ABSOLUTE SYNTAX AND STRUCTURE OF AN INDEXING AND SWITCHING LANGUAGE*

A. Neelameghan

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

1. TERMINOLOGY

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

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

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

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

---

* Based on S.R. Ranganathan's Postulates and Normative Principles: Applications in Specialized Databases Design Indexing and Retrieval. (1997)

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

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

2. STRUCTURE AND PATTERN

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

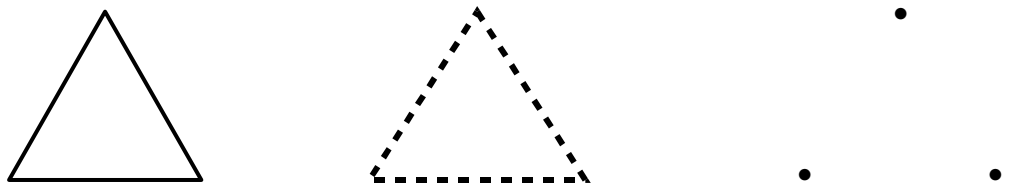


Fig. 1

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

Structure is the way in which the components of an entity are put together. Lancelot Whyte defines the concept of structure as "effective patterns of relationships in any situation"[2]. In a general sense, structure denotes logical form. The content of a logical form may be physical, musical, psychological, temporal, or in some other way non-physical. Anything that has structure has parts, properties or aspects, which are in some manner related to each other. Thus, in every structure one can distinguish the relations and the items which are related. The items may be qualities, values, or any conceptually distinguishable feature called elements of the structure[3].

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

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

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

3. REPRESENTATION OF SUBJECT

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

4. PROBLEMS IN THE EFFICIENT USE OF THE LANGUAGE OF SURROGATE

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

4.1 For various reasons, it may be difficult for the user to perceive precisely and express coextensively the subject of his interest at the moment. Therefore, the total semantic domain represented by the expression of the subject of his query may not be coextensive with the semantic domain of the subject of his interest as perceived by him.

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

4.3 A user may not and, perhaps, cannot be concentrating attention or work on at one and the same time on all the component ideas potentially falling in the subject of

his interest, even if it be a narrow one. He may study and concentrate attention only one or a few of the component ideas at a time. The recall value at the moment that is, the likelihood of being retained and recalled from memory for this component idea (or a few of them) would be relatively greater than that for the other component ideas in the subject of his interest. Therefore, he is more likely to bring up the name of this component idea (or of a few of them) in searching for information on the subject of his interest at the moment.

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

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

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

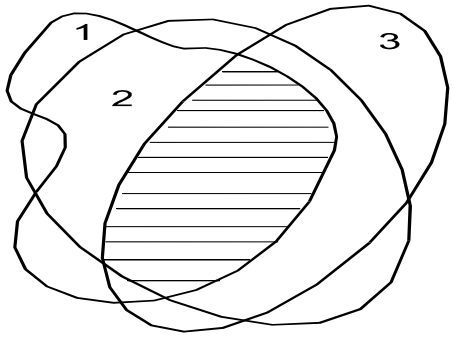


Fig. 2

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

▨ Common semantic domain

5. MINIMIZING THE CONSTRAINTS

5.1 Helpful Features of an Information System

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

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

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

1. Subjects greater in extension but subsuming the subject of the user's interest at the moment.

2. Subjects greater in intension but containing a substantial portion of it devoted to the specific subject of interest to the user at the moment.

3. Subjects in some other manner related to the specific subject of interest to the user at the moment (for example, collateral subjects, analogous subjects, and subjects studied in mutual relation to the subject of interest to the user).

5.2 Intersystem Connection and Compatibility

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

To use the same or very nearly the same information storage and retrieval language in all the information systems; and

To use an intermediate language or switching language through or by which one moves from one information system to another.

There can also be partial combination of 1 and 2.

