

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

*The Development of a Notational System for a Restructured UDC*

by

*Nancy J. Williamson  
Professor Emeritus  
University of Toronto*

**Introduction**

Since 1992 (McIlwaine and Williamson, 1993, 1994) a pilot project has been under way to examine the feasibility of restructuring the *Universal Decimal Classification* (UDC) into a full-faceted system using the framework of facets provided by the *Bliss Bibliographic Classification*. The research is being carried out on UDC class 61 (Medical Sciences) and Bliss Class H. During the first stage of the process the required elements from the Bliss framework have been identified, insofar as possible the Bliss notation stripped out, and a new database created. UDC topics were then merged into the new database to ensure that there was a place for every UDC topic presently contained in the UDC Machine Readable File (MRF) of the English Medium Edition. The next step is to provide a new notation for the new structure. This paper addresses the task of providing that notation.

**The present situation**

Since both Bliss and UDC are existing schemes both have systems of notation, but neither, as it presently exists, is suitable for use in the restructured Medical Sciences. Bliss has a purely alphabetic notation which is cumbersome to use and is relatively non-mnemonic. (e.g. Arterial aneurysms HUI HUI JK). It is hierarchical only to a very limited degree. For example under Blood vessels (Obstruction) the topic Embolism has the notation HUI KA and types of Embolisms are listed subordinately, each with an extended notation as illustrated in Figure 1 below. The major advantage of the Bliss notation is the base of the 26 letters of the alphabet which serves to minimize the length of the notation as much as possible when a number of facets are combined. Such a notation is also capable of greater

HUI KA	Embolism, thromboembolism
HUI KAP	Air embolism
HUI KAQ	Fat, oil embolism
HUI KAR	Amniotic fluid embolism\
HUI KAS	Foreign body
HUI KAT	(Blood clots) Thrombosis
HUI KAV	Thromboangiitis obliterans

**Figure 1. Example from Bliss Schedules**

hospitality than decimal systems. Long notations are a common characteristic of faceted classification and resulting class numbers are often lengthy. A primary disadvantage of an alphabetical notation such as this is its lack of independence from language, making it less than an ideal choice for a classification system which exists in many languages, some of the them with non-Roman alphabets.

On the other hand, UDC has the advantages of a decimal notation developed hierarchically (for the most part) and is language independent. However, because of the level of division and its highly pre-coordinate nature, UDC has lengthy notation, a problem which is exacerbated by the existence of numerous auxiliary tables which are used in building complex numbers using a colon combination. In the application of the existing schedules the following example is typical:

616.322-002-036.87 Recurrence of tonsillitis

In this example, 616 is Pathology, Clinical medicine; .322 is Tonsilia; -002 denotes inflammation; and -036.87 denotes recurrence. In schedule terms 616.322 is a schedule number, -002 is a special auxiliary subdivision and -036.87 is a general subdivision. The last two segments are derived from tables. Numbers can be expressive and are also more amenable to machine searching than an alphabetical notation such as that used in Bliss. Nevertheless, the base of ten characteristic of a decimal system has its limitations with respect to hierarchical expansion and hospitality for new topics. Moreover, because the project will result in a drastic reorganization of topics in the medical sciences the present allocation of UDC numbers has been

ravaged and cannot be used as it stands. A completely new notation is needed.

### **Basic notational requirements of restructured 61**

Since the Bliss framework is being used, the easiest solution to the notation problem for a restructured UDC Medical Sciences class would have been to use the Bliss notation or something similar. For a number of reasons, early in the project it was decided that the notation should be "UDC-like" and as machine searchable as possible. In addition to its expressiveness, hierarchical characteristics, its mnemonics, and its independence of language, it is a notation which is familiar to UDC users. Moreover it was necessary to recognize that 61 Medical Sciences is just one discipline which in most UDC libraries must be used with part, or all, of UDC as it presently exists. As can be imagined, this latter factor adds to the complexity of the problems of creating the notation.

Careful planning is necessary because the two systems are based on two different basic principles of classification. Bliss is based on facet analysis and UDC is semi-enumerative. Bliss topics are concepts and UDC topics are precoordinate. In using the Bliss framework, concepts should recur a minimal number of times throughout the schedules, and be imported to other locations in the classification as needed for detailed and complex division of subjects. This means that in many cases UDC terms need to be broken down into their concepts in terms of their position in the schedules. This has four major effects on planning and developing the final notation. Before the process can begin it must be clear what concepts are to be included in the schedules and what concepts are to be located in auxiliary tables. Then the notation must first be allocated at the highest level; at succeeding levels of subordination the notation must be assigned by moving systematically from the beginning of the class to the end; and as the work proceeds, careful attention must be given to the effect on the notation of both schedule order and citation order.

