THE COMBINATORICS OF STORYTELLING: *MYSTERY TRAIN* INTERACTIVE

Amy Bruckman
Interactive Cinema Group
The MIT Media Laboratory
April 1990

THE COMBINATORICS OF STORYTELLING: *MYSTERY TRAIN* INTERACTIVE

What is an interactive story? The traditional idea of a “story” is linear - it has a beginning, a middle, and an end. There is an apparent contradiction in the phrase “interactive story,” because non-linearity [1] is essential to interactivity.

It is difficult to imagine what a non-linear story might look like. Some modernist fiction can be viewed as non-linear. In a class on James Joyce's **Ulysses**, one professor advised his students to approach the work non-linearly saying, “when you get tired of Stephen's arrogance, read a little of Molly or Leopold.” [2] He was telling his students to guide their reading not by the order of the page numbers but by their desires. In a non-linear work, the viewer must be guided by his/her model of what he/she would like to know. This explains why many viewers find multimedia interesting only if they were already interested in the basic subject matter.

Non-linear writing is often difficult to understand; **Ulysses** is accessible to a limited audience. Interactive narrative need not be elitist. The designer must provide the viewer with sufficient tools for orienting himself/herself within the work. Interactive multimedia has the power to exist on multiple levels, making it accessible to a diverse audience. A successful work reveals more on successive explorations. [3]

Non-linear is not the same as unstructured. A work without any structure becomes a database --a laundry list of available information. The alternative to a linear story model is to use a combination of spatial, temporal, and thematic mapping to construct a *storyspace.* The viewer becomes an explorer of this storyspace. [4]

To date, interactive multimedia has been more successful with documentary subject matter than narrative. Documentary lends itself more naturally to non-linearity; the viewer is exploring a body of factual information, and can simply follow his/her interests. Narrative poses additional challenges. This paper explores some of those challenges.

---

[1] While an interactive story may offer more or less linear options, the viewer's experience of a work is always linear. The term “poly-linear” is perhaps more appropriate than “non-linear.”

[22]] Dr. Robert Keilly, Harvard University, 1986.

[3] Post-Modern art-historical theory discusses the value of works having multiple levels in a figurative sense. Interactive. multimedia can *literally* have multiple levels.

[4]] The viewer may be a constructor as well as explorer of the storyspace. The distinctions between viewer and maker, documentary and narrative begin to break down in interactive media.

COMBINATORIAL EXPLOSION

Rupert Holmes' play **The Mystery of Edwln Drood** creates interactivity within the traditional, linear story model by inviting the audience to select from a variety of possible endings. This type of interactivty is limited by the problem of combinatorial explosion.

Drood is based on Charles Dickens' unfinished novel. The audience "finishes" the work by voting on three questions: Is Edwin Drood dead? If so, who of the seven principals murdered him? [5] Who is masquerading as Dick Datchery? [6] The cast prepares for each possible ending -- which is made possible by the fact that there are a limited number of outcomes. Thus, the story is shaped roughly like the diagram on the following page.

The identity of the murderer and the identity of Dick Datchery are independent questions. Princess Puffer has a song she sings if she is declared the murderer. It does not matter who is disguised as Dick Datchery; she does not have six different songs. [7] There are twenty-two songs which can be combined to make a total of fifty possible endings (eight if Drood is alive; forty-two if he is dead). This is a technique to reduce the total amount of information needed. If music were never reused between endings, then there would be at least ninety-two songs (eight if Drood is alive; forty-two times two if he is dead) instead of twenty-two. Consider what would happen if the audience were allowed to vote one more time; for example, suppose that when the show opens the audience votes on who disappears - is it The Mystery of Edwin Drood or is it The Mystery of Princess Puffer? Then there would be 738 songs needed (eight times ninety-two). The rate of growth is exponential.

This style of interactivity can be called the *broomstick approach;* it consists of a linear story which branches into a number of alternative endings. The broomstick approach is used in many children's interactive fiction paperbacks, such as the "Choose Your Own Adventure" series. An analysis of number eighty-four, **You are a Monster**, [8] appears on the next page. After reading page 1, the reader is told to go to 65. The bottom of 65 tells the reader to go to 109. This hopping around serves no purpose except to increase the appearance of interactivity. After seven hops the reader Is finally given a choice: "If you come out of hiding, turn to page 87. If you stay hidden, turn to page 14." Each path through the story contains two to five questions. The growth is exponential, but the numbers stay manageable because each question offers only two alterative. A series of five questions each with. two alternatives would lead to $2^5 = 32$ endings. In fact, many paths terminate with fewer

[5] Rich, Frank. "Drood, a Musical in the Park." The New York Times. New York: August 23rd, 1985; C:3:3.

