3 The Problem

The world has arrived at an age of cheap complex devices of great reliability; and something is bound to come of it.

Vannevar Bush, As We May Think, 1945

3.1 The Problem Addressed by the Thesis

3.1.1 The Thesis

The last chapter examined a few of the many ways people have pushed the bounds of various media in a nonlinear or metalinear direction. While those noted explorers have progressed and expanded the media's means of expression, some problems still remain. For instance, those examples cited in the previous chapter are bound to their fundamentally linear media. No matter how talented the filmmaker or engaged the audience, film remains a linear medium. This thesis shows that in order to write a metalinear story, one must use a metalinear writing tool from the very beginning. Writing using a standard word processor and a multilinear presentation tool cannot yield a metalinear result. One requires a special tool to aid appropriately in the process of building a metalinear story system.

In general, the computer has the ability to be this tool. Yet the problem remains; how? How can a computer take instruction from a story writer such that the computer can "know" something about the story? How can a computer provide feedback to the writer during the grueling rewrite process? How can a computer fit into the metalinear story design system which includes story creation, presentation and writer/audience feedback?

If a computational system is to make narrative sequencing decisions, it will have to have access to some portion of the author's intentionality. It will have to have some representation of how the author understands his/her own story. The metalinear narrative model accomplishes this by representing the authorial intentionality through various types of links. By connecting one piece of story to another, and stating a simple reason for this connection, like causality or temporal precedence, some amount of the author's intention is represented in the linking system. Additionally, by providing the author with feedback in the form of a story, a writer will know if the connections they made between story pieces can be navigated in a coherent way. A specialized software tool for writing metalinear stories can provide such functionality. It can help the author construct metalinear story by capturing the relationships between story pieces, providing a means for constructing narrative structure, and by providing ways to reconstruct multiple linear narratives from the same collection of narrative material. This chapter addresses the issues involved in creating such software by examining different approaches others used in this domain.

3.1.2 Overview of the Problems

Any storytelling experience or system, whether electronic or otherwise, typically includes three important elements: the artist/storyteller/writer, the story itself, and the audience. For a commercial system these functions might be called creator, product, and consumer. The artist creates the story, which is then realized in a specific medium and presented to an audience or consumers. Once the audience experiences this single instance of the story (book, theater, movie, etc.), there may be some mechanism for the audience to respond to artist. Such a system is modeled in Figure 10. While a book is fixed, with audience feedback possible only after the fact (through the publisher), feedback in live theater is a little more fluid. One example is that actors can often read the audience for needed adjustments during their performance. Additionally, applause during or immediately after the performance serves as some measure of artistic success for performers and writers. However the audience's feedback is usually given only after they have experienced the story – this is especially true in the case of movies. In general for all media, the audience cannot give feedback during the story in a way which will change any of the narrative structure or its production elements. As discussed in chapter 2, this model does not hold for oral storytelling, which has a much more fluid artist-audience

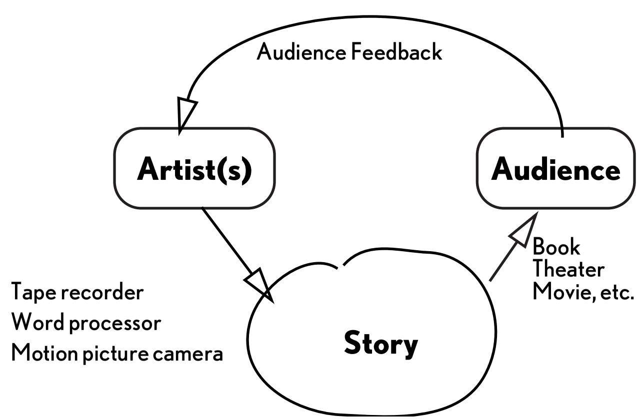


Fig. 10 Simple artist-story-audience structure, with feedback.

interaction. This model does hold true for all recorded media. For the benefit of reaching a wider audience through mass duplication, the recording process tends to distance the audience from the artist.

Is it possible for a computer mediated story to change this model? Computers have already been incorporated into the story writing and production process—most notably with the use of word processing tools. Word processors mainly offer a way of getting the story text written down quickly, and allow for fast and efficient editing. Beyond this common functionality, there are software tools that also give feedback to the writer based on some simplified knowledge the tools have about the written language. Spell checkers and grammar checkers have a representation of words and sentence grammar which allow them to use the written text as input for comparison with their rules. The software offers the writer feedback regarding where its representation of correctly spelled words or proper grammar does not match the writer's text. These actions compose a feedback loop between writer, the story/text, and the computer. However, these systems have little to do with the narration of the story and nothing to do with the audience. That is, they do not have any representation of what the story is about. Spell checkers and grammar checkers do not make the wolf bigger or badder, or Red a smarter child, but just "cleans-up" what the writer has done. Figure 11 illustrates such a system.

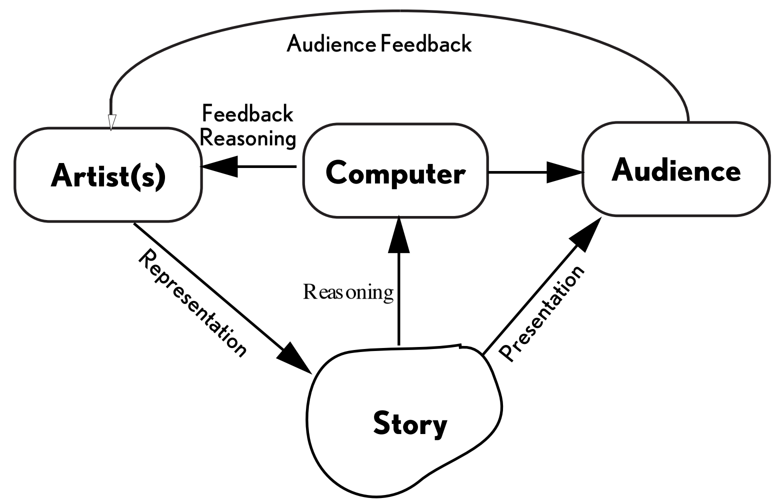


Fig. 11 Computer assistance in the simple artist-story-audience structure.

Many researchers and software designers have tackled the problem of trying to place the computer within the creative process in a more meaningful way. These attempts have usually not been directed toward general solutions, but instead toward application specific solutions. This chapter examines some representative examples which demonstrate the limitations of currently available systems.

