The Rhetoric of Generated and Adaptive Hypermedia

Lloyd Rutledge, Jim Davis, Jacco van Ossenbruggen and Lynda Hardman*

CWI (Centrum voor Wiskunde en Informatica)
P.O. Box 94079
NL-1090 GB Amsterdam, The Netherlands
Tel: +31 20 592 41 27
E-mail: Firstname{.van}.Lastname@cwi.nl

*CourseNet Systems
170 Capp St.
San Francisco, CA 94110, USA
Tel: +1 415-431-2250
E-mail: jrd3@alum.mit.edu

ABSTRACT

Meta-data of a document can guide the process of generating hypermedia presentations for the document. It can also guide the adaptation of a hypermedia presentation. One type of meta-data that can exist for a document is its rhetorical structure. Rhetoric is an important aspect of a document because it embodies how the author intends to convince the user of his or her message.

Much work has been done on analyzing existing documents to determine their rhetorical structure. This earlier work has been applied to the analysis of text as well as a variety of other media. In this paper, we apply the process of rhetoric-based analysis for media in general to the consideration of rhetoric-based synthesis of synchronized hypermedia presentations. Here synthesis is the semi-automatic generation of hypermedia presentations, and the process of making hypermedia presentations adaptive. Processing rhetorical structure in the automatic synthesis of presentation enables the author to represent rhetorical intent independent of presentation, and preserves the rhetorical intent through all variations in the document's presentation.

KEYWORDS: Rhetoric, adaptation, presentation generation, hypermedia

1. INTRODUCTION

The field of rhetoric studies how argumentation within texts can be made persuasive. Recently, rhetorical analysis has been applied to modern media such as film [5][8], hypertext and the hyperlinking aspects of multimedia [10]. Rhetoricians have developed specific types of rhetorical structure in terms of which communication can be defined. They have also described techniques and patterns with which such rhetoric is conveyed.

Most rhetorical research has focussed on the analysis of documents that were made entirely by human authors. With the advent of automatic processing for adaptive and generated presentation, the possibility opens for rhetoric to be one type of input to the automated synthesis of document presentation, rather than an output of analysis of documents. This allows authors to write their documents directly in terms of the desired rhetorical structure and let the details of how to present that structure be encoded and or processed at a later time. When applied to adaptive presentations, processing of rhetoric can ensure that the rhetorical structure is conveyed consistently with all adaptations of the presentation. When applied to generating presentations, the possibility opens for document readers to define their own tailored style for how rhetorical structure is to be conveyed.

There are several aspects of rhetorical study used in analyzing documents and presentations. Mann and Thompson have specified a frequently cited collection of rhetorical relations called Rhetorical Structure Theory (RST), which define the rhetorical structure a document can follow [11]. Grice has described implicature, a conversational mechanism whereby one may imply something without directly expressing it [7]. While we do not expect software systems to deliberately employ implicature in a positive manner (i.e., to hint at something obliquely) we do see a need for them to take pains to help users avoid *unintended* implicatures.

In this paper, we describe communicative devices for conveying rhetoric that are unique to synthesized hypermedia, just as devices specific to other media have been described in earlier work. We also discuss RST as a basis for guiding the presentation of a document.

2. HYPERMEDIA COMMUNICATIVE DEVICES FOR RHETORIC

The rhetorical analysis of a communication consists of breaking it down into its communicative acts, which are the units with which information is conveyed. Communicative devices are patterns of communicative acts that use or express a particular rhetoric. Since a hypermedia presentation is a form of communication, it can be broken down into a collection of communicative acts [12]. Furthermore, communicative devices exist that can be used in hypermedia presentations for conveying rhetoric. Understanding these hypermedia communicative devices is useful not just for analysis but also the synthesis of hypermedia presentations. This section describes some hypermedia communicative devices and how instances of these can be generated as part of hypermedia presentation synthesis.

2.1 Rhetoric and Communicative Devices of Hypermedia Compared to Those of Other Media

Media for communication include speech, text, graphics, film, comics and hypermedia. Much has been established for the rhetoric of speech and text, and also for the communicative devices used in them. Film theory has set communicative devices for film [5] and for the use of rhetoric in film [8]. Communicative devices and some uses of rhetoric have been established for graphics [18]. Communicative devices have also been established for comic strips [6][13], but this media has received comparatively little attention. Research has been done, however, on the use of comic strip presentation devices for other aspects of multimedia, such as spatial layout [19]. Communicative devices have been established for hypermedia [12], as have rhetoric-based guidelines for human authors of hypermedia [10].

Hypermedia, and in particular synthesized hypermedia, is distinguished from the other media by the following characteristics:

- Interaction
- Adaptivity
- Temporal and Spatial Structure

The hypertext community has extensively explored interaction as hypertext navigation and has established several interactive devices. Metric analysis of hypermedia navigational structure has been established as a basis for analysis and also, potentially, for synthesis [2]. Landow has explored rhetoric devices for hypermedia, with a focus on navigation [10]. His work has shown how certain patterns of navigational structure act as rhetorical communicative devices.

Adaptivity in hypermedia has also been widely researched [4]. In earlier work, we incorporate rhetoric with the adaptive hypermedia technique of presentation constraint systems [15]. An important aspect of good adaptation is that the intent of the author is consistently maintained and conveyed through all variations in the final presentation.

