

# ConText: Towards the Evolving Documentary

## **Glorianna Davenport**

Associate Professor of Media Technology  
Massachusetts Institute of Technology  
Media Laboratory  
gid@media.mit.edu

## **Michael Murtaugh**

Research Assistant  
Massachusetts Institute of Technology  
Media Laboratory

---

## **ACM Copyright Notice**

Permission to make digital/hard copy of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of publication and its date appear, and notice is given that copying is by permission of ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright © 1995 by Association for Computing Machinery, Inc. (ACM).

---

## **Keywords**

content-based indexing and retrieval, hypermedia, interactive television, entertainment, digital libraries, databases, information storage and retrieval

---

## Abstract

The advent of digital technologies has enabled the emergence of a new type of database or "Media Bank" that supports an evolving collection of media elements. [1] This opportunity suggests the design of ConText, a system by which content, description, and presentation are separated into interconnected pieces, redefining the relationship between the story, the viewer and the author. For the viewer, repetition and revisitation of the story experience is encouraged and no constraints are placed on the duration of a session. For the author, the tasks of content gathering and sequencing take on new dimensions because the content base is extensible, and the author is separated programmatically from the exponentially complex task of explicitly sequencing the material for each viewer visitation. We call this new form the "Evolving Documentary."

A new and crucial authorial role becomes defining the core methodology that governs the story presentation and viewer interaction. The foundations of our proposed model have been developed and implemented in conjunction with an evolving story about urban change in Boston. This story features a 7 Billion Dollar public works project to rebuild the Central Artery (I-95) and the project's impact on surrounding neighborhoods. The project will be on-going through 2004 which makes it a practical story for an Evolving Documentary investigation.

This work has been funded by the News in the Future Consortium at the Media Lab and by the Eastman Kodak Company.

---

## Table of Contents

ConText: Towards the Evolving Documentary .....	1
<b>Introduction: The Evolving Documentary .....</b>	<b>3</b>
Conditions for an Extensible Content Base .....	3
Video Annotation and Narrative Retrieval: An Overview .....	5
<b>The Artery Story .....</b>	<b>6</b>
<b>The Author's Activities .....</b>	<b>7</b>
Annotating the Content.....	7
Presentation Methods .....	9
<b>The Interface .....</b>	<b>14</b>
The Viewer's Activities .....	16
Continuous Playout .....	16
<b>Conclusion .....</b>	<b>17</b>
<b>References.....</b>	<b>17</b>

## **Introduction: The Evolving Documentary**

For the past 100 years, motion picture and sound elements have been collected in order to be edited together to form a single program - a movie, a commercial, an MTV video, a documentary. Typically editing decisions are made on the basis of plot in the case of feature film, or on some combination of relationships which allow the author to determine the "best" sequence of shots in a commercial or documentary.

New digital technologies can support evolving collections of media elements which are stored and accessed non-linearly. A growing body of research looks at the problem of how video should be indexed. Research areas include automatic parsing and matching, as well as human supported annotation activity. Once an author or logger has attached descriptors to a set of media elements, these materials can be retrieved according to simple or complex directed queries.

While the retrieval-by-query mechanism suits particular users, it will not satisfy users who do not know what is in the database or what they want to see. Consider the difference between the editor who is looking for certain content in an historical shot, and the consumer of news who wants an in-depth review of a public works project.

This latter circumstance suggests both a new model for content and the need for a narrative approach to browsing which incorporates some partnership between the viewer and the presentation engine. Assuming a rich content database of media materials, such a browser needs to suggest narratives by taking advantage of descriptive associations and methods for temporal progression. This paper presents ConText, a model browser.

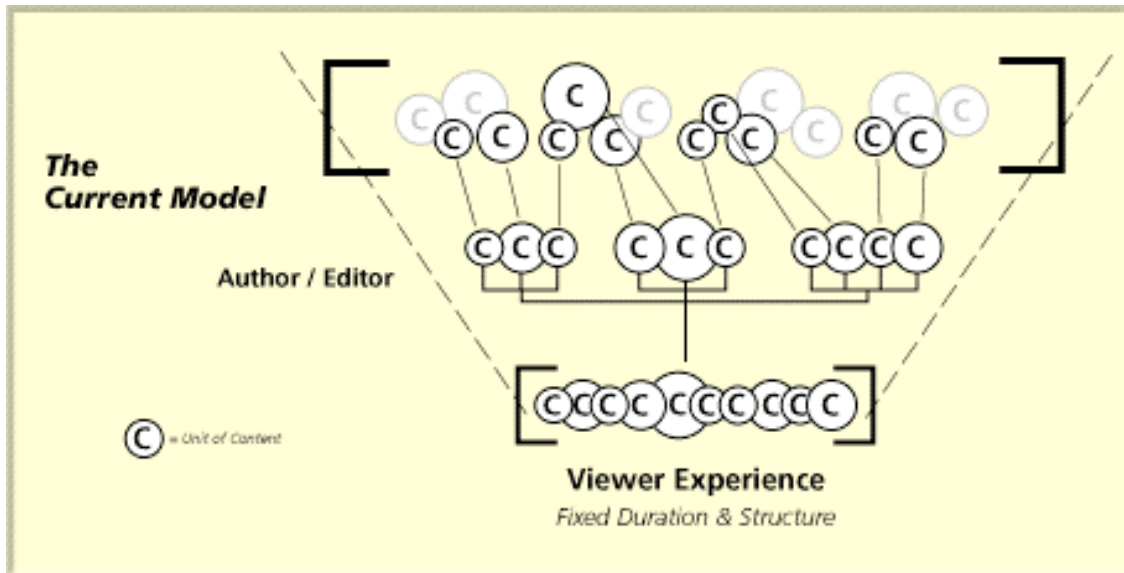
In the next section, we discuss the "evolving documentary" model for extensible content and review previous research in the area of annotation and story modeling. Next we discuss the author's activity with explicit reference to an example data set we are in the process of developing. We then introduce a set of presentation methodologies which can take narrative advantage of a description space. Finally we present the look and feel of ConText as a functioning browser. Our conclusion summarizes the current state of the work, underscores idiosyncratic aspects of the method, and suggests future directions.

