



© IBM Corporation

Vol 36, No. 3 - Nontopical issue

[**Table of contents: HTML, ASCII**]

[**This article: HTML, ASCII**]

0018-8670/97/\$5.00 © 1997 IBM

Automatist storyteller systems and the shifting sands of story

by G. Davenport and M. Murtaugh

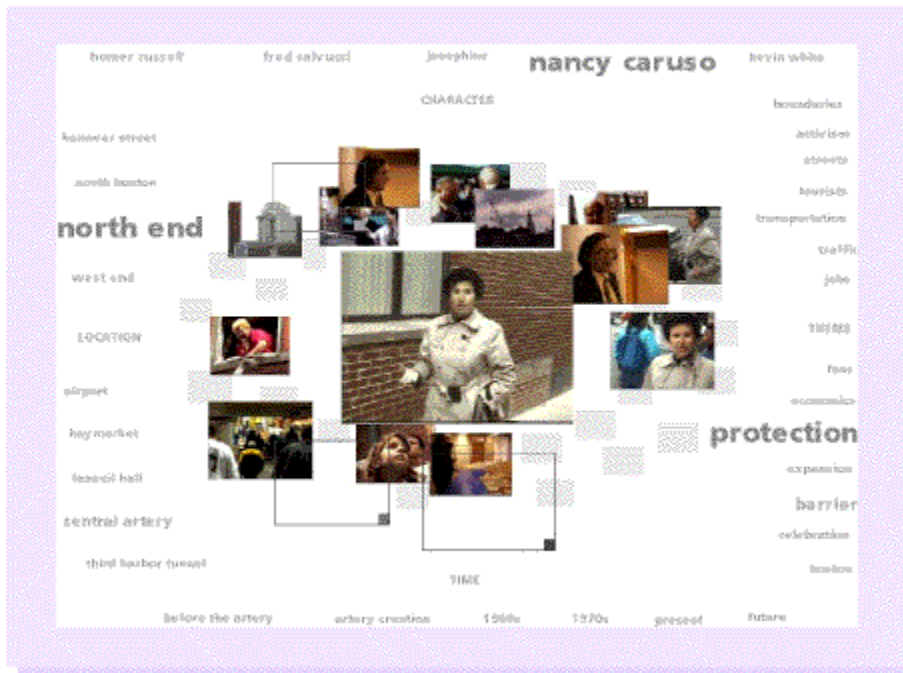
Reprint Order No. G321-5652.

We present a novel approach to documentary storytelling that celebrates electronic narrative as a process in which the author(s), a networked presentation system, and the audience actively collaborate in the co-construction of meaning. A spreading-activation network is used to select relevant story elements from a multimedia database and dynamically conjoin them into an appealing, coherent narrative presentation. The flow of positive or negative "energies" through associative keyword links determines which story materials are presented as especially relevant "next steps" and which ones recede into the background, out of sight. The associative nature of this navigation serves to enhance meaning while preserving narrative continuity. This approach is well-suited for the telling of stories that--because of their complexity, breadth, or bulk--are best communicated through variable-presentation systems. Connected to the narrative engine through rich feedback loops and intuitively understandable interfaces, the audience becomes an active partner in the shaping and presentation of story.

For over a decade, research in interactive cinema has posited the idea that a computational system [1,2] that incorporates the decision-making process of the documentary video editor could serve the needs of interactive storytelling. Recently, an extremely interesting "editor-in-software" system for interactive storytelling was designed by Michael Murtaugh, who received his master's degree from the Massachusetts Institute of Technology in 1996. [3] In developing his approach, Murtaugh was strongly influenced by Pattie Maes's work in autonomous agents [4] and by visualization systems developed in the Visible Language Workshop of the MIT Media Lab in 1993 and 1994. [5] These systems rely on a decentralized model of computing and the use of a spreading activation network to dynamically determine relationships between agents or elements. However,

Murtaugh's work represents a departure from work in the "autonomous agents as characters" camp. His prototype systems-->[6]--do not simulate the story object itself; instead, by selecting the next story element based on the context of what preceded it, they simulate the very process of storytelling and story understanding. (See Figure 1.)