In addition to navigational structure, hypermedia presentations also have temporal and spatial structure. A simplified representation of this combinational structure is shown in Figure 1. Visual media objects are positioned spatially on the display window, as shown in Figure 1a. This spatial structure can borrow rhetoric-related communicative devices from print layout, some of which have been established by Reichenberger [14]. The timing of the presentation of individual media objects is coordinated with that of other objects, as shown in Figure 1b. There is also navigation structure, shown in Figure 1c, and the selection of media content.

The progressive combination of these structure types is shown in Figure 1d, and is then combined with navigation in Figure 1e. The issues of timing and its combination with layout and navigation are unique to hypermedia. In this paper, we introduce communicative devices and rhetorical issues that are unique to timing and these structural combinations.

2.2 Avoiding Accidental Implicature in Synthesized Hypermedia Presentations

Grice shows how communicated ideas can be implied as well as expressly conveyed [7]. His work shows how important it is to make sure that each communicative act conveys what the author intends, and that none convey an unintended message. These guidelines for human authors can also be applied to the process of automated hypermedia synthesis. This section applies the principles of Grician implicature to the synthesis of hypermedia presentations.

If an implicature were desired, it would be hard enough to generate a presentation conveying it. However, it would be still harder for an automatic process to avoid or detect and correct accidental occurrences of such implicature in presentations. One example of this is if alphabetical order is used and incidentally groups items in categories beyond the alphabet. For example, a screen display of a tour of Amsterdam presentation could display nothing but “coffeeshops” (Amsterdam’s marijuana cafes) because “coffeeshop” comes before “museum” and “restaurant” in alphabetical order. Nevertheless, the reader may accidently perceive that an importance of coffeeshops in particular was intended to be conveyed.

Landow [10] presents an example where authors using a software system that was technically unable to present color

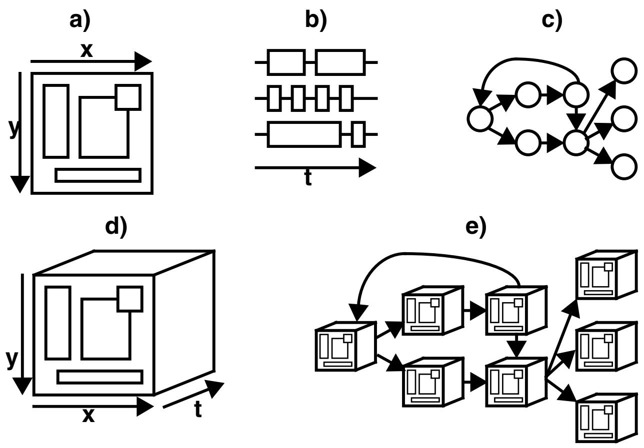

Figure 1: The Hypermedia Presentation Structure of a) Space, b) Time, c) Links, d) Space and Time Combined and e) Space, Time and Links Combined

images well chose to omit such images from their presentations, which led users to conclude that there was no color art relevant to the subject in question. To avoid such unwarranted implicatures, a software system would be well advised to use the device Grice terms OPTING OUT, where one makes an explicit mention of one's inability to present a piece of material.

In building semi-automatic presentation generators, it is important to account for the danger of unintended communicative acts. In encoding automatic processing, it is easier to make presentation structures that are helpful than to avoid presentation structures that are hurtful. One difficulty is that it is technologically impossible to represent all the conveyed semantics of media objects. It is therefore impossible to prevent adjacency of objects with semantics whose combination conveys undesired messages. It is much easier for a human author to detect and avoid unintended or implied messages than a computer.

There are some things a computer can do to avoid accidental implicature. One is to try to find as many undesired semantic implicatures as possible, perhaps with constraint systems one certain themes in media objects that should not be shown together. The more the computer can find, the better, but the programming of themes, related ontologies and undesired matches is labor-intensive and would only detect a small portion of undesired semantic implicature.

2.3 RST Constructs as Meta-data Input to Hypermedia Presentation Synthesis

Mann and Thompson's Rhetorical Structure Theory (RST) defines a frequently cited and used rhetorical structure that a communication can be described in terms of [11]. It is designed to represent the author's intended effect on the reader and the means by which the reader is convinced. The original intended scope of RST is that of a written monologue. RST consists primarily of *relations*, a list of which is provided in Figure 2. The units that make up members of RST relations are typically sentence clauses.

For the purposes of this paper, we assume that, for the most part, the same RST constructs that apply text apply also to hypermedia. This assumption has been made in related work [12]. It is debatable, and arguing it could be the subject of future work. Nonetheless, some variations of RST do occur in the course of this paper that apply differently to hypermedia than to text, and these are discussed herein. However, this paper focuses not on what rhetorical constructs exist for hypermedia, but how rhetorical constructs should be presented with hypermedia. The means of conveying rhetoric with hypermedia are much different than that of text. Furthermore, this paper explores not just the presentation of RST with non-text media and hypermedia, but also the synthesis of new presentations from RST, rather than just the analysis of existing ones.

