Viewpoints on Demand:

Tailoring the Presentation of Opinions in Video

by

Gilberte Houbart

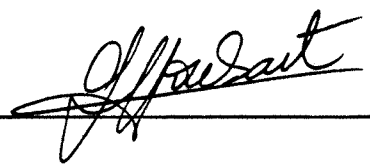
Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in Partial Fulfillment of the Requirement of the degree of

MASTER OF SCIENCE in Media Arts & Sciences

at the

Massachusetts Institute of Technology

September 1994



Signature of Author
Program in Media Arts & Sciences

August 5,1994

Certified by
Glorianna Davenport

Associate Professor of Media Technology
Thesis Supervisor

Accepted by
Stephen A. Benton

Departmental Committee on Graduate Students
Program in Media Arts and Sciences

Viewpoints on Demand:

Tailoring the Presentation of Opinions in Video

by

Gilberte Houbart

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in Partial Fulfillment of the Requirement of the degree of

MASTER OF SCIENCE in Media Arts & Sciences

at the

Massachussetts Institute of Technology

September 1994

Signature of Author
Program in Media Arts & Sciences

August 5,1994

Certified by
Glorianna Davenport

Associate Professor of Media Technology
Thesis Supervisor

Accepted by
Stephen A. Benton

Departmental Committee on Graduate Students
Program in Media Arts and Sciences

Viewpoints on Demand:

Tailoring the Presentation of Opinions in Video

by

Gilberte Houbart

Thesis readers

Reader

Michael Hawley
Assistant Professor
of Computer Science
and Media Arts and Sciences

Reader

John S. Driscoll
The Globe Newspaper Co.
Vice President

Viewpoints on Demand:

Tailoring the Presentation of Opinion in Video

by

Gilberte Houbart

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning on August 5, 1994
in partial fulfillment of the requirements for the Degree of Master of Science

Abstract

A model of interaction is presented for a home based system generating evolving documentaries that can be tailored according to the viewer's interests in a given viewpoint or story. This model takes advantage of the strong points in the linear experience that television and movies have traditionally offered by letting the story flow while allowing the viewer's intervention to constrain it.

The journalist works in this personalized context by shaping the material for the video database using graphical annotations for video content and story structure

The framework developed is used to reveal a special angle on the Gulf War: information technologies turned this event into a landmark in the history of not only media but also warfare. The amount of information available to the troops and the public reached levels never seen in previous conflicts.

By adjusting "content knobs" and selecting headlines the viewer sets the way the story will be presented. You might favor the view of a particular journalist or of the view of the former Director of Information for the Pentagon on the selected issue. Maybe you had planned to spend two hours but changed your mind in the middle of the documentary, asking for a shortened version. All these adjustments increase control over content. The scenario I just described could be seen as a component of a "smart VCR".

Thesis Supervisor: Glorianna Davenport
Title: Associate Professor in Media Technology, MIT Media Lab

---

This work was sponsored by the News In the Future Consortium

Acknowledgements

I wish to thank with this thesis my advisor Glorianna Davenport for her moral support and for giving me the opportunity to grow ideas in such a unique place as the Media Lab, since the very first day she let me contribute to one of her projects. I thank her also for letting me pick and develop the topic I became interested in.

Thanks to Walter Bender and all of the NIF members for making NIF happen and turning it into a social event within the lab.

I hope my readers, Jack Driscoll and Mike Hawley, will forgive the wait for the first drafts. It was difficult for me to hand them a document I was not completely satisfied with. Jack is the archetype of a true journalist and I have been an admirer of his talent and modesty. I find myself very fortunate to have him as a reader. Mike made the terrible mistake when I first visited the Media Lab in February 1990 of helping me out when it mattered. Now that he is a professor not only he does find himself reading my prose but he will have to tolerate my friendship for a long time, I hope.

I am very grateful to Walter Robinson, David Greenway and Jonathan Wiggs from the Boston Globe to have accepted to be interviewed. Help finding footage from Manulel Ortigoza at Televisa was also very much appreciated.

I had a wonderful time with the Interactive Cinema gang: David Tamès who held the camera for my interviews and helped me so often in the editing room, not to mention the passion for movies, the great strawberry margaritas and comforting words; David Kung was a precious and never ending source of silliness and knowledge about television and movies; I will not forget the Badoit and madeleines Scott brought back to treat my home sickness; Stuart Cody always answered my questions about video equipment and never refused to help; Mike Murtaugh and Dan Hartman were a great team to work with; I hope Lee Morgenroth will forgive me for not having tried yet the "Files of Dr. Bern":I promise it's on my list after this document is finally accepted; Kevin Brooks was the best office mate one can dream of, his ready smile and kindness helped a lot; Betsy Brown's humor and understanding of MIT bureaucracy contributed to make life much easier; Ryan Evans was a great source of toys of all kinds, not to mention helping me with Lisp; Mark Halliday shared with us his passion for cinema, I will remember with nostalgia the lunch movies; and I guess Eddie Elliot will remain the author of the Streamer and Mr. Eddie for the rest of eternity(!)

I had so much fun taking breaks at Wave and Elizabeth's or Judith and Joe's parties, enjoying endless attempts to predict the future of media. David and Sue Sturman were of great help on critical things. The rest of the Media Lab community, faculty and students, also contributed to deal with all the obstacles paving the road. Thanks to Tod Machover for keeping me posted with French news in media and for the great talks. Thanks to Pattie Maes and Ken Haase for the advice on agents and knowledge representation. I will never forget the wonderful and unusual gatherings at Marvin and Gloria Minsky's house. Marvin and Pat Winston are probably those from whom I learned the most at MIT because they teach the most important which is humor and to think in many different ways.

Many thanks to Alan Wexelblat for helping debug the FrameMaker layout for my thesis; to Atty Mullin for loaning his high quality microphone; to Warren Sack and Jennifer Gonzales for the support and great discussions; Bob Murray, Ben Lowengard and David Blank-Edelmann helped numerous times with Mac and network trouble shooting. The Media Lab would not be the same without Linda Peterson and Santina Tonelli, the incorruptible presence reminding us the many rules and deadlines of MIT life, and occasionally offering an understanding ear.

Without all the members of the "2d floor" we would be worrying much more about the cost of research and education and would certainly have much less electronics per square feet. Thanks in particular to Chris Gant, John Hynes, Rebecca Prendergast, Valerie Eames, Susan Browne, who tried their best to make the demo routine happen smoothly, to Pat Peterson for her patience with the issue of Frames, and to James Blanchflower who was so helpful.

Thanks to Zane Vella and Coco Conn for their moral support. I hope some day we can work together. In the last year, Nicolas de St-Arnaud and Sandra Gordon provided refreshing French conversation and a newborn friend-

This work is dedicated to the memory of Cristina Bernal, the "grande dame" whose wisdom and creativity I always hoped to inherit.

Contents

1 What is this all about?

In May 1968 I was 6 and riots were setting Paris on fire. We were living in a small village near Paris and my parents bought our very first television to watch the reports on the events. Twenty-six years later I was in Boston watching CNN broadcasting live images of the Gulf War. Technology had brought me closer to the action in time for the second event, but I was as shocked to realize that violence is always around the corner. It had been one hour away from me, it was now thousands of miles away. Immediacy was not as crucial as awareness of potential dangers and how they take shape. To me, the real purpose of news is to keep you informed so that you can form an opinion and make informed decisions in your own life at a personal and global level.

Looking back in time is equally important: it is only after it was over that key aspects of the Gulf War were understood, in particular that it could be considered as the first information war [see chapter 5]. Documentaries and history books serve a particular purpose. They give us a long term in-depth perspective that news doesn't have and help us to better analyze news itself. It is to be hoped that the tidal wave of digital media will bring a revival of the documentary. However it would be a mistake to put the technology first. How can we add value to the TV audience of the 500 channels?

nal. The Internet already serves as a model for communication and publishing. Information is made accessible to Internet users (published) on servers from which it can be retrieved on demand, or it is broadcast digitally to members of discussion list and readers of electronic publications.

Retrieving viewpoints on a given issue from an online video library will be similar to checking an *encyclopedia* that has a *perspective*. On the one hand one will have the benefit of storytelling and perspective as in a newspaper, magazine or documentary. On the other hand, one will request expert answers with any depth on a given issue as in an encyclopedia.

The ultimate goal is to create virtual characters that embody the perspective of an expert and answer questions (or argue with you) on the fly without pre-recording answers. This means being able to represent someone's viewpoint and knowledge of the world and then have the system infer from that an answer. What if you could request from a virtual representation of Marvin Minsky or John F. Kennedy their view on war-related research and the conflict in Bosnia? Marvin Minsky was probably never interviewed on the topic and JFK is not available anymore for questions. However both have a certain way of approaching an issue, a certain style, various principles and opinions -- a viewpoint in short -- that could be applied to a new situation.

3. Viewpoints on demand: the specific problems addressed

In this thesis, I have addressed one specific aspect: how to dynamically personalize the presentation of a story to put the emphasis on a given viewpoint. At a more global level, it requires an understanding of how to reshape a documentary on the fly as the story unfolds. How can a video story be transformed to adapt to external constraints applied to the content?

An appropriate paradigm might be the smart VCR in the sense that it helps control the linear presentation of video content by making the machine smart about the content itself, not just the physical medium it is recorded on. A standard VCR only lets you fast forward, pause, stop and rewind. A videodisc player will do better by letting you directly access a

progression in the content. There are many ways of applying the principles described.

In the fall of 1993 I gathered information and interviews about the role of information technologies during the Gulf War. This original material was used to create a database of viewpoints and to develop an authoring model for such databases. The model of interaction for the viewer was based on 1) defining levels of interest in viewpoints and issues that can be changed as you watch 2) defining levels of interest in a given headline 3) content-based playback controls 4) browsing arguments. One important part of the work was to create an engaging interaction for the viewer. Another important part was to understand and rethink the role of the author in a personalizable environment.