6. FRAMEWORK FOR REPRESENTATION

6.1 Problems of Transformation

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

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

As mentioned in Section 3, there are various methods of representing subjects, such as, class numbers, subject headings or strings of words, multi-dimensional arrays, tree-structures, etc. These arise from the process of analysis of subjects of discourses into constituent elements; recognition of the relevant relations among the elements as they obtain in the context of the subject concerned; and assembling the elements in a preferred pattern so as to represent as coextensively as possible the subjects. (See Fig. 3). Representation of subject by a subject heading or a class number is equivalent to transforming the n-dimensional configuration of the subject into a linear configuration[6]. An arrangement of the component elements in each subject falling in a subject-field among themselves, in a sequence helpful to a majority of users requires keeping invariant every Immediate-Neighbourhood relation among all the subjects while transforming or mapping the n-dimensional configuration of subjects on to a line.

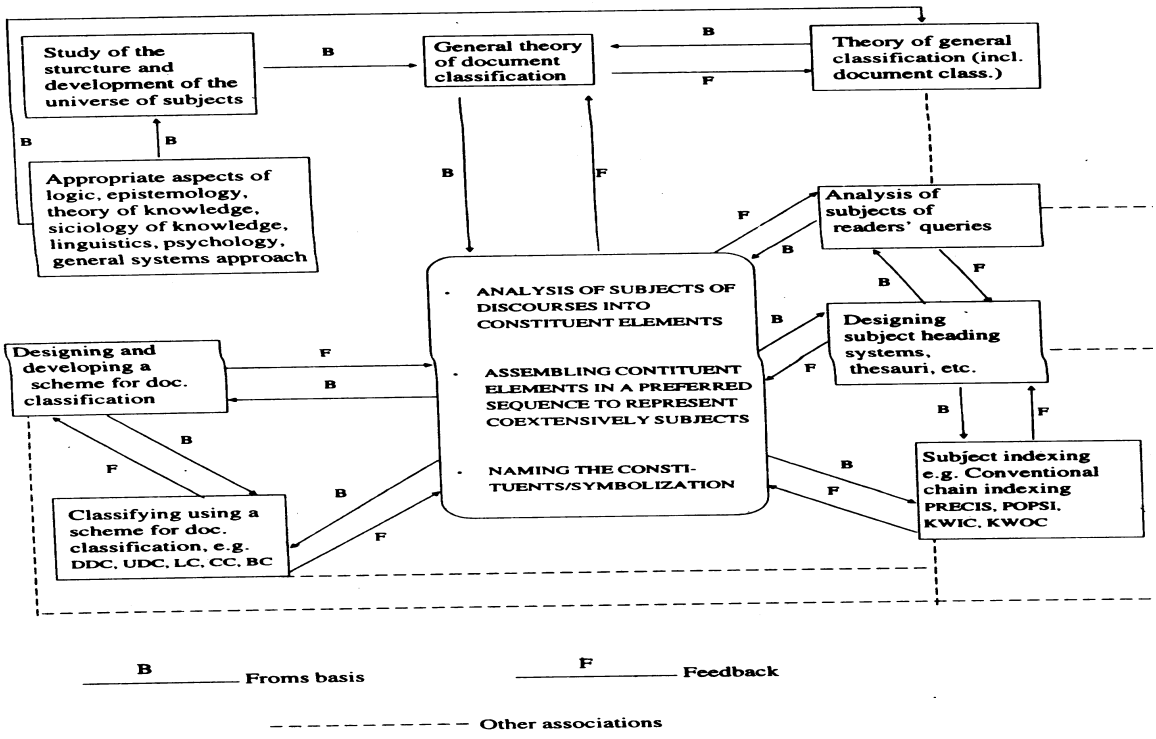


Fig. 3
Interrelation between discourse discourse/subject analysis, classification, subject indexing, subject heading work, thesauri, etc.

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

components of all subjects. But, such a consistency in pattern is helpful and necessary to the users, as well as the designers of the information system.

6.2 Criteria for Choice of Frame work

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

> "The long-term task is not merely to analyse the problems but to design methodological instruments for carrying out practical researches into problems of communication. These may be treated as problems of mapping. Given an original territory of factual phenomena how does this territory become mapped in the brain of the investigator? How does this map become transformed into a verbal or symbolic expression a linguistic map? How is this map transformed again into a language adopted to the needs of the ultimate recipient, the learner? Finally, how is this third map introjected into the learner's brain to form a pattern of knowledge? If we can establish a cartography for these maps, we can formulate "projective equations" leading from one map to the next. Each map will be a pattern of definable variables under appropriate controls and observing changes in the next map, the equations can be solved and laws of projection may be discovered."

