Towards a Theory of Narrative for Interactive Fiction

Sean Smith Joseph Bates

February 20, 1989

CMU-CS-89-121

Abstract

The current paradigm for synthetic reality based interactive fiction features familiar kinds of worlds realistically presented. This approach is unsatisfying in that it excludes several interesting classes of worlds, such as worlds where the user experiences the subjective reality of a character with a substantially different personality. We wish to extend the current paradigm to include classes of worlds such as these. To achieve this, we survey cinema technique and develop a theory of narrative and roughly sketch a system architecture to support enriched interactive fiction based on analogy with film techniques. This work is significant because it brings us closer to our goal of making interactive fiction into a rich, high quality artistic medium.

The authors can be reached at the School of Computer Science, Carnegie Mellon University, Pittsburgh PA 15213.

S. Smith was supported by an Office of Naval Research Graduate Fellowship.

The views expressed in this paper are those of the authors alone.

Towards a Theory of Narrative for Interactive Fiction

Sean Smith Joseph Bates

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

February 17, 1989

1 Introduction

In the seventies, researchers at MIT, Stanford, and elsewhere created computer-simulated *Dungeons and Dragons*-style fantasy worlds, in which users could fight monsters and search for gold. On a practical level, this hacking gave rise to a whole genre of adventure games, such as *Rogue*, as well as *Zork* and nearly everything else produced by Infocom. On a theoretical level, however, this work opened up some interesting research areas. How can we create high quality synthetic realities (SR's) [8] [21]? What kind of artistic expression is possible[1] through interaction with such synthetic realities [9]? What do the systems behind the synthetic realities have to be like to achieve such expression [6] [7]?

From an artistic viewpoint, these adventure games were quite primitive (but then so were Thomas Edison's pioneering motion picture experiments). The field of artificial intelligence is providing many powerful tools with applications in world simulation and interaction—tools not used by these existing primitive games. "Interactive Fiction" (IF) refers[2] to the artistic medium of which existing games are primitive examples; current IF research at CMU consists of imagining what some non-primitive examples would be like, building AI-based SR systems that move in this direction, and preparing to recognize and use appropriate AI tools as they become available.

To date, IF researchers have concentrated primarily on improving the quality of the simulated worlds,

[1] Chapters 11 and 12 of [25] present an interesting non-technical discussion of these issues.

[2] Here at CMU when we refer to "interactive fiction" we mean SR-based interactive fiction, as opposed to hypertext systems.

on plot management, and on improving the basic interaction tools (while idly speculating about full sensory interaction). While the systems (real and imagined) coming out of this work are promising, they suffer from a certain flatness: the user remains the user, wandering around in a realistically presented simulated world. Such systems can be interesting, but numerous interesting systems lie completely outside this paradigm. Consider worlds where the user is the homunculus inside some other character's head—the user interacts through this different personality and partially assumes it, for the duration of the IF system. Consider experiences with storylines not linear in space or time—they do not adapt well to the simple "wander in this" approach.

In order not to exclude these worlds, our IF paradigm has to be extended beyond mechanisms for world, plot and interaction to include mechanisms for influencing the user's mood[3] and for communicating in ways that lie beyond the scope of ordinary world presentation. We need a theory of narrative, a theory of how IF systems should select and order details presented to the user. In many ways, this task of selecting and ordering is similar to that of film-editing. We exploit this similarity to provide a basic narrative theory by analyzing and adapting some basic cinema techniques. We then consider the enhancements and constraints necessary for a distributed character-driven synthetic reality architecture (such as the OZ system in development at CMU [6]) to implement these adapted techniques.

2 Why: the Need for Another Dimension

2.1 Motivation

2.1.1 Magic Helmets

To date, the most serious attempt to place IF in an artistic (rather than a game-playing) framework is Laurel's thesis [21]. Drawing heavily on Aristotle's *Poetics*, Laurel analyzes IF in terms of theater principles and proposes a corresponding IF system, complete with a playwright expert system for plot management. One might interpret this model—a stage built any way one wants, with actors doing whatever one wants—literally and take it as the single sensible architecture for IF.[4]

For a number of reasons, we feel that this approach is too limiting. The literal theater model fails to exploit fully the potential of interactive fiction. IF need not happen on a real stage where the world follows the laws of physics and the viewer sees everything that happens nearby, but no more. Rather, the world involved is the world of the synthetic reality, which follows whatever logic it pleases; the choice of

[3] Apparently the role of emotional reaction is underestimated in literature as well—see [23].

[4] Laurel does not appear to take such a view, but neither does she address particular ways to broaden the framework.

which details are presented to the user, and how and when they are presented, need not be constrained by the rules of our universe. The theater approach excludes worlds which incorporate such alternate logics and presentation mechanisms.

The limitations imposed by the theater paradigm are especially rankling when one envisions some interesting IF systems that lie squarely among the excluded worlds. Suppose one doesn't want to see actors on a stage, or even be one of them (even making up one's own lines). Suppose one wants IF to be a magic helmet that would trigger the right neurons so that the wearer dreams he is someone else in another reality. Laurel uses examples from *Hamlet*; yet in these situations the user never ceases to be himself, in Hamlet's position. What if the user *is* Hamlet (at least partially)—what if the user's mind is manipulated by the system to try to make the user think/feel like Hamlet, not just experience Hamlet's objective experiences?