### **Conditions for an Extensible Content Base**

The development of an extensible content base requires some organizational framework. What is being collected? Under what conditions are elements added? With what consequence? How is the storage and retrieval structured? There are many ways to answer these questions. We will focus on a particular content model which we call the evolving documentary. [2]

Certain ongoing news stories - including wars, elections, public works, science - fit this content model. The news story evolves as follows. A body of material is collected by a journalist for an editor who shapes the story. The journalist designs the story so that something is disclosed and if possible a cause is highlighted. The story rarely includes all

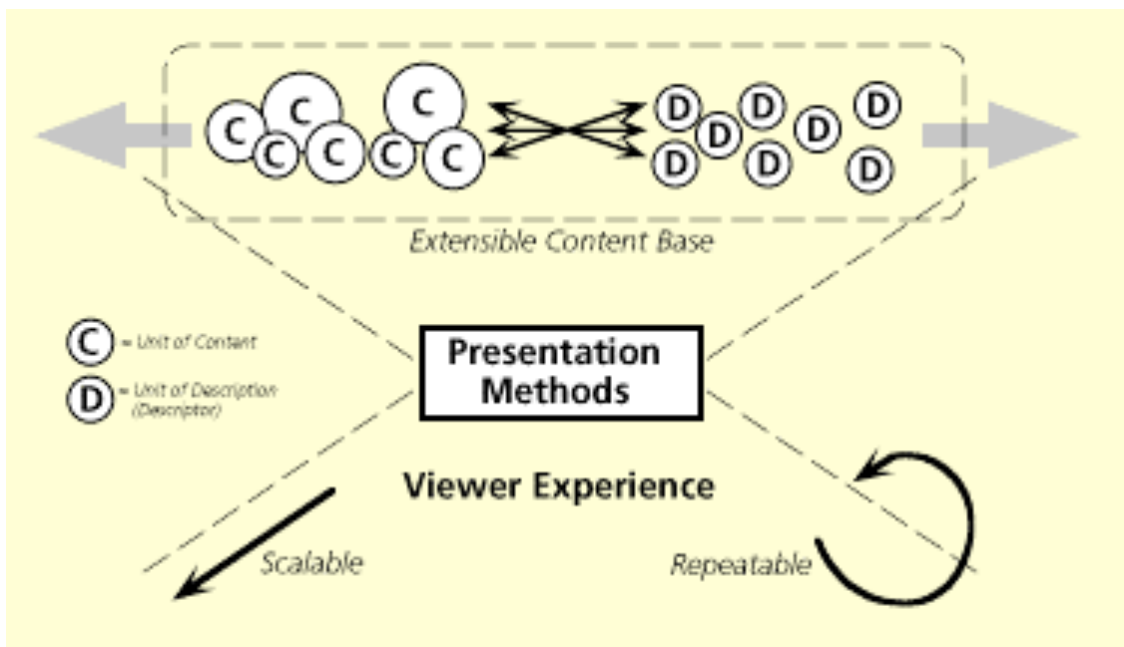
relevant material; rather the "best" material for the chosen story is selected and sequenced. This selection reflects available editing time as well as available air time. Figure 1 shows an overview of this process whereby an editor filters the available content down to a rigid form for presentation to viewers.



**Figure 1: The Current Model**

In the case of developing or "evolving" stories, both the journalist and the viewer's knowledge of the situation changes as additional stories are filed. Eventually, the "bigger picture" begins to emerge. The first story now becomes a mere fragment of a larger whole. Different reporters will often develop very different stories around different themes on different days. This breadth provides us with context. Both editor and viewer draw on these older stories to shape their understanding of a new event.

In the world of digital media systems, the materials for an evolving story may be stored immediately, on collection, in a content repository. The materials grow as the story evolves. For this reason the storage and descriptive architecture must be extensible. We must be able to add new materials easily without noticeably impairing retrieval. In addition, we must be able to retrieve materials without the viewer knowing explicitly what is there. In the case of the "Evolving Documentary," the media materials and the description space in which they are mapped must be scalable, and the experience afforded the viewer must be interesting enough to encourage repeated exploration. Figure 2 shows an overview of this process.



**Figure 2: The "Evolving Documentary"**

## **Video Annotation and Narrative Retrieval: An Overview**

Description of content materials is a critical aspect of authoring an evolving documentary. While contemporary research focuses both on human and automated annotation processes, our particular need for annotation is bound to our computational browsing model. The following review is intended to indicate directions which provided us with a starting point for the current approach.

In past work, we have developed systems for content annotation of video. Parallel development has looked at both class/keyword and knowledge-based systems. [3] In addition, we have considered the value of stream-based [4] and clip-based annotation methods. In most annotation schemes we have found the journalist's w's -- especially, who? what? where? when? -- to be critical annotations for editorial query; encoding the fifth w, why?, is still problematic.

