Visualising Discourse Structure in Interactive Documents

Clara Mancini **Christian Pietsch** **Donia Scott**

Centre for Research in Computing
The Open University
Walton Hall, Milton Keynes, MK7 6AA, UK
{c.mancini,c.pietsch,d.scott}@open.ac.uk

Abstract

In this paper we introduce a method for generating interactive documents which exploits the visual features of hypertext to represent discourse structure. We explore the consistent and principled use of graphics and animation to support navigation and comprehension of non-linear text, where textual discourse markers do not always work effectively.

1 Introduction

There is a long and well-established literature within theoretical, computational and psycho-linguistics on textual devices that function to signal the coherence structure of a discourse to the reader. This work has addressed the traditional conceptualisation of text: a two-dimensional array on a physical page, traversed in a set pattern (e.g., left to right, top to bottom in the Western tradition). As texts are increasingly being read on a computer screen, the act of reading is progressively becoming one of hypertext navigation. However, hypertext presents a strikingly different conceptualisation of text, and brings with it new challenges for conveying discourse structure.

Hypertext is distinctive in that it is interactive and non-linear, with several reading paths available through the document. It is organised around *nodes* and *links*; the reader moves from node to node by mouse-clicking on links. A node can be the equivalent of a traditional text page or can contain just a few sentences; links can be words or graphical elements. Since nodes typically contain more than one link, the author can only partially control the order in which the reader will access them.

With hypertext, then, a new conceptualisation of text has emerged as a three-dimensional array on a computer screen, which can be traversed in any number of ways. One of the challenges this poses for text research is that coherence markers of the traditional notion of text often do not work for this new medium. We are exploring new possibilities for signalling coherence in non-linear documents, exploiting the graphical features of a visually rich medium yet to be systematically exploited as a dimension of signification. This work is set in the context of the textual presentation of medical records from a repository of data-encoded medical histories.

2 Structure representation in non-linear text

As we discuss elsewhere (Mancini et al., 2007), discourse markers such as adverbials, pronouns and connectives cannot reliably be used to signal the discourse relation between hypertext nodes, since nodes can be accessible in more than one way, via paths that reflect different relevancies. This restriction is less likely to apply to graphical features, because they are visual and work in space. Owing to its technical characteristics, hypertext is a spatial medium (Carter, 2000) as well as a temporal one (Luesebrink, 1998), in which spatial structures have a temporal dimension and realisation: both space and time can be exploited in hypertext to express discourse coherence through space-temporal configurations in a three-dimensional space.

At present, most hypertexts (especially on the Web) make no use of graphical features to signal discourse relations between nodes, and nodes often consist of long text pages with a few links targeting other pages, from where the source page can no longer be seen. We take a different approach whereby the text is made more readable in two ways: by making hypertext nodes much smaller and by using graphical features to signal the coherence relations between nodes. We use the screen as a visual field across which the text can dynamically distribute, as links are clicked and new nodes appear, composing meaningful patterns. The presentation and distribution of the nodes are intended to signify the rhetorical role that their content plays within the discourse. To achieve this, coherence relations are used as document structuring principles during discourse construction to define hypertext links. These are then dynamically rendered during navigation through the consistent and concurrent use of the medium's spatial and temporal graphic features connoting the nodes. We refer to this paradigm as *cinehypertext*.

Having established a parallel between textual and visual processing, research informed by Gestalt theory has proposed relevant principles of document design (Campbell, 1995; Riley and Parker, 1998). Additionally, a number of representational rules for visually expressing discourse relations between hypertext nodes can be derived from the semiology of graphics, according to which graphic features can be employed to express conceptual relationships of *similarity*, *difference*, *order* and *proportion*, exploiting the properties of the visual image (Koch, 2001). Following these principles and rules, we have designed and begun testing a series of prototype visual patterns expressing coherence relations in non-linear discourse (Mancini et al., 2007).

