Rebecca Green (OCLC, Washington, DC)
Michael Panzer (OCLC, Dublin, Ohio)

The ontological character of classes in the Dewey Decimal Classification

Abstract
Classes in the Dewey Decimal Classification (DDC) system function as neighborhoods around focal topics in captions and notes. Topical neighborhoods are generated through specialization and instantiation, complex topic synthesis, index terms and mapped headings, hierarchical force, rules for choosing between numbers, development of the DDC over time, and use of the system in classifying resources. Implications of representation using a formal knowledge representation language are explored.

1: Introduction
At the heart of any classification scheme is a set of relationships between topics (broadly construed) and classes. The foundational assumptions of a scheme cannot be fully comprehended without first understanding the character of its classes as mediated by these relationships. The ontological character of classes in the Dewey Decimal Classification (DDC) system will be examined as a case study to illustrate this point.

Classification schemes can be viewed as directed graphs, with classes as nodes and relationships between classes as edges. If the relationship between topics and classes is ignored, these nodes function as points and are without extent. But if the relationship between topics and classes is taken into account – as it must – the classes of a knowledge organization system (KOS) can each be seen as a semantic space defined by associated topic neighborhoods. While classes are nodes within an overall system graph, each class is also internally structured as a graph.

Section 2 of the paper investigates the nature of topical neighborhoods associated with classes in the DDC. The character of these internally structured classes has wide implications for the ontological representation of the system and for operations on that representation (e.g., abridgment, derivation of subject-specific classification systems). These implications will be explored in section 3.

2: DDC classes as neighborhoods
In the DDC, class-topic and class-class relationships are seen as together forming a category description of a class. Our focus in this paper is on the structure of that description, specifically on the configuration of topics associated with a class; as a shorthand, we will speak of topics being in the class. It is not simply the case that a class-node contains a set of points, each of which represents a topic. Instead, we find a set of focal topics, each of which may motivate the inclusion of other topics. These focal topics may each expand to become the hub of a topical neighborhood, where a neighborhood consists of all topics (nodes) related (adjacent) to a specific node and any relationships among those topics (nodes). The entire configuration corresponds to the neighborhood of the set of focal topics, that is, the union of the individual neighborhoods for each focal topic. To fill out our picture, we need to investigate (1) how topics are associated with classes and then (2) how focal topics are extended, so that related topics are also associated with classes.
2.1: Association of topics with classes

In the DDC topics are associated with classes primarily through captions and notes (especially, but not limited to, including and class-here notes). Consider, for example, 782.292 Chant, shown below with its relevant upward hierarchy. (Figure 1 shows the class network around 782.292; figure 2 shows the topical configuration within 782.292, capturing points from both this and the upcoming section.) The caption indicates that works on chant are classed in this number (if the caption had expressed more than one topic, further analysis would be required to determine if the topics mentioned should be classed in the immediate number or in a subclass thereof). The including note specifies that works on responses, of which litanies and suffrages are examples, are to be classed in 782.292, while the class-here note indicates that works on plainsong (that is, monophonic chant) are also classed here.

Two class-elsewhere notes round out the entry. These notes indicate that, although works on chant in general are classed in 782.292, works on Gregorian chant (a type of plainsong) and Anglican chant (which, in contrast to plainsong, is polyphonic) are classed elsewhere, in classes for liturgy and ritual of specific Christian denominations. This displacement reflects the multifaceted nature of the topics, which combine liturgical form (chant) and association with a specific denomination (Roman Catholic, Anglican). Throughout the 780s, works on aspects of two or more of its subdivisions are to be classed in the number coming last; the class-elsewhere notes here make the general principle explicit. In this case, the class-elsewhere notes are subtractive in nature, indicating that the semantic space at 782.292 excludes two subtopics of the topic named by the caption. See references are always subtractive in nature. For example, the see reference note at 782.295 Biblical texts indicates that psalms, which qualify unequivocally as Biblical texts, are classed elsewhere, in 782.298 Psalms.

