Medicine and the UDC: the process of restructuring Class 61

I. C. McIlwaine
University College London
Nancy Williamson
University of Toronto

ABSTRACT: This paper gives an account of the project to reconstruct and update the UDC schedule 61 - Medicine Class, that has been conducted since 1993. An overview of the work already done in Phase 1 of the project is provided as well as its findings. The authors identify problems still to be addressed and set out the methodology for Phase 2.

KEYWORDS: UDC revision; Medicine

Introduction

Over its long history, the UDC has gone through a number of changes in its management and the location of its headquarters. Its foundations are in the *Dewey Decimal Classification*, edition 5, but over the years it has developed independently from Dewey with different objectives. It is a bibliographic classification designed to be used to organize bibliographic listings, rather than a book classification to organize books on shelves. It inherits partial enumeration from Dewey but with its auxiliary tables and symbols "it has always been a synthetic classification" (Maltby 1975, p. 34) in which "the enumerated classes are the building blocks whereby compound and complex classes may be denoted by various synthetic notational devices." (McIlwaine 2007, p. 9). It was an inspiration for the development of Ranganathan's Colon Classification, a fully faceted scheme based on categories of simple concepts. UDC is not a fully faceted scheme. "A truly faceted classification does not permit the listing in schedules of any combination of terms and states each concept in simple terms. While it is not fully faceted in that sense, the principles of facet analysis are inherent in the structure of UDC." (McIlwaine 2007, p. 33). For it to become fully faceted, it would need to shed its historical connections with Dewey. One UDC class where this has already been achieved is the recently approved Religion Class 2. This project is a further attempt to move in that direction

The purpose of the project is to reconstruct and update the UDC Medicine class 61 as part of an experiment in the ongoing development and updating of the *Universal Decimal Classification* in general. In the late 1980s the management and development of UDC underwent major changes. As part of that process, in 1988 the UDC Management Board recommended the setting up of a limited life Taskforce on UDC System Development to advise on the future development of UDC. It was the recommendation of the Taskforce that a study be carried out to determine the feasibility of converting the classification system into a fully faceted classification. It was decided that the best way to do this would be to actually convert one class.

Medicine (Class 61) was chosen for the experiment for two reasons. First of all, this class was a part of UDC that was most out of date and greatly in need of revision and secondly it presented an opportunity to consider an approach to the organization of the subject which would be more in keeping with modern methods of the study of medicine. Beginning in 1993, this project was started as a direct response to that recommendation. Phase I of the project is now complete and

Phase 2 is underway. This paper describes Phase I and its findings, sets out the methodology for Phase 2, and identifies questions and policies that still need to be addressed.

Intellectual support for the project comes from the work of the Classification Research Group (CRG) in the UK. Following from Ranganathan, the CRG refined his principles and applied them to a number of special subject schemes and to the development of the *Bliss Bibliographic Classification* (BC2). BC2 is fully faceted and applies sound principles. While this system is still in progress, Class H: *anthropology, human biology, health sciences* (Mills and Broughton (1981) has been published and is being used in this project.

Phase I

Phase 1 began with the establishment of the following objectives:

1. To address the question of the feasibility of converting UDC to a fully faceted system;
2. To provide an up-to-date UDC class for medical sciences;
3. To create a table that could be used in conjunction with the existing UDC;
4 To achieve a display of topics commensurate with the modern approach to the study and practice of medicine;
5. To provide a table with a depth of analysis commensurate with the recently published Standard Edition of UDC (2005-2006).

In support of these objectives, the health sciences, biomedical sciences section of Class H (Mills and Broughton, 1981) of the *Bliss Bibliographic Classification* 2nd ed. was chosen as the basis for the framework of the proposed UDC class 61 table. Its suitability for the task was judged to be its sound principles of facet analysis and its emphasis on the grouping of the subtopics of medicine by systems of the body. This approach is currently used in the study and practice of modern medicine. Because of the age of Class H (1981) it was recognized that additional support would be needed from other tools such as Medical Subject Headings (MeSH) and *ICD 10: International Statistical Classification of Diseases and Related Health Problems*.

As the first step in the process, phase 1 was to establish a framework for the class using BC2 Class H and to convert it into a UDC-like system. Class H was not in machine-readable form. The printed document was not of high quality having been produced on a typewriter using elite type font. However, it was scanned to create a machine readable file. The result was a file that could be read but needed considerable rearrangement to return it to its original display.