To make this vague notion a little more concrete, let us consider a storyline with some simply stated, explicit goals achievable through being a character. In Ken Kesey's *One Flew Over the Cuckoo's Nest* [19], the narrator Chief Bromden is an inmate of a mental hospital who regains his sanity as a result of interacting with the rowdy new patient McMurphy. The thrust of the novel is Bromden's cure; Kesey illustrates this cure by showing us the increasing clarity of the narrator's perception. Kesey plays all sorts of tricks to try to place the reader inside Bromden's mind. (Kesey commented in a letter that "you need to take the reader's mind places where it has never been before to convince him that this crazy Indian's world is *his* as well."[5]) In this spirit, what better way to experience the story than to be Bromden? *Cuckoo* as IF is aided by the fact that what really changes in Bromden is a superficial world-filter, the way external events and sensory impressions get processed before being fed to the homunculus. The sick Bromden suffers from fog episodes, muteness, hallucinations, and mistaken belief that his large, strong body is in fact frail and weak. The sick Bromden describes things with comparisons to machinery, the healthy with nature.[6] Another feature of *Cuckoo* that makes it convenient for IF is the basic irrelevance of Bromden's actions to the flow of the story.

Something that is quite infeasible as IF but is still challenging to think about is Joyce's *A Portrait of the Artist as a Young Man* [16]. Written as stream-of-consciousness, this novel reports the intellectual development of the young Stephen Dedalus. (Indeed, in both *Portrait* and *Ulysses* [17] Joyce succeeds, by using text only, in placing the reader not right behind a character's eyes, but a little further back.) The changes in Dedalus are no longer confined to sensory processing, but deal rather with deeper thought patterns. What kind of interaction mechanisms do we need to achieve this effect in IF? One imagines a computer dual of the user's character; the user has to act in consensus with this dual. There may very

[5] [28], p. 23.

[6] [28], p. 28,32. Also [24], pp. 14-15

well be no feasible answers to this question, but it is interesting to think about.

2.1.2 Issues Raised

Consideration of this vision raises a number of issues, which, while related, are sufficiently orthogonal to allow substantial independent research. We present the issues in order of increasing generality.

- *The innermost homunculus should remain unaltered.* In *Portrait* the reader sees into Dedalus' mind in order to understand *and judge* his point of view ([18], p. 45). The value of this experience would be destroyed if the user were completely Dedalus; while we want to foster user identification with a certain character, we don't want that identification to be so strong that it inhibits the user's ability to judge and appreciate.
- *We must allow the user at least an illusion of free will, even in the presence of plot.* As a deaf-mute (partially feigned), Bromden's role in *Cuckoo* is largely passive, so the modeled plot can be largely linear. A more active user character greatly complicates these issues. (Of course, if we've sufficiently convinced the user he's Hamlet, then the user (believing he is acting freely) will do exactly what Hamlet did, thus simplifying the SR's plotspace.)
- *The synthetic reality must have some type of significance.* The alternative worlds suggested aren't just abstract worlds sufficiently interesting to motivate exploration by the user. Rather, we have specific emotional and mental states we want to bring out in the user. These states need to be mapped out; the realm of plot enforcement has to deal not just with having the right physical events occur but also with having the user get to the right mental states. (In essence, we need to model both the physical world and some invisible underlying mental/moral/emotional world).
- *We need mechanisms to bring about interesting mental states in the user.*

This paper presents research into the last two issues.

2.2 The Need for Another Dimension

When objecting to the theater paradigm for IF, it is tempting to use terms "realism" and "involvement." The theater model stresses *realism* too much. It doesn't personally *involve* the user.

However, it soon becomes apparent that these terms blur a great deal of the picture. Doesn't a world where the user remains himself fully involve the user? Doesn't a world where the user is immersed into

someone else's subjective reality stress realism? After all, the interaction would feel more real than a be-yourself-in-another-world system.

Our initial description of some interesting but excluded worlds (such as Bromden/*Cuckoo*) similarly obscures the real issues by being overly simplistic. Do we really want the user to experience some character's complete subjective reality? No. Huss and Silverstein in *The Film Experience* [15] fault an experimental camera-behind-the-eyes film for being too real—the blinks and jostles recorded by the camera are distracting because while human eyes record these things, the human mind doesn't actually see them. How Bromden eats his food and where he keeps his toothbrush and whether he puts socks on both feet before putting his shoes on are details that are entirely unimportant. We don't want the user to be Bromden arbitrarily; we want the user to be Bromden so he can see how X and Y change with certain actions, and can see the symbolic aspects of the hospital confrontations.

What we really want to achieve in the excluded worlds is the communication of certain significant aspects of a subjective experience. In order to do this, user interaction needs to occur on the level of this significance. Synthetic realities (currently, physical world plus plot) must be deepened accordingly to include this context. We thus need to develop a theory to to deal with installing and manipulating this context.

Such a theory can conveniently fill the primary weakness Laurel sees in IF systems: the limited modality of the user/system interaction. Displayed text, simple graphics, and keyboard input fall pathetically short of the ideal you-are-there sensory exchange. Is this really a weakness? Huss and Silverstein observe that "the art of the film is born out of the triumph over limitations in the medium."[7] Such a statement also holds true for the art of interactive fiction—these limitations are what sets IF apart from other media, and by focusing on these limitations we will be able to discover the new forms of artistic expression that interactive fiction makes possible. In this spirit, we can regard the narrow communication channel as a tool. Controlling the information flow between the SR and the user can be a powerful way of influencing the user's mental state. ("Language is a Trojan horse by which the universe gets into the mind."[8]) Hence, the theory we need is a theory of narrative: a set of valves on this narrow channel and devices to automatically control them.

As we have observed, the excluded Bromden-style worlds were classified as ones where the user experiences strictly subjective reality (that is, the user *is* Bromden). However, by eliminating the blinks and jostles we've already moved away from that—we have sacrificed some of the strictness in order to let the user achieve some greater understanding of the character's situation. Inventive IF researchers may wonder what happens if we use the narrative tools proposed for the Bromden worlds but sacrifice all of

[7] [15], p. 145.
[8] [18], pp. 34-35.

the strictness. What greater understandings would be communicated? Attaching this axis of narrative tools to the basic theater paradigm creates a whole new space of possible worlds. Bromden worlds are finally admitted; so are a wide range of worlds we haven't thought of yet.