Figure 1 Nancy Caruso talks about protecting Boston's North End in our evolving documentary



To the extent that Murtaugh's work represents an interesting direction in our progress toward "interactive cinema," it is useful to situate this work within the broader context of story and content construction.

The beginnings of automatist storyteller systems

For centuries, artists, mathematicians, and engineers have demonstrated their interest in reconfigurable or "automatic" storytelling. In the 15th century, Gutenberg's innovative printing press--which featured movable and reusable type--drove home the notion of machine-reconfigurable text. In the late 17th century, Jonathan Swift published a delightfully tongue-in-cheek account of an automatic "wisdom-generating device": an army of clerks turned cranks that spun vast arrays of lettered blocks; meanwhile, a supervisory bureaucrat scanned the scene and jotted down any sentences that accidentally appeared.

In the 1920s and 1930s, the Surrealists and Dadaists delighted in experiments and parlor games that took stream-of-consciousness text--often generated by several distinct voices--and merged them into a single composite entity. They believed that by suppressing the conscious mind, they could release spontaneous, intuitive imagery from the subconscious--a process they termed "automatism." Other artists advocated a chance mechanism for authorship, such as pulling words out of a basket. In attempting to set

expression free from the conscious control of a maker, these artists were responding to contemporary cultural concerns and scientific opportunities.

With the invention of the computer, the creation of random processes for text generation became a "no-brainer." Randomness quickly proved itself to be an uninteresting storyteller. The difficult problem was how to shape a system that could massage a pool of existing elements into coherent stories. Harder still was the challenge of generating good stories.

In the 1970s, the emergence of frame-addressable videodisks ushered in a period of experimentation with fixed branching through pre-made audiovisual story materials. The limited amount of "real estate" available on these disks--about half an hour of video and synchronized audio per side--forced authors to offer a minimal set of interactive choices. Typically, these interactive videodisk-based stories were built from a small inventory of traditionally crafted scenes; a large, monolithic chunk of story would play out, the presentation would grind to a halt, and the audience would be offered two or three predetermined choices of where to go next. The large granularity of story chunks, the small number of chunks the disk could hold, the desire to make the story experience coherent and well-crafted, and the cumbersome and disruptive nature of the control interface all served to place the task of narration squarely on the author's shoulders.

The 1970s and 1980s saw the emergence of hypertext systems, which offered opportunities for interactive branching at a single-word granularity. As text was placed in a vast web of interconnected links, the assumptions of visionary authors like Roland Barthes and Michel Foucault were called to task. [7] What was intended to be encyclopedic became kaleidoscopic. [8] In retrospect, as Murtaugh's work will assert, "sparse and basic" are what may save us from becoming the lost traveler.

The late 1980s will be remembered as the age of consumer video. Suddenly, video cameras were becoming cheap, light, and small. The convergence of affordable video capture and playout devices, plus the availability of large, randomly accessible disk memory, created the revolution in nonlinear editing. With video edit suites becoming the *quid pro quo* of the personal computer desktop, amateur early adopters challenged professional editors for the business. By the mid-1990s, editing in the entertainment industry had entered the digital age.

Over the last two decades, digital storytelling has matured but has yet to flourish. Before it becomes a mature art form, several stumbling blocks remain to be conquered, particularly: the need to create a truly systemic approach to narration and story structure; the need to derive a flexible, universally applicable representational schema that describes the form, content, composition, and subtext of media elements; and the need to establish conventions for interaction that are acceptable within the story framework.

What's wrong with the television documentary?

As the coauthor and a former documentary filmmaker who made programs ostensibly for television, I am often asked, "What's wrong with TV?" From my perspective, the

greatest limitation of television is that rather than causing the viewer to think, television consumes the viewer. Sitting passively in front of a TV screen, you may appreciate an hour-long documentary; you may even find the story of interest; however, your ability to learn from the program is less than what it might be if you were actively engaged with it, able to control its shape and probe its contents.