The next step was to strip out the Bliss notation. Then the captions were edited and UDC notations added. In so far as possible, the captions were made complete in themselves to facilitate machine searching. The facet indicators were used as headings and worked into the captions to created a format commensurate with UDC. Built numbers were displayed as *Example(s) of combination* and contained modified BC2 captions, BC2 facet indicators and class numbers from elsewhere in UDC as needed. Figures 1 and 2 below illustrate the transition.

Figure 1 - **BC2 Format**

```
HUH      Heart, cardiology
           (Disorders by manifestation)
           (Dysfunction)

*For arrhythmia see disorders of
                physiological processes HUH OX
           (Enlargement)
HUH JK       Heart enlargement
HUH  JL      Cardiac oedema
           (Rupture)
HUH JO       Heart rupture
           (Obstruction)
HUH K        Congestive heart failure
               (Paroxysmal dysporoea)  see  Respiratory
                     disorders HWE OXX
           (Neoplasms)
HUH ME     Carcinoid heart disease
```

Figure 1 above illustrates the BC2 format. The facets are shown in parentheses. Figure 2 below shows the translation to UDC format. The BC2 notation has been stripped out and replaced with UDC class numbers and the topics have been arranged and subarranged. Built topics appear as *Example(s) of combination* under main captions. The facets are worked into the examples as parts of built numbers and the remainder is built up from other subclasses in the 61 class and, where appropriate, from other parts of the UDC classification.

Figure 2 - **UDC Format**

```
617.231    Heart disorders by manifestation
             Example(s) of combination:
             617.231:615.2-021.473  Neoplasms (malignant). Carcinoid heart disease
             617.231-044.88         Enlargement of the heart
             617.231-214.11         Heart dysfunction
                                        → 617.233:617.211 Arrhythmia
             617.231-216.12         Cardiac oedema
             617.231-216.22         Heart rupture
             617.231-216.42         Obstruction. Including: Congestive heart failure.
                                    For Paroxysmal dysporoea see 618… Respiratory disorders

             → 617.243:616.72       Congenital disorders of the myocardium

617.232    Heart disorders by cause
             Example(s) of combination:
             617.232:615.3          Heart infections
             617.232:615.4:579.862  Streptococcus haemolytica. Rheumatic heart disease
```

Figure 2 illustrates the handling of facets in four ways – in captions by such phrases as "617.231 – by manifestation" and "617. 232 – by cause"; as class numbers following the colon and taken from 615 *Diseases and Pathology* (e.g. 617.231:615.2 – *Neoplasms (malignant)*); in the numbers (e.g. -044.88) taken from the common auxiliary tables; and from the special auxiliary table of types of diseases in 615 (e.g. -216.12 *oedema*; -216.22 *rupture* and -216.42 *obstruction*). In the

built numbers, the last number represents the facet. The number 579.862 represents the name of the virus taken from 579 – *Microbiology* in UDC and the cross reference "618 …" signifies that the class number was not available when class 617 was prepared. In all, 14 subclasses have been developed in this manner. All have been published, one at a time, as proposals in the annual publication *Extensions and Corrections to the UDC* (E&C) between 1993 and 2007.

Findings from phase 1

The resulting table is not a final scheme ready for use. Rather, it is basically Class H interpreted in the light of UDC and designed to act as a framework for the final product. The result of phase 1 is a flexible and workable system which will accept new built numbers as needed and should not require frequent major revisions. The use of facets works particularly well in the tables on the various body systems (616/619) where many of the diseases and other medical problems are contained under the label *Example(s) of combination*. When new topics emerge, additional examples can easily be created. The tables are rich in terminology. In BC2 the level of analysis is very deep and an effort has been made to preserve the terminology. Where there are long lists of related diseases some have been placed in "including" statements to keep the level of analysis to something close to the UDC Standard Edition. Importantly, as the project has developed it has had some effect on UDC as a whole. In particular, Phase 1 was instrumental in identifying some of the deficiencies in the tables of common auxiliaries and the work on class 61 has provided considerable input to the development of tables 1k -02 *Common auxiliaries of properties* and -04 *Common auxiliaries of relations, processes and operations*. These general auxiliary tables were developed during the course of the project.

