Classification

7.1. Classification schemes

Overview

This module presents brief summaries of classification schemes that demonstrate principles of the two major approaches to classification: hierarchical and faceted.

In **hierarchical approaches**, subjects and their relationships are prearranged in classes and subclasses; whereas in **faceted approaches**, potential facets for subject classes are predetermined, but classes and subclasses are not prearranged.

Also recall that all classification schemes have **facets** (types of classes, such as topic and language) as well as the **classes** themselves (such as the languages English and French). At the same time, remember that the major classification schemes were developed for certain information environments and have somewhat different terminology. For example, faceted approaches use the term "foci" for classes (although for simplicity's sake, this module uses mainly "classes").

Note that, while most classification schemes are mainly intended for subject description, they also can accommodate different kinds of subjects and often other attributes such as literary form and physical format. Major schemes have published **schedules** of classes with alphabetical **indexes** to help locate concepts in the schedules.

As you read these summaries, follow the Web links to see examples of the schemes and learn more about their history and creators. People like Melvil Dewey and S.R. Ranganathan were fascinating characters who contributed far more to library and information science than their classification schemes.

Hierarchical Approaches

- Subjects and their relationships are prearranged in hierarchical classes and subclasses
- Schemes are usually also enumerative, which means they attempt to include every possible concept.
- Schemes are relatively inflexible, although classes can be added carefully within the existing hierarchy.
- The classifier (1) analyzes the information object to determine its subject, (2) identifies the location of the object in the scheme, and (3) applies a predetermined notation for each class.

**Project Alert!** For more information, see the tutorial on hierarchical classification. Note that hierarchical schemes are difficult to develop, and most projects are not extensive enough to warrant them. That's why we recommend that most of you take a faceted approach.
Dewey Decimal Classification (DDC)

**History**
Created by Melvil Dewey, published in 1876

**Current use**
School, public, and small academic libraries

**Schedule**
Three volumes
Symmetrical hierarchical class structure

**Index**
Relative Index, 1 volume

**Notation**
Pure (one type of character only: numbers)
Simple decimal expansible system
(add more decimals for narrower subclasses)

**Advantages**
- Widely accepted
- Frequent revisions
- Compact
- Easy to use

**Disadvantages**
- Arbitrary division by tens
- Tendency toward long notation
- American orientation

DDC’s division by tens is called **arbitrary** because it forces classes into groups of ten instead of the number of classes naturally suggested by the topics themselves. When there are fewer than ten classes, some numbers are not used. When there are more than ten classes, other subdivisions of ten are created.

If you look at the current structure of DDC (the DDC 21 Summaries), you will see that the areas of greatest growth, such as technology, have gained many subclasses. Computers are in the 000s, with other general works, including those of library and information science.

DDC has often been criticized for its American orientation. One example is in the 200s, Religion, where all classes except "other" originally related to Christianity. This reflected U.S. library collections of the 19th century. DDC now includes other major religions such as Judaism and Hinduism.

The **Relative Index** is an important contribution by Dewey: it is an alphabetical index of classes that points to the location of a class in the hierarchy.

DDC classifiers call the process of assembling notation **number building**.

To see how DDC works, go to Dewey Decimal Classification System at [http://www.oclc.org/dewey/about/](http://www.oclc.org/dewey/about/)

Because DDC is used in school libraries, there are some excellent sites about it for—and by—children:

"Do We" Really Know Dewey? at [http://tqjunior.thinkquest.org/5002/](http://tqjunior.thinkquest.org/5002/)
Let's Do Dewey at [http://www.mtsu.edu/~vvesper/dewey.html](http://www.mtsu.edu/~vvesper/dewey.html)

For Dewey with rhythm, check out two rap songs at [http://web.utk.edu/~dgobbell/melvildewey.html](http://web.utk.edu/~dgobbell/melvildewey.html)

To learn more about Melvil Dewey’s contributions to librarianship, see Melvil Dewey: The Father Of Modern Librarianship at [http://www.booktalking.net/books/dewey/](http://www.booktalking.net/books/dewey/)
Library of Congress Classification (LCC)

<table>
<thead>
<tr>
<th>History</th>
<th>Developed largely by Herbert Putnam; Z schedule published in 1898</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current use</td>
<td>Medium to large academic libraries</td>
</tr>
<tr>
<td>Schedule</td>
<td>Forty-seven, separately published</td>
</tr>
<tr>
<td></td>
<td>Asymmetrical hierarchical class structure</td>
</tr>
<tr>
<td>Index</td>
<td>No comprehensive index</td>
</tr>
<tr>
<td></td>
<td>(although LCSH substitutes to some extent)</td>
</tr>
<tr>
<td></td>
<td>Indexes in individual schedules</td>
</tr>
<tr>
<td>Notation</td>
<td>Mixed (both numbers and letters)</td>
</tr>
<tr>
<td></td>
<td>Some decimal expansion possible</td>
</tr>
<tr>
<td>Advantages</td>
<td>Widely accepted</td>
</tr>
<tr>
<td></td>
<td>Specific, relatively short notation</td>
</tr>
<tr>
<td></td>
<td>Constant updates</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>Lack of comprehensive index</td>
</tr>
<tr>
<td></td>
<td>Inconsistencies among schedules</td>
</tr>
<tr>
<td></td>
<td>Bulky, difficult to use</td>
</tr>
</tbody>
</table>

The first schedule developed, Z, encompasses library and information science (comparable to 000 in DDC). Subsequent schedules were developed gradually over many years. Law (K), for instance, started with U.S. law and has expanded to law in other countries. Schedules are developed by committees of subject experts and follow the organization of the discipline rather than one standard overall pattern.