3.2 Problem Examples

3.2.1 The Knowledge-based Approach

Figure differs from Figure in that there are paths marked with the words *Representation*, *Presentation*, and *Reasoning*. *Representation* conveys the existence of story content as well as story description. The story is described to the computer in a way which allows the computer to understand it enough to facilitate simple manipulation. (Brachman & Levesque, 1985, p. xiii) *Presentation* conveys that the form of media for the story is not predetermined. That is, the story exists first in an amorphous, unrealized state. It is the presentation process which forces a choice of medium on the story, and thereby a cascading series of further decisions like editorial style, treatment of sound, etc. *Reasoning* conveys the existence of something which makes logical inferences about the story based on the description (representation) of the story. The reasoning engine reads the story description, makes its inferences, and feeds the results of those inferences back to the artist.

A sophisticated example of this structure would be the software program *Dramatica*, from Screenplay Systems.[9](Screenplay-Systems, 1994) Dramatica provides the screenwriter with a gigantic sophisticated questionnaire. As the writer fills out the questionnaire, making high level narrative choices, Dramatica searches through its list of known narrative structures for matches. Its main goal is to force the writer into making the most detailed decisions possible about the story, so that Dramatica's matches for story structure and style come down to just one. When there is only one match, the system can provide the user with additional information about their screenplay according to the description

9 See: http://www.screenplay.com/

of the found match. The theory behind Dramatica suggests that writing screenplay dialog is a lot easier if all the necessary character decisions and attributes are already decided, like how they think, their strengths and weaknesses, etc.

Dramatica seems well suited as a narrative feedback mechanism, as long as the user's goal is to create a type of story in line with Dramatica's "expertise" – linear Hollywood styled movies. Dramatica's knowledge is stored as static rules about linear screenplay structure and character definition. In artificial intelligence (AI) terms, Dramatica takes a knowledge-based approach to the problem domain of screenplay structural analysis. It is well known in the AI community that one of the weaknesses of Knowledge-Based AI (KBAI) is that its structures become brittle when faced with a dynamic problem domain or any problem domain which it was not specifically designed to handle (Kolodner, 1993; Maes, 1992).

An alternative approach would be a behavior-based approach. This approach is described well by professor Pattie Maes of the MIT Media Laboratory. In her paper Behavior-Based Artificial Intelligence she compares and contrasts these two forms of AI by listing characteristics which typify the knowledge-based and behavior-based approaches. (Maes, 1992) What follows are those characteristics I have found most appropriate in the domain of narrative structure:

A Knowledge-Based Approach

- Models isolated and advanced or specialized competences (i.e. medical diagnosis or chess playing). The knowledge-based approach would rather provide "depth" than "width" in its expertise.
- Solves one problem at a time, usually with no time constraints for solving that problem.[10] Also the problem domain is static, remaining unchanged.

[10] This is certainly true in Dramatica's case. In the Hollywood film industry, the writing process is so far removed from the presentation process that, as far as the Dramatica software is concerned, the writer has all the time in the world to create and analyze her screenplay before production begins. Therefore, Dramatica's reasoning system really does not need to work any faster than what good human interface design dictates.

- Usually is not concerned with developmental aspects, how the knowledge structures got there in the first place, or how they change over time. Therefore, the knowledge-based approach does not have to be adaptive.

Alternatively,

A Behavior-Based Approach

- Has multiple integrated competences, such as those needed in locomotion or navigation for robots. For stories, these competences may choose conflicts, decide on story resolutions, or, with regards to presentation, decide how to smooth out audio transitions, for example.
- Is a system "situated" in its environment. For robots, this means that they are directly connected to their problem domain through sensors and effectors. For computational narratives, it means that the system navigates an environment of story representation and is "open" in that it is always accepting of user feedback.[11]
- Emphasizes the behavior exhibited by the system rather than the system's knowledge.
- Emphasizes the system's adaptive ability, which means that the system improves over time.

Behavior-Based AI (BBAI), as an alternative to knowledge-based AI, represents a fundamentally different way of thinking about a problem domain. Where the knowledge-based approach makes an a priori attempt to capture the rules for successfully solving or navigating a domain, the behavior-based approach instead relies on a set of lower level competences which are each "experts" at solving one small part of the larger problem domain.

BBAI constitutes the theoretical groundwork for the notion of Autonomous Agents (Maes, 1990). Autonomous agents are intelligent software modules that embody the ideas of BBAI. Autonomous agents are typically designed to control some sort of mobile robot or computer screen character, maintain certain "personalities" as members of a MUD

[11] That is, the system is open to user feedback whether it is responsive to that feedback or not.

(Multi-User Dungeon or Domain (Foner, 1993)), control user response behaviors in a software interface, or perform certain financial or search tasks on the internet.(Maes & Kozierok, 1993) (Tecuci, 1998) An agent must be able to maneuver around obstacles without getting stuck in an awkward space or lost in an endless loop of co-dependent tasks, oscillating between multiple goals. They are designed to deal with domains that are not entirely known, where unexpected things can happen. To this end, care must be taken to ensure that there exists a set of competences within the agent which handle the low-level tasks necessary for the operating environment; i.e. stepping, talking, or communicating over a network.

Metalinear narrative research employs BBAI through the use of software agents. The software agents are less brittle and more adaptive to a dynamic narrative representation environment than a KBAI approach would be. Using software agents, story domains can grow or change, while the agent remains the same.

3.2.2 The Simple-Link Approach

As mentioned in Chapter 2, hypertext systems like StorySpace offer a web-like structure for text, providing navigable connections between small granules of story. With its graphical representation of nodes of text and the connecting links between those nodes, StorySpace encourages a spatial conception of writing that goes beyond what we have come to expect on a simple 2D writing surface.

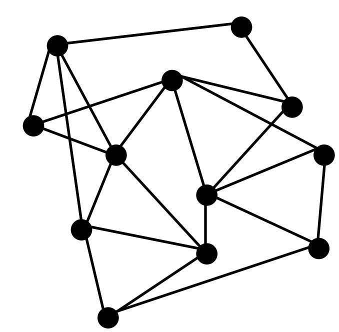


Fig. 12 The web of simple linked nodes.

> *All forms of writing are spatial, for we can only see and understand written signs as extended in a space of at least two dimensions. Each technology gives us a different space. For early ancient writing, the space was the inner surface of a continuous roll, which the writer divided into columns. For medieval handwriting and*

modern printing, the space is the white surface of the page, particularly in bound volumes. For electronic writing, the space is the computer's video screen where text is displayed as well as the electronic memory in which text is stored. (Bolter, 1991, pg. 11)

The next step past hypertext is hypermedia, which goes beyond simply text as a medium of expression to include a diversity of media forms including still pictures, video, and sound as well. Both hypertext and hypermedia have been defined and examples built, as Stuart Moulthrop states, "with electronic cross references that move the reader instantly from one piece of information to another." (Moulthrop, 1990, p.7) Other hypertext or hypermedia systems, like Apple Computer's HyperCard and the World Wide Web, offer a broader range of uses for hypertext and hypermedia with roughly the same level of spatial representation.

