ADAPTIVE PRESENTATION STYLES FOR DYNAMIC HYPERMEDIA SCRIPTS

BY

MICHELLE ANN FINEBLUM

Bachelor of Arts, Swarthmore College
(1980)

Submitted to the Media Arts and Sciences Section
in Partial Fulfillment of the Requirements for
the Degree of

MASTER OF SCIENCE
in Visual Studies

at the Massachusetts Institute of Technology

September, 1991

Signature of Author

Michelle A. Fineblum
Media Arts and Sciences
August 12, 1991

Certified by

Muriel R. Cooper
Professor of Visual Studies
Thesis Supervisor

Accepted by

Stephen Benton, Chairman
Departmental Committee on Graduate Studies

Adaptive Presentation Styles for Dynamic Hypermedia Scripts

by Michelle A. Fineblum

Submitted to the Media Arts and Sciences section on August 12, 1991 in partial fulfillment of the requirements for the Degree of Master of Science in Visual Studies.

Abstract

Most research in hypermedia has focused on navigation and links, but does not adapt the presentation of information at each node to the individual needs of the user. The described prototype, called HyperAdaptive, addresses this lack by supporting the adaptation of dynamic multimedia presentations within a hypermedia application. The style of the presentation varies both in content (selection, ordering and emphasis among components) and form (visual emphasis, coordination of various media, and layout). HyperAdaptive automatically generates a multimedia presentation by using knowledge-based representations of different presentation styles as guides in its search through a database of richly described multimedia data. The underlying representations of presentation styles also guide the multimedia design decisions.

Thesis Supervisor: Muriel Cooper
Title: Professor of Visual Studies

This work was supported in part by grants from the Hewlett Packard Company, NYNEX Corporation and PAWS, Inc.

Acknowledgements

Many people have contributed to this work, and have helped to make my two years at the VLW exciting and formative.

The faculty and staff of the VLW and the Media Lab have created a lively space for interdisciplinary research. I am grateful to my advisor, Muriel Cooper, for all her support, keen observations, and great ideas. I want to thank Henry Lieberman, Ron MacNeil and Patrick Purcell for two years of thoughtful feedback and direction, and for helpful reviews of this thesis. David Small's inventiveness, warmth, hard work and visual wit have inspired me. Anne Russell, Linda Peterson, and Marie Sobalvarro provided critical support and many good talks.

My fellow VLW grad students, David Young, Laura Scholl, Bob Sabiston, Grace Colby, Steve Librande and Didier Bardon, have greatly enriched my experience with their good company and creativity. "Merci bien" to Didier for countless hours of graphic design coaching and discussion. I am indebted to Laura Robin for building a robust system, for being such a good teacher and for being a delight to work with. "Hats off" to Robin Kullberg, who wrote an elegant program to support hypermedia links. M-lab buddies Laura Teodosio, Molly Bancroft, Janet Cahn, Debby Hindus, Stan Sclaroff, and Ben Lowengard deserve high praise for providing such high quality distraction.

Two research groups played an important role in shaping many of the ideas in this thesis. The Media Lab's Narrative Intelligence study group introduced me to the application of narrative and rhetoric to user interface design. I am especially grateful to Edith Ackermann, Marc Davis and Alan

Ruttenberg for their time, ideas, and enthusiasm. Barbara Grosz and her AI research group graciously opened their meetings to me, and provided insightful feedback on my research proposal. Thanks to Joe Marks for his ideas about knowledge-based graphics and the commonalities underlying linguistic and graphical communication.

I am indebted to Wendy Mackay for first suggesting I come to the Media Lab, and for teaching me so much over the past ten years about user interfaces, adaptive software, and multimedia. I am grateful to John McDermott for his steadfast support, enthusiasm, and illuminating feedback.

I am most grateful to my parents, Carol and Sol, and sisters, Deborah, Tobi, Susan and Wendy, for their love, creativity and humor. Finally, I thank my incredible Michael: I'm very happy to be travelling through life with you.

Table of Contents

I. Introduction

A. Problem and Motivation

Hypertext and hypermedia are quickly becoming dominant paradigms for computer-based information systems. The idea that one can click on a region of interest for more information on that subject is no longer surprising to users. This simple notion's popularity, in all its variations, is based on its intuitiveness and the role it serves in human-computer communication.

However, hypermedia as it is implemented to date fails to exploit the computer's potential for dynamics and flexibility. Although important progress has been made in enhancing the navigational process (Meyrowitz, 1989; Zellweger, 1989), the information presented at each node tends to ignore the fact that different users would benefit from different presentations of the information. Most of the current systems, in this respect, are too similar to their print predecessors -- the encyclopedia, the book and the newspaper.

The current thesis project adapts both the content and form of a multimedia presentation viewed within a hypermedia application. As the user moves through a hypermedia application, a dynamic multimedia presentation is shown at each node. Such a presentation is similar to a multi-format slide show, but is displayed on the computer screen (with audio) and can take advantage of the computer's power to adapt to changing conditions and user interaction.

The changes in content and form are made to suit the user's informational needs and preferences. The content decisions include the inclusion and exclusion of pieces of information, as well as the ordering and emphasis of information presented. The form decisions include the graphical emphasis given to the components of the presentation, the layout, and the coordination of the various media (voice, static and dynamic images and text). These multimedia design decisions are made to reinforce the priorities and relationships in the content.

To clarify the notion of adaptivity in hypermedia scripts, consider how expert designers and writers might create two multimedia presentations on a museum exhibit: both presentations give an overview of the exhibit, but one is intended for a lay person, and the other for students of art history. Extra explanations might be added in the lay person's presentation, which the art student already understands. The creators would avoid specialized vocabulary in the lay person's presentation, or be sure to define any jargon used. They will feel free to make more obscure references to other artists and art scholarship in the presentation for the art history students. More fundamental than these local modifications to the presentations, the entire structure of the presentation will be different for the two audiences -- each story has its own coherence, ordering and flow. The designer will reinforce the overall structure and emphasis of each presentation with the use of multimedia design techniques, which we define as graphic design expanded to include sound and dynamics.

In this project, the content's subject matter is research at the Visible Language Workshop. The system automatically generates multimedia presentations at run-time based on the user's style selections. The user specifies whether she is a researcher or a non-researcher, and whether she is technical or non-technical. The system has a knowledge-based representation of presentation styles linked to these user types, along with a pool of richly described multimedia data. The system conducts a recursive search through the database, using the knowledge of the selected styles as guides to the search and to the multimedia design of the presentation.

In addition to supporting presentations which are better suited to a variety of individuals, this approach allows a presentation to evolve as multimedia information is updated and augmented. This is possible because HyperAdaptive's script generator will use whatever semantically relevant multimedia objects are currently in the database. If data are updated or added, the new data are used.

B. User Scenario

Imagine a visitor enters the lobby of the MIT Media Laboratory and sees a computer with something interesting on the screen. As the visitor approaches, she sees that the screen is lit up with an illustration of one of the laboratory's projects, and text reading, "Come learn about our research." The curious visitor takes a seat at the machine and starts exploring this hypermedia application on the laboratory's research. She sees a diagram of the research groups within the laboratory, and chooses one (by clicking with the mouse or stylus, or other input device) that particularly interests her: the Visible Language Workshop.

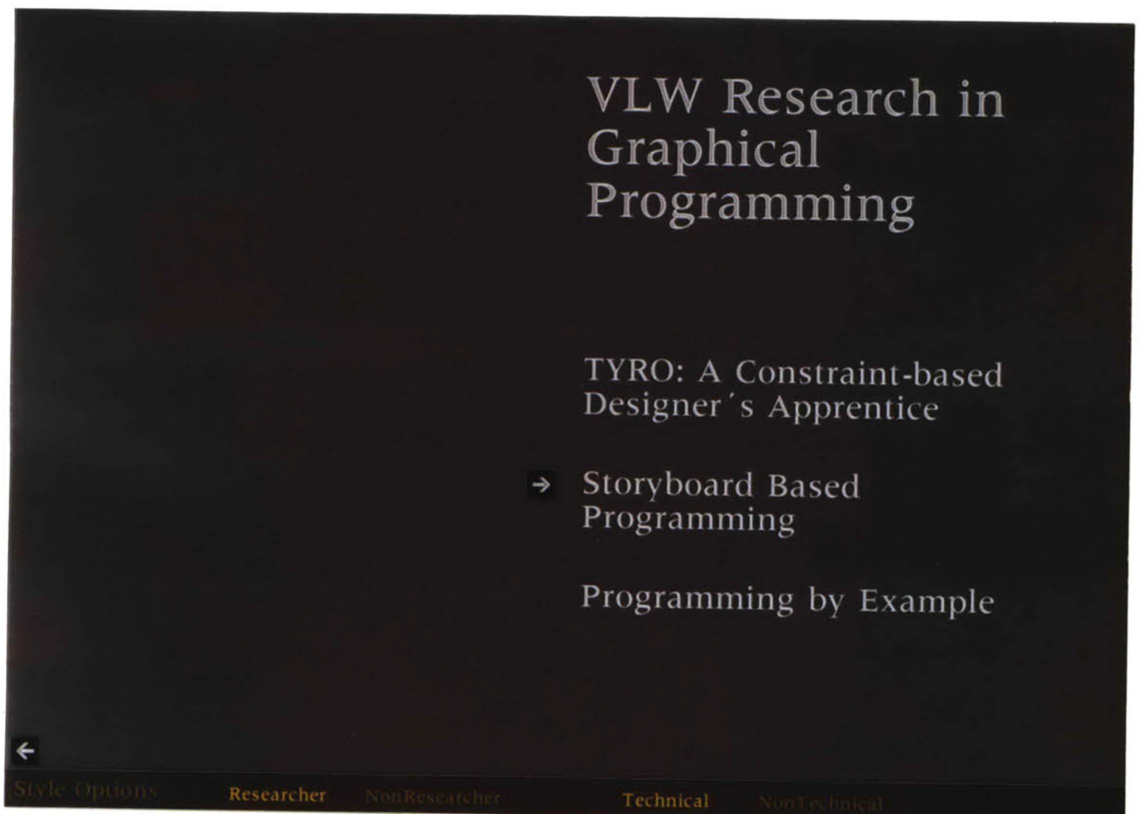

Figure 1 [Color Photograph]: The *VLW Research in Graphical Programming* menu.

After viewing a 40-second multimedia overview of the VLW's work, she selects "Research in Graphical Programming at the VLW." A menu appears, as shown in Figure 1.

A horizontal strip runs across the bottom of the screen, with two pairs of buttons on it, as shown in Figure 2. The user is a researcher, and she is technical in the software area, so she clicks on the *Researcher* and *Technical* buttons. She then selects a particular project, "Storyboard-based Programming."

