

Terms and Classification Systems  
—An Introduction to TOCS Principle—  
用語システムと分類システム  
—TOCS 原則の応用—

*Yukio Nakamura*  
中 村 幸 雄

要 旨

インデクシング言語の二大潮流の一つである用語システムでは当初階層構造は不要と考えたが、現在の標準的シソーラスでは用語相互間に階層を備えることが普通となってきた。シソーラスが用語カテゴリーと階層構造を持つ場合には、これを分類表の形に書換えることは容易である(表1, 2)。

分類法では階層構造がその記号の中に表示されているが、シソーラスではそれは拒否されている。一方シソーラスでの階層性は一般に部分的断片的である。また分類法でもファセット式となれば、一体系の中に独立した何箇かの階層構造が併立することとなり、また複数箇の分類記号(標数)の併置(co-ordination)で複雑な概念を示すこととなる。この併置は従来用語システムの専売と誤解されているが、分類法でもそれがファセット方式であるなら、用語の場合と同じく併置を利用することとなる。分類表を見る手間はシソーラスを見る手間と同じ程度である。

残る問題はファイル作成時のインデクシングと、検索式の不一致だが、これは原則的になくならないものであり、ここに生ずるギャップを克服するため、上位検索の利用が必要となり、これは併置法と階層構造を持つ表記法とを共に用いることにより有利な解決ができる。

Ⅱにおいては、このような要求に会う分類表を構成するに当って、従来比較的無視されてきた概念出発点方式を用語出発点方式に切り換えるため、表現様式に必要な変更を加えることにつき実例を示して考察し、これができれば分類法における用語索引の作成が極めて楽になることを示した。

Ⅲにおいては上記のことを実現するに必要な、分類法記述の用語とその表現に与えるべき約束ごと(これを TOCS 原則と呼ぶ)を確立する方法を論じ、これを UDC に適用することには困難があるが、ある程度実施して UDC 日本語中間版の分類表および索引のコンピュータ作成を容易にすることを狙っている。新しく作成する分類法の場合なら TOCS 原則の適用は容易で、既に完成した二つの例(CST および CAST)を述べている。

Yukio Nakamura, Authorized Consulting Engineer, Nippon Telecommunications Consulting Co., Ltd.  
中村幸雄: 日本通信協力株式会社 技術士

Terms and Classification Systems

- I. Classification and Terms Systems for Indexing
- II. Terms in Classification Schedules
- III. TOCS principle and its application
- VI. Conclusion

**I. Classification and Terms Systems for Indexing**

**A. Difference between Classification and Terms Systems**

It was thought, at one time, that a terms system for indexing requires no hierarchical relationship among the terms. Now the terms system recommended as a standard is going to contain a sort of hierarchy in it.<sup>1)</sup>

Basically terms correspond to concepts and many concepts relate to each other by a hierarchical structure—generic-specific and/or

whole-part relations. A classification system adopts, without much consideration, this hierarchical structure in its notation structure. The majority of thesauri existing now have categories and subcategories as well as BT-NT relations which are nothing but a hierarchical structure. Then a thesaurus is very similar to the combination of schedule and index of a classification system.

Table 1 reproduces a part of *Macrothesaurus*<sup>2)</sup> edited by OECD mainly for social science problems. This thesaurus is equipped with categories (shown by centesimal notation) and BT