We have also explored strategies for expressing and filling higher level narrative queries to a database of video materials. Such queries will inevitably add value to both the materials and the storage bank itself. In previous work, we have explored the idea of cascading multiple filters [5] as well as the idea of creating a temporal query which relies on class/keyword content descriptors. [6]

The limitation of these approaches is that they are explicit over a duration. The query cannot grow exponentially complex. If the story changes -- which could happen if for instance an interview became available which contradicted an earlier opinion -- the story model could not know to modify itself. Due in part to this limitation, neither of these

strategies provides us with a dynamically steerable, progressive narrative. The program computes a narrative whole and the viewer must watch all or quit. This may be an appropriate strategy for some multivariant fiction stories, but it is insufficient in the case of an "evolving documentary" where the database contains a rich mixture of breadth and depth.

Our conclusion, based on these experiments, is that an "evolving documentary" needs a generalizable approach to description and presentation. Moreover, the audience needs to be able to dynamically steer the presentation so that it may remain relevant to their interests.

---

## **The Artery Story**

In accordance with a belief that computational systems for information or entertainment are best developed in parallel with model content, we began to build our own "evolving documentary." Using our Workshop in Elastic Movie Time to garner a group of student journalists, we began to construct "Boston: Renewed Vistas." This evolving documentary on urban change in downtown Boston will continue to develop through 2004 and possibly seed stories in other cities.

Boston's Central Artery stands as "a green monster," a relic of highway construction in the 1950's. Built to revitalize the downtown, this large iron, rusting structure visually separates the commercial Faneuil Hall area from the North End, Boston's most protected historic neighborhood.

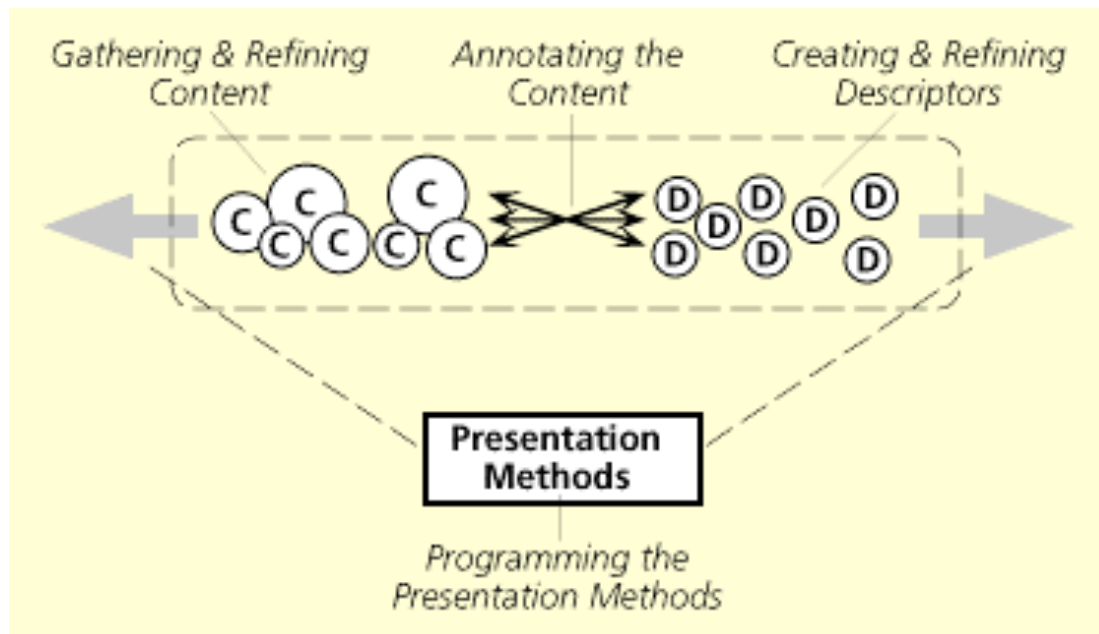
Throughout the 1980's, a group of politicians devised a plan to replace this overhead highway with 37 lane miles of underground roadway. The estimated cost of this project is 7 Billion Dollars, most of which comes from Washington under the auspices of the Interstate Highway bill. This is currently the largest public works project in the United States. Several neighborhoods, including Boston's North End, will be affected in ways we do not yet know by these developments.

Currently MIT students are following a variety of stories in different media. These include portraits of particular places and people, interviews with residents and politicians, coverage of events and meetings. Stories are developed with more or less depth depending on interest, but the stories crisscross with a rich variety of themes. In addition to original material, the project uses materials from the Boston Globe and historic photographs. Some sequences are edited with a particular intent relative to narration. Shorter segments of material, edited by individual authors, are described according to an established set of descriptors and the general editorial guidance of the principle software designer Mike Murtaugh.

---

## The Author's Activities

In the model of the evolving documentary, the author's activity also evolves. The author may still function as a journalist/editor working directly with the materials. However, in addition to these tasks the author must create a description space which contributes narrative coherency to the materials. While the content model of an "evolving documentary" and the corporate publication structure may vary, there is a clear editorial function which consists of critically evaluating the boundaries and style of appropriate content and in parallel mapping the types or categories which make up the content/description space. As the "evolving documentary" develops, this role may be complicated by the existence of multiple authors, a growing collection of characters, and the emergence of new themes. Figure 3 shows an overview of the author's activities in an evolving documentary.