	Multi-nuclear Relations	
Sequence	Constrast	Joint
	Nucleus-satellite Relations	
Evidence	Circumstance	Restatement
Concession	Background	Antithesis
Elaboration	Volitional Cause	Solutionhood
Motivation	Non-volitional Cause	Enablement
Condition	Volitional Result	Purpose
Evaluation	Non-volitional Result	Interpretation
Justify	Otherwise	Summary

Figure 2: The Relations of RST [11]

In the remaining sections of this paper, we present hypermedia communicative devices for the different types of RST relations. These devices are listed with the relations they convey and the types of hypermedia presentation structure they use at the end of this paper in Table 1. These devices are not universal or required for all presentations. Often, individual presentations or sets of presentations will establish their own devices for conveying rhetorical structure. These can be in addition to or override the devices described in this paper.

2.4 Automated Hypermedia Synthesis

For the purposes of this paper, we consider automated hypermedia synthesis to be the process by which a computer creates or affects a hypermedia presentation. We break automated hypermedia synthesis into two categories: generation and adaptation. *Generation* is when a computer program automatically creates a hypermedia presentation from a source document or data store. A final presentation, whether written by a human or automatically generated, must contain all information necessary for the presentation to adapt to varying circumstances. *Adaptation* is the process by which a presentation renderer processes this information to determine how to adjust the presentation to these varying circumstances that were not determinable when the encoding was generated.

3. HYPERMEDIA COMMUNICATIVE DEVICES FOR CONVEYING SEQUENCES

The RST relation for which it is easiest to find hypermedia communicative devices is, perhaps, the *sequence*. Sequence is a multi-nuclear relation, meaning that it has any number of components, none of which is more important than any other. What distinguishes the sequence from the other RST multi-nuclear relations is that its members have an ordering that is significant. The reason that sequence is so easy to convey is that there are several aspects of the hypermedia presentation interface that are strongly and inherently sequential themselves, and thus readily lend themselves to conveying rhetorical sequences.

The sequence by itself does not have strong rhetoric significance. Other relations, particularly the nucleus-satellite relations such as *evidence* and *justify*, have a more specific meaning for how the author intends to convince the user. However, the ordering of these more specific relations is often an important factor, representing how the author wishes the argumentation to progress. The sequence provides this function. Moreover, at times the focus is on the ordering itself, rather that on the contents of the ordered sequence.

3.1 Devices from Each Type of Hypermedia Presentation Structure

The spatial structure of hypermedia presentations has an often-used device for communicating a sequence: *bookshelf order*, the arranging of visual displays from left to right and then top to bottom. Our application Fiets, which provides a hypermedia tour of Amsterdam, uses the bookshelf order device for some of its screen displays [16]. Figure 3 shows Fiets' use of the device for showing houses along the Herengracht in Amsterdam and how their architecture changes as one progresses along the canal. The houses need to be shown in order of their street number in order for this concept to be presented. Here, the sequence of their physical placement along the canal is conveyed using bookshelf order.

A variation of the bookshelf order device's use of spatial structure is the *manga* device [19]. This is a technique adapted from comic strips. It is primarily bookshelf order, but allows some vertical stacking within rows. Thus, within the left-right progression of a row there will sometimes be two or more objects stacked vertically. For such a stack, the reading order switches to top-down for the items within it and then continues from left to right after the stack. This allows for more efficient spatial placement of objects that are small enough to stack within a row, but still maintains a readily perceivable sequence.

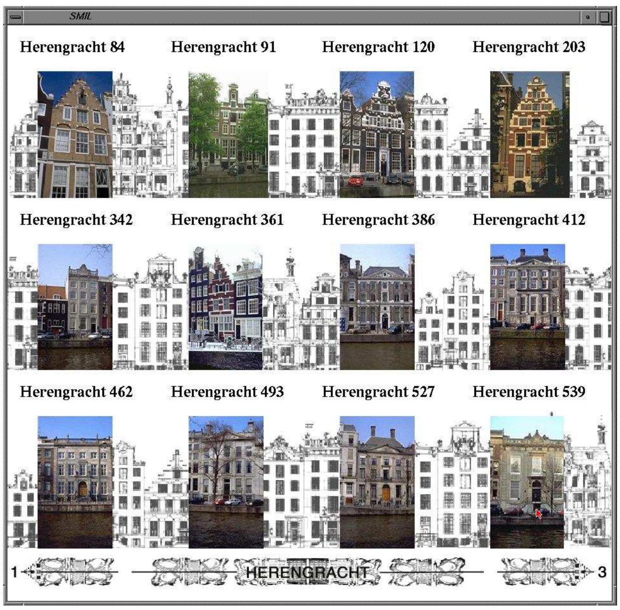

Figure 3: A Fiets Screen Display

Time is, in terms of human perception, intrinsically a sequence. Seeing one item displayed after another conveys to the human perceiver a strong sense of sequence. This *temporal sequence* device was used in several presentations in Fiets. The same street sequence conveyed spatially in Figure 3 is conveyed temporally in an alternative mode in Fiets, with each house shown one a time in order of address.

Navigational structure can also be used to communicate a sequence. One common link-based device is the use of *previous/next buttons*. With this technique, one item of a sequence is presented, along with the starting points for links to the previous and next items in the sequence. The resulting navigation structure is that of a chain, with each adjacent node connected by two links, one in each direction.

