

Digital Cinema

An Environment for Multi-threaded Stories

by

Mark David Halliday

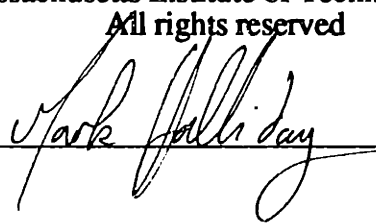
B. A. Computer Science
Dartmouth College
Hanover, NH
1985

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

MASTER OF SCIENCE IN MEDIA ARTS & SCIENCES
at the
Massachusetts Institute of Technology
September 1993

© Massachusetts Institute of Technology, 1993
All rights reserved

Signature of Author



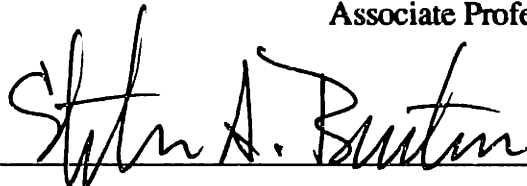
Media Arts & Sciences
August 7, 1993

Certified by



Glorianna Davenport
Associate Professor of Media Technology
Thesis Supervisor

Accepted by



Stephen A. Benton
Chairperson

Departmental Committee on Graduate Students

ARCHIVES

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

OCT 18 1993

Digital Cinema - An Environment for Multi-threaded Stories

by

Mark David Halliday

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning
on August 7, 1993
in partial fulfillment of the requirement of the degree of

MASTER OF SCIENCE IN MEDIA ARTS & SCIENCES
at the
Massachusetts Institute of Technology

Abstract

This thesis explores how simple interactive narratives can be constructed from a large collection of digital video clips described in a database. Constructing such narratives means encoding story structures that can allow the computer to choose appropriate clips from the database as the story progresses over time, referred to as orchestration, and establishing rules controlling how and where images appear on-screen, referred to as positional editing. The aim of this thesis is to use orchestration and positional editing to create an environment within which it is possible to view, interact with and understand multi-threaded narratives.

Thesis Supervisor: Glorianna Davenport
Title: Associate Professor of Media Technology

This work was supported in part by Kansa and Bellcore.

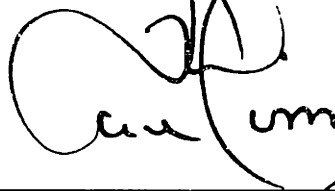
Digital Cinema

An Environment for Multi-threaded Stories

by
Mark David Halliday

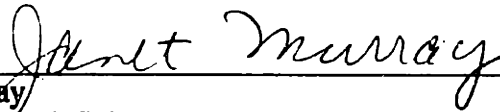
The following people served as readers for this thesis:

Reader:

A handwritten signature in black ink, appearing to read 'Errol Morris', is written over a horizontal line. A diagonal line is drawn across the signature from the bottom right towards the middle of the line.

Errol Morris
Film Director

Reader:

A handwritten signature in black ink, appearing to read 'Janet Murray', is written over a horizontal line.

Janet Murray
Senior Research Scientist
Director, Laboratory for Advanced Technology in the Humanities

Acknowledgments:

The work described in this document would have been impossible without the encouragement, support and patience of many people. Thanks to all of them.

Glorianna Davenport for believing in me from beginning to end and giving me the freedom to experiment with images and stories over the last two years.

Janet Murray for your invaluable commentary in shaping this document and your enthusiastic support over the past two years.

Errol Morris, watching your latest edits on the Avid reminded me what I'm really trying to achieve with moving images.

Ryan Evans you've been a great friend and wonderfully patient when it comes to showing me what computers can really do. I hope we can work together again someday.

Betsy Brown and David Tames are real friends for suffering through early drafts of the document and scribbling lots of comments. James Seo, thanks for the late nights that yielded the latest and greatest interface code. Thomas Aguiere Smith, Kevin Brooks, Amy Bruckman, Stuart Cody, Eddie Elliott, Stephan Fitch, Scott Higgins, Gilberte Houbart, Hiroaki Komatsu, David Kung, Lee Morgenroth, Brian Robertson, Kiochi Yamagata, Hoony Youn have been great friends to work with in Macondo, our beloved lab.

The lab would have been very different without the musical tastes of David Kung, John Shipple, and James Seo. Thanks for all the taping opportunities.

My roommates Paul and Henrietta took me hiking and cooked great meals and made my 2 years at the Lab a wonderful living experience.

Finally, Alessandra Prosperi thanks for supporting me emotionally and otherwise over the past two years.

How to read this document

This document is designed in a modular fashion. There are a number of different ways you can read it:

- | | |
|------------------------------|---|
| The 1 Minute Reading | Read the abstract on page 2 and look at all the pictures. |
| The 5 Minute Reading | Read the abstract on page 2, read the summaries at the beginning of each chapter and look at the pictures. If you are a fast reader, you might also have time to read the conclusion on page 83. |
| The 30 Minute Reading | Read the abstract, and Chapter 1. This will give you a detailed outline of the problems addressed by this thesis. Follow this by a reading of the summaries at the beginning of each chapter. You might then go on to Chapter 6 which presents the results of the research and finish with Chapter 7, the conclusion. |
| The 2 Hour + Reading | Read the whole document from beginning to end. There is a summary at the beginning of each chapter that gives an overview of what each chapter is about. With 2 hours, you'll have enough time to get into a detailed account of all the experiments that were carried out as part of the research. |
| The Fun Reading | Appendix A gives an account of the production process and includes a bio of all the people who worked on the various shoots. If you have time, read Chapters 1, 6 and 7 for an overview of the research. |
| The Mystical Reading | Throw document in the fireplace and contemplate the ashes. |

Table of Contents

1. Introduction	9
1.1. The Digital Age.....	10
1.2. The Aim	11
1.3. The Road.....	12
1.3.1 Process.....	12
1.3.2 The Stages.....	12
2. The First Experiment - IC Portrait.....	15
2.1. Concept.....	15
2.1.1 Unstructured positional editing and orchestration	15
2.2. Background and a little theory.....	17
2.2.1 A Definition of a Story.....	17
2.2.2 The Maker's Activity.....	18
2.2.3 The Viewer's Activity.....	18
2.2.4 The Frame	19
2.2.5 The Viewer Becomes Maker	20
2.3. Interaction, Interactive Cinema and Previous Work	21
2.3.1 Aspen (1979-1981).....	22
2.3.2 New Orleans (1982-1987).....	22
2.3.3 Elastic Charles (1988-1989).....	22
2.3.4 Rows Wharf Intrigue (1988-1989).....	23
2.3.5 The Electronic Scrapbook (1990-1991).....	23
2.3.6 Homer (1991-1992).....	23
2.3.7 Windows on the World (1981)	23
2.3.8 The Video Streamer (1991-1993).....	24
2.3.9 LogBoy Meets FilterGirl (1992-1993).....	24
2.4. Where does my research fit in?.....	24
3. Digital Micromovie Orchestration - An Endless Conversation	26
3.1. Concept.....	26
3.1.1 A Filter-Based Turn Taking Conversation.....	26
3.2. Production	27
3.2.1 Software.....	27
3.2.2 The Logging Process.....	27
3.2.3 Filtering.....	29
3.2.4 Scripting.....	30
3.2.5 Additive Construction.....	31
3.2.6 Interface.....	32
3.2.7 Extensibility	34
3.3. Criticism.....	34
3.3.1 Results.....	34
3.3.2 Problems.....	35
4. Orchestration through Linked Story Structures	36
4.1. Concept.....	36
4.1.1 An impressionistic collage through a linked story structure	36
4.1.2 A Multi-threaded Structure	36
4.1.3 Granularity	36
4.1.4 An Inspiration for Content and Structure	37
4.1.5 Sequencing Story Units.....	37
4.1.6 Documentary Content	38
4.1.7 Fictional Content.....	38
4.2. Production	39

4.2.1 An Impressionistic Collage.....	39
4.2.2 The Story.....	40
4.2.3 Casting & Shooting	41
4.2.4 Software	42
4.2.41 Links.....	42
4.2.5 The Interface: a visual context for the viewer	45
4.3. Criticism.....	47
4.3.1 The Experience.....	47
4.3.2 What works	48
4.3.3 Feedback from selected users	48
4.3.31 Perspective.....	48
4.3.32 Two choices	48
4.3.33 Content and Structure.....	49
4.3.34 Problematic links.....	49
4.3.4 Lessons Learned.....	49
5. Story Models and Story Agents - Train of Thought v 2.0.....	51
5.1. Concept.....	51
5.1.1 A linked story model and a filter-based story agent	51
5.1.2 Thoughts on Structure	52
5.1.3 Story Template.....	52
5.1.4 Quantifiable Structure Exists in Films.....	53
5.2. Production	54
5.2.1 Software.....	54
5.2.2 Data Structures.....	55
5.2.3 Sequencing the Story.....	56
5.2.4 The Story Agent	57
5.2.5 Logging.....	58
5.2.6 Inference.....	58
5.2.7 The script	60
5.2.8 The Interface	62
5.2.81 Two Modes	62
5.2.82 Annotation Browsing.....	64
5.2.83 Inferred Material.....	65
5.3. Criticism.....	65
5.3.1 The Experience.....	65
6. Results - Insights on Content and Structure.....	67
6.1. A Brief Review.....	67
6.2. Story Models	68
6.2.1 Kernels and Satellites	68
6.3. Interface - Story Visualization through Positional Editing	71
6.3.1 Non-interactive.....	71
6.3.2 Interactive.....	71
6.3.3 Design.....	72
6.4. The Interactive Experience	73
6.4.1 Exploration in Fictional Narratives	73
6.4.2 Exploration in Documentaries.....	74
6.4.3 Thoughts on Dreams, Emotions and Interaction.....	75
7. The Future.....	77
7.1. New kinds of distribution, new kinds of stories	77
7.1.1 Broadcasting	77
7.1.2 Networking.....	77
7.1.3 Personalizing.....	78
7.1.4 Open Ended	79
7.1.5 Free Form Content in a Controlled Structure.....	79

7.2. Research directions	80
7.2.1 Scalable Story Structures	80
7.2.2 Interfaces	81
7.2.3 Interaction	82
8 Conclusion	83
References	85
Appendix A - Digital Guerrilla Movie making.....	87
On Guerrilla Filmmaking	88
Writing	88
Actors	89
Shooting	89
Sound	90
Post-production	91
In Closing.....	92

1. Introduction

This chapter introduces the main problem addressed by this thesis: how can digital technology enhance cinematic story-telling? We look at how computation can assist the movie maker in two ways: *positional editing*, which refers to how the positioning of story-elements within the frame can provide a way to visualize a multi-threaded story, and *orchestration* which refers to how the computer can control the sequencing of cinematic material. The chapter concludes with a detailed outline of the four experiments carried out as part of the research that will be described in the rest of the paper.

On two deaths and three births

My movie is born first in my head, dies on paper; is
resuscitated by the living persons and real objects I
use, which are killed on film but, placed in a certain
order and projected on to a screen, come to life again
like flowers in water.

Robert Bresson

The ultimate goal of any filmmaker is to tell a story to an audience. In traditional film, the creator weaves images and sounds into a complex but fixed tapestry that grabs the viewers and draws them into an engaging display of sounds and images. The experience can be very powerful and absorbing.

The process of making a movie is a collaborative effort that involves the talents of many people and the use of complex equipment and tools. In a fictional film, a screenwriter lays down the plot, dialog and action, creating a detailed template of the entire story. During production, the director works with the actors, cinematographer, art director, sound person and production manager to capture on film or video the material that will shape the story. They use cameras, sound recorders, lights, dollies and other sophisticated equipment to record the material. During post-production, the director works with the editor, sound editor, composer and other crafts people to shape the images and sound into a cinematic narrative using editing machines, mixers and special effects equipment to bring the movie to life. The production of a movie is an iterative process that involves the constant give and take of many people as they slowly bring the story to life. In some cases, the makers will realize that the movie is missing certain crucial elements and they will rewrite and re-shoot entire scenes to make the story flow. The final product is a movie.

Linear movies, whether created and released on film or video, are made from a collection of shots and sounds arranged along one unchanging strip. The structure, style and content will remain the same every time it is seen by an audience, although the interpretation of the events happening on screen is different for each viewer. It is an experience that draws the viewer into a state of reverie, completely absorbing them into the events happening on-screen.

1.1. The Digital Age

Digital technology gives the movie maker new capabilities and tools. It is possible to start thinking of different kinds of narrative environments that give the viewer new ways to experience stories.

The frame can be broken down to encompass multiple streams of video. The decision of where to position those streams on the screen, an activity we call *positional editing*, becomes as important as temporal editing, deciding how to sequence images out over time.

Digital video stored on a hard disk is just like any other piece of data that can be accessed immediately by the computer. This gives the movie maker options for deciding how to *orchestrate* the images and sounds on the screen. Computation can be used to help this process.

Viewing has traditionally been a passive activity. The viewer has no control over the playout of a movie. The mind is active as the eyes and brain process the incoming audio-visual signal, but there is no activity in reshaping the playout of the story. In the digital domain, the viewer can affect the viewing experience through interaction, thereby becoming an active participant. This is a novel feature that requires new ways of writing, filming and structuring cinematic stories.

This thesis explores how movie making and movie viewing change when the activities take place in a digital environment. Digital technology makes it easier to do things that were difficult or impossible to do before. I want to explore how the instant access of video frames and the ease of positioning them on-screen changes the making and the viewing of movies. The research is motivated by my desire to create cinematic narratives that involve the active participation of the viewer.

1.2. The Aim

The research presented in this thesis looks into two ways in which digital technology can change the way we can arrange audio-visual information over time to tell stories. Two organizing principles are investigated:

- **Positional Editing** looks at how to break out of the frame-based presentation of movies. The computer can display multiple streams of video on the same screen, offering the maker new ways to position story elements so as to structure a story in a visual manner that can help the viewer's story building activity.
- **Orchestration** explores the ways in which computation can be used to assist the story construction process. As computer technologies convert video into digital data that can be easily accessed through computation, what used to be a linear strip of frames can be broken up and recombined to form new sequences. Computational orchestration offers the maker flexible ways to structure the content and the playout of a story.

Technology has always been a driving force behind the art of cinema. “Each new medium modifies and extends the linguistic possibilities of the moving image, subsuming the syntaxes of previous media without negating them” (Youngblood, 1970). The digital medium gives the movie-maker new options for manipulating the moving image to communicate to an audience. Things that were possible but difficult to accomplish become easier; images can be accessed within a split second and multiple images can be displayed on the same screen. The problem is to find out how to effectively use these capabilities to tell engaging stories. Movie makers have always used the latest technological advances to push their art forward. Color, sound and special effects have all extended the art of cinematic story telling. Computation pushes movie making in new directions.

In 1980, *Aspen* investigated the use of computation to present a collection of video frames in a spatial context. This videodisk-based movie allows the viewer to travel around the city of Aspen by driving up and down the streets (Lippman, 1980). *Aspen* literally puts the viewer into the driver's seat. It is a pioneering experiment that took advantage of instantaneous access to any frame of video and of simple computer graphics, to create an active cinematic experience. *Aspen* was the first of many experiments at the Media Lab, including *Windows on the World*, *New*

Orleans, *Elastic Charles*, and many others ¹ that all explore the problem of using computation to create new kinds of viewing experiences.

The research in this thesis grows out of these experiments. I use computation to assist in the positional editing and the orchestration of cinematic material to drive the viewing experience. The viewer should be able to sit back and enjoy the show or to change what is happening on-screen in a fluid manner. My hope is that the computer can be used to put viewers into the driver's seat, but not to require them to drive all of the time.

1.3. The Road

1.3.1 Process

I am a filmmaker by training. The process I follow during my research is very much like the iterative process of producing a linear movie. In collaboration with other researchers, I make movies using digital technology. I start with a *concept* of what I want to see on the screen. I then move on to the *production* stage where the concept is brought to life. This stage often involves changes and rethinking the concept; just as in a linear movie, certain things are found to work and others are found to fail. The final outcome is invariably an experience in which an audience can participate. During this stage, *criticism* feeds back information to the maker on what works and what doesn't. This criticism often motivates the maker to go back and change the concept, initiating new production and criticism. It is through this iterative process of creative feedback that movies are made and it is through this iterative process that I research the production of digital cinematic experiences.

1.3.2 The Stages

This thesis tracks the production of four experimental interactive movies, outlining each production as it goes through the three stages mentioned above: *concept*, *production* and *criticism*. The four movies each explore varying degrees of structure in the way they are displayed and played out using computation. They range from little structure in either positional editing or orchestration, as in *IC Portrait*, to more structure in both orchestration and positional editing as in *Train of Thought v 2.0*.

¹These and other experiments will be discussed further in Chapter 2.

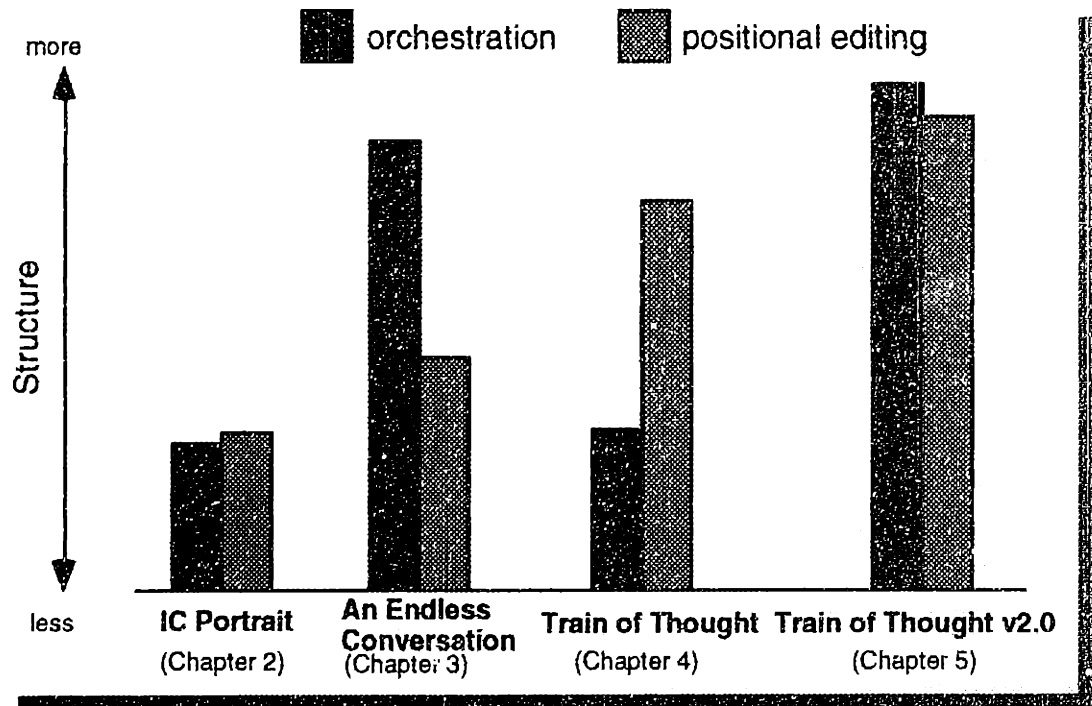


Figure 1 - A subjective representation of the amount of structure present in the positional editing and orchestration of each of the four experiments.

Chapter 2 introduces the two main problems tackled in this thesis, orchestration and positional editing, by analyzing the first experiment, *IC Portrait*, a multi-threaded movie about students in the Interactive Cinema Group created by my colleague Ryan Evans and myself. What new factors have to be taken into account when creating for the digital environment and for interaction? How have other experiments tackled the problem?

Chapter 3 looks at how the second experiment, *An Endless Conversation*, addresses the problem of creating a cinematic sequence from a collection of digital video clips. How can we achieve real-time shot selection to build movies?

Chapter 4 analyses *Train of Thought*, a fictional multi-threaded narrative, filmed with the help of Ryan Evans and Lee Morgenroth, that tells the story of the separation of Jack and Nicole as she leaves him for a job on the other side of the world. It is a multi-threaded story where links are used to build the narrative structure. A 2 1/2 D interface attempts to give the viewers a sense of where they are and where they are going in the story. Links did not serve our purposes, but the interface is an interesting structured approach.

Chapter 5 tracks the process of creating the second iteration of *Train of Thought*, *Train of Thought v 2.0*. A formal story structure is developed and the interface becomes more dynamic.

Chapter 6 discusses what was discovered from the whole process of trying to tell alternative stories in the digital domain. What does the viewer want? What does the maker have to take into account? What does interaction mean?

Chapter 7 talks a little about the future and frames some of these ideas in the context of what is happening with the distribution technology. What kinds of new viewing experiences might emerge from a networked digital cinema environment?

2. The First Experiment - IC Portrait

This chapter looks at the first experiment, *IC Portrait*, carried out by Ryan Evans and myself. *IC Portrait* used unordered positional editing and orchestration to build a cinematic sequence on-screen. The lack of structure in *IC Portrait* created a difficult viewing experience but identified the problems to be dealt with in subsequent research. This chapter goes on to define what is meant by *story*, and how the maker's and viewer's roles affect the structure and content of interactive movies. It also places this research in the context of other work that has been carried out in the area of interactive cinema.

2.1. Concept

2.1.1 Unstructured positional editing and orchestration

IC Portrait was a first attempt at creating a multi-threaded experience out of a large collection of digital video clips. The movie is about Interactive Cinema; what is it? who is it? what does it mean? We filmed all the members of the Interactive Cinema Group talking about themselves, interactive cinema and various objects that were part of the lab environment. They were filmed against a black background and were encouraged to make funny and wild statements. The video was then converted to the digital format and edited by hand to create coherent statements from each character.