3 Graphics devices to visualise interactive text structure

Our empirical work so far shows that graphics can indeed be used to express abstract relational concepts (Power et al., 2003; Mancini, 2005). But can graphics usefully support the expression of discourse structure in hypertext navigation? Can visual discourse markers support comprehension, substituting or complementing textual discourse markers in non-linear documents? If so, how do these textual and graphical features interact: can they express the same semantic meaning by following the same principles or do they, due to their different semiotic characteristics, function in different ways to the same end?

Our starting point is two NLG systems that generate, as linear text, summaries of a cancer patient's medical history (respectively generated to be read by clinicians and medical researchers, or by patients) from a repository of data-encoded medical reports (Hallett et al., 2006; Williams et al., 2007). We are extending these systems to generate interactive text on the one hand, and its animated graphical presentation, on the other. Since textual and graphical discourse markers will interact with each other,

> You had a consultation with your doctor on September 20th 1993.
>
> On September 27th you did a self examination and you found that you had a lump in your right breast. A self examination is an examination of the breasts by running your hand over each breast and up under your arms and checking for changes to their size, shape or feel.
>
> On October 4th you did another self examination and you found that you had a lump in your right breast.
>
> On October 11th you had a radical mastectomy to treat cancer in your right breast. A radical mastectomy is an operation to remove the breast, along with the lymph glands under the arm and the muscles of the chest wall. Cancer is a tumour that tends to spread, both locally and to other parts of the body.

Figure 1: Example section of a linear report generated for a patient.

our interactive documents may not present some of the textual features that occur in the corresponding linear form. We aim to provide a principled account of this interaction, and start by studying the effect of using visual markers on an interactive version of the linear reports already produced by our generators. Figure 1 shows an example section of a linear textual report.

At present, we transform a linear medical report into an interactive one by dividing the text into related chunks within text nodes whose graphics and animation features signal the relations holding between the chunks. In our example, the relations holding between the nodes are SEQUENCE, ELABORATION, MOTIVATION and RESULT. We generate graphical presentations of these discourse relations, using motion trajectory, distribution, colour value and dimensions of text windows. Red arrows and bold fonts in the windows signal active links, whereas followed links are signalled by dark grey arrows and bold fonts.

Specifically, the SEQUENCE of events described in the report is expressed by the vertical alignment of text windows appearing one under the other and having the same width and background colour (Figure 2).

Causal relations, like RESULT and MOTIVATION, holding between events are expressed by the horizontal alignment of text windows having the same height, where the one representing the result slides out, moving from left to right, from behind the one representing the cause, having a darker background (Figure 3)

The expression of MOTIVATION is similar to that of RESULT, with the difference that this time the window representing the motivation moves from

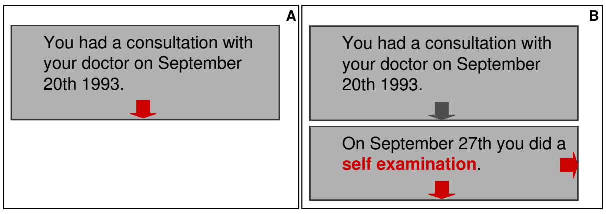

Figure 2: Representation of SEQUENCE: the red arrow link in the first node (A) is activated and the second node appears underneath it (B).

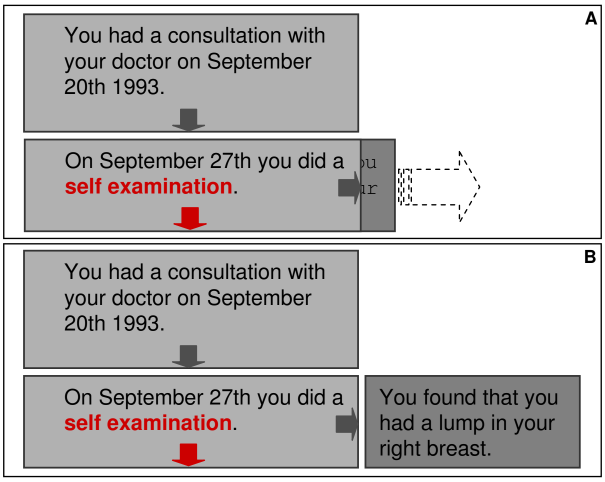

Figure 3: Representation of RESULT: the result-node slides out from behind the cause-node (A) placing itself next to it on the right hand side (B).

right to left and has a lighter background (Figure 4).

