of Newcastle, which runs an English MEDLARS unit in collaboration with the National Lending Library of Science and Technology. Even when the search topic is quite precisely named, hundreds of references are retrieved, and the average approaches three hundred.

This sort of result is usually justified on the grounds that a user wants maximum recall, that is, a complete bibliography of his subject. Both Lancaster and the Newcastle group doubt this; Lancaster questions very strongly whether users will wish to scan a large number of references in order to pick up a few that are useful. On the other hand, one needs to be able to alter the terms of a search if the first attempt draws blank. This is where the hierarchies come in. Broadening, or narrowing, the terms of a search (say, from Eagle up to Bird, or down to Golden Eagle), depends on having

well-organised hierarchical chains already in position in the computer – entered in the index language. If they are not, they need to be remade for every enquiry, with all the attendant risk of inconsistency.

This ability to "explode" a search must extend to all facets of the subject, but it is absolutely vital to keep the facets free, so that the searcher can explode on facets one at a time. Clearly, if one draws a blank on a combination of two specific terms, one may very well reach a good result by broadening only one or the other, while broadening both together will result in a very general search bringing out many unnecessary references.

Another research project on the use of a computerised indexing service is the Chemical Society Research Unit in Information Dissemination and Retrieval under A. K. Kent at the University of Nottingham. In his *Search Manual*, Kent writes that search questions should be analysed into "main conceptual elements", such as:

a group of compounds
a process
a type of organism
a type of material

This is a sort of facet analysis, of course; and the resemblance is underlined by Kent's technique of "expansion of concepts", which is the construction of hierarchies of co-ordinate and subordinate terms within each facet in order that the search may be refined so as to achieve optimum results for the user.

Both MEDLARS and the Chemical Society project are dealing with large fields of knowledge, but with a more or less specific vocabulary of technical terms. It is an avowed aim of the scientist, after all, to have a precise terminology. If in fields like these we find the need for an entry vocabulary or a search strategy that depends on a scientifically structured language, how much more will this be the case in fields whose terminology is not so precise? Indeed, in MEDLARS, where the literature ranges over the boundaries of several branches of the social sciences, it is here that the problem gets worse. This has been noted most emphatically, perhaps, by Richard Coward, head of the *British National Bibliography* team engaged in the MARC Project for producing machine readable cataloguing, in collaboration with the Library of Congress. These teams have gone to great trouble to analyse and identify every single item in a bibliographical description, and it seems likely that all problems connected with recording the contents of an entry on tape have been solved. From the subject analysis point of view, however, it is quite another story. Coward has written:

> "I would suggest that a MARC service can only be fully effective in the context of a standard catalogue code and standard methods of conveying subject information. New editions of Dewey appear with

monotonous regularity and new editions of Library of Congress hardly at all. The British MARC record will actually contain five different sets of subject information. This is a frightful price to have to pay for our inability to reach some sort of agreement in this field. However, there is one thing that can be said with certainty about Dewey or the Library of Congress classifications. They are totally unsuitable for machine systems. This might turn out to be a good thing. The field is open for a general classification designed for use in computer systems. If one were developed now there is a good chance that it could become established."

Thus even where computers are used, and where a whole file can be searched in a matter of moments, there still emerges a clear necessity for some kind of structured vocabulary of concepts to provide guide lines both for the indexers recording new documents and for searchers using the store. If computer or any other indexing techniques depart radically from the usual pattern of human thought processes, which involve perception, concept formation, and expression in language, then we reduce our chances of providing a service that will offer help to the user, instead of only responding mechanically to his often ill-expressed needs.

## ORDERING OF ENTITIES

In order to control large quantities of highly specific information, we need to have a structured indexing language capable of expressing the terms found in the literature and relationships between them. Ranganathan was the first to apply scientific method to the arrangement of micro-thought, and his *Prolegomena* ends with the construction of a Generalised Facet Formula which could be interpreted in terms of any field of knowledge, and which would give, by means of a preferred sequence of facets, a helpful order to the literature of that field. The GFF is a summation of the facet formula based on Fundamental Categories, with Rounds and Levels added; and may be found in full on p.461 of *Prolegomena* (2nd edition).

A full exposition, with comments on his later techniques called Zone and Sector Analysis, can be found in Ranganathan's seminar on CC given at Rutgers University in 1965. Applications in great detail to various Basic Classes are published from time to time in the journal *Library Science with a slant to Documentation.*

But assuming that we wished to make use of the GFF, how would we set about it? CC itself can hardly be claimed to do justice to the full range of

Ranganathan's theories; on the other hand, if we take a specific subject, as the contributors to *Library Science* do, we are back in the same position as the CRG found itself, after its work on special schemes. Ranganathan admits the necessity for defining a Basic Class: it is a subject generally recognised as having sufficient separate identity to provide a satisfactory starting point for facet analysis. A Main Class, on the other hand, is not defined, but postulated. Thus Production Engineering is a Main Class, but Screw Production Engineering and Motor Vehicle Production Engineering are recognisably specific enough to be able to apply the GFF with consistent results – as has been done by A. Neelameghan and Abdul Rahman and their colleagues at the Documentation Research and Training Centre in Bangalore. These Basic Classes seem to be of the same order as the subjects of CRG schemes: not perhaps completely arbitrary, but defined by the visible needs of some organisation or branch of industry. They have no philosophical justification as separate starting points but, says Ranganathan, "we cannot wait and hold up further work on classification until clear-cut definitions emerge. Therefore, it is desirable for a Scheme for Classification to postulate its own provisional schedule of Basic Subjects .... the same thing holds good also with regard to the distinction between a 'Main Subject' and a 'Basic Subject'.

One way, then, of solving the great problem of what are our starting points, is to postulate an arbitrary selection of "separate" fields, or to accept traditional "disciplines". The early editions of CC chose the latter course, but later editions have added more and more Main Classes simply by postulation. The CRG has felt that this procedure, being highly subjective, does not merit public acceptance so long as there remains the possibility of finding a more systematic foundation for the setting up of starting points for analysis. Can we find a theory, or a set of principles, which will guide us in our selection of Main Classes, that is, in setting up the sequence in which terms should be enumerated in a structured IR language? what sort of terms ought to be enumerated?

The essence of the difference between enumerative and analytico-synthetic classification schemes is that the latter lists in its schedules the elementary terms, and not the complexes, that are to be found in any context. Elementary terms may be combined by the classifier to fit the document to be classified, and so an analytico-synthetic scheme does not have to rely only on those headings actually found in the schedules, and can describe the subject of a document as accurately as is necessary. Enumerative schemes, on the other hand, have most of their relationships already built in, or pre-coordinated; classifying them becomes, in most cases, no more than an approximation. This can easily be seen from any issue of the *British National Bibliography*, where the symbol [1] has been added to any book class number where DC is unable to provide a specific match. The proportion of [1] books has

risen steadily over the years, proving that the DC becomes increasingly unable to meet the demands of modern documentation.

Now all schemes for the structuring of knowledge are temporary and may be superseded in due course, but a scheme enumerating only elementary terms is less liable to change. Iron, Gold, Man, Woman, House, City are elementary terms with a unitary meaning which has been clearly recognised since the beginnings of civilization, and which are not likely to change. The complex processes in which they become involved are the ones that change and develop, and it is in the mode of combination to form complexes that we must have the greatest possible flexibility.