[6] Gussow, Mel. "The Park Drood: Spine-Tingler Turned into Rib-tickler." The New York Times. New York: September 1st, 1985; 2:3:1.

[7] I believe that it was not allowed for the same person to be both the murderer and Dick Datchery.

[8] Packard, Edward. **You are a Monster**. New York: Bantam Books, October 1988.

than five questions, so there are only thirteen endings. [9] With five questions and seven alternatives per question as in **EdwIn Drood**, there would be 7^5 = 16,807 endings. Even limiting the story to two alternatives per question, with ten question there would be 2^10 = 1024 endings. This broomstick style of interactivity is limited by combinatorics.

INFORMATION CONSERVATION

You are a Monster limits the amount of information needed by having some paths rapidly lead to dead ends. Another approach is to introduce *looping* - whatever the participant decides, he/she will end up in the same place in a few pages or segments. Looping is used in the videodisc game Maze Mania. [10] The game consists of a mixture of short video segments and trivia questions (on still frames). Correct answers advance you through the story. Incorrect answers send you on a detour. If you complete the detour question correctly, then you move ahead; otherwise, you are sent back.

There is one central story line. Interactions may send the viewer on a detour, but he/she will always end up back in the central story line. The basic technique of looping is described in the diagram on the following page.

Looping is common in behavior-modeling interactive videos. The participant is shown part of a conversation, and then is given a list of possible replies. If he/she picks the “wrong” choice, he/she is briefly shown what would happen, and then is thrown back on the central path with a statement like ”Let's see what would've happened if you had selected response B”.

For interactive training, looping serves a purpose -- it reinforces the correct answer. (It also assures the client who paid for the interactive training that every student or employee will see every precious bit of training that has been paid for.) This faked interactivity becomes tedious. The participant gradually discovers that no real choice is being offered, and may feel manipulated.

However, there are useful lessons to learn from this technique. In narrative, part of what makes looping possible is the interchangeability of certain story elements. It doesn't matter whether you find the blaster or the space suit first - as long as you have both before you meet the space squid. And it doesn't matter at all whether you find the diamond in the space squid's stomach. Thus, there are:

- events which are obligatory but can occur in a random order, and there are
- events which are optional.

[9] One question has three possible alternatives.
[10] Maze Mania. New York: Optical Programming Associates, 1982. This laserdisc game has no computer software. Instead. the player is instructed to jump manually to disc frame numbers.

Both of these types of events are used in what I will call the *cocktail-party approach* to interactive narrative. In **The Name Game**, a videodisc program by MindBank Inc., the participant is invited to stop a spy plot. To stop the plot, you must obtain clues from people at a cocktail party. Clues are earned by remembering people's names. At any time, you may leave the cocktail party to open your attaché case, which contains "agent training" -- a set of short tutorials on tricks for remembering names. [11] You may approach anyone you see at the cocktail party. There are many that you don't need to speak to in order to solve the mystery, and most can be approached in any order. This is diagrammed below.

The result successfully simulates the feeling of being at a cocktail party. **The Name Game** is entertaining. (I cannot comment the effectiveness of the memory tricks it teaches, but it does reinforce the importance of learning names.) The cocktail-party style departs from a linear story model.

FORM AND CONTENT

The laserdisc game **Murder, Anyone?** uses aspects of both the broomstick approach and the cocktail party model.[12] **Murder, Anyone?** is played by two teams. In an opening video segment, the character Derrick Reardon is murdered. The object of the game is to guess the murderer, the motive, and the method. After the introductory video, the first team selects either a one or a two. No significance is attached to the choice: players select a number randomly. There are a total of four such decision points. This creates sixteen different plots, allowing the game to be played multiple times.

The sixteen plots are not represented by sixteen different video presentations. The designers used the technique of voice-over narration to conserve space on the disc. Players are instructed to listen to audio track one if they have chosen story one, or audio track two for story two. For example. here are the two different audio tracks for a scene in which the private detective, Stew Cavanaugh, is seen sitting in the library interviewing the family physician, Dr. Theodore Morfield:

Audio track one:

Detective Cavanaugh (voice over):

> If you like your doctors slightly seedy, then Theodore Morfield Is right up your alley. He had spent most of his years tromping through the jungles of South America, but now he was staying with the Reardons, a kind of Ilve-in medico) only the rich and the sickly can afford. It seems that Derrick was both.