What such simple-link systems lack is any meaningful definition for their links. Their links are constructed simply with the knowledge that one part of one piece of text is connected to one part of another piece of text. The reasons why a connection is in place or the notion of a particular type of connection between two pieces of text is absent from hypertext.[12] In a sense, the reader of a hypertext is more active than the reader of traditional text because they are having to mentally resequence pieces of text so they can have a better understanding of character, plot, temporal relationships and physical/geographic relationships. Hypertext offers a flexible arrangement of narrative pieces at the cost of the user bearing the burden of creating a coherent narrative.

12 There are starting to appear specifications for extensions to World Wide Web which include adding metadata to links, including link typing. See the HyTime Standard: http://www.hytime.org/papers/htguide.html or XML: http://www.textuality.com/

> *Hypertext... provides an infinitely re-centerable system whose provisional point of focus depends upon the reader, who becomes a truly active reader in yet another sense. (Landow, 1992)*

While this decentralized or distributed mechanism of writing does offer some readers the pleasure of finding their own way through the task of constructing a story (as Landow implies) for many readers this process can be quite troubling or even frustrating. Readers are presented with the difficult task of constructing narrative from a cloud of interconnected text nodes, with few (if any) hints of how to do so. (Murray, 1997, p. 133) Each link is essentially a blind alley from one point to another. The terms "hypertext" and "hypermedia" have been used to define a story form with rich networks of story pieces, but with the authorial intentionality of those network connections safely tucked away in the author's mind. Yes, the reader can move from one piece of information or story to another, as Moulthrop states – but why? Much of a hypertext user's experience is based more on impulsive decision making and not on narrative reasoning.

Hypertext does not help the author to build the hypertext system, other than through its graphical user interface. As the network of nodes and links become more complex, the author must pay more attention to organization and the many possible paths a reader may take. The only way to check these sequences is to follow them "by hand" – methodically checking each node and choice path, making sure that what was written actually makes sense when placed together. Fortunately for hypertext, story coherence is less of a problem for readers than it is, say, for viewers of a movie or some other cinematic experience (i.e. television narratives and movie rides)[13]. A reader, rather than a viewer, is much more flexible and forgiving of seeming discontinuities in text. We are familiar with poetic

[13] Radio dramas might also fall into this category. To what extent the radio drama experience is a cinematic experience is completely dependent on the listener and their ability to picture the events described.

models of text which use discontinuity as a means of expression (i.e. James Joyce's ULLYSES). In some ways, textual discontinuities can even act as an asset and appear as part of the art form. Hypertexts, even with these advantages and flexibility, cannot create true meta-linear narratives. For this a tool is needed.

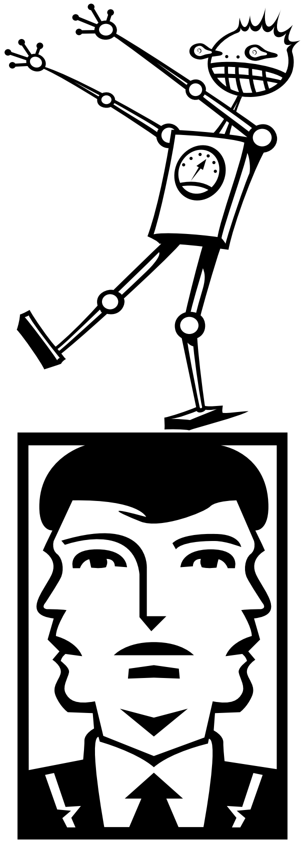


Fig. 13 From robot locomotion to personal assistants, software agents have influenced the way we think about computing by emphasizing behavior and personality.

3.2.3 The Multiple Character Approach

Much work has been done in computer science in the area known as artificial agents or software agents. As discussed earlier in this chapter, software agent research began as a robust method of controlling physical robots as they dealt with an operational environment that was not completely known. (Maes, 1994) (Tecuci, 1998) Further work has yielded agents which control the behavior not of physical robots, but virtual robots called animants. (Blumberg, 1996) (Galyean, 1995) Animants are graphically represented robots on a computer screen which exist and react to each other and events inside of a virtual world. The agent behavior modules of the animants allow them to interact with each other and with humans with varying degrees of complexity and realism.

Professor Joe Bates and his graduate student team at Carnegie Mellon University worked on a research project called Oz which used intelligent animated characters. (Bates, 1992) While Bates' research encompassed a number of projects having to do with software agents and dramatic experiences, *Oz* is particularly significant because it evolved after Bates did some basic experiments with live improvisational actors and a generalized story script. Bates created a type of artificial narrative using small animated potato-shaped characters called Woggles on the computer screen. The characters inhabited a rocky landscape and each had their own individual set of behaviors. The user, controlling one of the characters, could drive their character up to another character and engage them in any number of activities of communication or play in believable ways. One of Oz's goals was to keep the user engaged in the system by making the characters believable. They accomplished this by using personality and interaction models. Bates created a potentially narrative play space, in which the human user could direct the actions and events between characters – actions and events chosen according to a narrative being constructed in the user's mind.

Professors Barbara Hayes-Roth and Daniel Rousseau of Stanford University took Bates' work further by providing agent characters with improvisational models of interaction, creating synthetic actors which can produce performances that are theatrically interesting, believable and diverse.(Rousseau, 1997) Rousseau and Hayes–Roth's virtual actor agents perform stories by providing the agents with a set of abstract directions in the form of a script which describes what to do, and a set of personality behaviors which describe how to do the scripted actions.

Further research such as that by Peter Wavish and David Connah of Philips Research, included ways for making agent actors appear more intelligent or capable than they actually were, by implementing scripts. (Wavish, 1997) While they implemented their agents in a way similar to Hayes–Roth and others, Wavish and Connah were able to achieve this deception of the audience/user in part by having the actors engage in diectic dialog; that is, the agents pointed at objects in their environment and referred to them relative to other objects, rather than referring to them by name.(Cremers, 1996) Their Communicating Deictic Agents (CDAs) perform a written script of actions, improvising background activity along the way. By including a script in the actions of agents, Wavish and Connah opened the door a little wider for a more authored interactive agent experience, yet with the additional ability to react and do the right thing when events are dynamic and unexpected.