If, then, we make a start on classifying the universe about us, as we must do to understand and be able to live in it at all, we find that what we have the most certain and common knowledge of are the objects with which we come into contact in daily life. One of Aristotle's primary categories was Substance, Ranganathan has Matter, philosophers speak of "entities", or like Peter Caws, of "pure phenomenological wholes" such as tables and chairs. This is inevitable. Life is a process of continual detection of identity in things which have a separate existence from one another. Piaget's long series of books gives details of all his experiments to show how children learn to form concepts by investigating objects, interfering with them, arranging and re-arranging, in order to arrive at an understanding of similarities and differences. Percepts derived from these interactions are turned into concepts by a process of classification – the beginning of abstract thought, which is the ability to identify and sort ideas about things into categories. The same applies to memory; the great American psychologist J. S. Bruner has said that "perhaps the most basic thing that can be said about human memory, after a century of intensive research, is that unless detail is placed into a structured pattern, it is rapidly forgotten".

We might, therefore, begin our universal scheme of classification with the familiar objects of daily life. But we should very soon find that we are once again back in the position of making a special, and not general, scheme, for these objects, or entities, all depend on the pre-existence of other entities. The table and chair depend on the wood and the tools, the wood comes from the tree, the tree from the seed. Are we in an impossible situation, what philosophers call an infinite regress? Or can we find some principle of action which will lead us to a genuine starting point, involving no more than an irreducible number of prior assumptions?

The CRG came to the conclusion, some years ago, that the theory of integrative levels showed great promise of providing such a principle. The progress of knowledge has shown up the shakiness of traditional disciplines, or Main Classes, because these tend to be names for collections of ideas that are very much coloured by their historical context. Accepting that the listing of entity terms corresponds largely to Ranganathan's Personality facet, we

ought to be able to find a better way of ordering a series of [P] facets, based on some intrinsic characteristic of the entities rather than on some particular historical grouping or paradigm, which may well be overturned, as Copernicus overturned the Ptolemaic conception of the movement of heavenly bodies.

The notion of "integration" is inseparable from the notion of a "whole". We can readily see, from the world about us, that bits and pieces of this and that can be put together to make one single new thing with quite different and characteristic properties. The Bicycle is more than a collection of bits and pieces of steel, rubber, plastic. All these in a mere heap, could not transport a person; but when they are arranged in a certain way, when the bits and pieces are joined and fixed in a certain set of relations one with another, then they turn into the new thing, which can transport a person provided that the right sort of motive power is applied. And this sort of power could not be applied at all to the heap of parts. So an entity exists by virtue of the internal relations of its parts, and it can also enter into external relations with other entities – with you or me, when we ride our bicycles. The parts have become integrated into a new whole.

The levels theory proposes that entities exist in a recognisable sequence which exhibits an increasing complexity in the way in which parts are organised to become wholes. It is easy to find examples over the whole field of knowledge. In Euclidean geometry, for example, a point has position but no magnitude, a line has one dimension, length but no breadth, a plane figure has two dimensions, a solid figure has three dimensions. The following sequence equally illustrates the notion, from another field:

E

This letter E is an entity, one of 26 in our roman alphabet, which has a significant existence, by itself, for people such as calligraphers, type-designers, and printers.

Earth

This word is made up of five letters, but the meaning of the one word is vastly different from the sum of the meanings of its five constituent letters, and would be altered if the relation between the letters – their sequence – were altered; anagrams are not synonyms.

Earth has not anything to show more fair.

This sentence has eight words, which together have a distinctive meaning; and this sentence is the first line of the well-known poem by Wordsworth, on Westminster Bridge. Furthermore, the poem itself is a new kind of whole entity: it is a sonnet. This means that the relations between the parts (the individual lines and words) are organised in one of a special set of ways; it is not just a collection of fourteen lines, or even of fourteen lines of iambic

pentameters, but fourteen lines of iambic pentameters showing a particular rhyme-scheme. As the great French linguist Ferdinand de Saussure wrote, "The signs that make up language are not abstractions, but real objects; signs and their relations are what linguistics studies, they are the *concrete entities* of our science".

So far, I have been describing levels of integration as a process of building up, but clearly this will not give us a starting point. If we start on the opposite process, that of cutting up, or dividing, we find the same transformation of quantity into quality; we can "cut up" our familiar objects into smaller and smaller units, and at various points a change of kind occurs and a new unit is isolated. Masses reduce to groups of molecules, molecules to atoms, atoms to fundamental particles. Here, however, we have to stop, for in "nuclear physics" the distinction between entity and energy begins to disappear. The first scheme of levels, a very general statement, might be:

Fundamental particles
Atoms
Molecules
Masses
Cells
Organisms
Humans
Societies

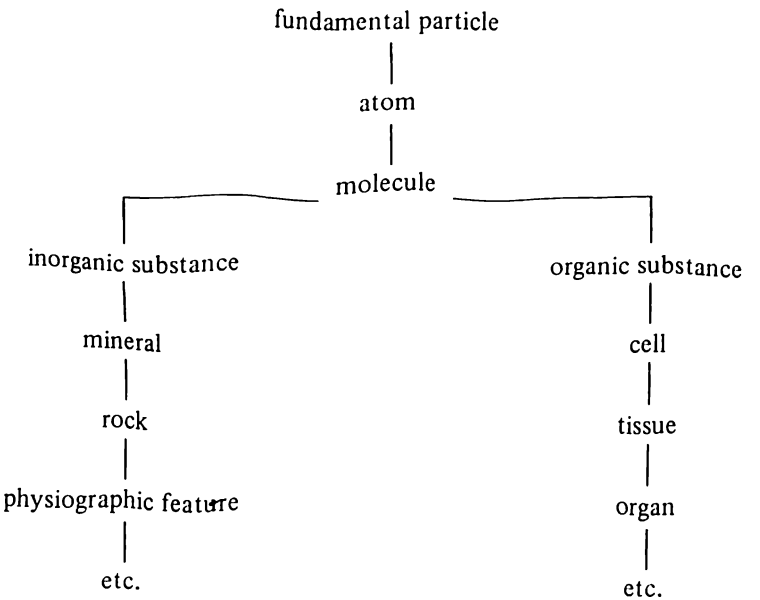
This was the position of CRG discussions on integrative levels at the "London Conference" of 1963, on *Some problems of a general classification scheme.* The papers presented at this conference were intended to give a general account of the problems that the CRG had set itself to solve, and for which it had obtained a grant from the Science Research Fund of NATO. This was also an international conference, attended by representatives from the U.S.A., France, Denmark and Sweden. The Danish member was Rasmus Mölgaard' Hansen, chairman of the F.I.D. Committee on Classification Research, FID/CR. After the papers and discussions, an agreed statement of aims was produced, and a Research Assistant, Mrs. Helen Tomlinson, set to work to carry them out.