The rest of this thesis is composed of eight chapters: *chapter two and three* analyze the roles of the viewer and of the author, *chapter four* explains how the engine works selecting the clips and ordering them, *chapter five* is about the content itself (the role of Information during the Gulf War), *chapter six* relates this thesis to previous work, *chapter eight* is dedicated to future directions for this work, the last chapter concludes.

2 The Viewer

Presenting Content to the Viewer:
User Interface & Consumer Behavior

Customizing viewpoints & issues
with content knobs

"No interface at all is better than any interface"
Jean-Marie Hullot (NeXT)
Steve Jobs (NeXT)
Nicholas Negroponte (MIT Media Lab)[1]

This chapter takes the point of view of the viewer, explains the design issues about the interface and describes how it works. The first section tries to support five points that determine the model of interaction described in the second section. The third section describes the actual scenario of an interaction. The last section is a series of Questions & Answers.

The five main points supporting the model of interaction are:
1- Smart agents will contribute to eliminate unnecessary interfaces.There still will be a need to control and shape a given program and that task will be handled by interface agents.
2- The common assumption that home and work products should obey different rules of design is wrong.
3- The combination VCR/remote control is a first attempt at constraining the flow of video. It is limited because it lacks knowledge *about* the con-

1. I heard this first from Jean-Marie Hullot around 1987, then from Steve Jobs and finally from Nicholas Negroponte. It is not clear which one influenced the other. Any one of them could have come up with it.

"There is always an interface. Some we identify as such, some are so completely part of our vernacular that we do not think about the interface as something other than the thing itself" (Glorianna Davenport 6/93). The current interface for "Viewpoints on Demand" uses a mouse for input and gives feedback visually. Having a conversation instead might feel more natural. Feedback could then be verbal or visual. In all cases the function offered by the interface does not change. The interface still exists.

2- Work and home products don't obey different rules of design

There is a common assumption that in order to understand customer behavior with technology-oriented products it is essential to make a distinction between the home and work environments. A typical argument is that at home we don't have the strong motivation that we have at work to accept complexity. At home we want to relax and use simple and useful or entertaining appliances like food processors or television. At work we have a task to accomplish and we need to be productive.

Paul Saffo (1992)
Consumer Technology Purchase Behavior in the 1990's - IFTF

Paul Saffo from the Institute for the Future sees a dichotomy between the consumer electronics industry and the computer industry (Saffo 1992). According to him: "We use tools to accomplish tasks, and we abandon them when the effort required to make the tool deliver exceeds our *threshold of indignation* -- the maximal behavioral compromise we are willing to make in order to get a task done". An average consumer has a low threshold while a teenage hacker has a high threshold.

The reward can be getting work done or being entertained. It does not matter to the customer where the product comes from (computer industry or consumer industry). Increasingly more people use on-line services like America Online or Compuserve, because they have a strong need to communicate. They are ready to go through the technical difficulties inherent in getting connected.

This doesn't apply just to consumer electronics but to all human activities. Hiking or windsurfing also require a tremendous investment from those who decide to participate in them. They do it, because the expected reward is high.

Upside, April 1994
An interview with Eric Nee of Scott Cook, CEO of Intuit

An essential difference between the consumer industry and the computer industry is *the willingness to take into consideration consumer feedback.* Consumer products are extensively tested and studied in surveys in order to track reactions from customers. Computer products don't go through such thorough testing (see interview of Scott Cook, CEO and founder of Intuit which dominates the personal finance and small business software market). High tech products like the Newton or the NeXT computer did not pay sufficient attention to the customer feedback when they were first brought to market. This is not to say that the technology itself was not good, especially for the NeXT: while the first Newton suffered in its ability to transcode handwritten input which had been marketed as a key feature, the technology of the NeXT was sound. Rather, the market reaction reflects that some key aspects which matter to customers were not taken into account; as a result, it was impossible to turn high quality technology into a successful product.

Work and home products obey the same rules: they must answer a need and provide a reward worth the effort and the cost.

3- VCRs are a step in the right direction but they lack knowledge about the content

VCRs (or video cassette recorders) are basically affordable consumer electronics devices that are hooked to a television and have many interesting features:

• you can record a TV program or movie on tape which implies in turn that:

- you can watch a program when you want to watch it
- you can watch it many times

Brenda Laurel 1989
A Taxonomy of Interactive Movies, New Media News, Vol.3 issue 1

What she seeks to solve is the problem of *"creating compelling interactive plots"*. Her solution is to "endow the system with dramatic expertise and to give it the ability to manipulate, through an internal 'action calculus', the shape of the whole action. Such a system would incorporate the brainpower of a good playwright, writing his script in real time, where one of the characters literally has a life of his own".

With news and information, you do not expect to modify the events you want to be informed about. You want to be able to experience them as an observer, as closely as they actually happen. A second-person type of interaction (or no interaction as all) is more relevant in this context. By no interaction at all, I mean that agents that know you would search information on your behalf.

That is not to say that you don't ever want to simulate your participation in real world events. However, the goal is then different: you want to explore "what if" scenarios, observe the outcomes and eventually compare them to what actually happened. This is extremely appealing for education purposes. If schools had "holodecks"[1] that could simulate the French Revolution, it is expected that more children would, for instance, be more receptive toward history.

5- Content knobs allow linear media to become interactive

We already established that VCRs would benefit from meta-knowledge about the content to adjust linear narrative. What type of interface would allow this?

Before answering the question we should keep in mind that there are alternatives to linear playout of motion picture sequences. We need to understand the type of interaction they provide. Most of interactive media has been based on models in which the experience is constantly interrupted ("suspension of disbelief"). How to characterize the various alternatives?

Brenda Laurel proposes in the same paper mentioned earlier a taxonomy of interactive movies which distinguishes 3 basic forms: *navigational* (like movie maps), *narrative* (like a branching story) and *dramatic* (in

1. Virtual reality system used as plot device in a TV show called "Star Trek, the Next Generation".

Ryan Evans, 1993
Logboy and Filtergirl, MS
MIT Media Lab

In this model there are no predefined threads. The final arrangement of the shots (or clips) is the result of applying various constraints to the content and high-level organizing rules. These help select and order the shots (see chapter 4 for more details).

Benjamin Rubin, 1989
Constraint-Based Cinematic Editing, MSVS MIT

This fits in the narrative form as defined in Brenda Laurel's taxonomy:

- it offers a control *activity* in which the viewer controls the resulting story with content knobs.

- the *frequency of interactivity* depends on how often the viewer wants to modify the parameters of the story; the *range* is constrained; the *significance* is great but not maximal, as switching viewpoints on demand changes the way the story is told but not the events themselves (which is in fact the whole point). Maximal significance would go against the goals of the system.

- The *personness* is second.

2. The resulting model of interaction

As a result of this inquiry, I have built a smart device, a sort of smart VCR, that can adjust linear content based on the various points stated in the previous section. The interface uses content knobs to offer more control over the content without getting in the way when the viewer experiences the story. It provides the means of controlling the selection and shape of the *content* of video pouring on the screen. Video logging makes it possible by providing the system with *information about the content.* Stories are "advertised" to the viewer using **headlines**. The mode of interaction is second person.

In chapter 7 I propose a specific type of branching structure based on text connectors (like BUT, AND, OR) that can be super-imposed to navigate through opinions at the shot level, the shot becoming a node.

The following section describes an actual scenario of interaction. I designed it to be used more specifically in the context of presenting viewpoints on given issues.

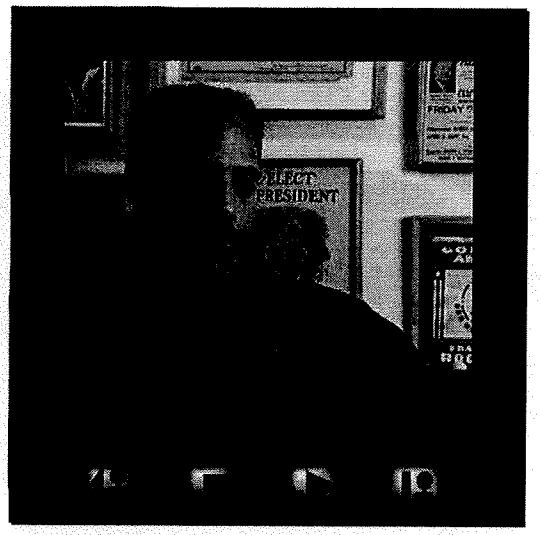
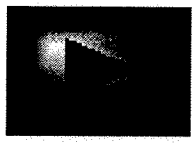


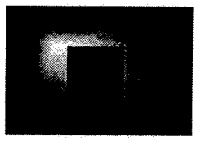
Figure 3 - Once the story is generated, a small panel pops up.

The control buttons look like VCR controls:

play



stop



next speaker

previous speaker

Figure 5 - The Boston Globe logo gives a hint of where the people interviewed come from.

journalists interviewed - give a hint about the person interviewed [fig.4]. Starting with a short version of one of the stories gives the viewer a sense of the contents but does not give background on the various speakers. It is as if you were watching a trailer.

Dynamic changes of content

By hitting *play* the viewer can watch the story generated based on his or her input. If the settings are changed - in particular the viewpoint - the selection will be different, but in the current version overlap cannot be avoided as the system does not keep track of what the viewer has already seen. In particular the fact that the narration remains the same, makes this limitation more apparent.

How different will each playout be as the user changes the settings of the knobs? There are various content-related factors:

- how well a viewpoint is represented in one story (in short how much footage is available).

- how key is one viewpoint versus another in a story (a native from Ecuador can hardly have a position on how life is in Alaska unless s/he lived there). If the viewpoint of a key expert in a story is eliminated, the story might lose its interest.

For instance, in the specific database I created, while all viewpoints are related to the role of information during the Gulf War, different opinions stress different aspects. Two major views are information as seen on the

story structures and the selection of the main topic and of the viewpoints presented. However, this approach can get the audience to become more conscious about the world and the many ways it can be seen. Or one might want to have stories told only in a certain way, with a certain style we like.

Is it different for video or text?

Tailoring viewpoints remains obviously valid for text. It requires different solutions to select, organize and assemble the content.

