D. Grant Campbell
University of Western Ontario, Toronto, Ontario, Canada

FRBR and Models of Semantic Cognition

Abstract or Résumé:

This paper will use research in cognitive science, particularly the mechanical modeling of semantic cognition, to suggest ways in which the model of the Functional Requirements of Bibliographic Records (FRBR) can be mobilized in a fashion that is consistent with linked data, knowledge discovery and artificial intelligence. The shift from hierarchical models to those based on parallel distributed processing suggest new and innovative ways in which FRBR can enhance library catalogues in the future.

1. Introduction

This paper will use research in cognitive science, particularly the mechanical modeling of semantic cognition, to suggest ways in which the model of the Functional Requirements of Bibliographic Records (FRBR) can be mobilized in a fashion that is consistent with linked data, knowledge discovery and artificial intelligence.

2. FRBR and Markup Languages

IFLA's Report on the Functional Requirements of Bibliographic Records emerged in 1998. At the time, the World Wide Web, with its HTML encoding environment, was rapidly spreading, but had not yet attained the pre-eminence it was to gain in the following decade. FRBR emerged out of an older study: an ambitious attempt to model the Anglo-American Cataloguing Rules in terms of Entity-Relationship modeling, thereby making the Rules consistent with standard relational databases (IFLA Study Group on the Functional Requirements of Bibliographic Records 1998, 12). As such, FRBR rests on a series of one-to-many relationships, in which a single work can have multiple expressions, each expression supporting multiple manifestations, and each manifestation producing multiple items. After its release, however, the FRBR paradigm attracted considerable interest, not as a database structure, but as a potential document structure. The FRBR structure supported a major assumption of bibliographic control at the end of the 20th century: that information would be mediated through markup languages such as Standard Generalized Markup Languages (SGML) and its offshoots Extensible Markup Languages (XML) and Extensible Hypertext Markup Language (XHTML). The original persuasiveness of the FRBR paradigm, therefore, lay in its amenability to the nested hierarchies of markup languages. As the Library of Congress's FRBR Display Tool (2009) showed, MARC records could be converted to XML format, and then displayed with an XSLT stylesheet as a sequence of editions nested within works. And when RDA finally emerged in 2014, it proudly proclaimed that its FRBR-based structure could support both existing legacy systems and newer systems (RDA, Rule 0.1): a claim based, at least in part, on the fact that the

Semantic Web standards, developed by the World Web Consortium in the 2000s, frequently defaulted to an XML encoding environment that made similar use of nested elements in whole-part hierarchical arrangements.

As happens all too often in information systems, development and innovation took an unexpected turn. The primitive Web of the 1990s turned, not, as expected, into the rich Semantic Web envisioned by Tim Berners-Lee, but into a complex tapestry of makeshift interactions through Web 2.0 platforms, wikis, RSS feeds and weblogs. The rise of mobile technology shifted the default Internet environment from a browser-supported Web to a network of applications in Apple or Android environments. Above all, the rise of interest in linked data and big data led to a shift in focus from documents to data. Markup languages have their origin in the concept of the document, defined as an assemblage of data elements joined together through a document-type definition and then rendered through a stylesheet. The ascendance of "data" over the document has fueled a preference for temporary and provisional assemblages of data, visualized in highly transitory fashion. XML has, ironically, attracted the same kind of criticism that we have long leveled at MARC. XML, its critics cry, is too complex and cumbersome for what it achieves; it is mistakenly characterized as a data storage system rather than as a data coding system; it fosters unnecessary repetition; it is inconsistent with relational database structure (XML Sucks 2014).

What, then, are we to make of a paradigm like FRBR, which appears to fit so well into the XML model? One possible interpretation: that libraries have linked their future hopes to a falling star, and, in their efforts to move with the times, they have become even more outmoded than they were before. Another possible interpretation: that libraries have re-established their role on the periphery of the modern information environment. While the rest of the information world continues to chase its tail, libraries have established a slower, more elaborate system of information organization, designed to preserve the core bibliographic universe. We're updating our systems to prevent technological obsolescence—XML has a better chance of tech support in 50 years than MARC does—but in so doing we're positioning ourselves as the monastic preservers of the best of things.

3. Cognitive Models

This presentation will present an alternative to both these interpretations. Grounded in a project on the information aspects of dementia, this interpretation rests upon the work of cognitive scientists in the area of semantic cognition. It rests particularly in the practice of modeling semantic cognition according to mechanical models inspired by the rise of computer programming in the 1970s. In such models, computer programming is used as an analogy to explain the way human beings process and carry out cognitive tasks, how this ability develops in childhood, and how it changes in cases of dementia.