Using the fundamental categories of Ranganathan (P, M, E, S, T) as a basis, Helen Tomlinson set out first of all to identify the original [P] facet to which an entity should be assigned, without having to have recourse to traditional main classes. The arrangement of entities should ideally be such that each entity has, as its primary place, the one where its relations with neighbouring entities are constant – in the term of J. E. Farradane, its place of "unique definition", where all characteristics essential to the definition, and no more, are available. Fewer would be inadequate, more would be superfluous. For example, a horse is only sometimes a sporting animal or draft

animal, but is always a perissodactyl mammal. Its taxonomic status is the only definition which is always true, and all its features which are always present will relate to its place in a systematic taxonomy, near to related mammals, farther away from other vertebrates, and farther still from invertebrates.

A single sequence of entities grouped according to their first appearance as "wholes" – their place of unique definition – would not of course be a classification scheme itself. It would not show the horse in the context of sporting or draft animals, for example. But it could be used as a reservoir of terms, already structured by means of characteristics essential to their definition, and would thus go far to solve the special scheme problem of marginal fields. Thus the original idea of constructing a classification scheme for the arrangement of documents or their proxies (such as index cards) began to give way to the idea of a universal structured language for all information control purposes; this would, when equipped with a suitable notation, provide an arrangement for a general library or bibliography (such as *BNB*), but it would also provide a great deal of the non-core material for special schemes tailored to suit a particular organisation.

In practice it soon became clear that a branching structure of levels, rather than a single sequence, was required by the characteristics of the entities. The most obvious example is the differentiation of organic from inorganic substances:



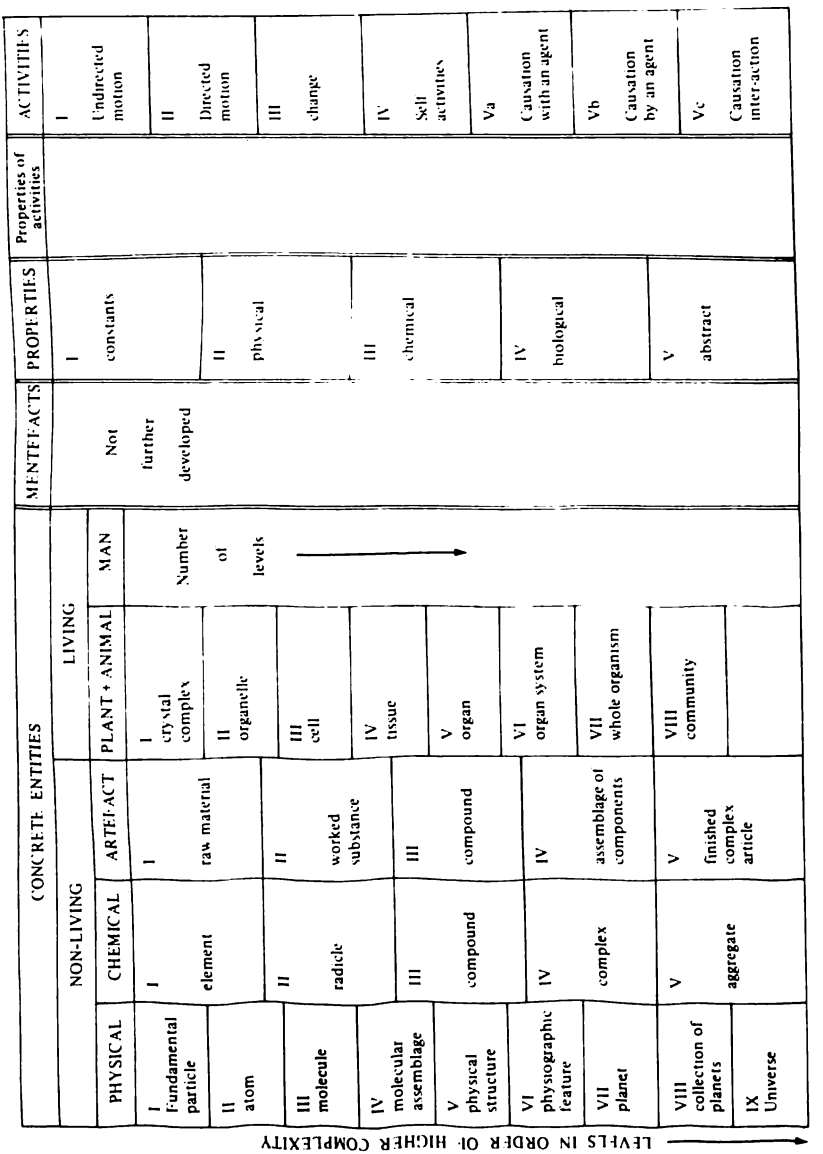
At another point there is a branching between higher forms of organic substance which evolve naturally, and higher forms of organisation in man-made artefacts. The levels above Man similarly divide; along one branch we find more complex forms of organisation made up of individual human beings, such as families, communities, nations and international bodies, and along the other there are intellectual products of individuals, such as language, literature, art, religion.

A crucial question thus emerges: how should we define the nodal points at which such branching occurs? Applying the idea of a system of levels to the lowest forms of matter, Helen Tomlinson found that a fundamental distinction could be drawn between physical and chemical entities. At the particle level, one can enumerate individual names such as proton and electron, or group names such as baryon; but at the atomic level we have terms like atom and isotope, and others like oxygen, copper, carbon. The same thing occurs at higher levels, with terms like element, radicle, liquid, colloid, on the one hand, and iron, zinc, nitrate, chloride, and the various individual compounds and complexes on the other. These examples show that an entity may be discussed sometimes in terms of its physical form, and sometimes in terms of its chemical content. The distinction between form and content (also noted by Aristotle) is a fundamental one, and arises in many other contexts. Of course, it must always be remembered that, like matter and motion, these two cannot be separated in practice; but they can be separated conceptually, and it is essential for clear thinking that they should be. The universe is not a set of static entities, each in its own place, but we have to stand still, as it were, and examine each entity in isolation, to establish its characteristic features, before we can examine its relations with other phenomena. In poetry, the sonnet form has to be understood before we can come to grips with the actual content of a sonnet by Shakespeare, or by Wordsworth, or compare the use of the form by different writers.

Branching points can now be seen to depend not on the inner structure, or form, of the entities at any level, but on the external relations that the entity is capable of maintaining. There is a constant interplay between form and content, but confusion between the two (as in, the medium is the message) cannot lead to clear thinking when we are concerned with the analysis of structure. For Plato, Form was the ideal, the perfection, and actual objects were only imperfect approximations striving towards the ideal; but even Plato had to concede that the ideal chair had a quality of "chairness" without which it could not be distinguished from the ideal table. Form without Content is empty, but Content can only be realised through Form, and to this extent is determined by Form. A poet accepts the restrictions of the sonnet form when he feels that it will embody his ideas and feelings better than the epic; but even when one wants to study the sonnet as a form, it must be by means of some poet's realisation of it.

The working out of these ideas in depth was made by Helen Tomlinson in the fields of Geology, Mining and Sculpture. She drew up a table of general categories based on sequences of levels of entities shown in the Table:

**PROPOSED GENERAL CATEGORIES FOR GENERAL CLASSIFICATION SCHEME**



The Form/Content discussion relates to the conceptual distinction that has to be made between Matter and Motion, Things and Processes. While concentrating in the first instance on the listing of entities, the CRG was also bound to consider the activities in which the entities engage. These arise out of the structure of the entity and what relations its structure enables it to establish with other entities. The structure of the bicycle enables it to carry a person along the ground by means of one rotary action (of the pedals) transferred by a chain to cause another (of the wheels). An attempt to define levels of activities independently of entities did not amount to much, and, just as levels of properties tended to reflect the levels of the entities, so the activities category could be divided into a succession of types of activity showing increasing complexity. A first tabulation of all these Energy terms was divided into three major headings: Properties of entities, Activities of entities, Properties of activities. A more detailed analysis provided several possible sub-divisions.