Researcher **Non-Researcher** **Technical** **Non-Technical**

Figure 2. HyperAdaptive Style Strip for the two dimensions of Researcher/Nonresearcher and Technical/Nontechnical.

The system automatically generates a multimedia presentation on the chosen subject, using the style selections (in this case, Researcher and Technical) to guide the creation of the presentation. (See Figure 3 for a sample screen from the presentation.)

The user's selection of Researcher causes the system to generate a presentation which emphasizes research history and process, whereas the Nonresearcher presentation would emphasize the features and benefits of the Storyboard-based Programming project. The user's selection of Technical causes the system to generate a presentation which includes technical concepts and specialized terminology, and which excludes explanations of fundamental technical concepts which would be needed by a non-technical user. The user's style choices not only affect the emphasis and inclusion of certain material but also the ordering of the material included.

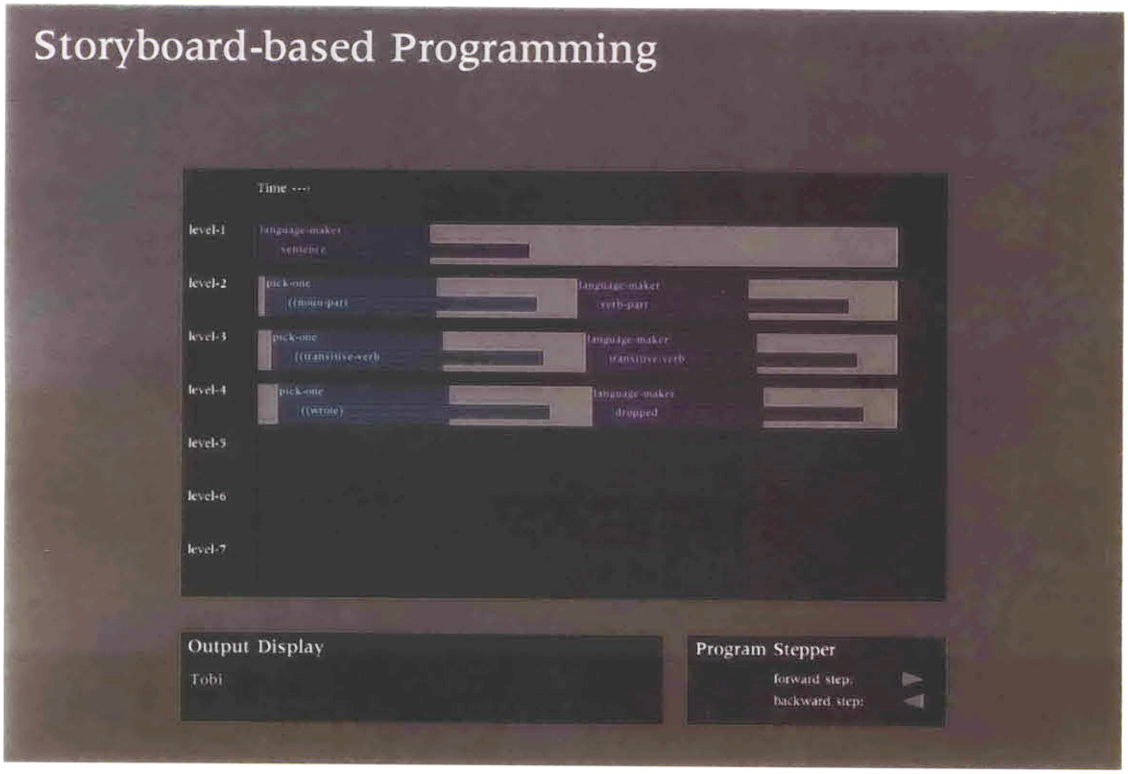

Figure 3 [Color Photograph]: From a HyperAdaptive Multimedia Presentation on Storyboard-based Programming.

C. Approach

Everytime a user selects high-level styles from the style strip, and selects a subject of interest, the system automatically generates a multimedia presentation on the chosen topic. Each style selection is associated with a knowledge-based script structure, which is a hierarchical list of information goals. A composite script structure is generated from combining the selected styles. Each information goal can be made up of subgoals, and the entire sequence reflect the elements of the story to be told. In the example, the choice of Researcher vs. Non-researcher is called a "high-level style determiner," in that it determines the list of high-level information goals, i.e., the overall structure of the presentation's content. The choice of Technical vs. Non-technical, by contrast, is treated as a "local style determiner," in that it has a more low-level control over the content. For instance, the Technical/Nontechnical choice may determine the presence or absence of certain technical information and language, as well as the addition of certain background technical information (for the benefit of a non-technical user).

See *Section III: The System* for a more extensive description of the system's implementation.

II. Research Context

This work sits at the junction of several disciplines. On the computer side, this work is in the traditions of adaptive hypermedia, knowledge-based graphical interfaces, and intelligent tutoring. Many of the work's theoretical foundations, however, reside with disciplines that pre-date the technology, most notably graphic design, narrative and rhetoric. This section describes relevant work in all of the above fields.

A. Adaptive Hypermedia

Most of the research in adaptive hypermedia focuses on navigation, paths and links. Two examplary papers in this area are by Meyrowitz (1989) and Zellweger (1989). Robin's work (1990. 1991) represents the most relevant effort because it supports adaptation of the contents of the presentation found at a hypermedia node. This section provides summaries of these three works, and their relationship to the current work.

Intermedia (Meyrowitz, 1989)

Brown University's Intermedia project is a premier example of hypermedia systems. Norman Meyrowitz, principal architect of Intermedia, reviews the state of hypermedia and outlines challenges for the future. The emphasis in this paper is on the nature of navigation, types of links (hot, warm, virtual and automatic), anchors and wayfinding. Meyrowitz also proposes a system-level set of hypermedia primitives, which would support a standardized implementation of links and buttons. Although all of these issues and features do play a crucial role in the personalization and improvement of hypermedia environments, there is little to no attention paid to adapting the contents of the information presented at the node itself.

Scripted Documents (Zellweger, 1989)

Zellweger (1989) describes Xerox PARC's "Scripted Documents" system, which supports prescribed paths to guide a user through a hypermedia application. One feature of particular interest is the notion of "active entries." This term indicates that the contents of a node can be dynamic or computational. For instance, the entry may play an animation, include some conditional tests or user interaction. All of these are relevant to the current work.

MMSE (Robin, 1990, 1991)

The Multimedia Scripting Environment (MMSE) was developed by Robin (1990, 1991) within the same group as the current research, the Visible Language Workshop of the MIT Media Laboratory. The system architecture assumes that a writer/designer will use the scripting tools to create multimedia presentations, each of which will be accessible at a node of a hypermedia application.

MMSE is composed of two interwoven systems: authoring tools for creating on-line multimedia presentations and a hypermedia component for demonstrating temporal adaptability in hypermedia applications. The designer makes detail-level and priority specifications while creating the presentation, and the system refers to these specifications during run-time adaptation to the end-user's interest level.

When the end-user selects a node in this system, a multimedia presentation is played which is dynamically tailored to her level of interest in the topic, and to her time constraints. If the end-user is very interested in the topic, she'll see the entire presentation. If she is less interested, the more detailed elements of the presentation will be omitted. The omission decisions are based on the designer's assessment of the detail level for each visual element of the presentation. The designer sets these ranks as she builds the presentation. An element, such as a text block or an image, marked "most detailed" will be the first to be omitted. The variability of the script reflects the number of detail levels defined by the designer. The user can even adjust her interest level in the middle of the presentation, and the rest of the presentation will automatically adjust.

HyperAdaptive is an extension of MMSE, using MMSE's multimedia driver to realize the multimedia presentation on the computer screen. HyperAdaptive, like MMSE, supports adaptation at a node of a hypermedia presentation. Whereas MMSE's adaptation is quantitative (varying how much detail is presented), HyperAdaptive's adaptation is qualitative, varying the structure, emphasis and design of the presentation to suit different types of users.

B. Knowledge-based Graphics and Multimedia

HyperAdaptive's approach and architecture are strongly informed by previous work in the field of knowledge-based graphics and multimedia.

An important feature of each of the projects described below is the representation and use of semantic relationships among the graphical elements as fundamental to the output decisions. This plays a fundamental role in the variation in presentation styles in HyperAdaptive.

Although there are too many relevant works to describe here, the works described below are representative of the most closely related work.

COMET: Feinner and McKeown

COMET (Feinner and McKeown, 1990) is a knowledge-based system which provides the user with multimedia explanations of equipment repair and maintenance. The content, as well as the choice of media, are determined as the system is being used, so that the explanations suit the user's particular circumstance and informational needs. Upon receiving a user's request for information, the system identifies the informational need, and then decides whether this would best be communicated in text, graphics, or both, based on heuristics which map certain types of communication tasks to certain media. Feinner and McKeown cite research which states, for example, that location information is best communicated by graphics only, whereas actions are best communicated by a combination of text and graphics. Once the media are selected, the informational goal is then sent as input to the text and/or graphics generators. Both media generators share a common knowledge representation, which allows cross-reference. For instance, the text generator can output a sentence which refers to the knob currently highlighted by the graphics generator.

COMET's approach is relevant to the goals and approach of the current research. COMET's use of one underlying knowledge representation to drive all media served as a model for HyperAdaptive. In addition, both systems use automatic generation in order to better adapt to the particular user's informational needs. One difference between the system is COMET's adaptation occurs at a lower level, including the use of natural language processing techniques to automatically compose sentences and the use of 3D graphic models to automatically generate graphics. HyperAdaptive's adaptation is focussed on a higher level, automatically composing more complex dynamic multimedia presentations out of a pool of existing textual, voice and (dynamic and static) graphical objects.

Beach and Stone: Towards High Quality Illustration

Beach and Stone (1983) described their knowledge-based system for generating scientific illustrations for different contexts. The importance of this work resides in the notion of a semantic layer underlying varying presentation styles. The system, for instance, can generate a graph for either a 35 mm slide or for a journal article. The elements of the illustration are realized differently depending upon which of the two styles is selected. For instance, the x and y axes will be drawn more thickly for the 35 mm slide. The power is in the semantic description underlying the actual output (a horizontal or vertical line of a certain thickness). Because the axes are known to the system as axes, not simply lines, style rules can specify how these objects should be realized under different circumstances.

HyperAdaptive adds at least one or two more layers of semantics to its generation of multimedia presentation. In addition, HyperAdaptive deals with dynamic multimedia, as opposed to static print-based output. Nonetheless, the work of Beach and Stone provides the relevant and powerful notion of having one underlying semantic representation of the world which can be communicated differently depending upon the high-level style specification.