Progressing from the more general use of AI and agent technology toward a domain specific application, the term story engine is used to describe a set of software algorithms designed to make decisions regarding how a computer-based story should proceed. That is, the story engine decides what's next in the story, embodying some of a human author's reasoning for doing the same task. Story engines are construction engines, deciding the sequence of each small detail, major event, and opposing or supporting position of the story. While some story engine research has taken a more traditional approach, focusing on narrative structures and models, other research stems from an area which is not typically based on narrative, that of computer gaming. Text-based story engines tend to maintain a look and feel resembling their ancestors, the early computer-based adventure role playing games.

Computer game designer Chris Crawford's *Erasmatron*[14] is one example of a text-based story engine which has taken a step toward graphic representation. In general, the Erasmatron story engine offers a way of constructing a text and image story experience by navigating a collection of narrative material, with no exclusive ties to that material. A single story engine can construct many stories using different sets of story material. Erasmatron works by having the writer program each character into the system. Each character has behaviors which control how they react emotionally to other characters in various situations, and graphical representations expressing their mood or emotional response. The user interface is composed of a still image of the currently active character bearing a mood specific facial expression, text describing what is currently happening in the story (which

[14] Information about Erasmatron ca be found on their web site. http:.www.erasmatazz.com/company.html

could be the dialog between the currently active character and the user's character), and a menu of computed choices. Based on the user's menu choice, the active character on the screen reacts both graphically and through dialog depending on their programmed behavior, thus sending the narrative in a new direction.

What each of the above researchers have essentially focused on in their own special way, is the intelligence of the story characters themselves, relying on that intelligence and the way that the characters react to each other as a way to create a narrative experience. The characters are physically and aurally responsive to each other and so appear to create narrative by means of cause and effect. As writer E. M. Forster points out in his book *Aspects of the Novel*, cause and effect go a long way towards defining plot.(Forster, 1954) Forster's example is that to say, "The King died, then the Queen died," is not a plot. To say, "The King died, then the Queen died of grief," is a plot because causality is shown. In a similar fashion, the Multiple Character Approach researchers establish causality in their characters by defining reactionary behaviors. Characters are programmed to react in particular ways given particular situations or preconditions. Often, one character's reaction to one set of preconditions establishes a different precondition, which other characters then react to. In this way the characters interact and are responsive to each other.

Though the characters react back and forth to one another, they establish no over arching direction. To say, "The King died, then the Queen died of grief," is only a true plot if it is an overarching description of a larger work or if just that statement is the entire story. If that statement is "the plot" of a single sentence within a much larger work, then its narrative effect is minimal. Plot requires some amount of overarching intentionality and fore thought. Plot is a force that drives characters and events forward through time and space.

To reduce plot to a reactionary flinch response is not as effective. In the case of Oz, there is no overarching narrative plot at all. Any interpretation of an Oz experience as having a plot is solely in the mind of the audience/user. While the user's mind is not a bad place for the plot to exist, Oz makes little effort to place a plot there.

3.2.4 The Puzzle Approach

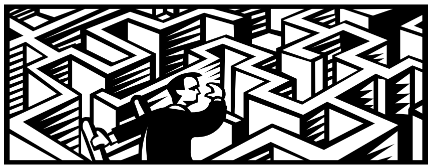
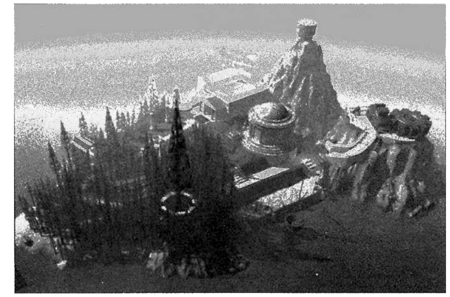


Fig. 14 For some, puzzles are an interesting challenge. For others, puzzles have all the allure of a trap.

Some narrative software projects take a puzzle approach to telling a story. That is, like a game, they present a world of interconnected puzzles which the user solves. These puzzles can be elaborately devised, using recurrent characters, adventurous journeys and exotic landscapes. *MYST* is an excellent example of using the computer to tell a story with the use of puzzles or problems.(Miller, 1993) The entire Myst experience is a puzzle. Its user interface includes a simple yet powerful employment of different media types including digital video, beautiful and elaborately designed graphics, and subtle continuous sound. The premise in Myst is to travel back and forth in time and space on a deserted island, solving puzzles in order to collect clues. These clues allow the user to find out what happened to two brothers mysteriously entrapped in two books in the library. The player must decide their fate.

Myst communicates an enormous amount of story with relatively little data by providing a rich landscape, an elaborate soundtrack, and a premise which ensures player safety. The documentation for Myst makes it clear that there will be no mad slasher bad guys jumping out from behind a tree, so the user may explore the landscape freely. This sets Myst apart from many other computer games. It is an adventure game which allows for interactivity based on intellectual pursuit rather than survival.

Because Myst is a puzzle, or actually a large collection of many different puzzles woven together, cause and effect play a strong part in the Myst experience. When one of the puzzles is solved, more knowledge is gained that helps solve the next puzzle. When a puzzle is partially solved, it is immediately clear to the player that they are close to the solution. The Myst world is highly reactive to player actions. Because of its reactivity, over arching goals and rich aural soundscape to draw in the player, the Myst experience works as an elaborate narrative. It is a narrative, however, which does not simply offer, but requires, a high level of interactivity. To gain access to the narrative material, the player must try to solve the puzzles. There can be no passivity with Myst.[15]



[15] One way in which Myst was strikingly interactive is that some of the puzzles within the game were so difficult it required multiple people to work together to solve them. Thus, some of the most interesting and fun interactivity happened in front of the computer screen between the players.

3.2.5 The Traffic Circle Approach

Another way to structure a computational narrative is to create an environment where one travels through, yet always returns to the same central spot. The user would go off exploring small side paths of narrative, discovering characters which may or may not color the remainder of their experience, always returning to the central “traffic circle” of the program. The advantage of this structure is that the user is presented with a narrative environment in which they feel as if they have free control over what they see and do. By giving the user the ability to freely explore, the authors give users more opportunity to construct (narrative) meaning from their experiences.

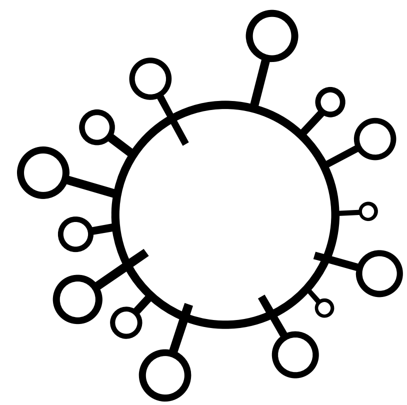


Fig. 15 The structure of The Traffic Circle Approach.