Finally, the definitions for which an explanation is available constitute links whose activation triggers an ELABORATION node, which appears slightly overlapped along a virtual third dimension. While the background of the other nodes is grey, the background of elaborative nodes is a lighter tone of red (Figure 5).

We have been evaluating the significance of these and other visual patterns, while exploring new designs to express discourse relations in interactive documents of different formats. We are also implementing a prototype system that employs a selection of patterns to carry out more empirical studies. The architecture of the prototype is described below.

4 System architecture

Few existing text generators which produce HTML output actually use the properties of non-linear text fully. Sometimes HTML is used as a simple text formatting language. Where hyperlinks are provided,

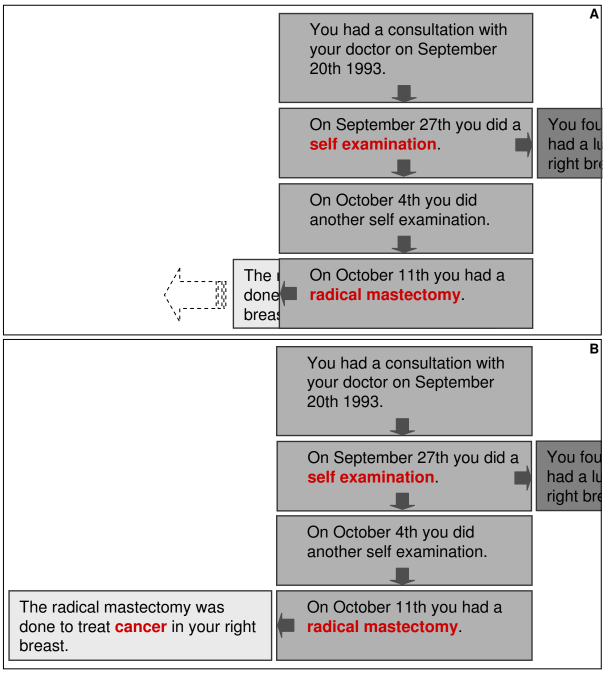

Figure 4: Representation of MOTIVATION: the motivation-node slides out from behind the consequence-node (A) placing itself to its left (B).

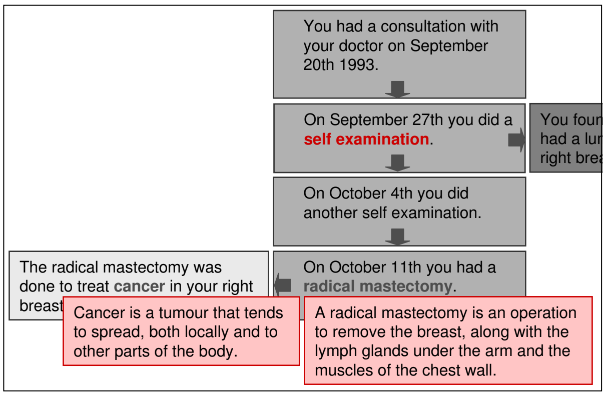

Figure 5: Representation of ELABORATION: the elaboration-node appears overlapped to the node containing the definition that is being elaborated on.

the rhetorical, semantic or pragmatic relation of the link target to the anchor is rarely made explicit. Our approach is a radical departure from this practise.

We propose an additional layer of abstraction embodied in a well-defined data format we call XCH (XML for CineHypertext), and describe a prototype architecture that extends previous approaches by providing a principled way of hyperlinking and animating content, thus encouraging the user to interact with the information dynamically presented. In fact