Schank : Scripts

Schank (Schank and Riesbeck, 1981), a researcher in Artificial Intelligence, introduced the notion of a "script," which has served as an influential idea in the field. A script is defined as a hierarchical structure used to represent experiences (such as going to a restaurant) as a series of events occurring over time. This construct has been widely applied to natural language understanding and generation, among other tasks. Schank's script representation served as an important inspiration for HyperAdaptive's script style representation, which is a hierarchical representation of the information goals that will be fulfilled over time, during the course of the multimedia presentation.

C. Intelligent Tutoring

Researchers in intelligent tutoring create computer-based educational software which uses artificial intelligence techniques to adapt to a variety of user needs (Sleeman and Brown, 1982; Polson and Richardson, 1988; Self, 1990). Most works in this field build and update a user model, which is used to determine the content and level of instruction needed at that moment. The user model may be informed by the user's background knowledge, but is even more often based on her history of interaction with the system.

Work in intelligent tutoring is relevant to the current research, because of the common application of knowledge-based techniques for the purpose of adapting to different user needs. One difference between HyperAdaptive and most intelligent tutoring projects is that HyperAdaptive does not have a representation of the user, per se. Rather, the presentation is what is represented. It should be noted, however, that assumptions about a user description are implicit when we deem certain presentations suitable for certain users. This means that the areas of intelligent tutoring and user modelling (Neal, 1991) are pertinent to HyperAdaptive's aims. In fact, user modelling could be plugged into the front end of HyperAdaptive, replacing the user's selection of style descriptors with a model of the user.

D. Presentation Style

The above sections describe research in the software domain which relates to the notion of varying presentation styles, with a strong emphasis on a semantic description underlying the output decisions. It is also important to consider relevant theories and formalisms from fields which pre-date computer interfaces, including graphic design, narrative and rhetoric. It is striking just how much these time-tested traditions have to contribute, and how much their respective messages and theoretical constructs reinforce eachother.

The word "style" in this thesis refers to variations in the presentation. This is to be distinguished from the common use of "style" in cognitive and epistemological literature, where it refers to cognitive style, or learning style, which refer to differences among people. [1] The source of confusion lies not only with the ambiguous meaning of the word "style," but also in the fact that these forms of style are certainly relevant to one another. Although this thesis manipulates presentation style, it is informed and motivated by the variety of cognitive styles present in users.

This section presents research most relevant to variations in presentation style.

---

[1] See Turkel & Papert (1990), Goldman-Segall (1990), Strohecker (1991), for discussions of learning and cognitive styles as they pertain to computer user interfaces and computer-based learning environments.

1. Graphic Design

Independent of whether a hypermedia system adapts, it is crucial to consider its multimedia design. What design principles should be applied to insure that the hypermedia presentation communicates effectively? Within the context of an adaptive system, it is important to encode multimedia design principles to reinforce the varying content relations of the different styles.

The dynamic and interactive nature of the computer medium presents new graphic design challenges and opportunities, as described by Cooper (1989). Cooper, director of the MIT Media Laboratory's Visible Language Workshop (site of the current research) addresses many of the implications this new medium has on graphic design methods. The article describes the group's experiments in this area, including use of position, transitions, dynamics, and translucency to exploit the computational context.

Given that our presentations are being automatically generated, it is important to formalize some of these multimedia design principles. This section addresses the aspects of multimedia design that HyperAdaptive's methods consider.

Form and Content

Style should not be divorced from content, but should grow out of it. We learn from the influential Bauhaus school of design (Wingler, 1969) that a good design weds form and function. In this project, this means that the system should use design techniques to reinforce the semantic, or rhetorical, role of each element.

Layout: *The Grid*

An important concept in graphic design is the *grid*. In Foundations of Graphic Design, Gatta et al. (1991, page 230) provide this definition:

> Grid: A set of horizontal and vertical lines used as a guide for alignment of type and photos; creates a uniformity of design.

The regularity of the grid serves to orient the user to the material, guide the user's eye and communicate where functional components of the layout can be found. Figure 4 shows two pages from the book, Foundations of Graphic Design. The book designer created a five-column grid which provides uniformity and predictability, while allowing for variety within its constraints.

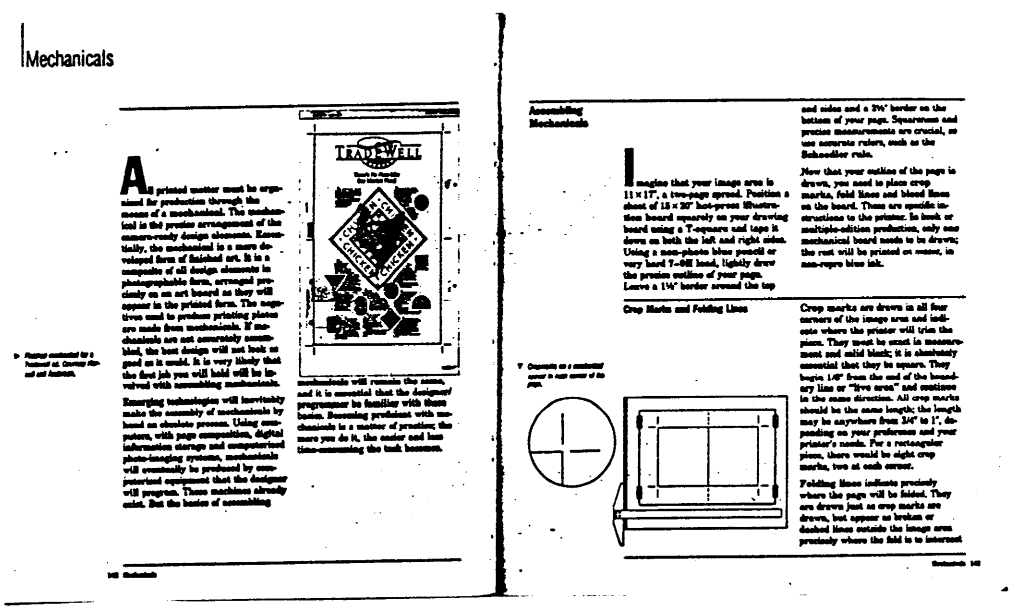

Figure 4: Two pages from the book, Foundations of Graphic Design (Gatta, et al, 1991, pp. 142-143) reflect the book's grid, which is divided into five columns per page. From the left, column 1 is where titles, topic headings, footnotes and small illustrations appear. Columns 2 and 3 are used (as one wide column) for text and illustrations, as is the wide column formed by columns 4 and 5.

HyperAdaptive's output uses a set of simple functional grids, where portions of the screen are assigned to one function or another; among these functions are: illustrations, titles, topics, and animated simulations. There are a few different pre-defined layouts, each with its own variations on the basic grid. See Section *II.B.2. Style Scripts and Methods* for examples of HyperAdaptive's layouts.

Dynamics

Today's high-quality graphical computing environments allow for a powerful experience of dynamics. In fact, dynamics are intrinsic to computer-based multimedia. The presentations created by HyperAdaptive are defined by the appearance, disappearance, and tranformation of a variety of graphical, textual and sound objects over time. Dynamics exist both at the local and global levels. At the local level, a particular object may become an animation, itself changing from static to dynamic. Such local dynamics calls attention to that part of the screen, and the particular piece of information conveyed by the animated image or text. There is also a global dynamic, defined by the change over time in which a set of objects is present (in what arrangement). Both types of dynamics, in addition to calling attention to particular parts of the screen, can communicate certain ideas.

Memory

In dynamic presentations, things move, change and disappear. This places an extra demand on the user's memory. The multimedia design should employ techniques to support the user's memory of what has gone by.

What cues can be used to remind user of what she's seen, where she is in the larger presentation? Graphic design techniques for supporting the user's memory of what has passed include leaving visible traces of what came before, such as key words of phrases, or miniature versions of still images. The current prototype leaves a trail of key words (one to three words per subtopic) for every narration. Since sound is by definition fleeting, the visible traces are all the more important. In addition to boosting the memory, the simulataneous presentation of these key words and the narration reinforces the message as it is delivered.

Attention and Focus:

Various methods from graphic design (and more recently, multimedia design) have been developed to control the user/viewer's attention. The Visible Language Workshop has conducted extensive research in applying and advancing such techniques for the dynamic nature of the computer interface. These techniques include the use of highlighting, position, translucency, size, timing, and dynamics to emphasize, and de-emphasize graphical (including images, text, and sound) elements of a presentation over time. The current project explicitly employs highlighting, position, size, timing and dynamics to reinforce the message of the presentation.

Timing

An important aspect of computer-based multimedia design is the relative timing of narration, text, graphics, and sound effects. A lack of coherent temporal coordination can be very disconcerting to the user, and even be confusing. There is a danger of *unwanted Gricean implicatures,* which Marks and Reiter (1990) address in the context of automatic text and graphics generation. The philosopher H.P. Grice (1975) first described this phenomenon in terms of linguistic communication. Grice proposed that speakers and hearers share certain assumptions about their communication, and that unintended violations of these assumptions can cause the hearer to infer unintended meanings. For instance, if the speaker makes a reference to an object with more detail than the hearer would normally expect in the context, the hearer is likely to assume that there is some importance to this additional information.

Marks and Reiter show the relevance of implicatures to both computer-based generation of both text and graphics. When a computér, untutored in such subtleties of human communication, automatically generates a sentence that is overly detailed, this can cause an unwanted implicature. In the graphical context, graphical features, such as bolding, clustering or alignment serve as details which the user/viewer tries to make some meaning of. If they are created unintentionally, this could easily distract and confuse. Figure 5 shows an example of a network diagram where ordering by size "implicates that there must be a similar ordering relation among the vertices in the network model, which is not true" (p. 450).

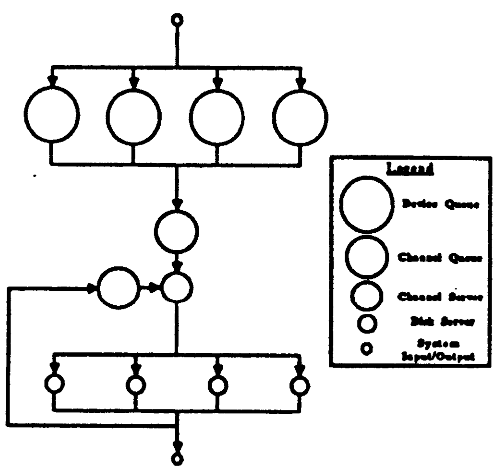

Figure 5: An example of unintended graphical implicature. The ordering by size falsely implicates an ordering among the vertices. Figure copied from Marks & Reiter (1990, page 456, Figure 3).

Although professional graphic designers are trained to avoid such implicatures (even if they have never heard of Grice), the computer needs to be so programmed. Marks and Reiter have built automatic text and graphics generators which incorporate constraints to prevent the creation of unwanted graphical and textual implicatures.