### Allocation of notation to major subclasses

The decision to restructure using the Bliss framework, and therefore its schedule order, results in a complete reorganization of the UDC main subclasses of 61. The existing organization of the major subclasses of UDC 61 is as outlined in Figure 2.

61	Medical Sciences
611	Anatomy
612	Physiology, Human comparative physiology.
613	Hygiene generally, Personal health and hygiene
614	Public health and hygiene. Accident prevention
615	Pharmacology. Therapeutics. Toxicology
616	Pathology. Clinical medicine
617	Surgery. Orthopaedics. Ophthalmology
618	Gynaecology. Obstetrics
619	Diseases of domesticated animals. Veterinary medicine

**Figure 2. Class 61 in Existing UDC**

In contrast, the basic schedule order in Bliss is the location of general medicine, followed by general anatomy and physiology together in the first subclass, followed by general health matters related to preventive, curative and clinical medicine, and a general section on diseases. The remainder of the discipline is then ordered by systems of the body. The result is that such topics as pathology, surgery, and physiology are located with the individual "parts, organs and systems of the body". Thus pharmacology is with clinical medicine (general), orthopaedics is with the locomotor and skeletal system, ophthalmology is placed with the nervous system; and gynaecology, pregnancy and obstetrics are with the female reproductive system.

There are at least two alternatives for allocating notation at the subclass level for the restructured class. The allocation being proposed by the principal investigators at the time of writing as exemplified in Figure 3. This proposed allocation is roughly proportional to the

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

amount of material which is expected to be included in the UDC schedule.

61-	Generally of medicine
611	Anatomy. Physiology
612	Cell development
613	Health science
613.1	Preventive medicine
613.3	Curative medicine
613.6	Clinical medicine
614/615	Diseases
616/619	Systems and persons
616	Locomotor system. Cardiovascular system
616.1	Locomotor system
616.3/.9	Cardiovascular system
617	Nervous system. Glandular system
617.1/.7	Nervous system
617.8/.9	Glandular system
618	Respiratory system, Digestive (incl. Dentistry) system. Urogenital system
618.1/.4	Respiratory system
618.5/.7	Digestive system
618.8/.9	Urogenital system
619	Special persons, etc. (incl. Gynaecology and Obstetrics, Pediatrics, Geriatrics, etc.)

**Figure 3. Proposed Restructured UDC**

For example, there is a great deal of material in the section of the schedule on "Diseases". Hence the allocation of two major subclass numbers. Also there is much less material on the locomotor system than there is on the cardiovascular system.

A few topics are under consideration for removal from 61 to other parts of the classification. These include some aspects of management and administration, aspects of personal hygiene (which are not related to health) and veterinary medicine. At the time of writing this proposed allocation is very tentative and adjustments will probably be needed as the work proceeds. This particular provides for maximum hospitality and minimum length of notation. It may have some limitations for machine searching.

An alternative pattern for allocation would follow the above schedule order but would

place all of the systems in 616 is shown in Figure 4.

616	Systems and persons
616.1	Locomotor system
616.2	Cardiovascular system
616.3	Nervous system
616.4	Glandular system
616.5	Respiratory system
616.6	Digestive system
616.7	Urogenital system

**Figure 4. Alternative Proposal for Notating Systems of the Body in UDC**

This alternative might produce a notation which is more amenable to machine searching. However, it has the disadvantages of longer notation, and wasted space at the end of the class. A choice between these alternatives will involve a major policy decision from the UDC Consortium (Goedegebuure 1993) and may be decided on the basis of whether or not shorter notation or ease of machine searching is more important to users. Faceted classification systems tend to require the use of lengthy notation. However, if the major future use of UDC is to be the ordered retrieval of bibliographic records as opposed to the arrangement of physical documents, it is more important that the notation reflect the subject content than that it be brief.

### **Schedules vs. tables**

One of the major differences between Bliss and UDC is the kind of content found in the Bliss auxiliary schedules and the UDC special auxiliary tables. In both systems there are tables to take care of general concepts which apply to any part of the classification scheme. However, the restructured class 61 in UDC will need to accommodate the existing UDC general "auxiliary tables". In addition, both classification systems have special auxiliary tables which apply specifically to the Medical Sciences. However, the list of concepts which are included in those tables is very different in the two schemes.