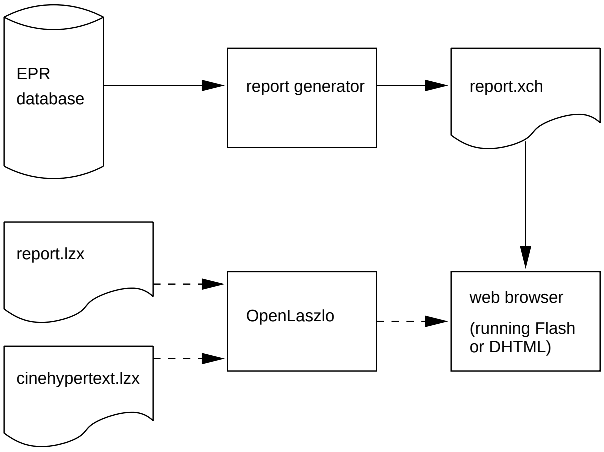

Figure 6: Prototype architecture. Solid arrows signify online data flow, dashed arrows signify data flow which occurred earlier.

we are crossing the border from document generation to application generation because the user interacts with a Rich Internet Application (RIA) which is very much data-driven.

Our XCH format is an instance of XML defined by two XML schema files. One is domain-independent and captures the range of semantic/pragmatic relations that can be encoded between chunks of information. The other is domain-specific and defines the document structure of (in the current prototype) a medical report.

Often, XML is processed using XSLT. Indeed a simple XSLT style sheet can be used to convert XCH to linear text. However, to transform XCH into the interactive and animated Web application we are aiming for, we prefer to use a Web toolkit. We chose OpenLaszlo which since version 4 is no longer restricted to one particular run-time platform for content delivery (Flash) but can also compile source code into AJAX-like DHTML code which is natively supported by most current browsers. OpenLaszlo also makes it very easy to animate any of its visual components (widgets). Its XML-based programming language (LZX) is as declarative as XSLT but adds object-orientation and supports creating reusable class libraries.

Figure 6 shows an overview of the current system architecture. Its upper half depicts an existing NLG system modified to produce XCH output. Its lower half outlines the new presentation module which takes XCH as input. Here, report.lxz is the domain-specific part of the presentation module implementation, and cinehypertext.lzx is the domain-independent, library part. We expect that our emerging cinehypertext library and format will prove useful in other domains and systems.

5 Conclusion

This work aims to identify ways of presenting hypertext discourse which employ graphics to signal discourse structure in a systematic and principled way, by making articulate use of the space-temporal dimensions of the electronic medium. The work is part of a larger effort in natural language generation, aimed at producing different renditions of the same semantic content for different purposes and for different media. One of the novel aspects of our work is that we are generating ‘paraphrases’ that vary not just along the traditional dimensions (discourse, syntax, lexicalisation) but also in terms of graphical presentation (e.g., as textual reports in different styles – including linear vs. non-linear – or as slides for a presentation).

References

K. S. Campbell. 1995. *Coherence, continuity, and cohesion. Theoretical foundations for document design*. Erlbaum, Hillsdale, NJ.

L. M. Carter. 2000. Arguments in hypertext. A rhetorical approach. In *Proceedings 11^{th} ACM Conference on Hypertext and Hypermedia*, pages 87–91, New York. ACM.

C. Hallett, R. Power, and D. Scott. 2006. Summarisation and visualisation of e-health data repositories. In *UK E-Science All-Hands Meeting*, Nottingham, UK.

W. G. Koch. 2001. Jaques Bertin’s theory of graphics and its development and influence on multimedia cartography. *Information Design Journal*, 10:37–43.

M. C. Luesebrink. 1998. The moment in hypertext. In *Proceedings 9^{th} ACM conference on hypertext and hypermedia*, pages 106–112, New York. ACM.

C. Mancini, D. Scott, and S. Buckingham Shum. 2007. Visualising discourse coherence in non-linear documents. *Traitement Automatique des Langues, Special Issue on Computational Approaches to Document and Discourse*, 47(1). To appear.

C. Mancini. 2005. *Cinematic hypertext. Investigating a new paradigm*. IOS, Amsterdam.

R. Power, D. Scott, and N. Bouayad-Agha. 2003. Document structure. *Computational Linguistics*, 29:211–260.

K. Riley and F. Parker. 1998. Parallels between visual and textual processing. *IEEE Transactions on Professional Communication*, 41:175–185.

S. Williams, P. Piwek, and R. Power. 2007. Generating monologue and dialogue to present personalised medical information to patients. In *European Workshop on Natural Language Generation*, Prague. Submitted.