VZG Project Colibri/DDC

DDC Number Logicology

August 2017

Ulrike Reiner

VZG Colibri/DDC Report 1/2017

Name of the document

colibri07.doc

photographer: Ulrike Reiner (Fundy National Park, New Brunswick, Canada, May 25, 2008)

Versions

colibri07-17-08-02.doc   First draft
colibri07-17-08-17.doc   Continuation
colibri07-17-08-22.doc   Final version

Contents

> "Logic takes care of itself;
> all we have to do is to look
> and see how it does it."
> *Ludwig Wittgenstein*

1. Introduction

The VZG-Project Colibri/DDC deals with automatic procedures for the Dewey Decimal Classification (DDC). Since 2003 this project deals with following research questions:

Colibri/DDC Research Questions

Q1. Is it possible to classify automatically GVK PLUS bibliographic title records using DDC?

Q2. Is it possible to decompose automatically *molecular DDC numbers* into atomic DDC numbers?

Q3. Is it possible to improve automatic classification and retrieval by means of *atomic DDC numbers?*

Colibri/DDC Definitions

D1. An *atomic DDC number* is a semantically indecomposable string (of symbols) that represents a DDC class.

D2. A *molecular DDC number* is a string that is syntactically decomposable into *atomic DDC numbers.*

In this report syntactical features (first-order predicates[1]) of DDC numbers are considered. Some work has been done in 2004: "Ulrike Reiner: DDC-Notationsanalyse und –synthese, VZG Colibri/DDC Report 2/2004, Göttingen[2].

---

[1] https://en.wikipedia.org/wiki/First-order_logic (retrieved on 17-08-22)
[2] http://taipan.dyndns.org/~ul/colibri03-05-04-11.pdf

2. Preliminaries

Following abbreviations are used:

bano   base number:

1. "Number building begins with a base number (always stated in the instruction note) to which another number is added."[3]

2. "A number of any length to which other numbers are appended"[4] (extended meaning of bano 1.)

dno   DDC number (notation);

dnos: DDC numbers
For all abbreviations their plural is built analog, i. d., **banos**, **dno_abbrs**, **dno_atoms**, …

dno_abbr   dno of the abridged edition. The abridged edition is "A shortened version of the Dewey Decimal Classification (DDC) system that is a logical truncation of the notational and structural hierarchy of the corresponding full edition on which it is based. The abridged edition is intended for general collections of 20,000 titles or less."[5]

dno_atom   dno atomic -> definition D1

dno_comp   dno complete -> **dno_full**

dno_corr   syntactically correct (built) dno

dno_bui   dno built: "A number constructed according to add instructions stated or implied in the schedules or tables."[6]

dno_div   2nd level dno: **"**The second level of subdivision in the DDC, represented by the first two digits in the notation, e.g., 64 in 640 Home and family management."[7]

---

[3] https://www.oclc.org/en/dewey/features/summaries.html (retrieved on 17-08-22)
[4] http://www.oclc.org/content/dam/oclc/webdewey/help/glossary1.pdf (retrieved on 17-08-22)
[5] l.c. (loco citato)
[6] l.c.
[7] l.c.

dno_full   dno of the full edition. The full edition is "The complete version of the Dewey Decimal Classification (DDC) system"[8]

dno_main   1st level dno: "One of the ten major subdivisions of the DDC, represented by the first digit in the notation, e.g., 3 in 300 Social sciences."[9]

dno_mol   dno molecular -> definition D2

dno_opt   optional dno: "*Optional number*: (1) A number listed in parentheses in the schedules or tables that is an alternative to the standard notation. (2) A number constructed by following an option. *Option*: An alternative to standard notation provided in the schedules and tables to give emphasis to an aspect in a library's collection not given preferred treatment in the standard notation. In some cases, an option may provide shorter notation for the aspect."[10]

dno_part   part of a **dno_full** (**dno_comp**, **dno_mol**), e .g., **dno_atom**, **dno_short**.