This database of clips was then incorporated into a multi-threaded presentation. The system was based on a message-passing structure where the encoded representation of each character could send and receive messages to other encoded characters within the movie world. There was also a "master story teller" who coordinated the overall playout by broadcasting messages into the environment at specific time intervals. It was a rather primitive system where we were attempting to create a flexible environment within which a range of events could unfold. We wanted to create a multi-threaded story that the viewer could mediate by clicking on a particular character or object.

IC Portrait addressed a couple of issues. It tried to create a multi-threaded structure where events are not pre-scripted to replay in one particular order. Instead, certain actions such as an utterance from a character or intervention from the viewer were supposed to trigger other events, such as an utterance from another character or the

appearance on-screen of an object. This environment was loosely modeled around a micro-world² model where a collection of objects influence each other's activities in a closed environment. It was not a strict interpretation of a micro-world however, because the master story-teller guided the conversation in a predetermined way.

Another issue addressed by *IC Portrait* was the display environment. All characters and objects were filmed against a black background. This was done so that all the visual material could then be smoothly incorporated into a black background to allow us great flexibility in positioning elements within the screen. We were thinking of the environment as a non-linear collage of shots in space. This model takes the idea of a movie a step back in the direction of theater: we are experiencing a narrative where the characters are interacting with each other and with the viewer within the confines of a defined space, the screen.

IC Portrait was a first attempt at breaking away from a pre-scripted linear display of audio-visual material. What were the results?

Watching and interacting with *IC Portrait* was a rather confusing experience. Characters and objects appeared in various parts of the screen in a seemingly random manner. There was no narrative flow. As I said at the time:

Some basic elements of a narrative are missing from *IC Portrait*. There is no conflict between the characters. The interaction between them is quite limited. One character might say one thing and cause another to come up and respond in a different part of the screen. Unfortunately, the reaction is not always obvious. The result on-screen is of a somewhat random collage of characters saying various things about themselves and about interactive cinema. It does not quite add up to a story. (Halliday, 1991)

IC Portrait got things going on the right track, and identified the problems to be dealt with in subsequent research.

Keeping *IC Portrait* in mind, it is useful to analyze the problems we were trying to address. There are two basic areas of interest:

- The position of movie elements on-screen, which we refer to as positional editing

²The term micro-world was originally coined by Seymour Papert in his book "Mindstorms". The use of the term here refers to a system where objects influencing each other within a closed environment.

- The sequencing of movie elements in time, referred to as orchestration.

The basic question being addressed is this: How do I get the computer to assist in the story building activity by positioning and orchestrating digital video such that a random-access, on-the-fly playout can make narrative sense?

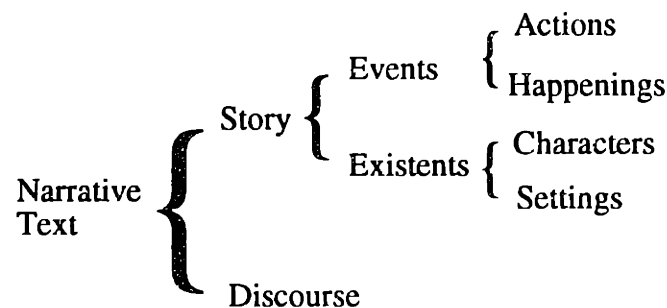
2.2. Background and a little theory

IC Portrait tried to maintain the fluidity of the linear medium while still allowing the viewer to intervene in what was happening on-screen. It was possible to sit back and just watch what was happening because the experience played out without prompting from the viewer. The computer was orchestrating the experience, but not in a narrative way. The data representing the video did not help the computer select segments to build a coherent narrative. Similarly, the video clips were being positioned in different parts of the screen, but not in a cinematic way. A brief look at some of the theory behind stories, makers and viewers will give us a good perspective on what I am trying to accomplish.

2.2.1 A Definition of a Story

The first important point to clarify is to define what constitutes a story.

What are the necessary components - and only those - of a narrative? Structuralist theory argues that each narrative has two parts: a story (*histoire*), the content or chain of events (actions, happenings) plus what may be called the existents (characters, items of setting); and a discourse (*discours*), that is, the expression, the means by which the content is communicated. In simple terms, the story is the *what* in the narrative that is depicted, discourse, the *how*. The following diagram suggests itself:



(Chatman, 1978, p. 19)

I take Chatman's definition of a narrative as the basis upon which to build stories using computation. The basic elements of a story (actions, happenings, characters

and settings) are organized into a narrative. The organization of these elements (the discourse) is at the heart of this research. I want to use computation to create this organization.

“A narrative is a communication; hence it presupposes two parties, a sender and a receiver. [...] Whether the narrative is experienced through a performance or through a text, the members of the audience must respond with an interpretation: they cannot avoid participating in the transaction” (Chatman, 1978, p. 28). This relationship of the sender to the receiver is central to my approach. It is the synthesis of their activities that creates the flowing cinematic experience.

2.2.2 The Maker’s Activity

The movie maker will carefully build a sequence of images and sounds to draw the viewer into a state of reverie and in so doing, is able to communicate something to the audience. Using cinematic conventions that have evolved over time, the maker constructs a story. A script clearly delineates the actions and scenes required for the story that needs to be told. The cinematic material is gathered using a camera and a sound recording device. This material is then edited together using specific rules that allow the maker to articulate a story. The finished product is a collection of shots that have characters carrying out actions in various settings. These shots play out and build a narrative experience for the viewer. The essence of the maker's activity involves building up this experience and then passing it on for viewing. What does the viewer do?

2.2.3 The Viewer’s Activity

“In comprehending a narrative film, the spectator seeks to grasp the filmic continuum as a set of events occurring in defined settings and unified by principles of temporality and causation” (Bordwell, 1985, p. 34). The viewer “grasps the filmic continuum” by taking the incoming visual and auditory information and processing it to build up a story. “For schooled perceivers in contemporary Western culture, narrative comprehension and recall are centrally guided by the goal of creating a meaningful story out of the material presented.” (Bordwell, 1985, p. 34). This goal of building up a meaningful story is a continuous and subconscious activity, something the audience does as soon as they see a sequence of shots play out in front of them, whether on a movie screen or on a television tube. The viewer's story building activity is part of the viewing experience. This is where a

skilled filmmaker will provide the necessary clues at the right point in time such that the viewer is able to construct a narrative out of the separate events appearing on-screen.

If we are to create engaging experiences in the digital medium, we have to start on familiar ground. I would like to take some of the features of the linear medium that we understand and see if we can extend them to the digital medium. We understand how stories are built up in movies; images and sounds are carefully selected and sequenced to build an engaging experience. Although *IC Portrait* did attempt to create a flowing experience on-screen, there was no narrative in the flow. This was one of the fundamental elements lacking from the experience. The viewer was not able to create any kind of meaningful story out of the material presented on the computer screen. The experience resembled a collage of unrelated utterances. One of the aims of this thesis is to demonstrate how the computer may orchestrate a meaningful sequence that builds up a narrative from a collection of clips that have been assembled by the movie maker.

2.2.4 The Frame

The digital video sequences presented to the viewer in *IC Portrait* appeared in different parts of the screen. One character might say: "Hi my name is Thomas" as he pops up into the middle of the frame. The image of a light might then appear on one side of the frame. While these digital video clips played out in various parts of the frame, they did not provide any visual cues to the viewer as to what was happening in the narrative. If anything, there was a general feeling of surprise anytime a new clip appeared, because it popped up in an unexpected place. This was a little confusing. The experience would probably have been more coherent if the video had been displayed within a frame as linear video separated by cuts. In a movie, every transition is motivated. A cut, a dissolve, or a wipe each signify something to the viewer (as long as they have become accustomed to the basics of cinematic language). In a digital environment, the positioning of multiple video streams can be used to add meaning to the playout of the story. Instead of just positioning video in different areas of the screen, the way in which a particular piece of video is displayed can help the viewer's activity. In particular, I am interested in using perspective to show what is happening over time. Since storytelling is an inherently time-based activity, representing time in a graphical manner might enhance the viewing activity. In an interactive environment visual cues that

show where one is in the story and where one can go next will help the viewer make decisions. It will help the story construction process.

2.2.5 The Viewer Becomes Maker

The viewer's activity and the maker's activity as described above are quite distinct from each other. Once a film is made, it is handed over to the marketing and distribution people rarely to be touched again by the filmmaker. What happens when the viewer suddenly is given a chance to intervene in the on-screen activities?

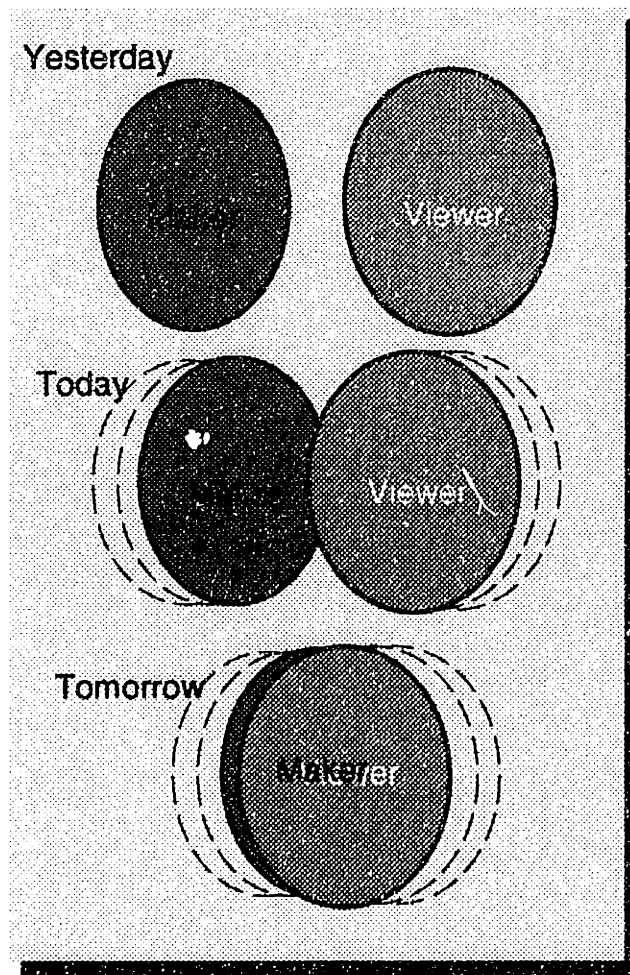


Figure 2 - Shows how the viewer's and the maker's roles merge when interaction is introduced. The viewer becomes an editor. This changing role happens over a continuum as the viewer is able to affect in different ways what is occurring on-screen. The viewer might only decide where to go next in an interactive movie but might shoot and kill a character in a video game. In both cases, the viewer is making decisions about what occurs next.

The viewer's activity and the maker's activity begin to overlap. As the viewer starts making choices about what to watch or which thread to follow, he starts making choices that have traditionally been part of the maker's activity. A big difference between the maker and the viewer is that the maker usually has a clear idea of the story and how to tell it. A filmmaker knows what material has been shot and how to

edit it all together to tell a story. A viewer does not usually think about these things. Giving the viewer a visual representation of the story can help make sense of where in the story they are and where they are going.

2.3. Interaction, Interactive Cinema and Previous Work

Until the advent of computers, story telling had largely been a one way activity. The audience, reader or listener had very little control over the arrangement of what plays out on TV or in a book. An exception might be found in environments where the oral tradition still prevails. A child listening to a parent tell a story will often interrupt to ask to hear more about the elephant or to ask why the ham is green. These interruptions redirect the narrative flow. This usually happens in a fluid manner because the parent is able to process the child's request and adapt the story in real time.

A completely different kind of experience is a video game where the player is constantly deciding what should happen next. If he stops playing, the action usually comes to a rapid end (usually through death). In games, there is usually very little story building activity (in the context that I have laid out). There is a lot of interaction largely aimed at controlling split-second actions - like throwing a punch or shooting a bullet.

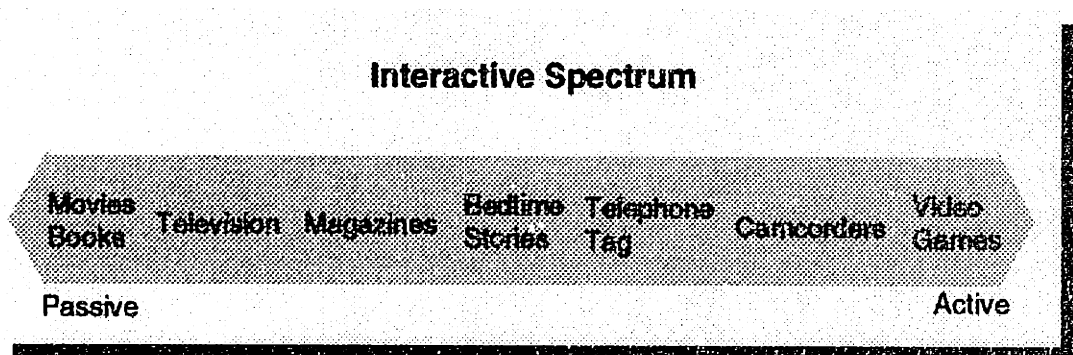


Figure 3 - Shows how some interactive activities relate to each other along an interactive spectrum

Figure 3 shows an interactive spectrum along which different story activities might spread according to the kind and amount of viewer activity (interaction). At one end are books and movies, where viewing is the primary task. At the other end are games where interaction is the main task. In between, we find magazines which invite non-linear reading, bedtime stories which invite interruption, telephone tag

which demands some story construction, and camcorders which require story acquisition and construction. Where does digital cinema fit in? Much research has been done in different areas of the interactive spectrum. A brief introduction to some of the work will help frame the research of this thesis in the context of what has been done before.

A brief overview of some of the projects follows. Some investigate the way audio-visual material is displayed. Others look into how the computer might help arrange the material as it plays out over time.

First, we take a look at various models for building stories.

2.3.1 Aspen (1979-1981)

Aspen is a video-disk based application that was described in Chapter 1. It is an early example of an environment that allows the viewer complete control over what is seen. A story is created in the sense that *Aspen* creates a spatially-referenced narrative where the viewer traveled through a visual representation of a space. Short video pieces could be seen upon entering certain buildings, such as a police officer talking in the police station.

2.3.2 New Orleans (1982-1987)

New Orleans Interactive is a documentary which begins to explore the importance of attaching computer readable content descriptions to video clips (Davenport, 1987). While watching a linear movie about New Orleans as the city prepares for, holds and concludes the 1984 Louisiana World Exposition, it is possible, using keywords attached to people, places and themes, to get more information relating to what is currently being watched in the linear movie. Multiple narrative threads are created through the material through the use of thematic and chronological filters.

2.3.3 Elastic Charles (1988-1989)

Elastic Charles is an example of a hypermedia journal that uses micons as dynamic representations of video segments. The viewer experiences different stories along the Charles River that are each represented by micons and can be accessed by clicking on them. The choice of which micons to bring forth to the screen is based on a pre-defined, fixed table of contents and on dynamic links that can be made by the viewer or the author (Davenport & Brondmo, 1990). One of the important research issues to emerge from Elastic Charles was the demonstration that a

significant amount of information about the content of a video sequence can be packed into a short segment of repeating frames: a micon. An arrangement of these micons on the screen provided an intuitive way for the viewer to interact with new material related to what is currently being viewed on screen. The micons appear and disappear as the viewers travels along the river, allowing the viewer to follow links to related material.

2.3.4 Rows Wharf Intrigue (1988-1989)

Ben Rubin explored the use of layered filters for building a personalized movie. His constraint-based editing system selects shots based on limited input from the viewer and builds an edit list to create a linear story for subsequent viewing (Rubin, 1989). This is an example of how the computer might select shots based on user input to build a story. In this system, the processing is done ahead of time, precluding any real-time interaction on the part of the viewer.

2.3.5 The Electronic Scrapbook (1990-1991)

Amy Bruckman developed a system to help the editing of home videos. Computer readable descriptions are attached to the video material using Arlotje, a knowledge representation system developed by Ken Haase (Haase, 1992). This representation allows material to be retrieved from the database using complex queries that might be able to find all shots of a character at a particular location, or all "important firsts" made up of key events in a child's development. A story-model is then used to process the raw list based on a set of criteria. "The goal is to transform a long, unorganized list into something that is enjoyable to watch in sequence." (Bruckman, 1991, p. 62).

2.3.6 Homer (1991-1992)

Homer is a system designed by Lee Morgenroth that uses a specific story model created by the movie maker to create a video story from a supplied database of logged video (Morgenroth, 1992). The aim is to automatically build a coherent narrative from video that has been extensively described either by an automatic logging system or manually by the movie maker.

The following projects explore the display of moving images:

2.3.7 Windows on the World (1981)

Dr. Richard Bolt developed a project called *World of Windows* that simulates the display of multiple streams of video on a large screen (Bolt, 1981). The viewers

controls the size and audio level of each stream by looking at a particular stream for long enough. The aim was to help the viewer make sense of a large collection of audio-visual material through a more passive interface (pointing and looking). This experiment looks into how the arrangement of multiple streams of video material on-screen can assist the viewer's activity.

2.3.8 The Video Streamer (1991-1993)

The Video Streamer is a powerful visualization tool that uses the z-axis of the screen to represent a sequence of video over time (Elliott, 1993). It is an intuitive environment that allows the viewer to watch video as it streams off the top left portion of the screen in a 2 1/2 D manner and then to grab chunks of it. Simple manipulation tools allow the viewer to grab video blocks and to drop these blocks into a collage environment where it is possible to arrange the video by creating categories for the grabbed material. The Video Streamer demonstrates the intuitive nature of a 2 1/2 or 3D representation of time-based media such as video. It also demonstrates how a simple technical leap (displaying the video in a streamed fashion) can mean a huge conceptual leap for the viewer who can now see an order of magnitude more information.

2.3.9 LogBoy Meets FilterGirl (1992-1993)

LogBoy and FilterGirl together form a tool kit that enables a movie maker to create personalizable movies suited to the digital environment. LogBoy is a logging module that allows movie elements to be logged in a graphical manner; the entire database can be viewed at once, making it easy to see how content is distributed throughout the database. FilterGirl is a filter-based real-time shot selection system that uses the descriptions logged using LogBoy. The system builds sequences on-the-fly based on viewer preferences and is a powerful environment for the creation of interactive cinematic experiences (Evans, 1993).

2.4. Where does my research fit in?

The research described above covers many areas of the interactive spectrum, from the making of home-movies in Bruckman's system to the viewing of multiple streams of Bolt's system. My research concentrates on the viewer's activity as it relates to narrative movies. The situation I am interested in is one where the cinematic material is collected and annotated by the maker and the story models are generated by the maker. The maker is very much in control of what appears on screen and the manner in which it appears. The viewer interacts during playback by

constructing the story and controlling payout. I like to think of the viewer as being in the position of the child listening to a parent (the maker) tell a story. The child can interrupt and redirect the payout and build a story in its mind as it hears and sees the narrative material. It does not originate the story in any way, but rather listens to and redirects what the parent is telling. Likewise, the viewer of the movies I want to make can redirect and review what is playing out on-screen, but does not actually originate or build the story.

IC Portrait tries to create a flowing narrative with multiple clips of digital video that appear in different parts of the screen. The viewer navigates the clips by choosing which one to view. Both the orchestration and the positional editing occur in an unstructured manner. The experience lacks narrative and visual structure. The rest of this thesis looks at one way of creating structure to help the viewer's story building activity.

3. Digital Micromovie Orchestration - An Endless Conversation

One of the problems with *IC Portrait* is its lack of narrative structure. We realized this at the time and decided that the next step in the research agenda was to tackle a simpler problem rather than to try and solve all the problems at once. This meant concentrating on orchestration and getting the computer to build a simple conversation between two characters on-the-fly. The result of this effort is a movie called *An Endless Conversation* and a system called the Digital Micromovie Orchestrator (DMO). The DMO uses a system of layered filters to create a "turn-taking" conversation between two characters. This chapter outlines the process used to create this movie and delineates the positive and negative features of the approach. The DMO works well at the micro level of building sequences, but makes it difficult to build more complex narrative structures.

3.1. Concept

Instead of working with ten or more characters, we decided to narrow our cast down to two. The aim was to create a conversation between two characters. *An Endless Conversation* is composed of a collection of utterances by David Kung and Thomas Aguiere Smith, two students from the Interactive Cinema group. The computer decides how to sequence their utterances so as to build a conversation. Even though a conversation between two characters is not considered a story, it is a good narrative structure with which to start. Cinematically, it need not be that complicated, yet the content can be interesting because we rely on the scripted dialog of two characters to entertain the viewer.

3.1.1 A Filter-Based Turn Taking Conversation

We wanted to create a cinematic sequence where the decision about what to show next is decided on-the-fly computationally. If we want an environment where the narrative can adapt itself to changing circumstances, it is important to achieve on-the-fly shot selection. The full sequence of clips cannot be scripted ahead of time if we eventually want to give the viewer some control over what happens on-screen. As a first step toward achieving this, the machine needs to know what a cinematic sequence is and how to string it together. This knowledge is encoded by the Digital

Micromovie Orchestrator (DMO), a system that uses machine readable annotations and a layered filter structure to sequence digital video clips.³

3.2. Production

3.2.1 Software

The DMO has two elements: the video database which consists of video clips with sketchy descriptions bound to them, and a layered filter structure that uses the descriptions to sequence the video clips in a cinematic manner. A sketchy description is an annotation, determined by the maker based on the needs of the story, that contains the minimum amount of information for use in a particular movie (for example, there is no need to describe the weather in a particular shot if the system is not going to use this information to build a sequence). The video is shot by the movie maker and edited down into coherent clips with beginnings and endings that make sense (they do not cut off a character in mid-sentence for example). A logger designed as part of the system by Ryan Evans is then used to bind the sketchy descriptions to the video clips. The shot selection module makes use of these descriptions to filter out unwanted clips using a collection of filters designed by the maker. It is the layering of multiple filters, each filtering the database of clips according to its own particular rules, that gives the system its sophistication and allows it to build coherent cinematic sequences on-the-fly.