Table 1. An Extract from the OECD *Macrothesaurus*

09.03	MARKETING		Sales promotion/Promotion des ventes	055050
09.03.01			USE Marketing	055060
	Cooperative marketing/Vente en coopération	054830	Selling	055070
	BT Marketing	054840	USE Marketing	055080
	Market access/Acces aumarché	054850	Wholesale marketing/Vente en gros	055090
	RT Commercialization	054860	BT Marketing	055100
	Market	054870	RT Wholesale trade	055110
	Market study/Etude de marché	054880		
	RT Market	054890	09.03.02	
	Marketing/Vente	054900	Purchasing/Achat	055130
	UF Sales promotion	054910	Turnover/Chiffre d'affaires	055140
	Selling	054920		
	NT Cooperative marketing	054930	09.03.03	
	Retail marketing	054940	Buffer stock/Stock régulateur	055160
	wholesale marketing	054950	BT Stock	055170
	— FAO	054960	Inventory/Inventaire	055180
	RT Market	054970	Stock/Stock	055190
	Salesman	054980	NT Buffer stock	055200
	Marketing board/Office de vente	054990	Storage/Emmagasinage	055210
	Marketing cooperative/Coopérative de vente	055000	NT Cold storage	055220
	BT Cooperative	055010	Food storage	055230
	Retail marketing/Vente au détail	055020	RT Silo	055240
	BT Marketing	055030	Storage capacity/Capacité d'emmagasinage	055250
	RT Retail trade	055040	Wharehouse/Entrepôt	055260
			RT Silo	055270
			09.03.04	

Table 2. OECD's *MACROTHESAURUS* converted into classification schedule form

<b>09</b>	<b>Commerce. Trade</b>
<b>09.03</b>	<b>Marketing</b>
09.03.01	(Entry unnamed)
.1	Marketing. Sales promotion. Selling
.12	Cooperative marketing see Cooperative 03.03.04
.14	Retail marketing see Retail trade 03.03.04
.16	Wholesale marketing see Wholesale trade 03.03.04
.3	Marketing board
.5	Market access see also Market 09.01.02
.6	Market study see also Market 09.01.02
09.03.02	(Entry unnamed)
.2	Purchasing
.4	Turnover
09.03.03	(Entry unnamed)
.2	Inventory
.3	Stock
.33	Buffer stock
.4	Storage see also Silo 07.04
.42	Cold storage
.44	Food storage
.5	Storage capacity
.7	Warehouse
09.03.04	(Entry unnamed)
.2	Publicity
.22	Advertising
.4	Public relations
.6	Trade marks
09.03.05	(Entry unnamed)
.2	Display
.4	Exhibition

NT relations. If the terms in a category are exclusive of each other, it is an easy task to transform their contents into a classification schedule form. As shown in Table 2, this style conforms to the existing UDC schedule practice except that centesimal notation is partly used.

### B. Difference of Hierarchy in Terms and Classification Systems

One of the differences in the classification and the terms systems is that the former is associated with a notation involving hierarchy, whereas the latter is not provided with such notation, and the hierarchy is shown by BT-NT relations or in the "hierarchical index" appearing in a thesaurus. It is possible, of course, to have a notation or code involving a hierarchy with a terms system, such as proposed by J. W. Perry and his followers in their semantic factoring method, but it is rejected by most people in the United States for the reason that the use of such codes nullifies the advantage of the terms system—the direct use of terms in their natural form and requiring no detailed knowledge about the subject fields and the system from the user.

Another difference is that the hierarchy found in a thesaurus is local and fragmentary comparing with the hierarchy in the traditional classification systems which cover the whole of knowledge with a single tree-structure.

However, the introduction of facet concept admits the parallel existence of several independent tree-structures (facets) simultaneously. The objects to be classified relate to, in principle, each of these hierarchical structures with the result that these objects are represented by the co-ordination of notations corresponding to each hierarchical structure.

When the retrieval is made by visual inspection, the co-ordinated notation gives nothing but a single train of complicated characters, and the advantage of co-ordination is lost. It is only with the use of a mechanical sorting device or a computer that the advantage of co-ordination (irrespective of the order of co-ordinated elementary codes) is fully realized.

This advantage has been erroneously associated with a terms system. The co-ordination has a common feature whether the co-ordinated elements are terms or classification notations.