The notion of unwanted implicatures applies to many aspects of HyperAdaptive's automatic generation, including layout and emphasis, as stressed by Marks and Reiter, but also the notion of timing. If a narration, for instance, begins just before one image disappears and another appears, the user could easily make the incorrect inference about which image in intended to relate to the content of the narration. The current project attempts to avoid such unwanted implicatures by simple timing templates for different communication contexts. There are standard temporal relations enforced for different communication contexts, such as *overview explanations* versus a *guided simulation* of software.

Relative role of sound (narration, music, sound effects) , images, text

Related to the issue of timing are the relative roles of the various media in the communication task. Different roles are appropriate to differing communication goals.

Schapiro, in Words and Pictures (1973), addresses this issue in the context of illustrations of biblical text:

> Illustrations of a text may be a mere emblem of the story, or may enlarge the story, adding details not given in text. An interpretation of the text (not found in the literal text) may determine an illustration's detail.

The roles of various media is clearly an issue of great complexity. However, as Feinner and McKeown (1990) show[1], it is possible to generate certain heuristics about what combination of media suit various types of information goals.

For HyperAdaptive's goals, informal user feedback revealed that people preferred narration over extensive blocks of text, especially in a presentation that is moving along. Therefore, the system is biased toward narration over text blocks.

When using narration, HyperAdaptive uses text (in the form of captions or short titles or keywords) and images to reinforce the narration.

This multimodal reinforcement is supported by extensive evidence in cognitive and perceptual psychology that a message received both visually and aurally is more effective than one transmitted through only one of the modalities (Osherson, 1990).

As mentioned in the "Memory" section above, HyperAdaptive leaves a brief textual trace (one to three key words) of each topic on the screen to overcome the transient nature of narration with the potentially "permanent" nature of the visual.

---

[1] See Section II.B. of this thesis for a discussion of Feinner and McKeown's relevant work.

2. Narrative

Developers of computer interfaces are now exploring and applying notions of storytelling and narrative in computer systems (Laurel, 1990, 1991; Don, 1990; Schank, 1990). These researchers argue for the use of the time-tested principles of story-telling to better engage computer users. This section first explores fundamental concepts of narration, and then discusses their relevance to the current work.

Fabula, Syuzhet, and Style

In Narration in the Fiction Film, Bordwell (1985; chapter 4) describes fundamental constructs of the narrative form, and how they relate to films. Although the book focuses on the medium of film, the principles of narration, *fabula, syuzhet* and *style,* are derived from ancient sources, and are largely medium-independent. These distinctions, traced to Aristotle's Poetics, were expanded upon by the Russian Formalists (Tynianov, 1978).

Fabula (usually translated as "story") is defined as the underlying series of events; what actually occurred (even in a fictional world) and in what order. For instance, a detective story's fabula might be: shopkeeper murders competitor; shopkeeper bribes competitor's clerk to silence him; competitor's wife hires private detective; detective is originally misled, but eventually discovers the true murderer.

The *syuzhet,* (translated as "plot") by contrast, is how the events and facts of the fabula are revealed over time to the reader, audience, or, in the case of user interfaces, the user. For instance, the sequence of the syuzhet for the example detective story may be: competitor's wife walks into private detective's office and hires him for the case; detective searches for the truth, finding evidence that the competitor's customer committed the murder, then that the hiring wife did it, and finally three bits of evidence, each presented separately over the last hour of the show, come together to force our conclusion of the shopkeeper's guilt. Each piece of evidence is revealed over time, and its order may be completely independent of the actual order of events. It is the syuzhet that controls the amount, order, and emphasis on the information contained in the fabula. Each genre has its own syuzhet techniques and traditions. The detective genre is known for withholding or postponing the presentation of key information, or for burying important clues among irrelevant details. The genre's particular balance of information, misinformation and delayed information is part of what draws its fans.

Bordwell describes *style* as an important third variant in narration. Style, unlike syuzhet and fabula is a set of medium-dependent techniques. In film, style decisions include editing, cinematographic approach, sound, and lighting.

Each of the three narrative constructs can easily be related to a multimedia presentation generated by HyperAdaptive.

The fabula in one of HyperAdaptive's presentations on a VLW research project might be the historical sequence of events in the described research project, such as: the researchers joined the group; they had a meeting in which the idea first arose; they had further meetings refining the idea, including other group members, they designed and developed an early version of the prototype, they discovered problems with the prototype and created a new version to overcome these problems.

The syuzhet can vary, and should vary, for different user groups. A presentation for a fellow researcher, for instance, might emphasize the research process, whereas a presentation of the same fabula to a non-researcher might emphasize the features of the final prototype, and how it would be useful.

The style, or medium-dependent decisions such as layout, transitions, and timing (as described in the Graphic Design section above) should vary to reinforce the chosen syuzhet.

HyperAdaptive's variation of presentation style is a variation of both syuzhet and style, in Bordwell's terms. The script style structures (described in more detail in Section III.B.2.: Style Scripts and Methods), are mapped to what Bordwell calls syuzhet. The multimedia design decisions reflect differences in style, using Bordwell's definition of the term.

As Don (1990) points out, narrative models provide "multiple constructions of knowledge." She writes that such models should be used in computer interfaces to better provide users access to a variety of views of information. Based on her experience developing multimedia artworks, as well as the Guides interactive video application at Apple, Don writes of some important variations in storytelling, which apply to computer user interfaces as well. Don differentiates between the story being told, and the conditions for its telling. The conditions include such factors as *who* is telling the story, *to whom*, and *why* the story is being told. All of these conditions would naturally affect what Bordwell calls syuzhet and style. Don makes the connection between these principles of narration and variations in constructions of knowledge at the user interface.

An important aspect of applying narrative techniques to the user interface is that the principles of narrative and those of other relevant fields, including graphic design and computer-based graphic design, reinforce each other. The notions of form and function in graphic design (as articulated by the Bauhaus school), and the notion of a semantic layer underlying graphical output in computer-based graphic design, both mesh well with the ideas of fabula, syuzhet and style in narration. The fabula maps to the semantics represented in knowledge-based graphics, and the syuzhet maps to the ordering and emphasis among the facts in a presentation. Both of these represent the function discussed in graphic design. The "form" of graphic design maps well to the "style" of narrative, as presented by Bordwell.

The following section on rhetoric describes another field which reinforces and adds to these constructs.

3. Rhetorical Constructs and Relations

Rhetoric, like narrative, is a set of age-old human communication skills and techniques. Rhetoric is defined[1] as "the art or science of using words effectively in speaking or writing." Aristotle's Rhetoric emphasizes persuasive oratory, but the principles presented are broad enough to also bear on multimedia communication, as well as communication whose main goal is to inform rather than persuade. Aristotle's work still serves as a basic work in the field. Despite the obvious differences between Aristotle's time and our own, it is striking how many of The Rhetoric's principles remain relevant.

The methods of rhetoric, like narrative, allow the formation of, as Don (1990) writes, "multiple constructions of knowledge." The methods of rhetoric can be applied to support successful communication to different users.

In The Rhetoric, Aristotle advises the would-be rhetorician to prepare an argument on three levels: *ethos, pathos* and *logos*. This section describes each of these, and how it relates to the generation of multimedia presentations on the computer.

Ethos refers to the speaker's credibility.[2] In order to be convincing, first convince the audience that they can believe you. Aristotle outlines several methods for communicating such credibility. Although there are obviously differences in cultural context, not to mention the difference in medium,

---

[1] According to Webster's New World Dictionary, 2nd College Edition, 1984.

[2] Aristotle presumes the form of presentation to be the speech. No multimedia yet, nor even printed books.

the notion of creating a credible presentation is certainly relevant to the creation of computer-based multimedia. Such credibility is communicated both by form and content, and the appropriate measures will vary with the audience. For instance, an academic may find that academic references, scientific data, affiliation with respected institutions, and the look of a scientific journal all help create an air of credibility. A business person, however, may associate credibility with the management or financial know-how shown in the presentation, or the look of a corporate presentation. In HyperAdaptive's presentations, it is important to establish credibility with different audiences, and therefore to consider who the groups are, and how to build the credibility of the multimedia presentation as the stand-in for a human presenter.

Pathos, Aristotle writes, is mood; the rhetorician must put the audience in an acceptable mood by manipulating their emotions. The desired mood depends on the goals of the presentation. In trying to convince voters to approve a strict crime referendum, a politician may promote fear. A goal of HyperAdaptive's presentations of VLW research is to promote curiousity and concern about the issues which our research explores. Curiousity is encouraged by creating presentations which are visually compelling, involve meaningful user interaction, and which address issues of interest to the particular user. Concern is encouraged by communicating at the start of the presentation that there is a problem in the world which needs to be solved, and that the project described helps to solve it.

Logos refers to reasoning or logic. The reasoning of rhetoric is related to, but different from, formal logic. A rhetorician may use formal logic in order to prove a point, but most of the rhetorician's tools are of informal logic. Aristotle describes the *enthymeme,* an important rhetorical construct which, though not a strict logical syllogism, has structural properties of an argument that convinces; it *appears* to be logical. For instance, making comparisons is a good example of the informal logic of an enthymeme. Here is one such enthymeme you will find (even if not in exactly these words) in a HyperAdaptive presentation on a graphical programming project in the VLW:

> Graphical interfaces and representations have been very useful for end-users of computers.
>
> Most programmers' tools are textual, and lack graphical representations.
>
> Therefore: Graphical interfaces should be provided for programmers, who would benefit from them just as end-users have.

This gives the appearance of a logical argument, but a formal logician would find that it is not a valid syllogism because the conclusion does not strictly follow from the two premises. The premise that graphical interfaces are helpful for end-users does not logically necessitate that graphical interfaces would be useful to programmers. The first premise may very well suggest that the conclusion is reasonable, but that does not make it a logical necessity. (This does not mean that the conclusion is false, of course.) The comparison of two ideas, phenomenona or things which are similar in certain ways, in order to support an assertion that they are similar in other ways, is a classic Aristotelian enthymeme still quite common in everyday communication.

Constraint-based Hypertext for Argumentation:

The work of Smolensky et al (1987, p. 215) makes a link between the rhetorical and logical principles explored in rhetoric, the presentation of research, and the computer user interface. The authors describe their software, which supports the representation of logical arguments within a well-structured hypertext context. The system helps a researcher to manipulate, test and build a reasoned discourse for research papers. The system's implementation tackles two hypertext problems: managing user interaction and managing the screen to reinforce the logical relations within an argument with graphical representations.