Television severely limits the ways in which an author can "grow" a story. A story must be composed into a fixed, unchanging form before the audience can see and react to it: there is no obvious way to connect viewers to the process of story construction. Similarly, the medium offers no intrinsic, immediately available way to interconnect the larger community of viewers who wish to engage in debate about a particular story.

Like published books and movies, television is designed for unidirectional, one-to-many transmission to a mass audience, without variation or personalization of presentation. The remote-control unit and the VCR (videocassette recorder)--currently the only devices that allow the viewer any degree of independent control over the playout of television--are considered anathema by commercial broadcasters. Grazing, time-shifting, and "commercial zapping" run contrary to the desire of the industry for a demographically correct audience that passively absorbs the programming--and the intrusive commercial messages--that the broadcasters offer.

Documentary production zeitgeist

As a documentary filmmaker, I find what takes place off-camera as fascinating as what the camera actually captures: the editing room is as interesting as the screening room. Documentary filmmakers are driven by a passion for exploration; in contrast, documentary editors are "bricoleurs" who fit together the often disjunct bits and pieces of media into a coherent story experience.

Many years ago, when I was working on my first "interactive" documentary, I was introduced to the concept of relational databases. From that time forward, I have had the sense that if we could only find the right way to index documentary film segments, then we could design an "editor in software" that would emulate the processes and expertise of the film editor. Such a system would support the human user by offering relevant suggestions, or could navigate a large database filled with many aspects of a complex story and "make choices about what the viewer would like to see next."

Over the past decade, students in the Interactive Cinema laboratory have developed many systems that attempt to solve the "editor-in-software" problem. For the most part, I have suggested that they build content and delivery systems simultaneously, taking note of the constraints and powers that characterize each. In addition, I strongly urge students who have minimal production experience to turn their attention to documentary storytelling.

My rationale for this is twofold. First, I have a strong intrinsic sense of how a documentary is produced and constructed, and of what makes an interesting documentary story "work." In addition, the real world is intrinsically complex and multifaceted: a particular type of organization is required to follow a story as it emerges. The methodologies of investigation, which have been refined and proven through extensive

use, offer valuable insights to system designers. For example, a documentary filmmaker exploring some aspect of social change might begin his or her inquiry by asking, "Why did this happen?" or, "How does this work?" However, in order to discover the "why" or "how" of a situation, the filmmaker must get very specific, organizing an investigation around questions of "who, what, where, and when." As the ever-optimistic filmmaker/explorer collects material, he or she hopes that a network of composite observations ("Who did what where?" etc.) will provide a way of understanding the larger "whys" and "hows" of the matter.

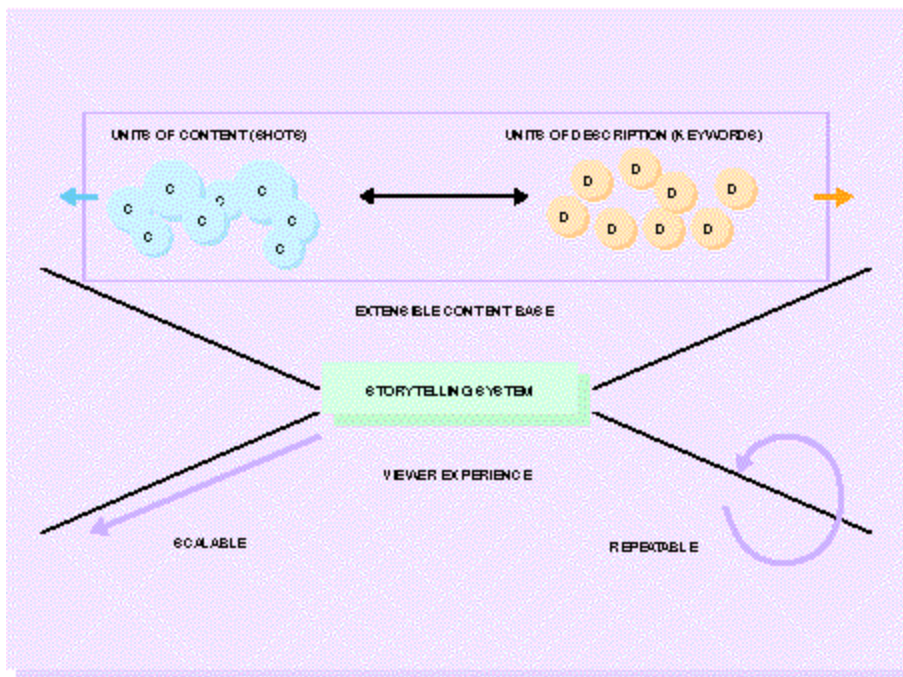
Editing, emergent stories, and the evolving documentary