3.2.2 The Logging Process

Descriptions are attached to the clips in the form of slots and values using keywords as descriptive elements. Each clip has several slots associated with it and each slot can have one or more values. The slots used in *An Endless Conversation* include:

- character
- type of utterance
- subject of utterance
- pacing
- rating.

³ The original DMO system was developed by Ryan Evans and Mark Halliday. The work was supervised by Glorianna Davenport and sponsored by the Kansa Corporation. The DMO extends previous research on automatic assembly carried out in the Interactive Cinema Group. The system assumes that the movie maker explicitly designs a story for interactive playout using simple annotations and layered filters. The concepts were worked out by Mark Halliday and Ryan Evans. Most of the initial coding was done by Ryan, with later additions being carried out by Mark. (Davenport, Evans, & Halliday, 1993).

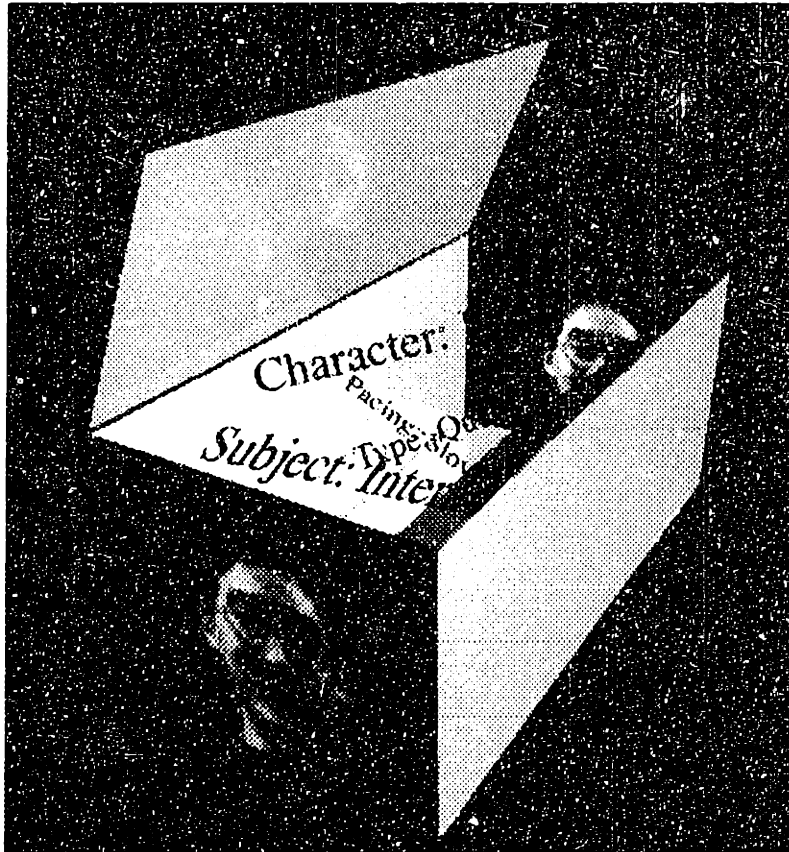


Figure 4 - A visual representation of a clip with slots and values attached. For example, the slot *Character* has the value *Dave* attached. The slot *Type* has a value *Question*. The values are attached using a logging module. (This video box representation of a clip was inspired by Eddie Elliott's Streamer).

To indicate that a particular shot has Dave in it for example, the *character* slot has *Dave* as its value. To indicate that Dave is talking about a light, the *subject* slot has the value *light* bound to it. Only those values determined to be important to the filtering process are logged. The filmmaker manually binds these descriptions to all the clips using the logging module. At this stage, it is a manual process, and quite a laborious job, but at some time in the future, as image recognition technologies improve, logging may be partially automated. It is these kinds of sketchy descriptions that form the basis of the video annotations.

Other researchers in the Media Lab have been looking at more comprehensive descriptive schemes. Thomas Aguiere Smith, working on an anthropologist's video notebook, developed a description system called stratification, a context-based layered annotation method which treats the descriptions of video as objects. Stratification encodes *thick descriptions* (Smith, 1992), meaning that a video clip is extensively described for multiple uses, as opposed to the DMO which uses sketchy descriptions for use in a single domain. Marc Davis has created video annotations that seek to describe a large subset of all the recognizable features of a video clip (Davis, 1993). For the purposes of the DMO, this approach is unnecessary. The

DMO relies only on the descriptions that are used by the filter structure. All other annotations are superfluous. This is closer to the kinds of descriptions used in Amy Bruckman's Electronic Scrapbook (Bruckman, 1991), which seeks to describe home movie footage that will be used only in the context of a home movie editing system. A limitation inherent in the sketchy description approach is that the video can only be described for a single purpose. It is not intended for re-use by any other application because the descriptions are specific to the DMO movie they were created for.

From the perspective of a movie maker there is little demand for re-using video to tell stories. The images that most movie makers include in their films are carefully shot to tell a particular story. I am taking a similar approach to the construction of multi-threaded stories by working with a collection of clips that have been shot and edited specifically for the movie being made.

3.2.3 Filtering

An Endless Conversation uses three filters for orchestrating the sequences:

- **a story template filter** which defines a conversation as a *question* followed by an *answer* and then a rebuttal
- **a character filter** that makes sure that if Dave asks the question Thomas answers it and vice-versa
- **a continuity filter** that makes sure the subject of the answer matches that of the question.

These filters make use of the sketchy descriptions to select appropriate shots. Just as a human editor uses logs that have been created by an assistant to find clips and uses knowledge of story structure, cinematic rules and available video to create sequences which build up to stories, the layered filter structure mimics this process in a primitive way. It takes as its input the entire database of video and successively filters out clips that do not meet the specification of each filter. At the end of the filtering process, assuming the database is complete, the system returns a list of one or more appropriate clips. If there is more than one choice, it randomly chooses one of them. This process is carried out each time a new cut is needed. It is true real-time shot selection.

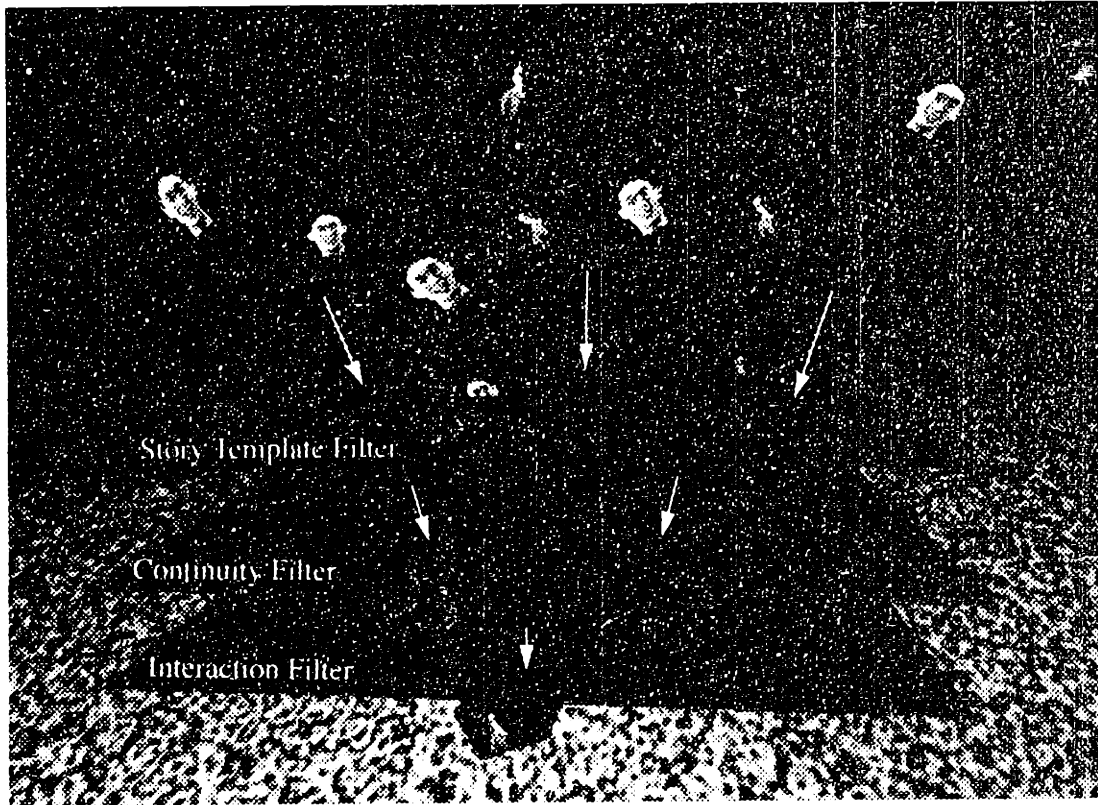


Figure 5 - A visual representation of the filtering process. Individual filters successively narrow down the list of possible shot to only those that satisfy the constraints of each of the filters.

An individual filter is actually quite simple. Each one contains a simple rule or a description against which a clip description is matched. For example when Dave asks a question about the light ("What's this light?"), the continuity filter saves the *subject* of the question; the saved value *light* is then used to filter out all other subject values when the story template filter is at an *answer*. Every time a new clip is needed, each filter is successively applied to the database in this way, narrowing the list of clips down each time. As long as the database has the right material for the filter set, and as long as the material has been described in the correct manner, the DMO system will orchestrate a sequence on-the-fly.

3.2.4 Scripting

Making a movie that fits within the constraints of the DMO is very different from making a fixed linear movie. It has to be scripted to take into account the new ways in which it will be watched. We have constructed a multi-threaded narrative structure that the viewer can traverse in various ways. The structure of the content has to be flexible because the shot selection system needs to have a variety of clips to choose from for each part of the movie. A question from one character could

have many possible answers and an answer can elicit many responses. The movie is different each time it is watched. Eventually, interaction should be able to change what is happening and it should change things in real-time, creating an interaction mode that we call "fluid interaction".

The actual structure of *An Endless Conversation* is a question and answer dialog session. Dave and Thomas are asking each other questions about interactive cinema. A typical interchange follows:

Thomas: "Hi, what's your name?"
David: "David Kung"
Thomas: "I like that"
David: "What are you doing here?"
Thomas: "I'm a graduate student in the Interactive Cinema Group."
David: "That's true"
Thomas: "What's this light here?"
David: "Well this light fixture actually comes from the first movie-house in Beijing and my grandfather took it as a souvenir when the theater closed down"
Thomas: "Don't listen to this guy. He's an undergraduate. He doesn't know anything."

Not the most riveting of conversations, but it can be followed and it makes sense. This conversation will go on forever, hence the title, *An Endless Conversation*. Dave and Thomas can make three different kinds of utterances: ask a question, give a response or give a rebuttal. They can talk about a variety of different subjects, including interactive cinema, the light, the pterodactyl and various other objects and subjects remotely associated with the Interactive Cinema Group⁴. The database of content is made up of approximately 120 clips which are accessed by the DMO system.

3.2.5 Additive Construction

The success of the playout relies on having a rich database of material. When we first ran *An Endless Conversation*, we had a large number of questions from David but we only had two questions from Thomas. This meant that when we ran the movie, the system only had the choice of two questions from Thomas which severely restricted the breadth of the conversation. To remedy this, we shot Thomas asking the missing questions. After adding these new clips to the database and

⁴The light was a bathroom fixture introduced to the lab by Thomas during his tenure at the Media Lab. The pterodactyl was a graduation present that Stephan Fitch, another graduate student, bought when he graduated from NYU Film School. The lab, which we called Macondo was littered with various objects introduced by different graduate students during their residencies. *An Endless Conversation* provided an introduction to many of the obscure but interesting features of the lab where we spent two years of our lives.

describing them using the logger, the movie changed dramatically as Dave's responses, which had lain dormant in the database, suddenly appeared on-screen automatically. This was one of the most exciting moments for me as a filmmaker. This means that a movie can now be an open structure into which new clips can be easily incorporated (as long as they have been described correctly for automatic playout). The movie accommodates this new material without re-scripting or re-cutting.

3.2.6 Interface

The look and feel of this digital movie is very different from a traditional linear film. The video is small, the frame rate is slow and the sound is of poor quality⁵. This is an unfortunate temporary limitation. The DMO system is entirely separate from the display environment and as the quality of the image improves, the filter-based shot selection scheme can be adapted to any new display environment.

To establish an understandable environment, as well as to mask the poor quality of the video and to compensate for its small size, we designed an interface that frames the image in a familiar environment: a television monitor. The effect is to set up certain expectations for the viewer, who sees a representation of a television and expects a linear movie because the device is familiar. Those expectations are met because the presentation on-screen is a linear sequence (even if it is a little small and slow). It goes on and on no matter what the viewer does, which is similar to what happens on television. The small size of the image is less apparent because it is placed in the context of a larger image and is motivated by being framed within in the image of a TV set.

⁵ This is a limitation of the digital compression scheme being used to display the video (Apple's QuickTime™). Another noticeable limitation of the system is the delay between shots as the filtering takes place. This makes totally seamless cuts between clips difficult, but there are schemes that will make this possible. Simple parallel processing is one; one processor can be used to display the digital video and another can be used to compute the sequencing of the content using the DMO software.



Figure 6 - The interface to *An Endless Conversation*. Clicking on the + or - of the remote controls the pacing. Clicking on the rating card changes the pacing. Clicking on the plug quits the program.

There is one major difference with what happens on television: interaction. Because shot selection is happening in real-time, we are able to take one small step toward making fluid interaction possible. Fluid interaction in this case means that the viewer can intervene in what is happening on-screen at any point during the experience. In *An Endless Conversation*, the viewer can control pacing and rating. To do this, we added two filters to the layered filter structure: a pacing and a rating filter. We then added two kinds of controls to the interface: a remote with a '+' and a '-' on it, and an MPAA rating card that can switch between PG and R ⁶. When the pacing is on +, the filter rejects all shots that have been logged as *slow*. When the rating is on PG, the filter rejects all shots that have been logged as R. With these two controls, the viewer can exert limited control over what happens on-screen by changing the rating and the pacing of the sequence. It is a simple form of fluid interaction. The significant feature of the interface is the fact that the flow is never interrupted and that the interaction can happen at any time.

⁶ The Motion Picture Association of America rates movies to restrict their viewing to certain age groups. PG is for children accompanied by adults. R is for adults only. The system is based on subjective decisions about the appropriateness of movies for certain audiences.

3.2.7 Extensibility

There is another advantage of the DMO filtering system: not only can appropriate clips be added, but new filters can be added anytime. If we want to add functionality to the interface or change the way the movie is structured, we do two things: add the appropriate descriptions and incorporate a new filter into the filter structure. As an example, I wished to have *An Endless Conversation* come to an appropriate ending. This was a first step toward making the narrative structure a little more complex. To do this, I chose to add a time-based filter. This filter specifies that certain things have to happen at certain times during the viewing. I chose to specify a *beginning*, *middle* and *end* to occur during the first third, second third and final third of the playout. Interaction on the part of the viewer consists of entering the playout time of the movie before it starts. Using the existing clips in the database, I described some questions as *beginning* (What's your name?) and others as *middle* (What's this light?). Unfortunately I didn't have any appropriate ending questions. *An Endless Conversation* now had a *beginning*, *middle*, *end* structure; it was no longer an endless flow. The movie starts with Dave or Thomas asking the other one his name and goes on from there. It can be of any length specified by the user but the better movies happen when the length is set below 90 seconds; with longer settings, the clips start being repeated.

3.3. Criticism

3.3.1 Results

What does *An Endless Conversation* demonstrate? It is a promising approach to toward the ideal of creating a complex interactive movie which maintains viewer reverie. It is a first step toward computer-controlled story-telling. The DMO does this by modeling at a very simple level the job of the editor. It starts with a collection of movie footage that has been carefully logged in a machine readable format. It then uses its knowledge of cinematic orchestration to create a sequence of clips that makes cinematic sense to a viewer. Granted, a conversation between two characters is not much of a story, but it is a first step toward making more complex narratives.

The important features of the system are:

- **real-time shot selection** - the movie is assembled in real-time and plays out without user intervention
- **fluid interaction** - the viewer can change the playout at any time during viewing

- **extensibility of the filter layers** - this allows the structure of the movie to be changed and extended without a complete rewrite
- **effective use of the video database** - the system relies solely on material that has been shot and described

3.3.2 Problems

Filters do have some drawbacks:

- **database incompleteness** - if all the clips are not in database, holes appear in the sequences
- **combinatorics** - too many filters means too many shots to satisfy all possible combinations of filters
- **micro-structure** - it is difficult to describe higher level story structures

If the database does not contain material for all the possible filter combinations, holes are created in the narrative sequence. If one of the filters in the system returns nothing, meaning that there are no clips in the database that satisfy its constraints, the DMO backs up one level and returns the list of available clips from the previous filter. This means that there have to be shots to match every permissible combinations of filters. This is a simple matter to resolve in theory: shoot all the possible shots that are needed for all the possible combinations of filters and the movie will make narrative and cinematic sense. In practice it is very difficult. It is almost impossible to know which clips are necessary. The research carried out by Ryan Evans makes this easier by allowing the movie maker to "run the DMO in reverse". Instead of producing a sequence of clips on-screen, it returns a list of descriptions that need to be satisfied for the system to work effectively (Evans, 1993) This is a necessary tool for making the production of this kind of movie feasible. I use a different approach.

The database for *An Endless Conversation* consists of individual shots. Filters orchestrate these shots to build sequences. They work at a low level of abstraction (micro-level), building narrative structure by filtering in a layered fashion; each filter has a small determination on the final shot that is selected by the system. The problem is that the filters have to be carefully selected and arranged to orchestrate shots into a sequence that looks cinematic. The research described from here on describes how sequences (instead of shots) can be orchestrated into more complex stories.

4. Orchestration through Linked Story Structures

The Digital Micromovie Orchestrator builds sequences from shots. As a movie maker interested in creating more complex stories, my question is: Now that we can orchestrate shots into sequences, how do we orchestrate sequences into cinematic stories? *Train of Thought*, described in this chapter, achieves orchestration through a linked story structure. Links build multiple threads through a collection of short sequences that have been pre-edited to tell different parts of the story. The chapter also describes a 2^{1/2}D interface that uses the positional editing of the sequences to give the viewer a sense of where they are in the story and where they are going. It is difficult to author multi-threaded narratives using links, but the 2^{1/2}D interface starts to give the viewer a sense of the overall narrative structure.

4.1. Concept

4.1.1 An impressionistic collage through a linked story structure

If the computer is going to construct a story rather than just a sequence out of material stored in a database, I have to be able to define what constitutes a story so that I can then represent it in a computer readable format. The dialog as we defined it for *An Endless Conversation* is quite a simple construct. A more complex narrative needs more thought. What attributes are necessary to build a story? What are its constituent parts? How is it structured?⁷

4.1.2 A Multi-threaded Structure

The DMO selects multiple paths through a database of cinematic material consisting of shots. It builds up a multi-threaded structure in the form of sequences that are different each time the movie is watched. I want to build a slightly more complex narrative from a collection of cinematic material consisting of short sequences stored and described in a database. The computer will then select multiple paths through the material so as to build a story.

4.1.3 Granularity

The first question to address is the granularity⁸ at which the story should be structured. *An Endless Conversation* operates at the shot level. The story is

⁷This is where the actual work carried out for this thesis starts. It grows directly out of what has been described so far with *IC Portrait* and *An Endless Conversation*.

⁸ Granularity refers to the length of the video clips with which the story is constructed. Granularity ranges from a single frame to an entire movie.

described at quite a low level, with much computation going into building sequences. As the research progressed, I realized that to tell stories, I had to work with sequences rather than shots as the basic story units. A story-unit that is a sequence tells a small part of a story; the maker creates individual sequences by editing together individual shots to build a small element of the story (introduce a character, present a dream, give a short conversation, etc...). These story units are structured so as to cover different events in the narrative. Some are introductory sequences, others are endings. The computational problem is to orchestrate these story units to build higher level story-structures. The aim is to build up narrative meaning in the viewer's mind.

4.1.4 An Inspiration for Content and Structure

An original inspiration for the story structure was Ricky Leacock's film *Oeuf a la Coq de Ricky Leacock*, a linear movie with a non-linear structure. The film is made up of a collection of short sequences that each tell a small part of the story. *Oeuf a la Coq* is a personal journey filmed by Leacock and his partner Valerie Laconde as they travel around France visiting friends and family. The unifying theme is eggs: how to shell them, how to cook them, how to eat them. The resulting story is an exploration of delightful personal moments that add up to an impressionistic portrait of two peoples' personal lives. Each vignette gives a little impression of the characters, actions and settings in the story. Ricky stated that the movie is meant to be watched on television where the viewer can jump between channels. It is not necessary to watch the entire movie in one sitting to understand the story of Ricky and Valerie; it should be possible to understand and appreciate the story by viewing parts of it. This seems to be the kind of narrative structure suited to a multi-threaded narrative.