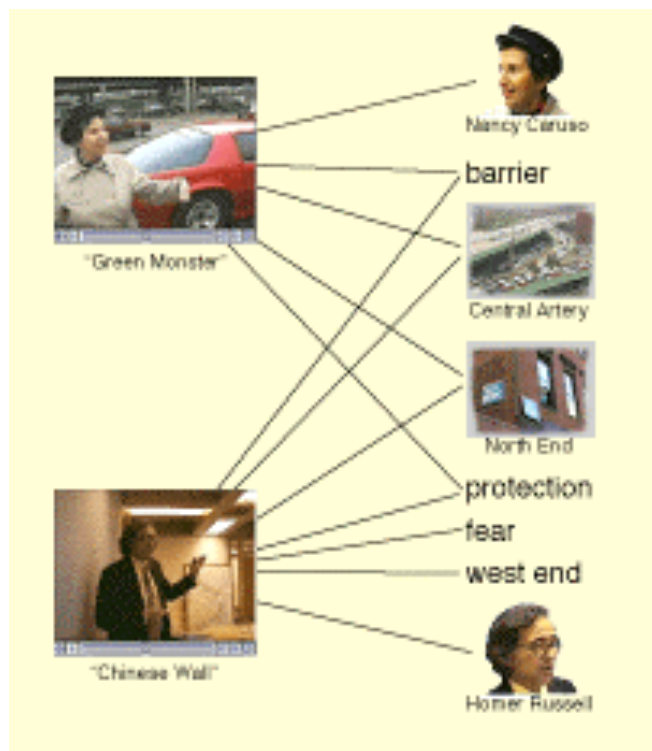


**Figure 3: The Author's Activities**

### Annotating the Content

The first step in building a ConText database after gathering a pool of content is to establish the set of descriptors. From the journalist's tradition of who?, what?, when?, where?, and why?, we conceptually divided our descriptors into the categories: character, time, location, and theme, where theme becomes some combination of what? and why?. The ultimate goal of the description process is to extract the complete set of abstract ideas or elements relevant to the big story. Too few descriptors limit the ability of the browser to group like elements properly, too many or overly specific descriptors may form clumps of content difficult for the browser to smoothly bridge. In sum, the creation of appropriate descriptors is a crucial part of the authoring process and vital to the proper functioning of the presentation system.

The second and final step is to describe the content using the established set of descriptors. The knowledge representation used in ConText is a simple bidirectional mapping between units of content and units of description. That is, the author associates a set of relevant descriptors for each piece of content added to the system. As the author connects each piece of content with sets of descriptors, they also define a mapping from descriptors to sets of content. Thus, for each descriptor, we know the set of content units that it describes. Note that relationships between content and descriptors are unqualified; the links are not weighted. Thus the core representation is equivalent to a simple keyword system. Descriptor weights, or prominences, are a function of payout during the ConText browsing session as described in the following section on Presentation Methods.



**Figure 4: Annotation of Content with Descriptors**

Figure 4 shows two video clips in the left column and their associated descriptors in the right. The first is of North End resident Nancy Caruso describing how the Central Artery currently serves as a "green monster," dividing the North End from the rest of the city and protecting the community's residents from outsiders. In the second clip, urban planner Homer Russell expresses a similar idea about the Artery functioning as "a sort of Chinese Wall" or barrier. In addition, he notes how the North End community was justifiably frightened based on the experience of the West End, a community once like the North End that was leveled and replaced by highrise buildings in the 1960s.

Authors currently use a separate tool to define these relationships and establish graphical representations for descriptors (either text or a picture file). This information is then saved out in a format that the ConText browser reads to begin a browsing session.



## **Presentation Methods**

Evolving the extensible base of content requires the isolation of sequencing decisions from the content itself. In ConText, we define a core set of simple methodologies based on fundamental principles of storytelling. Each method must operate solely on the content and description space as established by the author. In addition, the methods may rely on any additional state introduced by the presentation. The goal is that when used in combination, these simple programmatic pieces will result in a sequencing of content with narrative coherence and meaningful structure.

## **Continuity**

Continuity is a fundamental principle of cinematic storytelling. In a sequential presentation of content, meaning is derived from the way in which individual shots and sounds are connected to previous shots and sounds. The important function of story progression is that the viewer can construct a meaningful sense of the whole. [7] In ConText, the method for ensuring continuity, called description feedback, forms the foundation of the content sequencing mechanism.

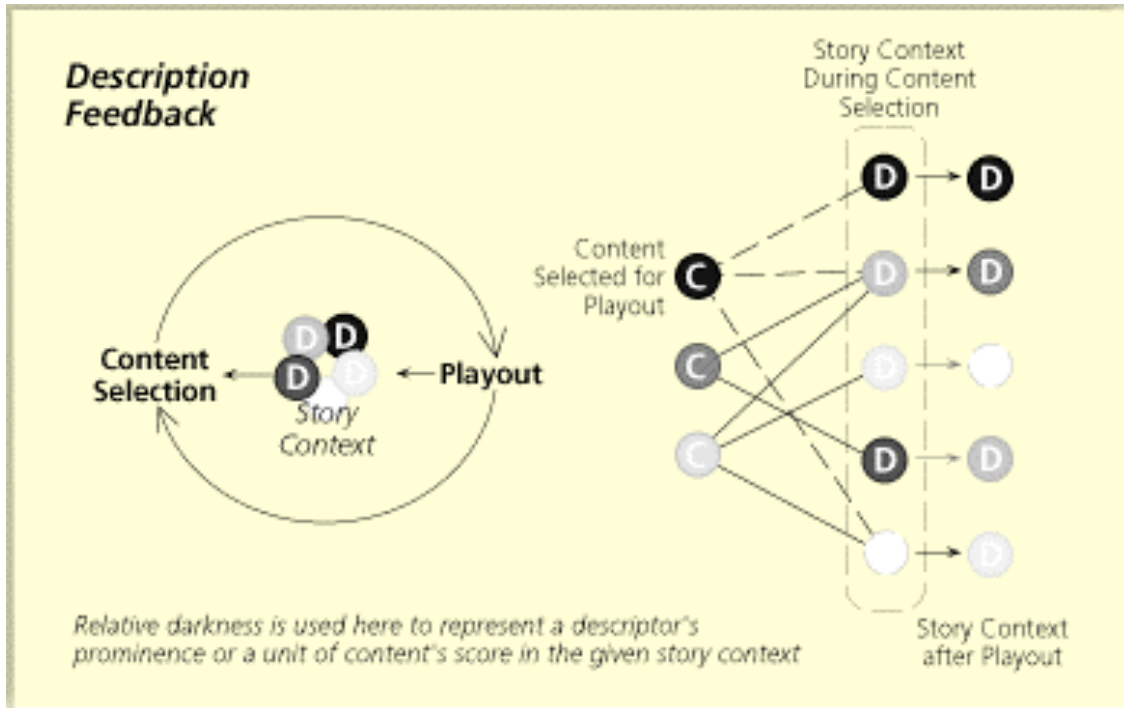
As its name is meant to underscore, a central property of ConText is the idea of a story context. A story context is defined to be a set of descriptors where each descriptor is qualified by a numerical value. That value, its prominence, represents the level of that descriptor's importance within the given story context. For example, a story context including the character descriptor "Nancy Caruso" with a prominence value of 100, the location descriptor the "North End" with prominence 80, and the thematic descriptor "Protection" with prominence 20, represents a moment in the story where Nancy Caruso and the North End are quite prominent while the theme of protection is only slightly prominent.