As an example, suppose a subject "Operating cost of supersonic jet carriers" is given. The terms system (e.g., TEST<sup>23</sup>) will give a co-ordinated descriptors sequence as:

### Terms and Classification Systems

Jet aircrafts \* Passenger aircrafts \* Operating costs (1-1)

while the classification system (e.g., UDC in Medium Edition level<sup>4)</sup>) gives the following compound notation

629.7.036.016.54 : 656.7.003.3 (1-2)

629.7.036	Jet propulsion aircraft
.016.54	Supersonic (aircraft)
656.7	Air transportation
.003.3	Cost in accountancy point of view

(1-1) and (1-2) show similar structures and the advantage of co-ordination can be fully utilized in the both representations.

The inconvenience of consulting a schedule and an index in a classification system is equal to that of a terms system which requires consulting with a thesaurus containing the "Thesaurus of terms," "Hierarchical index" and "Subject category index."

As everyone is well aware of, the indexing and retrieval activities are closely related and inseparable. But this does not mean that the indexing level and the level of search instruction formulation should always be the same.

Some system designers simply postulate that both levels of indexing and search instruction formulation should be the same or that they try to level off both levels. However, there exists a certain uncertainty in these two.

It is reasonable to assume, from the very beginning, that there is a certain uncertainty

the indexing levels of stored records and a discrepancy between the level conceived by the searcher and the indexer. It is therefore necessary to devise an effective strategy to overcome these gaps. This fact has been recognized and kept in mind by many experienced searchers. It is regrettable that system designers, especially for computerized systems, are apt to neglect this fact because of their habit to exclude any uncertainty in handling logics.

One of the useful strategies for this, has proven to be efficient through experience, is the combination of generic search and co-ordination at a less deep level.

For an example, a document dealing with "the insulation test of porcelain (as electrical insulation material)" is sought. Different indexers will formulate their subject in different ways (see Table 3).

The strategy to use generic search, i.e., the search at higher hierarchical level, will increase the recall ratio but also the noise ratio. If a single search is postulated the increase of noise ratio is not advisable for a better retrieval. On the contrary, if the output is checked by human judgment or if the number of relevant output records is automatically given for each search instruction (which is always possible at the retrieval by computer), this gives the searcher a chance to choose an adequate instruction in consideration of the characteristics of the existing database.

Table 3. Different ways in indexing

Level	Subject conceived	Indexed by UDC	Indexed by TEST
1	Insulation testing of porcelain (as an electrical insulating material)	621.317.333 : 621.315.612.2 Insulation testing      Porcelain (as insulating material)	Insulation testing * Electrical porcelain
2	Electrical insulation testing of an insulating material	621.317.333 : 621.315.61 Insulation testing      Electrical insulating material	Insulation testing * Electrical insulators
3	Electrical measurement of insulating material	621.317.3 : 621.315.61 Electrical measurement      Electrical insulating material	Electrical measurement * Electrical insulators or Electrical measurement * Electrical insulation or Tests * Electrical insulation

When formulating a generic search instruction, if the indexing language is a terms system, a completely different term must be chosen for designating a broader concept. For instance, in moving to level 2 from level 1 in Table 3, changes of the terms, (1) from "Insulation testing" to "Electrical measurement" and (2) from "Electrical procelain" to "Electrical insulators," can only be made after consulting a thesaurus.

In a classification system, the change (1) is from 621.317.333 to 621.317.3, and the change (2) is from 621.315.612.2 to 621.315.61, and both are made by deleting the last two digits in the notations. This can be done very easily by a small addition in the instruction.

If a deeper level of indexing is obtained by increasing the number of co-ordinated elements (or notations), the co-ordinated terms system and the faceted classification system have an equal adaptability to a generic search.

Thus it can be concluded that generic search can be facilitated by the notation of indexing language containing hierarchical structure therein.

## II. Terms in Classification Schedules

### A. Example Terms in a Schedule Entry

The text of a classification schedule entry must give enough explanation on the content of the entry. This means the entry must contain its definition. The definition may be a single word, if it is possible, but in many cases it necessitates a rather lengthy expression, a noun phrase.

Even if the definition is technically sound and complete, the addition of example(s) is desirable for better understanding by users.