dno_sec   3rd level dno: "The third level of subdivision in the DDC, represented by the first three digits in the notation, e.g., 641 in 641 Food and drink."[11]

dno_seg   dno containing segmentation marks. "The indication of logical breaks in a number by a typographical device, e.g., slash marks or prime marks. Segmentation marks indicate the end of an abridged number."[12]

dno_short   intellectually selected/created → **schedno** with a certain (intended) level of specificity. Its length varies between the length of **dno_sec** (min) and the length of **dno_full** (max); for the DDC class "610 Medicine and health" there exist 137 **dnos dno_short** for medical dissertations, the length lies between 3 and 9 (including the Dewey dot)[13]

dno_sub   dno subdivision "A subordinate member of a class, e. g., 519 Probabilities and applied mathematics is a subdivision of class 510 Mathematics, and 519.3 Game theory is a subdivision of 519."[14] . <u>Remark</u>: in the project Colibri/DDC any four- and more digit **dno** is defined as a **dno_sub**.

---

[8] l.c.

[9] l.c.

[10] l.c.

[11] l.c.

[12] l.c.

[13] http://www.dnb.de/Subsites/ddcdeutsch/SharedDocs/Downloads/DE/anwendung/ddcGliederungMedizin.html (retrieved on 17-08-22)

[14] l.c.

dno_sum **dno** is element of the DDC Summaries:

"The _first summary_ contains the ten main classes. The first digit in each three-digit number represents the main class. For example, 600 represents technology.

The _second summary_ contains the hundred divisions. The second digit in each three-digit number indicates the division. For example, 600 is used for general works on technology, 610 for medicine and health, 620 for engineering and 630 for agriculture.

The _third summary_ contains the thousand sections. The third digit in each three-digit number indicates the section. Thus, 610 is used for general works on medicine and health, 611 for human anatomy, 612 for human physiology and 613 for personal health and safety."[15]

schedno schedule number (notation)

schedno_add add table of schedno, e. g., "617:05 Preventive measures and surgery"

tabno table number (notation)

tabno_add add table of tabno, e. g., "T4--8642-T4--8649:024 Readers for people in specific occupations"

3. DDC Number Logicology

DDC numbers can be characterized via multiple non-exclusive one-digit predicates. This is shown in table 3.1 (dnos characterized by first-order one-digit predicates). The first-order predicates can be true ("×" marked in table 3.1) or false and are defined as follows (two-character abbreviation in the table due to lack of space):

```
 ba(dno):   dno is a bano 1.
 ab(dno):   dno is a dno_abbr
at(dno):    dno is a dno_atom
bu(dno):    dno is a dno_bui
co(dno):    dno is a dno_corr
di(dno):    dno is a dno_div
fu(dno):    dno is a dno_full
```

---

[15] https://www.oclc.org/en/dewey/features/summaries.html (retrieved on 17-08-22)

ma(dno):   dno is a dno_main
mo(dno):   dno is a dno_mol
op(dno):   dno is a dno_opt
sc(dno):   dno is a dno_sec
sg(dno):   dno is a dno_seg
sh(dno):   dno is a dno_short
su(dno):   dno is a dno_sub