The use of some devices described in this paper add media objects to the presentation beyond those of the underlying content, for example, navigation icons. This device also affects the media content selected or generated for integration into the presentation.

3.2 Comparing and Combining these Devices

There are trade-offs in choosing among these devices. With the bookshelf order technique, the entire sequence is displayed at once, but there is little visual space for showing each item. Presenting the same information with the temporal sequence device takes more time, of course, but more data can be given for each item. Using previous/next buttons has the same display space advantage as the temporal sequence device, though perhaps with little less room due to the display of the buttons. This technique also takes more time than bookshelf order, but with it, the user controls the pacing and direction of the presentation.

These devices can also be used in combination to convey a sequence. With bookshelf order, for example, there may be only so many items that can be shown in one display. Where there is overflow, additional displays are used. These multiple displays can be sequenced temporally or with previous/next buttons. We describe in more detail the use of presentation constraint systems to handle this type of overflow of bookshelf layout to temporal sequences in earlier work [15]. When using these types of device combination, a balance can be set based on how much display space is desired for each item in a sequence. If less space is desired, then more bookshelf order can be used. This minimizes the number of display screens, and thus also the time involved, and perhaps the number of button presses as well. If more space is desired for each item, then the temporal sequence or link chain becomes longer.

A combination of these devices is used in Fiets, as shown in Figure 3. Here, one screen display of multiple items is shown, with next and previous buttons on the bottom for seeing the group of items that precedes or follows those current shown in the sequence. Numbers are shown with the buttons to show which displays are shown previously and next. Also in this Fiets display, the next display is shown after a certain amount of time passes, thus handling the overflow of the spatial display with a combination of temporal and navigation devices.

3.3 Strict and Loose Sequences

In analyzing existing text for sequences, it suffices to say that a sequence exists. In synthesizing new hypermedia from sequences, a sequence has two ways of affecting the presentation. One is with *strict* interpretation, which means that the items must be presented to the user in that order. The other is *loose* interpretation, which means that the user needs only to know what the sequence is, but can access the items in it in any order. In earlier work, we discussed this distinction between strict and loose RST sequences when applied to constraint systems for hypermedia presentations [15].

With the Fiets display shown in Figure 3, the sequence is loose. The user should understand that there is an ordering of the buildings that is significant, but the buildings do not necessarily need to be presented in that order. With educational hypermedia, on the other hand, authors do often want a strict ordering of presentation to be maintained for portions of the application, even while other portions are more freely adapted [3]. For these, strict sequences would be more appropriate.

A device for communicating a loose sequence is the *page index*. This uses a combination of bookshelf order and navigation structure. It shows one display of one or, in bookshelf order, multiple items of the sequence. A bookshelf-ordered sequence of display numbers is also shown on the display, typically along the bottom. The addition of this numeric text affects the presentation integrated media content. Each number displayed can be clicked to go the display of the sequence corresponding to that number. The number of the current display is typically highlighted or otherwise visually distinguished. Previous and next buttons are often displayed with the page index. As such, the page index device conveys to the user what the order of the sequence is, but allows the items to be accessed in any order the user chooses. This same device is used by thumbnail browsers, such as the widely used Acrobat™ reader.

The use of previous and next buttons without the page index presents a stricter sequence since the user typically starts at the beginning and cannot jump ahead. Removing the previous button makes the relation between rhetorical sequence and presentation order even stricter.

A variation of the page index device is the use of a *central node*. Typically, the page index is shown with one or more of the sequence members in each temporally or navigationally sequenced display. Alternatively, one central node for the sequence can be displayed that does not show any information of any member but instead shows a page index to the member items. When a member link is activated, the resulting display shows information on that member with a link back to the central node. The central node variant conveys the sequence even more loosely than the typical page index method because the user never perceives the members sequentially — the central node is always seen between the displays of any two members. The central node technique may also be desired when the information about each member is not considered important for the user to know.

3.4 Conveying the Extent of a Sequence

A second aspect of a sequence that is sometimes significant is its length. This can be important in its own right (as in there are exactly three primary colors, not more or less), and to allow the user to estimate the remaining presentation time. It can also counter false implicature, as in the example presented earlier, where alphabetical presentation combined with a limited screen size, misled the user into believing that coffeeshops were all Amsterdam had to offer. (The user may still infer an unintended significance to the ordering, taking the alphabetical order to bear meaning.) Proportional scrollbars serve this function well.

The page index is a clear communicative device for denoting the extent of a sequence. By highlighting the current page in the index, the user can track the presentation time involved in conveying the whole sequence. When link structure is not used to present a sequence, a piece of text saying the equivalent of “page *x* of *n*”, letting the user track the sequence’s progression even when not being able necessarily to control it. This device is easy to program the creation of in presentation generation systems.

There are also hypermedia communicative devices for conveying the end of a sequence. When previous/next buttons are used, the final page will often have the next button ghosted out. When a sequence is shown using multi-page bookshelf order and the total number of sequence members is not divisible by the number of members shown on each page, the last page shown will not be full. The empty space on this page to the right and bottom of the last member shown communicates that this page is the last shown. When a temporal sequence is used, each page will typically be shown for the same amount of time, except for the last, which stays on the screen. This breaking of the rhythm of presentation the last page brings communicates its finality. It is also easy to program these devices into presentation generation systems.