3 Presenting and Representing Content by and for the Author

Sharing Control with the viewer

"Distributed" Design and Graphical Annotation

1. Redefining the role of the author in personalizable media

... I have brought them as it were into a continued relation, which as I take will be the pleasanter, because you need not trouble your remembrance with looking backe after former matters.

Thomas Gainsford (1623), considered to be one of the first news editors

A major research area in electronic publishing is "The Daily Me" (Bender et al.) or how to adapt content to the needs and interests of an individual. This implies collecting information from various sources and doing the layout *automatically*. A great concern in the press is to understand how to redefine the role of the author, and particularly the editor, in helping shape the information delivered to the reader or viewer.

The job of the newspaper (or magazine) editor involves identifying the most important stories according to the interests of the audience (or average reader/viewer). An editor also takes material brought in by reporters and puts it in perspective, fitting it into the overall layout designed to better communicate the message.

Bender et al., 1991
Newspace: Mass Media and Personal Computing, USENIX conference

J.T.Harrigan, 1993
The Editorial Eye SMP

The focus of this chapter is to give one example of collecting and editing information for an interactive personalizable medium: what are the tools needed and how can the editor apply a certain pattern while allowing freedom for the viewer? A plausible methodology seemed to be going through the actual process of identifying a potentially interesting content, investigating it, gathering interviews and shaping the result. Chapter 5 gives more details about the content itself. The present chapter

M.Murtaugh, 1994
ConArtists, a System for Graphically Representing Story Knowledge, MIT

Mike started the development of this tool in the summer of 93. Until May of 94 we went through many versions of the database as I was providing feedback about the needs and as Mike was coming up with new interface designs. In January of 94 as my interviews were all digitized, I started working on my "real" database and we had to freeze the design of ConArtist. A major obstacle had been that copying one frame from one location to another would take several minutes. Saving and loading the database would take up to 10 minutes. This was optimized in the version I stuck to.

K.B.Haase, 1993
FRAMER Reference Manual Internal Report, MIT Media Lab

ConArtist was written on top of Framer, programming language designed to be used for knowledge representations. In the manner of object oriented languages (OOL), Framer uses a hierarchy of "objects", known as *frames,* but it does not apply the classical class/instance model. In the OOL model, classes describe the general attributes of an object. All the instances of a class share the same attributes (color could be an attribute of the class car). Classes at the bottom of the hierarchy inherit attributes from upper classes (*Mercedes* would inherit from *car*). In Framer, attributes of a frame are other frames (the frame car would be annotated by the frame color which would in turn be annotated by the frame brightness). Inheritance can be provided by specifying certain frames as being prototypes: several frames can then inherit from a common prototype. Framer has also built some inference mechanisms that make it particularly well suited for Artificial Intelligence applications. Finally all Framer objects are *persistent* which means that they can be stored in a database file and reloaded.

A graphical interface to "draw" annotations

In ConArtist all video annotations are graphical. Taking advantage of that I defined a graphical language to categorize the content, describe relations between clips and the structure of the story itself. It is a language using links and nodes. It is applied at two levels: to define a story structure and to define relations between clips.

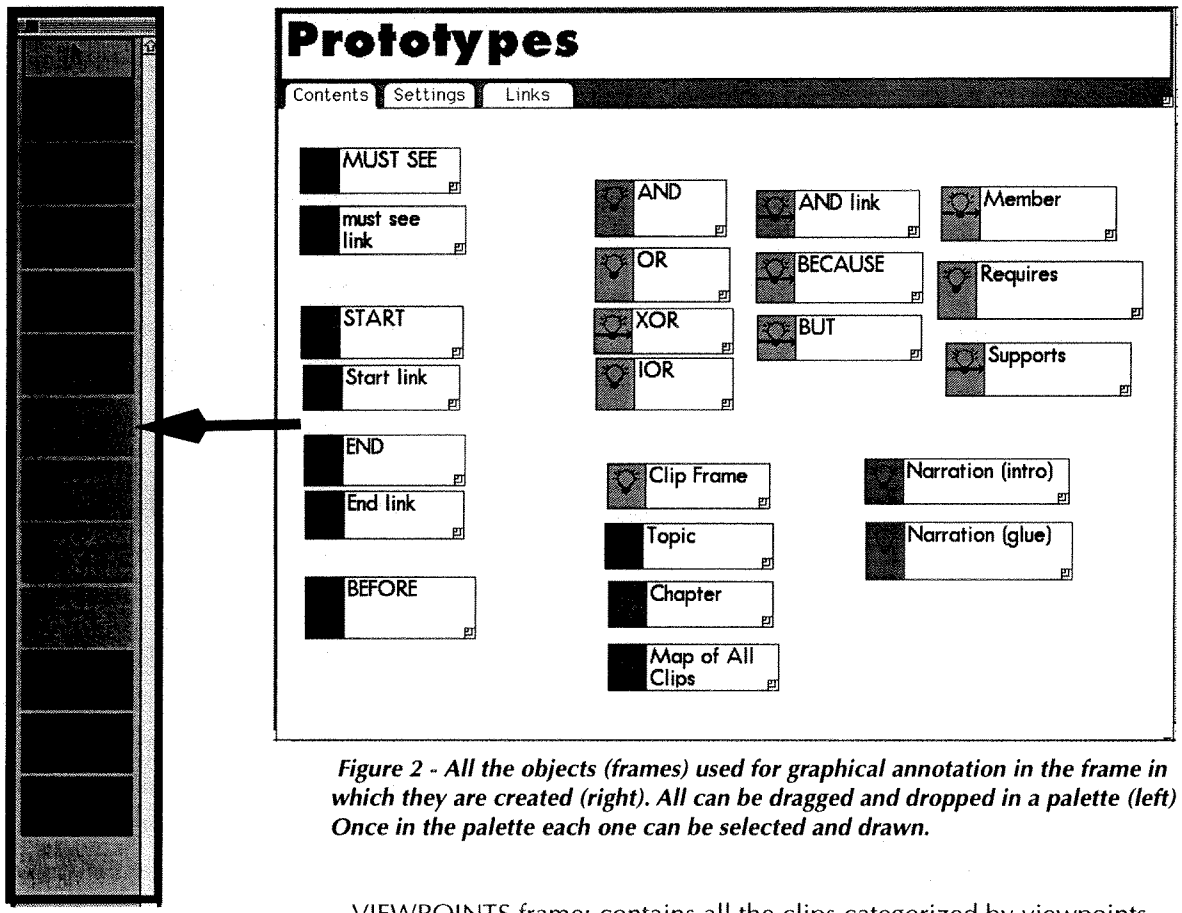


Figure 2 - All the objects (frames) used for graphical annotation in the frame in which they are created (right). All can be dragged and dropped in a palette (left) Once in the palette each one can be selected and drawn.

VIEWPOINTS frame: contains all the clips categorized by viewpoints and topic covered. In other words, it contains all the answers collected in the interviews. Answers are grouped in topic frames that contain in turn the clips.

STORY frame: contains the frames needed for the design of the story as well as the graphical description of the story.

COMMERCIALS frame: contains all the commercials available by category.

PALETTE ITEMS frame: contains all the graphical symbols that are used to annotate the video by tagging, grouping clips and drawing relationships between them.

3. Definition of a catalog of graphical annotations

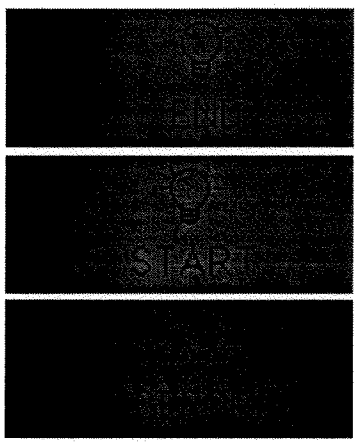
This section describes each notation used: links, containers or labels. They are all created and stored in a special frame [fig.2]. Once created

might be used.

CLIP frame: contains a movie frame and provides a text summary of its content.

Scheduling annotations

While a set of rules are embedded in the software in order to handle the ordering of the sequences, it is sometimes useful to impose a certain ordering:



END: To make sure a given clip is the first one shown in a chapter (if necessary).

START: To make sure a given clip is the last one shown in a chapter (if necessary).

BEFORE: In some cases, an order between clips has to be imposed. A typical example is the following: the speaker uses a pronoun ("*It* was important", "*she* knows what is important"). It is necessary to explain first to the viewer what "it" and "she" are referring to. [see chapter 7 on Future Work for a possible way of generating this automatically].

Logical annotations

Logical annotations serve two purposes:
1) to give the scheduling/selecting program an indication about how clips relate to each other logically.
2) to suggest to the viewer possible ways of branching to follow a thread of opinions or get additional detail.

left wing view for instance).

4. Inspiration from print media

Much of this authoring tool is inspired by print media.

Stories are announced with ***headlines*** asin a newspaper. In an earlier design, the viewer would specifiy a story by selecting a combination of words scattered across the screen. These keywords would then be used by the system as ingredients to produce a story. This was abandoned as keywords would lack context and the risk was that the viewer would select randomly words. This would not give the user a sense of being in charge. Headlines give the viewer a better hint at what to expect.

Stories are decomposed in ***chapters*** like a book. Chapters make it easier to set up interactivity;: the author uses them to define the general structure of a story; a chapter is like a box in which relevant clips are dropped; the final ordering and selection is not decided by the author (but the story structure *constrains* the final output).

An important result of experimenting with the interactive shaping of the content was that customizable content calls for "distributed design". By distributed design I mean that a clip might end up used in different contexts in different stories. It might have various meanings depending on the viewpoint. As the content was to be tailored to the viewer on-demand, I needed to be able to *visualize* various possible results. I needed also to instantly jump from one set of clips to another as the first set would inspire ideas for the second. ConArtist and Framer allow multiple copies to be updated in parallel.

These are characteristics of design general. When editing a traditional movie, an editor might want to visualize multiple possibilities prior to fixing a sequence. In non linear media supporting this process becomes a necessity as the final choice is specified by the actions of the user.