4.1.5 Sequencing Story Units

My approach is to build up a story from a collection of story units (or sequences) that each tell a part of a story. The aim is to get the computer to sequence these story-units in a similar manner to the way the DMO sequences shots to build sequences. The sequencing should build up a story in the viewer's mind as the material plays out. A movie maker thinks about movies at different levels of granularity: frames, shots, sequences, and stories. Similarly, the problem of orchestrating material with the computer needs to be abstracted at different levels of granularity: the DMO abstracts at the shot level; the research described from here on abstracts stories at the sequence level.

4.1.6 Documentary Content

What kind of narrative material is best suited to this kind of structure? Documentary footage appeared at the outset to be ideal, because a lot of footage is shot, documenting many different threads and perspectives on the characters, actions and settings of the subject being filmed. It is then the job of the editor to construct a main thread through the material. In this editing process, many threads of the story are left on the cutting-room floor as the movie is edited down to one linear strip of material. Instead of cutting the material down to one linear thread, I want to cut the footage down into little vignettes that tell different parts of the story and then have the machine sequence them into multiple threads instead of leaving some of them on the cutting room floor.

I started filming a modern dance company, following the choreographer and the dancers as they prepared for a performance in New York city. I captured many different perspectives on the events: dancers rehearsing, the choreographer directing them, dancers talking, the choreographer explaining the piece and the final performance itself. I ended up with many hours of material that follow many potential threads.

As I started to cull the material down to vignettes, I was confronted with the realization that even though I had shot a lot of material, it was not the right kind of footage for what I was attempting to do. The material was not engaging enough to sustain the story; the main character was not as good on-camera as I had hoped. The dance footage, although nice to watch, was difficult for me to edit down into a clip suitable for small vignettes. A dance piece is usually a few minutes long and the vignettes I had in mind should not be longer than 30 to 45 seconds⁹. I had to rethink my exact needs for the content. One of the criteria was that the material cut well into relatively short sequences. The other was that I have more engaging characters. I felt more comfortable working with fiction, as it provided me with more control over the content.

4.1.7 Fictional Content

I was enrolled in Ken Haase's *Projects in Knowledge Representation* class. I was interested in some of the basic concepts of knowledge representation that formed the underlying structure of *An Endless Conversation*. My project in that class,

⁹ QuickTime™ also presented some technical limitations. It was difficult to work with longer sequences because of the difficulties of digitizing (lack of disk-space) and editing (interface and speed restrictions).

working with Ryan Evans and Lee Morgenroth, was to explore how narrative structures can be encoded using knowledge representation. The project involved shooting material for the story and then creating a computational framework that would enable the computer to sequence the material into a multi-threaded narrative.

4.2. Production

4.2.1 An Impressionistic Collage

To explore narrative structures, you need to have a story. I needed one that would allow flexibility in the telling as well as one that would have engaging characters. One of the strengths of *An Endless Conversation* was the fact that we concentrate on two characters. I wanted to develop a fictional story which concentrates on the events and actions surrounding two characters.

One film that has always intrigued me with its flexible structure, or even anti-structure, is *Un Chien Andalou* by Salvador Dali and Luis Buñuel. According to Buñuel, a kind of automatic writing was used to generate the film, a process he referred to as a “conscious psychic automatism” involving the probing of the two authors' dream images (Kuenzli, 1987, p. 144). There is no conscious effort to build connections between the different scenes of the film. Man Ray, a prominent member of the Dada movement, says it best when he describes his effort to make a Surrealist film, *Emak Bakia*: “I had complied with all the principles of Surrealism: irrationality, automatism, psychological and dramatic sequences without apparent logic, and complete disregard for conventional story-telling” (Kuenzli, 1987, p. 3). This automatism and irrationality provided me with a forgiving environment within which to explore narrative structure.

Watching *Un Chien Andalou* is like looking in on someone's dreams. Analyzing dreams takes work and imagination as the significance of the events has to be worked out. The viewer of *Un Chien Andalou* has to work hard to build connections between the supposedly disconnected events happening on-screen. Watching a multi-threaded movie is a similar experience as we will find out later.

I did not want a structure that was as disconnected as Bunuel's film, but I did want a story that would allow me some flexibility in sequencing the narrative. In *An Endless Conversation*, the utterances of each character were scripted and shot in such a way that they could be re-used in various contexts. Similarly, the vignettes should be re-usable in different contexts within this story. Out of these ideas, and

from extensive conversations with Lee and Ryan, the concept for *Train of Thought* was born.

4.2.2 The Story

Train of Thought is a love story about two characters: Jack and Nicole. Jack is a slacker¹⁰ who is in between jobs, having just been fired from his job as an assistant at a camera store. He asks Nicole out after seeing her in some of the pictures he was developing. Nicole is a young woman who knows what she wants. She is a buyer at a large clothing chain and loves her work. The story enters Jack and Nicole's lives as they are about to be separated. Nicole has been offered the opportunity to go and work in London for a year. Jack is torn between wanting her to go to advance her career, and wanting her to stay with him.

The film is scripted to straddle the line between dream and reality, following the main actions of the story as well as the characters' subconscious thoughts. The viewer can follow either the dream or the reality thread of the story. The title of the movie, *Train of Thought*, suggests a thread that the viewer's mind might follow while watching the movie. The word *train* invokes the feeling of travel and movement, a central theme of the story.

I decided upon this structure to have the flexibility of putting the material together in a variety of ways. I did not want to be bound by the stricter linearity of a cause and effect story (where the events and actions have to follow each other in a realistic manner). I needed material that could be viewed in different sequences and even better, that could mean different things in different parts of the narrative. For example, if there is a shot of Jack driving in his car, it should be able to come before the scene where he picks her up to take her to the station or it should be able to come after he has dropped her off and is driving to the diner.

¹⁰ The inspiration for Jack's character was influenced by Richard Linklater's film *Slacker*, which also has an unusual structure. Linklater's film follows a hundred or so characters in and around Austin Texas. The film starts with one character (played by Linklater) riding a taxi back from a bus station. He witnesses a hit and run accident. The camera follows the driver who drives home without looking back at the old lady he just ran over. The film jumps from one character to the next in this manner throughout the movie. *Slacker* is a journey through the lives of a collection of oddball characters (slackers) who all exist on the fringes of society. It is similarly made up of a collection of little vignettes; the difference is that they are all linked in a linear manner. We leave one character to follow the next one for a few minutes, following 100 or so slackers in this manner.

4.2.3 Casting & Shooting

The movie was scripted and cast so as to make the most of the limited resources available. We only had two days to shoot with a Hi-8 camera, no lighting equipment or crew and no location permits. It was guerrilla filmmaking at its best.

Jack and Nicole are young twenty-something characters, products of the TV age, very different in their aspirations, but happily in love because of the freshness that their divergent views bring to each other. We found two actors who were able to bring these characters to life. Jack is played by Ian Dowell and Nicole by Sharon Cinnamon. Both of them were found through the wonders of the network. We sent electronic mail to the local theater mailing list and received numerous responses. Jack and Nicole were chosen after on-camera screen-tests which consisted of reading passages from Lewis Carroll's *Alice in Wonderland* and Jack Kerouac's *Book of Dreams*.

An important attribute we were looking for in the actors was improvisation skills. We were shooting on a low budget with no time for rehearsal and had to rely on the actors to go with the flow at unprepared locations with limited script details. The movie lives on the fuzzy edge between documentary and fiction. Having them read from the slightly unusual texts quickly showed us how they could deal with the unexpected.

The script itself consisted of location information (place, time), character information (who was in the scene), actions (what were the characters doing) and limited dialog lines (what were they talking about). This script format may appear a little skimpy when compared to a traditional Hollywood script, but it served our purposes beautifully. Its loose format gave the actors, the cinematographer and the director room to improvise and the opportunity to take advantage of unplanned events (such as an unattended Amtrak train sitting in South Station).

Using this style of filmmaking, we were able to capture the footage necessary for approximately 35 vignettes, ranging from Jack waking up from a nightmare, to an argument over spoiled milk with Jack and Nicole, to Jack dropping Nicole off at the train station on her way to Kennedy Airport in New York City. These clips were edited together into scenes ranging from 15 seconds to 90 seconds in length. We also interviewed the actors (in character) after filming each scene. We asked them how they felt about the events in the context of their lives. The inspiration for this

came from Woody Allen's film *Husbands and Wives*, where the characters are interviewed about the events taking place in their lives. These interviews give an added perspective on the story. The 35 vignettes form the collection of clips in the database.¹¹

4.2.4 Software

4.2.4.1 Links

The filter structure that drives *An Endless Conversation* uses a story template to decide how to structure a conversation: a question is followed by a response and then comes a rebuttal. This template is repeated alternating with a question from each character. The story is rigid, but the utterances from each character give the movie its richness. This first attempt at computationally building *Train of Thought* takes a step back from this rigid story structure and tries to create a framework that allows the encoding of multiple paths through the video material in the form of links.

The system uses Ken Haase's knowledge representation system called Framer (Haase, 1992) to encode the narrative structure¹². Framer is a database system for knowledge representation and media annotation. It provides a persistent object storage facility (a data structure facility where today's changes are visible tomorrow), and recursively annotated objects (descriptions have annotations which are themselves capable of having annotations recursively). There are a number of other capabilities built into Framer including an extension language: FRAXL, and a representation language language, ARLOtje, providing inferential capabilities. The story structure representations described in this thesis uses only the most basic functions available in Framer. These include the data structure facility of Framer and the inferential capabilities of ARLOtje.

The software is divided into two distinct parts: the logger and the sequencer. Ryan Evans built the logger that allows the maker to attach descriptions to the video material. The logger allows the movie maker to view the entire database of movies at once and to create a graphical description space where changing the value of a

¹¹ For more information regarding the way in which this video material was gathered, please refer to Appendix A

¹² Framer is a persistent knowledge representation system with extensive inference and other capabilities. The first version of *Train of Thought* described in this chapter only uses the simple database functions built into Framer. More extensive capabilities will be exploited and described later.

slot is accomplished by moving an icon representing the movie into the appropriate value area.¹³

The sequencer uses the descriptions attached to the clips by the logger. The logger allows link-descriptions to be attached to a clip. The links create a hypertext-like structure with multiple paths through the material. Viewing the linked video clips is equivalent to taking one of many possible paths through the material. Instead of only logging slots and values (Ex. character: Dave) as was done in *An Endless Conversation*, the movie maker also logs links between clips. In *Train of Thought*, the maker logs two kinds of links: dream and reality. In this manner, a reality scene always links to another reality scene and to a dream scene. The same is true of a dream scene. This scheme allows the viewer to follow the *dream* thread or the *reality* thread through the material. The logging process thereby builds up a narrative structure as the maker links up all the possible narrative paths through the material.

¹³ For a complete description of the logging space, please see (Evans, 1993)

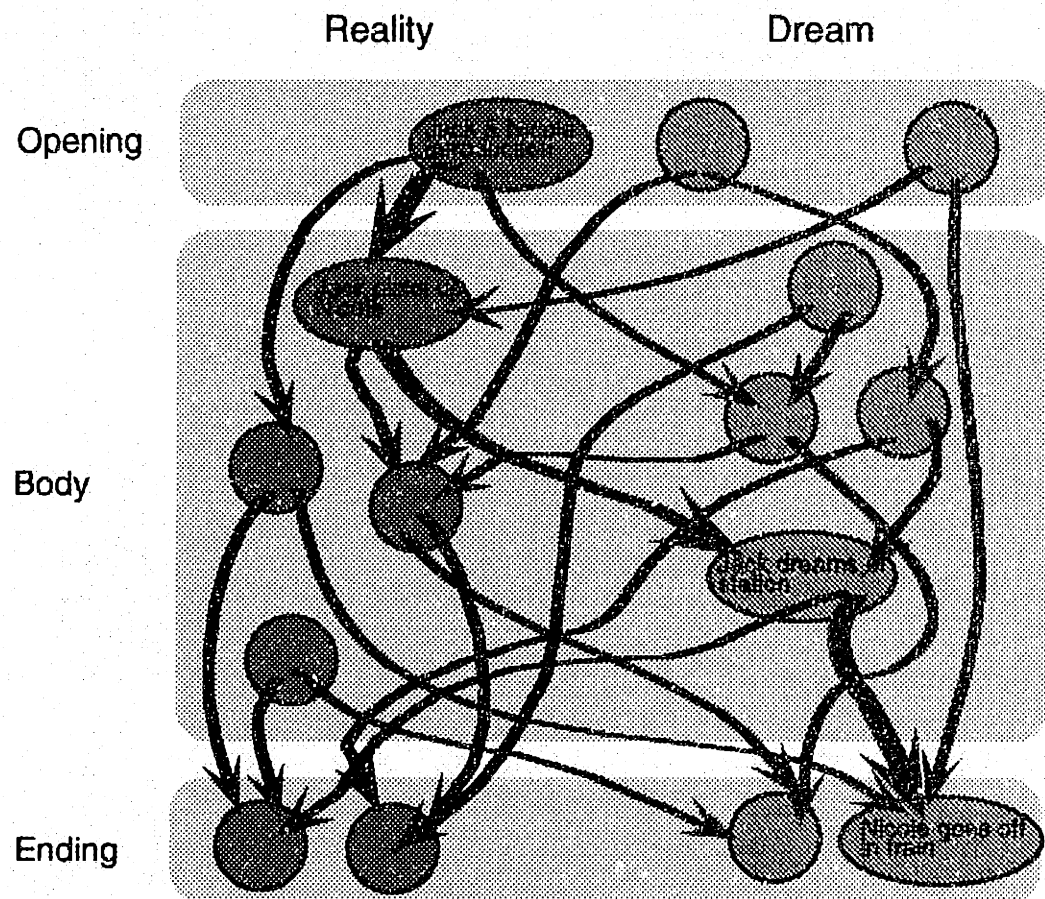


Figure 6 - A visual representation of linked narrative threads. Nodes represent story units in the form of sequences that have been pre-edited by the maker. Each node has one or more *Dream* (light arrows) and *Reality* (dark arrows) links that create paths through the material. The movie ends when an *Ending* clip is reached. A sample path through the material is shown. The confusing appearance of the diagram reflects the difficulty of orchestrating using links.

Figure 6 shows how the narrative threads are structured. The nodes represent the video clips. The arrows represent the links, with *Reality* on the left and *Dream* on the right. In addition to the links, simple descriptions are attached to clips that denote good *opening* or good *ending* narrative elements. Using this approach, the sequencer makes sure that a story always opens with a scene that has been logged as *opening*. This clip links to one or more *body* clips. If there is more than one link off a clip, the sequencer randomly picks one of the possibilities. The links are traversed in this manner until an *ending* scene is reached. The links form multiple threads through the material in this manner.

Using this linked structure, the maker creates a multi-threaded narrative that the viewer can follow by choosing to follow either a *Dream* thread or a *Reality* thread. The viewer actually moves the story forward by choosing to follow either the *Dream* or the *Reality* thread. This approach allows for the creation of less structured narratives than those in *An Endless Conversation*. The maker is not bound to a specific story template structure and can log a less restricted path through the material.

4.2.5 The Interface: a visual context for the viewer

In *Train of Thought*, the interaction consists of deciding which thread to follow. As was discussed in Chapter 2, positional editing, the manner in which a narrative is displayed, is important if the viewer is to make sense of the story and to make choices. The story needs to be displayed so as to give the viewers a sense of where they are and where they are going. This puts interaction into context and creates a visual structure for the viewer to follow. *Train of Thought* does this by giving the viewer a sense of what is happening over time.

I want to make a story easy to visualize. The DMO and the link-based sequencer attempt to create a story structure by representing the narrative information in a computer readable format. The aim is now to represent the story in a human readable format. The viewer should be able to look at the screen and immediately see where they are and where they are going. This becomes important when the viewer is expected to make decisions about where to go next. When a reader picks up a magazine or a book, there are cues that immediately inform him about where he is and where he is going: the layout of the page, the thickness of the magazine as compared to the current position in the publication. How can some of these physical cues be carried over to a story as visual cues in the digital domain?

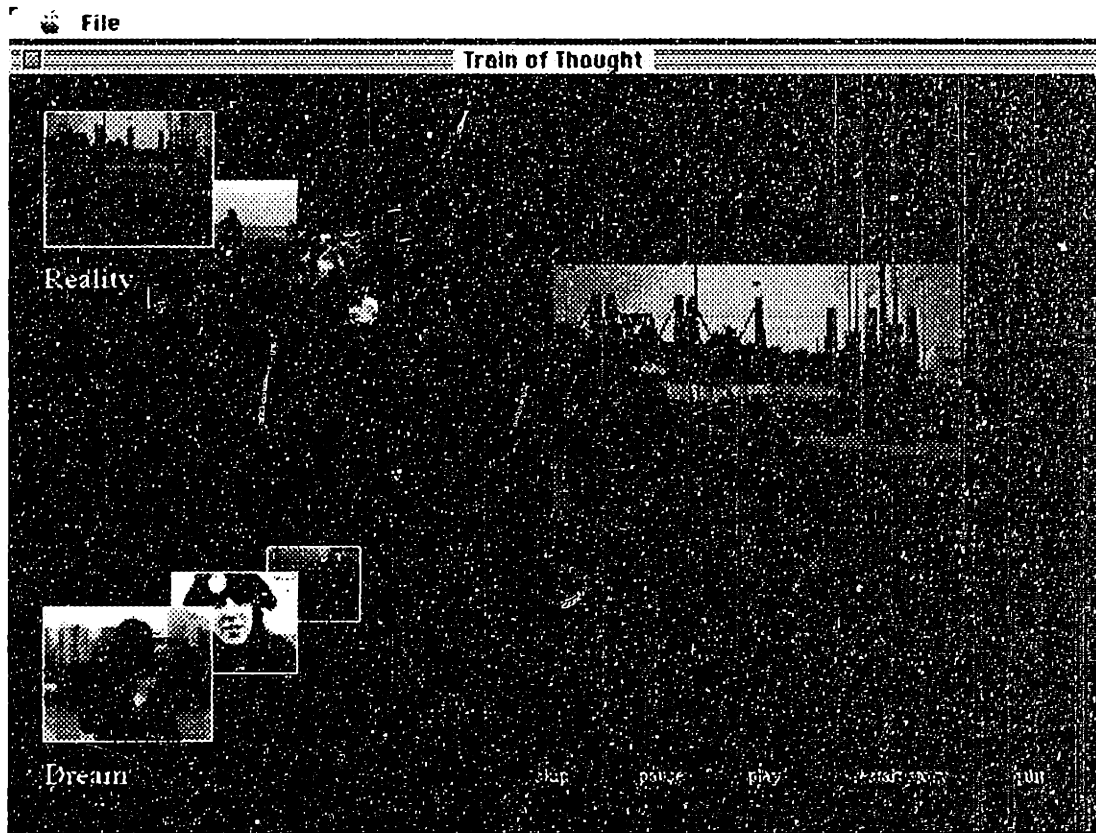


Figure 7 - The interface to *Train of Thought*. The left side of the screen shows the story units as they stream off toward an imaginary vanishing point. The right side shows the main playout area, where the clips can be seen.

The screen environment for *Train of Thought* takes advantage of perspective in a similar way to Eddie Elliott's streamer. What has been seen in the past becomes smaller as the story progresses and streams off toward an imaginary vanishing point at the center of the screen. The figure above shows two streams on the left-hand side of the screen, *Dream* and *Reality*, as they recede off to the right. The large frame on the right is where the chosen clip plays out.

As the story progresses, new choices of vignettes appear in the form of icons on the left side of the stream. The other icons recede toward the imaginary vanishing point. In this manner, the previous stages of the story remain on the screen in the form of icons that get progressively smaller. The icons are positioned in such a way that their relation to the overall story is maintained. It is possible to see what came before and after a particular stage in the narrative just by looking at the screen.

In this manner, the cinematic frame is broken down into two main parts: the story layout and the story playout. The story layout (on the left) provides an overview

function, while the story playout window (on the right) provides the traditional frame where the chosen vignettes can be viewed.

With a display environment that makes use of perspective and the imaginary third dimension of the screen, it is possible for the user to make choices in the context of what they have seen before. It is also possible to jump back in time and view a part of the story that was missed. Thus, if the viewer is following the reality thread, but wants to see what the dream vignette contained, they can jump back and view the dream sequence without losing track of the current place in the story.

This display interface is the first step toward giving the viewer a visual understanding of what they are viewing on screen. It gives an overall picture of the story structure that can easily be navigated back and forth in time. What does the actual experience feel like?

4.3. Criticism

4.3.1 The Experience

Train of Thought starts out with a standard title sequence that plays in the main playout window on the right side of the frame. It finishes with the caption "Once upon a time, Jack loved Nicole...", immediately setting up the main premise of the story. When the sequence finishes, one icon automatically appears on the left side of the screen in the reality stream. This is the standard introductory sequence. The user chooses this vignette and sees Jack and Nicole talking about each other to the camera. At its conclusion, two new icons appear and the old one recedes toward the vanishing point. The choice of what appears next is determined by the links that lead off the chosen clip. The viewer chooses to follow either the dream or reality thread and in so doing, determines what thread of the story will play out. The story goes on like this until an *ending* clip is reached (which could happen in either stream). At this point, a closing "The End" sequence appears as an icon. The viewer can choose to end the story or continue by choosing the other icon. At some point (there are only about 35 clips in the database) there is one choice: to end the story. The experience is over.

The viewer can choose to restart the story and to experience a different sequence. This is possible because there is usually more than one link off each clip, making each viewing slightly different. Each viewing gives a different perspective on Jack and Nicole's story.

4.3.2 What works

Now that we've seen an objective description of the system and how it works, we can talk a little about how the system actually performs and what people think of it. The comments are based on observations from a select group of individuals who sat down with the system and actually used it. I wrote down their observations after discussing the experience with them.