When the system is required to select a unit of content to be displayed, the choice is made based on the current story context. Excluding those already displayed, each content unit is assigned a score equal to the sum of the prominences of its associated descriptors. The unit with the highest score wins and is selected, and is later marked as having been played. If there's a tie, the choice is made at random from among the high-scorers. Thus, given the story context above, a video clip associated with the "Nancy Caruso" descriptor and the "North End" descriptor would have a score of 180, while one tied to the character "Fred Salvucci" and the theme "Protection" would have a score of just 20 (since "Fred" has a prominence of 0). Given the choice between just these two clips the system would select the first.

The final crucial component to this system is that as the selected unit of content is presented to the user, the prominence value for those descriptors associated with the content is increased while the prominence value of all other descriptors is decreased (unless already zero). Thus the selection of a given piece of content influences the story context in a way that makes the future selection of similarly described content more likely. This property is called description feedback because of the way the sets of

descriptors associated with a selection "feedback" on the selection process by influencing the story context. In this way, content acts as a bridge between story contexts. Figure 5 shows an overview of this process.

As further clips are selected and presented, the story context continues to develop with repeated elements rising in prominence while others recede. Thus, at any point, the story context reflects the path the viewer has taken during their particular browsing session. Later, in the section, "The Interface," an actual progression of story contexts from the Artery story is shown and described.



**Figure 5: Description Feedback**

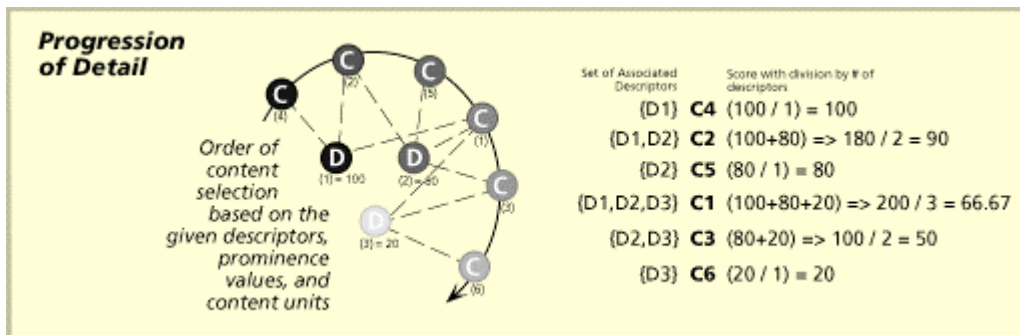
## Progression of Detail

Another fundamental principle of storytelling that governs the sequencing of content involves level of detail. Typically, general information about the components of a complex idea are presented as "background" or "context" before the presentation of that idea. By establishing simpler ideas first, the storyteller is able to establish a foundation or context against which more complex ideas may be set. This progression from simple to more complex may also be seen as a progression from general to more specific and detailed content.

Given the mechanism described above for continuity by description feedback, a progression of detail is surprisingly easy to implement when one makes the observation that: The number of descriptors associated with a unit of content is representative of the

content's specificity. To favor relatively general units of content before more detailed ones, we simply need to divide the unit of content's "score" by the number of its associated descriptors. Thus we "penalize" content for having a large number of descriptors and cause less heavily described content to be presented first. The effect is relative in the sense that as the "general" content is played, it's removed from the pool of potential content. The previously penalized pieces then become relatively general to the content still available for payout that is even more heavily annotated. Thus, the movement from general to specific is truly a gradual progression.

As an example of this principle, consider a piece of content associated with only the character descriptor for "Nancy Caruso" and the location descriptor for the "North End." Such a piece of content might simply be audio narration stating that "Nancy Caruso is a resident of the North End." In contrast, a clip of Nancy Caruso talking about the North End in the 1950s would also be tied to the time descriptor representing the "1950s" and the theme "Memory." This clip is clearly more specific and has more relevance when presented after the former "establishing" material.



**Figure 6: Progression of Detail**

Figure 6 shows a progression of content based a given weighted set of descriptors representing a story context. Note how the progression builds from general to more specific content while featuring those units of content associated with the more prominent descriptors.