3 How: Importing Cinema Techniques

3.1 Borrowing a Kernel

Rather than trying to build a theory from scratch, we decided to explore importing mood manipulation and communication tricks from other media. Although transplanted, this core once established could then grow and become unique to IF. (This approach is not new—when inventing cinema, D.W. Griffith studied and adapted literary techniques used by Dickens.) The areas considered for harvesting included poetry, psychology, advertising, film, and music; Joel Bloch (a CMU graduate student in English with particular knowledge of film, and a member of the CMU IF group) finally suggested film-editing.

Upon reflection, cinema is an appropriate parent medium. IF is similar to cinema in that its presentation of data is bound to real time (unlike the presentation in a novel, where the reader can change his speed and jump back and forth in the text) and it is very limited in the data that can actually be presented.[9] In film, a few facial expressions that have nothing to do really with a person's character nevertheless must serve as characterization—and they often succeed. Such simultaneous effectiveness and thrift are what we need in our narrative tools. Further, film is quite adept at communicating within a context of significance. The same scene edited five different ways will present one set of physical information but five different moods and sets of deeper ideas. This approach is promising—we need to catalog some basic tools in film-editing and to try fitting them to IF.

3.1.1 Basic Cinema Techniques

The following collection of standard film techniques was culled from Huss and Silverstein's *The Film Experience*, an art-film text taken at random from the shelves at CMU's library, and enhanced through discussion with Joel. *Experience* presents the techniques along with examples of their use and the resulting effect on the viewer. We summarize these, occasionally adding an example or two from our own experience, and quickly mention some possible applications.

[9]The relative poverty of synthesised worlds and interaction modes limits the amount of data IF systems even have to consider for presentation.

- Lap dissolve ([15], p. 22). This is a transition from scene X to scene Y consisting of simultaneously fading out X and fading in Y. The lap dissolve establishes a conceptual link between the two scenes; in *Frankenstein* a burial is dissolved into a grave robbery.

- Pan shots ([15], p. 22). In a pan shot, the camera smoothly scans the scene. *Frankenstein* also used these—"what new horror will the camera discover?" Pan shots could also be used to present material for which smoothness is appropriate.

- Strange camera angles. ([15], p. 22) In *Frankenstein*, fear and unease are elicited by "scenes rarely photographed from normal lines of vision." Unusual viewpoints suggest unusual situations.

- Close-ups (close shots of details) and cross-cuts (parallel development of physically separated scenes) ([15], p. 24). (D.W. Griffith got these from Dickens.) Close-ups give a scene a sense of detail, of reductionism. Cross-cuts allow the film to establish links between the scenes presented in parallel.

- Repetition of a device ([15], p. 27). In Hitchcock's *Marnie* the lead character has for various reasons some psychological problems linked to young blond female children. To underscore these problems, such children show up in the background at points in the film significant to these problems.

- Quick flashbacks ([15], pp. 25,27). Brief but important scenes from a character's past are replayed at selected times. In *The Loneliness of the Long Distance Runner* the obsessed runner keeps recalling the face of his dead father. Like repeated devices, repeated flashbacks become a sort of "rhyme or refrain." (Singular flashbacks allow for non-linear time sequence, which allows incidents from the past to be presented in a manner which better stresses their relevance.)

- Symbolic camera movement and positioning ([15], pp. 30,37). *Citizen Kane* provides numerous examples. As Kane's mistress sings, the camera pulls higher (mimicking the soaring of her voice) to reveal two stagehands in the catwalks showing their disapproval; the camera shoots down at Susan, forcing the viewer to consider her condescendingly.

- Visual rhythm ([15], pp. 42,43). *Experience* uses this term to describe the regular, coordinated linking of things like images, movement, and actions to time. Huss and Silverstein cite several examples of battlefields and marching soldiers; the purpose of the technique is apparently to provide some deeper aesthetic consistency.

- Conflicting movement ([15], p. 46). A somewhat pretentious example of this occurs in the conclusion of Fellini's $8\frac{1}{2}$, where Guido confronts and makes peace with his past, which consists of a ring of people. These people dance in a circle while Guido dances the other way; he finally changes his direction to coincide with the that of the others.

- Distortion of natural rhythms ([15], pp. 48,107). Lindsay Anderson's *This Sporting Life* features a nightmarish rugby match. That the scene is in slow-motion and without sound enhances its nightmarish aspect.

- Zoom-freeze ([15], p. 52). The camera zooms in on some important facet of the scene and freezes there. This technique lends extra emphasis to that facet by arresting the viewer's attention.

- Iris ([15], p. 59). Irising to some important detail—like the anguished mother's hands in a courtroom where her son is on trial—means physically masking out everything else in the scene. This is similar to the close-up (indeed, for purposes of our analysis the close-up is subsumed by irising) except the exclusion of the non-emphasized details is more deliberate. (Irising in fact is somewhat archaic, and is not often seen in contemporary cinema.)

- Imagery. [15] has a whole chapter on this, most of which is naturally irrelevant to the non-visual realm of IF. However, the effect of a film can be greatly enhanced by visual allusion—Huss and Silverstein cite echoes of Christ's passage to Calvary in *On the Waterfront*. Put simply, we can obtain interesting inheritance along the lines of reference. It would be nice to duplicate this effect in IF.