The system separates the content of an argument from its formal structure. The computer has access to the formal aspects of the argument; that is, to its logical relations. The user must make the logical relationships between assertions explicit. Once she does, the system can identify which claims are supported or unsupported, and which are refuted by other claims. The system graphically represents such relationships, through a set of graphical constraints on the screen management. Below are two examples of such graphical constraints:

An unsupported claim must be in large bold type or in red type.

A claim refuting another claim should be connected to it with an arrow labelled "refutes."

This system's use of graphic techniques to represent relations in the content is relevant to HyperAdaptive's need to map content relations to graphic presentation. HyperAdaptive is not so concerned with making the formal logic explicit, but rather with reinforcing the less formal semantic relationships existing within its presentations.

E. The Subject is Research

The subject matter of the hypermedia presentations generated by this system is research at the Visible Language Workshop. This is a rich subject matter, because research can be viewed in so many ways. In fact, a lively academic debate now focuses on the points of view employed when researchers discuss and evaluate research (Latour, 1987; Keller, 1985).

Although this system can only begin to explore the complexities of this issue, it is intriguing to consider what role automatic generation of research presentations could play in the larger academic debate. At any rate, an awareness of the larger academic debate is important to the process of formalizing styles of presenting this subject matter.

Media Lab researchers spend a good deal of time presenting their research, usually in the form of software demonstrations to visitors. This method of presentation, although varying to some extent among research groups, individual researchers, and the needs and background of the visitor, has its own particular types of rhetorical and narrative structures which often contrast with other traditions for presenting research. The "demo mode" is typically characterized by a strong emphasis on the features of the research prototypes. The researchers typically omit discussions of method and background expected in a scientific journal article.

The immediacy of the demo experience makes this mode of presenting research a powerful tool for communication. The demo allows the visitors to experience the research first-hand, and serves as a platform for two-way interaction between researcher and visitor. Papert[1] observes that, despite the advantages of this mode, it is an evolving form, and its evolution may benefit from the incorporation of new features.

Ackermann[2] suggests that one could consider research presentations along two dimensions: product-oriented vs. process-oriented and self-promoting vs. self-reflective. The scientific journal article can typically be described as process-oriented and self-promoting. The demo mode typically lies near the product-oriented and self-promoting ends of these spectrums. Contemporary scholars, such as Keller (1985), propose that the scientific world would benefit from discussing their research in more self-reflective, process-oriented terms.

Even in the "demo mode," there are interesting variations. The HyperAdaptive script styles support two dimensions of such variation, whether the visitor is a researcher or non-researcher, and whether she is technical or non-technical. The underlying script styles were created based on observations of how presentations vary for people on the basis of these two dimensions. The researcher mode is more oriented toward the research process, whereas the non-researcher mode focuses almost exclusively on the features of the product. The technical mode assumes software expertise, whereas the non-technical mode does not. (See Section III for more on HyperAdaptive's treatment of these dimensions.)

---

1 From videotaped recording of interview with Seymour Papert about demos in the media lab, conducted in fall '90 by Marc Davis for Pro Seminar, MIT Media Lab.

2 Edith Ackermann, personal communications, 1991

III. The System

A. Subject-matter: Media Lab Research

The subject matter for the multimedia presentations in this prototype is research at the Visible Language Workshop, as described in the immediately preceding section, *II.E.: The Subject is Research,*

For the purposes of demonstration and experimentation, there are now two dimensions along which a presentation can vary: researcher/nonresearcher and technical/nontechnical. The system's architecture could support the addition of other style dimensions.

The assumption underlying the researcher/nonresearcher dimension is that researchers will have a greater interest in the research process, whereas nonresearchers are more likely to want a presentation which focuses on the features of the product of the research: the prototype. Therefore, the underlying knowledge structure for the researcher script includes a description of an earlier version of the prototype, its limitations, and how the current prototype attempts to overcome these limitations. By contrast, the nonresearcher script structure never even mentions the research process, or the existence of an earlier, problematic, version.

It is not claimed that this represents a universal definition of the types of presentations suitable for a researcher versus a nonresearcher. This definition suits the current context, and is a good test bed for experimentation and feedback.

References to "historical version," and "current version" in the System Architecture section refer to the versions of the VLW research prototypes, which are the subject matter of the automatically generated presentations.

B. System Architecture

The architecture of the system, called HyperAdaptive, was influenced by the work described in the research context section, especially the work of Feinner and McKeown, Beach and Stone, and that of Schank.

1. Software Context and Components

HyperAdaptive is an extension of MMSE, a system developed by Robin (1990, 1991) at the Media Laboratory's Visible Language Workshop, where the current research was also conducted.

<u>MMSE</u>

The adaptivity in MMSE, as described in section II.A., is based on temporal contraction of a script designed by a multimedia designer. There are two components of MMSE, as shown in Figure 6.

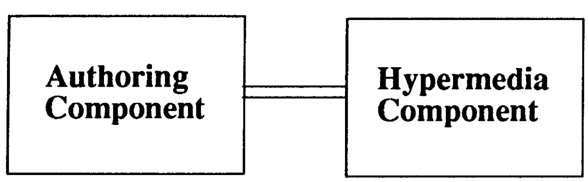

Figure 6: MMSE's two components. Figure adapted from Robin (1990, p. 13, Figure 2.2).

Within the authoring component, the designer uses MMSE 's scripting tools to create a multimedia presentation, and saves the description out to an

ASCII file. At run-time, this script file is read in, and the dynamic multimedia presentation plays back. (See Figure 7.)

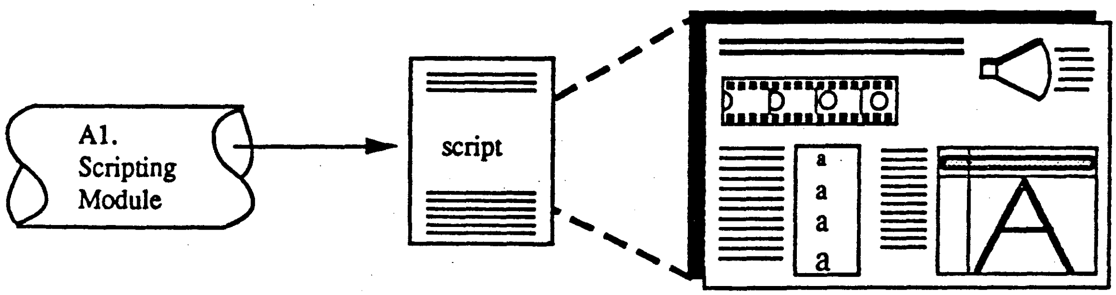

Figure 7: The designer uses MMSE's scripting tools to produce a script file. This file, describing the contents, sequence and layout of the presentation, is loaded in and played back as a dynamic multimedia presentation. Figure from Robin (1990, p. 15, Figure 2.4).

Figure _ shows a portion of the MMSE script file for the "Research in Graphical Programming at the VLW" menu. See Figure _ (in *Section I.B. User Scenario*) for a picture of this screen. The portion of the file shown in Figure 8 describes the background information, the title and the first menu choice (TYRO). Because the file exists, HyperAdaptive simply loads it in, instead of automatically generating a script.

As described in Section II.A., MMSE provides tools for the designer to specify detail rankings associated with the multimedia objects making up the presentation. These detail rankings are used to determine how much of the presentation is played back, according to the end-user's specification of interest level.

```
script_duration=4
timing=0_75
compaction=0.000000
area_xpos=20
area_ypos=85
area_width=1200
area_height=810
area_rgb=90_90_90
area_border=1
area_border_rgb=150_150_150

number_of_track1_children=4

trackicon=vlwgra_prog_trackicon
rel_starttime=1
rel_endtime=7
duration=6
detail_level=0

x_position=600
y_position=60
width=550
height=194

text_graphics=0
fontname=Latin_medium
fontsize=56
red=216
green=217
blue=216
trans=0

trackicon=TYRO_trackicon
rel_starttime=1
rel_endtime=7
duration=6
detail_level=0

x_position=600
y_position=360
width=491
height=94

text_graphics=0
fontname=Latin_medium
fontsize=36
red=216
green=217
blue=216
trans=0
```

Figure 8: Portion of an MMSE script file. The top section provides general information, including the presentation's background color, width, height and screen position. The following sections provide temporal, spatial and typographical information for two of the presentation's text elements. A full script file includes this information for all media types.

Relationship to MMSE

This prototype is an extension of MMSE. As noted above, a designer would use a set of interactive tools to create an MMSE script. This capacity is retained by the current system, to support those cases where the designer wishes to handcraft a presentation, and have it adapt only in terms of amount of detail included. Such a presentation would vary according to the user's expressed interest level, but not with the user's style selections.

In most cases, however, a script file will not exist, and the presentation will be automatically generated at run-time. Figure 9 shows how HyperAdaptive's automatic script generator outputs the equivalent of an MMSE script, which is performed using MMSE's playback engine.

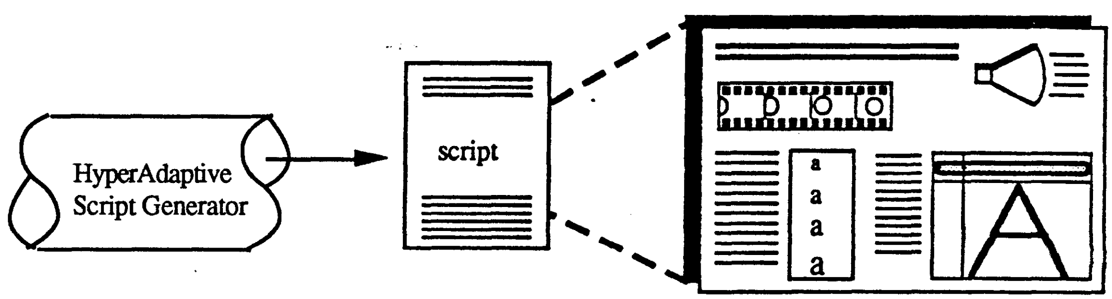

Figure 9: HyperAdaptive's Script Generator outputs the equivalent of an MMSE script description, which can be performed using MMSE's playback engine.

See figure 10 for a more detailed view of the HyperAdaptive Script Generation process.

The generated multimedia script is represented internally just as an MMSE script would be. This allows for co-existence between MMSE scripts and HyperAdaptive scripts, and full integration of the two.