4.3.3 Feedback from selected users

4.3.31 Perspective

One general observation was that it isn't necessarily easy to see that both streams move back in time together. There are not enough visual cues to explicitly show that both streams recede together. This is probably related to the fact that there are not enough depth clues to give the eye a good idea of how the icons are laid out in space. Specifically, the top and bottom streams do not have any visual link to indicate that they move together. It is also a static display. The icons representing movie elements do not move smoothly from one stage of the story to the next, but rather jump in fits and spurts.¹⁴ The 2¹/₂D display of the story structure is helpful. It is not something that people necessarily understand immediately; but after a brief explanation, they can interpret the basic layout. It becomes clear that what is in the past is smaller and offset toward the vanishing point.

4.3.32 Two choices

One positive feature of the display is that it clearly shows that there are two threads to the story. It is evident to the viewer that there two icons to choose from. Making a choice is simple to accomplish; clicking on the icon plays the vignette in the playout window on the right. The significance of the choice is a little more complex. In this piece, the viewer is made to choose between two different paths through the material. This adds a new dimension to experiencing the story. As we discussed in Chapter 2, allowing the viewer to interact puts them into the role of the maker, something for which most viewers are not prepared. Some people tried to watch all the vignettes, jumping from dream to reality and back. Others followed only one of the treads. A common problem seemed to be the fact that people did not realize the significance of their act of choosing. The problem is probably best demonstrated by Larry Friedlander's comment that "with interaction, thought on the viewer's part is necessary". The maker is no longer doing all the story construction

¹⁴ This is a limitation of currently available computing power. In the future, it will be much easier to display multiple images that move smoothly and independently of each other.

activity for the viewer. The viewer has to do some of his own. Most people are not used to this viewing mode.

4.3.33 Content and Structure

Some of the problems users of the system had relate to content and structure. The dream vignettes are rather abstract and do not necessarily add anything to the story. Glorianna Davenport said "dream is not a metaphor for anything [in *Train of Thought*]". Watching a dream sequence does not always add to the understanding of the story. It is often more of an impressionistic fantasy that requires too much analysis on the part of the viewer to extract any meaning. On the reality thread of the story, Janet Murray mentioned that "there is no clear sense of a narrative thread." Seeing the vignettes play out in an impressionistic fashion does not always create enough of a narrative flow for the viewer to follow. These are two problems that point to how the viewer understands a story. Bordwell talks about "the goal of creating a meaningful story out of the material presented". This point is even more true in an interactive environment. Every event and every action should be geared toward the viewer's story construction activity, something which is not always true of *Train of Thought*.

4.3.34 Problematic links

An impressionistic collage of vignettes does not always help the viewer's story construction activity. Some of the stories that result from the linking process are coherent. Unfortunately, more than a few are not. Part of this is due to the abstract content of the dream sequences. It is also a result of the authoring process. The truth is that creating links is a painful and unclear activity, even with a logger that allows the maker to visualize the linking process. As more links are added, it becomes very difficult to keep track of the threads that have been created. This is true of a story such as *Train of Thought*, where there are only two threads. It is also very difficult to add or remove clips from the database. Once a clip has been linked, it becomes an integral part of a large web. Breaking one link means that you have to rebuild all the associated links. This does not allow for very flexible authoring.

4.3.4 Lessons Learned

A few things were learned from the *Train of Thought* experiment. In attempting to achieve narrative and visual structure to aid the viewer's story building activity, I achieve some visual structure but little narrative structure. The viewer can easily make sense of the screen interface once it has been explained. The positional editing

of the narrative elements aids in the story understanding. Unfortunately, the story isn't always orchestrated in a manner that conforms to the viewer's expectations. The original aim was to create an impressionistic collage of video vignettes in the belief that the viewer would make the necessary connections to construct the story. This did not happen and points to the need for a more structured approach to the orchestration of the narrative elements. The nature of the interaction is also a little problematic. The viewer is expected to choose between *dream* and *reality* without being given any reason to do so. The outcome of the story may be affected by the choices made by the viewer, but this is not reflected in the playout. There has to be more of a reason to intervene in a story. Some of the problems that surfaced as result of *Train of Thought* are addressed in the second version of Jack and Nicole's story, *Train of Thought v 2.0*.

5. Story Models and Story Agents - Train of Thought v 2.0

This chapter analyses the second iteration of *Train of Thought*, and describes how a story model working in conjunction with a story agent can give us the structure we need to orchestrate multi-threaded narratives. The story model creates the overall story structure, while the story agent navigates this structure, keeping track of where we are in the story, which thread is being followed, and what characters have been introduced. Interaction is made possible through an inference mechanism that brings forth related material that the viewer can choose to watch. Positional editing combined with text annotations creates a more dynamic interface. In *Train of Thought v 2.0*, the experience plays out automatically, but the viewer now has the option of interacting at any point during the experience.

5.1. Concept

5.1.1 A linked story model and a filter-based story agent

Based on what I learned with *Train of Thought*, I went back to the drawing board, and started *Train of Thought v 2.0*, taking into account the successes and failures of the first version. The impressionistic collage of vignettes of Jack and Nicole, originally inspired by Ricky Leacock's *Oeuf à la Coq* and Buñuel's *Un Chien Andalou* did not present the viewer with enough narrative structure to build a story. The premise that a movie might be analyzed and reconstructed in a manner similar to a dream did not apply to the kind of material arranged into a story as a linked hypertext structure. A different approach is needed. A good question to ask is: what does a viewer expect when watching a movie? We can then use this as a starting point for structuring a multi-threaded movie.

When talking about the filmmaker Howard Hawks, Gerald Mast framed the problem of telling an engaging story in the following manner:

The narrative that is insufficiently spontaneous and surprising is familiarly condemned as contrived, overplotted, unnatural, and stilted; the narrative that is insufficiently patterned is familiarly condemned as random, wandering, arbitrary, and formless. (Mast, 1985, p 567)

In this statement, we can see the formal, predictable structure of *An Endless Conversation* as insufficiently spontaneous and the loose, rambling structure of *Train of Thought* as insufficiently patterned. We need to find a middle ground that

will allow the encoding of a narrative structure that is both spontaneous and patterned.

5.1.2 Thoughts on Structure

Discussions with Janet Murray set me on the path of thinking about what the constituent parts of a story might be and how it might be possible to arrange them. The vignettes that constitute *Train of Thought* tell different parts of the story between Jack and Nicole. The more engaging narratives produced by the link structure usually followed a coherent thread through the material. One of the possible threads follows:

- Jack is dreaming of Nicole
- he talks to her about moving into her apartment
- he picks her up at her house and drops her off at the train station
- she leaves by train
- Jack wakes up from a nightmare.

This little narrative follows a nightmare thread of Jack dreaming of having to leave Nicole. Even with the impressionistic collection of vignettes in the database, this more or less coherent narrative came together. This sequence is more satisfying than some of the other more wandering sequences. The viewer can quite easily build a story as the thread progresses; there is narrative closure as the story starts and ends with a dream. The question is: how can we reproduce this kind of structure regularly? As Glorianna Davenport put it, “Many systems can produce engaging stories some of the time; we want a system that produces engaging stories most of the time.”

5.1.3 Story Template

Early sketches pointed to a scheme that encodes more defined narrative threads through the material stored in the database. As a result of discussions with Janet, I decided to break the story template into distinctive story elements:

- introduction
- conflict
- flight
- discovery
- ending

This template creates a simple structure that lays out a story in a consistent manner. It always starts with an *opening* scene, introducing the situation and the main characters. A *conflict* follows; most narratives will introduce some kind of conflict early on in the story. The *flight* element expands the story slightly as the characters

avoid the conflict. *Discovery* follows the final resolution of the conflict. The *ending* tries to put a closure on the narrative.

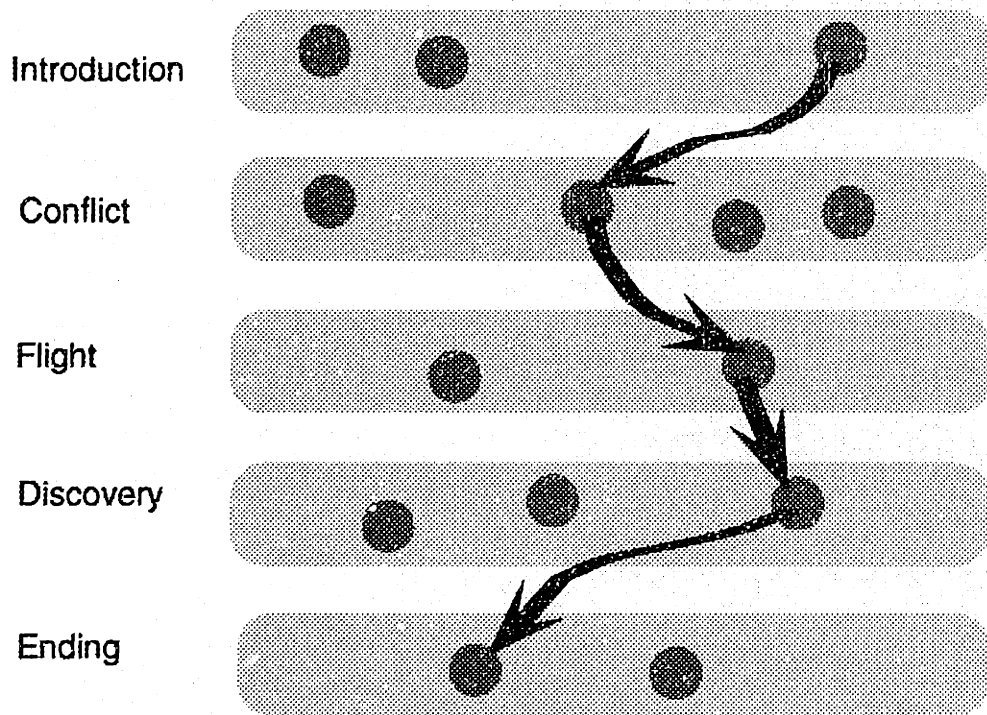


Figure 9 - A story structure. The story-units are organized to fit into distinct sections of the structure. The story-agent picks the path through the material.

Figure 9 shows what the layout of a story structure looks like. The story always starts with an *introduction*, follows a *conflict*, *flight*, *discovery* path and leads to an *ending*. The actual path followed by the system is determined by what I call the *story-agent* (to be described more in-depth later in this chapter). The story-agent carries out a function analogous to the layered filters of *An Endless Conversation*: It keeps track of the current thread being followed, the characters in a scene, and any other element chosen by the maker. It filters out inappropriate clips at each stage of the story; it allows only clips which follow a particular thread, with characters that have already been introduced, to appear on-screen. This orchestration, based on a formal story structure, goes back toward the more rigid structure explored through the DMO in *An Endless Conversation*.

5.1.4 Quantifiable Structure Exists in Films

According to Gerald Mast, Howard Hawks builds every story in an identical four part structure: a prolog establishing the conflict, development of the conflict from

one character's point of view then the other's, resolution of the conflict, and occasionally a brief epilogue to return the narrative full circle to its beginning. This structure "gives a Hawks story the firmness of shape, the elegance, economy, and symmetry that allow surprising events to transpire within the logic and structure of a controlled pattern." (Mast, 1985). Hawk's structured approach lends credence to a more rigid story model template.

A controlled pattern is very important. Computers are very good at recognizing patterns. In the scheme that grew out of *An Endless Conversation* patterns are encoded in a layered filter structure. The story sequence is represented in the form of a template: question, response, rebuttal. Filters then take care of making sure that if Dave asks a question about the light, then Thomas answers the question about the light. In *Train of Thought*, links build up the patterns that allow multiple paths through the material. We need to combine the rigidity of the conversation template with the unpredictability of the links.

When writing about Aristotle's Poetics, Hardison put things clearly with regard to our quest for a multi-threaded narrative structure:

The author can arrange the incidents in a story in a great many ways. He can treat some in detail and barely mention or even omit others [...] He can observe chronological sequence, he can distort it, he can use messengers or flashbacks, and so forth. Each arrangement produces a different plot, and a great many plots can be made from the same story.
(Hardison, 1968, p 43)

The story of Jack and Nicole is not an epic in the Aristotelian tradition, but the main point to draw from Hardison is that one story can be told through many plots. What I would like to do is to get the computer to orchestrate the sequence of incidents so as to build up different plots all relating to the same story. The story template described above lays down the foundation for creating a such structured narratives.

5.2. Production

5.2.1 Software

The story structure is encoded using Framer¹⁵. In the first version of *Train of Thought*, Framer was used to create a linked web of vignettes. The content and the links are attached to each other and create a structure that is difficult to construct and

¹⁵ See (Haase, 1992).

to modify. We now have a better idea of the various elements that are necessary for building a story and can therefore more easily encode them.

5.2.2 Data Structures

Framer is used to build a data structure that encodes the story-template described above: opening, conflict, flight, discovery, closing. In Framer, a frame is the basic building block. A frame can have any other frame(s) as annotations; such a frame is said to be within the other frame and inherits its features. The story-structure is encoded such that each frame corresponds to a story element. I want to encode the story in a flexible way so that more complexity can easily be introduced to the story. To achieve this, I break the structure into nested frames, with each nested frame inheriting the features of its parent frame. The top level story frame has three children frames: *introduction*, *body* and *conclusion*. These frames can be broken down further. Body has three children: *conflict*, *flight* and *discovery*. The conclusion frame has two elements: *closing-scene* and *credits*. Thus *conflict* can easily be seen to be in the *body* of the story. Using this approach, I can create a story template that can be expanded indefinitely to accommodate more detail or more depth. Thus if I later wish to break the *introduction* down into *Titles*, *Intro-character1* and *Intro-Character2*, I can do this easily by nesting the appropriate frames in the *introduction* frame.

Figure 10 shows the basic story structure with each story-element corresponding to a frame. Every part of the story is within the main *story frame*. Lower level story elements (frames), such as *conflict* or *credits*, are nested within higher level story elements (frames), such as *body* or *conclusion*.

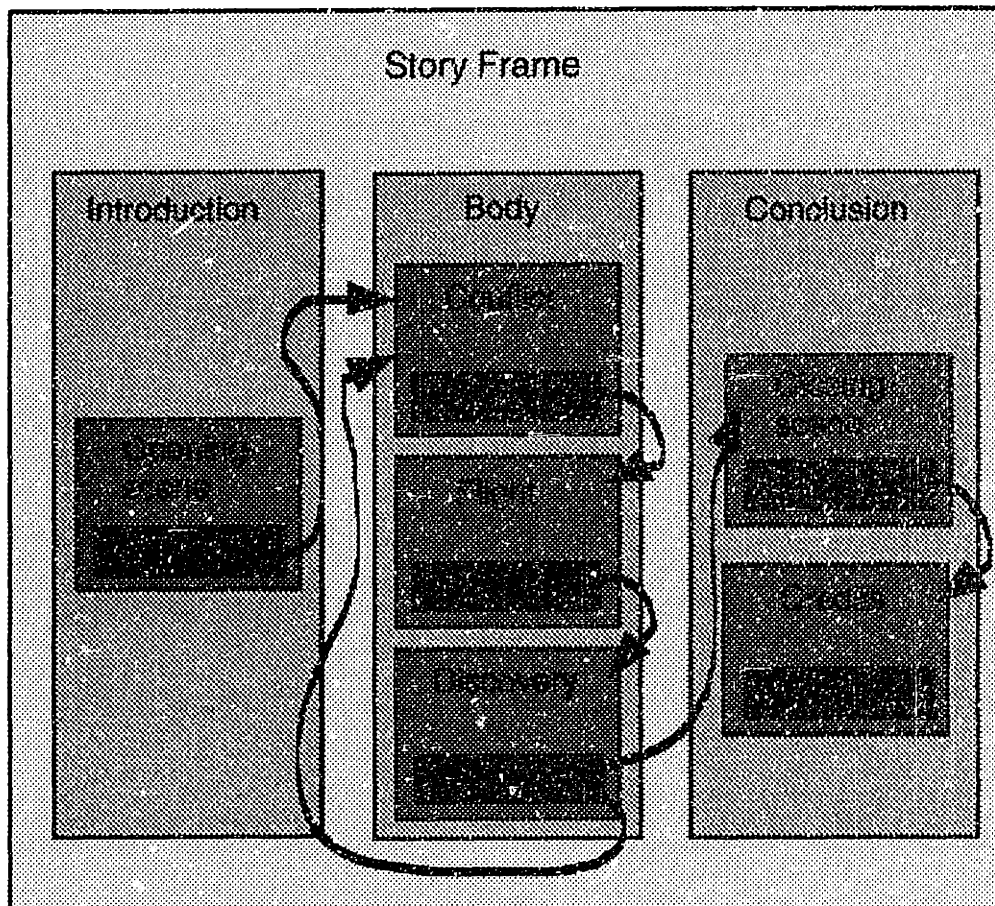


Figure 10 - The story-structure is organized using Framer. Each frame corresponds to one story-element. The path through the structure is determined by the *next-element* pointer.

elements, such as *body* or *conclusion*. It can clearly be seen that the structure is flexible enough for me to easily increase the complexity of any story element.

Once the structure is built, we need to figure out how to go through it to play out the story.

5.2.3 Sequencing the Story

The story has to be sequenced in a specific order. *Conflict* cannot come before *flight*, *credits* cannot come before *conflict*. Sequencing the story is accomplished by annotating each story-element frame with another frame that points to the next stage

in the story: *next-element*. The ground¹⁶ of these frames is represented by the arrows in the figure above. These pointer frames are analogous to links, except that they are now completely separate from the content and an integral part of the story structure. It is possible to have more than one pointer linking one frame in the story-structure to another. Thus, *discovery* points to both *conflict* (introducing a new conflict) and to *closing-scene* (ending the story). The story can go either way, depending upon how the system randomly decides to go.

5.2.4 The Story Agent

We now need to traverse the story keeping track of elements that are important to understanding the plot. In *An Endless Conversation*, this is done by the DMO, a layered filter structure in which each filter successively narrows down the list of available clips to only those that meet certain criteria (character: Dave, subject: Light). In *Train of Thought v2.0*, I develop what I call the story-agent to control the story payout. An agent is an autonomous entity with goals, a repertoire of actions and a means of sensing appropriate data. The story-agent navigates the story-structure, keeping track of the current point in the story (sensing data) and where to go next (the goal), and, through a filtering action similar to the DMO, decides which clips are appropriate at each stage of the story (the action).

Navigation is easy: the story agent looks at the current point in the story, checks where the *next-element* pointer leads to and follows that thread. If there is more than one pointer, it randomly picks one. This produces the name of the next story element that needs to be found in the database of vignette: *introduction*, *conflict*, etc...

The next stage is to pick the correct story-element based on what has been seen before. To do this, the agent keeps track of the following annotations: *story-element*, *thread*, and *characters*. During the logging process (described below), each clip in the database is annotated with the following attributes:

- *story-element* - what part of the story structure do the clip fits into?
- *thread* - which thread(s) is the clip appropriate for?
- *character* - what characters are in the scene?

The story-agent picks a clip with the correct *story-element* slot to make sure we get the right part of the story (*conflict*, *flight*, etc...). The *thread* slot makes sure that

¹⁶ In Framer speak, the "ground" of a frame is a particular value associated with that frame. In this case, the ground is a pointer to another frame.

we follow one thread through the story at a time. The *character* slot keeps track of which characters have been introduced and makes sure that new characters are not introduced in the middle of a story. This last slot is less important to keep track of, but does matter when it comes to story playout; it would be confusing to introduce a new character in the *closing-scene* for example. These annotations were chosen for this particular story structure, but just as in *An Endless Conversation*, other annotations can be added to reflect different characteristics that need to be tracked for a particular story.

Once the current *story-element* has been determined, the story-agent picks the clip that meets all the criteria of the annotations as described above. Through a filtering process analogous to the one described for the DMO in Chapter 3, the story-agent checks for clips that are part of the current thread, with characters that have been introduced and discards clips that do not meet the appropriate criteria. Thus, orchestration in this instance can be seen to rely heavily on a formal linked story-structure; filtering is what keeps the narrative on the right track. We are taking the best features of links and of filters to orchestrate the playout.

5.2.5 Logging

Ryan Evans' logger, *LogBoy*, is used to log slot/value descriptions for vignettes in a manner very similar to that used in *An Endless Conversation*. As an example, let's analyze a vignette, *The Refrigerator Incident*. In this scene, Nicole scolds Jack for keeping sour milk in his refrigerator. The value of the *story-element* slot is *conflict* because a problem develops between Jack and Nicole. The *character* slot is *Jack & Nicole*; they are the only two characters in the scene. The *thread* is *refrigerator-incident*; there will also have to be a *flight* and a *discovery* to this *thread* if we want a coherent story. The story agent keeps track of these three descriptions to build the narrative sequence. Other descriptions can easily be added using the logger, should we want to filter according to other criteria.

5.2.6 Inference

The story model and story agent take care of producing the main thread through the material. An inference mechanism adds interaction by bringing up material related to what is currently being watched. The viewer can choose to view this material or decide to move on with the story.

As I mentioned above, we filmed the actors being interviewed in character about their feelings and thoughts on a particular situation. This interview material is not included in any narrative thread by the story-agent because it doesn't fit into the story-structure. It is interesting material that can add meaning to the story. The problem is to get the computer to bring up these sequences at the appropriate times.