In this system, the editor does not focus anymore on what is the best final edit but s/he defines a general shape or structure for the stories and how they need to be constrained.

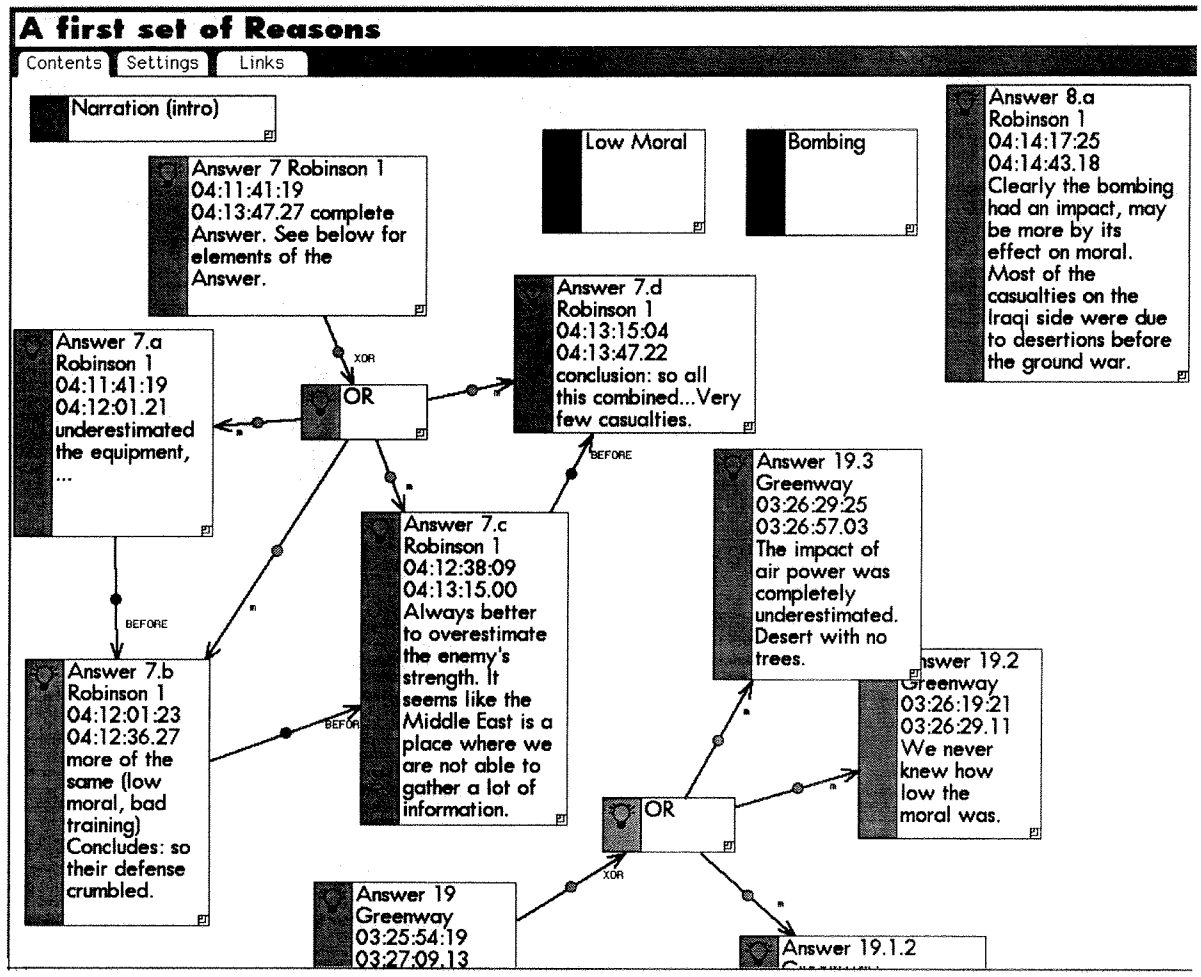


Figure 4 - Sample of the clips in the first chapter. Two OR frames indicate that two clips can be decomposed into several clips, giving more flexibility to the story generator. BEFORE links force an order on some clips.

Looking at the links existing between clips in the first chapter [fig.4], we note that some clips can be decomposed (OR links). In some cases, an additional ordering constraint is imposed (BEFORE links). There are two subtopics under which answers can be grouped: Low Moral and Bombing.

"Story Hunting under Pentagon Rules"

Any journalist remembers from the Gulf War the rules imposed on the press by the Pentagon in order to filter out any crucial information. This story tries to answer various questions: What were these rules? How were they applied? How did journalists react to them? What was the motivation of the Pentagon?

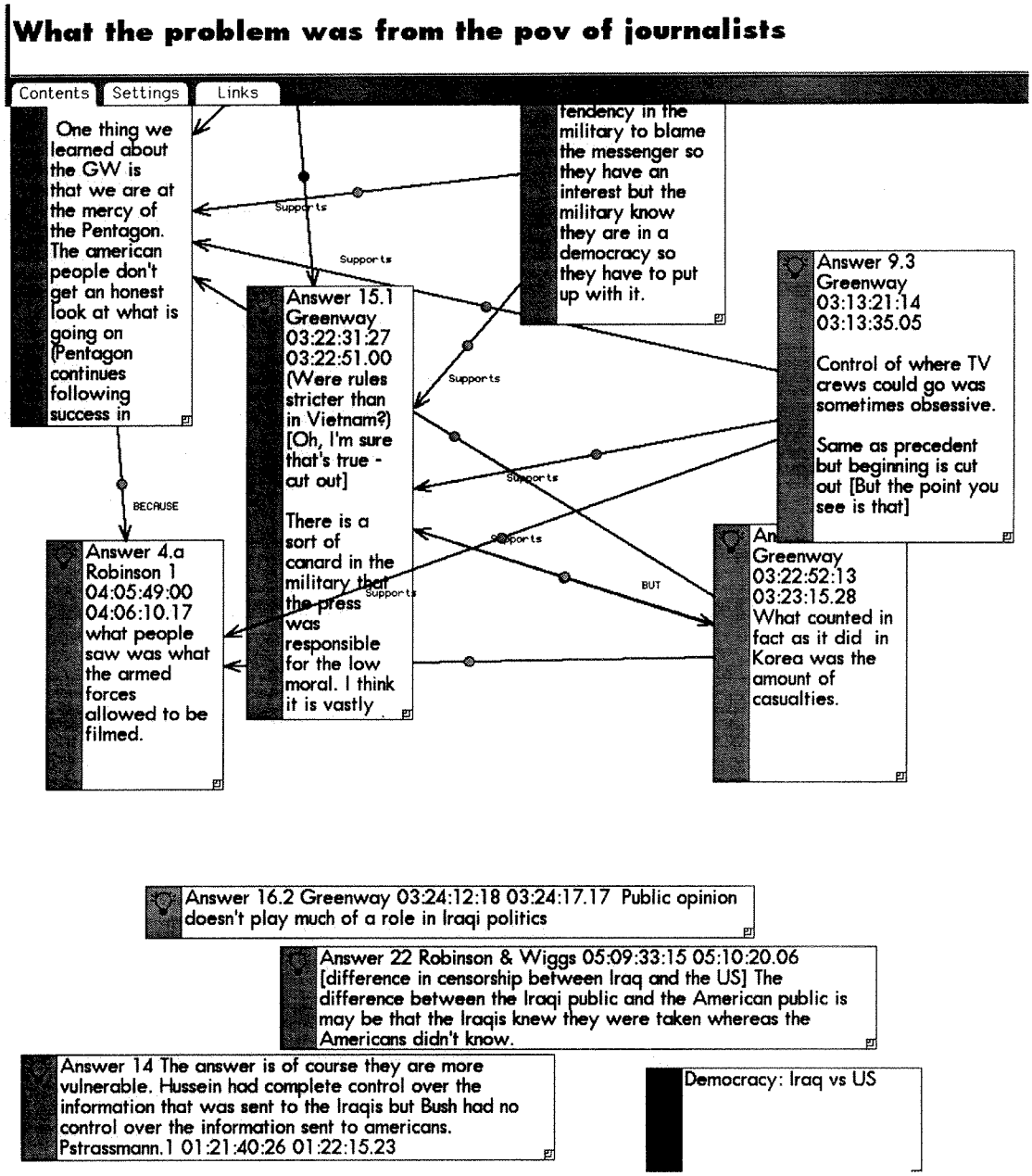


Figure 6 - The clips of the chapter "What the problem is about" and existing links between some of them.

3) What the problem is about

Presents reactions explaining why the press rejects partially or in total the rules defined by the Pentagon. Figure 6 shows clips supporting the point made by other clips ("supports" links). This is to be used in particular to support a given opinion with a combination of answers from dif-