Properties of entities: constants, physical, chemical, biological, abstract.
Activities: undirected motion, directed motion, change, self-activities, causation.
Properties of activities: rate, velocity.

The term "Mentefact" was introduced by Miss Barbara Kyle to describe intellectual entities, the products of the mind of man, to distinguish letters, words, works of literature and philosophy, and so on, from the products of the hands of man, artefacts. So far, there has been no detailed analysis of this category.

Some work has been done, however, on the other branch of levels arising from the level of Man – the levels of social groups. Barbara Kyle's outline classification of the Social Sciences was based primarily on a sequence of Activities of Man:

| | | |
|---|---|---|
| Biological | – | family, tribe, race |
| Political (internal) | – | local and national government |
| Economic | – | Commerce and industry |
| International | – | war and peace, United Nations |

Barbara Kyle's original idea was to abandon Main Classes altogether, and divide the whole of the Social Sciences into two facets, Activities and Personalities. These had themselves to be ordered by some principle, however, and the sequence above corresponds to levels of groups as well as of activities. The biological level comprises activities of Man as an individual, the political comprises activities concerned with bringing about the organisation and administration of a community. Economic activities are those

entered into by social groups concerned with the production and exchange of the means of life. Politics and Government deal with the internal relations that exist within Nations, but politics of course extends to influence relations with other Nations (who also have their own internal politics), and the use of the term International Relations does not imply something wholly separate. As with the Matter/Motion and Form/Content discussions, it is often necessary, if we are to think clearly, to separate and identify conceptually things which in the real world are inseparable.

Thus the work on integrative level theory applied to single entities led up to a scheme which may be given in outline as follows:

*Physical entities*

| | | |
|---|---|---|
| Level | I | Fundamental particles |
| | II | Atoms, isotopes |
| | III | Molecules |
| | IV | Molecular assemblages, e.g. solid |

*Chemical entities*

| | | |
|---|---|---|
| Level | I | Elements |
| | II | Compounds |
| | III | Complex compounds |

*Heterogenous non-living entities*

| | | |
|---|---|---|
| Level | I | Minerals |
| | II | Rocks |
| | III | Physiographic features |
| | IV | Astronomical entities |

*Artefacts*

| | | |
|---|---|---|
| Level | I | Raw materials |
| | II | Processed raw materials |
| | III | Components |
| | IV | Finished articles |

*Biological entities*

| | | |
|---|---|---|
| Level | I | Viruses |
| | II | Organelles |
| | III | Cells |
| | IV | Tissues |
| | V | Organs |
| | VI | Systems, e.g. digestive systems |
| | VII | Organisms |
| | VIII | Communities, e.g. shoal, herd |

*Man*

| | | |
|---|---|---|
| Level | I | Individual |
| | II | Group |
| | III | Local community |
| | IV | National community |
| | V | International community |

*Mentefacts*

| | | |
|---|---|---|
| Level | I | Units, e.g. digit, note |
| | II | Words, numbers, bars, etc. |
| | III | Sentences, formulae, musical phrases, etc. |
| | IV | Paragraphs, themes, etc. |
| | V | Complete works, philosophical systems, etc. |

## ANALYSIS OF SYSTEMS

In applying integrative level theory to various subjects, Helen Tomlinson uncovered several problems which seemed to be beyond the scope of the theory to solve. It is clear that some "wholes" are on a higher level of organisation than their constituents, but in a different relation to their constituents from others. For example, a molecule of oxygen consists of two atoms of oxygen, a committee consists of several individuals, and in each of these entities, the members are of the same type; they are, so to speak, interchangeable. A bicycle, on the other hand, is not made up of a number of similar sub-units: the pedals and the wheels are obviously not interchangeable. Thus the part/whole relation produces a higher organisation, but in a different way from the building up of similar units. Can it be said that the change of level produced by an accumulation of interchangeable "members" is of the same order as that produced by the assembly of different parts, or as B. C. Vickery further distinguished, of different organs?

Again, it was found that the theory did not offer a solution to another problem that had frequently appeared in special schemes. Many terms in common use do not represent one entity, or a whole class of entities, but only some members of the class. In social science literature for example, we find the terms "Men" and "Women", representing the basic biological division of the human species. We also find terms such as "Women and Children", "Women and Girls", and "Housewives". These have all been called "partially-comprehensive headings", but it will be clear that the last term is of a different type from the others. Plenty of other such classes can be found: "Metals", "Solvents", "Pets", and so on. These classes are groups of

entities which possess certain characteristics in common, but which do not arise out of a grouping process by integration. They are aggregates, rather, based on certain special properties. D. W. Austin preferred to call these "fields", rather than sub-levels. Perhaps this was not the best term, since the word "field" already has other connotations, even in the area of documentation, as in "fixed fields" applied to punched cards, or in such a use as the book by R. F. Piper and P. W. Ward, *Fields and methods of knowledge,* in which the "fields" are the traditional disciplines or what would be called Main Classes in a scheme of classification. But the term is better than "sub-level", in the important particular that it does not imply any process of *integration*, and thus fitted in more unambiguously with the general line of the work.

When Austin took over the research project from Helen Tomlinson, he made a detailed examination of one of the most explicit formulations of integrative level theory, "The integrative levels in Nature" paper by J. K. Feibleman in the collection of papers entitled *Focus* edited by Barbara Kyle. Austin concluded that while the theory could explain the appearance of entities as emergent phenomena in Nature, it fell short when it came to the complexities of modern specialist documentation, with which a classification scheme has to cope. In the search for a more universal theory, the first notion to be found was that of consecutiveness, in which the entities might be related to one another according to their time of emergence. This led to a possible solution for another question, raised by Helen Tomlinson, that of a place for artefacts in the sequence of levels. The principle of emergence-in-time would relate artefacts directly to their makers, as nests to birds, industrial products to Man, and so on. It would mean that Technologies would not appear in the scheme at least until after Human Biology; and also that human artefacts would have the same ultimate point of origin as what Barbara Kyle called Mentefacts.

This conclusion also fitted in with the idea of "purpose" which Helen Tomlinson proposed as a basic ordering principle for human constructions, as Sculpture illustrates the aesthetic purpose. It seems to point to not one, perhaps, but two series of levels. .One contains naturally-occurring entities in a sequence of levels of complexity of organisation; the other sequence is based on Man's needs or drives, and contains terms denoting entities or techniques according to the "purpose" of the grouping. Thus Sculpture might be regarded as Man/Aesthetic needs/Three dimensional realisations.