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

1. The representation should be capable of expressing any conception which a human can formulate or understand.

2. The representation should allow for relatively efficient search for and retrieval of information. That is, specific information should remain relatively accessible even when the data-files grow to encyclopedic proportions.

3. The representation should saliently exhibit the substantive information extracted from a given input. It should not be influenced by the peculiarities of the particular natural language in which that information was communicated. This hope for language-invariance amounts to a wish for a universal interlingua in which any conception in any language could be expressed, but for which the format would not be specific to a particular language...

4. For reasons of parsimony, the representation should involve a minimum of formal categories. That is, it should make a minimum of formal (structural or syntactic) distinctions at the outset; more complex distinctions would be built up by the construction rules for concatenating primitive ideas.

5. The representation must allow for easy expression of concatenation operations, by which "duplex ideas" can be constructed out of "simple ideas". This means, for example, that the representation should allow easy expression of conceptual hierarchies, or multiply embedded predications, or allow one to predicate new information on any old information- structure".

7. ABSOLUTE SYNTAX

7.1 A Postulate

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

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

1. Helpful to majority of normal intellectuals;

2. Consistent in pattern in subjects falling in different subject-fields;

3. Relatively more stable and continue to be helpful to a majority of normal intellectuals so long as there is no mutation in their mode of thinking;

4. Free from the aberrations due to variations in linguistic syntax from the use of the verbal plane in naming subjects;

5. Capable of representing and indication of subjects co-extensively with a minimum number of variety of component elements;

6. Helpful in recognizing the less explored and unexplored regions in the universe of ideas; and

7. Helpful in probing deeper into the pattern of human thinking and modes of combination of ideas.

7.2 Analogy from Search for Linguistic Universals

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

1. A Semantic Component which will define the relations obtaining between semantic (including categorical) units or, rather hierarchically ordered clusters of semantic features (such as, (THING), (CONCRETE), (COUNTABLE), (ANIMATE), (HUMAN), (PERSONAL), (MALE), (ADULT); (PREDICATION), (AGENT), (DEFINITE), (ACTION), (PATIENT — ORIENTED), (TIME-DETERMINED), (ASPECT — DETERMINED), ETC.,

2. A Transformational Component which will convert the semantic deep structure representations into surface structure representations.

3. A Phonological (or Symbolization) Component."

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

"*Agenti ve* (A), the case of the typically animate perceived instigator of the action identified by the verb.

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

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

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

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

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

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

Leibniz's ideal language [16] and the Whorfian hypothesis [17] that "Every language contains terms that have come to attain cosmic scope of reference that

crystallize in themselves the postulations of an unformulated philosophy.. such are our words ‘reality, substance, matter’ and .. ‘space. time, past, present, future”, are worth noting here.

7.3 Biocybernetic View

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

> “..It is also becoming evident that all men, regardless of the culture they happen to belong to, have basically similar nervous systems, are equipped with analogous sense receptors, command like patterns of response, and use patterns of thought (whether rationally or emotively motivated) which obey very similar laws or regularities. In other words, there appear to be some ”universal" traits underlying cultural cognitive relativities: Chomsky could locate “linguistic universals” and Kluckholn discovered a number of" universal categories of culture
>
> “Finding such universals is rendered difficult if not impossible, by arguing out of one’s own culturally or individually relativistic categories. In that light, every other world-model becomes but a special case of one’s own, and is forced into the latter’s structural scheme. But, in using the neutral frame work of a cybernetic mode, one is no more arguing out of his own culture-categories than out of that of a thermostat. Conceptualizing the cognitive process with such categories, we can reach universal structures, for we are not dealing with particular contents. Regardless of whether a person conceives a sensory pattern as trees, meaning ”standing peoples, in whom winged ones built their lodges and reared their families" or interprets (presumably) the very same pattern as obstructions to be cut down and burnt; he is using a construct (or gestalt) which endows his perceptual input with meaning. And the development of constructs and gestalts obeys some general regularities, already manifest in biological evolution and set forth in cultural development".
>
> Lazlo further points out
>
> “..Regardless of the genetically and empirically induced differences, however, basic modes of thinking characterize all human beings, and indeed all higher biological species. These are rooted in, and explained by, the fact that all such organisms are self-maintaining open systems using a specific mode of reproduction, and forming part of some similarly specific social structure. The mental capacities needed to maintain such systems in their environments are adaptive functions; they crystallize as cognition in the more evolved species, and culminate in man.
>
> ”..The most immediately pertinent to human cognition make up an ascending ordering of categories, universally human in principle but variously evolved in different real individuals. These categories may be listed as follows:

1) Gesta lt (invariant patterns with established meanings to which the input patterns are assimilated);

2) Rational constructs (theoretic entities postulated through abstract reasoning and connected to the input patterns by means of some established rule of correspondence); and

3) Aestheti c construct (non-discursive meanings discovered in the input and illuminating some part of the knower's "felt experience"). These are the types of constructs which represent the limits of human cognition, given the kind of perceptions, cognitive organizations and effective output channels at our disposal. I argue that many forms of human experience do not constitute disjunctive culture-conditioned categories, but a set of universal structures which transcends individual and cultural differences and relativities, and accommodates as subclasses, the many varieties of cognitive patterns as environment mappings and constructions of natural cognitive systems on the specially human level of nature's hierarchy"

7.4 Syntax of Knowledge and Epistemics

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

> "At a multi-lingual conference.. with a community of disciplines, experience and thought, the translators have no difficulty in transforming, virtually instantaneously, the most elaborate syntactic forms of one language into the quite different forms of another whilst reserving the essential structure of information and conceptualization in the speech. Thus, there is a 'syntax of knowledge' which, even if not entirely independent of the particular languages, can and does, in practice, follow its own course alongside the syntactic sequence of language. It may serve to sharpen the difference if, provisionally, we think of the latter as governed by temporal relations (by the sequence of words in the sentence) and the 'syntax of knowledge' as primarily a spatial structure only shredded into temporal filaments in order to conform to the sequential character of speech.
>
> "This is a big step forward. Even though the syntax of language cannot be entirely divorced from the syntax of knowledge, we can pragmatically separate them by treating the one as a temporal sequence and the other as a spatial pattern. But, it may be objected, what about the temporal character of knowledge itself? Our knowledge of history, our under- standing of sequential operations, of industrial processes, of astronomical events etc., all of which involve time. Two points may be noted here: (1) Even though in a narrative the sequence of paragraphs normally (though by no means in every case) follows the time-sequence of the events narrated, this correspondence scarcely holds at all within the limits of a single sentence. And what is called linguistic syntax is largely based on the analysis of the single sentence. 'The assassin shot the President at the end of his speech'. In the actual event, the speech came before the shot; in the sentence after it. Thus, 'epistemic time' and 'linguistic time' are partially independent. (2) We speak a sentence sequentially, that is, at the beginning we have not yet spoken the end but what we are talking about even though it may be temporal event, is known to us throughout .. 'Epistemic time' is in fact 'dead' time, the completed past history, fossilised, and hence not "time" at

all in the linguistic sense. It has a discernible sequence but no flow. Our knowledge of it is a geometric knowledge of evidence spread out in space or held in memory".

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

7.5 Common Structure

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

"When a person hears a symphony, the messages sent by the orchestra reach the listener as air vibrations. These vibrations stimulate mechanically the receptors of the organ of corti, and these receptors set up nerve impulses along the fibres of the VIIIth nerve. It is clear that at this stage the physical events that are taking place are of an entirely different kind from those occurring in the instruments of the orchestra. Yet the message is preserved because there are similarities in certain features of the two series of events sounds emitted by the orchestra and nerve impulses traveling over the auditory nerves. The existence of these similarities of relations is precisely what is called a common structure. The mental decoding, which is the perception of the symphony, again preserves the corresponding relations. A common structure thus implies the quantitative preservation of the relations that exist between the independent constituents of an event or message through a set of trans- formations".

