Bliss on the Web

Leonard Will

<www.willpowerinfo.co.uk>

Slides from a talk given at the 2008 AGM of the
Bliss Classification Association

The following slides were used to illustrate a talk discussing ways in which the Bliss classification might be made available on the Web. They do not represent the complete content of the talk, but the information that they contain should be understandable on its own.

The first slide shows an example of how another classification scheme, the Dewey Decimal Classification, has been made available on the Web, showing a hierarchy, notes and rules for synthesis.

Search　　|　　Browse　　|　　Create　　|　　Show

* WebDewey　☑　Browse Options ☑　Create Options ☑　Show Options　　　　　　☑

◉ Search ○ Browse　hosiery　　　in　All Fields　　　☑　　Quick Search　　Search History

◀ 2 ▶　Search Results　Record 2 of 9

Centered Entry: 646.42-646.48　　　　　　　　　Browse　Tables　Create Note
Caption:　　　Specific kinds of clothing　　　　　　　　　Notes　Terms

	Main Classes
600	Technology
640	Home & family management
646	Sewing, clothing, management of personal and family life
646.4	Clothing and accessories construction
646.42-646.48	**Specific kinds of clothing**
646.42	*Underwear and hosiery
646.43	*Specific kinds of garments
646.45	*Outer coats, sweaters, wraps
646.47	*Garments for special purposes
646.48	Accessories

Notes

Add to each subdivision identified by * the numbers following 646.4 in 646.4001-646.406 , e.g., construction of shirts for persons with physical disabilities 646.43501

Class comprehensive works in 646.4 .

The next three slides show:

• a list of some formats that have been used to represent classification data in computer form

• the format in which the Bliss schedules are currently prepared

• the format in which the UDC schedules are currently stored

Computer formats for classification data

Exchange formats
- Tagged text files
- XML files
  - BS DD8723-5
  - SKOS

Internal structures
- Text files
- Relational databases
- XML databases

Problem: no satisfactory format to represent compound concepts made up of combinations of concepts from different facets

Representing a classification in a format that can be processed by computer

Bliss schedule source file format (with notation omitted)
Whole schedule represented as a continuous list

05(objects)
06clothing
07(clothing by type)
08socks, stockings, hose
09(properties)
10<colour>
11<blue colour>
11<green colour>
11<red colour>
11<yellow colour>
09(materials)
10<cotton>
10<wool>
11(properties)
12<blue colour>

Numbers indicate logical and physical indentation levels

No distinction is made between node labels showing characteristics of division and labels showing change of facet, both being shown in round brackets (...).

Synonyms shown as a list separated by commas or equals signs.

Concepts "brought down" from other facets have been shown in angle brackets <...>. Originally reversed brackets)...(were used, but these caused problems in processing.

Additional codes show which terms may appear in the index or thesaurus.

UDC export format

Each concept stored as an individual record

Tag	Field	Content
<01>	001	UDC-number
<02>	100	Description
<03>	105	Verbal examples
<04>	110	Note
<05>	111	Note
<06A>	115	Example of combination: direct addition (1st subfield: ^a)
<06B>	115	Example of combination: colon addition (1st subfield: ^b)
<06C>	115	Example of combination: full notation (1st subfield: ^c)
<07>	011	Instruction for parallel division
<08>	120	Example of parallel division
<09>	125	References

The next three slides show the structure of a faceted classification scheme that has to be represented in any computer format, including:

• hierarchical relationships within a facet, with node labels

• combination of concepts from more than one facet to express compound subjects, with facet labels to show when a change of facet occurs.