In UDC, for instance, the entry:

659.132.43 Publicity letters

in the class 659.1 Publicity, appearing in the Medium Edition level, does not contain such a word as "Direct mails," but it is without doubt an example term of this entry. It is, therefore, desirable to have an index entry:

Direct mails 659.132.43

in the index of the Medium Edition. However,

how is it possible to get that index entry from the schedule entry when automatic production of schedule and index is envisaged? For this question the author has given a solution in his paper<sup>5)</sup> and has suggested the use of the master tape including such entry as:

659.132.43 Publicity letters; Direct mails.

Here the term "Direct mails" is an example term. If many example terms are deemed necessary for the index, mentioning of such example terms should not be neglected even if the entry becomes lengthy.

In the same UDC edition the entry:

331.8 Hours, conditions, facilities, organizations, etc.

has no clear single definition. Even if the broader concept 331 is for "Labour. Work. Employment," these enumerated terms represent nothing but a logical sum of the concepts, working hours, labour conditions, labour facilities and labour organizations.

This kind of enumeration cannot be regarded as an appropriate listing of the example terms. It should be considered as an "indirect definition by enumeration (IDE)," for the case where no adequate single term can be found. On the contrary, in some of the entries by enumeration, there is a doubt if they are necessary if the limitation for the number of subdivisions due to decimalism is removed.

If the two cases of example terms and of IDE are mixed, the situation becomes worse. The users of a classification system have difficulty understanding the real meaning of an entry.

It is, therefore, necessary to distinguish the definition phrase from the example terms in the text of the schedule entries, by using some simple devices. The author has suggested to put semi-colon between the definition phrase and example terms, the last of which ends with a period.<sup>6)</sup>

### B. Production of Indexes

The alphabetical index of a classification system is essential since it is an indispensable half of the system. If the publication of the index is made at a later time than that of the

schedule, there arises inconveniences not only from the discrepancy of the publication dates but also from the discrepancy in the contents of both due to a continuous revision work, as in the case of UDC, which adapts itself to the incessant change of the outer world.

Some classificationists, however, regard the compilation of such an index as a burden to be avoided if possible, since the compilation work needs the comparable labour to that for the schedule. However, the labour required for compilation can be minimized by the use of a computer with proper considerations.

This possibility suggests that it is worthwhile to consider the desirable structure of a classification schedule which simplifies automatic production of both the schedule and the index. For further detail on the subject, refer to another study<sup>7)</sup> by this author.

### C. The Relation between Schedule and Index Entries

#### 1. Addition of Example Terms

If an index has as many entries as possible, it will benefit the users in their searches, but there is a doubt whether it is necessary to keep one-to-one correspondence in the entries of both the schedule and the index. Some hold the opinion that if it is done, the text of the schedule will become so crowded with similar but not identical terms that it will bother the users.

On the other hand, if the text of the schedule is too short and simple, the users will suffer from the difficulty in interpreting the content when they use it for indexing and searching purpose. It is our common experience that the schedule with the texts in abstract expressions cause much difficulty in its use unless the user is well versed in the subject field.

The compilers of the existing schedules, such as of UDC, are well aware of this fact and the insertions of example terms are done extensively where it is necessary. However, the schedule does not distinguish the definition phrases from the example terms explicitly, and it is left to the user's interpretation. But this very fact causes difficulty to the users in many

cases, especially when editing a new edition in a language different from that of the original edition.

Consider the case of class 378 in UDC :

378	Higher education. Universities. Colleges (of advanced study)	(2-1)
-----	---	-------

In this entry, the users can easily judge that the term, "Higher education" is a definition and the terms, "Universities" and "Colleges (of advanced study)" are example terms though there is no explicit sign in the text.