In the following, the output (*vc_daygram*) of four analyzed DDC numbers via the program *vc_day*[16] is shown. These **dnos** "371.1024", "618.92238", "371.102409481", and "617.605" respectively are characterized in table 3.1 by the above defined predicates:

```
371.1024 <ul97_to_analyze; length: 8>
3------- Social sciences <dno_main>
37------ Education <dno_div>
371----- Schools and their activities; special education <dno_sec>
371.1--- Schools and their activities <dno_sub_span:371.1-371.8>
371.1--- Teachers and teaching, and related activities <dno_sub>
371.102- Teaching <dno_sub>
371.1024 Classroom management <dno_sub>


618.92238 <ul98_to_analyze; length: 9>
6-------- Technology <dno_main>
61------- Medicine & health <dno_div>
618------ Gynecology, obstetrics, pediatrics, geriatrics <dno_sec>
618.9---- Pediatrics and geriatrics <dno_sub>

618.92--- Pediatrics <dno_sub>
618.922-- Specific diseases <dno_sub_span:618.921-618.929>
618.922-- Respiratory tract diseases--pediatrics, . . . <dno_bui>
618.922-- Respiratory tract diseases--pediatrics <RI_bui>

618.92238 Asthma--pediatrics <RI_bui>
618.92238 Asthma--pediatrics, . . . <dno_bui>
---.--2-- Diseases of respiratory system <p1_2_7->dno_sub:616.2>
---.--2-- Specific diseases <p1_2_7->dno_sub_span:616.2>
---.--2-- Severe acute respiratory syndrome--medicine <p1_2_7->RI:616.2>
---.--23- *Diseases of trachea and bronchi <p1_2_7->dno_sub:616.23>
---.--23- Tracheitis--medicine <p1_2_7->RI:616.23>
---.--238 *Asthma <p1_2_7->dno_sub:616.238>
---.--238 Bronchial asthma--medicine <p1_2_7->RI:616.238>

ul98[1]: 765 0#  $b618.92 $a618.921 $c618.929 $r616 $s238 $u618.92238
```

---

[16] http://edug.pansoft.de/tiki-download_file.php?fileId=144 (retrieved on 17-08-22)

```
371.102409481 <ul99_to_analyze; length: 13>
3------------ Social sciences <dno_main>
37----------- Education <dno_div>
371---------- Schools and their activities; special education <dno_sec>
371.1-------- Schools and their activities <dno_sub_span:371.1-371.8>
371.1-------- Teachers and teaching, and related activities <dno_sub>
371.102------ Teaching <dno_sub>
371.1024----- Classroom management <dno_sub>
---.----0---- Table 1. Standard Subdivisions <tabno:T1--0>
---.----09--- History, geographic treatment, biography <tabno:T1--09>
---.----09--- Regional treatment <RI:T1--09>


---.----094-- Specific continents, countries, localities;
              extraterrestrial worlds <tabno_span:T1--093-T1--099:094>
---.------4-- Europe <p20_5->tabno:T2--4>
---.------4-- Modern world; extraterrestrial worlds
              <p20_5->tabno_span:T2--4-T2--9>
---.------4-- Modern world <p20_5->RI:T2--4>
---.------48- Scandinavia and Finland <p20_5->tabno:T2--48>
---.------48- North Calotte <p20_5->RI:T2--48>
---.------481 Norway <p20_5->tabno:T2--481>

ul99[2]: 765 0#  $b371.102409 $z1 $a093 $c099 $z2 $s481 $u371.102409481
ul99[1]: 765 0#  $b371.1024 $z1 $s09 $u371.102409



617.605 <ul100_to_analyze; length: 7>
6------ Technology <dno_main>
61----- Medicine & health <dno_div>
617---- Surgery, regional medicine, dentistry, ophthalmology, otology,
        audiology <dno_sec>
617.6-- *Dentistry <dno_sub>
617.605 Surgical complications--dentistry <RI_bui>
617.605 Surgery,. . . <dno_bui>
---.-05 Preventive measures, surgery, therapy, pathology,
        psychosomatic medicine <p9->dno_sec_1:617:05>

ul100[1]: 765 0#  $b617.6 $a617.6 $a617 $w617 $y1 $t05 $u617.605
```

dno	ba	ab	at	bu	co	di	fu	ma	mo	op	sc	sg	sh	su
3			X		X		X	X						
37			X		X	X	X		X					
371			X		X		X		X		X		?	
371.1			X		X		X		X				?	X
371.10													?	X
371.102		X	X		X		X		X				?	X
371.1024[17]			X		X		X		X			X	?	X
371.10240														
371.102409			X		X		X		X					X
371.1024094			X		X		X		X					X
371.10240948			X		X		X		X					X
371.102409481[18]			X		X		X		X					X
6			X		X		X	X						
61			X		X	X	X		X					
618			X		X		X		X		X		?	
618.9			X		X		X		X				?	X
618.92	X	X	X		X		X		X				?	X
618.922			X	X	X		X		X			X	?	X
618.9223														
618.92238[19]			X	X	X		X		X			X		X
6			X		X		X	X						
61			X		X	X	X		X					
617			X		X		X		X		X			
617.6	X	X	X		X		X		X					X
617.60														
617.605[20]			X	X	X		X		X			X	X	X
19M[21]			X		X				X	X	X			
dno	ba	ab	at	bu	co	di	fu	ma	mo	op	sc	sg	sh	su