3.5 Presentational Sequences vs. Rhetorical Sequences

Often, a sequence is inherent in the presentation structure without conveying any sequence relation in the rhetorical structure. While this is arguably best avoided where possible, it is often unavoidable and seems not to cause problems for readers. Printed text, for example, is perceptually sequential, and yet readers use semantics, context or common sense to tell the difference. Often devices that are specific to a particular collection of documents are developed that use presentation sequences but do not convey rhetoric sequences. Users learn what these devices mean and readily avoid confusing them with rhetorical sequences.

3.6 Prerequisites as Input for Generating Dynamic Sequences

The progression of printed text conveying rhetorical structure is linear. Thus, the matching of a rhetorical sequence to a sequence of text phrases is straightforward. Hypermedia presentation, on the other hand, is not always linear. The progression of media presented can vary with user selection. Relating rhetorical sequences to the varying sequences of hypermedia is thus more complicated than relating it to fixed sequences of text. This necessitates a more complex variation of rhetorical sequence when applied to hypermedia, which we call *prerequisites*.

We base the processing of these prerequisites on that of the AHAM system [3]. In AHAM, the user is allowed to navigate to certain parts of the presentation only if certain other parts, the prerequisites, have already been presented. When applied to authoring rhetorical structure from which to synthesize hypermedia, it provides the representation in the document of information that allows for more varied final presentation. This increased variation is the larger number of sequential navigational paths that can be presented to the user.

4. HYPERMEDIA COMMUNICATIVE DEVICES FOR CONVEYING OTHER MULTI-NUCLEAR RELATIONS

Presentation must also distinguish between the different *multi-nuclear relations*. Section 3 discusses sequence, and hypermedia communicative devices for conveying it. This section discusses the other two multi-nuclear RST relations *joint* and *contrast*, and presents hypermedia communicative devices for conveying them. The impact of multi-nuclear relations in general on synthesizing hypermedia is also explored.

4.1 Joint

The *joint* relation indicates that its members form a group but that there is no significant ordering among these members in terms of rhetorical progression. While there are several aspects of hypermedia presentation structure that are inherently sequential, and thus readily communicate the rhetorical relation of sequence, it is harder to conceive of an aspect of presentation structure that is inherently unordered and avoids implying a sequence.

The example of a rhetorical joint from the Fiets application is the collective presentation of Amsterdam buildings built in the Louis XIV style. It is important to show all the buildings to communicate to the user a sense of what the Louse XIV style is. However, the order in which these building is presented does not affect how well this sense of style is conveyed, nor is there any sense in which one building is "more Louis XIV"-ish than any other.

Time is inherently ordered. This presents a danger of accidental implicature because members of a rhetorical joint shown sequentially in time may be perceived as a rhetorical sequence. Thus, the *simultaneous display* of objects is the use of temporal structure that best conveys joint, or at least best avoids having them perceived as in a sequence.

There will be times, of course, when all members of a joint cannot be displayed simultaneously, and thus temporal sequencing may be desired. When displaying members of a joint in a temporal sequence, it is important to establish an *exception* in how the user perceives the temporal sequence. The user should be made to understand that the order of the members' presentation is not significant. Often, the user can tell that there is no meaning to the order. However, there is the danger in presentation synthesis that a perceivable semantic order appears that matches the presentation order. In these cases, the user will think that the author intended that semantic order, and may thus perceive a different message than the author intended. (Note for example the devices that authors of multiple-authored papers employ to indicate the relation, or lack there of, between the order of names and the relative contributions or importance of those authors.)

Similarly, navigational structure has a strong impact on perceived order, since the order in which the user activates links determines the order in which portions of the presentation are presented. Navigational structure can avoid conveying unintended sequential order by providing equal access to all components of a joint, and by making sure no members of the joint are shown sequentially. The *central node* device described for conveying loose sequences in Section 3.3 applies here as well. A link to each member is shown simultaneously on one page. The user can activate a link to see the corresponding joint member, but must traverse back to the central node before seeing another member. This prevents any one member from being seen directly after another in sequence, and thus reduces the risk of implying a rhetorical sequence.

The best use of spatial structure to communicate a joint is to avoid any structures that convey a sequence. Bookshelf order is one such pattern to avoid, since it strongly conveys a sequence. *Random arrangement* of the items on the screen is the most clear means of using spatial structure to avoid convey a sequence. This device, however, can waste much screen space. The *patchwork* device for spatial layout finds more efficient means of playing the visual objects in a joint. It, however, can sometimes incidentally align in a pattern resembling bookshelf order. A problem that arises with both random and patchwork arrangement is that they can leave the screen looking cluttered and unappealing. The *grid* arrangement is a more orderly-appearing device that conveys sequentiality less than bookshelf, but still more so than the random and patchwork arrangements.

Because the time-based and link-based devices presented above both rely on simultaneous display of objects, both rely on spatial structure to handle this simultaneous display. Whether items are shown simultaneously in their entirety, or simply links to the joint items are shown simultaneously, making the spatial structure of their display random will enforce the perception of a joint rhetorical relation and avoid implying a sequential rhetorical relation. Thus, both the simultaneous display and central node devices end up being combined with the random arrangement device.

