

# Bridging Across Content and Tools

*Glorianna Davenport*

*Associate Professor of Media Technology*

*Media Arts and Sciences Program*

*M.I.T. E15-433*

*20 Ames Street*

*Cambridge, MA 02139*

*(617) 253-1607*

*Email: gid@media.mit.edu*

Over the past 10 years, I have contributed to the development of two kinds of "multimedia environments." In the first, the human-machine partnership evolved as a means of accessing specific content for discovery and enjoyment - this was true for works such as "New Orleans in Transition" (1987), "Elastic Charles" (1988) as well as for several fictional narratives. In the second, the human-machine partnership was designed to allow users facilities for describing, managing and sequencing content. Both types of environments encourage the invention of richer human-readable pneumonics and incorporate machine-readable description as tools for content-look-ahead. What has fascinated me over the years is the complementarity which binds the generation of content and the design of tools. In fact we cannot talk about form without discussing content and the tools for accessing that content.

## **The content framework and interaction**

In linear cinematic storytelling, we have a 100 year history of discovering how to immerse the viewer in story framework. Computational environments, particularly those which offer the viewer editorial tools, challenge our understanding of immersion. As designers and content makers we ask: What entices the audience to interact? How does the system reinforce meaning for the viewer?

One key to the successful design of random access environments has to do with positioning the audience (usually an audience of one at a time) to engage in role playing. This is important because it can motivate active participation. A role that has been extensively used is that of a student or researcher. The definition of the role allows the audience to anticipate what sort of experience they are likely to encounter. Typical functions which a student or researcher will require include search, show me that, show that again, mark that for future consideration, link this to that. Models of tools which support these actions can be found in environments for library indexing and in the process of film and video editing.

At the outset, the role of researcher offers a compelling user footprint because as a researcher, the audience of one, is willing at least in theory, to hunt and poke through windows and keywords in order to discover some new aspect of the story. This solution was used in my own research with "New Orleans in Transition," in Ricki Goldman-Segall's "Learning Constellations" and in Thomas G. Aguiere Smith's work with the Mayan Indians. The descriptive base of classes and keywords possibly combined with some forward motion through chronology offers a crude form of limited look ahead. For example, what is the next instance of video which has been tagged as the Jackson Brewery Development, or what is the next instance of John in class? As we move into the future, these search functions need to be expanded to include relationship and analogy. This area of content description is a hot bed of research activity. In addition to extending the descriptive base, we can easily extend the top level interface and provide the audience with a notion of the layered complexity of content. The idea of navigation through a tool kit which includes facilities for searching, comparing, annotating and editing provides clear access to critical modalities of interaction.

Ultimately, however, the research intent offers limited condolence for those who enjoy good stories. Too much energy must be spent in making interesting connections, which in turn limits time which could be available for viewing. In designing for entertain-

ment as opposed to education, we cannot forget that most people switching on TV for a few minutes do other things with their lives! For these viewers the MORE button in an interactive news story is perhaps more appropriate than the research paradigm. The problem for representation is to understand what we mean by MORE in a given context. As we grow large media banks which can incorporate many levels of media, the problem of representation becomes more difficult. In such environments, the database may grow over time as users participate in the annotation. In future environments which transform communal memory into bits, we will be severely challenged by the need for personalization.

## **Tools for payout**

Perhaps the most important aesthetic of a computational content environment is fluidity—fluidity of expression and fluidity of manipulation devices. Editing is the traditional means of sorting the "ins" from the "outs," thereby shaping the meaning of a particular motion picture. Editing is a wonderful example of what Seymour Papert and others call a "heads in, hands on" activity of learning. As we make choices about content, we think through multiple interpretations of the story. Editing tools facilitate (or hinder) the ability of the editor to discover and shape a meaningful story.

Having edited many stories in film and video, it is not surprising that I am interested in how we embed editing knowledge in software in order to allow stories to payout on the fly in interactive viewing environments. In fact, over the years of building environments for content, it has twice happened that the resources which have been provided for viewing activities have inspired the development of more generic tool applications.