7.6 Logic of Exposition and Linguistic Syntax

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

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

7.7 Concept and Conception

Suzanne Langer points out that the psychological context of our thoughts may be private and personal. Therefore, two persons talking about the same thing may perceive it in different ways [22]. They are then said to have different conceptions. But, if they understand each other, then their respective conceptions embody the same

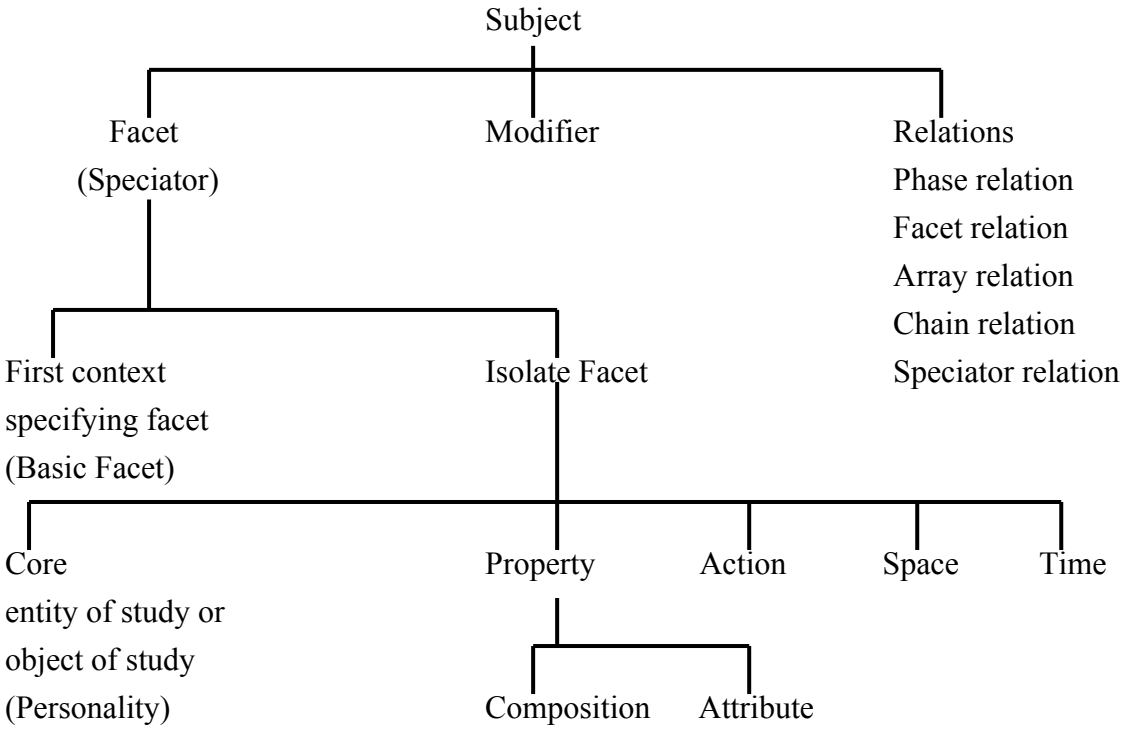
concept. A concept is an abstracted form. Abstraction is the consideration of logical form (structure) apart from content [23].

8. GENERALIZED FACET STRUCTURE FOR SUBJECTS

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

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

For Example:



3. Developing a typology of Basic Subjects, the modes of formation of Basic Subjects, and the arrangement of Basic Subjects[24].
4. Developin g a typology of Modifiers (speciators) for basic facet and for isolate facet in different subject-fields.
5. Recogniz i ng the relative strength of bond (relation) between the first context-specifying element (basic facet) and other types of facets in subjects[25].
6. Formulating principles for helpful sequence among

   (a) Facet s of a subject
   (b) Speciators to a facet

(c) Compound subjects falling in a particular subject-field
(d) Subjects falling in different subject-fields (26).

7. Developing a Generalized Facet structure (Model) of subject, with specific models for different subject-fields.

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

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

The chart in Fig. 3 shows the interrelation between subject- structuring, designing a classification scheme, generation of subject indexes, etc. Depth classification schemes for over a hundred subject-fields have been designed and several hundreds of articles, technical reports etc., have been classified using these schemes in each subject-field [28]. The structuring of subjects and the sequence in which the subjects get arranged have been found to be acceptable to a large number of users.