Concepts in a single facet, subdivided by characteristics of division

clothing
  <clothing by type>
  jumpers
  shirts
  socks
    <socks by colour>
    blue socks
    green socks
    red socks
    yellow socks
    <socks by material>
    cotton socks
    woollen socks
      <woollen socks by colour>
        blue woollen socks
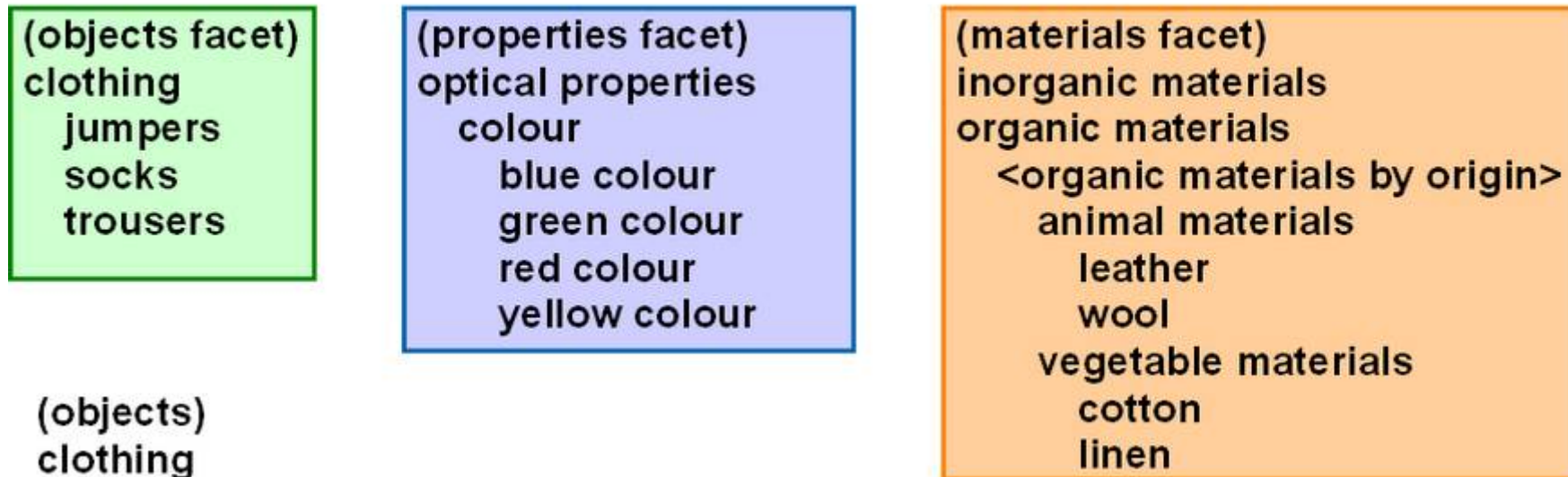   green woollen socks
  trousers
   . . .

Every item is a specific type of 'clothing',
i.e. the relationships are all BT/NT

Node labels showing *characteristics of division* are not labels for concepts and do not have a hierarchical relationship to the concepts.

They contain the word "by". In Bliss sometimes the parent term is omitted or replaced by "kinds",
e.g. <by colour> or <kinds by colour>.

SKOS can encode this type of structure

Concepts in multiple facets, pre-coordinated as necessary

(objects facet)
clothing
  jumpers
  socks
  trousers

(properties facet)
optical properties
  colour
    blue colour
    green colour
    red colour
    yellow colour

(materials facet)
inorganic materials
organic materials
  <organic materials by origin>
    animal materials
      leather
      wool
    vegetable materials
      cotton
      linen

(objects)
clothing
  <clothing by type>
  socks
    (properties)
      colour
        blue colour
        green colour
        red colour
        yellow colour
    (materials)
      cotton
      wool
        (properties)
          blue colour

Node labels showing names of facets – "facet labels" – show different facets from which concepts have been *combined*. The relationship of concepts across these labels is not hierarchical; it *may* be RT/RT, but not always.

This class represents the compound socks : wool : blue colour

Classification schedules show *examples* of pre-coordination. Users are expected to create others as required. SKOS can not yet encode this type of structure.

Pre-coordination which cannot be represented as arrays within a single facet

(objects facet)
clothes
   jumpers
   socks
   trousers

(activities facet)
manufacturing
   knitting
   sewing
maintaining
   darning
   washing

(materials facet)
   <materials by application>
    cleansing materials
     detergents
     soaps

(objects)
clothing
  socks
   (activities)
   washing
    (materials)
    detergents
    soaps

Should anything be listed *here* that is not listed *here*?

This class represents the compound concept "soaps for washing socks" expressed as:
socks : washing : soaps

Facets are combined here in the citation order objects – activities – materials (as agents)

Citation order is difficult to automate because it depends not only on the facets themselves but also on the role which each plays in the compound concept, e.g. "as agents" here.

How to move forward?

Web representation of printed pages?

Existing format for current developments?

Restructure schedules to unify facets?

Give node labels and facet labels distinct, standardised formats?

Consider changing notation to make facets explicit?

Experiment with database format?

Cooperate with UDC / SKOS / anyone else?

Projects to devise and test formats

Projects to restructure schedules?

Projects to move schedules into new formats?