In the index of the Abridged Edition level, the index entry is :

Higher education	378
------------------	-----

This is correct, but if the next two entries are given as :

Universities	378
Colleges (of advanced study)	378

by inverting the text of 378, these are obviously incorrect since the following subdivisions of 378 exist in the same edition :

378.4	Universities	} (2-2)
378.6	Colleges, Schools, institutes of higher education and advanced study. College of Advanced Technology (CATs). Technological universities	

The correct index entries in this edition level are :

Colleges	378.6	} (2-3)
Higher education	378	
Universities	378.4	

In converting the schedule text into index entries, if the same terms appear in two or more different positions, only the lowest one in the hierarchy should be adopted for the index. This preference process should be done automatically by an adequate programme, or if it is not advantageous, by providing special signs, e.g., \$ sign at the beginning and the end of the terms to be deleted in the inversion process. Thus the text (2-1) shall be changed (in TOCS representation which is discussed in III. A.) into the new text shown below at least in the

master file :

378 Higher education ; \$ Universities,  
Colleges (or advanced study) \$

2. Logical Sum Entries

In an entry formed by the logical sum of many subentries and where these subentries are represented by subdivision numbers in the schedule, it is of no use to have an index entry for the logical sum type entry (LSTE). It is, therefore, necessary to stop the inversion of text-classification number pair for index production by the use of the same sign as mentioned in the above example.

Example. The Case of UDC 645.42  
 645.42 Beds. Cots. Divans. Couches  
 645.421 Beds  
 645.422 Cots. Cradles  
 645.423 Divans. Couches. Settees  
 645.424 Sofas  
 645.425 Folding beds. Camp beds  
 English Abridged Edition  
 (1963) with corrections

In this example, 645.42 is a LSTE and the whole text of 645.42 should be deleted when inverting the schedule into an index. Thus the entry in the master tape should be in TOCS representation :

645.42 \$ Beds, Cots, Divans, Couches \$  
 This text gives such correct entries  
 Beds 645.421  
 Cots 645.422  
 Cradles 645.422  
 etc.

and not

Beds 645.42  
 645.421  
 Cots 645.42  
 645.422  
 etc.

3. The Case of Very Long Expression

The production of index entries from schedule entries which are complicated in text expression causes a problem. For instance, the entry

659.43 Public relations according to goal or objective.  
 PR aimed at, directed to...(any entity, organi-)

zation, person or group, product, fact or event) e.g.

:372.168 PR directed to family planning  
 :502.7 PR aimed at nature protection  
 :614.23 PR directed to doctors

is created to give such compound UDC numbers as :

659.43:362.178 PR directed to family planning  
 659.43:502.7 PR aimed at nature protection  
 659.43:614.23 PR directed to doctors.

However, a consideration must be given to produce index entries. There are solutions such as shown below :

a) To have such index entries :

Family planning,  
 PR directed to 362.178:659.43  
 Nature protection,  
 PR aimed at 502.7:659.43  
 Doctors,  
 PR directed to 614.23:659.43

b) To have such index entries :

PR aimed at nature protection 659.43:502.7  
 PR directed to doctors 659.43:614.23  
 PR directed to family planning 659.43:362.178

c) To have the combination of a) and b).

Three examples shown at the entry 659.43 (2-6) is meaningful but having colon combination numbers at each of 362.178, 502.7 and 614.23 is too arbitrary a choice. Thus it is better to have such index entries as

PR aimed at..... 659.43..... :  
 PR directed to..... 659.43..... :

besides the generic index entry

Public relations ... 659.4.

For this purpose the schedule entry should be changed, in TOCS representation, into :

659.43 Public relations according to goal or objective ;

expl subdiv  
 659.43...PR aimed at..., PR directed to...  
 since 659.43...is a subdivision of 659.43.

**III. TOCS Principle and its Application**

**A. The Importance of Terms and Words in Classification System**

The importance of concepts has been stressed in classification system and the terms and words to represent the concepts have had only a secondary importance.

On the contrary, there is a tendency that the human daily thinking and action are based on terms and words and not on concepts. Especially among pragmatic people, even on nationwide scale, this tendency is prevailing. Classificationists may dismiss such a tendency as being careless or thoughtless, but it is a fact that every human action can not be made after a rigorous concept analysis. This fact allows for an assertion that classification systems are useful and yet it is not "practical." Classification systems must be improved in this respect.