4.2 Contrast

A *contrast* multi-nuclear relation establishes that its members are intended to be compared with one another. Since no ordering among these members is considered significant, the communicative devices for the joint relation apply to contrast as well. Communicative devices that apply particularly to contrast are those that facilitate comparison.

Members of contrast relations will often have components with corresponding components, or fields, in other members. For example, in comparing Amsterdam buildings, each might have a year of construction, an address, an architect and various images such as those of the full building, the doorway and the gable. In presenting these buildings as a contrast, it helps the user to compare corresponding fields one at a time. The devices we present for the contrast relation all facilitate field-by-field comparison.

The most important way to facilitate comparison is to have the corresponding components of each contrast member displayed in visually corresponding areas of the screen at the same time. For example, if each Amsterdam building in a contrast is shown in a column, then corresponding fields could be along the same row. This is the concept of *parallelism*, described by Tufte [18]. In the context of hypermedia, we call this use of spatial structure to maintain parallelism the *symmetric layout* device.

The *consistent fields* device affects the media content selected. It facilitates comparison by ensuring that media objects for as many corresponding fields as possible are selected. Furthermore, when alternative media objects can be selected for individual fields, the best choice is that which provides the best comparison between corresponding fields in different members. For example, if images of different aspect ratios are available for the gables of several Amsterdam buildings in a contrast, then the best choice is the one in which the aspect ratios of the images are closest to each other.

If the members of a contrast relation collectively have too much information to display on the screen at once, then multiple screen displays for this information must be made that are separated by temporal structure, linking structure or both. In order to facilitate comparison under these circumstances, the presentation of each member should be *consistent* in its presentation structure. When each member is shown on a separate screen display, the same spatial layout should be used for showing corresponding information for each member.

Sometimes the information to present for each member may also be too much to show on one display, and then individual members must be broken up with temporal and navigational structure. If each member is broken up into a temporal sequence, then unifying the timing of each with *consistent pacing* will help the user identify corresponding components of each member to compare. Similarly, the device of consistent linking should be used if navigational structure breaks up individual members into multiple displays. For example, if a central node is used to access components of a member, then the information displayed in the central nodes for all members should be placed uniformly in same spatial layout.

Although the techniques of consistent pacing and timing facilitate the comparison of corresponding information from contrast members, the lack of simultaneous display of this corresponding information — the lack of parallelism — is still disadvantageous. The devices of symmetric layout and consistent timing and pacing can be combined into a device that we call *parallel presentation*. The parallel presentation device shows corresponding portions of all the contrast members on the screen at once. This provides the easy simultaneous comparison of corresponding fields provided by symmetric layout. Subsequent screen displays show other corresponding portions of the same contrast members. The display of other screen displays is triggered either by the passage of time, with temporal structure, or by the user activating links, with navigational structure.

4.3 Communicating the Hierarchical Structure Defined by Multi-nuclear Relations

The members of any RST relation can be themselves multi-nuclear relations. This enables document rhetorical structure to be hierarchical. For example, the satellite in an evidence relation may be a sequence of logical steps supporting the relation's nucleus. Another example, from the Fiets application, is that the Louis XIV and Louis XV architectural styles may be members in a contrast relation. These members may in turn consist of a joint relation of images of sample buildings from the architectural period. A multi-nuclear relation may also contain in turn multi-nuclear relations of the same type. An example in printed text is the multi-layered numbered item list, which is the display of sequences containing other sequences.

The hierarchical grouping defined by multi-nuclear relations has an impact on synthesized hypermedia presentation structure. It affects how grouping is conveyed in the presentation. It also affects how much and what to show simultaneously, and how to break the content displays apart with temporal or navigational structure.

Visual grouping is a device for conveying hierarchical structure. Screen display items that are visually distinguishable as being located together will be perceived as belonging to the same group. This technique for conveying RST hierarchical structure with spatial layout is described in other work [14]. Multiple layers of a hierarchy can be conveyed on a single display by having visually discernible groups being placed, in turn, in visually discernible groups.

5. HYPERMEDIA COMMUNICATIVE DEVICES FOR CONVEYING NUCLEUS-SATELLITE RELATIONS

Nucleus-satellite relations are binary relations where one node, the nucleus, is of greater significance than the other, the satellite. Nuclear-satellite relations form the greater portion of the RST relations. An example of a nucleus-satellite relation is given in Figure 4. Typically, the nucleus of a nucleus-satellite relation cannot be removed from a presentation without making the rhetorical flow of the presentation incoherent. Deleting the satellite, however, does not interrupt the rhetorical flow [11].

5.1 Different Types of Nucleus-Satellite Relations

There is little theoretical justification to support different renderings for different nuclear-satellite relations. Although in some applications, it may be useful to establish conventions for marking such relations (e.g., that an Evidence relation be conveyed by a red arc, a Elaboration by a blue one, and so on.), in general, individual presentations and presentation sets will establish their own devices that users learn to identify the different relation types. Furthermore, presentations will often use media objects that convey particular rhetorics. The best example of this comes from text, with phrases like “for example” and “in other words”. But although there are no universal presentation devices for the individual nucleus-satellite relation types, there are still some devices that apply to nucleus-satellite relations in general. These are discussed in the following subsections.