LCC schedules are so large that different schedules have their own indexes. It is possible to use Library of Congress Subject Headings (LCSH) as an overall index of sorts.

LCC notation is shorter because mixed notation (letters and numbers) allows more options for codes than pure notation (numbers only).

To see how LCC and its notation work, go to Library of Congress Classification System at [http://geography.miningco.com/library/congress/bilc.htm](http://geography.miningco.com/library/congress/bilc.htm)

The Library of Congress was founded by Thomas Jefferson in 1800 to serve Congress; over many years it evolved into its role as the national library of the United States. Jefferson's own collection served as its foundation, but in 1851 two-thirds of this collection was destroyed by fire. Librarians of Congress are Presidential appointees. Herbert Putnam was one appointee who was instrumental in developing the current classification scheme, which he based in part on Jefferson's personal system.


Details of Jefferson's classification scheme can be found in a description of his 1789 Catalog of Books: [http://www.thomasjeffersonpapers.org/catalog1789/catalog1789_chapters.html](http://www.thomasjeffersonpapers.org/catalog1789/catalog1789_chapters.html)


A brief bio of Herbert Putnam is at [http://web.utk.edu/~jgambill/putnam/](http://web.utk.edu/~jgambill/putnam/)
Faceted Approaches

- Potential facets for subject classes are predetermined, but classes and subclasses are not prearranged.
- Faceted approaches are also called *analytico-synthetic* because the information object is analyzed and then the notation is synthesized (combined) to suit the object.
- Schemes are flexible and hospitable (easily expandable):
  - Not every facet must be used and
  - Classes (foci) can be added as needed.
- The classifier:
  - analyzes the information object to determine its subject,
  - identifies each facet applicable to the object, and
  - synthesizes notation by drawing together notation codes for different facets.

Generic example

A table is a good way to illustrate the structure of a faceted scheme, as in this example for literature:

<table>
<thead>
<tr>
<th>Language</th>
<th>Form</th>
<th>Period</th>
<th>Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>C</td>
<td>E</td>
<td>G</td>
</tr>
<tr>
<td>1 American</td>
<td>Poetry</td>
<td>18th c.</td>
<td>Nature</td>
</tr>
<tr>
<td>2 English</td>
<td>Epic Poetry</td>
<td>19th c.</td>
<td>Politics</td>
</tr>
<tr>
<td>3 French</td>
<td>Essays</td>
<td>20th c.</td>
<td>Sports</td>
</tr>
</tbody>
</table>

Columns represent facets and rows represent classes (foci). The notation codes are letters for facets and numbers for classes. To synthesize notation, the classifier finds the applicable facets and classes and combines the codes in order of the facets (citation order). For example: Collection of 19th-century French literature: A3E2 and Modern American essays on politics: A1C3E3G2

Note that **not all facets must be used**, but that the **facets that are used must be combined in order**. The actual codes, plus punctuation, vary widely among schemes.

**Project Alert!** The generic approach in the example above is a good model to follow. For more information, see the tutorial on faceted classification.

Colon Classification

<table>
<thead>
<tr>
<th>History</th>
<th>Created by S. R. Ranganathan, published in 1933</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current use</td>
<td>Rarely used, but an inspiration for development of all schemes since</td>
</tr>
<tr>
<td>Schedule</td>
<td>One volume containing many schedules</td>
</tr>
<tr>
<td></td>
<td>Includes broad disciplines in traditional hierarchies</td>
</tr>
<tr>
<td></td>
<td>Subclasses are based on universal facets (<em>PMEST</em>)</td>
</tr>
<tr>
<td>Index</td>
<td>Included in single volume</td>
</tr>
<tr>
<td>Notation</td>
<td>Mixed</td>
</tr>
<tr>
<td></td>
<td>Noted for use of punctuation (colon),</td>
</tr>
<tr>
<td></td>
<td>which includes Roman and Greek characters</td>
</tr>
<tr>
<td>Advantages</td>
<td>Innovative; much imitated</td>
</tr>
<tr>
<td></td>
<td>Hospitable, flexible</td>
</tr>
<tr>
<td></td>
<td>Many revisions, adaptations</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>Too compact, terse for ease of use</td>
</tr>
<tr>
<td></td>
<td>Long, complex notation</td>
</tr>
<tr>
<td></td>
<td>Poorly suited to shelving books</td>
</tr>
</tbody>
</table>
Colon Classification is noted more for its innovative approach rather than its usefulness: it is complicated to explain as well as to apply. It is not solely faceted; the scheme contains hierarchies of disciplines within its facets.