In this way, the research showed that while integrative level theory provided a considerably firmer scientific basis for the ordering of natural entities than for the ordering of artefacts, it also enabled the problems of the latter to be visualised more clearly, even though it failed to solve them. Further examination of "fields" has shown that many fields are identifiable because their entities behave in a certain way under certain circumstances,

and the properties of entities can be divided into two quite different kinds. When we buy a bicycle, we do so primarily because we want to benefit from its property of carrying us about with comparatively little effort on our part. The organisation of the parts gives it the property of converting the rotary motion of the wheels into horizontal progress along the road. But given a choice from several bicycles all having this property, we make our final choice on other grounds. This one looks nice, but is too big, or too heavy; that one is the right size, but looks ugly; here is one that looks nice and is comfortable and this is the one we select. But who decides what "looks nice", or what exactly is "heavy"? These are not intrinsic and necessary properties of the bicycle; they are qualities which we have attributed to the bicycle from our own special point of view. Someone else might easily say we were wrong, and that the bicycle did not have these qualities. We might ourselves pick up a large book and say "This is a heavy book"; but if we needed a weight for something, we might pick up the same book and say "this is not heavy enough". The weight of the book, its objective property, would not have changed, only our subjective estimate of that property in relation to other external factors. Thus a "property" exists in an entity even when considered in a hypothetical static condition; an attribute arises out of the processes or activities engaged in by the entity in dynamic relation with other entities.

This differentiation between entities as described by objective and subjective properties foreshadows the difficulty of building up a sequence of levels of process or activity terms, on the basis of the levels of entities. For while certain process terms can be related to an entity level, for example, "Breathing" to the biological level, other process terms cannot be so related. Even those which at first appear to be, such as "treatymaking" to the International Relations level, can easily be seen, on closer examination, to be practical manifestations, at that level, of an activity which can be expressed as more general notion. "Treaty making" is clearly the same sort of activity as trade agreements between firms or contracts between individuals. Similarly, attributes such as "beautiful", "dangerous", "heavy" are not among the intrinsic properties of an entity; an entity can objectively be only of a particular shape, weight, volume, etc., and therefore attribute terms are less concrete, or of a more general application. In planning the sequence of all these terms, the less concrete will be better placed earlier than the more concrete, in the same way as less complex entities appear earlier than more complex.

This technique follows the same lines as Ranganathan's treatment of his [P] and [E] facets. In citing facets for one particular document, [P] precedes [E]:

| Agriculture<br>= Main Class | Crop<br>= Fruit | Process<br>= Harvesting |
|---|---|---|
| J | 37 | 7 |
| Medicine<br>= Main Class | Organ<br>= Lung | Process<br>= Tuberculosis |
| M | 45 | 421 |

On the shelves, the order is reversed, so that all books on Harvesting (in general) appear before any books on specific crops, and all books on Diseases (in general) appear before any books on specific organs. While citation order is:

Main Class [P] ; [M] : [E]

shelf order is:

Main Class: [E]
Main Class; [M] : [E]
Main Class [P] ; [M] : [E]

Thus the more generalised the application of the [E] facet, the earlier it appears in shelf order. If we abandon the idea of postulated Main Classes, as the CRG has done (at least *pro tem*), we arrive at the result that [E] facets, that is process, activity, property, attribute terms, should be placed in the sequence in front of the first level of entity terms to which they apply.

It quickly becomes clear, in considering attribute terms, that a category of very high generality indeed needs to be created, in order to place terms like "very" or "quasi" or "maximum". Austin has called these "relative terms" because they may be needed to qualify all other process terms. Other terms needed to identify any entity with the highest possible accuracy, such as location, colour, shape, etc., now fall into the following sequence according to this principle of decreasing generality or increasing specificity:

*General relative terms*
Degrees of amount
Degrees of order or rank
Degrees of kind or substance

*General evaluative terms*
Favourable reaction terms
Unfavourable reaction terms

*Positions*
- Time
- Space
- Person as user or possessor

*Physical measure*
- Mass
- Linear measure

*Shapes*

*Appearances*
- Light
- Colour

*Sounds*

*Tactile sensations*

*Flavours*

*Odours*

*States*
- Mechanical states
- Energy states

*Structures*

We have been unable to discover any other analysis which goes into this degree of detail. The work at the U.S. Patent Office (particularly that of S. M. Newman) used a similar approach, with reference to relations existing in and between artefacts, in order to classify inventions in great detail, but was not concerned with such a degree of generality, and did not make use of the integrative level structure.

Applying the principle of increasing specificity of application to the ordering of activity terms, the basic idea was to apply a similar notion to the one used for entity terms, namely that increasing interaction between parts or members leads to the point where the integrating force of the inter-relations overcomes the opposing force of the integrity of the individual parts or members. A new whole entity is formed when the interrelations become fixed, no longer subject to change from inner forces, but only by the

imposition of superior force from outside. Once our bicycle is all bolted together, it should not fall to pieces of its own volition. In the activity sequence, static conditions are considered to be more general than dynamic, and the interactions are ordered as, first those leading to mixtures of entities, then aggregates, and finally newly emerging wholes. This leads to the following outline order:

*General Activity concepts*

*General static and kinetic conditions*

*Equilibrium*

*General kinetic conditions*

*Contacts and disturbances*

*Motions and transfers*

*Assembly and disassembly*

What has emerged is the notion of recurring patterns of processes which appear at many different entity levels. At each level, the process naturally receives its appropriate specific name (such as "treaty-making"), but it can be seen to be an example of a general phenomenon. Ranganathan has a similar idea, not only in his fundamental categories from which facets are derived, but more particularly in his "seminal mnemonics". These are used for identifying the sequence of terms in facets, and are groups of associated ideas given the same notational symbol, somewhat after the style of Roget's thesaurus, but much wider ranging. These are discussed both in *Prolegomena* and in *Philosophy of Library classification.* The numbers used in CC include :

2 meaning Structure and its analogues
3 meaning Function and its analogues
4 meaning Fault and its analogues
6 meaning Correction and its analogues

The first paper to examine in detail the application of integrative level theory, which appeared in *The Sayers Memorial Volume*, had briefly mentioned General Systems Theory as a possible source of classification of the notion of generalised [E] facets. General Systems Theory uses the word "system" in a highly abstract way, meaning any set of elements whose inter-relationships bind them together so that the set has a continuing and recognisable existence. The notion of "integration", that the whole is more than the sum of the parts, is central to the theory. It was first advanced

in the 1930's by Ludwig von Bertalanffy, who looked for a unifying set of ideas that would account for the appearance of similar fundamental conceptions in all branches of science "irrespective of whether inanimate things, living organisms or social phenomena are the objects of study .... Not only are general aspects and viewpoints alike in different fields of science; we also find formally identifical or isomorphic laws in completely different fields". He has given instances of a dozen different applications of the exponential law, or law of compound interest. He has also emphasised that "systems" means "wholes, or unities". General Systems Theory is "a logico-mathematical field, the subject matter of which is the formulation and deduction of those principles which are valid for 'systems' in general. There are principles which apply to systems in general, whatever the nature of their component elements or the relations or 'forces' between them. The fact that all sciences .... are concerned with systems leads to a formal correspondence or logical homology in their general principles, and even in their special laws if the conditions of the systems correspond in the phenomena under consideration".