This research parallels research in knowledge organization in library and information science, in certain fruitful ways. Research in semantic cognition exhibits an underlying tension between two interpretive approaches. The first approach models human cognition in terms of hierarchical relationships: taxonomic structures provide a provisional means of explaining how general category membership influences processing of more specific category information (Rogers & McClelland 2004, 13). Important concepts from this modeling approach have found their way into knowledge organization: particularly the concept of the basic level of categorization (Rosch 1978, Lakoff 1987, Iyer 2012). Other such relevant connections include the concepts of back activation, which parallels Olson's discussion of hierarchical force (1999, 247), and naïve classification (Beghtol 2003).

Parallel distributed approaches in cognitive science have emerged as a means of explaining some curious features that eluded hierarchical approaches (Rogers & McClelland 2004, 2). They also serve to bring some system and order to the concept of "association," which, as in KO, tends to be a catch-all for all those "related terms" that defy hierarchical theorizing. In addition, because of its close link with computer science and computer programming, the PDP approach bears a close resemblance to the ways in which associative and hierarchical relationships are modeled in the Semantic Web and linked data. PDP models posit the existence of a semantic network, consisting of nodes and arcs that highly resemble the data triple of node-arc-node which characterizes the Resource Description Framework.

There exists, however, a split in cognitive science between two different ways of modeling the semantic network on PDP lines. Some argue that each node represents a black box: a complete concept with a fixed address (Hinton 1981, 161). Others argue that each node subdivides into an array of units which can be fired in multiple combinations. The same nodes and arcs, therefore, can sustain a variety of triples, permitting a semantic network that can sustain massive growth (Hinton 1981, 162).

This willingness to see nodes and arcs as flexible collections of sub-units bears a subtle relationship to the historical trends of development in bibliographic description. Ever since Charles Cutter articulated his famous set of objectives of a library catalogue, cataloguing theorists have been steadily enriching articulated concept of "the book." Julia Pettee addressed the multiple nature of literary works by introducing the concept of the literary unit, an assemblage of editions and translations, all representing a single "work"(1936). Seymour Lubetsky (2001) argued that instead of conceptualizing a "book," we need to see the library catalogue as a collection of related works and editions, a concept which was incorporated into the Paris Principles of 1962. The Anglo-American Cataloguing Rules posit the existence, not of simple resources, but rather of resources that consist of a bibliographic unit and a work. And with FRBR, the resource splits into four elements: the work, the expression, the manifestation and the item.

If we link this history to the developments of cognitive science, we can see the FRBR relationships in two different lights. We can see them as nested hierarchies, in which case we can see FRBR as the enactment of a taxonomic paradigm that resonates, not just with XML, but with taxonomic paradigms of cognition, complete with back propagation, inheritance, and base levels. On the other hand, can see FRBR as a means of enabling of parallel distributed processing system hierarchies, in which each core unit that we define as a "resource" exists, not as a black box but as an array of subunits within the resource, each of which can be fired separately.

4. Conclusion

A model which exploits the internal components of a bibliographic description provides some fascinating possibilities. Such a view gives us the potential to enrich the concept of "resource" which lies at the heart of the Resource Description Framework, beyond the relatively naïve notion that has prevailed in much of the W3C standards; it permits the formation and definition of relationships, not with "resources" as a whole, but with particular combinations of the sub-elements of various resources. Different resources in the catalogue can be linked in node-arc-node relationships in a multitude of ways, triggered by the different relationships between the sub-elements within the resource, and between sub-elements across different resources.

FRBR, then, may be the best thing that happened to library catalogues, for reasons that never occurred to those who designed it.

Reference List:

Beghtol, C. (2003). Classification for information retrieval and classification for knowledge: relationships between "professional" and "naïve" classifications. *Knowledge organization, 30/2*: 64-73.

Hinton, G.E. (1981). Implementing semantic networks in parallel hardware. In G. E. Hinton and J.A. Anderson, eds., *Parallel models of associative memory*. Hillsdale: Erlbaum, 161-187.

IFLA Study Group on the Functional Requirements of Bibliographic Records. (1998). *Functional requirements for bibliographic records: Final report,* München: Saur.

Iyer, H. (2012). *Classificatory structures: Concepts, relations and representation*. Würzburg: Ergon-Verlag.

Joint Steering Committee for the Development of RDA. (2014). *RDA: Resource description and access.* Chicago: American Library Association.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lubetsky, S. (2001). *Writings on the classical art of cataloging*. Ed. D. McGarry, E. Svenonius. Englewood: Libraries Unlimited.

Olson, H.A. (1999). Mapping beyond Dewey's boundaries: Constructing classificatory space for marginalized knowledge domains. *Library trends, 47/2,* 233-254.

Pettee, J. (1936). The development of authorship entry and the formation of authorship rules as found in the Anglo-American code. *Library quarterly, 6/3,* 270-290. https://doi.org/10.1086/613854

Rogers, T., & McClelland, J.L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge: MIT Press.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd, eds. *Cognition and categorization*, Hillsdale: Erlbaum, 27-48.

XML Sucks. (2014). WikiWikiWeb. http://wiki.c2.com/?XmlSucks