- Nearly subliminal "flash shots" ([15], p. 108). In the development of scene X, the film quickly flashes some image Y that recalls or emphasizes some important idea. *Easy Rider* foreshadows important scenes such as the funeral pyre via such flashes. The most extreme example of this technique is probably Friedkin's use of actual subliminal shots to try to heighten the horror of *The Exorcist*.[10]

- Momentary excursions into subjective reality. An example comes to mind from Lindsay Anderson's *If...*: the bookworm Peanuts allows the virile Mick to look through his telescope, focused on the stars. Mick instead trains the telescope on a window across the way, where a girl with whom he is involved appears and waves at him. It is completely beyond the physical reality of the film for the girl to be in that building, let alone to be able to see Mick; nevertheless, the scene helps to characterize the boys and is acceptable in the slightly experimental context of the film.[11]

- Reminding viewers that they are watching a movie ([15], p. 115). *Experience* discusses the actors' asides to the camera in *Tom Jones*; Joel (in conversation) cites asides and subtitles in several Woody

[10]Whether subliminals actually work is an open scientific question; [22] contains a good survey of this subject.

[11]It is interesting to consider the aptly named *If...* [3] [27], a somewhat surreal discussion of alienation, repression and rebellion in an English boarding school. The occasional subjective intrusions (not following any hard set of rules, like "black and white = fantasy," but flowing organically) contribute to a whole that strikes some (such as one of the authors of this paper) as powerful visual poetry that achieves several levels of resonance. It strikes others (such as a certain housemate of that author) as uninteresting nonsense. The vast difference between these reactions radically demonstrates the importance of making the user comfortable with the context of a non-realistic IF reality.

Alien films. These examples all feature a bit of camp; use of this conceit heightens the sense of camp and allows direct communication of characters' thoughts.

- Contrast between shots and subject matter ([15], p. 118). Godard presents a couple having a Serious Intimate Discussion, except the camera shows only the backs of their heads. This suggests that the conversation isn't as important as it sounds, or that the participants are unnaturally detached from or unware of its import.

- Presenting a scene through a sequence of myopic close-ups ([15], p. 119). Huss and Silverstein cite a love scene presented as such a mosaic and claim the technique rendered the scene even more erotic as it provided no sense of gestalt—forcing the viewer to use his imagination. *The Graduate* presents one of Mrs. Robinson's attempts to seduce Ben in this fashion and thus allows the viewer some firsthand insight into the heady, aroused, and confused state in which the incident leaves Ben.

- Voiceovers. Joel cites the theory of narration in films—no narrator, omniscient external narrator, character as narrator, etc. Each of these has its own purposes in communicating information to the user; for example, character-as-narrator voiceovers communicate directly the information spoken and indirectly the beliefs and opinions of the speaker.

3.2 Adapting these Techniques to IF

We now want to consider how these techniques could be adapted to and implemented in interactive fiction systems. These implementations may very well require extensions and modifications to currently proposed architectures; hence, before we can sensibly discuss these implementations we need to discuss briefly the base architecture.

3.2.1 Basic Synthetic Reality Architecture

Laurel proposes a synthetic reality architecture that combines a playwrighting expert system with a distributed world/character model [21]. At CMU we are building such an architecture—OZ—that de-emphasizes the monolithic playwright control in favor of a more distributed approach [6] [7] [8].

Currently, an OZ synthetic reality consists of a physical world graph, object models, and character models. The world graph indicates rough 3D arrangement, plus relationships between objects such as containment. In a *Deadline*-style world [5] [10], this graph specifies the rooms in the mansion and their connectivity; the contents of the rooms are listed with containment and other structures but without explicit physical coordinates. Characters are rule-based entities that operate in parallel but also in lockstep: they

all perceive, then they all choose an act, then the system performs the chosen acts. Plot enforcement is currently embedded in the behavior of objects in the world and in the rules driving the characters (in the form of scheduled actions, for example). Our guiding principles include maximizing the liveness of the world; that is, we want to minimize the inclusion of author-supplied text so that the world can react properly to the great variety of possible user actions.

Related research at CMU (besides what this paper presents) includes exploring ways to enrich the characters, user interaction, and world model, and developing better methods to express and enforce the plot. Our ultimate goal is a computer-assisted writing environment with which authors can interact directly (rather than through programmers) to create high quality, artistically legitimate interactive fiction.

3.2.2 Implementing the Techniques in IF

How can we implement the film techniques in interactive fiction? To answer this, we must first imagine how the implemented technique should affect user interaction and then consider the set of mechanisms necessary to bring about this behavior. (Notice that this set should include not only mechanisms to produce this behavior, but also mechanisms to decide to produce it.)

Our analysis of implementing some film techniques presented earlier suggests a core of architecture extensions necessary for the entire set. We introduce this core by discussing a motivating sample technique.

Story Context

The implementation examples will largely be drawn from the same (hypothetical) world, loosely inspired by Gagne's discussion [13] of George Romero's film *The Crazies*.[12] A strange plague has broken out in town. Infected people become insane and go around killing people and doing bizarre things. Not everyone exposed becomes infected. Many of the infected people appear quite normal in between their random acts of violence.

Things are in a state of chaos. The National Guard is running around herding people into the gym; soldiers in white suits and gas masks are everywhere. It is rumored that the plague is the result of some mishap in the military's handling of a biological weapon.

The overall tone of the story is chaos. One overriding theme is the fragility of modern life: on a social level, the proximity of normal life to disorder; on an individual level, the proximity of the mental states labeled "sane" and "insane." A lesser theme is the ineptness and inefficacy of the scientific and

[12]Indeed, Gagne makes the point that the film is effective because Romero unifies the film's subject with the way it is presented.

military establishments, those bastions of authority and order.

The user character (UC) is Our Hero, a young scientist connected to a local university who is defying martial law and scientific establishment by sneaking around discovering the plague's cause and pursuing a cure.

A Sample Technique: Architectural Implications

Imagine a scene where Our Hero is trying to do something against a chaotic background. Suppose he's got his hands on some classified military document and is trying to extract the relevant information while hiding behind a tree. Soldiers are everywhere; there are only a few minutes before Our Hero will be discovered and have the document removed from his possession.

