Integrating Semi-formal Knowledge Organization Structures

FALQUET, Gilles, et al.
Integrating Semi-formal Knowledge Organization Structures

5th ECDL 2006 NKOS workshop
Alicante, Spain, 21SEP2006

Gilles Falquet
Luka Nerima
Claire-Lise Mottaz Jiang
Jean-Claude Ziswiler

Centre universitaire d’informatique
University of Geneva – Switzerland
Outline

Introduction: Building ontology-based Digital Libraries

Context

Hyperbook Alignment Process

Hyperbook Alignment Process (findings)

Selecting / validating relations through social networks

Conclusion
Introduction: Building ontology-based Digital Libraries

Hyperbooks

- Break the monolithic aspect of the documents
Introduction: Building ontology-based Digital Libraries

Digital Library built of hyperbooks
- Compare books
- Extend a book with the contents of others
=> New access methods / reading possibilities
=> Synthesize books
Introduction: Building ontology-based Digital Libraries

Generating virtual documents out of the hyperbook structure through an interface specification

[Crampe], [Garlatti], [DeBra], [Brusilowsky], [Falquet]
Introduction: Building ontology-based Digital Libraries

Example: Virtual document of a Digital Library
Context

Wikipedia

encyclopaedia

architecture:
- articles (sub-articles)
- categories

collaboratively created (contains discussions)

content in plain text, difficult to process it automatically

Wikibooks (Wikiversity)
  -> reuse of existing content?
Context

Semantic Wikipedia

[Völkel et al., WWW 2006]

encyclopaedia

architecture:
  - articles (typed links between articles)
  - ontology (RDF, RDFS, OWL)

RDF export

=> Where is the wiki in Semantic Wikipedia?
### Context

**Electronic versions of documents**  
Book chapters  
Journal articles  
Web pages

**Weakly structured KOS**  
Glossaries  
Directories  
Indexes, Folksonomy  
Metadata annotated models (Learning Objects)

- + availability  
- – no or weak semantic structure  
- – not or hardly machine-processable

**Highly structured KOS**  
Ontologies  
Thesauri  
Concept maps  
Taxonomies

- + strong semantic structure  
- – not available  
- – difficult to construct  
- + machine-processable
Hyperbook alignment process

Integrating the ontologies of the hyperbooks

Compute **semantic similarity** between concepts of the hyperbook ontologies
Hyperbook alignment process

Alignment method

Alignment method that works with less formalized (incomplete) ontologies, adapted from similarity function of [Rodríguez&Egenhofer03] by involving fragments:

\[ a \cdot WM + b \cdot FragM + c \cdot NeighM(r) > t \]

- **Word Matching** (terms of the concepts, often 0.00)
- **Fragment Matching** (terms of the fragments, often > 0)
- **Semantic Neighborhood Matching** (terms of the concepts and fragments in the neighborhood)

Experiment with students that wrote hyperbooks about the course topics (within the hyperbook system that allows creating fragments / concepts / links)
Hyperbook alignment process (findings)

Relations that indicate equal concepts

-> depends on the setting (variables a,b,c,r / threshold t)

Setting S1: 19 relations were found
  -> 12 of 12 equal relations (precision 63%, recall 100%)

Setting S2: 11 relations were found
  -> 9 of 12 equal relations (precision 82%, recall 75%)

Relations that indicate similar concepts

-> with S1 and S2, precision around 50%, very low recall
Hyperbook alignment process (findings)

Strategy

- focus on high precision concerning relations indicates equal concepts?
- few, but good quality relations
Hyperbook alignment process (findings)

Strategy

-> focus on high precision concerning relations indicates equal concepts?
=> few, but good quality relations

=> alternative access path (different from IR)
Hyperbook alignment process (findings)

Strategy
Hyperbook alignment process (findings)

Strategy
Selecting / validating relations through social networks

Starting point

-> All links automatically generated through the alignment process are presented

3 possible interactions of the user:

-> Follow these links
-> Definitely accept a link in the user’s view
-> Definitely reject a link in the user’s view

=> not only a static validation of links (by statistical data)
Selecting / validating relations through social networks

Processes

=> Following or definitely accepting a link reinforces the quality of it (the quality of a link is measured by the similarity value calculated in the alignment process).
- Local effect (one user): validation bottom disappears
- Global effect (all users): rises the link quality

=> An explicit acceptance of a link by a user is stronger than the simple click on a link

=> Rejecting a link:
- Local effect (one user): link disappears
- Global effect (all users): decreases the link quality

-> not limited to generated similarity links, applicable also to manually created links
Conclusion

- Hyperbook: small domain ontology and fragments
- Building Digital Libraries by aligning hyperbook ontologies

=> Alignment process to find equal concepts works also with small ontologies if fragments and the semantic neighborhood are included
=> Setting of variables / thresholds (= [Villa et al., 2004])

=> Instead of finding the right strategy for the alignment process settings, we use Social Networks to select / validate relations (!= [Villa et al., 2004])
Conclusion

Explicit vs. implicit representation of semantics

-> Explicit representations are costly to produce
-> Re-use of existing sources difficult, often not appropriated

=> Explicit representations in hyperbooks are useful
   -> weakly structured KOS (not expensive to create)
   -> allows (automatic) interlinkage of hyperbooks
   -> pedagogical skills
   -> Internal re-use of resources is possible (Wikipedia -> Wikibooks)

=> Next step: Integration of the hyperbook system into a (semantic) Mediawiki
Integrating Semi-formal Knowledge Organization Structures

Thank you!

Gilles Falquet
Luka Nerima
Claire-Lise Mottaz Jiang
Jean-Claude Ziswiler

Centre universitaire d’informatique
University of Geneva – Switzerland