The "traditional" process of making a documentary film could be roughly described in the following way: The filmmakers collect a large amount of raw material--original film footage, audio recordings, archive photographs, and text articles. These raw materials are organized into progressively larger chunks of story: shots, scenes, and sequences. Finally, the finished sequences are edited together to form the final "cut" of the movie. The resulting experience, as presented to the viewer, is rigid and uniform; every viewer sees the same presentation, no matter when or how they see it.

Described in this way, the filmmaking process may be seen as a kind of leaky funnel. A large collection of content elements--frequently an order of magnitude larger than the final piece--is gradually refined and reduced to form the program. As editing decisions are made, the program becomes more and more determined; as each shot is placed in position, the demands of coherence, context, and continuity dictate to some degree which shots and scenes can meaningfully precede or follow. As the various pieces fall into place, a specific story--with its own particular themes, central characters, and motivations--begins to emerge.

The experience provided by a storyteller system might be described as hourglass-shaped--open on both the authoring and the viewing sides. In this model, the storyteller system does not allow the author to explicitly sequence story elements into a finished tale: there is no "final cut" of the film. Instead, editing decisions are deferred until the moment of ployout. Thus, storyteller systems might consider the context of a particular viewing experience, the preferences and interactions of the current viewer, and what databased material is presently available when selecting content for display. (See Figure 2.)

Figure 2 The storytelling system model



In such a system, the viewer's experience is no longer rigid or uniform. The experience itself is extensible; viewers are free to stay with the story for as little or as long a time as they wish. The experience is also repeatable; viewers could leave having only seen a portion of the available material and then return later to see more.

The system is open-ended on the author's side as well. Real-world stories are seldom complete in themselves; a detailed picture of circumstances may only emerge over the course of days, or weeks, or even several lifetimes. Thus, the resources of a story (and its associated descriptive database) can grow and evolve as newly discovered information is added, or as users add their own commentary and evaluations of quality and veracity. Stories of this type may be described as having "emergent" or "evolving" properties over time. Instead of "sealing off" the story with the release of a particular program or film, the base of content is free to grow as the story grows.

Furthermore, as structural decisions are deferred until playout time, the story remains to some degree undetermined and thus free to support variable presentations. When the viewer is faced with a substantial mass of accumulated information, looking at it through a particular focus--such as choosing to follow a specific individual, or by adopting a particular philosophical point-of-view, or by taking information culled from only one source among the multitude--can yield many widely differing stories.

Two heuristics for designing storyteller systems

Before endeavoring to build a storytelling system, it is useful to identify heuristics that add constraints to the design. Here, the ideas of "autonomous playout" and "direct access" are embraced as highly desirable design characteristics.

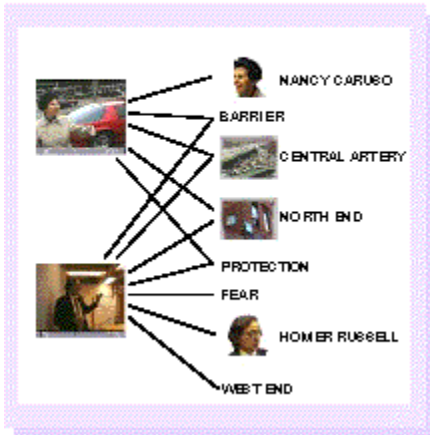
A common experience when viewing contemporary CD-ROMs seems to be an increasing frustration with having to use the story's interface to "get at" the content. Eventually, if you are actually interested in the content, you just want the thing to "play out by itself." The ability for automatic- or self-playout, therefore, serves as a powerful design heuristic for building a storyteller system. Designing around the potential absence of a viewer requires that a system be built with enough base-level competence to present its content autonomously. The addition of interactivity poses an interesting challenge, as the role and value of the interaction must always be gauged against its absence.