How do excluded topics fit into the neighborhood model, when the graph of a class is a representation of topics in the class and relationships among them? One possible strategy for handling them is to make all inclusions explicit, with exclusions being only implicit (that is, any topic not explicitly included is excluded). Another strategy for excluding topics is to include their negations, e.g., not-Gregorian-chant. This strategy is both easier to implement (it would be difficult in many cases to enumerate all inclusions) and a more faithful representation of the classification system.
Figure 1. Segment of class network around 782.29
Figure 2. KOS-based neighborhood of 782.92

[Printed music]

Treatises about

Sound recordings of

Modified standard subdivisions

General principles

Musical forms

Modified standard subdivisions

General principles

Musical forms

Figure 2. KOS-based neighborhood of 782.92

[Chant specific to Liturgy of a Christian Denomination]

Gregorian chant

Anglican chant

Chant

Plainsong

e.g., Salve Regina

Responses

Litanies

Suffrages

[Chant specific to Liturgy of a Christian Denomination]
2.2: Development of neighborhoods

How do neighborhoods develop around focal topics? Two mechanisms, specialization and instantiation, are so fundamental they are often taken for granted. Thus, “chant” in the 782.292 caption expands to include by implication all types of chant (e.g., plainsong) and all examples of chant (e.g., Salve Regina).

Although including notes and class-here notes both associate topics with a class, in other ways they differ. Topics in class-here notes are said to approximate the whole of the class. Standard subdivisions can be added for such topics; they can also participate in other number-building techniques. Topics in including notes are less extensive in scope than the full meaning of the class and are said to be in standing room; standing room topics generate subclasses when literary warrant justifies expansion. Standard subdivisions are not added for topics in standing room, nor do such topics participate in other number-building techniques. The add instruction referred to under 782.292 (“Add as instructed under 782.1–782.4”) can therefore be applied to plainsong, which approximates the whole of the class, but not to litanies, which are in standing room. Hence, notation of plainsong is classed in 782.2920148 and performance of plainsong in 782.292078, but notation of litanies and performance of litanies are both classed in 782.292. The neighborhood of a standing room topic includes complex topics that would go in other class numbers generated by the addition of standard subdivisions or other number-building techniques were the topic to approximate the whole of the class.

Index terms are another potential source for expansion around a focal topic, including both the Relative Index terms assigned to a class and any headings from other sources (e.g., LCSH, MeSH) mapped to the class. As Relative Index terms should be assigned for all indexable topics in the class, it is frequently the case that the focal topics drawn from the captions and notes of a class and the topics named by its Relative Index terms are in one-to-one correspondence. However, Relative Index headings may be added for implied topics; index terms may also name topics that are broader or narrower than the topics mentioned by the caption and by including and class-here notes or that overlap with those topics. For example, there are three focal topics for 782.295 Biblical texts: Biblical texts, amens, and canticles. Each is matched by a Relative Index heading. A fourth Relative Index term, Lord’s Prayer – music, names a specific Biblical text, a more narrow topic. Mapped headings can amplify the topic space of a class. For example, two LC subject headings have been mapped to 782.3222 Gregorian chant, which has only the one focal topic named by the caption: they are Ambrosian chants (a distinct Roman Catholic chant tradition) and Prosulas (Music) (a kind of text created for Gregorian chant). The topics added to the neighborhood by the mapped headings include a sibling topic and a topic of a different semantic type (text), whose relationship with the focal topic is built into the basic vocal music concept.

Up to this point, we have ignored the effect of hierarchical force, which plays an important role in developing neighborhoods around focal topics. The fundamental structure of DDC classes is defined by the notational hierarchy, in which most classes are a specification of the class whose notation is one digit shorter (exceptions to this principle are handled by special kinds of headings, notes, and entries). The nature of this specification varies, including subsumption, partonomy, and complex topic synthesis. Additionally, see references provide for polyhierarchy in the system by defining hierarchical relationships in addition to those defined by the notational hierarchy; many
see references lead from classes where topics are given interdisciplinary and/or comprehensive treatment to classes where specific aspects of the topic are treated or where the topic is treated in another discipline. The full set of these hierarchical relationships constitutes the structural hierarchy of the system.