Once the UC has hidden, a conventional IF system would summarize what's going on in the surrounding area and then let the user read the document until he is caught.

> You begin to read the document. You only have a few minutes. The town green presents a scene of disorder and chaos. Soldiers are marching around, shouting orders as they round up the inhabitants. Some of the townspeople struggle. Sweet old Mrs. Buckman sits knitting on her front porch. A soldier approaches her; she kills him with her knitting needles.
>
> *The user interacts with the document...*

In a situation like this, the myopic close-up technique would be appropriate. The staccato sequence of details would help convey the urgency of the UC's situation; the lack of overall coherence would mimic the lack of order in the town and the failure of the UC to figure out the cause and cure of the plague.

> You are holding the document. You know that you only have a few minutes.
> OPEN THE DOCUMENT AND READ.
> *a sentence or two from the document*
> "What the hell's going on? You can't do this!"
> "I'm sorry sir; we have our orders." The soldier clubs the man.
> Mrs. Buckman knits on her porch across the way.
> *a sentence*
> "MOVE SQUAD 17 OVER BEHIND THE POST OFFICE!"
> A youth is kicked into submission.
> *a sentence*
> A soldier approaches Mrs. Buckman.

The Amberson family and some friends scream when they see the corpses in the gazebo.

a sentence

"WHERE ARE THOSE ROUNDS WE NEEDED?"

Mrs. Buckman is splattered with blood.

"JENKINS—CHECK THOSE TREES OVER THERE!"

a sentence

Tear gas is lobbed into the church picnic.

"My baby! How could they have done this?"

A strand of yarn from the unraveling sweater is anchored in the knitting needles planted in the soldier's chest, as he crawls through the daffodils.

a sentence...

What's going on here? The user's perusal of the document is interleaved with details from the surrounding environment. These elements are chosen and assembled automatically by the IF system so that the resulting sequence is grating and lacks coherence.

Even this quick analysis suggests a number of necessary architectural constraints. What's doing the interleaving? We need some kind of Editor that takes care of selecting and presenting these details.[13] Why should the Editor choose this technique and not some other? In addition to being able to implement this technique, the Editor also needs some mechanisms for choosing to implement it.

How does the Editor get a handle on the surrounding environment? Previously we imagined the environment modeled by a physical world graph, with character interactions arbitrated and performed by a central loop. This architecture (perhaps with decentralized control) is a good base for implementing myopic close-ups. For one thing, the Editor needs to be able to discover what scene locations are sensorily accessible from the location where the UC is. For another, what's happening in scenes beyond the perception of the Editor is essentially irrelevant (more on that later). Interaction, like character position, is localized. This suggests[14] that each character should be tied to the particular node in the world graph corresponding to its physical location (the Editor in some sense would ride piggy-back on top of the UC). Each of these nodes would then have an interaction arbitration procedure attached to it, responsible for the interactions of the characters attached to that node. Both interaction cycles (gathering and processing basic user reactions) and Editor cycles (selecting and collating data and sending it on to the user) would require a small local traversal of the graph to see what's going on around it.

[13] One may conceive of the Editor as a sportscaster describing a hockey game, or as a TV director looking at 16 screens and selecting in real time which view to transmit. As we shall see, this conceptualization is slightly flawed, as the Editor's interaction with the SR may not be this passive—the sportscaster analogy may only be appropriate if the sport is something like roller derby or professional wrestling.

[14] In fact, the OZ architecture conveniently moved in this direction, independent of the research in this paper.

This idea of a local traversal raises the issue of what relations should be local in the graph. Research is already underway here at CMU into enriching a basic scene graph to indicate more complex physical (the post office is across the town square from the park) and sensory relationships (it is possible to hear very loud noises through the wall).[15] Both these scene nodes and the links between them all have very physical bases. It's very intriguing to consider what other kinds of information could be embedded here: edges between nodes containing parallel characters? The graph is supposed to make local the information needed for interaction—what information besides sensory data might be needed for interaction in a synthetic reality?

What are these elements that the Editor is choosing? Each of the neighboring scene nodes has to be able to talk about what's going on there. Since this information not only is computer-generated but also needs to be manipulated and semantically analyzed by the system, it cannot be in text form. Hence, we need to install some computer-palatable system of knowledge representation. Schank's conceptual dependency machinery [26] suggests itself, but this is only suitable for basic action data. We see that this issue will require a great deal of thought. In our motivating example, data pieces are chosen for their shock value—we need the ability to perform rough semantic analysis, so that the Editor can call heuristic evaluation functions to estimate how shocking or soothing or whatever a data piece is. We also want the Editor to select a set that is not quite coherent—hence we need an estimate of how completely a data piece represents what's going on in a scene node. Notice that coherence is not a static notion—"what's going on" involves events over a window of time. Notice also that one could present complete information about a scene, but with sufficient fragmentation it will still be incoherent. Making these evaluation functions simple to compute will be a major issue in designing our knowledge representation system.

The need to determine how one particular data-piece from a given scene-node fits in the context of all the data-pieces coming out of that node suggests that what we really want the node to produce is not a set of discrete pieces but rather a single hierarchy. The deeper a piece in this hierarchy, the more detailed the information: the root might indicate the characters present, the first generation of children that they are having a conversation, the next generation a summary of the conversation, the next the actual sentences exchanged. A benefit of using such hierarchies is computational efficiency—we imagine that frequently only a few nodes in this hierarchy will need to be expanded. This illustrates a principle of lazy evaluation which will apply often in our IF system sketches. Besides yielding computational efficiency, this principle allows greater thematic flexibility—not specifying the color of the couch in advance means that the Editor can color objects red when it needs red objects. Setting up a system to generate details when needed, rather than specifying all in advance or calculating everything at every simulated moment,

[15] One challenge is how to deal with telephones, for example—that for the duration of the phone conversation there is an audio link between the telephone in Juan's office and the telephone in Su-Yuen's office. An aspect of a physical world graph that our scene graph fails to inherit is permanence over time.

also may protect us against pathological users who want cut the away the couch's upholstery and scatter the stuffing around the room—the user can dive ridiculously deep in the detail hierarchy, if he so chooses.