As with self-playout, the designer of a storyteller system might imagine a similar base-level functionality that provides direct and immediate access to all of the story's content (such as the way a file system might be used to directly browse the media files of a CD-ROM). Any additional functionality or control given to the viewer must then be gauged against direct access. In this way, the piece must prove its value by enabling a method of construction appreciably better than simple random access.

Using keywords for deferred sequencing and extensibility of content

In an automatist storyteller system, simple keyword descriptions associated with media objects provide the crucial function of isolating authors from the process of defining explicit relationships or links between units of content. Instead, by connecting a material (story element) to a keyword, the author defines a potential connection between the material and others that share that keyword. By connecting each material to a set of keywords, the author enables a material to be related to other materials in more than one way. (See Figure 3.)

Figure 3 Two video clips annotated with a set of keywords



Lacking explicit links, sequencing decisions are made during the viewing experience based on implicit connections via keywords. Deferring sequencing decisions in this way has two consequences: First, the base of content is truly extensible. Every new material is simply described by keywords, rather than hardwired to every other relevant material in the system. In this way, the potential exponentially complex task of adding content is managed and made constant. Second, because sequencing decisions are not precoded, viewers may play a more active role in the construction of the experience. Instead of

using predetermined links bound to a specific purpose or organizational scheme, viewers may influence how they want to move from one material to the next.

Autonomous agents and automatist storytelling

The approach taken in an automatist storyteller system is highly decentralized and draws on the techniques of autonomous agents. In her introduction to *Designing Autonomous Agents*, Pattie Maes describes a shift in artificial intelligence research from approaches based on "deliberate thinking" and "explicit knowledge" to ones based on "distributedness and decentralization." She notes how these new approaches avoid the "brittleness" and "inflexibility" of the former by using "dynamic interaction with the environment and intrinsic mechanisms to cope with resource limitations and incomplete knowledge." [9]

Maes goes on to describe an approach to programming the mechanical behavior of robot-based autonomous agents. Decisions about what action the robot should take at any given moment are based on an "action selection" algorithm. In this scheme, the "competency modules" are based on specific actions the robot arm can perform. The applicability or usefulness of each action is a function of the current state of the environment. When an action is selected and performed, its invocation alters the environment, thus influencing the selection of future actions. In this way, a sequence of actions--a plan--emerges. [10]

In an automatist storyteller system, editing decisions are made based on a similar action-selection algorithm. In this case, individual story materials (short video clips, pictures) and keywords act as modules with an "internal representation" consisting of a list of associated modules; materials are associated with a set of keywords, and, conversely, keywords are associated with materials. When invoked, both materials and keywords spread activation to their associated modules. The resulting interaction of the spreading activation forms the basis of how materials are selected and sequenced. Thus, the resulting structure of the story is an "emergent property" of the interaction of individual material presentations.

Although the approach taken in an automatist storyteller system closely conforms to the ideas of autonomous agents, it differs significantly from previous applications of this methodology to the area of storytelling. For instance, in Maes's own subsequent work, agents are applied in the following way:

Many forms of entertainment employ characters that act in some environment. This is the case for video games, simulation rides, movies, animation, animatronics, theater, puppetry, certain toys and even party lines. Each of these entertainment forms could potentially benefit from the casting of autonomous semi-intelligent agents as entertaining characters. [11]

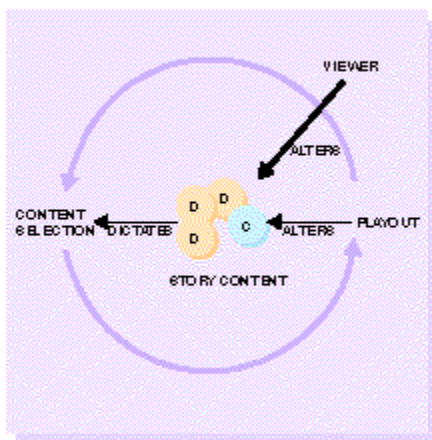
Thus, research originally developed in the context of coordinating the actions of a robot arm in an industrial environment is used to plan the actions of virtual characters in a fictional environment. Viewers are considered a part of the environment and thus, in a literal sense, "inside the story." The process of story construction is typically viewed as one of generating a sequence of events, or a plot, based on the potential actions of

characters' internal rules (or "motivations") while maintaining certain global rules (such as gravity or logical cause and effect). Ultimately, the challenge of constructing a "good" story is reduced to the process of creatively expressing a well-formed chain of events.