---

[11] This is an interesting example of the use of a physical object as the rationale for a branch point. Instead of "entering tutorial mode," the participant "opens a suitcase," which keeps the action within the context of the narrative and adds a James-Bond-like flavor.

[12] **Murder, Anyone?** Cincinnati: Vidmax. 1982. This laserdisc game has no computer software. Players are instructed to jump manually to disc frame numbers.

[Dark, murky video of Dr. Morfield examining Derrick Reardon. Audio is voice over.]
Dr. Morfield:
I don' know why anyone would kill him. He had only six months to live anyway.

Detective Cavanaugh:
Who all knew about this?

Dr. Morfield:
His wife, his sister, his brother.

Detective Cavanaugh:
Was it you who diagnosed the disease?

Dr. Morfield:
Of course. He was supposed to have gone to New York today for a second opinion. I had insisted on it.

Audio track two:

Detective Cavanaugh:
Dr. Morfield, you live here, is that right?

Dr. Morfield:
It's convenient. Derrick Reardon was my only patient.

Detective Cavanaugh:
He was sick, eh?

Dr. Morfield:
A hypochondriac. Not to the point where it inconvenienced him; just other people.

Detective Cavanaugh:
Huuh. That's interesting.

Dr. Morfield: -
(chuckles) You mean suspicious? Well, I guess I had a reason for hating him.

[Dark, murky video of Dr. Morfield examining Derrick Reardon. Audio is voice over.]
Five years ago I had a thriving practice. Derrick was one of many. But he was possessive. I didn’t think it strange at the time. But one by one my patients all found other doctors. He seemed eager to take up the slack. I found out he was spreading rumors about my competence.

The same video footage is used to tell two different stories. One audio track is voice-over while the other is sound from the actual scene: then they switch. Murky flashback shots are used whenever it is necessary for both audio tracks to be independent of the video.

At first glance, the structure of **Murder, Anyone?** resembles the broomstick approach. It is diagrammed on the following page. At each decision point (except the first), players are instructed to skip to a different chapter of the disc depending on what sequence of numbers has been chosen so far. However, note that video segments E, H, and I can follow different preceding plot lines.

For example, segment E can follow a sequence of either 12 or 21. Not requiring a different video segment for each of these choices conserves space on the video disc.

The nature of the subject matter is what makes this technique work. The factual information presented in these sequences is different, as the excerpt above shows. Each video segment contains a set of clues. Many of the most specific clues are in the final video segments (G, H, I, and J), because no video follows them. The players put the clues together to form a theory of what has happened. Which clues are significant changes between different story lines. The creation of a whole out of loosely connected pieces draws on aspects of the cocktail party model.

The content of **Murder, Anyone?** is inseparable from its form. The mystery theme allows for information conservation; combinatorial explosion is curbed by reusing segments in different storylines. The game format transforms the lack of narrative coherence into a challenge for the players.

NARRATIVE COHERENCE

In making a story non-linear, the storyteller relinquishes the power to control the flow of information to the viewer. In a sense, the viewer becomes a co-author of the story. A balance must be struck between giving the viewer freedom and maintaining narrative coherence.

Consider the tree-climbing sequence from Flaherty's film **Moana**. We see a boy with his arms and legs wrapped around a tree trunk, creeping upwards. In a sequence of close shots, we watch his muscles tense and his feet slide upwards. Finally, the sequence cuts to a wide shot, and we see that the boy has climbed alarmingly high.

The power of this sequence is in the withholding of information. Consider the difference in impact if Flaherty had begun with a long shot of the tree -- the story would be robbed of its punch line. The impact would be greatly reduced. What happens when the storyteller gives up the power to control the flow of the story? Will the impact be reduced, because viewers will always leap to the punch line?

One solution is not to give up the power to withhold information. In a computer-controlled interactive story, the viewer can be prevented from accessing key pieces of information until a set of prerequisite places in the storyspace have been visited -- you can't see the long shot of the tree until you have seen the boy climbing. In **The Name Game**, there are certain characters you must meet at the cocktail party before you can meet others. Rules of this nature can help to maintain narrative coherence.

SPATIAL AND TEMPORAL MAPPING

The organizing principles which have been discussed are temporal; they use the sequence of events in time as structure. Freedom from a linear model of time enhanced interactivity in the cocktail-party model.