The software program MIDNIGHT STRANGER is a good example of the traffic circle approach. In Midnight Stranger, the user adopts the character (avatar) of a white male in his twenties, investigating the Los Angeles night life and club scene. It is one of the best examples of a simple, flexible user interface for navigating through a narrative space, while gathering feedback from the user. Navigation through the LA streets in Midnight Stranger happens by clicking on active areas of the background, which depict buildings, doors, street corners to turn down, and so on. The player is not always allowed to freely walk around, however. Sometimes navigational control is taken away from the player to ensure they have or finish having an interaction with one of the game's characters.

The characters in Midnight Stranger make statements or ask questions of the user, to which the user responds through the use of a color bar. The color bar is a thin graphic rectangle at the bottom of the screen which contains a color spectrum that gradually changes from red to blue. By clicking toward the red side of the bar, the player is responding with emotional warmth to the character. By clicking toward the blue side, the player is responding coolly to the character. Characters respond to the player's reaction in different ways, depending on how the characters were designed. For example, if the player approaches a female character at a club and engages her in conversation by responding emotionally warm to everything she says, there is a chance that the conversation will continue, that she will become friendlier, and ultimately result in both character and player going together to the character's apartment. If the player responds emotionally cold to the character, the conversation will most likely be broken off and the player will not "get lucky."

The advantage of such an interface is that the system does not have to be concerned with understanding natural language. Instead, various emotional response segmentation patterns can be applied to the same color bar at different points in the program, allowing the different characters to respond in different ways. For example, one character may only have two types of responses to the player, warm or cold, giving the color bar only two active areas. Another character may have a broader range of responses, giving the color bar four or six active areas, and making it possible for the player to respond with too much warmth to the character and turn them off. While the story material, character dialog, and media production quality of Midnight Stranger leaves much to be desired, the method of character engagement is quite ingenious.

No matter what happens between the player and various characters, after the climax of an event, the player always materializes back on the same street, facing the same decision time and again, “Which direction should I try now?” Here is where the traffic circle structure comes into play. Part of the narrative granularity is based on travel to a physical location (to the club for example), yet after each event granule, the central granule comes back into play.

There is the notion of passing time within Midnight Stranger, i.e. the restaurant can close, the cute woman at the club could no longer be there, for example. Therefore travelling back to previously experienced locations can mean having a different experience than before. Yet the central traffic circle scene is always the same. It binds together the other elements like the hub of a bicycle wheel, making navigation through the narrative space simply a matter of always knowing how to get to the hub.

While it is conceivable that this technique could have a comforting effect on some players (providing a home-base of sorts) it can also be quite frustrating if not boring. If nothing ever changes in the central section of the game, a player could easily get bored with it. Also, it is very difficult for a story designer to create a scene, much less “the central scene”, which can effectively act as narrative glue between the many other scenes it connects. In order to make the central scene work in this way, it would have to be rather generic and bland (as in Midnight Stranger), otherwise it would eventually not work as a connecting tissue between certain narrative events, seem discontinuous to the player, and thereafter either be a point of confusion or simply ignored.

3.2.6 The Single-Stream Cinematic Sequence Approach

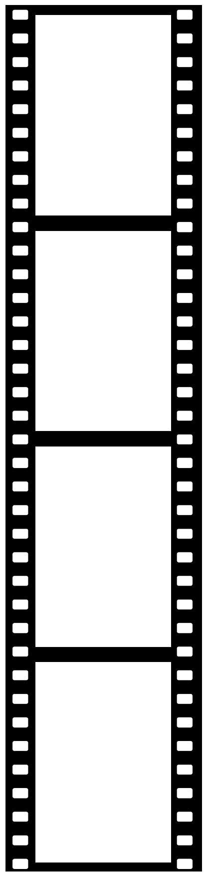


Fig. 16 Cinema has given us moving pictures and sound presented in a manner which we strongly adhere to, even when the technical and physical limitations are removed.

Cinema has provided us with a paradigm for experiencing narrative through moving pictures for the last century. Over the last fifty years, television has extended cinema's model with its own technical and aesthetic rules and styles. Presentation through a single stream or frame of moving image is one attribute of cinema that has not been extended by television, but remains as part of the definition of the medium. There are examples of computer-based cinematic narrative projects which have introduced new and unique ways of representing and sequencing video, but which have maintained the paradigm of single-stream cinematic presentation.

New Orleans in Transition, 1983-1986 is a project completed by Glorianna Davenport of the MIT Media Lab.(Davenport, 1987) It captured approximately fifty different people on film who were involved with the development of the city of New Orleans, Louisiana around the time of its 1984 world's fair. The film captures architects, developers, business people, politicians and residents in their effort to preserve the historic French Quarter of New Orleans, while also restoring and developing the deteriorating waterfront region along the Mississippi River.

New Orleans in Transition consists of three hours of footage stored on a set of laser-discs to allow for computer controlled access. With computer control, a user could navigate through the set of key players in the struggle between conservation and progress. The user could see the connection between one character and another, or click on the icon of a character on the computer screen to find out biographical information about them. Because of the laserdisc technology and video capture/display technology in use at the time, New Orleans in Transition was forced to exists as a single video stream project. While it made creative use of the computer screen in the way that video, text and icons for navigation where presented, it was limited in the way it could use video material, as it was all streaming in from a standard laserdisc player.

The Digital Micromovie Organizer (DMO) was able to shed the limitations of the laserdisc player by storing and playing digitized video. (Davenport et. al., 1993) The DMO took a large collection of short video clips that have been described in a video database and orchestrated them in real-time to create a flowing sequence of shots on screen. The shot selection used a series of filters to choose the most appropriate clip for each place in the sequence. A number of simple filters came with the system (for pacing, continuity, etc.) and authors had the power to create specialized filters for use in their stories. The project An Endless Conversation used the DMO to sequence clips of two characters in conversation, asking and responding to each other's questions.

While the DMO and Endless Conversation where less limited in their presentation methods due to their use of digital (rather than analog) media, they also chose to express narratives in a single video stream form, as in cinema and television. They maintained the cinematic paradigm of seeing only one point of view at a time, though it was possible to sequence through a stream of multiple visual points of view, possibly through the use of a specialized shot filter.

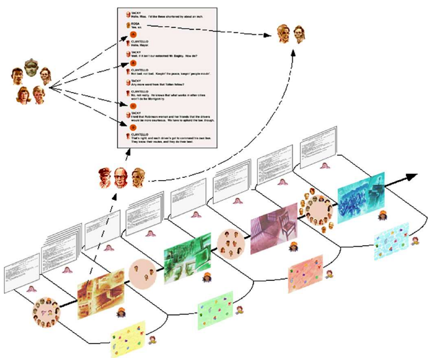


Fig. 17 The narrative unfolding structure of TOGI.