Subclass 615 *Diseases and Pathology* is the foundation for the remainder of the scheme (616/619 *Systems of the body*) and one of its most important features. It contains two types of data. The first type is class numbers for general pathology and the base numbers for the general diseases themselves (e.g. 615.2 *cancer*, 615.3 *Infectious and parasitic diseases*; 615.5 *Viral infections*; 615.7 *Genetic, hereditary and congenital diseases*, etc.). Each is expanded in great detail. The second type of data, is a major special auxiliary table listing types of diseases and illnesses (e.g. -216.21 *Deformities, malfunctions*; -3 *Diseases by cause*; -311 *Wounds, injuries, trauma*; -331.5 *Burns*, etc.) These data from the 615 subclass then feed into the major portion of the scheme at 616/619 *Systems of the body*. Here under each system the various organs are grouped with their physiology, symptoms, diseases, and treatment. In most cases these appear as *Example(s) of combination* as illustrated below in Figure 3. The resulting class numbers are long but the system permits detailed analysis and flexible development, since additional topics can be added easily. For the most part this feature works very well. Figure 3 illustrates this application.

Figure 3 - **Examples of combination under subclasses**

619.23	Clinical medicine and the mouth	
	Example(s) of combination:	
	619.23:614.8	Oral surgery
	619.23-11	Symptoms. Including: Oral manifestations
	619.23-111	Syndromes. Including: Burning mouth syndrome
619.231	Pathology and the mouth	
	Example(s) of combination:	
	619.231:615.211	Oral neoplasms. Oral leukoplakia
	619.231:615.3/.6	Infection of the mouth. Oral sepsis. Fusobacterium, fusiform. Including: Necrotizing ulcerative gingivitis, ulcerative stomatitis, Vincent's infection, trench mouth
	619.231:615.321	Mycobacterium tuberculosis. Oral tuberculosi
	619.231:615:4:579.867	Streptococcus of the mouth. Ludwig's angina
	619:231:615.72	Congenital abnormalities of the mouth. macrostomia. Microstomia
	619.231-216.271	Cysts of the mouth. Ranula
	619.231-216.413	Dental fistula. Oroantal fistula

Figure 3 illustrates a selection of topics from the subclass Clinical Medicine in the Digestive system. It illustrates the effect of schedule order caused by the notation and its symbols. This section could have been handled differently. All of the *Example(s) of combination* could have been listed under 619.23 and the class number 619.231 eliminated. However, this would cause a considerable rearrangement in the filing order. In that case, 619.23:614.8 would file at the beginning of the list shown at 614.231. 619.23-11 *Symptoms* and 618.23-111 *Syndromes* would file near the end of the diseases between the present 619.231:615.72 *Congenital abnormalities* and 619.231-216.271 *Cysts*. In developing the framework an effort was made to keep related concepts together in so far possible. Arguably "symptoms" and "syndromes" should precede "surgery." In such a short list the above solution appears to be tolerable. However in nearly all cases the longest lists of examples have to do with "diseases and pathology." Also, it is in this group that most of the additions and changes can be expected to occur. Thus, as a general policy, a decision was made to make a separate grouping for pathology in most cases while the other facets could be listed together and ordered as the notation requires. This is not a perfect solution but the one that seemed to make the most sense.

Not unexpectedly, the notation needs further attention. Nevertheless, the first phase has resulted in a workable base from which to finalize the 61 table. Yet there is still much work to be done.

Phase 2

The objective of phase 2 is to bring the proposed Class 61 to a workable conclusion. In doing so there a number of tasks to be accomplished. Among these are:

1. Reconsideration of the over all content;
2. Assessment of the impact on UDC in general;
3. Firming up of the policies used in development and maintenance of the subclass;
4. Completion and updating of the data to make it commensurate with the Standard Edition of UDC;
5. Rationalization of the relationships between the new medicine class, the existing 61 and the rest of UDC;
6. Reconsideration of the notation.