An alternative to temporal mapping is spatial mapping. The earliest example of this is **The Aspen Project**, done at the MIT Media Lab in 1979 to 1980. The participant is invited to "walk" around Aspen in different seasons. The area covered is finite, which limits the amount of information needed. By using sequences of still frames rather than running video, the amount of information needed was further cut down by a factor of 30. Short video "stories" were placed within certain buildings to add interest. The combinatorics of the problem were quite manageable - only one hour of video disc was needed. This style of interactive multimedia has come to be called "surrogate travel." No real story is told; the participant is simply allowed to explore a place.

The possibilities of both spatial and temporal mapping are limited. However, by using both simultaneously the author can create a *storyspace* for a true interactive narrative. Jim Jarmusch's film **Mystery Train** is ideal material for this approach.

MYSTERY TRAIN INTERACTIVE

Mystery Train is made up of three separate stories strongly linked both temporally and spatially. Spatially, characters of all three stories see the same statue, pass the same bar, and stay in the same run-down hotel. Temporally, the stories occur simultaneously. The temporal linkage is established by things that characters in all three stories hear: trains go by, the song "Blue Moon" played on the radio at 2:17 am, and a gunshot.

Imagine **Mystery Train Interactive**. At 2:17 am, the participant could find out what is happening in different places in the hotel. In front of the bar, the participant could move through time to see what takes place there. Thus, there are two basic ways to explore the story: to select a time and explore spatially, or select a place and explore temporally.

A third way to explore the story space is thematically. One could, for example, look for pictures of Elvis, or references to the legend of Elvis' ghost on the highway. Thus, the participant becomes a part of the same quest as the characters.

A variety of supporting materials would enrich the storyspace. Who has watched **Mystery Train** without wanting to learn more about Carl Perkins? Whenever a singer such as Elvis Presley or Carl Perkins are mentioned, the viewer would have access to a digital library of the artist's music. The library is supported by musical scores and historical analysis. [13]

Imagine a simple spatial map of **Mystery Train** -- that becomes surrogate travel of Memphis. Imagine a simple thematic map -- that becomes a database of Elvis trivia. A simple temporal map becomes a linear movie. Together, however, these three techniques create an interactive narrative. [14]

Something has been sacrificed in this scenario: the user no longer has the power to affect the outcome of the story. That feature could be added, but then the problem of combinatorial explosion would be reintroduced. Without the ability to affect the outcome, the amount of information available becomes crucial. The storyspace must be rich enough to engage the participant in exploration. This approach to multimedia works best when there is more information than the user could possibly peruse in one sitting.

COMPUTER UNDERPINNINGS

How would one construct **Mystery Train Interactive**? At the heart of the creation of hypermedia is the construction of links. The standard approach is to make each link by hand. In other words, looking at a scene in **Mystery Train** in which the bellhop swats the large toy fly, the author could sit and decide what links would be relevant. The author would probably give the viewer the option to see other scenes in which the bellhop plays with the toy fly. Linking spatially, scenes which take place at the hotel desk at other times should be accessible; linking temporally, what is happening to other characters and in other places at that time. But what of other shots that use the same camera angle? What of other shots of the bellhop? The author would need to evaluate a large number of possible links and select the subset which is most relevant.

An alternative approach is to give the system some knowledge about the material in it. A *knowledge base,* a type of database developed with techniques from artificial intelligence, could be the foundation for the system. A database stores data. A knowledge base uses a more sophisticated approach to representing information, and is capable of keeping track of complex relationships between data. How best to represent information of this nature is a current topic of research in artificial

[13] The end result is neither “documentary” nor “narrative”. Interactive technology begins to break down these traditional categories, which were developed for linear film.

[14] As is, Mystery Train contains sufficient material to be the basis for interactive multimedia. However, one is tempted to wonder, could more twists to the story be added? Are there scenes on Jim Jarmusch's cutting-room floor that could be added to the interactive version? Extra footage could be added within the framework of the spatial, temporal, and thematic maps but not appear if the participant should decide to view it linearly.

intelligence. One approach is called a "frame system" or "semantic net." A partial description of a shot from Mystery Train appears on the next page.

The author working with this scene would not need explicitly to select which links to include. Instead, he/she would simply describe each shot, and enter this information into the knowledge base. The system would contain a set of simple rules of inference and could construct links based on knowledge of the material [[15]].

This makes it easy for one to add new material to the system. The author doesn't need to search through the entire database to find the relevant places to construct new links, but can simply put information about the new entry into the knowledge base. The system simply needs to be told that in this new entry one sees the bellhop swat the toy fly. The system itself makes the connection that this might be relevant to the scene where the Japanese tourist swats a live fly. This approach is particularly suitable for applications which change frequently such as news archives or tourist information kiosks.