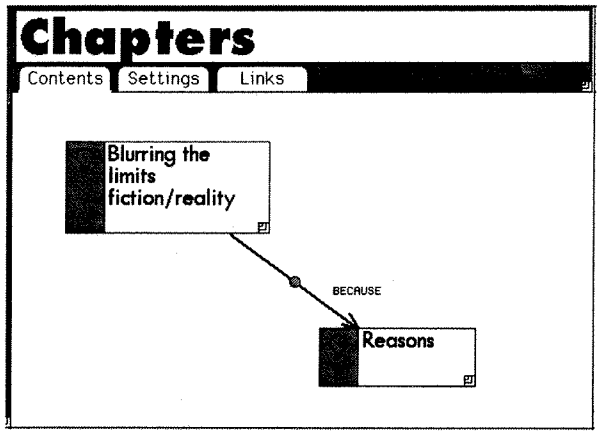


Figure 7 - Story structure

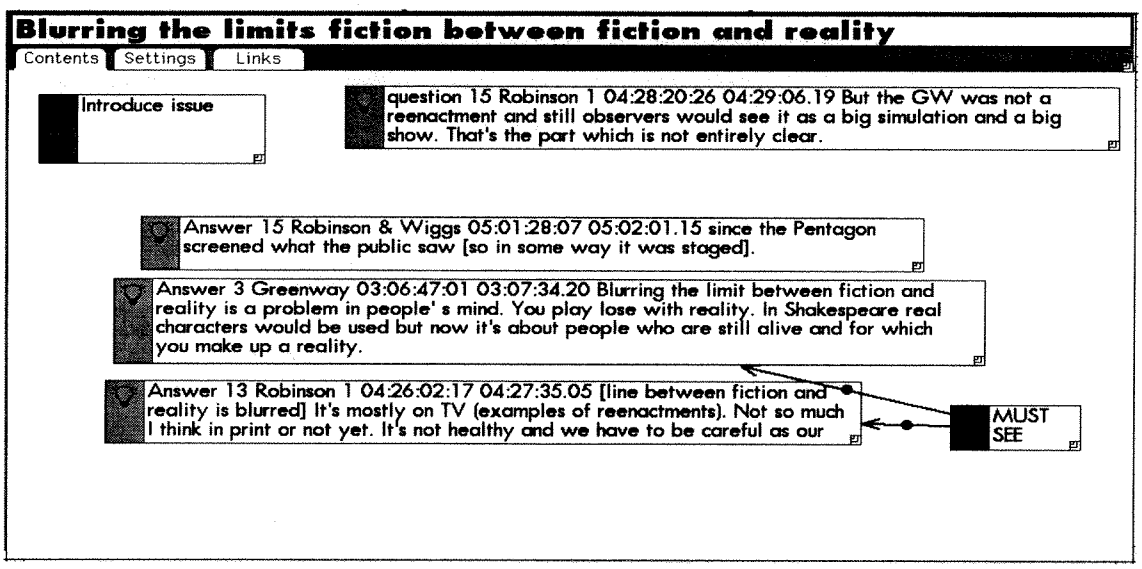


Figure 8 - Clips of the first chapter

"Being a Journalist during the Gulf War"

This story has also a simple structure with two chapters:the first one shows how information technologies were used by journalists, the other one reports anecdotes about safety issues. It does not matter in which way they are ordered.

2) My model of presenting viewpoints in relation to each other, in the context of a story defined by an author.

Concern for more participation from the viewer

Giving the audience more control over content might motivate many authors, specially if there is a market for it.

Linking a database to a discussion list

Such a framework would be of particular interest to a person (journalist?) creating a discussion list associated with a database of content. A tight relationship could tie together the database, the playback engine attached to it and the discussion list.

The Internet has been very successful at spawning virtual communities and discussion lists (the so called "news groups"). It has demonstrated that a visceral human need was not completely fulfilled, the need for communicating and exchanging opinions freely. A major problem faced by these discussion lists is the lack of tools for browsing a line of argument. It is also difficult to go back in time and browse the archived discussions. In addition, these groups are today limited to text because of the bandwidth.

The database could contain either past discussions or the equivalent of a documentary gathered by the owner of the list. In the second case it would serve as a base for discussion and actual links could be created between the database and the personal views expressed (or even gathered) by the members of the list.

7. Are we fooling the viewer?

One general concern might be that the viewer is fooled into thinking that because there is interaction and control of viewpoints, there are less chances for being misled and more control over information.

This is entirely in the hands of the author. The model for creating databases of viewpoints that I propose leaves an option that current media don't offer: interactive access to several viewpoints on a particular issue and the possibility to focus on one or another. However, the structure of the story, the selection of the content, the relations represented in the database between the viewpoints, how they are articulated one in reference to the other, are still in the hands of the author. The narration also

4 Behind the Scene: the Mechanics of the Playback Engine

Knowledge representation of opinions & viewpoints

First steps towards a storyteller agent

This chapter explains how the story generator works and describes the knowledge representation it relies on. The final section analyzes its weaknesses. Alternative solutions to improve the results will be described in the chapter on Future Directions.

1. Knowledge Representation

"Good representations make important objects and relations explicit, expose natural constraints, and bring objects and relations together."

"Once a problem has been described using an appropriate representation, the problem is almost solved".
P.H.Winston (Director of the MIT Artificial Intelligence Laboratory)

Using a semantic net representation

I used a semantic net representation, because it met the needs for expressing relations between various opinions. Each clip in my system is a comment, answer or bit of answer given about an issue. Several answers by one person (or group) contribute to build in the mind of an observer a model of that person's (group's) viewpoint. Answers or comments are denoted by nodes with links between nodes representing how the nodes can be *coordinated* to articulate opinions. These links need to be labeled and eventually handled as objects themselves. A semantic net also lends

The meaning of the nodes and links as well as their labels have been defined in order to let the system associate story and content information with clips. Story information was created as knowledge about how the content is orchestrated. Content information was focused on relations between clips that *coordinate elements of opinions.*

2. The Basic Principles

The current software presents viewpoints in a way that makes sense, informs the viewer and adapts to the parameters set by the viewer (given the eventual constraints imposed by the material available in the database). Clips are organized in a way that is inspired by TV news editing: a narration introduces an issue, asks a question and is followed by various reactions or answers. The main difference is that the narration is much shorter in order to leave more room for the actual expression of viewpoints. The story progresses through various "chapters". Each chapter may contain sub-topics. A selection of clips is based on the following rules or constraints:

- "must see" clips are all selected
- the order of the chapters is followed
- for each chapter the narration comes first and is followed by a reaction from each available speaker
- if there are several clips for a given speaker, they are grouped
- in the case where a chapter includes subtopics, the same rules applied at the chapter level are applied at the topic level
- the length of story requested by the viewer is taken into account
- two key organizing rules allow the focus on viewpoint:

A first, simple and quite powerful rule is to spend more time elaborating the viewpoint. This can be translated into a rule such as the following:

■ IF *A is the major viewpoint* THEN *maximize time allocated to A*

Another rule is to intercut characters' comments in a way that supports a given viewpoint. This can be translated into a rule like the following:

■ IF *A is the major viewpoint,* AND *B supports A* AND *C opposes A* THEN *juxtapose (A,B)* OR *juxtapose (B,C)*

The following example shows a context for applying this rule:

comment saying that *only* a few bombs missed their targets killing civilians. Losses could have been much more important without constant concern for avoiding highly populated areas. Some shots are intercut to show phone booths allowing the troops to call home free. Troops see images on television of supporters saying how much they are proud of them. They cheer. Images of successful bombing are shown with a voice-over commenting that the coalition was very successful at destroying the Iraqi's defense and indeed the war was finished in the wink of an eye.

3. Overview of the Story Generation Algorithm

The problem is to select the right clips and schedule them for playback according to parameters set by the viewer. The environment of the engine is composed of the database (clips and annotations) and the viewer's actions.

The action starts with efficiently parsing a tree: currently, the hierarchy has stories on top, composed of chapters, which are themselves composed clips. Some chapters have sub-topics. Tree search is a mechanism for which Artificial Intelligence has made major contributions, because quite often solving a problem can be translated into a tree search. The selection and scheduling algorithm is implemented as a "greedy" algorithm that traverses the story hierarchy. Greedy algorithms assume that local optimizations provide a fairly good global solution[1]. Since doing a global search of the database tree would take an excessively long time, the alternative is to optimize locally.

The pre-decomposition in stories and chapters helps minimize the breadth of the sub-tree to parse. Another advantage is that the sub-tree is not deep. This means that there is no need to evaluate and apply techniques like depth-first search or breath-first search. What makes the search challenging is the combination of constraints to be applied in the

1. The notion of "greed" as an algorithm comes from a simplistic model of the debates between socialists and libertarians. Socialists consider greed the source of all evil. Libertarians see greed as not necessarily negative. These debates cross over into economics, philosophy, government, and computer science, using jargon from all four fields.

within the threshold, the best speaker-weighted one will be chosen. Otherwise, the time closest to the desired will be chosen.

5 "It was a Knowledge War"[1]: Notes about the Content

"When we compare the new features of warfare with those of the new economy, the parallels are unmistakable. The day may well come when more soldiers carry computers than carry guns. Knowledge in short, is now the central resource of destructivity, just as it is the central resource of productivity".

— Alvin & Heidi Toffler

1. Information: a special angle on the Gulf War[1]

I became extremely interested in the topic of information and the Gulf War when I read about the impact of CNN's real-time coverage of the war, the heavy debate it triggered, the use of satellites and other information technologies by the media. But what really stimulated me was a conversation I had with Paul Strassmann whom I had previously interviewed in reference to his book, "The Business Value of Computers" .

1. The U.S. Army's Chief of Staff called Desert Shield/Storm the "knowledge war"

always a significant factor in the shaping of one's world." This approach to knowledge encouraging action and creativity contrasts with its Greek counterpart in which knowledge is about establishing a relationship between an idea and an objective world. *"One 'knows' the world not only passively in the sense of recognizing it, but also in the active shaping and 'realizing' of it".*

This philosophy of knowledge goes beyond warfare and appears very up to date in our information-loaded world. The major difference between today and the Warring State period is that information technologies now enable us to gather, communicate and process information extremely efficiently. The emerging problem today is to deal with the overflow of information. The difference between Vietnam and the Gulf War is stunning. This is why the Gulf War appears as a turning point.

On the military side

The Gulf War saw what has been called the largest single communications mobilization in military history. Starting with minimal capabilities in the region, a complex set of of interconnected networks were built at high speed. Numbers may be the best way to convey the massive presence of information technologies:

- 700,000 phone calls per day
- 30 million phone calls during the air war alone
- 30,000 radio frequencies
- By the end of Desert Storms, there were more than 3,000 computers in the war zone actually linked to computers in the U.S.

To put it into perspective, 700,000 phone calls a day is the traffic of a small town. There were 500,000 soldiers involved but obviously most of them were more busy preparing to fight than communicating. It was the first major conflict in which individuals brought their laptops to the battlefield. Troops were allowed to call home free (AT&T set up the phone booths).