Coherence evaluation raised the issue of time windows. The Editor's performance of a technique such as myopic close-ups also raises these time issues. The technique as presented encompasses quite a few simulated clock ticks.[16] This suggests that the Editor's basic cycle length (the expected duration of a technique) should be quite a bit longer than the user tick. The Editor should be prepared to discontinue the technique if the user takes some action that violates some precondition or makes the technique unnecessary for some reason. We've introduced at least two different clock cycles now, one for UC interaction (for interaction local to the scene-graph node where the UC is), and one for the Editor. Which one should be used in the interaction monitors of all the other scene-graph nodes? The principle of lazy evaluation would suggest something slower than the user-interaction cycle but quicker than the Editor-technique cycle; we want the system to produce as much information the Editor needs, and no more. (However, this approach will make it difficult to extend the architecture to deal with multiple users.)

Somewhat Feasible Techniques

When we consider adapting the film techniques outlined earlier, we find that for some of them actual implementation would be infeasible. Quite a few others seem possible, however.

The Decision Mechanism

Earlier we just glossed over the mechanism for deciding to perform some technique. The basic approach we had in mind was considering the techniques as a set of operators, each with its own set of results and user mood transformations, and each with its own set of necessary and suggested preconditions. Editing is then planning a path through mood/plot space using these operators. One might use classical planning techniques or more recent ideas on theories of activity ([1] [2]) to implement the decision mechanism.

This approach raises the issue of how exactly user mood and plot are embodied in the system, and what it is about them that the Editor is checking. There has been research on plot enforcement but little exploration so far of mood. In keeping with the principle of making our system character-driven, we believe it would be better to maintain a UC mood model rather than have a large explicit mood/plot space in which the user wanders. (We can achieve greater "liveness" in the system and greater complexity in mood space by concentrating on the functions generating the space rather than the explicit space graph.) Goal/plan information[17] could be incorporated into this model as well. This model would be essentially

[16] A "tick" is a basic user interaction cycle, in current OZ terminology. It is usually intended to represent a few seconds of simulated time.

[17] Frequently, characters' goals and plans are artifacts of the world's author and the fact that the world and a character's

the same model used to implement other characters.[18] The system first gathers information from the author regarding a character's destiny and personality quirks, what sort of moods and attitudes would be triggered by what sort of world events. The system then takes this information and brews up (using some unspecified but wonderful AI engine) a character model that monitors the world and keeps some system-visible set of mood registers up to date. It is these registers that the Editor checks to determine what mood the UC is in and thus what mood the user ought to be led to.

We suspect that some data important for shaping the user's experience will not be easily embedded in a UC model, and so imagine some plot monitor might be necessary. Again, though, to maximize liveness we'd like to install some AI engine that watches what happens, rather than an explicit plot graph. Perhaps certain Editor actions should be demon-triggered—when *foo* happens, an alarm in the relevant character model goes off, and the model interrupts the Editor and informs it. This device would be needed for events (such as critical plot developments on the other side of town) that the Editor should know about but cannot reasonably be expected to discover during the limited local search it performs. (The use of such demons is an alternative to installing non-physical thematic-link edges in the scene-graph.)

We suspect that despite our intention to make the system "live," we may need to include some canned plot information that cannot be easily embedded in the characters' destiny boxes. Hence, we refer to a plot box as well—we use this term to stress that this is more a register or set of flags than an explicit plot graph (something antithetical to the spirit of liveness). Perhaps much of the plot-box control can be built into the demon-triggered Editor actions mentioned earlier.

It soon becomes apparent that many of the film techniques can be easily adapted and implemented, but no easily implemented set of heuristic triggers can be specified. Use of each technique is linked to deeper aesthetic vision beyond the scope of the machine. Hence, we see a rough dichotomy arising between the directly automatable techniques (that would be supplied as part of a kernel package) and the ones whose calling mechanisms require substantial creative input from the world's author (these would be supplied as customizable options).

Automatable Techniques

- *Myopic close-ups.* See earlier discussion.
- *Lap Dissolve.*

existence in it have some pre-ordained meaning. As characters don't necessarily know or even intend their goals, we suggest that "destiny" be substituted for "goal" in such discussions.

[18] This approach would facilitate multi-user (or zero-user—see [8]) IF and systems where the user can change characters on the fly.

Here's a situation where the Editor must take into account interaction—it only seems sensible (from our current primitive perspective) to use it when the UC is going from one scene to another. Suppose Our Hero wants to walk from the tree to Mrs. Buckman's porch.

> WALK TO MRS. BUCKMAN'S PORCH.

A naive approach is to keep reporting the continuously varying scene.

> You walk through the softball field.
> You pass the bleachers and the dead boy scout.
> You hear the panicking church picnic crowd off to the left.
> You cross the street.
> You walk up to Mrs. Buckman's porch. She smiles sweetly and says "hello."

This approach greatly complicates scene-graph maintenance—just imagine all the possibilities for interaction as the UC walks. We could get around this by sometimes deciding to whoosh the UC right from scene-node to another (as in chutes and ladders) and not allowing for any interaction. But by not being explicitly announced, these transitions would appear to the user as troubling "dead spots."

An explicit way to perform this transition is to remove the dead spot—cut from one scene to another. (Indeed, the direct cut, something too trivial to include in the film technique list, should probably be listed here.) Explicit cutting also provides a nice way to mask the granularity inherent in the underlying synthetic reality.