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

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

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

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

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

REFERENCES

1. Ranganathan, S.R. Prolegomena to Library Classification. 3rd ed, Asst. by M.A. Gopinath, 1967. Webster's Third New International Dictionary of the English Language, Unabridged, 1966.

2. Whyte, L.L. Accent on form. 1954, B 28.
3. Langer, S.K. Introduction to Symbolic Logic. 3d ed, rev, 1967, Chapter 1.
4. Bruner, J.S. and others. Study of Thinking. 1956, p. 12
5. Anderson, J.R. and G.H. Bower. Human Associative Memory, 1973, p. 151.
6. Ranganathan, S.R. Classification and Communication. 1961, Sec 14632.
7. Meredith, P. Instruments of Communication: An Essay in Scientific Writing. 1966, p. 64.
8. Anderson, J.R. and G.H. Bower. op cit. p. 152.
9. Ranganathan, S.R. Quoted in ref 10. Sec 7.2.
10. Ranganathan, S.R. "Hidden Roots of Classification" Information Storage and Retrieval. Vol. 3, 1967, Sec.7.
11. Russel, B. Has Man a Future. 191, Penguin Special, S P KIilpatrick, W.H. "Philosophy of Education". 1951, Quoted in J. Park ed, Selected Readings in the Philosophy of Education. 3d ed. 198, p.107-117.
12. Neelameghan, A "Sequence of Component Ideas in a Subject", Library Science with a Slant to Documentation. Vol. 8, 1971, Paper Q.
13. Birnbaum, H Problems of Typological and Genetic Linguistics Viewed in a Generative Framework. 1970, p. 9-18, 58- 70.
14. Filmore, C.J "The Case for Case", In E. Bach and R.T. Hams (eds.), Universals in Linguistic Theory. 1969, p. 1-88.
15. Vleduts, G.E. and N.A. Stokolova. "About a Methhod of Constructing Information Languages having Grammar", Tr. from the Russian by joe Lineweaver DID/CR Report 13. 1974.
16. Neelameghhan, A. op cit. Sec 431.
17. Whorf, B.L. Language, Thought and Reality. 195, p. 61.
18. Lazlo, E. Introduction to Systems Philosophy: Toward a New Paradigm of Contemporary Thought. 1972, p. 200-204.
19. Meredith, P. op cit.. Chapter, Section 2.
20. Amsden, C.E. "Syntax — more Sinned Against than Sinning" In, Some Problems in Reading. Claremont, Calif. Claremont Graduate School. 193, p. 90-93.
21. Rosenbleuth, A. "Mind and Brain", A Philosophy of Science. 1970
22. Langer, S.K. op cit. Chapter 3.
23. Langer, S.K. op cit. Chapter 1.
24. Neelameghan, A. "Basic Subject", Library Science with a Slant to Documentation. Vol. 10, 1973, Papers F to N.
25. Ranganathan, S.R. "Colon Classification", Rutgers Series on Systems for Intellectual Organisation of Information, 1965, Chapter V.
26. Parkhi, R.S. Library Classification: Evolution of a Dynamic Theory. 1972, Chapter R. Neelameghan, A. Presentation of Ideas in Technical Writing, 1975.
27. Gopinath, M.A. "Classification Research (India): 1968-1973", FID/CR Report 14. 1974.
28. Gopinath, M.A. op cit. p. 37-44.
29. See ref 12.
30. Neelameghan, A. and Venkataraman, S. "Use of Computer for the Synthesis of Class Number: A case study with a Freely-Faceted Version of Colon Classification", Library Science with a Slant to Documentation, Vol. 5, 1968, Paper S.

