Are We Talking About the Same Structure? A Unified Approach to Hypertext Links, XML, RDF and Zigzag

Amir Pourabdollah[1] Helen Ashman[2] Tim Brailsford[1]

[1] School of Computer Science, The University of Nottingham
Jubilee Campus, Wollaton Road, Nottingham NG8 1BB, UK.

{amir.pourabdollah | tim.brailsford}@ nottingham.ac.uk

[2] School of Computer and Information Science, University of South Australia
Mawson Lakes, South Australia 5095, Australia

helen.ashman@unisa.edu.au

ABSTRACT
There are many different hypertext systems and paradigms, each with their apparent advantages. However the distinctions are perhaps not as significant as they seem. If we can reduce the core linking functionality to some common structure, which allows us to consider hypertext systems within a common model, we could identify what, if anything, distinguishes hypertext systems from each other. This paper offers such a common structure, showing the conceptual similarities between each of these systems and paradigms.

Categories and Subject Descriptors
H.5.4 **[Information Interfaces and Presentation]:** Hypertext / Hypertext – *architecture, navigation, theory*; E.1**:** *distributed data structures, graphs and networks*

General Terms
Design, Theory

Keywords:
Hypertext links, XML, Zigzag, RDF

1. INTRODUCTION
Although many of the major paradigms in modern hypertext superficially look to be very different, on a deeper level they do have a lot in. They all divide information into independent pieces of data (a set of nodes) and they all associate these nodes with each other (a set of links). This paper shows that a special kind of such a structure can be similarly found in Hypertext Link Models, Zigzag, RDF and XML. This ternary approach is not new, but it has not previously been used in this way to unify the fundamentals of different paradigms in hypertext community.

The simplest link-node structure is the binary mode, where each two nodes can be simply connected by means of a link. It has been suggested that a Binary Relational Model (BRM) [2] may be used to describe most classical hypermedia, with each binary relationship consisting of two linked nodes. However there are other hypertext systems and paradigms such as ZigZag and the Semantic Web which incorporate an awareness of *"the why"* of a link. A logical successor to the BRM is to use a ternary link-node structure – where the relations consist of not just of the two linked nodes but also a third node, which is the semantic of the link between them (Figure 1).

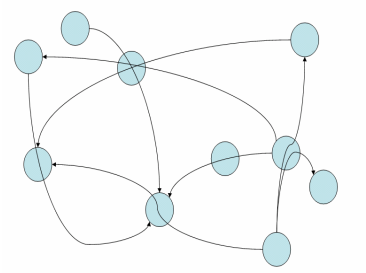

Figure 1: An example of a ternary node-link structure

2. A TOP-DOWN APPROACH
As the first approach, the possibility of reducing the structure of different paradigms to a ternary node-link structure is examined.

2.1 Hypertext LINKS
Reviewing different linking paradigms in hypertext, it can be seen that the semantic or associative elements of links have been considered in several ways – both implicitly and explicitly. It is possible to extract all the information about the associative meaning of a link into somewhere logically outside of the source and the destination nodes. This third node could be empty for an HTML link, or as complex as necessary (Figure 2).

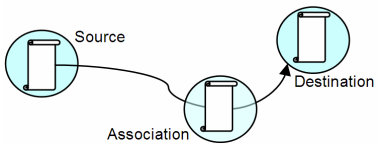

Figure 2: The ternary approach to hypertext linking

HT'08, June 19–21, 2008, Pittsburgh, Pennsylvania, USA.
ACM 978-1-59593-985-2/08/06.

2.2 XML

In addition to use the labelled graph conversion method [5] to represent XML in a ternary framework, a more general approach may be used to convert an XML listing to a complete ternary node-link structure. Here, each sub-element adds a branch to this hierarchy and connects the body contents to its super-element. An XML document can thus be converted to a set of ternary links, either in the form of (element, attribute-name, value) or (element, text, value) or (element, sub-element-name, sub-element). A simple example of this is shown in Figure 3.