The above described relationship between the number of descriptors and the level of specificity is actually an equivalence. If one asks what exactly "specific" means in this context, the answer is it means specific with respect to the description space created by the author. Thus, if we had a video clip of Nancy Caruso comparing the North End to Little Italy in New York, the clip's annotations would still consist only of "Nancy Caruso" and the "North End" if our description space didn't have descriptors for "Little Italy", "New York", or the idea of a "comparison." Thus, the apparently more general narration establishing Nancy Caruso as a resident of the North End would be considered equally general as the "Little Italy" clip with respect to our current description space. The point is, this measure of specificity is only as meaningful as the description space is complete. In this case, if Little Italy or New York were pertinent to the story, they ought to be added as descriptors; otherwise, the function of that piece of content in the current story is questionable.

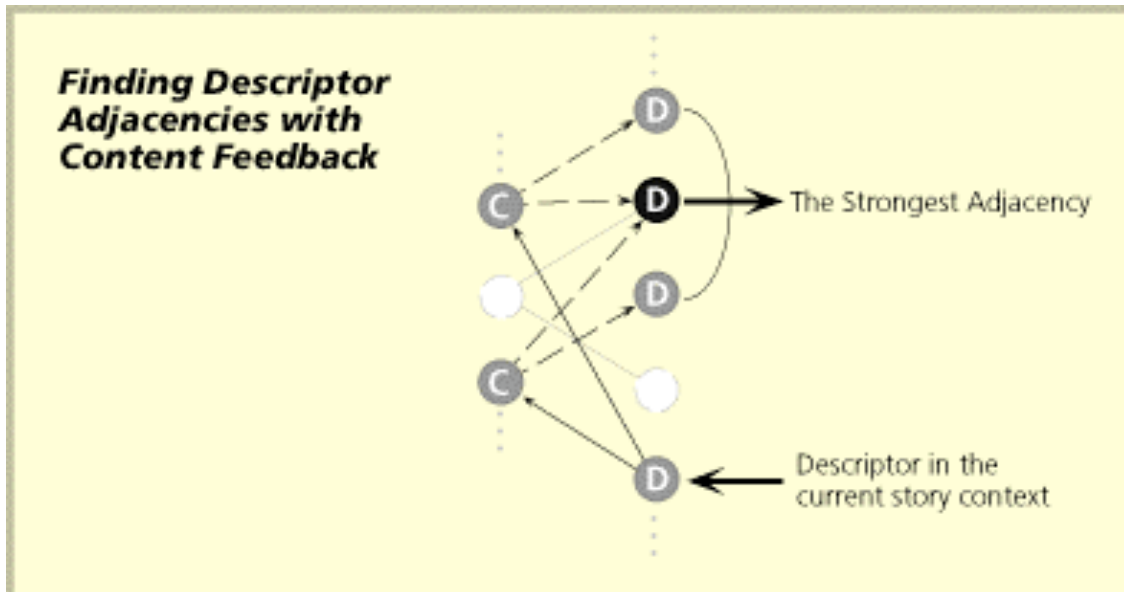
## Pacing

Frequently stories are thought of in terms of how "fast" or "slow" they seem to be moving. Far from referring to the literal speed of their presentation (the frame rate or the speed of the storyteller's voice), the pace of a story refers to the rate at which the content moves within the story's plot. Outside of the conventional Hollywood narrative, and especially in a form like the documentary, plot often plays a less prominent role. For our purposes, we make the simplification that pacing refers to the rate at which the presentation of content exposes or develops elements of the story. In ConText terms then, the pace relates to the rate of movement of the story context within the description space.

We've used the term description space repeatedly while in fact the term may seem rightly unjustified. Although we do have dimensions in terms of our categories of descriptors, directions along these axes or between any descriptors is not defined. In fact, implicit in the requirement of maintaining an extensible content base is that just as units of content must not be explicitly connected, neither may units of description. Thus, a relationship like "the North End is adjacent to the Central Artery" must not be explicitly represented in the database. However, such a fact is quite relevant especially when one wants the story to "pick up the pace" and possibly move away from a context including the Central Artery to adjacent locations.

Just as adjacencies of "continuity" are found between content based on their common connections to descriptors, meaningful adjacencies between descriptors may be found based on common connections to content. In the example given above, the existence of a unit of content annotated with descriptors for both the "Central Artery" and the "North End" would allow such a connection to be found. The content might be a video clip panning from the Artery to the North End, a visual expression of their geographic adjacency. Thus, the relationship between the two locations is available to the system from the content. If one wonders about relationships that exist between descriptors that aren't expressed by the content in the system, the fact is that they can't be captured. A unique property of the system is that only those relationships between descriptors articulated by available content could be used. In short, if the system can't demonstrate an idea, it can't know it.

By mirroring the structure and mechanism of description feedback described above, the system is capable of producing an analogous process of "content feedback" to explore possible movements within the space of descriptors. Given the set of descriptors with adjacencies to the those prominent in the current story context, the system can "increase the pace" of the story by making those adjacent descriptors more prominent than the current ones. In order to prevent immediate movement back towards the previous story context, a structure analogous to the "already shown?" tag used with units of content could be used with descriptors, capturing whether the descriptor had recently been invoked by the pacing mechanism. Figure 7 shows an overview of this process.



**Figure 7: Finding Descriptor Adjacencies with Content Feedback**

Finally, it is important to note that unlike continuity and the progression of detail which may always be active, this method for pacing requires the additional input as to when it ought to be invoked. One effective approach would be to simply give this control directly to the viewer, letting them decide if they wish to "stay in place" or "push the story forward." A second approach would be to start a ConText session with the pacing set to be relatively fast and gradually decrease it. This simple model corresponds nicely with the idea of allowing the viewer to explore the full breadth of the database first, then to dive in to depth as they find content of interest.

## Combinations

The hope is that by allowing the above described methodologies to operate in combination, a narrative coherence and structure will emerge from the sequencing of the content.