Why would we want to do a lap dissolve rather than a direct cut? Perhaps when something significant will be communicated by mixing the two scenes. Maybe a young guard had just discovered Our Hero, who was forced to struggle with the guard and knock him out. The UC then decides to walk over to Mrs. Buckman's porch. Link and contrast: violence against innocent men. Suggestion: the UC's nobly motivated action in knocking out a threatening guard is disturbingly similar to the murder just committed by the mad Mrs. Buckman.

> WALK TO MRS. BUCKMAN'S PORCH.
> Her turn-of-the-century house is charming, but up close you can notice the paint peeling.
> The young guard lies sprawled in the bushes.
> The white-suited soldier is twitching in the daffodils.
> The young guard's eyes are rolled back. Blood trickles out of his nose.
> The twisted yarn attached to the knitting needles is caked with blood.

Mrs. Buckman smiles sweetly and says "hello."

To implement such a lap dissolve, the Editor just has to recognize that the UC's destination scene has some data-pieces that are very similar to pieces in the UC's source scene. To make things less confusing for the user we could somehow signal that all these details are not occurring in the same place (in the example above we used different fonts to indicate this; perhaps opening a second window to correspond to the second location may be more effective). We once more see the need for a hierarchy of data-pieces: similarity could then be determined by comparing some higher level aspects of this hierarchy. Performing the transition just consists of selecting a few of these similar pieces and interleaving their presentation.

Lap dissolves could also occur for more complex reasons than just similarity of details—perhaps there are thematic links (such as the UC/Mrs. Buckman comparison we suggested). To recognize such links the Editor needs to be able to evaluate the scenes, and to examine the overall plot structure and the UC's goal and personality structure. It is difficult to see how we could have the Editor recognize these deeper links in the absence of explicit author preprogramming ("be on the lookout for *this*").

- *Pan shots.* We can communicate the proximity themes in this sample story when presenting a description by smoothly scanning the town square's jarring mixture of 4th of July Americana and chaotic violence. Mixing generally soothing details with occasional shockers makes the presentation jarring as well.

> You duck behind the tree. Beyond the tree is the softball field. There are wooden bleachers on one side, with a sign saying they were donated by the town's Rotary Club. There's a dark reddish-brown smear on the sign, and below it lies the crumpled body of a boy scout. There are some empty Coke cans on the ground. Further down the path is the picnic area. A charcoal fire is burning, and some hamburgers are sizzling away. A man with a cleric's collar and a chef's hat has his face buried in his apron as he stumbles about blindly. The scene becomes largely obscured by the clouds of tear gas. Sneakered feet pound through the grass, among the spilled potato salad, softball equipment, and yelping puppies. Across the street is old Mrs. Buckman's house, with her carefully tended garden and windchimes. Smiling sweetly, Mrs. Buckman wipes the blood from her face. A sweater at her feet is unraveling—a strand of yarn goes down the porch steps, around the milkbox, through the daffodils to a knitting needle anchored in the chest of a white-suited soldier lying there.

To implement such panning, the Editor scans the sensorily accessible nodes in some natural order—

perhaps left-to-right when viewed from the UC location.[19] From these nodes it selects pieces that are moderately coherent and fit the template of a generally soothing piece sequence interrupted at random moments by highly shocking pieces.

- *Cross-cuts (in-body).* Basically, cross-cutting is presenting two scenes in parallel by alternating images from each of them. This alternation is slow and deliberate—the intention is for the viewer to digest both sets of images, and to understand that these are separate scenes being presented.

  Our IF system is loosely bound to the principle of presenting the subjective reality of the UC. The idea of cross-cutting prompts one to ask, in light of this principle, whether the UC is to be in two places at once. Two types of cross-cutting arise from the two answers to this question. Here we consider "in-body" cross-cutting—presenting two scenes that could actually be observed from the same physical location. (Later we examine "out-of-body" cross-cuts.)

  Consider once more the situation where the UC is hiding and trying to read quickly a stolen document. Perhaps the Editor needs to raise a sense of urgency in the user without instilling as much panic as a technique such as myopic close-ups would. A solution? Choose a shocking scene within sight of the UC, and interleave pieces from that scene with pieces from the UC node in such a manner that both scenes come across in a somewhat coherent manner.

  sentence from document

  Across the green Mrs. Buckman sits on her front porch, knitting a sweater. A white-suited soldier approaches her.

  sentence from document

  The soldier starts to speak to Mrs. Buckman. Smiling sweetly, she stands up and, carrying her knitting, walks over to him.

  sentence from document

  Mrs. Buckman plunges the needles into the soldier's chest. Blood spurts onto the smiling old woman.

  sentence from document

  The soldier crawls away, trailing yarn from the unraveling sweater.

- *Zoom-freeze.* Perhaps the Editor needs to draw the user's attention to some object or detail. In our sample world, maybe it is the people that have been eating yogurt that have remained uninfected. Our Hero is talking to yet another uninfected person, and notices that there is yogurt in the room. Normally the yogurt's presence would be mentioned in passing (that it is important enough to be mentioned at all should be built into the SR somehow). We could draw the user's attention to to

[19]This suggests that scene nodes need to contain positional information as well.

the yogurt sentence by freezing the output of text for several seconds (if that wouldn't be confused with time manipulation), or perhaps by displaying the sentence in boldface, or something like that.

When actually implementing this technique, the Editor collects the data pieces it is considering for presentation. By examining the UC's destiny box and the overall plot box, the Editor recognizes that a certain item in this collection would be a good candidate for zoom-freeze emphasis. (Perhaps that recognition will give that piece higher priority for presentation.) When and if that piece is presented, it is presented emphasized.