3.2.7 The Folded Approach

Narrative involves an author setting forth predetermined paths or a general structure along which the narrative should proceed. Yet how should the author treat the user of a computational story system who is actively navigating these paths? That is, how does the author shape the participation of the user given feedback from the user? Midnight Stranger offered one method, utilizing character dialog and a form of response. Are there other methods which offer a significantly different result?

Tired of Giving In (TOGI) is a project conceived by Carol Strohecker of Mitsubishi Electric Research Labs and Larry Friedlander of Stanford University. It was produced by Strohecker and myself in 1996.(Strohecker, 1999)[16] TOGI tells the story of the beginning of the 1955 Montgomery bus boycott in Alabama, an event which ignited the American civil rights movement and changed the face of the United States. The project chronicles key events in the life of Rosa Parks, the black woman who refused to give up her bus seat to a white man and was arrested for what was then a crime.

TOGI was designed to tell this story in three different layers or folds, each fold revealing progressively more detail and allowing more user interaction. The first fold consists of a narrator simply and briefly relaying the major events of the story's four parts in chronological order: the town of Montgomery, the bus where Parks was arrested, the jail in which she contemplated her fate, and the church in which Rev. Martin Luther King spoke and first led the town's blacks to boycott. If the user clicks anywhere on the screen during any of the four sections, further story detail is presented in the form of major character dialog and choral commentary. TOGI uses a diverse set of voices as a "Greek chorus" to respond

[16] See also: http://www.merl.com

either individually, or in unison, to the events presented in dialogue by the major characters.

The chorus consists of twelve individuals from three different eras: the Past (ancient Africa) with 3 members, the Present (1955) with 6 members, and the Future (1990's) with 3 members. Within each of these groups there are chorus members who represent the voice of pro-boycott, the voice of anti-boycott, and the voice of someone who is torn between doing what is right and dangerous or conservative and enslaving. Akpan, Tebogo, and Udo are the ancient African chorus members of the Past, who are graphically represented by African masks. Winona, Isaac, Beulah, SallyJo, Jonah, and Bud are the voices of the story's Present in 1955 and are represented by black and white adults on both sides of the issue. Latisha, Natoya, and Ana are the voices of the story's Future and are represented by teenagers from three different ethnic groups.

The user has the opportunity to ask individual chorus members what they think about the events of the moment, thus gaining different perspectives. This user interaction represents TOGI's second fold of story. If the user wishes further interaction, they may click on a major character to have that character enter into a dialog with a system-selected chorus member for the story's third fold. The user unfolds the story through their interest, as communicated through their actions.

The question posed at the beginning of this section was: "How does the author shape the participation of the user-given feedback?" For TOGI, our specific question was: "Do we as authors give the user more of what they have demonstrated that they want or like, or do we give the user more of what they have experienced the least?" To some extent this question pertains to the goals of the project itself. Does the project have as its primary goal

Fig. 18 The twelve TOGI chorus members. Graphics designed by Steven Alexander.

to entertain or promote learning? Is there only one primary goal throughout the project? As some of the TOGI chorus members were pro-boycott and others anti-boycott, each side expressing their own worldview, the user is able to choose a narrative experience which generally weighs more heavily on one side of the boycott issue or the other. The binary nature of the issue makes navigating it a little like navigating the edge of a fence. The answer to the question of how to respond to user feedback still is not clear and perhaps cannot be answered definitively.

The issues addressed above also pertain to computational story aesthetics. The human storyteller has much more leeway than the computer for processing user feedback, both during and after a story presentation. Yet, in a computational story when a human author wants to make sure the audience gets a certain message, is it possible to ensure the delivery of this message computationally? Or must the author take the following into consideration as they write: *THE MACHINE DOES NOT HAVE THE POWER TO RETELL MY STORY IN SUBTLE WAYS, SO THAT IS WHAT I MUST ADDRESS WITHIN MY VARIOUS STORY CLIPS.*

The possibility of providing a substantive crafted computational narrative may be a level of difficulty beyond the scope of simple AI data structures and algorithms alone, especially when the issues being navigated are non-trivial and profoundly human. Part of the premise for this thesis is that the answer lies not in the computation alone, but in the coordination between computed structures and human craft. To tell a computational story well, a close working relationship is needed between the author and the author's tool.

3.3 Metalinear Story Coherence–The Art Form

They find it hard to grasp some things that come easy to us, because they simply don't have our frame of reference. I show them a can of Campbell's soup.
I say,
"This is soup."
Then I show 'em a picture of Andy Warhol's painting of a can of Campbell's tomato soup.
I say,
"This is art."
"This is soup."
"This is art."
Then I shuffle the two behind my back.
Now, what is this?
No,
this is soup
and this is art.

- From Jane Wagner, THE SEARCH FOR SIGNS OF INTELLIGENT LIFE IN THE UNIVERSE, 1986

3.3.1 Is There an Art Form?

The previous section examined different approaches to computational story systems and specific examples for each approach, including their strengths and weaknesses. Each example attempts to express a narrative form, through both computational structures and narrative material. Each example informs metalinear design in some way. Given the examples of knowledge-based, character-based, puzzle, traffic circle, single-stream cinematic sequence and folded approaches, the questions remain: *DOES THE METALINEAR*

Approach suggest an art form? If so, what is it? Is there an artistic form of narrative expression which is authored by a writer of words and a writer of computer code, and guided by the audience? Can such an art form provide a coherent narrative experience for the user/viewer.? The cited examples and the many hundreds of other story/computer/media projects suggest that there is indeed such an art form.

In the 1970's and 80's there were computer videodisc projects like *Aspen* which, while not narrative, did allow the user to explore a real space in a continuous cinematic flowing manner.(Mohl, 1982) In the 1990's, Tinsley Galyean's narrative guidance work and his *Dogmatic* project told a story in an immersive environment with a similar cinematic flow. (Galyean, 1995) From *Aspen* to *Dogmatic* it is clear that the evolved computer narrative art form is cinematic in nature. It is a dynamic visual art form which will take advantage of how the language and interpretation of cinema has pervaded our daily lives over the last century. When a character is hurt or insulted, we expect to see a close-up reaction shot of that person.(Zettl, 1990) When the music swells, we know that something important and possibly climactic is about to happen; gone are the days when the audience lurches at a close-up shot of an oncoming train. And as we have accepted this language in our media, we expect it to be spoken to us, and spoken to us well. We expect this to such a degree that when insightful projects like Midnight Stranger include shoddy video production quality, we find it jarring.