5.2 Conveying similarity of relation.

When two sections of the material have the same underlying rhetoric structure, this can motivate using a parallel presentation structure. Reichenberger [14] shows how to exploit this in laying out text and graphics. It can be used as an intended device, and must be prevented as accidental implicature. This prevention can be ensured for the most part with presentation constraint systems because this type of implicature is conveyed with presentation structure alone and does not necessarily involve the semantics of the document objects involved.

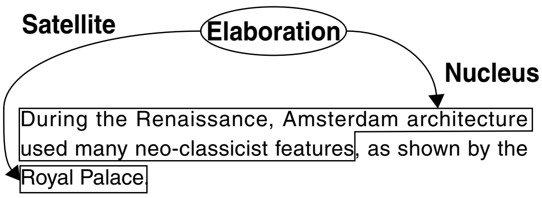

Figure 4: An *Elaboration* Nucleus-Satellite Relation

5.3 Relating Satellites with their Nuclei

One principle that affects these devices is that the presentation structure must associate a nucleus with its satellite, and vice versa. Reichenberger has stated that in laying out the spatial display of nucleus-satellite relations each nucleus may be positioned adjacent to its satellite [14]. We refer to this as the *adjacency* device.

Another principle is that since these relations are directional, the nucleus in each should be distinguishable from the satellite in the presentation structure. We have not determined a general technique for doing this. Within RST, there is no inherent order between a satellite and nucleus. The same pattern occurs in hypermedia presentation structure. While nuclei and satellites are typically displayed adjacently in layout, there is no overriding pattern for one to appear above or to one side of the other. There is also no overriding pattern for one to appear before the other in time or in the progression of navigation. Individual presentations and presentation sets will establish for themselves how nuclei are distinguished from satellites, and the user will come to learn these patterns.

5.4 Using Satellite Expendability

The expendability of satellites can be used to adapt the presentation by making it smaller when necessary. Making it smaller can mean having fewer objects on the screen, taking up less bandwidth, or ending within a certain period of time. The simplest way to use satellite expendability is not to show a satellite at all. We call this method of removing satellites altogether to control the size of the whole presentation *pruning*.

Alternatively, linking presentation structure can be used to make showing the satellite an optional choice of the user. Here, the nucleus is shown as a whole, and a link is provided to the satellite. The user can either access the satellite or move on without seeing it. We call this process *hiding* — the satellites are not removed, instead they are hidden but accessible.

Some types of satellites are more expendable than others are. The satellites of the relations *summary* and *restatement* typically provide no new information to the user; they just make information presented elsewhere easier to comprehend. Thus, when a presentation needs to be made smaller, the satellites of these types of relations can be pruned or hidden first. Then, if more shrinking is needed, other types of satellites can be removed or hidden as well.

6. CONCLUDING REMARKS

6.1 Implementation Issues and Design

The adaptation of presentation using some of the communicative devices presented in this paper was described in earlier work [15]. This earlier work introduces *inter-dimensional constraints* as a new type of constraint that extends adaptivity in a way that applies well to rhetoric and similar meta-data. These provide an adaptive encoding of hypermedia presentations that maintains the intended rhetorical structure through more widely varying adaptation than constraints only on either temporal or spatial structure provide.

The rhetorical structures discussed in this paper can be used as input for hypermedia generation as well. The hypermedia communicative devices described in this paper would be the output from processes using these rhetorical structures. The mapping directing this processing from rhetorical structures to hypermedia presentation can itself be a separate encoding. The XML and SGML communities refer to such mappings from source document to final presentation as *style sheets*. With this separation of rhetoric from the style of its presentation, the user is free to choose among different styles for processing the same rhetoric, and authors can concentrate on the rhetoric structure and, ideally, trust that it will be presented properly under all varying presentation circumstances. Our Berlage environment processes source XML documents with style sheets to generate hypermedia presentations [17]. These output presentations are in the XML Web multimedia format SMIL [9].

6.2 Future Work

We are extending the types of meta-data put into the hypermedia synthesis process, both adaptive and generated, by exploring other types of literary analysis beyond rhetoric such as narrative. These other types of meta-data we are exploring also come from the study of other media, such as montage in film theory. Finally, we are investigating the use of rhetorical structure as input to a generation process instead of just an adaptive process implemented with constraint systems. This investigation involves extending our Berlage hypermedia presentation generation environment to process rhetorical structure as input.

6.3 Summary

This paper took earlier work in the rhetorical analysis of text and other media and applied it to the use of rhetorical structure in synthesizing hypermedia presentations. The use of hypermedia communicative devices for conveying the different RST constructs was explored, as summarized in Table 1. Types of rhetorical structure that apply uniquely to synthesized hypermedia were introduced, such as strict and loose sequences, prerequisite-based sequences and multi-layered hierarchical grouping. The use of these devices in both adaptive and generated hypermedia was explored.