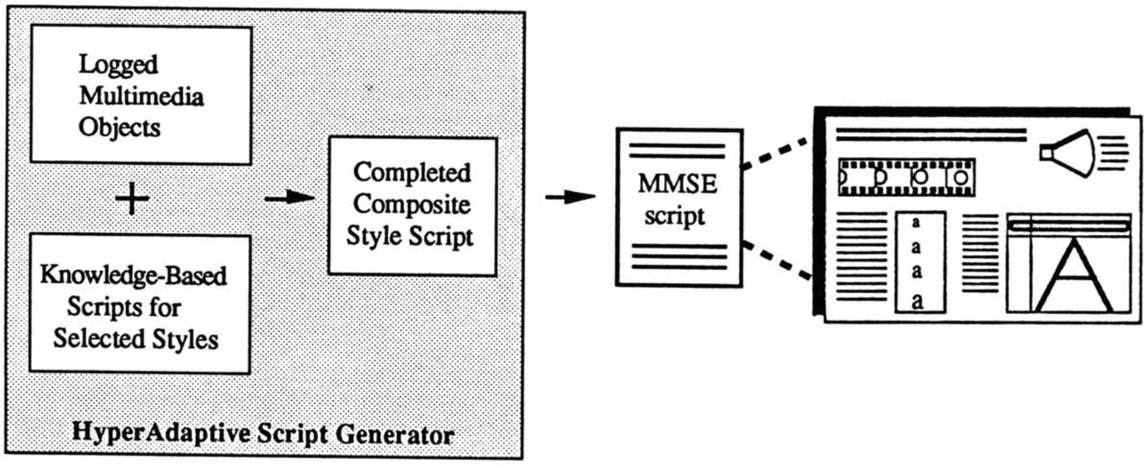

Figure 10. Stages of generating and displaying a HyperAdaptive script. This shows the same processes as Figure 9, but includes the components of script generation.

2. Style Scripts and Methods

The user's indicated style preferences (e.g., Researcher/Nonresearcher, Technical/Nontechnical) drive the system's selection of Style Scripts. This process is divided into two interleaved phases:

1. Search
2. Multimedia Design

Search Phase

The style scripts, which are nested linked lists, made up of hierarchical information goals and their associated methods, are used as guides for the search for suitable multimedia objects. In other words, the system searches through the multimedia database to collect the bits of text, sound, digitized images, and graphics to express the information goals of the Style Script.

As indicated in Figure 10, the search phase creates and uses a composite script structure, composed of the global and local style determiners. In the example, Researcher/Nonresearcher is a global style determiner, because it determines the high-level list of information goals.
Technical/Nontechnical (discussed in more detail below) is a local style determiner, because its choice constrains how the high-level goals are fulfilled.

Multimedia Design

Once the objects are found for each information goal, the multimedia design decisions are made. These decisions include the temporal coordination of the various media, spatial layout, choice of type style, size and color. See Section *II.D.1: Graphic Design* for a discussion of the graphic design principles which HyperAdaptive applies.

Representation of Style Scripts

A style script is implemented as a linked list of information goals. Each one of the goals may itself be made of lower-level information goals. This nesting allows for a good deal of richness and complexity within the list structure. Each information goal has a function associated with it which drives both the search and multimedia phases of its fulfillment. Figures 11 and 12 show the top-level information goals for the Researcher and Nonresearcher style scripts, respectively.

Information Goal	**Method**
Idea/Purpose	Compare-Contrast
Features of early version	Find/Describe Cross-version Features + Historic Features
Limitations of historic version	Find /Describe Problems with historic version
[Blind alleys][1]	[Find/Describe Blind Alleys]
Current Version: contrast with historic version	Find/Describe/Contrast Current Version Features
Conclusion	

Figure 11: Researcher Style Script

[1] Square brackets indicate that the information goal is optional (depending on whether relevant objects are found).

Information Goal	Method
Idea/Purpose	Compare-Contrast
Features of Solution (Current)	Find Cross-version Features + Current Features
Benefits	Find/Describe Benefits to Users
Conclusion	

Figure 12: Nonresearcher Style Script

Technical/Nontechnical dimension

Technical is understood as domain-specific. That is, whether the user is technically sophisticated in the subject-matter of the presentation.

The Technical/Nontechnical style scripts are classified as *local style determiners.* This means that they add local constraints to the search process conducted for the global style dimension, such as Researcher or Nonresearcher.

There are three ways that a selection of Technical or Nontechnical may affect the generated presentation. First, the Nontechnical selection may determine that extra explanation is added when fulfilling one of the information goals (or subgoals) specified by the researcher or nonresearcher style scripts. Second, the Nontechnical selection may mean that a narration lacking in technical jargon will be replaced for narration using jargon. Finally, the technical selection may mean that additional technical details are added, which would be of interest only to the technically-oriented audience.

Methods: Search and Multimedia Layout

Each information goal has a function which defines both the search process and the multimedia design decisions required to fulfill that information goal. The process is multi-leveled, since information goals may be nested two to four levels deep.

To clarify this point, consider the "Features of early version" info goal in the Researcher script style (Figure 11). This info goal's method is: "Find/Describe Cross-version Features + Historic Features." This method's search task is to find each feature of the described software which is either found exclusively in the early version, or which is common to both versions (i.e., neither version is specified).

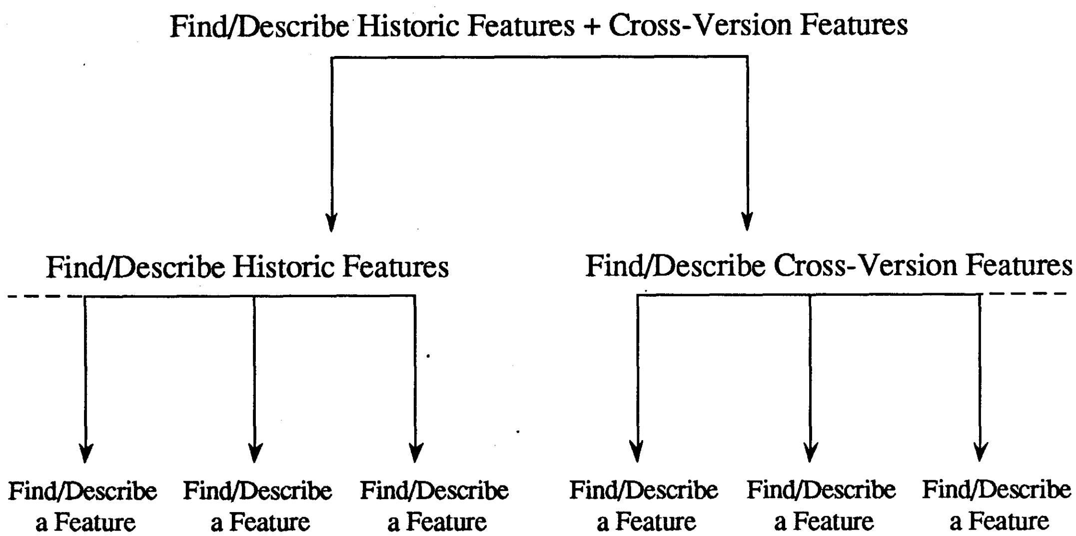

Figure 13. High-level description of search process. In this example, the goal is to fulfill the "Features of historic version" Information Goal in Researcher Script Style. See Figure 6 below for close-up view of each "Find/Describe a Feature" subgoal.

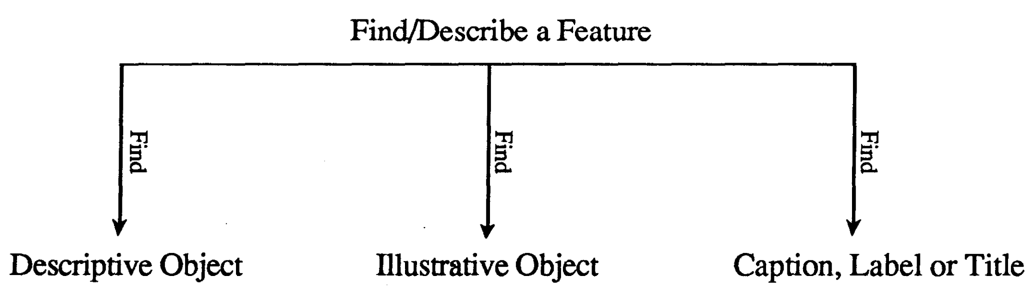

Figure 14. Search process for terminal level subgoal. In the example illustrated here, the goal is to describe a feature of the software.

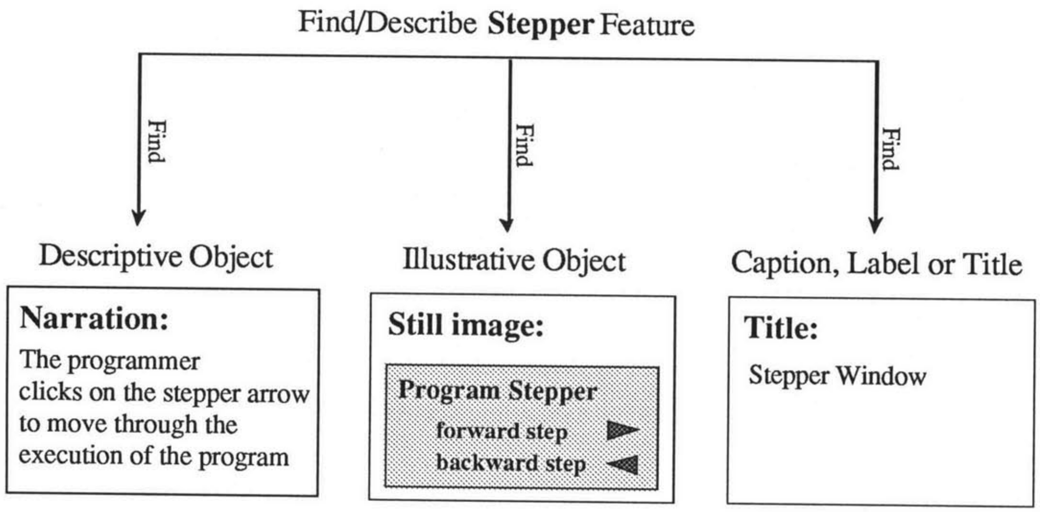

Figure 15: Illustration of the search process for an example feature, the stepper feature. This is a concrete version of figure 14 above.

Multimedia Layout Decisions

The search process is interleaved with decisions about multimedia design. These decisions include the selection of various media (if the objects are available in those media) and the temporal and spatial layout of the objects. Many of the design issues addressed in HyperAdaptive's automatic generation are described in *II.D.1. Graphic Design* .