The most interesting, and most often cited, aspect of this scheme is the philosophy of facets based on *PMEST* (by their first letters):

- Personality (primary characteristics, "essence")
- Matter (physical characteristics)
- Energy (operations, problems, processes)
- Space (geographical, topological)
- Time (date, period)

Many Web sites that describe Colon Classification also compare it to other schemes and include biographies of its creator, S.R. Ranganathan:

- Was Ranganathan a Yahoo?  

- Ranganathan's Colon Classification: A Selected and Annotated Bibliography  
  [http://www.slais.ubc.ca/people/students/student-projects/N_Aerts/517-group1/TitlePage.html](http://www.slais.ubc.ca/people/students/student-projects/N_Aerts/517-group1/TitlePage.html)

Two essays on Ranganathan as a father of library science, by Eugene Garfield:

- A Tribute to S.R. Ranganathan, The Father of Indian Library Science, Part 1. Life and Works:  

- A Tribute to S.R. Ranganathan, The Father of Indian Library Science, Part 2. Contribution to Indian and International Library Science:  

  (FYI, Garfield is himself a father of information science for his work in indexing)

**Art and Architecture Thesaurus (AAT)**

<table>
<thead>
<tr>
<th><strong>History</strong></th>
<th>Created by many people through the J. Paul Getty Trust; first released in 1980</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current use</strong></td>
<td>Current use by archivists, slide/photo librarians, museum curators, indexing services, architecture/design firms, art dictionaries and encyclopedias</td>
</tr>
</tbody>
</table>
| **Schedule** | Facets based on Ranganathan's PMEST approach  
  Has hierarchies within facets |
| **Index** | Really a thesaurus; a thesaurus is an index |
| **Notation** | Mixed  
  Each descriptor (controlled vocabulary term) has a unique code that can serve for classification |
| **Advantages** | Much needed for area difficult to handle with hierarchical approach  
  Potential wide acceptance |
| **Disadvantages** | Not universal for subject disciplines  
  Not designed for shelving |

AAT is a thesaurus, not a classification system. The visual arts and architecture have never been adequately covered by controlled vocabularies like LCSH. For many years museums developed their own vocabularies and classification systems for physical artifacts. These worked well within the museums, but were not standardized across museums, which greatly hindered the work of researchers.
AAT has proven useful not only for its controlled vocabulary, but also—and rather surprisingly—as a classification scheme because each descriptor (term) has its own code that can serve as classification notation. AAT is used for

- Physical description of museum objects, slides, photos, archival materials, etc.
- Subject cataloging and keyword indexing of books, images, periodical literature
- Database searching by scholars, students, museum professionals, librarians
- Classification integrated with catalog description

AAT is used as an example here because it is contemporary, widely used, and clearly representative of a faceted approach. Web sites about AAT emphasize its faceted structure and derivation from Ranganathan's approach.

Ranganathan and Facet Analysis:
http://www.mysticseaport.org/library/msitia/facets.html

Art & Architecture Thesaurus browser (try searching, say, “furniture”):
http://www.getty.edu/research/tools/vocabulary/aat/

Summary

Advantages of Major Approaches

As with all information organization, each approach has advantages. Here is a comparison of hierarchical (H) with faceted (F) approaches:

<table>
<thead>
<tr>
<th>Advantage</th>
<th>Coverage</th>
<th>Schedule and index</th>
<th>Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wide range of disciplines or broad subjects</td>
<td>Widely accepted scheme</td>
<td>Suitability for collocation</td>
</tr>
<tr>
<td></td>
<td>One or few disciplines or broad subjects</td>
<td>Shorter schedule</td>
<td>Fairly short and uncomplicated</td>
</tr>
<tr>
<td></td>
<td>Relatively few facets</td>
<td>Easy-to-use index</td>
<td>Easy-to-use expressive structure (shows level in hierarchy)</td>
</tr>
<tr>
<td></td>
<td>Complex and compound facets</td>
<td></td>
<td>Flexibility of citation order (can omit facets)</td>
</tr>
</tbody>
</table>
Choosing a scheme

This module shows only examples of major schemes. Hundreds of schemes exist and more are constantly being created to serve particular environments. New are usually based on principles of existing schemes. Here are the actions typically taken:

- Consider relative advantages of major classification approaches.
- Find out what scheme is used in information settings with similar problems.
- Decide whether to adopt, adapt or develop a scheme.
- If developing new scheme, consult subject area specialists and authority sources (people, reference materials).

An example of a major adaptation is *Universal Decimal Classification* (UDC), an international scheme, which was derived from Dewey Decimal Classification (DDC); see [http://www.udcc.org/about.htm](http://www.udcc.org/about.htm)

**Project Alert!** The assignment calls for an original scheme, not an adaptation of an existing scheme. Major schemes used as examples in this module are much more complex than necessary for the project. For example, your scheme has schedule and rules, but no index.