As the metalinear art form is heavily influenced by cinematic language, it is also clear that this computer narrative art form should not be static. There was a time when the dominant data storage media necessitated a locked and unchangeable narrative experience. Laserdisc and CD-ROM distribution methods inspire computational narrative structures which themselves are fixed and limited, such as branching algorithms and knowledge-based AI methods. With the promise of large scale high-bandwidth 2-way network access to the home from such companies as Motorola, Media One, and AT&T, fixed narrative structures become less necessary or desirable.

An over arching concern about the metalinear art form, which encompasses its cinematic and dynamic nature, is whether it will be coherent. The cinematic form is highly authored and thereby forced into coherency. The computational nature of metalinear narrative also provides authorship through its programming. This authored programming is more abstract in that algorithms and data structures provide a coherent potential for a multitude of combinations. Yet even with a multitude of combinations, the metalinear expression must still be coherent. It must allow a writer to not get lost in the writing and structuring process, while also making sure the audience experiences a story that makes sense.

There is an art form for metalinear narrative which draws inspiration from the projects listed in this document and others. Additional attributes of this art form are addressed in the sections below.

3.3.2 What the Metalinear Art Form Is Not

There are certain attributes that this art form should not have. For instance, a knowledge-based approach, as is used in Dramatica, is too brittle a structure to maintain a robust and evolving narrative. If one of the defining features of the metalinear narrative is that story material can be reused over and over again to construct different stories, then the inherent rigid structure and linear reasoning associated with a knowledge-based approach would be detrimental to this end.

As stated earlier, the art form would require that the author's reasons for connecting two pieces of story be somehow captured in the connection. By capturing the author's intentionality, the system should then have what it needs to make narrative decisions. It is the author's intentionality that fuels these decisions and recording that intentionality in a knowledge-based form is a cumbersome task, to say the least.

Bates, Wavish, Connah, Hayes-Roth and the others have demonstrated the power of focusing on the behavior and reasoning ability of narrative characters realized through software agents. While this can be a valuable attribute to have, care must be taken on a number of fronts. For instance, when designing a computational narrative tool, it is important to give the author control over the narrative in a form which they already understand or can easily appropriate. The story script (like a screenplay, for instance) is such a form. In computer science, the term *SCRIPT* usually refers to any structured procedural description of actions to be taken which can easily translate to something that a narrative writer would never touch. To a writer, a script is a complex collection of actions, events, characters, intentionality, dialog, and unspoken emotion. The writer's job is to use these items to craft a story world, and not exert a lot of energy translating them into a technical language.

The world does not need another piece of software which forces people to think in unnaturally rigid and narrow ways which happen to facilitate computation. There will have to be a balance between character intelligence and the story script which determines certain actions. An author will want to create a metalinear story focusing more on the craft of storytelling and less on the craft of software agent design. Yet both of these crafts must live together intertwined, as both are necessary components of the art form. The coordination and balance between the emotional art of story and the procedural art of designing software data structures, must be as smooth and productive as possible.

> *It is important to remember that any abstract story system ultimately refers to the sorrows and pleasures of human life and the story of any event depends heavily on who is doing the telling. A storytelling system that further calcifies the distortions of stereotypical thinking would be as destructive as the most bigoted and bloodthirsty bard. We humans already do enough mechanical thinking without enlisting machines to help us. (Murray, 1997, p. 199)*

Metalinear narrative is not a computer game. A story is not the same as a game, though many games can have narrative components in them. Computer games have been such a strong area of computer software research and development over the past few years that many of the advances in CPU design are based on the needs of the computer gaming community. Sega, Sony, Nintendo, and to some extent Motorola and Intel are all in an endless death match for biggest, baddest, fastest processor which can show more sprites, more realistic graphics, play better sounds, and all the while supporting more simultaneous game players. While certainly adventure games and first person shooter arcade games like *DOOM*, *QUAKE* and *TUROK II* offer some small amount of back story to their environments and characters, they have capitalized on the fact that it takes very little story to get a game player reloading their sub-machine gun and mowing down virtual soldiers. It is too bad that what has developed to be the most interactive electronic activity has little need for story. There are new video games starting to appear, such as *METAL GEAR SOLID* from Konami, which regularly stop game play for the purpose of engaging the player's avatar in developing narrative and revealing back story[17], unfortunately when new media developers with gaming backgrounds enter into the realm of story, they typically

[17] Pausing game play for the narrative is oddly reminiscent of musical theater and movie musicals when it is the narrative which regularly pauses for a song or dance number. Even so, it will still be the storyteller, not the programmer, who will make Metal Gear Solid's spy character as charming and witty as Fred Astaire.

attempt to do so with the same steering wheel or joy stick control system as their games. In so doing, there is little time for reflection, and therefore, little time for narrative.

> *The genre imitates action, rather than reflection, since if one makes a mistake and one's persona is killed, it is a simple matter to start over again. In a sense, the form intrinsic to the genre devalues the role of the individual persona and the need for reflection. (Niesz, 1984, p. 122)*

The art form is not "just" interactive. There is a difference between interactivity and agency. To have agency means to have meaningful control over a world, or even a destiny. The art form of metalinear story must give the user agency.

> *Because of the vague and pervasive use of the term interactivity, the pleasure of agency in electronic environments is often confused with the mere ability to move a joystick or click on a mouse. But activity alone is not agency. For instance, in a tabletop game of chance, players may be kept very busy spinning dials, moving game pieces, and exchanging money, but they may not have any true agency. The players' actions have effect, but the actions are not chosen and the effects are not related to the players' intentions. (Murray, 1997, p. 128)*

The agency that the metalinear form must have is two fold. First, the author must be in control of the story domain they are creating, whether they have absolute knowledge of every possible linear construction or not. Second, the audience must have some control of their narrative experience and know that the actions for navigating through the narrative material have specific effect. The audience must at least have the perception that they have control over the narrative, even if they do not always have the same level of control they believe they have or have their control taken away from them for brief periods of time. (Galyean, 1995) Without this two fold sense of agency, the computational narrative experience resembles something more akin to digital roulette than a story told with some thought and care.

The metalinear art form, therefore, is not a static knowledge-base of facts, nor a complex programming language or system of technical jargon that would put off a writer, nor a computer game with a high level of control and little narrative, nor a system where control is relinquished by author or writer. What the metalinear art form is, however, is an exciting new form of expression which holds a great deal of promise.

3.3.3 What the Metalinear Art Form Is

If we know what the art form is not, then what is it? What are the characteristics of metalinear narrative?