Notice in the diagram that "to swat" is a kind of "to hit." This allows "to swat" to inherit properties from "to hit." The system now knows that swat needs a subject (who is doing the swatting}, and two objects (what is being swatted and what the swatting is being done with.)

The process of describing (or "logging") an entry in the knowledge base often reveals new information. In making the diagram on the previous page, I added "silly" as a description of the appearance of the Bellhop's uniform. It then occurred to me that the Hotel Manager's suit is also rather silly. This is not a trivial point. The scene would have a different impact if the Hotel Manager were wearing an expensive, conservative, banker-style suit. The change in costume would effect the perceived relationship between the characters.

The knowledge-base approach makes it easier for personal annotation to be included on top of the public knowledge base. Thus, the cinematographer perusing *Mystery Train Interactive* might add a *slot* for camera angle to many of the frames, write text comments, and draw sketches on the use of camera angle. A later user could choose to load in the cinematographer's annotations.

The most important added power of a knowledge-base approach is that the viewer can define his/her interests. One viewer might want to move through spatial links only, exploring Memphis. The cinematographer might want to examine shot properties. An Elvis fan might follow the theme of Elvis trivia either spatially or temporally through the story. Each of these themes is like a thread in a complex web. When more threads are activated, the system begins to behave like a database of information. When fewer threads are activated, it begins to behave more like a linear film.

---

[15] In an ideal system, a set of image and audio processing tools would give the system some knowledge of the content of the scene automatically.

GRANULARITY

What is the smallest unit of video that can be manipulated -- a shot, a sequence, a scene, or a story? At the shot level, if the viewer of Mystery Train Interactive moves from the bellhop swatting the toy fly to the tourist swatting the live fly, narrative coherence might be lost. Although there is a logical basis for linking these two shots, they do not form an understandable narrative.

Could the computer recognize that they do not form a coherent narrative? That would require higher-level reasoning. The computer would need a model of the story and a model of the spectator. By tracking what the spectator does and does not know, the computer could evaluate which links would be meaningful.

Let us suppose that the link has been determined to be relevant. For example, suppose that the participant is not a first-time viewer and has expressed an interest in studying the use of violent imagery. Is the link made at the shot, the sequence, or the scene level? A cut directly from the toy fly to the live fly makes the connection explicit, but would probably look awful and make little narrative sense.

To make transitions at the shot level, the computer would need a complete theory of editing. Could one create an expert system to understand gesture, eyeline, narrative coherence, suspense, and style? Possibly. First attempts would be awkward, and through progressive refinement the system might achieve the level of a human novice. Whether it could be made to surpass a human novice is a central question of artificial intelligence: can computers ever simulate creativity?

Without a computerized editor, one way to improve narrative coherence is to match the granularity of links to the user's level. For a new user unfamiliar with the story, cuts should be made at the scene level. A more experienced user might want to explore at the sequence level. Only the cinematographer interested primarily in shot properties should explore at the shot level.

A great deal can be accomplished simply by associating a list of keywords with each entry in a multimedia system. In what ways is semantic information superior? Practical benefits can be seen in larger systems and systems that change frequently. Beyond that, better knowledge representation is the cornerstone of more sophisticated tasks -systems that answer questions, or that guide the viewer's experience by keeping track of what the story is about and what the viewer knows so far. Exactly what representation to use depends on the nature of the task, and is a subject for future research.

ACKNOWLEDGEMENTS

Credit for many of the ideas in this paper belongs to Glorianna Davenport. I'd also like to thank Ken Haase and Henry Jenkins for reading drafts. Lastly, Josh Kirschenbaum helped me negotiate my way through the various software applications I used to create the drawings.

BIBLIOGRAPHY

Backer, David S. “Structures and Interactivity of Media: A Prototype for the Electronic Book.- Ph.D. dissertation, Massachusetts Institute of Technology, 1988.

Gussow, Mel. “The Park Drood: Spine-Tingler Turned Into Rib-tickler.” The New York Times. New York: September 1st, 1985; 2:3:1.

Maze Mania. New York: Optical Programming Associates, 1982.

Murder, Anyone? Cincinnati: Vidmax, 1982.

Packard, Edward. You are a Monster. New York: Bantam Books, October 1988.

Rich, Frank. "Drood, a Musical in the Park.” The New York Times. New York: August 23rd, 1985; C:3:3.