The decisions are made on two levels: basic principles of effective multimedia design, and how can the design most effectively reinforce the semantic (rhetoric, narrative) relationships between the elements of the presentation. For instance, the "idea" section (Figures 11 and 12) of the

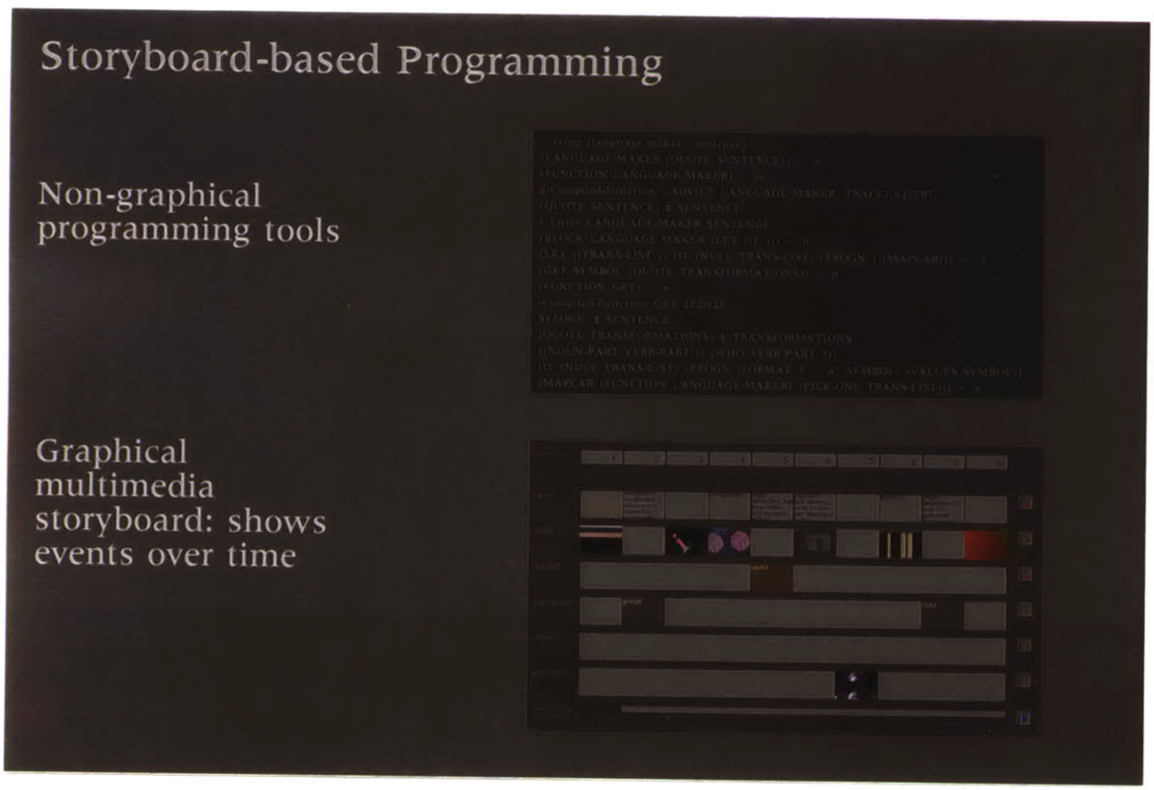

Figure 16 [Color Photograph]: A screen which uses the compare-contrast layout template. The layout reinforces the rhetorical relations of its elements. Narration describes what is shown.

script structure for a presentation on a VLW research project is represented as comparing and contrasting two things: non-graphical programming environments with graphical multimedia storyboards. Figure 16 shows this screen.

The comparison is shown as an example of an Aristotelian enthymeme in *Section II.D.3. Rhetorical Constructs and Relation.*) The system uses the fact that this presentation segment is a *comparison* to ensure that the images and text illustrating these concepts are shown on the screen together, and in a pre-specified layout. The system has a layout template for compare-contrast relations, illustrated in figure 17.

The "grid" is the graphic design principle which underlies the layout templates. The idea is that the layout should reflect pre-defined functional areas of the screen. The use of the grid serves to guide the user's eye, and associate meaning with the positions on the layout. (Wingler, 1969).

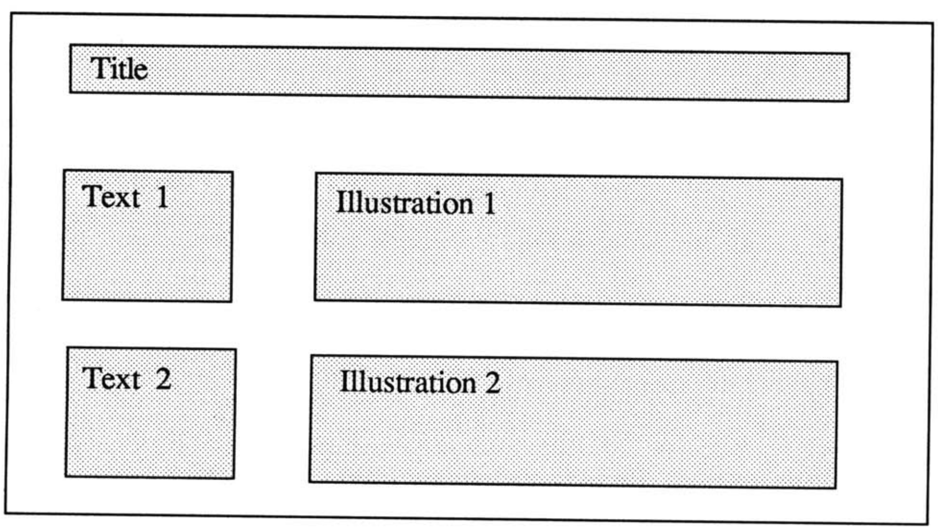

Figure 17. Layout template for compare-contrast relations.

Another layout template is used when the current information goal is the simulation of software. Figure 18 gives an example of the simulation portion of the presentation.

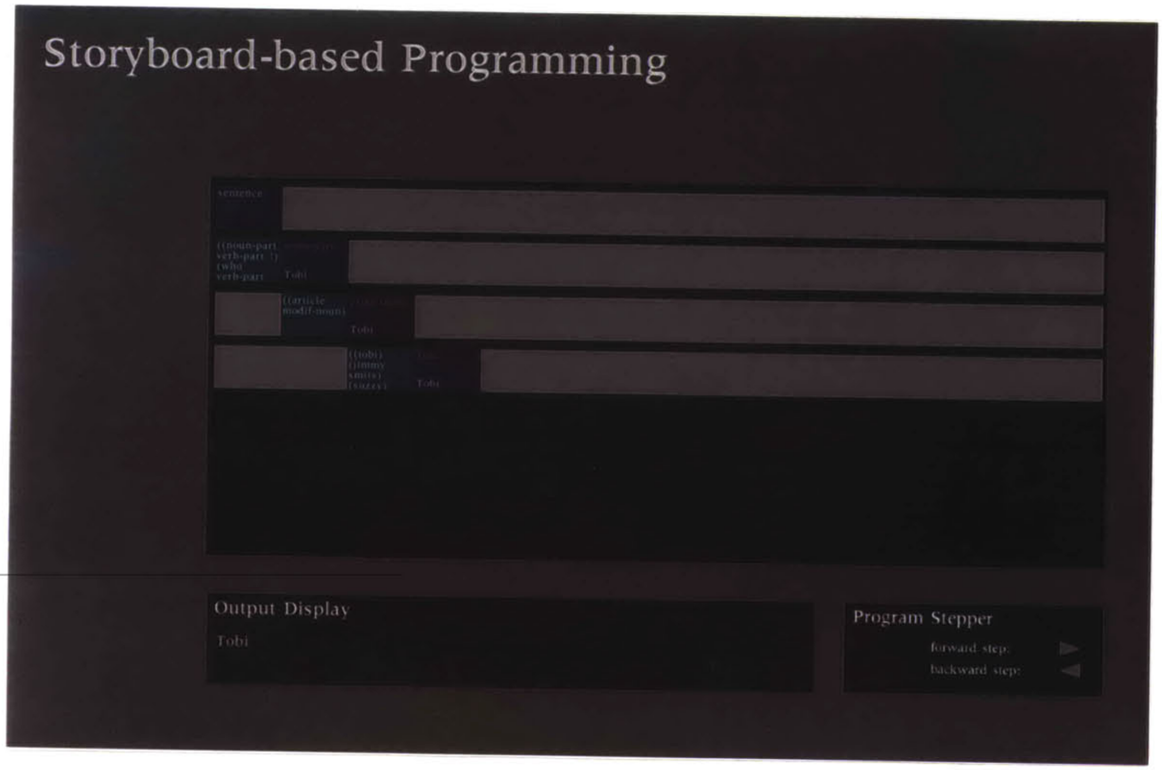

Figure 18 [Color Photograph]: Screen from a simulation portion of a HyperAdaptive multimedia presentation. Narration, describing what is shown, accompanies the visuals.

3. Multimedia objects

What does HyperAdaptive need to know about multimedia objects in relation to a style definition?

Each multimedia object is logged with various attributes, including media type, topic, domain (a superset of topic), what the object is an example of, what (if anything) it differs from, and how, or what it is similar to, and how. These attributes are sought out by the various methods associated with information goals. For instance, the "Find/describe features of the current software" goal seeks out objects which are examples of features of the current software. This would be indicated by the object's value for "topic" or "example_of." Figure 19 shows an example of an object description.

Object Name: Multimedia_storyboard_talk
Media Type: Narration
File Name: /u/michelle/DEMO/BROWSEDATA/mmedia_stbds
Topic: Multimedia storyboard
Domain: Interfaces
Example of: Graphical environments
Contrasts with: Non-graphical programming environments
Contrasts how: Graphical representation
Similar to: Non-graphical programming environments
Similar how: Events over time
Keyword of: None

Figure 19. A description of a particular multimedia object, the narration describing the multimedia storyboard.

An important element of HyperAdaptive's design is that objects are almost never described in terms of high-level style attributes. In other words, the objects will not be individually labelled, "researcher" or "nonresearcher." Such high-level labelling would mean that the knowledge about the objects could only be useful in the context of those style scripts. Instead, the objects are described in terms of medium-level semantic attributes, such as *example_of: current feature*. The knowledge represented within the style scripts bears the responsibility of translating from the high-level styles to these medium-level object attributes. This approach provides greater flexibility, reusability and inferencing power, allowing one to add a new style script which uses the same sets of objects used by another style script, but toward different ends.

C. Designer/Researcher Interface

HyperAdaptive provides two interfaces for the designer and/or user, in addition to the authoring tools provided by MMSE.

Logging Tool

The logging tool, shown in Figure 20, is used to create a description of a multimedia object, which would then be added to the multimedia database used during the search process. The logger can specify different sets of objects, which can be loaded in separately for different hypermedia applications. The information is stored in an easily readable (and editable) ASCII file.