According to Paul Strassmann, GPS[1] was *the* key technology of the war. It provided the mobility that was essential to the success of the vast operation that consisted in hitting by surprise the flank of the elite Iraqi Re-

1. Global Positioning System: portable device which accurately indicates the locationof a moving object anywhere on the earth surface

"The media itself became the star of the spectacle." (Alvin & Heidi Toffler, *War and Anti War*)

"In my wire service days, minutes made the difference between a scoop and being second. In live television, seconds made the difference [...] Superficiality is one of the perils of live television but hesitation is a mortal sin. Still, it is not easy to speak before you think [...] It was my turn again at the open mike. I said that all the bombs seemed to be hitting directly on target. I tried to qualify. We could see a lot from our ninth-floor perch but certainly not everything. We didn't know what was happening out there. I knew that in the critical first moments of the war, public opinion could form on our report. We had an open line to the world; there was no censorship at either end. The Iraqi minders had taken off for the bunkers." (Peter Arnett, *Live from the Battlefield*).

The radically new information technologies transforming the military were also transforming the media. All media increasingly rely on computers, fax machines, satellites, and telecom networks. This trend was illustrated in full scale during the Gulf War.

The media and the military

Control of information means allowing information to flow and controlling who gets what. A major conflict grew between the press and the military during the Gulf War. Many journalists criticized the Pentagon for censoring the information the press wanted to make available to the public.

For both the military and the press, satellites were playing a key role. Many satellites were repurposed to be used in the Persian Gulf. The military even used commercial satellites for non classified information.

3. Viewpoints: who was interviewed?

I interviewed three journalists from the Boston Globe who were foreign correspondents during the Gulf War and the former Director of Information at the Pentagon.

Two of the journalists had also taken part in the Vietnam war: David Greenway was a reporter at the time and Walter Robinson was in the army Intelligence. The third, Jonathan Wiggs, was a photographer in the Gulf pools set up by the Pentagon.

While working for the same newspaper, the three of them had different perspectives. Walter Robinson was extremely critical of the Pentagon rules. He also provided examples regarding the differences between

"We could see, hear and talk all through the war. After a few hours he could not" Air Force official Martin Faga

"An ounce of silicon may have been worth more than a ton of depleted uranium" Col. Alan D. Campen, USAF (Ret.)

"CNN went off the air at the very moment that the war called for a bomb from an F-117 stealth fighter to penetrate the city's main telecommunication center" Col. Alan D. Campen, USAF (Ret.)

"Frequency mangement challenges were enormous" Ernest May, "Intelligence Backing into the Future", Foreign Affairs Summer 1992

"We will never know how well Iraq's military might have done because the initial attacks took down its control network and, with it, any hope for Iraq to know what had happened, what was about to happen and what it might do about it". Col. Alan D. Campen, USAF (Ret.)

"Knowledge came to rival weapons and tactics in importance, giving credence to the notion that an ennemy might be brought to its knee principally through destruction and disruption of the means for command and control". Col. Alan D. Campen, USAF (Ret.)

6 Related Work

The purpose of this chapter is to draw a map of where "Viewpoints on Demand" stands in relation to the various research areas it overlaps with: personalizing and presenting information in electronic publishing, interactive storytelling, representing the content of video.

1. Customizing & Presenting Digital Information

W.Bender at al., 1991
Newspace: Mass Media and Personal Computing

The early nineties are marked by increasing interest in the digital delivery of information and entertainment, the so-called "information highway". By now it is almost impossible to find a magazine, newspaper or TV channel that has *not* approached the topic. Since the 70's the Architecture Machine group, which in 1985 became part of the MIT Media Laboratory, conducted research in the area of electronic publishing, digital television, interactive media and personalizing content. The initial concept that triggered the creation of the Media Lab was the convergence of television, publishing and computers. In terms of news and information, the central idea has always been that - as readers or viewers - in order to cope with the deluge of information, we need to get *personalized* information. For many years this made the Media Lab research very unique as no other lab was so invested in personalization.

shop (a Media Lab group that works on dynamic design). The user would specify whether s/he is a researcher or not, and whether s/he is technical or not. The system used a knowledge-based representation of presentation styles linked to these user types, along with a database of richly described media. The output was a presentation based on an ad hoc branching structure.

M.A. Fineblum, 1991
Adaptive Presentation Styles for Dynamic Hypermedia Scripts, MSVS MIT Media Lab

"Viewpoints on Demand" has in common with this system the idea that different users benefit from different presentation of the information. It has also in common the goal of customizing the presentation of information based on some initial user input parameter. A major difference is that "Viewpoints on Demand" is not based on a branching model (even though it can incorporate branching) [see chapter 7]. It is based on constraint based editing to generate a linear video documentary. In terms of content, another difference is that "Viewpoints on Demand" informs by using viewpoints and opinions. Michelle Fineblum's project is not meant to introduce people and their opinions on a topic. It is meant to present the results of research.

Graphical design

The research direction taken by "Viewpoints on Demand" is not only to present viewpoints based on initial input from the viewer but also to suggest opinions *related* to the sequence played on the screen at a given moment. In the context of graphical design, the Media Laboratory's Visual Language Workshop addresses similar issues. Its research exploiting its **Typographic Space** concept explores the graphical presentation of text information in a three dimensional space and deals in particular with *point of view in space*: how to avoid the distorsion of the typographic form caused by perspective and an arbitrary viewpoint, but also how to communicate visually shifts in viewpoint that indicate a change in focus to a new topic and its related information. Ultimately what is needed is an active sensor system and the dynamic display of related stories

D. Small, S.Ishizaki, M.Cooper,
Typographic Space, MIT Media Lab, Internal Report

E. Rennison, 1994
Galaxy of News, An Approach to understanding Expansive News Landscapes

Galaxy of News by Earl Rennison shows relationships within large quantities of information (news stories) by automatically building and visualizing links between related articles. It explores the semantic relation between graphical design *and content* in a 3D space. With semantic zooming and panning, the user is able to select a topic and explore it in more detail: as s/he zooms in, headlines to appear first and then the ar-

en story ought to be accessed and then told".

T. Oren, G. Salomom, K. Kreitmen, A. Don, 1990
Guides: Characterizing the Interface, in The Art of Human-Computer Interface Design, Addison Wesley

The **Guides project** developed at Apple shows in particular how narrative and information retrieval can be combined. The major difference is that Viewpoints on Demand presents views (opinions) in relation to each other instead of using viewpoints as guides to navigate through the material.

K.E.Steiner, T.G.Moher,
1992 Graphic Story Write: An Interactive Environment for Emergeent Storytelling, CHI'92

Graphic Story Writer uses graphics and text with no video but it does include knowlege about story structure which is precisely what it wants to teach to children in addition to helping them learn to read. With Graphic Story Writer they create interactively a story by dragging characters and objects in a set. Selecting initially a characteristic for a character like "greedy" helps define his/her goal. The system progressively generates the text of the story as more objects are added. It tries to guide periodically the user in order to work toward the resolution of the central conflict. It would insert for example: "the girl still looked for something fun to do" if her goal is indeed to have fun.

The elements of story are simple and help generate simple stories that children can manage. The structure of the story itself is not pre-determined as it is in "Viewpoints on Demand". It is created by interaction between the various elements added to the set at a given time. This direction is worth pursuing in fiction in order to give the user freedom of creativity. As I said in a previous chapter, with information the goal is different and so the influence of the user on the content must be of different nature.

Ryan Evans, 1994,
LogBoy Meets FilterGirl, MS MIT Media Lab

Mark Halliday, 1993
Digital Cinema: An Environment for Multi-threaded Stories. MS, MIT Media Lab

Train of Thought and **Just One Catch** are two examples of interactive movies in the fiction domain. Just One Catch plays back differently if the viewer asks for more dialog or more action. This could be applied in the information domain: a viewer might want to see more interviews about an event than descriptive reports of the event itself.

7 Future work: a Navigational Paradigm and a Storyteller Agent

Text connectors like BUT, AND, OR, BECAUSE...are meaningful navigational text connectors to explore opinions in video.

Techniques that enable dynamic reactions to changes in the environment pave the way toward a storyteller agent.

This chapter describes possible extensions. There are two major directions: navigating opinions and developing a storyteller agent. The last section positions this research in relation to research in automatic logging.

1. Navigating opinions

Navigational problems are a difficult but essential issue in interactive media. A major problem is to offer the user a navigational tool that will not prevent immersion in the content or encourage meaningless exploration.

What I foresee is the ability to interact with opinions while you are watching the story develop. This suggests viewer intervention at a clip level in addition to intervention at a global story level. Interaction with opinion requires *structure* but how much structure is there in the raw content?

An opinion, argument or debate expresses views about the world. Believability depends on how well the statements are supported, how much a given view takes into account other views, what emotions are conjured up and how much.

sional quotes. As with video, the journalists build relations between various opinions, events, and facts.

In academic discussions structuring arguments is particularly important. Arguments are presented with *text.* A text-based system called Euclid lets users state a position and defend it using a hypertext type of branching. Other users can then attack arguments by extending this net. This provides an extremely rigorous framework for debate.

This need for structuring appears also in email discussions. In fact, Euclid was created to make text debates more efficient. Text turns out to be a very convenient, flexible and malleable medium for this type of interaction. Even if asynchronous video conferences were widely available - allowing you to react when convenient to you - text is easier to edit, making it more flexible when you are shaping an argument that requires preparation.

Opinions: Text vs. Video

TV and film brings people to the screen. *"Camera allow us to get to know someone,"* comments Glorianna Davenport. Even if video is a representation of reality, and not reality itself, it can create a very strong sense of "being there" and bring to the viewers people with their sensitivity, their tone of voice and gestures. *Only with the moving image is it possible to show a debate to an audience.* Debates have taken place and are broadcast to an audience on television, not in newspapers. Newspapers on the other hand provide the context and the background information to fully understand the debate. They are more interactive: the reader is guided and gets oriented with front pages, headlines, tables of content, sections, page numbers.

"Literature often has the problem of making the significant somehow visible, while film often finds itself trying to make the visible

The major differences are that 1) TV is extremely limited in time and so has to reduce explanation and supporting facts to a minimum, 2) TV can show people, processes, places, events whereas newspapers describe them, 3) TV can report in real time. Newspapers need more time for distribution but this is changing with electronic delivery. They provide more in-depth analysis. Despite these three main differences, pros and cons are presented in a similar manner.