This is achieved through the inferencing capabilities of ARLOtje, a representation language built into Framer. Once the story-agent has chosen the next clip to view in the linear story, ARLOtje makes it possible for the computer to search for related material. All the vignettes have been annotated with a *type* slot. In *Train of Thought v2.0*, there are *reality*, *dream* and *interview* types. Other movies could have other types of related slots. By comparing the *type* slots and the *scene* slots, we can easily find *interview* clips for related scenes. For example, if the *Refrigerator Incident* is a kitchen *scene* and the clip is of *type* interview, then based on these simple descriptions, we know that the clip is an interview about the refrigerator incident in the kitchen.

The function is actually encoded as a special get-method in ARLOtje. Whenever the system tries to get the interview clips related to the current clip, the get-method runs a check function on all the clips in the database. If it finds a clip that is of *interview* type and is the same scene, it brings up the clip on-screen. The viewer can choose to temporarily leave the main thread of the story to view this clip, or can ignore the choice and move on with the story.

The implementation of this get-method to find related material in the database is flexible and simple. By writing a function to test whatever criteria the maker decides to check, it is possible to automatically bring up any kind of related material, not just interviews. I can imagine a movie-maker deciding to add all kinds of related material to add texture to the movie: a different point of view, an insight into the characters' motivations, or a clip that adds irony to the main thread.

The basic narrative structure is still linear, but it allows slight deviations from the main story. The viewer can choose to view this extra material and get a new perspective on the events. It can be thought of as an enhancement on the main events of the narrative.

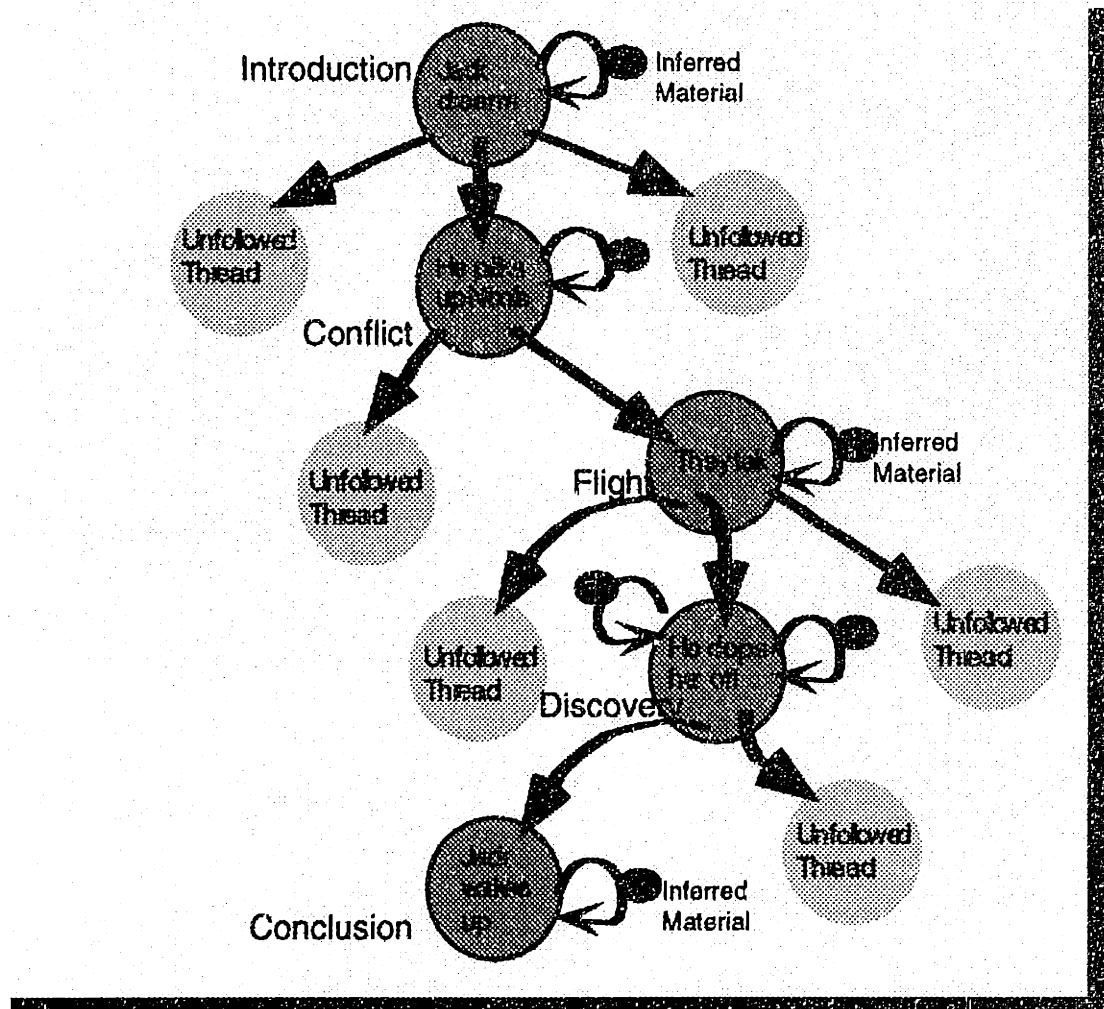


Figure 11 - This is what the story-structure looks like with the inferred material. The story agent selects the best path through the material and the inference mechanism finds related material. The viewer can choose to view the extra material or to go on.

Figure 11 shows what the story structure looks like with the inferred material taken into consideration. It can clearly be seen that the main thread of the story weaves through the collection of clips that fit into their respective parts of the story structure. Branching off this main threads are inferred clips that relate to what has been seen. This narrative structure is orchestrated on-the-fly based on a close relationship between the story model, the inference mechanism and the descriptions attached to the vignettes.

5.2.7 The script

This new narrative structure was tried out on the existing database of vignettes. Using LogBoy, some were described as *opening-scenes* (Jack in bed dreaming of Nicole) others were described as *conflict* (the refrigerator incident) and others were

described as *closing-scenes* (Nicole riding away by train). All the clips were described in this manner. The results are surprisingly good. The system produces much better stories. The sequences make narrative sense because the system is much more likely to follow a coherent thread through the material. It still does not produce engaging stories every time, because the database is not rich enough. For example, there is no resolution to the sour milk incident. The story-agent looks for another clip to resolve the story and the choice doesn't always make narrative sense.

I scripted and shot more material that is designed to fit into the story template structure. Instead of the impressionistic approach used in the previous version, for *Train of Thought v2.0*, each sequence was scripted to fit into a particular part of the story structure. There is a scene where Jack buys milk. This is an obvious resolution to the milk incident and falls in the *discovery* slot. Scenes were scripted so as to fill the various slots in the story structure.

I also added a new character, Beth. She is Nicole's roommate and best friend; they grew up together. Beth hates Jack and lets Nicole know this by trying to convince her to end the relationship. Beth, a materialistic, uptight lawyer, has her own problems and provides a different perspective within which to view Jack. She makes him more likable through her negative traits. There is a *conflict* scene where Beth tries to convince Nicole to drop Jack when she moves to London. Nicole points out that she loves Jack and that she is not about to leave him for any reason.

There is one *flight* scene, Jack cruising in his 1964 Ford Thunderbird. This is an escapist scene that could work with a number of threads. The basic aim when scripting these vignettes was to make them specific to a particular thread, but also to make them general enough to fit into other threads. For example, Beth talks on the phone to Michael, her boyfriend (whom we never see). This scene could follow a *conflict* where Beth berates Jack or it could come after a *discovery* scene where Jack and Nicole invite Beth to a diner, but she refuses, because she says she has too much work. It is important to get as much flexibility as possible out of the scenes; this creates a varied and flexible playout of the material.

The reality is that scripting scenes for multiple playouts is challenging. Even with a story such as Jack and Nicole's, where chronology and cause and effect are not as important as they might be in a tightly scripted story, there are some scenes that just

cannot fit in different parts of the story. For example, the *discovery* scene where Jack buys milk would be irrelevant if we never saw the refrigerator incident. Thus my aim when scripting the new scenes was to create a rich environment where the story is always the same (Nicole leaves Jack) but where many events leading up to the separation provide a variety of angles on the issues they are facing.

5.2.8 The Interface

The interface in the first version of *Train of Thought* had two confusing aspects to it. First, the two streams going off to the central vanishing point didn't always give the viewer a clear sense of how they relate to each other. Second, the interface was rigid in that it didn't allow the viewer to do anything beyond choosing to view a clip represented by an icon.

A major change with the new interface is that it has two viewing modes: interactive and non-interactive. The visual display is simpler and gives the viewer more options for interpreting the narrative. The nature of the interaction is changed because the viewer is no longer required to choose between two threads. The viewer can now choose to view more information about the story, adding to the interpretation of the narrative, instead of redirecting the narrative.

5.2.8.1 Two Modes

The system builds up the main thread of the story in real time. The viewer can just sit back and view the story in a linear fashion or the viewer can interact. A single check button, labeled "Interact" switches between the two modes.

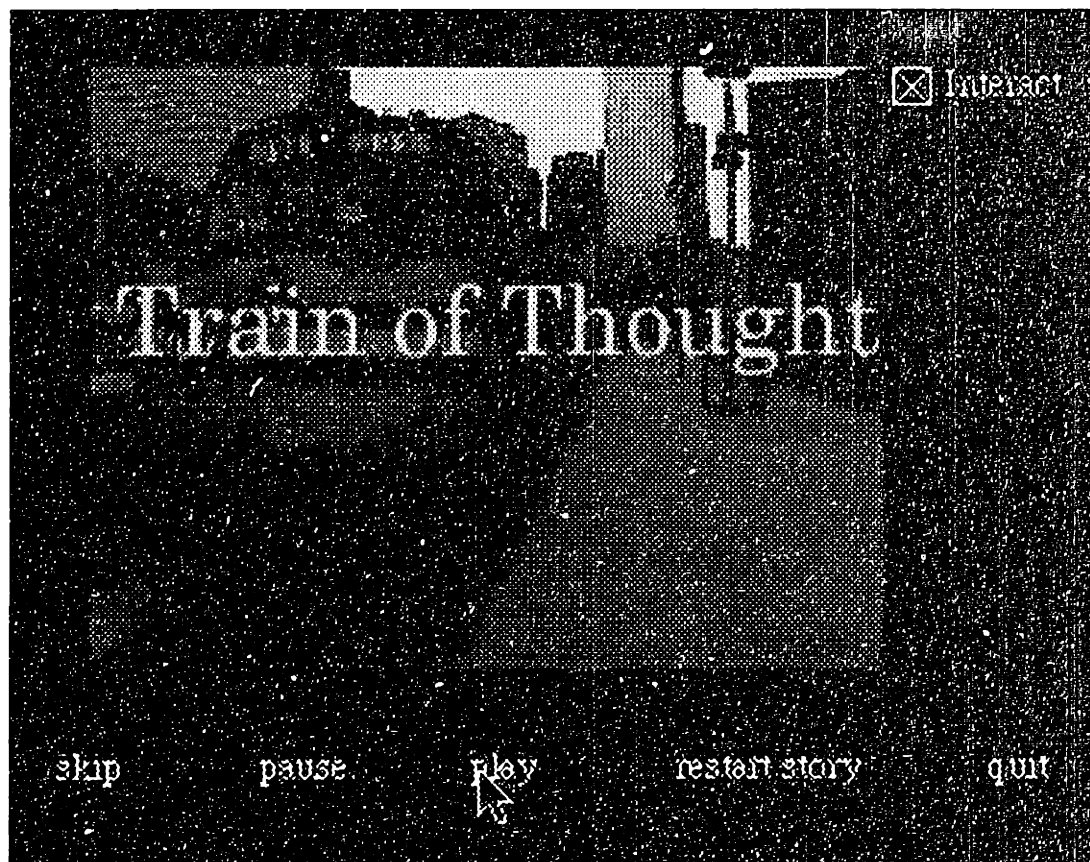


Figure 12 - In the linear mode, *Train of Thought* v2.0 fills the entire screen. The “skip” button allows the viewer to jump to the next story-unit. The “Interact” button switches modes.

In the linear mode, the frame of the moving image fills the entire screen¹⁷. Under the frame are five buttons: *skip*, *pause*, *play* and *restart-story* and *quit*. The viewer has limited control over the playout of the story. The *skip* button jumps the current vignette to the end and causes the next clip to be chosen by the story-agent. The *pause* and *play* buttons are self explanatory. Pressing *restart-story* resets the story agent and starts the story over again with the title sequence. Pressing *quit* quits the entire program. This linear mode is analogous to the experience of watching a video on a VCR. The main different is that because the story is broken up into distinct segments, the skip button allows the viewer to skip forward to the next section. This is like skipping forward on the tracks of a CD, except that what comes next is not fixed, but is computed by the story-agent.

¹⁷Doubling the image size is accomplished using the RasterOps 364 board which allows pixel doubling.

In the *interact* mode, it is possible to get more information on the narrative. Pressing the *Interact* button zooms the screen out to reveal the story structure on the left and the playout window on the right.



Figure 13 - In the "Interact" mode, the story structure becomes visible. The main narrative stream is in the upper left. The optional inferred material appears on-screen in the lower left, at a point along the time-line corresponding to the story-unit to which it relates. The main video plays out on the right side of the screen.

The figure above shows the layout in the interactive mode. The stream on the top left represents what has been seen so far. The positioning of the story units is similar to the previous version of *Train of Thought*: the icons get progressively smaller and shift toward the vanishing point as they go back in story time. The two icons on the lower left-hand quadrant are inferred material that relate to the story units represented by the icons directly above each one. There are two kinds of interactions.

5.2.82 Annotation Browsing

Passing the mouse over the stream brings up text annotations that allow the viewer to get more information about what has been seen. Associated with each vignette is

a text caption that gives a brief, poetic insight into the content of the clip. For example, the scene where Jack picks up Nicole to take her to the station says: "The day she left" or the scene where Jack drowns his sorrows in a bottle of whiskey says: "Jim Beam helps me out". These simple captions give the viewer a little insight into the content of a clip. It is also possible to look at the characters, the story element or the mood of the scene. The figure above displays the *story-elements*. This action of viewing text annotations can be thought of as browsing the descriptions associated with each clip. It is more of an analytical action, but can aid the viewer's story construction process by giving her the ability to see what is going on underneath the surface of the narrative. Seeing the *story-element* annotations gives a clearer sense of the structure of the story. This kind of insight becomes important as the viewer's role expands to include some making activities.

5.2.83 Inferred Material

The story-agent constructs the main thread to the narrative. The inference mechanism brings up associated clips. These clips are displayed on-screen on the second stream at the same point in the time-line as the vignette they are associated with. They move back toward the vanishing point as the story moves forward. These clips only become visible and available in the interactive mode. The viewer can choose to view them by clicking on the icon representing the clip. The short segment then plays in the main playout window. Viewing the extra material is a little like taking a detour from the main narrative thread and quickly coming back to the story. The story-structure is not affected by the interaction. This is a main difference with the previous interface, where the viewer actually chooses which thread to follow. This new interface changes the way the story is experienced.

5.3. Criticism

5.3.1 The Experience

Watching *Train of Thought v2.0* is a much more fluid experience. The scenes automatically play out in the non-interactive mode. They are almost full screen and the experience is analogous to a linear video experience. This is a very important feature; the viewer does not always want to intervene with what is happening on-screen. The system will play out a story automatically. The difference with a traditional linear experience is that the story is different each time it is watched. The story agent keeps track of what has been seen and makes sure that different threads are followed if at all possible. If the viewer decides to intervene and clicks on the *interact* button, the main playout window jumps back to reveal the story-structure

on the left side of the screen. In this mode, the viewer can choose to see extra material or to browse the annotations.

Viewing the interactive story is very different from sitting down to a linear film in two main ways. First, many different payouts mean that each viewing adds a different perspective to the story. The viewer doesn't get the whole story right away. Second, viewing is more of an analytical activity. The threads and connections have to be worked out by the viewer as the story progresses. This is very different from the experience of watching a normal linear movie. It is much closer to the experience of analyzing a dream, where things are not totally clear and where discussion and reflection leads to understanding. During one viewing session, the story might be seen two or three times; each time, a new thread adds to the understanding of the narrative. A traditional linear story has to pack all the relevant material into one main thread. The approach used in *Train of Thought v2.0* gives the maker the opportunity to explore new ways of presenting the material, where some clips are more important than others and are given prominence during the payout.

Computationally, the most important feature of *Train of Thought v2.0* is that it combines both links (from *Train of Thought*) and filtering (from *An Endless Conversation*) to achieve orchestration of the narrative. The links create the story structure independently of the content, and filtering (in the form of the story-agent) makes sure the right material appears on-screen based on where we are in the story and on what has been seen before.

6. Results - Insights on Content and Structure

This chapter reviews the four experiments and frames the results in theoretical terms. The story model and story agent orchestration, which combines a linked story model with a filtered playout, creates a narrative structure that is strikingly similar to the kernels and satellites proposed by Seymour Chatman. The approach fits into existing concepts of what constitutes a story. Positional editing starts to create a visual environment that displays what happens over time. It's use of perspective to give the viewer a conceptual idea of what has been seen so far is a first step toward creating an environment where multi-threaded stories can be visualized and understood. This chapter concludes with insights into what elements are necessary for making interaction an engaging experience.

6.1. A Brief Review

What do these experiments in digital cinema show? *IC Portrait*, *An Endless Conversation*, *Train of Thought* and *Train of Thought v2.0* are all short experiments, each one trying a different angle on the problem creating engaging interactive cinematic experiences.

Let us briefly review the four experiments:

- *IC Portrait* created a movie that used unstructured positional editing and orchestration. The experience was confusing to watch because characters and objects appeared randomly in various area of the screen.
- *An Endless Conversation* showed how a layered filter system can orchestrate clips into simple sequences. This is appealing because a familiar pattern is created: a back and forth conversation between two characters is a cinematic format familiar to viewers.
- The first version of *Train of Thought* used a linked story structure to build an impressionistic orchestration of video clips; there was not enough story structure for the viewer to grasp the narrative. Links attached to content are a difficult way to author a multi-threaded story.
- *Train of Thought v2.0* combines both links (from *Train of Thought*) and filtering (from *An Endless Conversation*) to achieve orchestration of the narrative. The resulting narrative structure meets certain expectations of the audience. There is a resolution to conflicts and the characters and their actions conform to expectations. This is a result of working with a story model.

6.2. Story Models

In the first incarnation of *Train of Thought*, the structure is attached to the content in the form of links that create multiple paths through the material. Changing one or the other is problematic. Adding or changing a piece of the story means breaking and recombining a large number of links. Visualizing the story formed by the links is impossible. It is a very inflexible approach.

The story model and the story agent in *Train of Thought v 2.0* keep track of the characters and events in the story. They create a framework into which the content can be dropped (with the appropriate descriptions attached). This approach has two advantages:

- The story structure is flexible. I can change the story model and keep the content intact. If I want a simple *conflict, resolution* structure for the middle of the story, I can make this change. I can also have a collection of story models for different stories. Thus far, I have been working with a more Aristotelian structure. It would be easy, however, to experiment with different narratives to suit different content. Some avant-guard twentieth century filmmakers such as Luis Buñuel and Jean-Luc Goddard have experimented with eliminating linear causality. Using different story models, it would be easy to adopt that kind of narrative structure.
- The content is completely independent of the narrative structure. In *Train of Thought v 2.0*, this means that I can add new content without affecting the structure in any way. I can add new conflicts and resolutions and new threads to the movie and watch them automatically be incorporated into the playout. This is a feature we first observed in *An Endless Conversation*, but it is equally applicable here.

6.2.1 Kernels and Satellites

The implementation of a machine orchestrated narrative in *Train of Thought v 2.0* points toward a promising direction for flexibly structuring interactive narratives. It also fits in with some existing theories of narrative structure. Seymour Chatman breaks a narrative into two kinds of events: kernels and satellites. "Narrative events have not only a logic of connection, but a logic of hierarchy. Some are more important than others. In the classical narrative, only major events are part of the chain or armature of contingency. ... Kernels cannot be deleted without destroying the narrative. ... A minor plot event - a satellite - is not crucial in this sense. It can be deleted without disturbing the logic of the plot, though its omission will, of

course, impoverish the narrative aesthetically" (Chatman, 1978, p 53). There are a number of similarities between kernels and satellites and the approach taken to create *Train of Thought v 2.0*. The narrative structure created through the use of a story model in my system corresponds to the *kernels* and the inferred links correspond to the *satellites*.

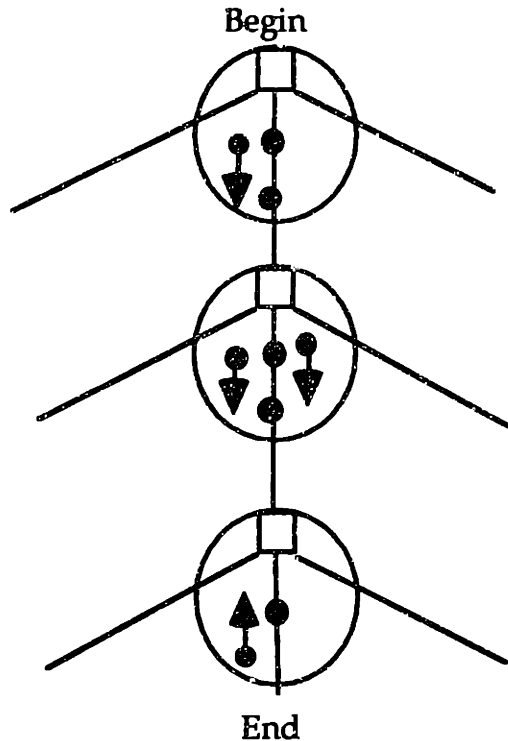


Figure 14 - This diagram is by Chatman and shows how he visualizes a narrative around kernels and satellites. There are many similarities between this diagram and Figure 11 in Chapter 5.