We know that the metalinear art form has two sides: that of textual design and computational structure design. The writer's role is expanded here to include a technological component, a necessary part of this new craft. Writer's expressive form goes beyond words and sentences by merging with the structures of computer software. This characteristic is one Janet Murray refers to as procedural.

> *Authorship in electronic media is procedural. Procedural authorship means writing the rules by which the texts appear as well as writing the texts themselves. It means writing the rules for the interactor's involvement, that is, the conditions under which things will happen in response to the participant's actions. It means establishing the properties of the objects and potential objects in the virtual world and the formulas for how they will relate to one another. The procedural author*

> *creates not just a set of scenes but a world of narrative possibilities. (Murray, 1997, p. 152)*

The metalinear art form encompasses a balance between authorial and audience control. The author can offer a narrative to an audience in which the audience is more empowered than they are with the functions on their VCR remote. The author's craft is one of story potential, where the final form is many possible forms. This makes the audience more active than ever, making choices on a level and scale never before accomplished. Striking this balance of authoring story potential and authoring audience activity is not an easy one.

> *If we give the interactor complete freedom to improvise, we lose control of the plot. But if we ask the interactor to pick from a menu of things to say, we limit agency and remind them of the fourth wall. (Murray, 1997, p. 190)*

Yet within that struggle to find balance there is also freedom. It is within the bounds of restriction that writers often find their voice, their power. In writing about textual electronic fiction, Jay David Bolter states:

> *Electronic fiction is technologically complicated in that it requires a computer and the sophisticated arrangement of text and graphics on a video screen. But it is conceptually simple – simpler than writing for print, where the writer must always force his or her text into a single line of argument or narrative. The computer frees the writer from the now tired artifice of linear writing, but the price of this new freedom for the writer is that the writer must allow the reader to intervene in the writing space. (Bolter, 1991, p. 145)*

The metalinear art form is born from a collaboration between writer and computer. The computer assists the writer in the process of structuring her story granules for sequencing. This process relies on the computer providing a means of representation for the granules. Writing a metalinear story means manipulating the computer's representation of the characters and events such that it can sequence the story granules according to the author's intent.

In freeing the writer from linear writing, metalinear narrative provides a structure into which multiple story fragments or granules can exist. Through this structure multiple construction is possible. The writer is free to write many different versions of the same story granule. The writer is not held to a single "right" scene, paragraph or sentence. They may create different versions of story granules in order to give characters and character relationships multiple possible sequences.

Metalinear narrative is an art form which extends narrative voice. The extended voice of the writer can say more things, in more ways, and in more contexts using the computer.

> *What the computer would provide would be a means for using formulaic patterning, in much the same way the oral bards did, as a system for assembling multiform plots. The electronic system might be able to generate more variants than the author could ever read in a lifetime (let alone write individually), but since she would have specified all the important details and all the rules of variation, the computer would be merely the instrument of the author, an extension of her memory and narrating voice. (Murray, 1997, p. 212)*

Metalinear narrative is also the embodiment of many voices. Through its embrace of multiplicity, those voices which are not usually heard are no longer silenced by the need to make the harsh choices of singular linear sequence. By no longer forcing writers to think, construct and edit uni-linearly, writers can be much more open and inclusive of the diversity of voices around them. Cognitive scientist Richard Lehrer refers to this type of construction as Hypercomposition.

> *...Hypercomposition encourages the composer to be aware of the multiple voices of his or her composition because there is always more than one path through the hyperdocument. (Readers need not cede control to the authors of hyperdocuments.) In principle, multiplicity of voice may make authors more likely to consider their audience when they design, and it may make them more likely to consider revision. (Lehrer, 1993, p. 201)*

In being aware of multiple voices, the metalinear narrative author has the freedom to present multiple simultaneous voices or points-of-view. Instead of holding on to the single video stream artifact of analog video, the metalinear art form can reflect cultural multiplicity through a presentation design of multiple simultaneous characters, events and sounds, engaged in expressing their view of the world.

The tool needed to create metalinear narrative is intelligent and empowering. But what does it mean to say that software is intelligent? AI researchers tell us that software intelligence is measured by the magnitude of a system's stored knowledge.(Lenat, 1991) In this case, however, intelligence should be measured not just in terms of knowledge stored in the form of data structures and algorithms, but also in terms of how well it fits the hand of its user. A software tool, or any tool for that matter, is only as intelligent as the user is with the tool in their hands. Therefore, the tool's empowerment of the user is of great importance.

> *The tools of the future will be intelligent, dynamically adaptive, customizable, and personalizable to a staggering degree. With experience, they will learn and grow and wear to fit the specific craftsman's hand. Their complex functionality will be deeply couched in metaphor or story, and their internal operations will be hidden from view, until demanded. (Davenport, 1997, p. 9)*

From this section it has been shown that the metalinear art form extends the writer's narrative voice so the writer can say more things in more ways. Metalinear narrative includes many character voices together, including typically marginalized voices. It balances agency between the author and the audience and frees the writer from a forced unilinear construction practice in favor of multiple linear constructions. Metalinear narrative requires computational assistance in the multiple construction process, because it is only with that computational assistance that the writer is truly freed to write without the concern of having to perform the laborious construction process themselves. Metalinear narrative is an art form which encapsulates multi-sequentiality with an authorship of multiple sequence potential.

3.4 The Need

The need, then, is for an intelligent and flexible writing tool which allows the writer to design the text, as well as the structures and algorithms which act on that text. The need is for a writing tool which can incorporate the diversity of voices we are becoming more aware of in our twenty-first century world. The need is for a writing tool which gives the writer the power to create these stories of multiple voices, while never losing track of narrative structure and character relationships. The need is for a writing tool which gives the audience an easy way of navigating the multi-dimensionality of multi-voice narratives, and see a differently constructed narrative made from the same narrative granules. The need is for a writing tool which can help the writer create metalinear stories without applying a qualitative metric to the writer's work.

A writing tool which offers the author knowledgeable feedback about narrative construction and context during the creative process, is essential to the task of creating metalinear narratives of significant dimension. By "significant dimension" I mean in terms of size as well as quality. A writer is able to keep only so much of a complex story structure in their head while they are creating. The task of remembering that structure while also thinking about the many different ways that structure can be applied to their story domain, is harder still. A specialized writing tool is needed to help manage a complex narrative structure and allow the writer to focus more of their energy on what they do best – writing good story material. There is a need for a software tool which maintains human creativity and authorship in the writing process, while also enabling the computer to construct multiple linear narratives. My contribution to this field is software which answers this need. The software tool is called *Agent Stories*.