In an automatist storyteller system, the fundamental units of structure are not events to be expressed but expressions themselves in the form of discrete units of content. Instead of characters interacting in an environment that is literally the "story world," individual expressions interact in an environment that is the process of storytelling.

In addition to enabling both an extensible base of content and an emergent story structure, the decentralized approach of an automatist storyteller system also consistently integrates the viewer's interaction. In a decentralized system, incorporating the presence of the viewer is straightforward: the viewer exerts influence over the emergent functionality of the system in the same way that any other component of the system does, by altering an aspect of the environment or influencing the operation of other components. (See Figure 4.)

Figure 4 Audience feedback is an active component of the storytelling system.



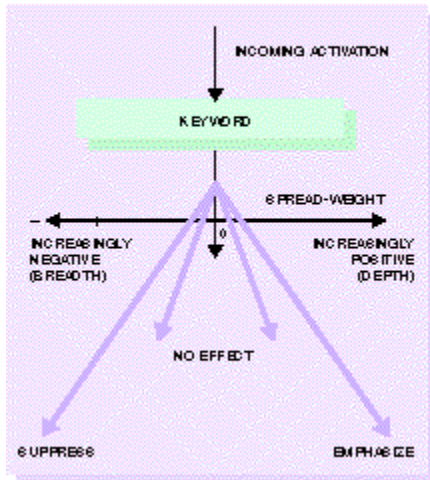
In this decentralized approach, the viewer is a full-fledged member of the system and consistently integrated into the experience. This contrasts with the model of hypermedia, where the consistency of viewer interactivity depends on the author's consistency in establishing links. In addition, while the operation of the system is open to the influence of viewer interaction, it is never dependent upon it. In this way, an automatist storyteller system allows viewers to exert influence only when they wish to and allows them to experience the immersive "reverie" of uninterrupted story construction.

ConTour: A design example

ConTour, a generalized system for producing continuous "steerable" presentations of keyword-annotated movies and pictures, provides us with a design example. ConTour is the result of several iterations of storytelling systems designed in conjunction with the story "Boston: Renewed Vistas," [12] and those materials are used in the illustrations. However, it is easy to replace one set of materials with another.

In ConTour, materials and keywords act as modules with an "internal representation" consisting of a list of associated modules; materials are associated with a set of keywords, and conversely, keywords are associated with materials. Both materials and keywords spread activation, when invoked, to their associated modules. The resulting interaction of the spreading activation forms the basis of how materials are selected and sequenced. (See Figure 5.) Thus, the resulting structure of the story is an "emergent property" of the interaction of individual material presentations.

Figure 5 The spreading of activation



The interface of the ConTour application was designed to demonstrate the effects of the spreading activation network on material selection. Although the visual principles had been seen in previous work in the Visible Language Workshop, [5] ConTour demonstrates the effects of spreading activation along a temporal axis that is appropriate to movie playback. Every keyword and material in ConTour has an associated activation value. When a keyword is clicked on or a material is presented to the viewer, the activation value of the element is raised (the element is injected with activation). Together, the activation values of every keyword and material in ConTour form a closed or "relative value system," which serves as the basis for both the automatic material selection algorithm and the system's graphical display.

Activation values are used to determine how elements are drawn on the screen; the element's size, depth or z-coordinate, and brightness are all derived from its activation value. The system uses activation to represent an individual element's relevance to the current "context" of the story playback. Elements with relatively high activation values are made visually prominent by making them appear brighter and closer than elements with lower activation values.