The diagram above from (Chatman, 1978, p 54), shows how Chatman visualizes a narrative structured around kernels and satellites. The kernels are the squares at the top of each circle. The circle is the complete narrative block. Kernels are connected by a vertical line to indicate the direction of the story-logic; oblique lines indicate the possible, but unfollowed narrative paths. Dots are satellites.

In *Train of Thought v 2.0*, a story model and story agent move the story along one focused thread from one *kernel* to the next. There is more than one possible thread, as demonstrated by the oblique lines. The kernels, corresponding to vignettes, have been written, filmed and edited to fit into a pre-defined narrative structure. The story model and a story agent sequence the vignettes according to how the model and the agent are programmed and how the content of the vignettes is described. This creates the main narrative thread.

The system automatically infers the *satellites*. At each main point (kernel) in the story, the system checks for related vignettes (satellites). The function of satellites is "that of filling in, elaborating, completing the kernel; they form the flesh on the skeleton" (Chatman, 1978, p 54). In *Train of Thought v 2.0*, these are usually in the form of interviews, but they could be any other related material. They form digressions from the main narrative thread in the form of interviews, although they could be any kind of supplemental material. They add to the understanding of the story, providing a different point of view or an insight into the thoughts of the characters.

The story structure resulting from the system I have implemented matches very closely the structure described by Seymour Chatman. The sequencing of the *kernels* and the search for *satellites* is computed on the fly. The format of the story produced by the system is predictable and fits into a more or less traditional narrative structure. The creation of a predictable narrative structure fits well into Bordwell's account of the viewer's activity: "In our culture, the perceiver of a narrative film comes armed and active to the task. She or he takes as a central goal the carving out of an intelligible story. To do this, the perceiver applies narrative schemata which define narrative events and unify them by principles of causality, time and space. Prototypical story components and the structural schema of the 'canonical story' ¹⁸ assist in this effort to organize the material presented" (Bordwell, 1985, p. 39). In other words, presenting the viewer with a familiar narrative structure helps them make sense of the narrative events. The story model and the story agent organize and assist this activity.

Some might argue that this approach is too rigid in its definition of what constitutes a story. It is important however, to start on familiar ground and to implement a story format that is known to be an accepted one. The main point to remember is that the story was flexibly implemented using Framer; a different story-model can be implemented from scratch or the existing story-model can be extended to accommodate a more complex story-structure. This research is a starting point from which to start implementing more complex story-construction mechanisms.

¹⁸ Bordwell's explanation of a canonical story-format: "Nearly all story-comprehension researchers agree that the most common template structure can be articulated as a 'canonical' story format, something like this: introduction of setting and characters - explanation of state of affairs - complicating action - ensuing events - outcome - ending." (Bordwell, 1985, p. 35).

6.3. Interface - Story Visualization through Positional Editing

The 2 ¹/₂ display explored in the *Train of Thought* movies is a first step toward giving the viewer a way to visualize an interactive narrative. In a linear narrative it is not necessary to think about how the story is displayed or played out. The viewer passively sits back and lets the story unfold uninterrupted and without worrying about where the story is or where it is going. The maker has done the entire job of assembling and sequencing the narrative material. In an interactive narrative, the viewer takes over some of the role of assembling and sequencing. Positional editing allows the maker to control the layout of the movie elements so as to aid the viewer's story building activity. There are two main modes to the experience: non-interactive and interactive.

6.3.1 Non-interactive

In the non-interactive mode, the story plays out almost full screen, as if from a video tape. The viewer experiences the main thread of the story without interruption. This is an important feature. At no point does the experience stop and wait for user intervention. This is essential for the creation of a flowing experience on-screen.

Should the viewer wish to intervene in the flow, an important control is the *skip* button, which is a simple form of interaction. It allows the viewer to jump directly to the next *kernel* in the story, skipping over a section of the scene currently being played out. This control is analogous to the *skip* control available on CD players; it allows a way of intervening in a story that was not possible before.

The viewer can jump to the main interactive mode at any point during the viewing, using the second important control, the "Interact" button. It is in this mode that positional editing becomes important.

6.3.2 Interactive

The main function of the 2 ¹/₂D interface is to give the viewer a sense of place and time in the story by positioning the story units in a pre-defined manner. The interface clearly displays what has been shown in the past. If something has been skipped over, it is possible to review that part of the story. It is also in this mode that related material can be accessed by the viewer. An interview clip found by the inference mechanism (a *satellite*) becomes visible at the same point on the 2 ¹/₂D

time-line as the *kernel*. In this manner, the viewer can easily place the individual elements of a story in the context of the entire story just by looking at the screen. It is easy to choose to review a particular part of the narrative (kernel) or to view a related part (satellite).

The main function of the display is to give the viewer more information about the story structure and content. Anything that gives the viewer an understanding of what is happening on-screen over time is important. The positional editing explored in *Train of Thought* is a first step toward taking the display of multi-threaded narratives out of the flat window-based interface and into a more dynamic environment that uses perspective to give the viewer a sense of time and place in the story.

6.3.3 Design

A noticeable feature of the 2^{1/2}D interface is that it does not have any grounding in a familiar metaphor. It does not look or feel like anything in the real world. At first glance, this may appear to be a problem, because "the theory is that, if the interface presents representations of real-world objects, people will naturally know what to do with them" (Laurel, 1991, p 128). This is certainly true of folders on a desktop except that the folders appear to have magical properties; for example, you can drop a file into them by dragging it over them. "The problem with interface metaphors, as illustrated by folders and other aspects of the desktop metaphor, is that they are like reality only different. Why should this matter? Because we don't know precisely how they are different" (Laurel, 1991, p. 129). The story interface I have designed presents another dilemma: I am trying to design an interface for a structure (an interactive story) that doesn't really exist in the physical world. Books and movies have evolved over the years to effectively display their content and linear structure. A multi-threaded interactive narrative is a structure that has only recently come into being. It is a major challenge to effectively display its content and structure.

"[Ted] Nelson offers an attractive alternative to 'metaphorics' in what he calls 'virtuality'. The design of a virtuality is driven not by its likeness to real-world phenomena, but purely by *conceptual structure* and *feel* [emphasis Laurel]" (Laurel, 1991, p 132). In my interactive story interface, I am trying to represent what is happening over time and space. What is the most effective way to give the viewer a sense of time and place in the story? Using perspective, I give the viewer a

conceptual idea of what has been seen so far and of related material that can be seen. By looking at the display, it is possible to get a *feel* for the general structure of the story. The 2¹/₂D interface is a first step toward solving the complex task of allowing the mind to visualize a complex multi-threaded narrative.

6.4. The Interactive Experience

The bulk of the work described in this thesis, as described above, has been geared toward computational orchestration and display of fictional narratives. The system, even in its crude implementation, does actually build a narrative on the fly from a collection of material stored and described in a database. It builds a flowing narrative experience on-screen. What has been learned from the implementation of the system and how close does it come to creating an engaging interactive experience?

6.4.1 Exploration in Fictional Narratives

The experience of watching *Train of Thought v2.0* is one of exploration. The viewer sits back and follows the main narrative threads created by the system. The story follows the lives of two characters, Jack and Nicole, as they live through a time in their relationship that tests the depth of their love for each other. The story is of a style that David Bordwell would call art-cinema.

In the name of verisimilitude, the tight causality of classical Hollywood construction is replaced by a more tenuous linking of events. [...] The viewer must therefore tolerate more permanent causal gaps than would be normal in a classical film. [...] the prototypical characters of the art cinema tend to lack clear-cut traits, motives and goals. Protagonists may act inconsistently, or they may question themselves about their purposes. [...] If the Hollywood protagonist speeds toward the target, the art-film protagonist is presented as sliding passively from one situation to another. [...] which is to say that 'inquiry into character' becomes not only the prime thematic material but a central source of expectation, curiosity, suspense, and surprise (Bordwell, 1985, p. 206-207).

The art-cinema structure of *Train of Thought v2.0* allows me as a movie-maker to explore the two characters, their motivations and their actions and to allow the viewer to do the same thing. The interaction supports this exploratory viewing mode.

Repeated viewings take the viewer through different threads of the narrative. They all relate to the fact that Nicole is moving away from Jack to follow her career. Clicking on the 'Interact' button allows the viewer to interrupt the narrative flow and to view more information on the events. This is viewing as exploration, not as engagement. It is a third person experience, where the viewer does not have an impact on the action, but instead is able to explore the actions and motivations of the characters. It is a small step beyond the traditional linear narrative where the viewer sits back and has absolutely no control over the events on-screen. In the 'Interact' mode, viewing means choosing what to watch - whether to just follow the main thread, to backtrack and review a scene or to delve deeper into some of the extra (satellite) material. That is the extent of the interaction in the system as it is currently implemented. The experience is one of “browsing” the story and of exploring more narrative material at certain points in the story.

Is it an engaging experience? Not yet. Yes, the system does give the viewer the option of exploring and viewing the story in different ways, but it still does not motivate the viewer to get into the characters' minds and really involved in the narrative.

An important insight to come out of the research is the issue of iterative viewing. At the beginning of this thesis, I emphasized the importance of the iterative process of making movies. The maker scripts and re-scripts, shoots and re-shoots, edits and re-edits. All this is done in the name of one seamless linear viewing experience, where the entire story is built up in the viewer's mind in one sitting. The viewing experience to come out of *Train of Thought v2.0* suggests that viewing might also be an iterative experience; the experience encourages multiple viewings; the story-agent builds different threads through the material each time. Some people like to re-read books, and see movies more than once. This experience can be enhanced if the payout is different each time, as in *Train of Thought v2.0*.

6.4.2 Exploration in Documentaries

The type of interaction described in this thesis has been one of exploration; the viewer can browse extra information related to what they are currently watching. As implemented in *Train of Thought v2.0*, viewing extra material gives the viewer the opportunity to get a different perspective on the narrative. The basic motivation for this kind of interaction is the search for extra information. In a fictional story, this is sometimes difficult to justify, because the reverie of the experience is broken as

soon as the viewer decides to interact. Documentary material is a better candidate for browsing interaction. The viewer is usually watching the movie to learn about a certain topic (as well as to be entertained) and interaction allows him to choose what thread to follow or what extra material to view. Browsing for extra information fits in well with the documentary mode of viewing.

6.4.3 Thoughts on Dreams, Emotions and Interaction

What would make interaction really engaging and give it meaning? Based on the experiments described in this thesis, I have a few thoughts on what might make interaction work in the context of fictional narratives.

A promising direction is to think of the relationship of interaction to dreams. Dreams are experienced in a linear format. They are often wild and semi-structured narratives that flow in an uncontrolled fashion. The interesting moment comes with the analysis of dreams. It often means going back over the action, reviewing the significance of the events and trying to make connections between them. Can an interactive experience be structured in this way? If the characters and their actions are intriguing enough, and the material is written so that the story can bear analysis, interaction might become motivated. If it is possible to get into the characters' mind and most private thoughts, and to have those thoughts bear significantly upon the understanding of the story, interaction might be motivated.

An important feature of dreams is that they are often analyzed with respect to one's own inner thoughts, feelings and emotions. In probing them, one often wishes to find out something about one's self - hidden fears, emotions and feelings. It is this involvement of the self that is probably most important in the act of interacting. Traditional movies are made for spectators. "The narrative film is so made as to encourage the spectator to execute story-constructing activities. The film presents cues, patterns, and gaps that shape the viewer's application of schemata and the testing of hypotheses" (Bordwell, 1985, p. 32). Yet the act of intervening in what happens on-screen means that the viewer is no longer a spectator. This means that an interactive movie should be made so as to encourage the interactor to execute self-reflexive activities on top of story-constructing activities. It is the involvement of personal thoughts, fears and emotions that will make the viewer care enough to want to interact.

This means that the maker will have to concentrate less on purely *manipulating* the audience's emotions as is the case now in linear movies. The maker should instead think about *involving* the audience's emotions in a way that makes them feel a certain responsibility for the actions and events happening on-screen.

Therefore, in the fictional domain, engaging interaction will probably have to wait until the viewer can become involved in the first person. Only when the viewer actually becomes a character and impacts upon the action will the self truly be involved in a self-reflexive way. When the viewer can interact and be made to feel responsible for the decisions made, where causing a death means loss, or where having a relationship means feeling love and hate, will interaction be truly meaningful.

7. The Future

This chapter looks a few years into the future when movies will be delivered over a network. How might their content and structure change using some of the ideas discussed in this thesis? Movies could be broken up into sections that are put together on the fly according to characteristics set by the viewer. We might also have the cinema equivalent of Network News: cinematic spaces where viewers contribute content which is orchestrated for playout. Also discussed are potential research directions for story structures and interfaces.

7.1. New kinds of distribution, new kinds of stories

7.1.1 Broadcasting

Movies are made to be watched by an audience. Any discussion of movies without a mention of the place and manner in which they are watched would be incomplete.

Feature films are worked on by many people for months and sometimes years, but there comes a time when the iterative process of production comes to an end. The movie is handed off for viewing. The movie maker will have put together what he or she considers the final version of the movie¹⁹, which is then sent out for viewing by various means. For theatrical release, the film reels are physically shipped off to the theaters. They are projected in a specified order and the film is viewed exactly as it was cut by the director and editor. For release on television, the movie is broadcast over the airwaves to an audience sitting in front of a television. The movie is sent to many households at the same time.

This distribution scheme is a direct result of distribution technology. With television, a central transmitter sends out a signal that can be received over a large area. The signal is identical for all viewers and the receiver does no more than decode the incoming signal. Whether the movie is sent out by truck, over the airwaves or over cable, the signal received in the home is the same for everyone.

7.1.2 Networking

Digital technology allows a completely different kind of distribution scheme.

Instead of everyone receiving the same signal at the same time, it is possible for one

¹⁹ On rare occasions, a movie maker will distribute a different cut of a film, sometimes called a director's cut. This happens if there was a disagreement between the studio and the director over the final cut of the movie.

person to get her own information at any time. Switched packet network technology allows individual bits of information to be sent to a specific receiver. These bits can be text, sound or images. The implications of this kind of distribution are far reaching. Instead of everyone receiving information from the same source at the same time, anyone can receive their own information at any time. Movies are just another type of information that fits into a networked distribution system. The movie data will be accessible from a television/computer sitting in the home. What could a movie designed for such an environment look like?

In some cases, they will be no different from existing films. "Movies on demand" is a concept that has been discussed for quite a while now. People can immediately grasp the idea that any movie can be ordered and viewed over the network at any time. There is a huge library of films that the studios want to make available to a paying public in this way. Instead of going to a video store, the viewer will be able to order it directly over the network.

7.1.3 Personalizing

I am more interested in thinking about the new kinds of movies that could be produced using some of the ideas explored in this thesis and digital distribution technologies. The first thing that comes to my mind is best illustrated by the following imaginary request: "I want a movie about a love story that ends happily, with no violence, a little sex and I have 50 minutes to see it." This kind of request will be possible with a switched network distribution scheme and computational orchestration. The movie elements (vignettes) will be stored on a server and described using a sketchy description scheme; only the features deemed important by the maker will be logged. The computational orchestrator will actually put the movie elements together to build a full narrative, the content and structure being based on the specifications of the viewer and the content design of the maker. This is a personalizable movie, where the interaction happens before viewing starts and is geared toward watching a linear piece that fits the needs of a particular audience of one or more. The maker decides ahead of time which options are available and what latitude of interaction to allow. With real-time shot selection, it is conceivable that the viewer would be able to change certain parameters of the movie while it is playing out. One format particularly suited to this approach is soap operas; there are a large number of characters whose lives intersect and intertwine as the multiple plot lines unfold. This multi-threaded format lends itself to personalization, where the

viewer could specify which character to follow (or emphasize) and how long to watch.

This approach to movie making takes advantage of distribution and orchestration capabilities made possible by digital technologies. To makers accustomed to creating a movie with a single thread for only one audience, telling a story with multiple threads for more than one type of viewer will be a new experience. Some will reject the idea outright. Others will embrace the new options open to them and create movies that take advantage of the new capabilities.

7.1.4 Open Ended

One of the more exciting ideas to come out of this research is the idea that movies can become open ended entities where the content is separate from the structure. I can easily add scenes to *Train of Thought v 2.0* and as long as they have been described correctly, they will automatically be incorporated into the payout. A movie can automatically grow and evolve as content is added, changed or deleted. Using this approach, a movie might never be finished. It is possible to imagine a movie where the maker adds content on a regular basis and even changes the orchestration software to reflect different story structures. I like the idea of a documentary where the maker could explore an issue over a long period of time. She films and edits material into vignettes that fit into a pre-defined documentary structure, maybe along character lines, or issues. Viewing such a movie would be an exploratory experience, where the viewer can follow different characters or issues through the material. One important feature of such a movie distributed over a network is that the maker can have a server of material that she constantly adds to, and that the viewer retrieves from. It is even possible to imagine a two way interchange between maker and viewer via e-mail or fax. Viewers could suggest, comment and criticize the movie as it evolves over time. Another important feature of such an approach is that computation takes care of the overall (macro) structure and distribution, while the maker takes care of content and the micro-structure (vignettes). Once again, this kind of making is not for everyone. Some makers will dread the idea that the piece they are working on may never end. Others will take advantage of their new relationship to their audience.

7.1.5 Free Form Content in a Controlled Structure

One other idea for a networked film form is one where the structure is formed by the maker, but the content is contributed by the viewers. This scheme is analogous

to the Network News bulletin boards that exists today for text and pictures. Here, a format is provided by the system in the form of news groups that are broken down according to subject: comp.multimedia.mac, alt.autos.antique, clari.news.tech. The users exchange information, anecdotes, and advice within the structure created by the system. The groups are sometimes moderated by a person who controls what appears and what does not. Very often, they are self moderated by the users and contributors themselves. The resulting environment is an informative and free-flowing exchange of up-to-date information; the important feature is that it is unlike any form of information exchange that existed before it became technically possible to exchange information in this manner.

I would like to see these ideas carried over to the moving image. With computational orchestration and two way networks, it is possible to imagine a scenario where people contribute their own video material to a movie space. The movie clips are described according to a standard description format. These descriptions are then used by the system to retrieve the appropriate clips for a particular movie group. Rather than having these clips only play through user-browsing (as in Network News) I would have the system orchestrate the clips into a personalizable linear playout similar to the one described above. The power of this approach is that the resulting cinematic space is free to flow, evolve and change in a way limited only by the imagination of the hundreds and thousands of people who could contribute to such a system. What kind of narratives would emerge? I would not hazard a guess. But I would say that the cinematic experience that emerges from such a system will be as different from linear television, as Network News is from magazines.

7.2. Research directions

7.2.1 Scalable Story Structures

The story structures created for *Train of Thought v 2.0* are rigid. The system follows more or less the same path through the structure each time it is watched. The only variation is left to chance; sometimes, a second conflict is introduced once the first one has been resolved. The system randomly decides whether to go from *discovery*²⁰ to a *closing scene* (and the end of the movie) or to a new *conflict*. More interesting would be a story structure with depth, where a conflict could be made up of other conflicts and resolutions. Under this approach, it would be possible to

²⁰ See Chapter 5 for a complete description of these terms.

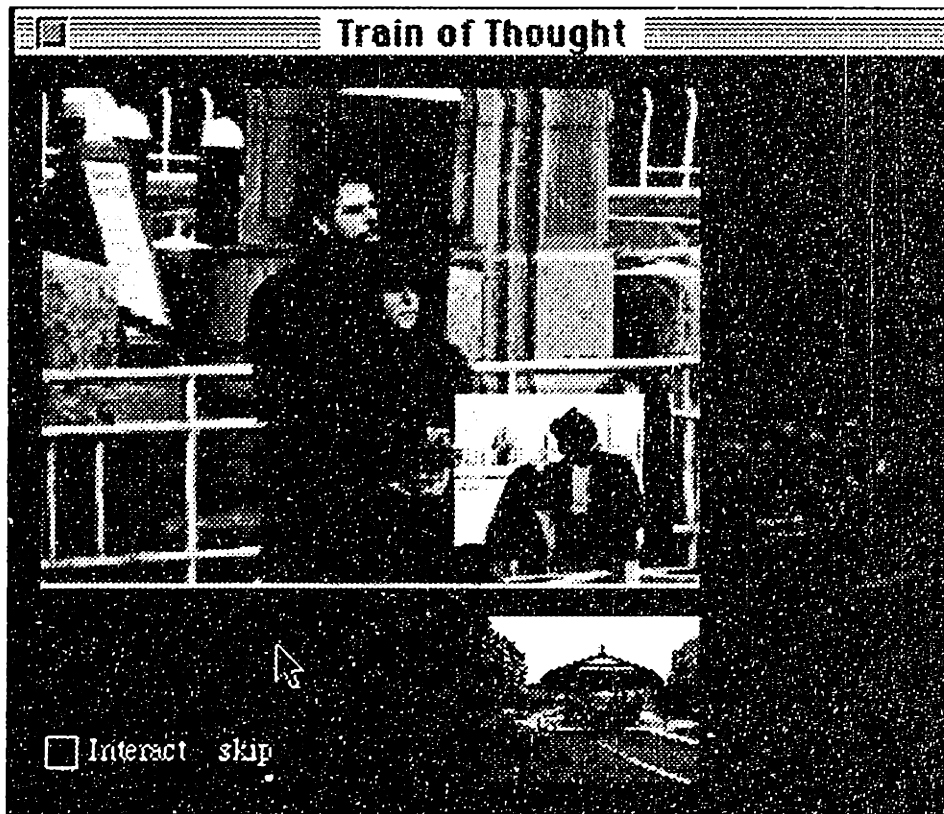


Figure 15 - Shows the positional editing of clips with the main playout window (the large frame) integrated into the inferred material (the smaller icons). The *Interview* clip appears as picture-in-picture, the *Dream* clip appears on the right and another clip appears below. Passing the mouse over any of the clips plays them. The strength of this interface is that it combines the interface into the movie through the positional editing of the cinematic elements themselves.

introduce different kinds of conflicts (minor and major) at various point in the story structure, not just at one specific point. It should then be possible to navigate the structure in a flexible manner, with the story agent deciding how to play out the story based on user preferences or other criteria (time of day, audience, amount of money paid). I would call this approach *scalable story structures*.