TOCS (Term-Oriented Classification System) principle is to make classification systems more easily accessible from terms and words by reforming the style and structure of the schedule so as to make the preparation of alphabetical index easy and useful.

To achieve this objective, several guiding principles must be adopted as shown below :

1. One term should not correspond to two or more classification entries. To make this principle workable, terms must be specific enough. Take an example; the word "bed" relates to many concepts in the schedule entries in UDC. These concepts with their representation in terms must be distinct.

There are two entries which are expressed by the term "beds."

These are

Beds as furnitures	645.421
Beds of machines	62-216

These should be changed into two distinct terms as

Beds (BC Furnitures)	645.421
Beds (BC Machine	

components)	62-216
Other terms including the word "beds" or "bed" are, e.g.,	
Ocean bed	(26.03)
Railway beds	625.12
Road beds	625.731.1
Seed beds	631.53.03
etc.	

The last group treated as single terms where the word "bed" or "beds" are not separable from the other combining words.

Thus, with this precaution, one term corresponds to one classification schedule entry. On the contrary, one classification schedule entry contains many terms, in principle, so that the reversed relation "one classification entry corresponds to one term" does not hold.

**2. Classification System with Concept Coordination**

A classification system based on the coordination principle uses mostly representation of complicated concepts by combining simpler concepts with shorter notations; thus, if the principle is applied to the case of UDC, the lengthy UDC main numbers will be replaced by a combination of two or more short UDC numbers.

The application of this principle is now encouraged in UDC by the Central Classification Committee of FID in every occasion to revise UDC. This means, instead of extending main numbers, one among the following combinations are preferred in revised entries :

(main number) (auxiliary number),  
 (main number):(main number), or  
 (main number):(main number with auxiliary number(s)).

Example :

In UDC class 624.6 Arch bridges, the old text was

624.6	Arch bridges
624.61	Timber arch bridges
624.62	Masonry arch bridges
624.621	Natural stone
624.622	Brickwork
624.623	Concrete
624.624	Reinforced concrete
etc.	

The above example was restructured, in the recent revision, into a new text by using 624.011/.014, the auxiliary numbers common to the class 624 dealing with the materials used in civil engineering.

- 624.01 Complete structures according to material and method of construction
- 624.011 Structures in materials of organic origin
  - .1 Wood
  - .2 Wood structures with sub-structures of other material
  - :
- 624.012 Masonry structures in natural and artificial stone and ceramic materials
  - .1 Natural stone
  - .2 Artificial stone
  - .3 Precast concrete
  - .4 Concrete

With the co-ordination of these auxiliaries with 624.6, the new text reads

- 624.6 Arch bridges
  - 624.6.011.1 Timber arch bridges
  - 624.6.012.1 Arch bridges of natural stone
  - 624.6.012.2 Arch bridges of masonry
  - 624.6.012.4 Concrete arch bridges
  - 624.6.012.45 Reinforced concrete arch bridges

:

### 3. Structure of Entries

In general, a schedule entry is composed of "definition phrase" and "example terms," as stated in II. A.

#### a) Definition Phrase

The definition phrase is sometimes a single term but in most cases it is a noun phrase which may be lengthy. At the end of definition phrases, there shall be a punctuation mark such as semi-colon by which users can clearly distinguish the definition phrases from the example terms.

Example :

- 629.4.053.3 Automatic train control, with repeater signal shown in the engine ; ATC.

In the entry, the use of definition phrase alone shall be avoided unless a) it is a very concrete and simple term, or b) it is followed by subdivisions directly below

Example :

- The case of a) 669.13 Cast iron
- 778.6 Colour photography
- The case of b) 332.8 Interest
- .81 Rate of interest
- .82 Usury
- :

But the following entry :

- 629.452 Coaches articulated at a common axle or bogie

shall be changed into :

- 629.452 Coaches articulated at a common axle or bogie ; Articulated trains, Talgo trains. (in TOCS format)

Sometimes there are cases where a definition phrase is given for an entry but for its subdivision entries, only the example terms are given without the definition phrases. In these cases, if the example terms are specific enough and, therefore, no misunderstanding is envisaged, the structure can be adequate.