The principle of hierarchical force in the DDC applies specific attributes of a class to its subclasses in the structural hierarchy. Most notes that associate topics with classes have hierarchical force. Specifically, definition notes; scope notes; former-heading, variant-name, and former-name notes; class-here notes; see and see also references; and class-elsewhere notes all have hierarchical force (including notes are the most notable exception). Notes with hierarchical force can expand or contract the semantic space of a class. For example, the class-here note at 782.1-782.4 Vocal forms ("Class here treatises about and recordings of vocal forms for specific voices and ensembles") has hierarchical force at 782.292. This class-here note expands the neighborhood around each vocal form topic to include not only (e.g., printed) music that is of that form, but recordings of the music, and treatises about the music. Synthesis is thus another mechanism by which topics of other semantic types may be associated with a class.

Several aspects of the KOS contribute to the growth of neighborhoods around the focal topics of a class. Class neighborhoods are further expanded by rules for the use of the classification system. In the DDC, the rule of application, the first-of-two rule, the rule of three, and the rule of zero all provide guidance on selecting a single class for works whose subject matter points toward multiple classes. The implementation of these rules will often result in associating subject matter with classes in a manner that expands or contracts their scope. For example, by the rule of application, the use of chant in music therapy will be classed in 615.85154, not in 782.292; by the first-of-two rule, works on chant and tropes will be classed in 782.292 (782.292 Chant comes before 782.297 Tropes), but by the rule of three, works on chant (782.292), tropes (782.297), and psalms (782.294) will be classed in their superordinate class, 782.29.

Development of a classification system over time produces unintended expansions in the neighborhoods of affected classes. A completely revised music schedule was developed for Edition 20; previously, chant was classed in 783.5. Neither 782.2 nor any of its subdivisions was used in Editions 15-19, but in Edition 14, Wagnerian grand opera was classed in 782.2. Thus, music collections that have not been reclassified to reflect the revised schedule in Edition 20 may have works on divergent topics classed in 782.2 (the Wagnerian grand opera of yesteryear’s classification contrasting with the nondramatic vocal forms of the current classification). Other modifications to the system – discontinuations, relocations, and expansions – will likewise produce largely unintended expansions in the neighborhoods of affected classes over time.

So far, our topical expansion has been limited to expansions grounded in the KOS itself. Additional expansion takes place as the system is used to classify bibliographic resources. Works assigned to the same class will inevitably expand the topical neighborhood. For example, works on plainsong classed in 782.292 include Plainsong in the age of polyphony (Kelly 1992) and Accompaniments to plainsong for schools (Allen 1930). A review of the former (Wegman 1993) indicates that contributions in this edited volume address sociopolitical motivations for the diversification of plainsong, despite ecclesiastical efforts to maintain the uniformity of plainsong; this work both introduces extramusical factors to our topic neighborhood and shows that
plainsong is not a simple concept, but a neighborhood unto itself. Since plainsong was traditionally sung *unaccompanied*, the latter work gives evidence that this liturgical form has continued to evolve. At the same time, variations in application practice by specific institutions will further affect the extent of a class’s neighborhood.

### 3: Representing neighborhoods using a formal knowledge representation language

When transforming a classification scheme into a concept-based representational model like SKOS, it is common practice to treat classes as concepts, i.e., instances of `skos:Concept`. This practice restricts the expression of relationships to the level of Dewey classes, with classes regarded as individuals of the domain. The ability to assert relationships between classes and topics is eliminated (topics are not recognized in such a formalization), unless a more expressive language in which instances can be treated as classes is used, e.g. OWL Full (McGuinness & van Harmelen 2004).

The conceptual model associated with Dewey classes and topical neighborhoods points in a different direction. Firstly, this notion of neighborhood is akin to the notion of classes in OWL (or similar) semantics, where classes work as an abstraction mechanism for grouping resources with similar characteristics together.

As discussed in section 2.1 above, captions and notes function to establish focal topics, which in turn define the extent of the topical neighborhood of a Dewey class. These mechanisms work similarly to class descriptions in OWL that place constraints on the class extension. In essence, each of these mechanisms can be expressed as one or more class axioms that conjunctively describe the topical extent of a Dewey class.