Phase 2 is already under way. As phase 2 proceeds, two fundamental questions are being addressed: What exactly should be contained in the medicine class? What implications does content have for the rest of UDC? The circumstances under which UDC 61 and BC2 Class H were created are different. With respect to UDC, class 61 is a part of an in integrated system; whereas in BC2 each class has been developed to be complete in itself and to stand on its own. Thus BC2 Class H contains topics which are dealt with elsewhere in UDC. For example, Class H has a substantial section on biochemistry. How much of this, if any, should be retained under medicine, or be completely covered in 57 with biochemistry? Some subjects are split. For example, the embryology of plants is in 581.3, of animals is in 591.3 and human embryology in the Standard Edition of UDC is in 611.013. Should all embryology be together in the 500s and the number pulled forward as needed? In a faceted classification, it would seem so, but it may not be practical. In Class H, numerous cases can be found of extraneous topics like car parks, furniture, etc. and it is heavily biased towards the UK Health Service – all reflecting the fact that it used to be used in the library of the Ministry of Health. The result will be the removal of some topics from medicine and major revisions to other areas of UDC.

Another aspect of content is that some data have been included in Class H as support data. For example, under "properties of anatomy and physiology and under "time factors in anatomy and physiology" in Class H there are lists of such terms as: strong, weak, continuity, periodicity, duration, etc. With the development of new common auxiliary tables -02 and -04 these terms can be drawn from those tables and do not need to be repeated as a list in Class 61. Subclass 612 has similar lists under "properties of medical materials, types of medical material and equipment." A brief preliminary survey of subclasses 611 to 615 suggests that the contents of 611 Anatomy and physiology and 612 Health sciences. Medical sciences are the most likely candidates for removing unnecessary material. Subclass 615 Diseases and pathology is more problematic and could grow with examination of ICD 10 and the discovery of new diseases not found in BC2 or the present UDC. For example SARS which has only recently been identified is now listed in ICD 10.

Much of the material to be moved out of the proposed Class 61 has been taken care of in general auxiliary tables -02 and -04. However, there are also special auxiliary tables. Whereas in the current UDC 61 there are a number of small auxiliary tables, there are only two tables in the proposal: 615-1/-3 in Diseases and pathology and 616-5 for regions and parts of the body. Are there other subtopics which should be dealt with as special auxiliary tables? Medical equipment has been suggested as one candidate? Should there be one special auxiliary table in Class 68 listing all equipment to be used across the scheme? The choice of terms for inclusion in a general auxiliary table may take some time. The answer to the question: what is equipment? may not be as easy as it sounds. Terms such as 'tables', 'forceps' and 'beds' are obvious candidates, but what about a term such as 'pumps'? A search of *UDC online* locates 86 hits on the word "pumps." Included are 6 class numbers for pumps in 6l and there are many types of pumps named in subclass 621.

The process of updating requires several kinds of editing - internal updating of the new tables updating with respect to the Standard Edition; updating of the diseases; and making sure that all the topics in the present Class 61 are present or accommodated for in the proposed tables. Updating of the internal tables is required because during phase 1 the various subclasses were created in serial order. Once a subclass was published, we did not return to revise it as the ongoing revisions were made in UDC. Some class and table numbers chosen early in the project were superseded by ongoing revisions to the existing tables and the development of the new common auxiliary tables -02 and -04. Hence, there is an unevenness of application and numerous blanks exist where appropriate class numbers were not available until part way through the project. There are numerous blanks and incomplete cross references to other parts of the proposed 61 which

were developed later in the project. In addition, while the proposal was being developed, the rest of UDC was in continuous revision. Moreover the topics contained in the proposal are based primarily on BC2 (with some reference to ICD10). Albeit, BC2 has more detail than the current 61, but are all the topics in the current 61 covered? What are the differences in terminology? For example, BC2 Class H uses the term "health professionals" and never mentions the term "doctors" which is the term used in UDC. In summary, the proposed medicine class needs to be brought into line with the latest Standard Edition. Captions must be searchable. For example, if a user were to wish to retrieve all occurrences of cancer the term neoplasms need to be repeated at every location even though this has not always happened in BC2.

The updating problem is even more complex, in that the present Class 61 is very outdated – a reason for selecting it for this project. Also BC2 Class H was published in 1981. Much has happened in medicine and the identification and treatment of diseases in the intervening 27 years. For example, topics such as "visualization and imaging" and "genetic engineering" need further investigation. One on the most important tasks in Phase 2 will be the updating of proposed subclass 615 Diseases and pathology. As indicated above, subclass 615 lists the major diseases and is fundamental to the development of the subclasses on the various body systems (616/619). General diseases listed here are taken down to the various systems as they apply to diseases and illnesses identified as attacking different parts of the body. Cancer (neoplasms), for example, has a number 615.2 and this number is used under each body system together with the name of the type of cancer appropriate to that part of the body.