	Types of RST Relations				
	Multi-nuclear				**Nucleus-Satellite**
Types of Hypermedia Presentation Structure	**Sequence**	**Joint**	**Contrast**		
Space	visual grouping				adjacency
	bookshelf order	grid	symmetric layout		
		patchwork			
	manga	random			
Time	temporal sequence	*avoid or establish exception*	consistent pacing	parallel presentation	*none given*
Link	next/prev buttons	central node	consistent links	parallel presentation	hiding
	page index				
Content	icons and labels for link starting points		consistent fields		pruning

Table 1: Hypermedia Communicative Devices from Each Presentation Structure Type for Different RST Relations

6.4 Acknowledgments

Jim Davis performed his research for this paper during a six-month visit with the CWI. The research performed at the CWI was funded in part by the Multimedia Information Analysis (MIA) project and by the ToKeN2000 project. Brian Bailey programmed the initial implementation applied to this work. Frank Cornelissen and Joost Geurts have also worked on the implementation. Cathleen Damplo Schaad, Laura Mitchell Foley and Scott Melchionda provided valuable information on comic strips and graphic design. Many images in Fiets come from the Amsterdam Heritage Website [1]. Finally, we thank our colleagues Frank Cornelissen, Joost Geurts, Craig Lindley, Frank Nack and Steven Pemberton for their valuable conversations and insights.

7. REFERENCES

1. City of Amsterdam Municipal Department for Preservation and Restoration of Historic Buildings and Sites. Amsterdam Heritage, http://www.Amsterdam.nl/bmz/adam/adam_e.html.

2. Botafogo, R.A., Rivlin, E. and Shneiderman, B, “Structural Analysis of Hypertexts: Identifying Hierarchies and Useful Metrics”, *ACM Transactions on Information Systems*, Vol. 10, No. 2, April 1992, pp. 142-180.

3. De Bra, P., Houben, G.-J. and Wu, H. “AHAM: A Dexter-based Reference Model for Adaptive Hypermedia”, *Proceedings of ACM Hypertext 99,* Darmstadt, Germany, February 1999, pp. 147-156.

4. Brusilovsky, P. “Methods and Techniques of Adaptive Hypermedia.” *Journal on User Modeling and User-Adapted Interaction*, Vol 6, 1996, pp. 87-129.

5. Eisenstein, S.M., *S.M. Eisenstein: Selected Works*, (Taylor, R. ed.), British Film Institute, 1991.

6. Eisner, W., *Comics and Sequential Art*, Poorhouse, 1985.

7. Grice, H. P. “Logic and conversation” In: *Syntax and Semantics* (Cole, P and Morgan, J.L. eds.), Vol. 3, Speech Acts, Academic Press, New York, 41-58, 1995.

8. Harrington, J., *The Rhetoric of Film*, Holt, Rinehart and Winston, Inc., New York, 1973.

9. Hoschka, P. (ed.). *Synchronized Multimedia Integration Language (SMIL)*, World Wide Web Consortium Recommendation. June 1998.

10. Landow, G.P. “The Rhetoric of Hypermedia: Some Rules for Authors”, , http://www.w3.org/TR/REC-smil/*Hypermedia and Literacy Studies*, (Delany, P. and Landow, G.P. eds.), MIT Press, May 1994, pp. 81-103.

11. Mann, W.C., Mattheissen, C.M.I.M and Thompson, S.A. “Rhetorical Structure Theory and Text Analysis”, *Information Sciences Institute Research Report*, ISI/RR-89-242, November 1989.

12. Maybury, M.T. “Planning Multimedia Explanations Using Communicative Acts”, in *Intelligent Multimedia Interfaces* (ed. Maybury, M.T.), AAAI Press / The MIT Press, 1993, pp. 60-74.

13. McCloud, S. *Understanding Comics: The Invisible Art*, Harper Perennial, 1994.

14. Reichenberger, K., Rondhuis, K.J., Kleinz, J. and Bateman, J. “Effective Presentation of Information Through Page Layout: a Linguistically-Based Approach”, *Proceedings of the ACM Workshop on Effective Abstractions in Multimedia*, San Francisco, USA, November 1995.

15. Rutledge, L., Bailey, B., van Ossenbruggen, J., and Hardman, L. “Rhetorical Structure and Hypermedia Presentation Constraints”, *Proceedings of ACM Hypertext 2000*, San Antonio, Texas, USA, June, 2000. (accepted)

16. Rutledge, L., Hardman, L., van Ossenbruggen, J. and Bulterman, D.C.A. “Structural Distinctions Between Hypermedia Storage and Presentation”, *Proceedings of ACM Multimedia 98*, Bristol, England, September, 1998, pp. 145-150.

17. Rutledge, L., van Ossenbruggen, J., Hardman, L. and Bulterman, D.C.A. “Practical Application of Existing Hypermedia Standards and Tools”, *Proceedings of Digital Libraries 98*, Pittsburgh, USA, June 1998, pp. 191-199.

18. Tufte, E.R., *Visual Explanations: Images, Quantities Evidence and Narrative*, Graphics Press, 1997.

19. Uchihashi, S., Foote, J., Girgensohn, A. and Boreczky, J. “Video Manga: Generating Semantically Meaningful Video Summaries, *Proceedings of ACM Multimedia 99*, Orlando, Florida, USA, Nov. 1999, pp. 383-392.

20. Weitzman, L., and Wittenberg, K. “Automatic Presentation of Multimedia Documents Using Relational Grammars”, *Proceedings of ACM Multimedia 94*, San Francisco, USA, September 1994, pp. 443-451.