Since Bertalanffy went to the USA at the end of the Second World War, most of the further explication of the theory has been carried out there, particularly by Kenneth Boulding, who teaches a course with Bertalanffy at the University of Michigan. The wider interest of the theory has been expounded by Dan Bergen, in his article, "Implications of the General Systems Theory for Librarianship and Higher Education". J. K. Feibleman has also extended his account of integrative levels into the area of the social sciences, calling the entities there "cultural circuits". The fact that the inter-relations between the members of a social group are reciprocal rather than one-directional implies a "circuit" (as in a molecular structure) and this in turn implies a "whole" or entity. In the world around us, the world of macro phenomena, one can indeed often see the inter-relations between parts and members of an entity. A committee in session, for example, displays the role of the chairman, the secretary, the several members who speak to each other through the chairman. Similarly, one can readily see and understand the role of the parts of the bicycle. Bertalanffy suggested that the same type of structure existed in the micro world which we cannot so easily see, and with the aid of instruments one can find out by experiment that this hypothesis can be confirmed.

The theory certainly offers itself as a unifying factor throughout the whole of knowledge, but it is in a positive role; as Bergen points out, in a college curriculum it would act not only as a means to unify knowledge but even more as a mechanism for structuring knowledge. The significance of this for librarians needs no emphasis, because in the most important part of professional work, information and reference service, one constantly has to come to grips with the structure of knowledge – often in strange subjects – in order

to provide the right information that the user requires. As operational experience in computerised systems such as MEDLARS shows, this becomes even more important when the user does not have direct access to the file of documents (or their surrogates), but has to pose his question in the light of the admittedly incomplete information in his own head. In such a situation, a structured index language, even the entry vocabulary to a particular system, can itself be creative, in the sense that it promotes thought, and is not, as are most alphabetical systems, merely able to respond in a question-answering manner, without being able to suggest to the enquirer what he might ask.

In systems analysis, we start with a definition of a system very similar to the notion already expressed of an integrated "whole". Austin cites Macmillan and Gonsalez:

"A system is a set of objects (subsystems) together with relationships between them and their attributes. Since a system never occurs without an environment, a characterisation on this basis is incomplete without reference to the total system, i.e. the system plus its environment. All systems have thus to be identified by means of constraints which delineate their true position and that of their subsystems vis-a-vis the environment. The only exception is the universe, which has no environment and thus constitutes the only true total system". From this, it emerges that a classification must at least start from the point of having systems, sub-systems, and their environments. The analysis of relations has long been familiar through the work of Farradane, and Austin borrowed Farradane's term "operator" to indicate the type of relation subsisting between the various cited entities or processes in a subject analysis.

This development of the research shows the inevitability of linking theory with practice in any scheme which actually purports to solve problems. The CRG has often been accused, particularly by those who imagine that computer systems have no need of a controlled language (presumably they can think for themselves), of making the study of classification an end in itself, of allowing interest in ideas to carry us into the realm of higher thought unrelated to the ordinary everyday world. The fact is that, just as experience with the construction of special schemes led to the search for a scheme for ordering entities, so the theory of a general classification scheme could not proceed beyond a certain point without further reference to the needs of practising classifiers. Austin's position as Subject Cataloguer at the *British National Bibliography* brought him constantly into hard contact with the problems set by current documentation, and this immediately took effect once the extension of integrative level theory into General Systems Theory had begun.

In constructing a general scheme of classification for specific subject analysis, one has always to face the question "How specific is specific?".

This means, briefly, that there must be points set at which analysis stops. We do not express the analysis of committees, Trade Unions, or bicycles, in terms of their chemical constituents, even though these constituents are vitally important to the behaviour of all those entities. Integrative level theory provides a framework within which systematic analysis at appropriate depth can be made consistently. General Systems Theory provides a framework within which the component parts of any subject analysis can be ordered consistently. For a subject such as "the abrasion of rock beds by glaciers", integrative level theory provides a place in a scheme to which the entities can be assigned, while General Systems Theory states the sequence in which the terms (or their notational symbols) should be arranged to provide the most generally useful approach for the user.

It does this by distinguishing between "active" and "passive" systems. In a given environment containing two systems, one will always be found to be the "agent" of change, the other will be the "patient" in which the change occurs. The fact that an agent, or active system, also may change does not alter the fact that it is the instigator, or material cause, of the change. Thus the glacier, moving along its bed, causes parts of the bed to be worn away, or abraded, and is the active system, while the rock bed is the passive system. Clearly, the same entity will not invariably play the same rôle, either active or passive, but can at times be one, at times the other, according to its relation with other entities in the environment. A coat of paint is an active system in relation to the metal it protects, but a passive system in relation to the sunlight which causes it to blister.

In setting out a subject analysis in citation order, one has usually a choice to make. One can say either:

Glacier (active system)/Abrading/Bed (passive system)

*or*

Bed (passive system)/Being abraded by/Glacier (active system)

Examination of many examples shows that it is the passive system which is the main focus of interest. The rock bed would be cited first, followed by the [E] facet or operator indicating the action performed on it, followed by the various agents which cause the action. In this way the different abrasive agents, which might be scattered in the classification scheme schedules, will be brought together following the term "abrasion", when actually classifying documents. It is the behaviour of the rocks that are mainly studied in this context; if the citation order were reversed, the rocks would be scattered under the names of the various abrasive agents such as wind and ice. It is worth noting that this confirms the well-known principle of "purpose" in facet analysis, or citing the "end product" first in the citation order of facets in any class number.

This division of the concept of system now leads to an increase in the

number of operators, to include active and passive sub-systems and their interactions. Further analysis of the concept of "environment" showed that it implied rather more than the notions of Time and Space, in that the whole environment may at times be playing the rôle of an active system instead of merely providing a location in which changes happen. The Earth itself is no more than one particular environment when we are in the realm of Space Science.

From this, Austin elaborated the idea of "change of system". The properties of a system are followed by activities directed towards the maintenance of the system, but a new problem arises with such terms as "dissolving", "disintegrating", and even "growing", all of which imply that the system is changing into another system. It seems reasonable to suggest that all concepts that involve the material alteration of a system should appear at the end point of a sequence of terms used to enumerate qualifications of the system. This leads in a recognisable sequence to the next stage in a sequence of systems, that is, the point at which one system, by an accumulation of differences, changes into another system; at such points a new cycle of analysis begins. The sequence of operators was therefore as follows:

(1) Properties of system
(2) Second system/Environment related to (3)
(3) Activities and interactions not involving material change
(4) Active subsystem
(5) Passive subsystem
(6) Interactions within system
(7) Second system/Environment related to (8)
(8) Interactions involving change of system

This still left out of account the notion of the "changed thing", which had often appeared in CRG discussions, and had been given the title of "the evaporated milk complex". If we take a raw material (such as Milk), and act upon it with a particular process (such as Evaporation), we arrive at a new product (Evaporated Milk) which may, nevertheless, still be subjected to the same sort of process (such as Canning) as the original raw material. Originally, it was thought that it was sufficient to define the new entity by an enumeration of the [P] and [E] facets, that is, Milk – Evaporation. But Evaporation of Milk is not identical with Evaporated Milk, any more than the Redness of Stars is identical with Red Stars. We must recognise, therefore, a class of things (Milk, stars) which is distinct from another class of things defined by an attribute, whether the attribute is of its nature (Red), or imposed by artificial means (Evaporated). The final order of operators proposed by Austin is as follows:

(1) Properties of system
(2) Second system/Environment related to (3)
(3) Activities and interactions not involving material change
(4) Active subsystem
(5) Passive subsystem
(6) Interactions within system
(7) Second system/Environment related to (8)
(8) Interactions involving change of system
(9) Attribute defining a sub-class

The introduction of the (9) operator opens up a whole new area of investigation which has yet to be explored in detail. The implications for classifiers are clear, however, and the stages in classing a compound subject then become:

(a) distinguish and set down the various concepts discerned, designating them as *entities, properties* and *activities.* Translate concepts into notation.
(b) designate *entities* as passive system, active system(s) or sub system(s). Class at the passive system, indicating others with appropriate operators.
(c) assign *activities* to the entities to which they apply, and *properties* to either entities or activities.
(d) consider the attribute which is the *principal focus* of the document, and indicate this by the appropriate operator.
(e) indicate all other attributes by adding (9) to the operator which declares which entity possesses the attribute or which activity possesses the property of an activity.