Neoplasms of the heart are in "617.231:615.2-021.473 *Neoplasms (malignant). Carcinold heart disease.*" Lung cancer appears as "618.531:615.2 *Lung neoplasms. Including: Bronchiolar carcinoma. Bronchogenic carcinoma. Pancoast syndrome.*" Other tools are needed to aid in this segment of the revision. The World Health Organization's *International Statistical Classification of Diseases and Related Health Problems* supplemented by the National Library of Medicine's *Medical Subject Headings* will be used. The WHO publication will be most helpful as it is in classified order and arranged by systems of the body. Both are up-to-date and available online. Also particularly useful will be the sections of the print edition of MeSH which list headings added annually to the system from 1981 to 2003.

Another aspect of the revision must be a review of the use of special auxiliary tables. In the present Class 61 there are numerous special auxiliary tables. In the proposal these have been kept to a minimum in line with the current general policy on UDC revision. There are only two such auxiliary tables in the proposal, -1/-3 in 615 for diseases and -5 in 616 for regions and parts of the body. Both are usable throughout class 61. Should there be more? Are there other topics, in addition to "medical equipment," that should be dealt with in this way?

As a major part of the revision process there has to be a matching between the proposed Class 61 and the whole of UDC to answer such questions as: What topics in the existing UDC are missing from the new version? What topics related to medicine in the proposed new class 61 are located elsewhere in UDC? What topics in other classes in UDC need to be revised or added to? This requires considerable cross checking among Class H, the two versions of class 61, and the whole of UDC to determine how best to handle them.

Finally there are concerns about the notation which must be addressed. At the very least, the notation needs to be adjusted. At the outset it was impossible to determine precisely how much space would be needed to handle each of the subclasses. They are unequal in content and, not unexpectedly, there is some crowding at the end of the class resulting in very long class

numbers. While it can be anticipated that some sections may be moved elsewhere freeing some space, it may not be sufficient and there is need to allow for the future. As far as the class as a whole is concerned, much additional growth may be due to the addition of new diseases. Some of this will be absorbed in the examples and may not require new class numbers but this cannot be counted on as a certainty. One suggestion is that medicine be moved to the vacant class 4, thus providing 100 number base instead of a base of 10. Long class numbers are a problem in some cases and some people feel that a full 100 number base would contribute to shorter class numbers. This may be partly so, but the depth of analysis also has bearing on the length of class numbers. While the intention is to equate Class 61 with the UDC Standard Edition in its final form, class 61 will still be very detailed.

Conclusion

Phase 1 has resulted in a usable framework based on sound principles which should make the next steps easier to accomplish. While there is still much detailed work to be done, all indications are that the move forward to the final result should be accomplished much more quickly than Phase 1, because there is now a base from which to work. Once the work is completed, it still has to be submitted to a panel of experts for review and analysis.

References

Extensions and corrections to the UDC (1996-2007). The Hague: UDC Consortium.

ICD-10: International statistical classification of diseases and related health problems (2007). Geneva: World Health Organization. Also available at: http://www.who.int/classifications/apps/icd/icd10onlineonline/navi.htm.

McIlwaine, I.C. (2007) *The Universal decimal classification: a guide to its use*. Rev. ed. The Hague: UDC Consortium. pp. 9, 33.

Maltby, A. (1975) *Sayers' manual of classification for librarians*. 5th ed. Andre Deutch, p.34, 37-38.

Mills, J.; Broughton, V. (1981) *Class H: anthropology, human biology, healthsciences*. London: Butterworths.

Medical subject headings (1981-2003). Washington: National Library of Medicine. Also available at: http://www.nlm.nih.gov/mesh/MBrowser.html.

Ranganathan, S.R. (1933) *Colon classification*. Madras: Madras Library Association. London: E. Goldston.

UDC: Universal decimal classification (2005-2006). 3rd ed. Standard edition. London: British Standards Institution. Also available at: http://www.udconline/net/.