However, the fundamental difference between a formal knowledge representation language like OWL (based on first-order logic) and the more associative way that most KOSs have been structured should not be underestimated (Zeng et al. 2010). We believe it is beneficial to apply a formal framework to parts of classic KOSs, following a “multilevel approach” (Priss 2004). It is possible to start by formalizing a conceptual model of coarse basic relationships, even if not every finer-grained means by which neighborhoods are developed can be mapped to a formal ontological counterpart. The semantics of this upper level should remain compatible with the full DDC, working like an abridgment in the same way the abridged edition of the DDC represents a specific “level” of the classification that is consistent with the full system.

The domain of the DDC recognizes two disjoint ontological classes, `DeweyClass` and `Topic`. This basic structure enables the important formal distinction of relationships between two Dewey classes (each an instance of the OWL class `DeweyClass`), between two topics, and between classes and topics, all of which are needed to establish topical neighborhoods associated with Dewey classes.

Interclass relationships are characterized by the structural hierarchy of the scheme. Ontologically speaking, the structural hierarchy can be modeled as implications, or, using OWL semantics, as class-subclass relationships. Part-whole relationships would have to be constructed as new properties, as OWL does not provide built-in primitives, and therefore would have limited impact on the inference of topical neighborhoods. If a rigorous definition of structural hierarchy as implication or subsumption were used, some currently asserted hierarchical relationships would need to be changed.

Intertopic relationships resemble thesaurus relationships. In the neighborhood of the focal topic chant, plainsong is construed as a narrower term, albeit in the stricter sense
that every instance of plainsong is also an instance of chant.

The description of the neighborhood of a Dewey class is achieved by expressing class-topic relationships as OWL class axioms with the domain of \textit{DeweyClass} and the range of \textit{Topic}. For example, many class-elsewhere notes can be expressed as complements of class-here notes, indicating that the negation of the topic should be in the Dewey class. The open world assumption made by OWL precludes the possibility of treating exclusions as implicit; the truth value of a statement not known to be true is not automatically assumed to be false. This assumption resonates well with the DDC and the nature of Dewey classes; the category description a class provides opens up a semantic space in which not every included topic has to be stated. In fact, some of the power of classification systems lies in the many ways this space can be shaped.

Moreover, OWL 2 provides an advanced tool for expressing the hierarchical force by allowing a property to be the composite of other properties. Thus, class-here note topics at 782.1–782.4 could be assumed to be in the neighborhood of its subclass 782.292.

In conjunction with the asserted class and topic hierarchies and the implications of inferred superclasses, these axioms should allow for limited inferences of topics that are in the neighborhood of a Dewey class. But as mentioned above, ontological class axioms are unlikely ever to reflect all of the complex processes of defining, extending, or restricting neighborhoods externally by either application rules or assigned works. They can nevertheless be an important tool to formalize specific levels of a KOS to enable automatic classification or consistency checking of the complex interplay between classes and topics.

4: Conclusion

As a KOS for organizing document-like objects, the DDC has intentionally used broad buckets to gather related topics, thus differing from an ontology that organizes knowledge directly. While this seems to be in conflict with the possibility of a “DDC ontology” (which would require a tightly constrained, well-defined domain model unlikely to be found in any bibliographic KOS), we believe that the construction of an ontological model that \textit{reuses} a certain level of knowledge in the DDC is feasible.

Notes

DDC, Dewey, and Dewey Decimal Classification are registered trademarks of OCLC: Online Computer Library Center, Inc. For help with DDC-specific terminology and usage, see glossary at <www.oclc.org/support/documentation/glossary/dewey>.

References

Allen H.P., 1930, \textit{Accompaniments to plainsong for schools}, Rushworth & Dreaper, Liverpool.
Priss U., 2004, Multilevel approaches to concepts and formal ontologies, in E.N. Efthimiadis ed., \textit{Advances in classification research}, v. 12: proc. 12th ASIST SIG/CR Classification Research


Web documents have been accessed 29 September 2009.