One notable property of the combination of the mechanism for controlling pacing and that of the "progression of detail", is that the control of pacing from "slow" to "fast" becomes the same as control from "depth" to "breadth" exploration of the content / description space. When the pace desired is slow, and the pacing mechanism is not invoked, the selection proceeds normally, with progression of detail tending to move from general to more specific -- exposing depth. When a faster pace is desired, however, adjacent descriptors are made more prominent than those of the current story context. The result is that by progression of detail, more general content relating to the newly invoked adjacent concepts becomes favored. Thus, movement is now decidedly upward and sideways -- exposing breadth.

By using a technique of "specializing by description", one can imagine applying the influence of the above techniques in more directed ways for specific story functions.

Imagine if we add a "weight" value in addition to the prominence of each descriptor in a story context. By using the multiplication of the descriptor's prominence by its weight instead of its prominence alone, one could imagine making certain types of descriptors more influential. For example, we might make location descriptors have a higher weight than the others to make locations the "focus" of our story.

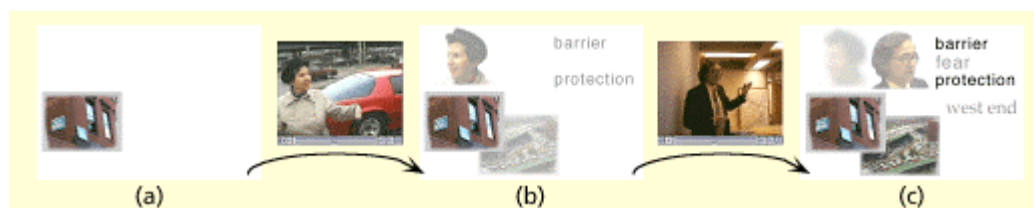
In a similar way, one can imagine applying the "pacing mechanisms" only to a certain subset of the descriptors. For instance, you could "increase the pace" of the character descriptors to cause movement across characters while other descriptors remained relatively stable. Using this technique in combination with the "location weighting" example given above would result in a story primarily about locations as told by many characters.

---

## The Interface

Given the ideas of the underlying methodologies, ConText's interface is surprisingly straightforward to describe and understand. The interface screen is a visual representation of the current story context. Each descriptor has a visual representation of itself, either in the form of text, a picture, or a combination of both. The descriptor's prominence in the current story context is represented by levels of brightness and focus. [8] Thus, the visual representation of a descriptor with prominence 100 would be bright and in sharp focus, while one at 50 would appear dimmer and slightly blurred. A descriptor with prominence 0 wouldn't be visible. The positions of each descriptor's visual representation is static and preset by the author. In sum, at any given time in a browsing session the interface displays a collage of pictures and text representing the current context of the story.

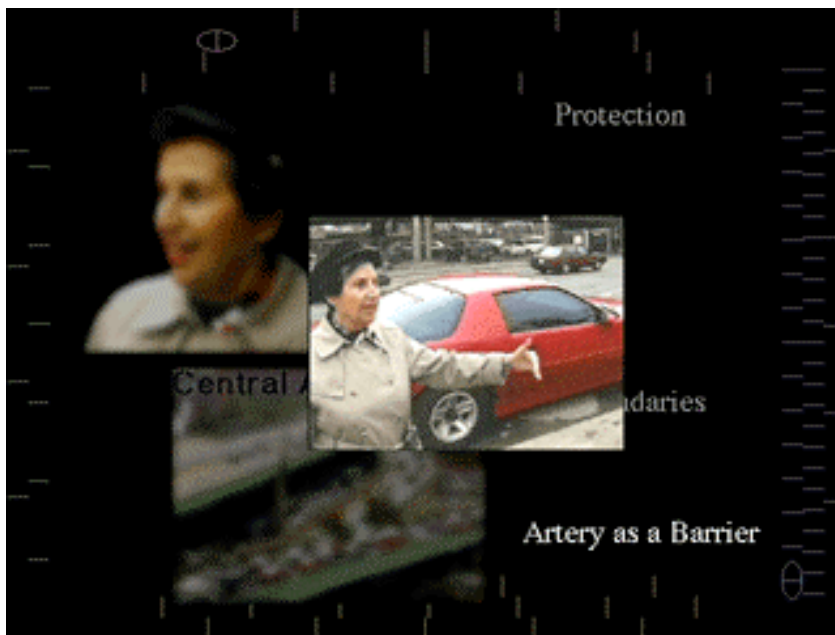
Units of content are displayed centered on top of the collage. The gradual influence a given content unit has over the story context is shown visually as the content appears on the screen. Thus, as a video clip plays out, the viewer sees its connected descriptors becoming brighter and more in focus while other descriptors fade away from view. An interesting property of this playout structure is that the longer the content is active, the more influence it has on the story context. Thus longer movie clips or pictures held on the screen by the user are more influential than shorter clips or materials that the user quickly dismisses (see "The Viewer's Activities" below).



**Figure 8: Progression of ConTexts**

Figure 8 shows a progression of three story contexts. In context (a), a picture representing the North End is the sole prominent descriptor. Recall the "Green Monster" and "Chinese Wall" clips and their associated descriptors shown in figure 4. Given this context and assuming neither clip has been seen by the viewer, both would have a non-zero score because of their shared connection to the "North End" descriptor. In accordance with the "progression of detail" methodology however, "Green Monster," with its five associated descriptors, has a higher score than "Chinese Wall" with its seven descriptors. Context (b) shows the effect of the "Green Monster" clip's playout, as the character "Nancy Caruso," the location "Central Artery," and the themes "Barrier" and "Protection" have become more prominent. Given this new context, the "Chinese Wall" clip has an even higher score than before and thus becomes the next clip selected for playout. Context (c) shows the resulting context after "Chinese Wall" plays out, the repeated themes of "Protection" and "Barrier," as well as the locations "Central Artery" and "North End" have become quite prominent. In addition, the descriptors for the theme "Fear," the location "West End," and the character "Homer Russell" have each become more prominent while the character "Nancy Caruso" has begun to recede.