By steering the user through the collection of materials, ConTour functions as a "digital editing assistant," interactively suggesting possible sequences of materials. At any time the user can influence the system by activating and weighting keywords. (See Figures 6 and 7.)

Figure 6 The initial state of "Boston: Renewed Views" before a selection is made

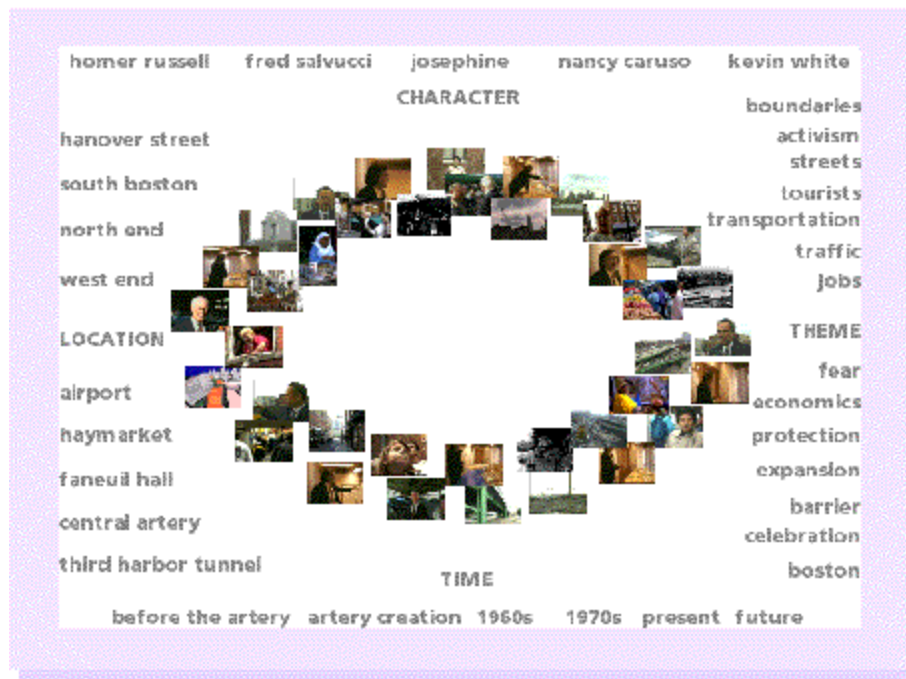
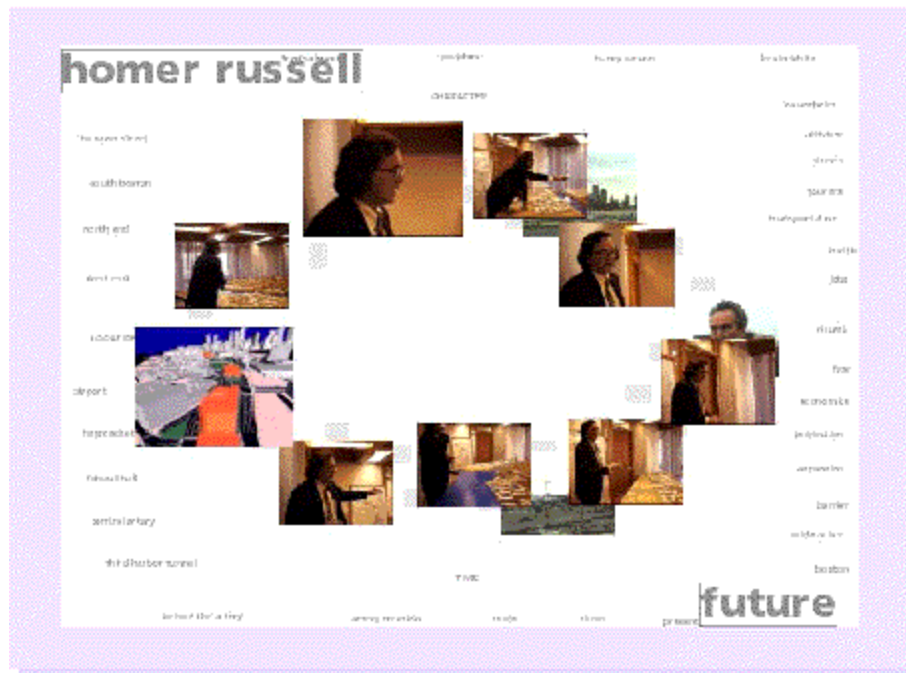