Like UDC, Bliss has the usual general "auxiliary schedules" for common formats,

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

common subject divisions, persons, places and languages. In Class H "Medical Sciences" tables are kept both small in number and small in size. There are three "auxiliary schedules" that provide for division under treatment, division under a disease or injury and division under a part, organ or system of the body, type of person. Thus each is designed to deal with a precise aspect of medicine. Specifically, "Auxiliary Schedule H1" for division under treatment provides for aspects of medical care (e.g. Primary care, Patient care, nursing, Patients' reactions and Actions on patients). "Auxiliary Schedule H2" provides for general topics related to diseases and injuries (e.g. Preventive medicine and Curative medicine), as well as treatment factors, disease processes (e.g. Swelling, Hardening and Neoplasms) and causes of injuries. Many of the concepts covered in this table are actually frequently recurring facet indicators. "Auxiliary Schedule H3" provides for the major divisions under "Parts, Organs, and Systems" of the body. It is the largest of the three tables and its application results in parallel arrangements of files across the systems. As seen above each of the systems will have a unique basic class number, but all will be expanded using the Auxiliary Schedule H3 for such topics as Anatomy, Pathology, etc. in each case. In general the auxiliary schedules operate at the concept level.

In contrast, in addition to its "Common Auxiliaries" UDC has many "Special Auxiliary" tables embedded throughout the schedules. Class 61 has special auxiliaries in subclasses 611 Anatomy, 612 Physiology, 613 Hygiene generally, 614 Public health, 615 Pharmacology, Therapeutics, Toxicology, 616, Pathology and 617 Surgery, Orthopaedics, Ophthalmology. Some auxiliary lists are very small (three or four topics) while others are very large. For example, the major list of auxiliaries in 616 requires seven and three-quarters double column pages of space in the English Medium Edition. Many of the UDC tables are highly precoordinated and many of the entries are for concepts and multi-concepts which are main schedule concepts in Bliss. For example the auxiliaries in pathology include such topics as "General effect of heat, including: Sunstroke, Heartstroke; Contused wounds; Fatty necrosis; and Tuberculous lesions, Tuberculosis.

As a result of these major differences between Bliss and UDC, prior to applying a new



notation it is essential to determine exactly what belongs in the schedules and what should be moved to the auxiliary tables. In the restructuring process, there has been a concerted effort to minimize the number and complexity of the tables. With this in mind, the auxiliary tables in 61 are being checked against the index of Bliss Class H. Where there are locations for topics in the Bliss schedules, these topics will be removed from the UDC tables.

The remainder of the UDC auxiliaries will be merged into the 3 auxiliary schedules in Bliss. It may be necessary to add an additional table, but findings at the time of writing suggest that many of the topics in the UDC auxiliaries will become schedule topics and additional tables may not be needed. In the process of creating a class number using a faceted system, many of the concepts can be brought from other parts of the schedules using the colon combination. For example:

System : Part : Disease [general facet indicator] :  
Disease of that Part of the Body : Treatment [facet indicator] :  
Specific treatment

## Conclusion

The development of a notation for any classification system is a major problem. Seldom is it possible to satisfy all the desirable requirements and compromises must frequently be made. However, the process becomes even more complex when trying to reconcile differences between systems as different as Bliss and UDC. The discussion presented here reflects the understanding of the problems and possible solution at the outset of the process. Some of these decisions may need to be adjusted as the process moves forward. As soon as it has been determined which concepts are schedule data and which are table data, the process of assigning notation can begin. It is absolutely essential that the actual allocation of notation begin with the tables in place and working from the beginning of the class. As each successive concept is assigned a notation it must be established whether the concept is a schedule or a table concept, whether the topic has previously been established earlier in the schedule and can be brought down using the colon combination rather than being repeated in the schedule; and whether it needs to be treated as a facet indicator. Facet indicators will appear in the schedules but it has not yet been decided whether to assign notations to them in



the schedules, or only in the tables. There must be a consistent order of facets throughout the schedule and their possible use in machine searching will be an important factor.

## References

Goedegebuure, Ben G. (1993) "The UDC Consortium." *Extensions and Corrections to the UDC 1993*. The Hague: UDC Consortium, pp. 7-9.

McIlwaine, I.C. and Williamson, N.J. (1993) "Future Revision of the UDC: Progress Report on a Feasibility Study for Restructuring" *Extensions and Corrections to the UDC, 1993*. The Hague: UDC Consortium, pp. 11-17.

McIlwaine, I.C. and Williamson, N.J. (1994) "A Feasibility Study on the Restructuring of the Universal Decimal Classification into a Fully-Faceted Classification System". *Knowledge Organization and Quality Management: Proceedings of the Third International ISKO Conference, 20-24 June 1994, Copenhagen, Denmark*.

Edited by Hanne Albrechtsen and Susanne Oernager. Frankfurt/Main:

Indeks Verlag, pp. 406-413