## PROBLEMS AND PROSPECTS

Several important questions have been dealt with only cursorily so far in this research, but they will need a good deal more work before we can say that we are ready to make a new general scheme of classification. On the other hand, experience has proved that we can already do two important things: first, we know how to make a scheme for a special subject, in great detail, and with a highly articulated structure that corresponds to the way the experts think about their subject; and second, we have a very fair idea of a model structure covering the whole of knowledge, which has thrown up quite sharply problems arising out of the relations between subject fields.

A brief discussion of some of these problems will show the direction of future CRG discussions.

*Levels of disintegration*

First of all, perhaps, we might look at the other side of integrative level theory, what Austin has called "levels of disintegration". A whole, or entity, is identifiable because its parts have become ordered into a more or less stable relationship, by means of bonds between them. This linking alters the characters of the assembly from a mere collection into a new kind of unit with its own characteristics. When the bonds are broken, the unit disintegrates back into the pieces that make it up. This does not always, however, mean that these pieces go back into their original state. If we take the pedals off a bicycle, they remain pedals. But such ideas as a pathological or "wrong" state, and the processes of birth, maturation and decay, indicate change of a totally different kind. A flock of sheep may be divided up in one way so that they cease to be a flock and become several separate sheep, or in another way so that they become mutton. A classification scheme could not, and probably need not try to set up a whole series of levels of disintegration parallel to its levels of integration, but some way must be found to identify the relations of disintegration, so as to be able clearly to distinguish "pork" from "pig", and "lumber" from "tree".

*Levels of artefacts*

From her work on Mining and Sculpture, Helen Tomlinson suggested that one at least of the basic notions identifying artefacts was "purpose". The entities in the traditional fields of physics, chemistry and geology occur in nature, and do not come into being for the sake of some man-made plan. In mining and sculpture, on the other hand, those entities become the raw materials out of which Man may construct new entities according to some pre-determined plan. The architect, unlike the bee, can erect his building in his imagination before ever a stone is laid. So far, only Farradane and his team at the City University have attempted to apply a levels sequence to artefacts, with this result:

Separated natural materials (e.g. sugar, cane, juice)
Derived single product (screw, brick, sugar)
Assembly of derived single products (book, machine, wall)
Complex of assembled products (factory, railway, house)

Farradane stresses that such a sequence of levels represents uniform steps in a series of one type of increasing complexity, and he has abandoned an earlier attempt to impose an arbitrary order such as "industrial usage", which produced the levels of Component, Product, Instrument, Machine, etc.

*Levels of mentefacts*

We have already seen an illustration, in the Wordsworth Sonnet, of a set

of mentefact levels. Barbara Kyle had made little progress in her analysis, having mainly directed her attention towards the social sciences in general. In linguistics (perhaps the first level of mentefact), we have letters, words, phrases, sentences, paragraphs, chapters, whole works. The whole works may be of literature, philosophy, religion, and so on. None of these entities are real, or concrete, in the sense that they can be picked up – since they must not be confused with the book or document in which they are embodied. We cannot pick up the play of *Hamlet*, though we may well pick up the volume which contains a text of it. I do not think that the analysis of mentefacts will present much problem, even though the [P] facet of the CC classes that contain mentefacts are perhaps the most indistinct in terms of Ranganathan's own notion of what constitutes Personality.

*Enumeration within facets*

By far the greatest task remaining, though by no means a difficult problem, is the mere labour of enumerating the terms for all the entities, properties, and activities that will be needed to fill up the various facets. This is mainly no more than a clerical task, since the filling up of facets is reasonably straightforward once the levels on which facets are to be enumerated have been defined. Fortunately, there exist already many classification schemes, thesauruses, and various other lists of terms, dictionaries, and so on, which can be used as reservoirs from which facets may be filled up. Once the new scheme has been brought to the optimum approximation to completeness, this process will become reversible. Of course, no classification of knowledge ever becomes complete and stationary, as knowledge ceaselessly advances; but the scheme should provide enough detail for anyone needing a special scheme to draw upon for most of his terms.

*Notation*

A notation is always required to convert a scheme of arrangement of terms into a system usable for arranging books on shelves or entries in catalogues and bibliographies. Basically, there are four main functions for notational symbols, to identify clearly and distinctly:

Main Class (Basic classes, levels)
Facets ( [P], [M] [E], or equivalents)
Terms in facets (foci, isolates)
Relations between terms

Although CRG members have experimented with several different systems in our special schemes, no single system so far commends itself as superior to the others. Some qualities appeal to some people more than others, and a good deal of work has been carried out on the psychological reactions to symbols; a bibliography compiled by Mrs. Lynn Quiney at University College

London School of Librarianship runs to 645 entries. In a special subject scheme, it may not be necessary to use any symbol for the Main Class itself, but to distinguish only facets, terms, and relations. Vickery used numbers and letters in Soil Science, with an oblique stroke for relation:

6s/4f = Salinity in relation to irrigation

Farradane and several others, including myself, have used upper and lower case letters for facets and terms, with strokes or punctuation marks, usually the colon, for relations. Thus, in P. F. Broxis' classification for Fine Arts:

Rym Fh En = French Gothic stained glass

My own London Education Classification carries this a stage further by making all symbols of three letters, in the form consonant-vowel-consonant, so that they are pronounceable:

Rur Ram Dab = Planning rural schools for children under seven.

E. J. Coates, in the British Catalogue of Music Classification, used only capital letters, but in a retroactive manner, by which any step backward in the alphabet indicates a change of facet:

K F V D W Songs for middle range female voices

where K is "vocal solos", FV is "Female voices – middle range", and DW is "Songs". Although this system, like the notation of CC is basically very economical of symbols, I myself find that it can easily lead (again like CC) to rather complicated looking symbols on the backs of books, and so arouse unfavourable comment in a library.

More research is certainly needed in this area, but it is a secondary problem. The main target is the production of the schedules of terms within a systematic framework.