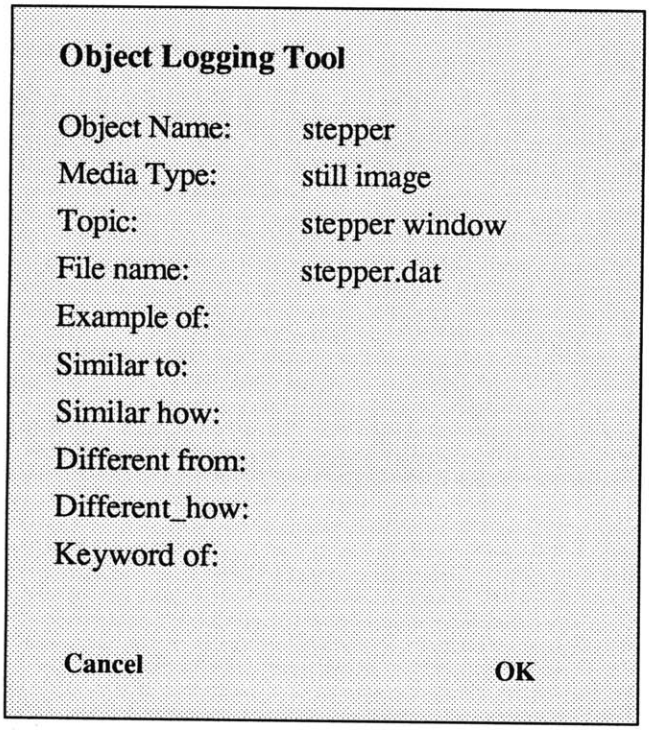

Figure 20: Illustration of HyperAdaptive's interactive logging tool, as it would appear in mid-use.

Script Inspector

The Script Inspector, as illustrated in Figure 21 below, presents the composite script structure. This window appears just after the script has been generated, and just when it is about to play back. It is displayed on the left screen of the two-screen display, so that it can be viewed while the multimedia presentation plays back.

Script Inspector

Researcher Technical

- Intro
- Idea
  - Compare and Contrast
    - Non-graphical Programming Tools
    - Graphical storyboards: events over time
- Early Version Features
  - Segue to Early Version
  - Find/Describe Version's Features
    - Tracks
    - Stepper
    - Function Icons
    - Function Call Complete
    - Output Window
    - Output Spoken
- Early Version Limitations
  - Problem: Vanishing Function Icons
- Current Version's Features: Differences
  - Smaller Function Icons
  - Use of Type Translucency
  - Return Value Displayed

Figure 21: The Script Inspector shows the composite script generated.

As each information goal (and subgoal) is reached in the presentation, its label is highlighted in the script inspector.

A "View All" button presents all possible composite scripts, based on the style selections currently available. This allows a comparison among the different composite scripts which could be generated.

D. Generalizability

How general are HyperAdaptive's scripts, methods and knowledge representation? Because the current implementation focuses on a particular domain, the generalizability has not been proven. However, the search techniques and basic knowledge representation scheme are based on artificial intelligence methods which have proven to be significantly generalizable. An open question, for this and many knowledge-based systems, is how much the specific domain knowledge could limit generalizability. Fortunately, HyperAdaptive's architecture should allow changes in domain knowledge to be reflected in object descriptions, and in adaptation of the script styles, while the underlying engine remains little changed.

IV. Conclusion

HyperAdaptive supports the adaptation of dynamic multimedia presentations within a hypermedia application. The style of the presentation varies both in content (selection, ordering and emphasis among components) and form (visual emphasis, coordination of various media, and layout). HyperAdaptive automatically generates a multimedia presentation by using knowledge-based representations of different presentation styles as guides in its search through a database of richly described multimedia data. The underlying representations of presentation styles also guide the multimedia design decisions.

The knowledge representation, search and inferencing of HyperAdaptive combines approaches found in the fields of knowledge-based graphics, graphic design, intelligent tutoring, narrative and rhetoric. HyperAdaptive's contribution lies in applying insights and formalisms from these fields in the context of adaptive multimedia presentations of research projects, and in creating a testbed for further experimentation.

The current work represents a limited foray into the extensive research area it enters. HyperAdaptive could be extended in several ways, to further experiment with and explore the multifaceted challenge of adaptive hypermedia. Below are several promising areas of extension.

Refining existing style structures

The current prototype includes two dimensions of adaptability: researcher/nonresearcher and technical/nontechnical. A logical next step would be to refine the existing style structures, and to create new ones. An important step in refining the existing structures would be to apply them to other domains. Such a step would bring out where the style structure should be adapted, or when different versions of the same structure should be available for different contexts. For instance, the "researcher" script may take different shapes in describing different sorts of research.

Extend internal knowledge representation of multimedia objects.

The addition of new and different style structures would bring the need for new and different descriptions of the multimedia objects used by these styles. As the system grows, the old representations could become drastically different to support greater generality. For instance, the objects could be described as part of a larger semantic network to represent the knowledge and of the domain at hand, and less as atomic objects with certain specific relations (such as *similar_to*) to specific objects. Such a network could allow some of the comparative relations to be computed, instead of a human logger being required to make them explicit.[1]

Extended visualization and modifiability

The current system provides some visualizations of the internal knowledge representations. These could be extended to allow direct manipulation. This would allow a non-programmer to modify and create script structures, and would make the system an interactive toolkit for building adaptive presentations.

Graphical logging of multimedia objects

HyperAdaptive currently supports interactive, but purely textual, logging of multimedia objects. The system would be greatly enhanced by converting

---

[1] Alan Ruttenberg, personal communication

this into a more graphical tool. Davis (1991) has developed a graphical logging tool for video clips, based on the use and manipulation of complex iconic representations of video characteristics, such as time and place of the video shot. The adaptation of Davis' system for the text, graphics, sound and images in HyperAdaptive's database is well worth exploring.

Bibliography

Aristotle. The Rhetoric. Published in The Rhetoric and The Poetics of Aristotle (1954) Translated by W.R Roberts. Random House Modern Library. New York.

Beach, R. & Stone, M. (1983) Graphical Style: Towards High Quality Illustration. Computer Graphics: Proceedings of SIGGRAPH '83. Vol. 17 No. 3 July 1983.

Bender, W. & Chesnais, P. (1988) Network Plus. SPSE Electronic Imaging Devices and Systems Symposium, Jan. 1988.

Bordwell, David (1985) Narration in the Fiction Film. University of Wisconsin Press, Madison, WI.

Bush, Vanevar (1945) As We May Think. Atlantic Monthly Vol. 7, 101-108.

Cooper, Muriel (1989) Design Quarterly. Vol. 142. MIT Press, Cambridge, MA.

Davis, Marc E. (1991) Director's Workshop: Semantic Video Logging with Intelligent Icons. Proceedings of the AAAI-91 Intelligent Multimedia Interfaces Workshop. Anaheim, CA.

Don, Abbe (1990) Narrative and the Interface. Published in The Art of Human-Computer Interface Design, Ed. B. Laurel.

Dondis, Donis (1973) A Primer of Visual Literacy. MIT Press. Cambridge, MA.

Feiner, S.K. & McKeown, K.R. (1990) Coordinating Text and Graphics in Explanation Generation. Proceedings of AAAI-90. Boston, MA.

Gatta, K., Lange, G., Lyons, M. (1991) Foundations of Graphic Design. Davis Publications, Inc. Worcester, MA.

Goldman-Segall, Ricki (1991) Three Children, Three Styles. Published in Constructivism, ed. by E. Harel, Ablex Publishers, Norfolk, VA.

Greenlee, Russell L. (1988) From Sketch to Layout: Using Abstract Descriptions and Visual Properties to Generate Page Layouts. MSVS Thesis in Media Arts and Sciences, MIT.

Grice, H. Paul (1975) The Logic of Conversation. Philosophical Review.

Keller, Evelyn Fox (1985) Gender and Science. Yale University Press, New Haven, CT.

Kince, Eli K. (1982) Visual Puns in Design. Watson-Guptill Publications, New York.

Klein, E. & Peneda, L (1990) Semantics and Graphical Information. INTERACT '90: Human-Computer Interaction.

Latour, Bruno (1987) Science in Action. Harvard University Press, Cambridge, MA.

Lieberman, Henry (1988) Design by Example. Unpublished paper.

MacKinlay, Jock (1986) Automating the Design of Graphical Representations of Relational Information. ACM Transactions on Graphics, Vol. 5 No. 2, 110-141.

MacNeil, Ron (1991) Capturing Multimedia Design Knowledge Using TYRO, the Constraint Based Designer's Apprentice. To appear in Proceedings of the 1991 SPIE Symposium on Electronic Imaging.

Marks, J. & Reiter, E. (1990) Avoiding Unwanted Implicatures in Text and Graphics. Proceedings of AAAI-90. Vol. 1, 450-456.

Meyrowitz, Norman (1989) Hypertext -- Does it Reduce Cholesterol, Too? Hypertext '89 Keynote Address.

Neal, Lisa (1991) User Modelling and Tailorable, Adaptable and Adaptive Systems. Tutorial notes from course delivered at CHI '91: Conference on Human Factors in Computing Systems. ACM/SIGCHI, New Orleans, LA.

Polson, M.C. and Richardson, J.J., eds. (1988) Foundations of Intelligent Tutoring Systems. Lawrence Erlbaum, Hillsdale, NJ.

Purcell, Patrick (1988) Computer Environments for Design and Designers. Design Studies. Vol. 9 No. 3 July 1988, 144-149.

Robin, Laura H. (1990) Personalizing Hypermedia: The Role of Adaptive Multimedia Scripts. MSVS Thesis in Media Arts and Sciences, MIT.

Robin, Laura H. (1991) Temporal Adaptation of Multimedia Scripts. To appear in Proceedings of the 1991 SPIE Symposium on Electronic Imaging, San Jose, CA.

Schank, Roger (1990) Tell Me a Story. Charles Scribner's Sons, New York.

Schank, R., and Riesbeck, C.K. , editors (1981) Inside Computer Understanding. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Schapiro, Meyer (1973) Words and Pictures.

Self, John (1990) Theoretical Foundations for Intelligent Tutoring Systems. Journal of Artificial Intelligence in Education, Vol. 1 No. 4, 144-149.

Sleeman, D.H. and Brown, J.S. (1982) Intelligent Tutoring Systems. Academic Press, London.

Smolensky, P., Bell, B., Fox, B., King, R., Lewis, C. (1987) Constraint-Based Hypertext for Argumentation. Hypertext '87 Proceedings. Association for Computing Machinery.

Strohecker, Carol (1991) Why Knot? Ph.D. dissertation in Media Arts and Sciences, MIT.

Weitzman, Louis (1988) Designer: a Knowledge-Based Graphic Design Assistant. MCC Technical Report Number ACA-017-88.

Wingler, Hans M. (1969) The Bauhaus. MIT Press, Cambridge, MA.

Zellweger, P. (1989) Scripted Documents: A Hypermedia Path Mechanism. Hypertext '89 Proceedings.