7.2.2 Interfaces

Interface design is a constantly evolving and changing process. The positional editing of the cinematic element in the *Train of Thought* interface has evolved as different schemes are implemented and used. Lately, I have been working with undergraduate researcher James Seo to evolve the interface one step further. We have integrated the playout window on the right side of the screen with the interactive stream on the left side to create a unified interface (Figure 15).

The main stream plays out in the large window and related material appears either as a picture-in-picture or as an attached frame. The nature of the related material is

indicated by its position; for example, *Interview* clips appear as picture-in-picture while *Dream* clips appear attached next to the main frame. The clips play out as the pointer is moved over them. In the interactive mode, as the story moves forward, clips seen in the past move off toward an imaginary vanishing point, as they did in previous versions of *Train of Thought*. Using this approach, the viewer immediately knows how different clips relate to each other by virtue of their position. A large amount of information about the current point in the story is conveyed to the viewer through the positional editing of the different clips.

One problem with the *Train of Thought* interface is that it is not very smooth in its operation. It moves jerkily from one story stage to the next. Movies convey a sense of motion and fluidity. An interface to cinematic story structures should do the same. *Train of Thought* is implemented in 2¹/₂D; it is not possible to "see" the story from the side, from behind or from a distance, to navigate it in a fluid manner. For the interface to really work, it should be implemented in true 3D. The viewer should be able to "fly around" the story in a smooth and flexible manner. Once such an interface is achieved, the viewer will be able to get a true sense of where he is in the story, by visualizing multiple threads and alternative ways through the material.

7.2.3 Interaction

Interaction has to leave the screen and happen in the head and body of the viewers. This does not mean virtual reality. Rather, the viewer should become part of the story by caring about the characters, the action and the story. This is a problem that will work itself out over time as makers and viewers explore their new roles in the digital realm. It was not immediately apparent during the early days of film that people would accept and understand the meaning of a "cut". It was also not clear to the maker what was meant by a "cut" from a wide shot of a person to a close-up of the same person for example. Similarly, it is through time and experimentation on the part of both the viewer and the maker that interaction will start to fit into the vocabulary of the new forms that will emerge through digital cinema.

8 Conclusion

I have dreamed of my film making itself as it goes
along under my gaze, like a painter's eternal fresh
canvas.

Robert Bresson

Making movies is an activity that involves the close collaboration of specialized machines and highly skilled people. The result of this collaboration is a stream of images and sounds that are brought to life as they play out over time. Traditional film technology has shaped movies into long linear threads where the content remains unchanged each time it is watched. Digital technology offers makers different ways of structuring and playing out images and sounds.

The research described in this thesis looks at two different ways in which the computer can assist the maker in the creation of multi-threaded stories: orchestration and positional editing. Four different experiments were reviewed, each one exhibiting different degrees of structure in the manner in which they are displayed and played out.

Positional editing becomes important when a story is no longer a single thread. The viewer needs to get a sense of where he is in the story and the 2¹/₂ display explored in *Train of Thought* starts to give a sense of what is happening over time as the story plays out. When the time comes to interact, the viewer has a way to visualize the current point in the story and the options available for interaction.

Orchestration demonstrates how the computer can open up exciting possibilities for creating simple narratives where the computer sequences many different threads through a collection of narrative material. Using a story-model, a story agent that knows how to navigate the model, and simple descriptions attached to video clips, *Train of Thought v2.0* demonstrates how computation can assist the maker's story-creation activities and enhance the viewer's story-construction activities.

The human tradition of story-telling is built upon thousands of years of experience, tradition and experimentation. The computer is just another tool that pushes that tradition into new and unexplored directions. For me, one of the most important concepts to come out of this work is that these thousands of years of tradition have to be respected and taken into account when creating for the digital medium. The

most interesting concepts emerged when I used the digital environment to build upon existing narrative theories such as Bordwell's or Chatman's. People are used to telling and listening to stories in very well defined ways. To stray too far from those ways is to risk losing the audience; that is the greatest fear of any movie maker. A very delicate balance holds the viewer glued to the screen. The balance is broken when the viewer is confronted with something that doesn't fit into the existing story-telling paradigm.

Interaction is a new paradigm that the digital environment has made possible. It has to be treated very carefully. I think that Robert Bresson was very insightful when he said: "The public does not know what it wants. Impose on it your decisions, your delights" (Bresson, 1975, p. 113). Interaction does not mean that the viewer always has to or wants to decide what to see or do next. The orchestration and positional editing methods described in this thesis mean that the viewer can interact if desired, but can just as easily sit back and "enjoy the show". This "optional interaction" approach offers a more flexible approach that takes into account the fact that the audience probably doesn't know exactly what it wants.

The work described in this thesis should be seen as a catalyst for further research into how digital technology can assist our story telling activities. The work is tightly focused on the specific problem of creating visual and narrative structure out of specifically shot content. It does not address such questions as: what does a movie with a first person participant viewer look like? what kind of new experimental narrative structures might be more amenable to a digital networked environment? or does the audience really want it all? If there is one thing I hope to have shown it is that coherent structure built upon understood concepts is the most likely path to the successful integration of computation and cinema. I have experimented with a few possible structures and hope that they point toward more interesting ones that can make digital cinema a fresh new art form.

References

- Bolt, Richard A. (1981). *Gaze Orchestrated Dynamic Windows*. Computer Graphics, 15(3), 109-119.
- Bordwell, David (1985). *Narration in the Fiction Film*. Madison, Wisconsin: The University of Wisconsin Press.
- Bresson, Robert (1975). *Notes on the Cinematographer* (Griffin, Jonathan, Trans.). London: Quartet Books Limited.
- Bruckman, Amy (1991) *The Electronic Scrapbook*. MS Thesis, MIT Media Lab.
- Calder-Marshall, Arthur (1970). *The Innocent Eye*. Baltimore: Penguin Books Inc.
- Chatman, Seymour (1978). *Story and Discourse - Narrative Structure in Fiction and Film*. Ithaca: Cornell University Press.
- Davenport, Glorianna (1987). *New Orleans in Transition. 1983-1986: The Interactive Delivery of a Cinematic Case Study*. In The International Congress for Design, Planning and Theory, Boston: Park Plaza Hotel.
- Davenport, Glorianna, & Brondmo, Hans Peter (1990). *Creating and Viewing the Elastic Charles - A Hypermedia Journal*. In R. McAleese & C. Green (Eds.), *Hypertext: State of the Art* (pp. 43-51). Oxford, England: Intellect Limited.
- Davenport, Glorianna, Evans, Ryan, & Halliday, Mark (1993). *Orchestrating Digital Micromovies*. Leonardo, 26(4).
- Davis, Marc (1993). *Media Streams: An Iconic Visual Language for Video Annotation* (To be presented at the 1993 IEEE/CS Symposium on Visual Languages, Bergen, Norway.). MIT Media Lab.
- Elliott, E. L. (1993) *Watch, Grab, Arrange, See. Thinking with Motion Images via Streams and Collages*. MSVS Thesis, MIT.
- Evans, Ryan (1993) *Log Boy meets Filter Girl*. MS Thesis, MIT.

- Haase, Ken (1992). *Framer*. In Cambridge, Massachusetts: MIT Media Lab.
- Halliday, Mark D. (1991). *Interactive Cinema Portrait: The Next Generation* (Work in progress report). MIT Media Lab.
- Hardison, Jr., O. B. (1968). *Aristotle's Poetics; a translation and commentary for students of literature*. Englewood Cliffs: Prentice-Hall.
- Kuenzli, R. E. (Ed.). (1987). *Dada and Surrealist Film*. New York: Willis Locker & Owens.
- Laurel, B. K. (1991). *Computers as Theatre*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- Linklater, Richard (1992). *Slacker*. New York: St. Martin's Press.
- Lippman, Andy (1980). *Movie Maps: An Application of the Optical Videodisc to Computer Graphics*. Siggraph.
- Mast, G., Cohen M. (Ed.). (1985). *Film Theory and Criticism* (Third ed.). New York: Oxford University Press.
- Morgenroth, Lee (1992) *Homer: A Story Model Generator*. B.S. Thesis, MIT Media Lab.
- Rubin, Benjamin (1989) *Constraint-Based Cinematic Editing*. MS Thesis, MIT Media Lab.
- Smith, Thomas A. (1992) *If You Could See I Mean...* MS Thesis, MIT.
- Youngblood, G (1970). *Expanded Cinema*. New York: E. P. Dutton & Co., Inc.

Appendix A - Digital Guerrilla Movie making

Slacker was made exactly opposite of the way you are suppose to do things. There was no real official-looking script, we used mostly non-professional actors, we never got permits to shoot on city property, no one got paid, etc. If you want to make movies that are different, I think it's okay to make them differently. A lot of independents limit themselves by imposing a studio structure of making a film onto a small budget production.

Richard Linklater on *Slacker*

In the following pages, I describe how new technologies such as Hi-8 cameras and desktop computers influenced the way I made movies during my two years at the Media Lab. I hope this will provide some useful insight into the process of producing movies with high technology and limited resources.

Making movies is a collaborative effort. The skills of many people go into making things look perfect on-screen. The film work described in this thesis was carried out with the collaboration of many people. Their roles and contributions are also included in the following pages.

Movies are made with machines. In the old days, filmmakers shot, developed, printed and edited their films, basically doing everything from scratch. Flaherty brought a camera, a projector and a printing machine with him to Port Harrison in Northern Canada, when filming *Nanook of the North* in 1920. After developing the film on his own, he projected the footage he had just shot for the community in which he was filming. "Nanook was to be a film of the Innuity by the Innuity, 'of the people by the people' insofar as that was possible." (Calder-Marshall, 1970, p. 82) Flaherty showed how technology could involve both maker and viewer in a shared experience that might be described as the ultimate interactive cinematic experience. That was in 1920. Miniaturization and digital technology is pushing the movies back in that direction. Lightweight, inexpensive cameras, desktop editing and, in the near future, network distribution, are making the production of movies more accessible to people who might never have dreamed of editing and distributing the video they shoot. This might be the biggest contribution that digital technology is making to the art of movie making.

On Guerrilla Filmmaking

Writing

One of the most important aspects of making narrative movies is writing. Aside from the obvious prerequisites of having a strong main character with which the audience can identify, a story that draws the viewer into a state of reverie and a minimum amount of action, you have to write material that can realistically be shot on the given budget and schedule: a couple of hundred dollars and a long weekend. Don't write lines like: "And Boston burned to the ground" or "The car exploded in a large ball of flames". The art of low-budget filmmaking rests upon making the most of very limited resources. Action scenes and difficult locations wreak havoc on low budget productions.

When it comes to dialog, unless you are willing to devote considerable time and effort, stick to creating situations within which the actors can improvise. That is exactly what I did for *Train of Thought*. It worked out very well. Writing lines that will sound good when delivered and that add up to a convincing story is an art that takes years to learn. Realistic dialog can fall out of good improvisation.

A good way to get convincing dramatic material is to write about what you know. *Train of Thought* was very loosely based on personal experiences. If you are close to



Name: David Lyn Kung
Born: 3/24/71
Role: Actor/Stunt Man in *An Endless Conversation*
Day Job: Grad Student
Real Job: Backup Singer for Sonny Bono
Movie quote:
"If you have any dignity, you will apologize to the rice right now!"
-- Chow Yun-Fat "A Better Tomorrow II"
Formative experience:
Attending Illinois Math & Science Academy
First memory: Wearing Footies
Last dream:
Listening to Answering Machine Message from Brother
Favorite movie:
Something Wild/Broadcast News/The Killer/Return of the Jedi
Last book read: Geek Love
Ambition: Be Rich
Recent accomplishment:
Graduated; Developed Callus from Playing Video Games
Fetish: Eyebrows and Backs
Most annoying shooting experience: Mark Halliday's Bad Breath; Shooting in the Rain and having steam rise from my actor's legs
Role model:
David Addison (Moonlighting), Chow Yun-Fat
Favorite director:
Jonathan Demme, John Woo
Drink: Coca Cola Classic
Last car driven: Nissan Maxima
Favorite (Internet) hangout:
a.t.music.alternative

the material, you will understand it that much better.

The script I worked with looked as follows: for each scene, there was one paragraph describing location, one for action and one for dialog. Each scene fit onto one piece of paper. This approach reflected the modular nature of the story-structure; different scenes could fit into various parts of the story. This approach did not make it easy for the actors, but made for good improvisational opportunities.

Actors

Good improvisation relies upon good actors. I used the Internet to find Ian Dowell and Sharon Cinnamon, the two main actors of *Train of Thought*. At MIT, the mailing list of interest is theatre@mit.edu. This list goes to all those affiliated with the MIT community who have an interest in theater. It is fast and effective and the only thing missing is the ability to easily get pictures across the network. Ian and Sharon's improvisational skills created very realistic looking scenes.

I gave them the basic situation with pointers about what I wanted to get across in a scene and let them take it from there.

Shooting

Guerrilla shooting means getting the most out of limited resources. When shooting *Train of Thought*, we didn't have time to get location permits from Amtrak or the Children's Museum; we didn't have the money to rent Betacam cameras; we didn't have the luxury of rehearsals.



Name: Ian Anton Dowell
Born: Keene, NH, 1964
Role: Jack
Day Job: programming educational software
Real Job: Misummerfest
Movie quote: "she turned me into a newt"
Formative experience: dropped on head from a bicycle by father
First memory: dropping out of kindergarten because the guy behind me in line threw up on me at recess
Last dream: four-wheel-drive vehicles were wrecking havoc on my woodland property
Favorite movie: Monty Python & the Holy Grail
Last book read: Winnie the Pooh
Ambition: find a good woman
Recent accomplishment: rigged up a hot shower which runs entirely off of natural power
Fetish: Macintosh computer
Most annoying shooting experience: the doughnuts got soggy
Favorite director: Mark Halliday, of course
Drink: chocolate frappe
Last car driven: 1986 Hyundai Excel
Favorite Internet hangout: National Weather Service

We shot everything in Hi8. The small, simple cameras allowed us to work with a very small crew: camera-person, sound person and director. This was a bare minimum. We used two camera for some of the scenes. Dave Tamés, the cinematographer would rather have had an assistant for such situations; instead, he was reduced to operating one camera himself and operating the second one remotely with the small infrared controller. Not an easy task. An assistant would also have helped carry the extra equipment around (a simple lighting kit, boom, microphone, gaffer tape, etc...). We got the shots, although next time I would give the cinematographer an assistant.



Name: Ryan Evans
 Born: 1/15/69 in Kansas City, Missouri
 Role: Audio Technician
 Day Job: Research Assistant at the MIT Media Lab
 Real Job: Blowing smoke in front of the mirrors
 First memory: Stealing a steak from my Grandfather. Unfortunately he hadn't cooked it yet.
 Last dream: Trapped in a tower full of video games.
 Favorite movie: Charles Laughton's "Night of the Hunter"
 Last book read: "The Selfish Gene" by Richard Dawkins
 Ambition: to graduate
 Recent accomplishment: actually beginning to write my thesis
 Fetish: "The Patty Duke Show"
 Most annoying shooting experience: realizing that the little red light in my PXL-2000 viewfinder meant "low light" rather than "now recording"
 Favorite director: Sam Peckinpah
 Drink: double espresso
 Last car driven: '71 Dodge Charger
 Favorite Internet hangout: asylum.sf.ca.us

Having such a small crew allowed us to change locations very quickly. The two actors, three crew and all the equipment fit (tightly) into the production vehicle, a 1964 Thunderbird. This vehicle doubled as the main character's car in the story. We would arrive at a location and take advantage of the situation as we arrived. For example, we filmed at the South Station train station; we arrived to find that an Amtrak passenger train was sitting idle on one of the tracks. No one was around, so we just set up the camera next to the open door and staged the entire scene where Nicole leaves Jack. We even went so far as to shoot a couple of scenes inside the carriage. After about 45 minutes of careful shooting, we left with all the shots we needed in the can (on the tape). Such a scene would have been impossible to arrange with a large crew or with the cooperation of Amtrak.

For interior locations, we had a little more equipment in the form of lights. These scenes were shot in my apartment and allowed us the luxury of setting up lights

and shooting with a little more control. The result is footage that looks and sounds remarkably good.

Sound

Sound is very important. Using the internal audio on the Hi8 cameras is not a viable option. Even with an external microphone, the sound is of poor quality because the automatic gain control raises the noise floor every time the ambient level drops. There is no manual audio level control on most Hi8 cameras. At David Tamés' suggestion we recorded sound using a double system. We used a TEAC DAT recorder with a shotgun mic and for scenes where booming was impossible (in a public diner) we used radio mics. With this setup the sound recordist is completely independent of the camera person.

Synching up the image to the sound is a little more difficult with double system. We had to use a slate during shooting, taking us back to the days of film. This means that the sound has to be synched with the image during post-production, adding to post-production time. The extra effort is worth it, however, because the sound quality we achieved with this setup is very high. It is probably close to if not better than that available on Betacam.

Post-production

The material was edited in two different ways. After the first shoot, we transferred the Hi8 video and audio directly to Betacam. This maintained the highest possible



Name: David Tames
 Born: August 27, 1960 in Lancaster, Pennsylvania
 Role: Videographer (Part 2)
 Day Job: Technical Instructor
 Real Jobs: Cinematography and Multimedia
 Movie quote: "I'm quite convinced that cooking is the only alternative to filmmaking" -- Werner Herzog
 Formative experience: Seeing "2001: A Space Odyssey" for the first time on the big screen as a nine years old.
 First memory: 48K of 16Kx1bit DRAM chips
 Last dream: Driving a shiny red sports car along the pacific coast highway
 Favorite movie: "Speaking Parts" (Atom Egoyan)
 Last book read: "Points of resistance: women, power & politics in the New York Avant-garde Cinema, 1943-71" by Lauren Rabinovitz
 Ambition: Get paid to shoot film, watch movies, read novels and make espresso
 Recent accomplishment: Figured out the mysteries of Photoshop channel operations
 Fetish: Chocolate and...
 Most annoying shooting experience: Mark would not let me have an assistant while we were shooting
 Role models: Nestor Almendros, Paul Feyerabend, Woody Allen, David Byrne
 Favorite directors: Atom Egoyan, Woody Allen
 Drink: Cabernet Sauvignon
 Last car driven: Mazda Miata
 Favorite Internet hangout: comp.multimedia, rec.video.production

quality and minimized the damage to the original material. After synching image and sound, we culled the best takes and transferred this material directly to QuickTime. The editing was then carried out using Adobe Premiere 2.0. This is a good approach if there is not too much material and if the edits are not too complicated. The main drawbacks are that a lot of material needs to be digitized and that the editing is a little clunky and slow. It is difficult to see much detail in the frame because of the low level of resolution of the image. The advantage is that it is possible to do all kinds of effects that used to be possible only in expensive on-line editing suites. This is a significant impact of the desktop editing revolution. It is possible to achieve professional looking results with a desktop computer. Right now, the frame-rate and size is limited, but it is only a matter of time before all the tricks are possible at full frame-rate and resolution.



Name: Sheila Egan
 Born: 5/31/68
 Role: Beth, Nicole's roommate
 Day Job: Publishing
 Real Job: Actress
 Movie quote: "That's it, we're suing" from
King of Comedy
 Formative experience: First play rejection in
 6th grade
 First memory: Being a princess
 Last dream: Bad wedding day
 Favorite movie: *Roman Holiday*
 Last book read: Monkeys
 Ambition: Be happy
 Recent accomplishment: Working with a
 hangover
 Fetish: Lipstick
 Most annoying shooting experience: You've
 been my only one!
 Role model: A sane and happy person
 Favorite actor: Katherine Hepburn
 Favorite director: Billy Wilder
 Drink: Maddress
 Last car driven: Toyota Camry
 Favorite hangout: A theatre

After the second shoot, I decided to edit on the Digital F/X, a non-linear editing system. The advantage to this approach is that it is possible to edit directly from tape – it is not necessary to digitize all the rushes. The quality of the final cuts is good and only the final cuts need be digitized, an important consideration if disk space is at a premium. The drawback with this approach is that special effects are more difficult. They have to be accomplished after the material has been digitized.



The Crew: David Tames, Ryan Evans, and Mark Halliday (right to left).

Full stomachs

I think probably the most important aspect of the *Train of Thought* shoots, and all the other shoots I have participated in, is food. Plenty of good food will keep people happy. Happy people do good work. Good work shows up on-screen. Never skimp on this aspect of any production.

In Closing

Making movies has always be a very personal thing for me. Whether shooting a documentary or a fictional film, I have always been fascinated by the images on the screen and by the synthesis of people

and machines that make it possible. I love the technology, which is one reason I came to the Media Lab. The work described in this thesis is more about machines and their workings than about people, but no matter what technology we use to make our movies, it is still the people in front of and behind the camera who make it all possible.



Working together filming the kite scene.