Example :

The following text is tolerable from TOCS point of view :

- 354 Government departments and ministries.
- 354.11 Ministries of foreign affairs,
- : Foreign office.
- 354.21 Ministry of finance.
- :

But the following is not :

- 354 Government departments. Ministries
- 354.11 Foreign affairs
- :
- 354.21 Finance
- :

It is further desirable to alter this in the way below :

- 354 Government departments and ministries.
- 354 : 327 Ministry of foreign affairs, Foreign office, State department.

354:33 Ministry of finance. Treasury  
: department.

b) Example Terms

The example terms must be as concrete as possible. Each term must be self-supporting, i. e., if this term is separated from the schedule and inserted into the index, the meaning must clearly be shown without the help of other words and phrases. The term must conform to the First Principle of this section.

Example:

Following are inadequate cases:

659.1 Publicity. Advertising and pro-  
paganda  
.01  
:  
.03 Costs. Efficiency. Tariff and  
rates

The words "Costs. Efficiency. ..." in 659.03 should be changed into "Costs of publicity, Efficiency of publicity, etc." to give more specificity.

If it is felt too awkward to repeat the words "of publicity" in the preceding example, it is necessary to have some means to transfer the word "publicity" from the generic entry, 659.1 and add it to the text of 659.1.03 in an acceptable form as "Costs (Publicity)," etc. One of the process is the use of the "supplementary phrase" suggested by the author elsewhere<sup>5)</sup>.

The number of example terms should be as many as possible since all index entries must be kept in the schedule or, if it is computerized, at least keep it in the master tape without worries of the text becoming too lengthy. Even abbreviated or acronym form must be given in the text as a different term.

Example:

621.352 Generation of electrical energy by  
electrochemical methods; Primary  
cells, Primary batteries, Standard  
cells, Weston Cells, Dischargeable  
cells, Polarizable cells, Daniell  
cells, Leclanché cells, Fuel cells.  
(UDC Medium Edition level)

**B. Alphabetical Index in TOCS**

The alphabetical index in TOCS does not differ much from the existing indexes of classification

systems in format. However, there are some principal features different in their preparation.

1. The index entries are taken from the schedule entries except for few exceptional cases. If the addition of a new entries to the index is found necessary, these entries are inverted and must be added into the schedule after examination. The preparation by computer of indexes is expected from the beginning. Some of the basic considerations are given by the author for this kind of processing<sup>8)</sup>.

The exceptional entries can be added, in the computerized case, by the use of display terminal keyboard in the editing process. Also a slight modification of wording in the entry text is possible through the same means, but these should be minimized to save the cost of index preparation.

2. An index may contain entries where the terms are concrete and simple but the classification notations are complicated because of co-ordination.

Example:

Power élite 323.396:301.188.2  
Concentration camp 341.346:348.819

3. Even if an IDE type definition (see II. A.) for an entry is unavoidable, and if there is no subdivision included, then the mentioned constituent specific terms shall be put in the index for the same notation.

If there are subdivisions, as in the case of (2-4), the entry shall not be put in the index by the addition of stopping sign (such as \$ sign).

4. In the index entries conforming to the TOCS principle, the addition of a scope note or comment is not necessary because the entry terms are chosen specifically enough. However, "see" or "see also" references can be added if necessary.

Example:

Existing style	UDC in TOCS
=60 Spanish	=60 Spanish
(Language)	language
(=60) Spanish	(=60) Spanish
(Nationality)	nationality

**C. Schedule Implementing TOCS Principle**

With the discussions of III. A. in mind, this section provides examples of implementation.

Examples are (1) UDC keeping its basic rule as it has been though modified in format to meet the requirement imposed by TOCS, and (2) a new classification system constructed in conformity with the TOCS principle.