Example

What follows describes an example of interaction using the database built for this thesis.

High level meaning: [fig.1] shows the net of the relations between all the elements of answers currently in the database that support (or explain) the statement "The Gulf War Ended in the Blink of an Eye". A sharp contrast appears between Paul Strassmann's set of reasons and the other interviewees of the database.

Terminology: an object of the net can be a link, a connector or an "opinion item" (or argument). Basically, an opinion item is a video sequence that shows the smallest element that can be cut supporting an opinion.

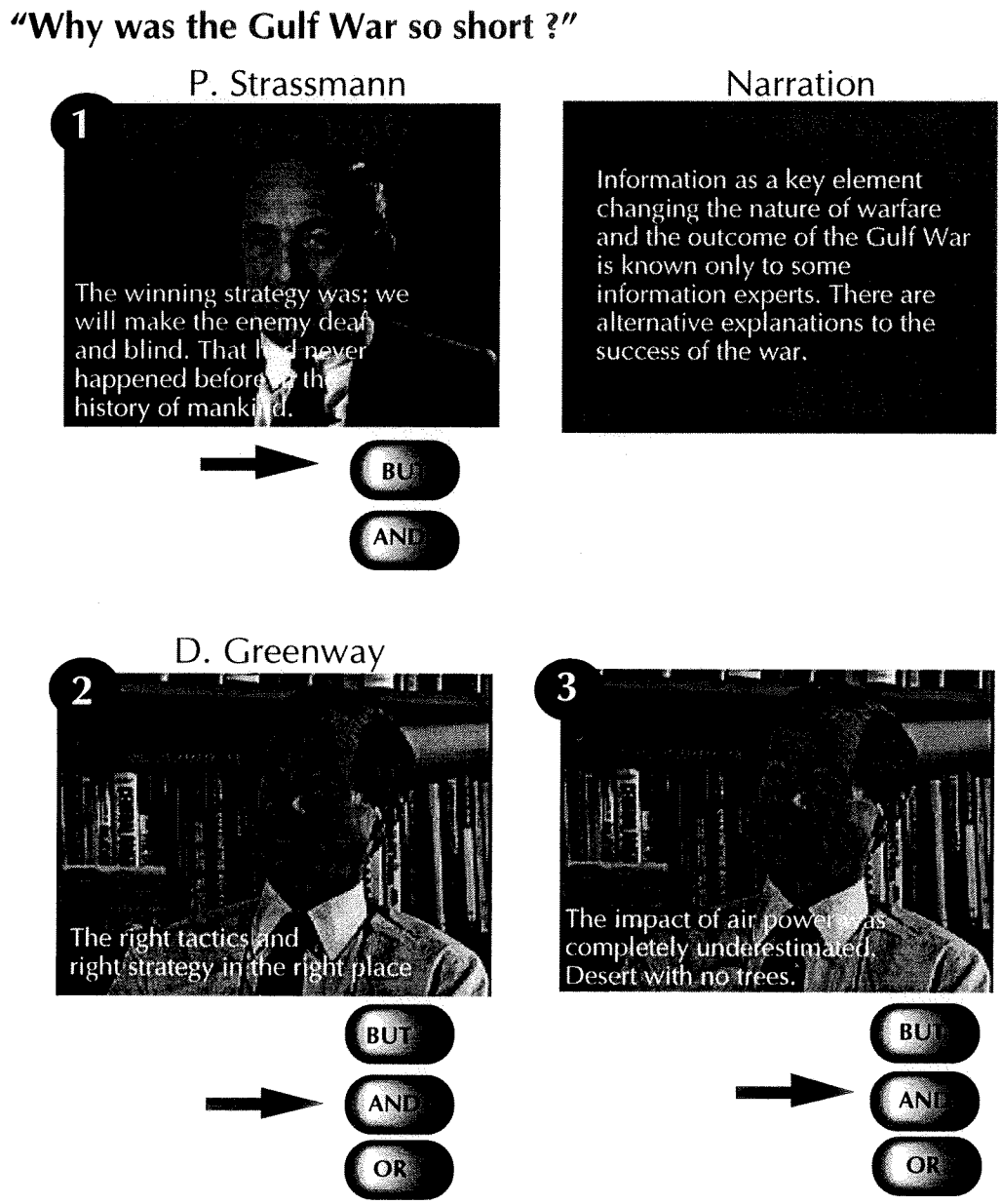


Figure 1 describes the internal representation of the relations between clips of explanation. Objects like the large gray OR box can be complex and contain recursively sets of connections. *Hierarchies* of logical connections need to be created. The BUT node shows an opposition between two sets of answers.

from Greenway, Wiggs or Robinson. In the first case, his options remain the same until the system runs out of clips for Strassmann. In the second case, s/he has a choice between getting more of Greenway (for instance) or switching to Wiggs or back to Strassmann or Greenway. The system keeps track of what has been shown and what remains relevant.

2. Towards a Storyteller Agent

Variations on a theme: "My storyteller knows me"

Under the concept "My storyteller knows me" [MIT Media Lab 5th Anniversary talk], Glorianna Davenport envisions a partnership between a viewer and a smart computer program. This is part of the more global picture in which intelligent agents who can learn about us will help us cope with the deluge of information we are more and more exposed to.

M. Minsky, 1986
The Society of Mind, Simon & Schuster, New York

As defined in the Society of Mind by Marvin Minsky, multi-agent systems use individual agents which each have a specific expertise. Solutions to problems emerge from the interaction among those agents. We can imagine agents expert in video editing and storytelling that would collect pieces of information in various places in order to assemble a customized documentary.

Why and how would an agent-like structure work

A very promising improvement would be to develop further the engine's ability to adapt by changing on the fly its plans about the selection and scheduling of the clips. Agent-based technology would be appropriate: adaptability to a changing environment is what makes a real difference between an ad hoc program understanding only a fixed data structure and an *agent* having some meta-knowledge about its environment and its task.

"Intelligence is determined by the dynamics of interaction with the world" Rodney Brooks[1]

The traditional A.I. approach has been to decompose into competence modules that are centralized and hierarchical. It does not work well for uncertain, unpredictable and changing environments. Under this model, a robot cannot find its way in a room if it is not lit in a certain way and

1. Rodney Brooks is professor at the MIT Artificial Intelligence laboratory. He has been advocating the use of more distributed knowledge representation models for robots to better adapt to their environment.

■ IF *expectation is created by a current sequence* THEN *check if new sequence meets this expectation (and act upon the result)*

Some rules can focus more specifically on enhancing a viewpoint:

■ IF *one main character is picked* THEN *maintain it as a strong character*

■ IF *character is afraid of X* THEN *make X look scary*

■ IF *character hates Y* THEN *show Y in a negative light*

■ IF *Y is main speaker* THEN *make Y talk last*

■ IF *Y is main speaker* THEN *insert a silence at end of clip* (to reinforce its impact)

Stylistic "tricks" that can reinforce a viewpoint:

- Audio from an "illustrative" clip starts as a person is still speaking, in a way that reinforces the meaning of what is said ("I hate street noise, I like to live in a farm" followed by a view of the farm with birds singing).

- Achieving the effect visually by using camera framing, camera position or lighting

Note 1: Rules involving cinematography instead of editing are much more difficult to achieve, because the technology for manipulating the content of a frame is still in its very early stages.

Note 2: There is a risk of bias if a speaker looks as though s/he is supporting another viewpoint when in fact s/he is not.

3. Automatic Logging of Video Databases

H.J. Zhang at al., 1994,
A Video Database System for Digital Libraries, Institute of Systems Science National University of Singapore

A question that often comes to mind with video databases is how to deal with large archives of footage. Currently news footage is archived on tapes and labeled using keywords. Keyword-based search applied to digital media is only a very limited retrieval mechanism. It also still implies manual labelling of the content which is a tedious process as video takes time to watch.

There is on-going research aiming at recognizing automatically moving and static objects in a sequence of images. Other projects are also attempting to do speaker recognition and natural language understanding.

8 Conclusion

Thinking about the future too often distracts us from the past as if everything was novelty. Personalization is hardly a new idea. In some cases, it is the presence of mass production that tends to make us overlook the human need for personalization.

Even before the industrial era, one can find examples of suppressed individuality. During his reign, Louis XIV consolidated French power and caused French craft to dominate Europe. However, in the second half of the 18th century, it was Paris, not Versaille that became the center of culture and fashion. The lesser nobility and the bourgeoisie were taking over in Parisian salons. As a result, urban taste developed, in opposition to official court style, a more intimate, *personal* style. Art was funded by private connoiseurs following their own taste instead of the academic style decided by the court (which had brought Molière and the Comédie Française). Individuality had become the rule.

A couple of centuries before that, books were usually sold unbound and most people liked to have them bound to their own personalized requirements.

One might see the automatic generation of personal news and information as the demonstration of automation pushed to an extreme, casting a

Il suffisait que Mme Swann n'arrivât pas toute pareille au même moment pour que l'Avenue fût autre. Les lieux que nous avons connus n'appartiennent pas qu'au monde de l'espace où nous les situons pour plus de facilité. Ils n'étaient qu'une mince tranche au milieu d'impressions contigües qui formaient notre vie d'alors; le souvenir d'une certaine image n'est que le regret d'un instant et les maisons, les routes, les avenues, sont fugitives, hélas, comme les années.

Marcel Proust, "A la Recherche du Temps Perdu", "Du Côté de chez Swann".