The final component to the interface is the existence of subtle "tick marks" running in lines around each of the four edges of the interface display. Each mark represents a different descriptor and is arranged in one of four color-coded groups along each screen edge. The four groups correspond to the four categories of descriptors: Character, Theme, Location, and Time. These marks give the viewer immediate access to any of the descriptors in the system, regardless of the current story context. Figure 9 shows an actual screen shot from the current prototype. The central image is the current frame of the "Green Monster" clip as it plays out. In the background, the clip's associated descriptors gradually change to reflect their rising prominence in the current story context.



**Figure 9: Screen Shot from the Browser**

## **The Viewer's Activities**

The interaction given to the viewer is simple but powerful. The user is able to guide the presentation by altering the context of the story. When the user moves the mouse cursor over a visual representation of a descriptor, that descriptor is assigned a very high prominence, making it the most prominent element of the story context and increasing the likelihood of the selection of content associated with it. By clicking on a descriptor's representation, the user "dismisses" it, immediately setting its prominence to zero thus removing its presence from the interface collage. Moving the mouse over one of the tick marks along the edges of the screen is another way to make the associated descriptor the most prominent element of the context, regardless of whether it had been visible. Finally, clicking on a descriptor's tick mark "locks" that descriptor in place, disabling any change in its prominence.

Returning to the scenario described in figure 8, after viewing the clip "Green Monster," the viewer might have chosen to move the mouse over the character "Nancy Caruso." This action would steer playout towards further content featuring that character (if any existed) and away from the "Chinese Wall" clip.

Currently, the viewer is given very coarse control over the playback. They may "pause" playout by moving the cursor over the material (or in the case of an audio clip, the center of the screen), and they may "dismiss" the material, stopping playout immediately, by clicking the mouse button.

## **Continuous Playout**

A top priority in the design of the interface and presentation engine was that the viewer not be required to interact for content to be presented. Too often in past and current multimedia, the user is asked to continuously push the content forward by consciously interacting. These interactions tend to distract the viewer from the experience of the story and threaten to weaken the reception of the narrative. In addition, the requirement that the browser allow access to those with little or no knowledge of the story, makes the potential for continuous uninterrupted playout vital.

In ConText, content is presented when the system detects idleness from the viewer. Thus, the story moves only when the viewer stops interacting. Content continues to be presented until the viewer stops it by clicking or moving the mouse over the interface to alter the story context. In this way, interaction in ConText follows the model of a one-sided conversation. As the story is told, the viewer is "passive" and attentive to the narrative. Only when the viewer wants to change the course of the presentation does he or she intervene, asking to delve into a particular aspect in more detail or perhaps urging to move on to something else.

Returning once more to the scenario depicted in figure 8 given the final state shown as context (c), the viewer may simply not interact, most likely resulting in more content related to the Artery and the North End and the themes barrier and protection. If,



however, the viewer chooses to intercede by moving the mouse over the emerging location "West End," the story would move towards further content describing that area and any related elements.

---

## Conclusion

The model described in this document has been implemented in the current working version of the ConText browser. Based on our research to date, the evolving documentary seems a viable and productive alternative to the current linear production model. In addition to allowing a new viewer to access a database they may know very little about, the browser has proved quite valuable for an author to begin to see possible connections between content, particularly when the system contains content from a several authors and sources. The ability for an author to quickly drop new content into the system and immediately witness the resulting playout is a prominent benefit. In the end, the success of our approach will be evaluated as we continue to develop our content.

The section on presentation methodologies raises the idea that by carefully weighing the influence of each methodology, the author could invoke specific types of story structures. In order for this to occur, the author must have some way to describe the operation of these structures, as well as the means for specifying when their use is appropriate. In order to adhere to the constraints of our extensible form, all of this must be done in a generalized way. In addition, our experience with the Artery story shows that individual authors need leadership when dealing with issues of content granularity and description. Limiting clips to a length of approximately 30 seconds, for example, was found to be helpful to both the description process and the resulting playout experience. Despite our attempts at "normalizing" the description space with four axes, we find that the task of maintaining the description space still grows as content is added. A more flexible and dynamic means of annotating content would be of great assistance. These problems will be the subject of future research.

---

## References

1. Lippman, Andrew. The Distributed Media Bank, in Proceedings of the First International Workshop on Community Networking, San Francisco, July 13-14, 1994.
2. Houbart, Gilberte. Viewpoints on Demand: Tailoring the Presentation of Opinion, MS thesis, MIT, 1994.
3. Davenport, Glorianna, Thomas G Aguierre Smith, Natalio Pincever. Cinematic Primitives for Multimedia, IEEE Computer Graphics and Applications, pp. 67-74, July 1991.

4. Davis, Marc Elliot. Media Streams: Representing Video for Retrieval and Repurposing, PhD Thesis, MIT, 1995
5. Davenport, Glorianna, Ryan Evans, Mark Halliday. Orchestrating Digital Micromovies. Leonardo, Vol. 26, No. 4, pp.282-288, 1993.
6. Davenport, Glorianna and Lee Morgenroth. Video Database Design: Convivial Storytelling Tools, Interactive Cinema Technical Report, MIT, May 1994.
7. Branigan, Edward. Narrative Comprehension and Film, Routledge, New York, 1992, especially "Chapter 1: Narrative Schema."
8. Colby, Grace, and Laura Scholl. Transparency and Blur as Selective Cue for Complex Information, in Proceedings of SPIE'92. 1992.