Appendix

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

$B - b_1 - b_2 \ldots - b_n$ $C' - c_1' - c_2' - \ldots c_n'$

$C'' - c_1'' - c_2'' - \ldots - c_n'' \ldots,$

$C^x - c_1^x \ldots - c_n^x;$ $M' - m_1' - m_2' - \ldots - m_n';$

$M'' - m_1'' - m_2'' - \ldots m_n'';$

$M^x - m_1^x - m_2^x - \ldots m_n^x;$ $P' - p_1' - p_2' \ldots - p_n';$

$P'' - p'' - p'' - \ldots - p'' \ldots;$

$P^x - p_1^x - p_2^x - \ldots p_n^x;$ $A' - a_1' - a_2' - \ldots - a_n' *:$

$A'' - a_1'' - a_2'' - \ldots a_n'' *: \ldots *:$

$A^x - a_1^x - a_2^x - \ldots - a_n^x.$ $. S' - s_1' - s_2' - \ldots - s_n'$

$. S'' - s_1'' - s_2'' \ldots - s_n'' \ldots.$

$S^x - s_1^x - s_2^x - \ldots - s_n^x \ldots 'T' - t_1' - t_2' - \ldots - t_n'$

$'T'' - t_1'' - t_2'' - \ldots - t_n'' \ldots$

$T^x - t_1^x - t_2^x - \ldots - t_n^x$

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

where

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

$P' P'' \ldots P^{x}$ represent respectively Property Isolate ofsuccessive removes.

$p' p'' \ldots p^{x}$ represent respectively Speciators to $p' p'' \ldots p^{x}$

$A' A'' \ldots A^{x}$ represent respectively Action Isolate of successive removes

$a' a'' \ldots a^{x}$ represent respectively Speciators to $A' A'' \ldots A^{x}$

2 Document: Cold storage studies on dusehri mango by K K Singh and ots. (*Journal of research (Ludhiana)*). 5,*N*4;1967; 516-22)

Chemical technology, Food Mango-Dusehri;

$B \qquad ,C' \qquad -c_1'$

Storage life influenced by Mango:Preservation-by Coating-

$;P'$ (phase relation) $C' \qquad : A' \qquad -a_1'$

Antimicrobial wax:Packaging-in Polyethelene bag: storing-

$-a_1'' \qquad : A'' \qquad -a_1'' \qquad : A'''$

at 2° C-for 3 Weeks

$-a_1''' \qquad -a_2'''$

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

Commodity production engineering, Lathe-Turret type-

$B \qquad ,C' \qquad -c_1'$

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

$-c_2' \qquad -c_3' \qquad -c_4'$

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

$-c_5' \qquad -c_6'$

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

$-c_7' \qquad -c_8' \qquad -c_9' \qquad -c_{10}'$

with Saddle movement maximum 330 mm-with Turret side

$-c_{11}'$

travel maximum 75 mm-with Attachment for drilling-with

$-c_{12}' \qquad -c_{13}'$

Program control

$-c_{14}'$

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

Banking, State Bank of India; Credit-by Lending

B ,C' ; P' $-p_1'$

-to Joint stock companies-for Working capital-

$-p_2'$ $-p_3'$

against collateral security-by Hypothecation-of

$-p_4'$ $-p_5'$

Moveable goods; Risk: Calculation-Debt equity ratio method

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

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

International trade, Commodity; Price; Variation; Trend

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

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

$-p_1''$ (phase relation) ; p' ; p''

:Evaluation-Elasticity of demand point of view

: A' $-a_1'$

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

International trade-Developing country, Export-

B $-b_1$, C'

Primary commodity; Grain-Long term: Determination-

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

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

$-a_1'$ $-a_2'$: A''

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

Medicine-Industrial, Human body; Disease-caused by

$B \quad -b' \quad , C' \quad ; P'$

Bacteria>Eubacterials>Brucella melitensis-Fever symptom-

$-P_1' \qquad -P_2'$

Intermittent

$-P_3'$

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

History, India; Independence movement-Indian National Congress;

$B \quad , C' \quad ; P' \quad -p_1$

Emergence *influenced by* United Kingdowm; Policy-Colonial

$; P''$ (Phaserelation) $, C' \quad ; P' \quad -p_1'$

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

$-'$ for a Time denoting speciator (e.g. John who came earlier than Smith)

$-.$ for a Space of Location denoting Speciator (e.g. Box on the Conveyor belt)

$-:$ for an Action denoting Speciator (e.g. Plastics produced by moulding)

$-;$ for a Property denoting Speciator (e.g. Paper coloured whi