C. D. Horner 1994 NewsTime: A Graphical Interface to Audio News, MS MIT Media Lab

Brenda Laurel 1989, A Taxonomy of Interactive Movies, New Media News, Vol.3 issue 1

B. Laurel, 1990, Interface Agents: Metaphors with character, in The Art of Human-Computer Interface Design, B.Laurel (Ed.), Addison Wesley

M. Murtaugh, 1994, Con Artist: A System for Graphically Representing Story Knowledge, BS MIT

H.W. Lie, 1990 The Electronic Broadsheet, All hre news that fit the display M.S. MIT Media Lab

J.Orwant, 1991, Doppelganger Goes to School: Machine Learning for User Modelling, M.S. in Media Arts and Sciences

Michael Rabiger, 1987, "Directing the Documentary", Focal Press

E. Rennison, 1994 Galaxy of News, An Approach to understanding Expansive News Landscapes, submitted to UIST'1994

B.Rubin, 1989 Constraint-Based Cinematic Editing. MSVS. MIT Media Lab

Aguierre Smith, T. G., & Davenport, G. (1992). The Stratification System: A Design Environment for Random Access Video. Workshop on Networking and Operating System Support for Digital Audio and Video. San Diego, CA: ACM.

W.Sack, 1994, Indexing Multimedia by Ideology, AAAI-94 Workshop on Indexing and Reuse in Multimedia Systems, Seattle Washington

W.Sack, 1994, Future News: Constructing the Audience Constructing the News, Proceedings of WRITE'94

W.Sack, 1993 Coding News and Popular Culture, IJCAI'93 Proceedings, Chambéry France

R.Schank et al., 1992 Agents in the Story Archive, TR 27, Northwestern University

D. Small, S.Ishizaki, M.Cooper, Typographic Space, MIT Media Lab, Internal Report

K.E.Steiner, T.G.Moher, 1992 Graphic Story Writer: An Interactive Environment for Emergeent Storytelling, CHI'92

Paul Saffo 1992 Consumer Technology Purchase Behavior in the 1990's - IFTF

B. Sheth & P.Maes, 1993, Evolving Agents for Personalized Information Filtering, Conference on Artificial Intelligence Applications, Orlando Florida

A.&H. Toffler, 1993 War and Anti War, Little Brown

Sun-Tzu, the Art of Warfare, c. 403-221 B.C. translation and comments by Roger Ames, Ballantine Books, 1993

Upside, April 1994 An interview with Eric Nee of Scott Cook, CEO of Intuit

sensor data: desired length of presentation

• add-clips

goal: make the selection longer

conditions: selection is available, selection-too-short, footage available

expected positive effect: longer selection, long enough

sensor data: footage available

• remove-clips

goal: make the selection shorter

conditions: selection is available, selection-too-long, footage is available

expected positive effect: shorter selection, short enough

sensor data: footage available

• enough on topic X?

goal: check if topic X is covered enough

conditions: selection is available

expected effect: True/False

sensor data: footage available

• enough on topic Y?

goal: check if topic Y is covered enough

conditions: selection is available

expected effect: True/False

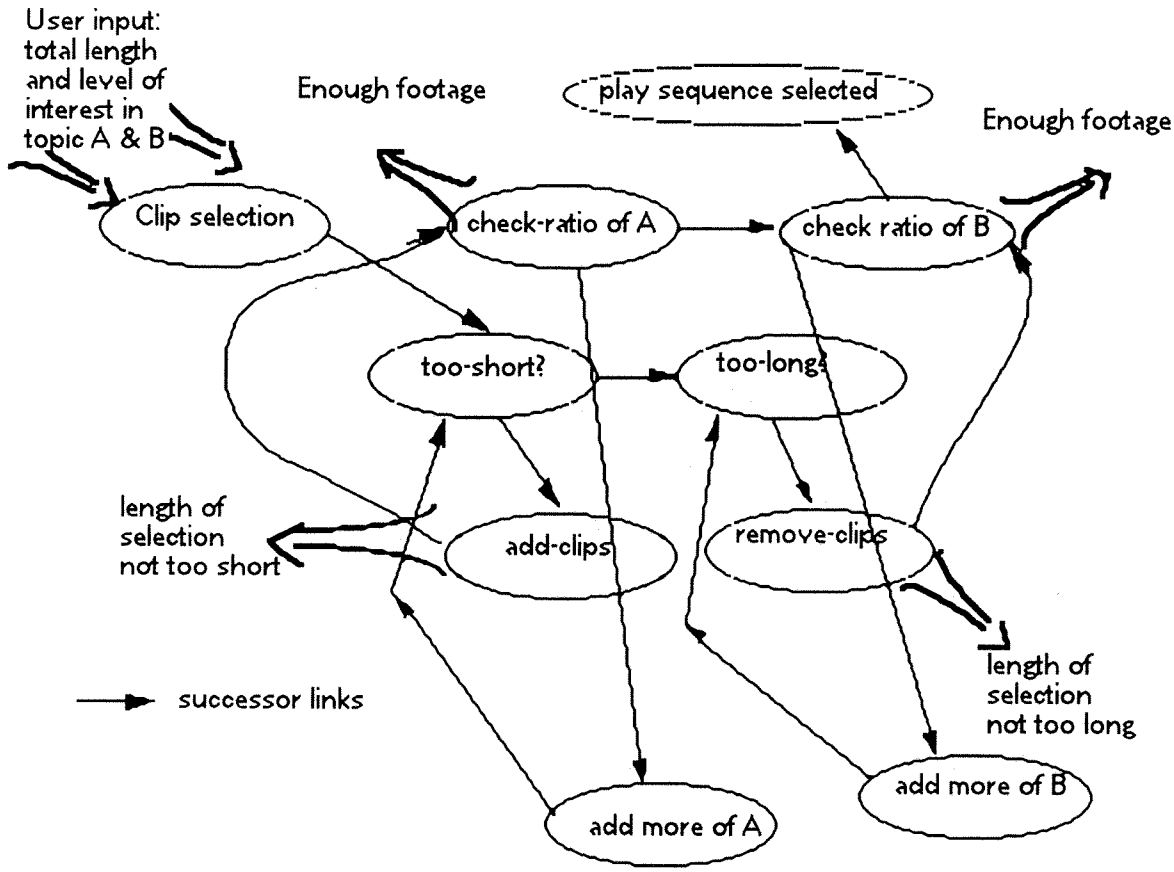
sensor data: footage available

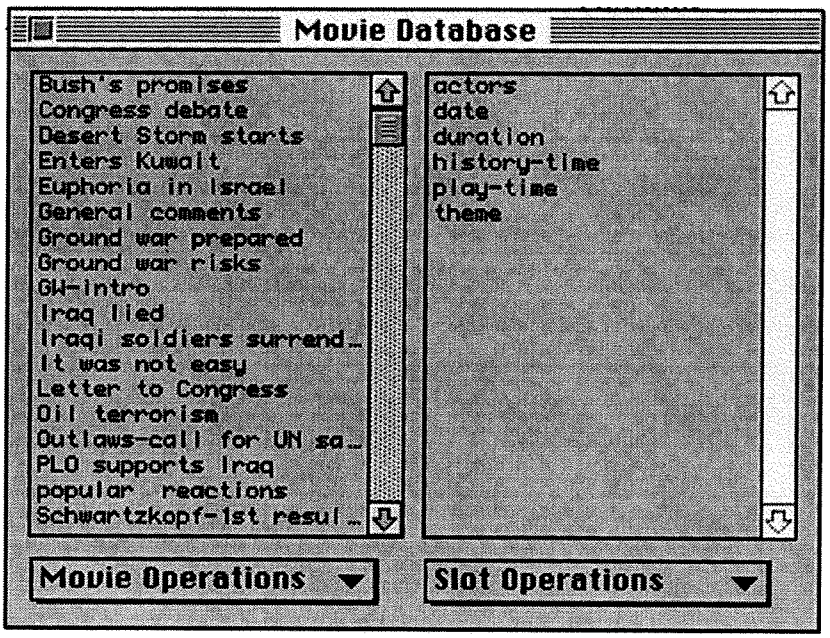
• add clips on topic X

goal: add clips on topic X

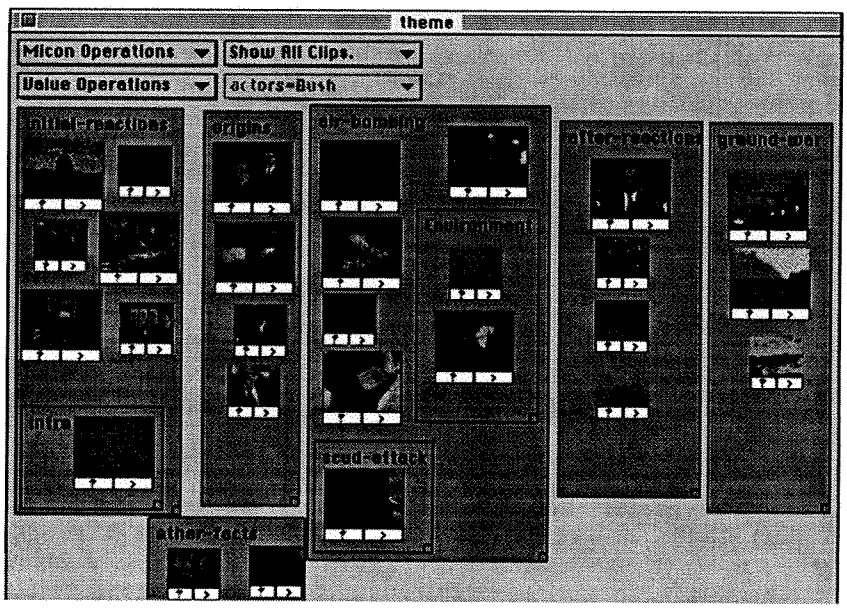
conditions: selection is available, footage on X is available

expected effect: more footage on X


User input:
total length
and level of
interest in
topic A & B
Enough footage
play sequence selected
Enough footage
Clip selection
check-ratio of A
check ratio of B
too-short?
too-long?
length of
selection
not too short
add-clips
remove-clips
length of
selection
not too long
successor links
add more of B
add more of A



The main database window showed on the left all the clips available and on the right the possible keywords.



Clips were annotated using Logboy* by associating to them a keyword (her clip dropped onto a rectangle is associated with a given value for the ke sible to represent relationships between clips which, among other reason: ation of a new tool.

(*) "LogBoy meets FilterGirl", Ryan Evans, MS 1993 MIT Media Lab