Table 3.1 dnos characterized by first-order one-digit predicates

"?" means "unknown", it depends on the definition of the shortening institution. As can be seen from the Table 3.1 the **dno** "617.605" is simultaneously a **dno_atom**, **dno_bui**, **dno_corr**, **dno_full**, **dno_mol**, **dno_seg**, **dno_short,** and **dno_sub**.

---

[17] ul_97

[18] ul_99

[19] ul_98

[20] "Chirurgie" http://www.dnb.de/Subsites/ddcdeutsch/SharedDocs/Downloads/DE/anwendung/ddcGliederungMedizin.html

[21] „Philosophie of Mexico" http://dewey.org/webdewey/index_11.html?recordId=ddc:190, in Notes

The defined predicates relate to each other, some of the (logic) relations are:

dno ∈ {**banos, dnos**, **dno_abbrs**, **dno_atoms**, **dno_comps**, **dno_buis**, **dno_divs**, **dno_fulls**, **dno_mains**, **dno_mols**, **dno_opts**, **dno_parts**, **dno_secs**, **dno_segs**, **dno_shorts**, **dno_subs**, **dno_sums**, **schednos**, **schedno_adds**, **tabnos**, **tabno_adds**}

dno_sum ∈ {**dno_mains**} ∪ {**dno_divs**} ∪ {**dno_secs**}

dno_abbr ∈ {**dno_abbrs**}

{**dno_abbrs**} ⊂ {**dno_fulls**}

dno_seg ∈ {**dno_abbrs**} ∪ {**dno_fulls**}

dno_bui ∈ {schednos} ∪ {schedno_adds} ∪ {tabnos} ∪ {tabno_adds} ∪ {still to be built dnos }

dno_short ∈ {**banos 1., banos 2.**} ∪ {**dno_abbrs**} ∪ {**dno_atoms**} ∪ {**dno_buis**} ∪ {**dno_fulls**} ∪ {**dno_mols**} ∪ {**dno_sums?**} ∪ {**dno_segs**} ∪ {**dno_subs**}

4. Conclusion

In this paper DDC numbers are characterized by features (predicates). A "full DDC number" is defined as a dno of the full DDC edition. The DDC full edition is "The complete version of the Dewey Decimal Classification (DDC) system" (cf. Section 2. Preliminaries). A "full DDC number" is not defined as a number that has not been shortened in any way because one cannot see by looking at a DDC number whether it is full or not, i. e. if it has not been shortened or has been shortened[22]. For instance, the DDC number "617.605" could be "full" ("complete") for the title "Principles and practice of laser dentistry"[23] or could have been shortened from "617.6050028" to "617.605". The title "Lasers in entistry: revolution of dental treatment in the new millennium; proceedings of the 8th International Congress on Lasers in Dentistry"[24] is assigned with the presumably full ("complete") DDC number "617.6050028" and not with the shorter DDC number "617.605" contained also in the full DDC edition. Therefore, we consider DDC numbers not as exclusively "full" or "short" – they can be both (a shortened number of the full DDC edition). Additionally, DDC numbers can have a lot more features (one-digit predicates) as has been shown for the DDC number "617.605".

---

[22] for the ease of use or by definition or as a result by an automatic assignment
[23] https://gso.gbv.de/DB=2.1/PPNSET?PPN=82592653X
[24] https://gso.gbv.de/DB=2.1/PPNSET?PPN=368456137