*Future plans*

The immediate plan is to complete the model of the structure, to test it and then fill out the basic list of terms. This will produce a controlled vocabulary covering the whole of knowledge which, when equipped with an acceptable notation system, could be used in libraries and bibliographies for arranging the items in useful order. But the CRG recognises two further objectives. One is the need, in special libraries, to be able to slant the arrangement to suit the local needs. For this, we intend to provide some rules for converting the basic arrangement into any other sequence by precise and easily understood methods. One simple example is the borrowing of facets from other subjects to enumerate marginal fields or second and subsequent Rounds of facets. Where a thesaurus is required, rules can be followed to convert the classification scheme into the form required. This has already been done, by Gordon Barhydt and his colleagues at Case Western Reserve University, with their Thesaurus of Educational Terms based on facet analysis, and also in the new Thesauro facet produced by Mrs. Jean Aitchison for the English Electric Company. The other need, perhaps the greater, is

to provide a basic vocabulary which will help to bring about compatibility between the many systems now in operation or being planned. It is obvious that, with the increasing application of computers to information control, efficiency and economy can only be achieved by close co-operation over all fields on an international scale. One of the working groups of UNISIST – the world science information system planned jointly by Unesco and the International Council of Scientific Unions – is studying "language problems"; this includes the control of index languages for information systems. These are just as necessary with computer processing as with shelf arrangement and manual indexes, as has been shown by two of the best and best-known mechanised scientific information systems – the Chemical Society's Research Unit in Information Dissemination and Retrieval at the University of Nottingham, and MEDLARS at the National Library of Medicine in Washington. In both of these, a higher proportion of irrelevant references appears in most search printouts, and can be largely attributed to defects in the entry vocabulary. When techniques of refining are introduced, as at Nottingham, the results can be much improved. Some at least of these defects are traceable to lack of systematic structure, and the refinements consist of introducing more structure and detail at the search stage.

Yet if this structure were present at the input stage, the quality of the indexes themselves would be improved, and the basic information problem of matching searchers' terms with authors' terms would be greatly simplified. Reality is the basis for the texts of documents; that is what authors try to describe, and what searchers are investigating. Classification, the investigation of structure, is the foundation of all our knowledge, and therefore should be the foundation of all our systems for controlling information. If the structure of the control vocabulary is based on what has won public acceptance, it will largely correspond to the way in which specialists think about their subject, and will therefore be acceptable to them as a genuine aid in their studies. A scheme such as the CRG proposes, being based on fundamental and more or less stable notions such as matter and motion, entities and activities, is not subject to the changes of paradigm caused by the advance of knowledge, to anything like the same extent as a scheme like DC or even UDC, which attempts to "find a place" for all the complexes that are likely to change even within short periods of time.

Now, moreover, is the time when such a scheme is required. The other general library schemes have proved intractable in face of the progress of knowledge; the new-style thesaurus, which overcomes some difficulties, fails to measure up to the challenge of the computer, and demonstrably requires a technique to help refine both index language and search language. Post-co-ordinate indexing can help to overcome the loss of recall due to distributed relatives, the secondary facets that are scattered throughout the linear

sequence produced by a classified arrangement. Yet the users of post-co-ordinate indexing, just as much as any other users of an information system, are still trying to form a pattern, or classification scheme, in their minds, and they will not be helped by the absence of system in their information store. In our techniques for information control the time is ripe for the overthrow of the existing paradigm; but we should not, at the same time, reject those aspects of it that can usefully contribute, for what we need now is not a blank slate, as was once thought, but a genuine synthesis.

# BIBLIOGRAPHY


Aslib. Proceedings of the International Study Conference on Classification for Information Retrieval, Dorking, 1957. London, Aslib, 1957.

Austin, D. W. Prospects for a new general classification. J. Lib., 1 (3) 1969, 149-169.

Batten, W. E. A punched-card system of indexing to meet special requirements. *In:* Report of the proceedings of the twenty-second conference. London, Aslib, 1947, 37-39.

Bergen, Dan. Implications of General Systems Theory for librarianship and higher education. Coll. and Res. Libs., 25 (5) 1966, 358-388.

Bertalanffy, L. von. An outline of General System Theory. Brit. J. Phil. Sc., 1 (2), 1950, 134-165.

Bruner, J. S. The process of education. N.Y., Vintage Books, 1960.

Caws, Peter. The philosophy of science: a systematic account. N.Y., Van Nostrand, 1965.

Classification Research Group. L.A. Record, 55 (6) 1953, 187-188.

Coward, R. E. MARC Record Service Proposals. Council of the BNB Ltd., 1968.

Documentation Research and Training Centre. Annual Seminar 4: Universe of knowledge, Depth classification, Documentation list. Bangalore, 1966.

Fédération Internationale de Documentation. Classification research: proceedings of the second international study conference, Elsinore, 1964. Copenhagen, Munksgaard, 1965.

Foskett, D. J. Classification and indexing in the social sciences. Butterworths, 1963.

*idem* Some fundamental aspects of classification as a tool in informatics. *In:* On theoretical problems of informatics. FID 435. Moscow, VINITI, 1969.

Foskett, D. J., *and* Palmer, B. I., *eds.* The Sayers Memorial Volume. London, Library Association, 1961.

Inhelder, Bärbel, *and* Piaget, Jean. The early growth of logic in the child: Classification and seriation. Routledge and Kegan Paul, 1964.

Kent, A. K. Chemical Society Research Unit in Information Dissemination and Retrieval Search manual. University of Nottingham, CSRU, 1967.

King, M. Report on the operation of the MEDLARS service in the Newcastle Region. University of Newcastle upon Tyne, 1969.

Koestler, Arthur, *and* Smythies, J. R., *eds.* The Alpbach Symposium 1968: Beyond reductionism – new perspectives in the life sciences. Hutchinson of London, 1969.

Kuhn, T. S. The structure of scientific revolutions. Chicago University Press, 1962.

Kyle, Barbara. Towards a classification of social science literature. Amer.Doc., 9, 1958, 168-183.

Lancaster, F. W. Information retrieval systems: characteristics, testing and evaluation. Wiley, 1968.

## BIBLIOGRAPHY – *continued*


**Library Association.** Classification and information control. 1969.

*idem* Some problems of a general classification scheme. 1964.

**Luhn, H. P.** H. P. Luhn: Pioneer of information science. Selected works, edited by Claire K. Schultz. N.Y., Spartan Books, 1968.

**Pantin, C. F. A.** The relations between the sciences. C.U.P., 1968.

**Piper, R. F., *and* Ward, P. W.** The fields and methods of knowledge. N.Y., Knopf, 1929.

**Puranik, K. D.** Field of knowledge and its repercussions on classification. Abgila, 2 (1), 1951, A19-A23.

**Ranganathan, S. R.** Colon Classification. Asia, 6th edn., 1960.

*idem* The Colon Classification. Rutgers State University, 1965.

*idem* Philosophy of library classification. Copenhagen, Munksgaard, 1951.

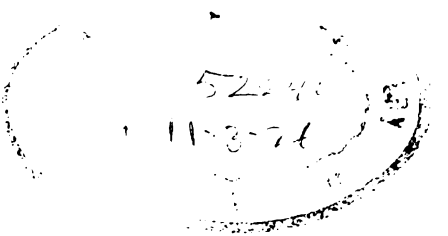
*idem* Prolegomena to library classification. London, Library Association, 2nd edn. 1957.

**Saussure, F. de.** Course in general linguistics. N.Y., Philosophical Library, 1959.

**Vickery, B. C.** Classification and indexing in science. Butterworths, 2nd edn., 1959.

*idem* Faceted classification schemes. Rutgers State University, 1966.

**Ziman, J. M.** Public knowledge: an essay in the social dimension of science. CUP 1968.

10s. (£0.50)
8s. (£0.40) to L.A. Members