In the case of the "New Orleans in Transition" project in 1987, the toolkit included a relational database description of all sequences, a sequencer and a method of linking sequences to buttons. Over time and through the ideas and work of three students, Thomas G. Aguiere Smith, Lee Morgenroth, and Erhhung Yuan, this toolkit was transformed into a system we now call Stratagraph. Stratagraph includes a logger which allows us to attach classes and keywords to the video stream, a graphical display which shows us how these descriptions relate to each other over time and which allows us to search on any one or a group of attributes, a sequencer and Homer, a story design application. Homer allows a user to roughly design a story which is then programmatically filled by a best guess method. As the story is fine tuned, the database is updated and so, in some sense, learns through the activity of use.

In the second case, Hans Peter Brondmo used methods of working with video clips which he initially programmed as a toolkit for the "Elastic Charles: A Hypermedia Journal" in the desktop editing package, VideoShop, which is owned and distributed by AVID Technologies. While VideoShop does not emphasize linking in a hypermedia mode, it does make use of the mnemonic "micon" as well as the organizational bins. Currently a logger is in development for use with VideoShop.

Still, none of these tools provide the fluidity for thinking with video which I feel will ultimately be possible. As a result, we are evolving yet another generation of tools which address issues of fluid manipulation, high level and low level descriptions and new methods for structuring story. One such tool, the Video Streamer, is essentially a capture utility which successively stacks incoming

frames of video by pushing the stack back and offsetting each frame as it is pushed back by one pixel width along the top and left side of the frame. What is unique about the Video Streamer is this 2 1/2 D representation which reveals the edge of video over time. If the capture rate is close to real time, approximately 30 seconds of video can be displayed on the screen at a time. The user can rapidly browse this stack of frames and easily define a shot or clip, saving it directly to QuickTime or to QuickTime by way of Hypercard.

In addition to video capture, some new tools for annotation and filtered playout have been designed by Ryan Evans. These tools presume that if we know the nature of the story we want the machine to play out, we use a limited set of descriptions which are designed in concert with the filter set which will be used in retrieval. While the story filters are not yet easy to use, they do allow us a scaleable way to build limited-look-ahead functionality into the content playout. For instance, in a portrait of the changing face of downtown Boston which we are currently building as a collaborative project, we have collected many shots of the cityscape. We can play these shots out from the perspective of moving through the city or by historically sequencing stories about Boston's history.

In order to gain more leverage from our descriptions in Stratagraph, we have recently hooked this application to Framer, a persistent knowledge representation data management system written by Professor Ken Haase. In a parallel development we are working on the idea of "concept frames" which will allow journalists to represent the idea behind their story as opposed to just describing the video elements. We hope that both of these developments will enable a future where limited-look-ahead can be based on broad issues and the "why" of journalism in addition to the "who," "what," "when," "where" which we make use of today.

## Conclusion

I have tried to point out that there is a symbiotic partnership between content presentation and tools for manipulation. In the digital age, developing one in the interests of the other is less interesting than finding a middle ground and working out in both directions. For me, this middle ground is content description. There are, however, clear tradeoffs in how we tailor our descriptive strategy for the eventual task of retrieval or playout.

## References

Brondmo, H. P. and Davenport, G., "Creating and Viewing the Elastic Charles: a Hypermedia Journal," *Hypertext II*, York, England, June 1989.

Davenport, G., Evans, R. and Halliday, M. "Orchestrating Digital Micromovies." *Leonardo*, Vol.26, No. 4, pp. 282-288, 1993.

Davenport, G. and Smith, T. A. "The Stratification System: A Design Environment for Random Access Video." Proceedings of the ACM Workshop on Networking and Operating System Support for Digital Audio and Video, San Diego, November 12-13, 1992.

Davenport, G., Smith, T. A. and Pincever, N. "Cinematic Primitives for Multimedia." *IEEE Computer Graphics and Applications*, pp. 67-74, July, 1991.

Elliott, E., "Watch, Grab, Arrange, See: Thinking with Motion Images via Streams and Collages," MSVS Thesis Document, February 1993.

Goldman-Segall, R. "Learning Constellations: A Multimedia Ethnographic Research Environment Using Video Technology for Exploring Children's Thinking." Doctoral dissertation in media arts and sciences, 1990.

Haase, K., "Framer: A Persistent Portable Representation Library," MIT Media Lab Technical Report, April 1993.

Mackay, W. and Davenport, G., "Virtual Video Editing in Interactive Multimedia Applications," *Communications of the ACM*, July 1989.