Figure 7 The state of "Boston: Renewed Views" after "Homer Russell" was selected



What I find most satisfying about Murtaugh's solution is that it mixes an author-centric approach to story creation with a generic approach to the selection algorithm. To the extent that the author creates the materials, including the keyword descriptions and hierarchies, the system reflects the human understanding of content. However, the algorithm that selects and presents possesses no deep domain knowledge, no far-reaching "common sense," and no special knowledge of the interrelationships among the available

story materials; instead, it operates on a statistical model of similarity. Likewise, the interface has no special knowledge of the content; rather, it presents all of the content, including the keyword representations and dynamic traces, to the user. In this way, the system functions both as a shape-shifter (by dynamically adjusting the signs of content) and as a mentor (by offering the "backstory" of what the viewers are watching).

On the cusp of story

Critical aspects of our autonomist storytelling system have been instantiated and substantially tested in two systems developed by Murtaugh: "ConTour" (a MacLisp implementation) and "Dexter" (a Java**²-based World Wide Web implementation). The current visual interface was designed primarily to communicate the computational principles of playout. Based on the casual observation of hundreds of demonstrations, the visualization of spreading activation is an extremely effective communicative device. Although not designed as commercial products, the ConTour and Dexter systems have both proven to be durable and easily extensible. Several new content sets are currently under construction by a variety of interested parties.

As with any "new" information type, it is difficult to evaluate the impact of this dynamically steerable, "evolving documentary" on an audience. Our review focused on three important aspects of these systems' use: communication, extensibility, and adaptability of the idea to existing interactive media channels. As Heidi Gitelman writes in her formative evaluation of the Dexter-based project, *Jerome B. Wiesner, A Random Walk Through the Twentieth Century*: [13]

Throughout the evaluation, respondents expressed strong positive feelings about "Random Walk." Respondents especially liked the nonlinear approach toward the subject and presentation of content. The combination, range, and quality of video and text from the Material Listing was important to them.

...Simultaneously, all respondents expressed strong and consistent concerns. These concerns fell into the categories of: user orientation, context, and to a lesser degree, the presentation of the information.

...Throughout the evaluation process, respondents were eager to understand and adopt the nonlinear approach to teaching and learning presented by "Random Walk." All commented that they liked this approach but also on the need to make "Random Walk" much more "user friendly" both navigationally and context-wise. Upon conclusion of the evaluation, all respondents indicated that although intrigued by the program, it is currently not accessible enough for them to use. All hoped this would change, as they are enthusiastic about the "idea."

...Finally, although respondent comments indicated that they were not able to develop an "emergent story," observations and further analysis of their comments indicate otherwise. All respondents were in fact, able to piece together concepts, ideas, and facts from "Random Walk" and make statements which indicated their assimilation of these ideas.

In evaluating a storyteller system, it is difficult to separate the form and content of a story from the system itself. Evaluating story, particularly a somewhat esoteric documentary story, will inevitably remain problematic--as Yeats observed, "How do we tell the dancer from the dance?" Despite the cognitive difficulties associated with the audience's "learning curve," this work points us along a course to a class of fully automated story engines. In selecting material for these systems, the authors have critically evaluated how the production and selection of story elements can circumvent the notion of a standard plot. The beauty of the system resides in "the editor-in-software" approach to story element sequencing and the contextual, associative nature of travel through story content.

**Trademark or registered trademark of Sun Microsystems, Inc.

Cited references and notes

Accepted for publication April 18, 1997.

Reprint Order No. G321-5652. -- -----
>

[[Journals home page](#) | [Subscribe/order](#) | [Current issue](#) | [Recent issues](#) | [Description](#)]

IBM Home	Shop	Contact IBM	Search	Privacy
Legal				

©1998 IBM Corporation