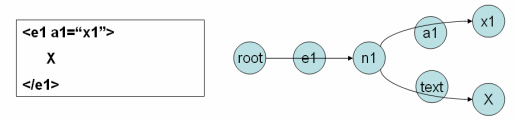

Figure 3: The ternary structure (right) of XML (left)

2.3 RDF

In the multi-layer architecture of the Semantic Web [1], RDF is used to write descriptive "statements" about each resource - by using other resources, in the form of resource-property-value triples. A similar ternary approach to the Semantic Web clarifies that RDF, as the basic data model of the Semantic Web, uses three URIs to build a relation.

2.4 ZigZag

ZigZag is an information paradigm that has been developed by Ted Nelson over the last decade[4]. In ZigZag, *cells* are atomic information units that can be interconnected with directed links along *dimensions* – which may have meaning or may be arbitrary. A cell cannot have more than an originating link or terminating link along any given single dimension (i.e. one in, one out). Using a ternary approach, a zzstructure may be reduced to a set of triples. Each triple consists of an originating cell, a dimension, and a terminating cell (Figure 4).

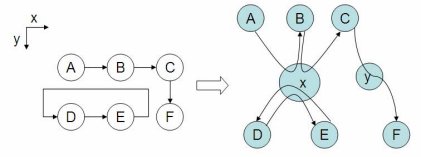

Figure 4: A ternary representation of a zzstructure

3. A BOTTOM-UP APPROACH

So, we now return to our original question – are we talking about the same structure? The answer is a definite yes. The fundamental information model of all of these hypertext paradigms is indeed a ternary node-link structure. A bottom-up approach can clarify why and how a common structure can bridge between different paradigms and also can rise new design ideas.

Each system may use a subset of a general ternary information structure (like the Zigzag main constrain or the static HTML links) in which each node can be connected to any other node through a third node, and while the selection of each of these three nodes can be a function of two others. Now in a down-top approach it can be seen how this flexible structure can benefit the top layers, how it can bridge between any of two upper layer applications, and even make some new applications' paradigms. For example, applying this to the hypermedia makes the links very flexible because these three nodes can be dynamic, i.e. functions of environmental variables or even functions of each other. This implementation called *functional linking* (like pE links of BRM [2]) inherits the benefits of functional links including the Turing completeness of the resulting hypertext system [3]. As another example, a new data model can be thought like Zigzag but with no constrain in one-to-many relationship.

This study has also highlighted new horizons of interconnectivity between these areas, which previously may have not been always noticed. For example, Zigzag can be used as an RDF repository, and XML and Zigzag information can be bridged and RDF, XML and Zigzag can all be used to manage hypertext links. By an extended view, a database can be rewritten in either Zigzag or RDF, and XML can be used to share those database contents.

4. CONCLUSION

The ternary modelling provides a flexible approach that can describe the data models underpinning all of the paradigms that we have examined: Generalized Hypertext links, XML, RDF and ZigZag. Because these, seemingly very different, systems can all be considered to be ternary node-link structures. The differences between these paradigms, which are very substantial, are not at the level of the information model, but in visualisation and user interaction. This approach does seem to be useful in bridging the gaps between different areas of research, and highlighting the potential for building mixed paradigm systems. There are also likely to be many applications of this approach in fields other than hypertext, such as database or workflow management. In the future we intend to formalise this into a new information model the Ternary Relational Model (TRM), which will be the true successor to the BRM.

5. REFERENCES

[1] G. Antoniou and F. van-Harmelen, *A Semantic Web Primer*: the MIT Press, 2004.

[2] H. Ashman, "Relations Modelling Sets of Hypermedia Links and Navigation," *The Computer Journal*, vol. 43, pp. 345-363, 2000.

[3] L. Moreau and W. Hall, "On the Expressiveness of Links in Hypertext Systems," *The Computer Journal*, vol. 41, pp. 459-473, 1998.

[4] T. H. Nelson, "A Cosmology for a Different Computer Universe: Data Model, Mechanisms, Virtual Machine and Visualization Infrastructure," *Journal of Digital Information*, vol. 5, 2004.

[5] F. Tian, D. J. DeWitt, J. Chen, and C. Zhang, "The Design and Performance Evaluation of Alternative Xml Storage Strategies," *SIGMOD Rec.*, vol. 31, pp. 5-10, 2002.