The classification schedule is characterized by :

(1) Notation of classification has no restriction (decimal or non-decimal), provided that it is hierarchical and has a faceted structure. In adopting such a structure, an elementary notation (in the case of UDC, a single UDC number<sup>4</sup>) becomes shortened.

(2) In the text, the definition phrase is followed by special sign (e.g., a semi-colon) which is followed by the example terms.

(3) The example terms are made specific, thus no two notations correspond to a single term. Thus the inversion of the schedule into an index is made very simple.

(4) Co-ordinated concepts are limitless, but as far as the corresponding term is concrete and useful, the entry with co-ordinated notation is included in the schedule.

The alphabetical index is characterized in TOCS by :

(1) Index entries are numerous and include all practical terms including abbreviated or acronym form.

(2) Each index entry is represented by a single term which corresponds to a notation, simple or compound.

(3) The classification notations quoted in the index are provided with a special sign if these have subdivisions.

#### D. The Implementation of Classification

The classification system discussed in this paper refers only to UDC. For one thing, UDC is well known and is equipped with a co-ordinated structure at least partly, and for another, this author and his colleagues are actually working on the computerized production of Japanese Medium Edition of UDC, the schedule as well as the index, from a single master file.

As UDC is a system established over many years, there is much difficulty in changing it rapidly. For this reason, the full application of the TOCS principle cannot be made easily.

On the other hand, for a new classification system, the application of the TOCS principle is much easier. The application has been tried by this author and his collaborators in several instances. Two among them will be mentioned here as the results of completed projects.

One of them is the schedule for the "Classification of Science and Technology (CST)" published in Japanese by the Japan Science Foundation in 1976. This Schedule called the REGISTER System is intended to cover the whole field of science and technology in retrieving current research items in scientific and technological institutes in Japan.

The other has similar characteristics but concentrates on the agricultural sciences and their peripheral fields. This classification system is called the Classification of Agricultural Sciences and Technologies (CAST), and forms a part of the Retrieval System for Current Researches in Agricultural Sciences (RECRAS), made for the Ministry of Agriculture and Forestry of Japanese Government by the committee headed by this author.

It is not the purpose of this paper to describe these two systems. For the latter system a full report is expected to come out in 1977.

#### IV. Conclusion

In conclusion, what this author wishes to stress is the basic way of thinking about classification in connection with terms used therein. The computer use has brought an appreciable quantitative change in the labour required for preparing classification systems, and to make a good use of this labour saving, it becomes necessary to have a qualitative change in the structure of classification schedules and the methods of preparation, especially, of the indexes.

The increasing tendency of UDC to adopt faceted principle classification can be combined with this change so as to produce a practical classification system to be used by people who prefer to be based on terms.

With this fact in mind, the application of the TOCS principle to UDC is accounted with

### Terms and Classification Systems

examples. The expected advantages are: the use of post co-ordination in the retrieval, the prompt preparation of the schedule and the index in printed from by computer, as well as an up-to-date UDC database to be consulted through a display unit in the computerized system.

- 1) *Documentation—Guidelines for the establishment and development of monolingual thesauri*, ISO 2788, 1974. p. 6-8.
- 2) *Macrothesaurus—A basic list of economic and social development terms*. Paris. OECD, 1972, p. 129-130.
- 3) *Thesaurus of Engineering and Scientific Terms*, New York, Engineering Joint Council, 1967, 690 p.
- 4) In this paper, the UDC numbers mentioned as examples are in the Medium Edition level unless otherwise stated. Since English Medium Edition is still in preparation, numbers are taken from German Medium Edition (Dezimalklassifikation, DK-Handausgabe, 1967) and the English text is translated from the German text with due consideration of the text in Abridged English Edition (1951) and the recent Full Edition.
- 5) Nakamura, Yukio. "Automatic production of indexes from schedules. The Case of Universal Decimal Classification," *International Classification*, vol. 2, no. 2, 1975, p. 81-88.
- 6) *Ibid.*, p. 84.
- 7) *Ibid.*, p. 84-85.
- 8) *Ibid.*, p. 84-87.