- *Iris (for emphasis).* Like zoom-freezing, irising is a way to emphasize a certain detail. Irising however emphasizes a certain piece in a scene to the deliberate exclusion of the other pieces. Conceptually, we can look on this as an unbalanced cross-cut between the big global scene and the emphasized aspect—we only present enough of the big scene to let the user know it's there.

  We see two purposes possible for irising: to stress some aspect of a scene, or to influence the user's mood by forcing him into some type of tunnel vision. We deal with the former—using iris for emphasis—here.

  Returning to our sample world, perhaps it's vitally important that the user recognize the importance of the yogurt and get on with solving the mystery of the plague, so when the UC walks into healthy Libby's kitchen the Editor says "You have entered Libby's kitchen" and then launches into a discussion of the yogurt, not letting up until the user finally walks over and picks it up.

  An implementation of irising for emphasis might have the Editor decide based upon the overall plot box and the UC's destiny box that thing X needs to be stressed. The Editor then performs the appropriate action (which we modeled as a cross-cut). (We imagine that what may in fact happen is that a demon in a scene node recognizes that thing X is going on and that it might be important, so either it informs the Editor or demons in the destiny and plot box do.)

Techniques Requiring Significant Author Customization

- *Cross-cuts (out of body).* This technique involves presenting two scenes—at least one of which is beyond the sensory reach of the UC—to the user via cross-cuts.

  In our sample world, perhaps another scientist discovers a valuable clue but is mistakenly killed before it can be revealed. One can conceive of many reasons why it is important to communicate this to the user—stressing the blundering theme, letting the user know that the newly late scientist's work was on the right track, etc.

  It must be clear to the user that this scene is happening somewhere else. We can use an alternate font, or perhaps have another window open up. In that window...

The chemistry lab. Professor Mait Jones removes a test-tube from an incubator.

Jones puts a few drops on a slide and examines it under a microscope.

"This is it!" says Jones excitedly. "The human cells remained untouched!" He scribbles down notes in his notebook.

A soldier turns the corner. Jones approaches him with the test tube.

The nervous soldier perceives Jones' agitated test-tube waving as a threat, and guns him down.

Two other soldiers throw the body—including notebook—into a pyre set up outside.

While it would be very nice for the Editor to be able to see the Lab node from the UC node (one imagines "important parallel character" links in the scene-graph), we cannot reasonably expect this will be the case. Hence, the Lab node has to have enough intelligence to realize "hey, it might be important for the user to know what's happening here" and go tap the Editor on the shoulder. This suggests that we need scouting demons attached to each scene node, demons sufficiently aware of plot and events and destinies to know when things are important. We fear that building completely general watch-demons is too infeasible; such demons need to be highly customized for individual worlds.

Note also that there is nothing to prevent the Editor from scanning out from the Lab node, once it has called attention to itself, in order to present that node's data pieces in a nongeneric fashion. (One also wonders what effects can be achieved by letting the user's actions in the UC node influence the events in the out-of-body node. What if the UC runs over to Mait's lab?)

- *Repetition of device.* In this technique we establish a link between some physical detail and some mood or deep idea; we then be sure that this detail shows up in the simulated world at appropriate times.

Perhaps the UC had a girlfriend who died of the plague. Say she had red hair and a distinctive accent. The UC is traumatized by the death. During stressful times, we'll have the UC notice that certain victims around him have red hair or a similar accent...

reading the document

Over by the monument, a soldier clubs a young figure. Scarlet blood stains the red hair.

This technique is different from finding and reporting Mait's lab. The girlfriend's death is an important scene that has already occurred—either in the SR itself or before the simulation started. Either way, it is sufficiently significant to be stored somewhere. Perhaps we need a Significant Event List in the system's UC model. Events that occur pre-simulation will be hardwired in;

other events need to be recognized and added (perhaps these need hardwiring too—"watch for the girlfriend's death..."). If the Editor decides the UC is feeling especially stressed, it can examine the data-pieces it's presenting to see if the appropriate details from a traumatic scene in the SE List can be introduced. (Here we see once more a need for lazy evaluation in world simulation, so the Editor can impose occurrences of red hair when it needs them.)

- *Quick flashbacks.* Perhaps the UC saw his girlfriend killed when, driven insane by the plague, she attacks someone. When he witnesses other violence against other victims, this scene is briefly replayed.

> The nervous soldier perceives Jones' agitated test-tube waving as a threat, and guns him down.
>
> (She brushes her red hair out of her eyes. Smiling, she raises the butcher knife. The shot rings out and she crumples to the ground. But the hand you held so often won't release the knife even in death.)

Such a flashback can be used in events such as Jones' death to suggest the UC's fear that perhaps his girlfriend was mistakenly killed. It can also be used in instances when the UC thinks that perhaps some character is infected and will have to be killed.

Perhaps the UC was the one who pulled the trigger. The flashback can be gradually extended[20] to include this information. The growth of the flashback can reflect the growth of the UC's anxiety. Maybe the discovery that it was the UC that killed her, along with some piece of information that proves in hindsight that she wasn't infected (she was defending herself against an attacker outside the UC's view), can parallel the UC's discovery of a cure.

To implement this technique, we once more need to assume the Editor has a handle on this critical scene. We suspect that this handle can only be provided by author hardwiring. The flashback uses discussed above suggest several ways that flashbacks can be invoked: when the Editor wants to suggest certain UC feelings to the user, when certain kinds of scenes are being presented, and when certain preselected events occur. (This latter activation approach also requires author hardwiring.)

- *Rhythm.* What is rhythm? Huss and Silverstein use this term to describe the regular, coordinated linking of things like images, movement, and actions to time.

It's fairly easy to see how the Editor could implement deliberate rhythmic behavior: it would pick a class of objects that are going to be rhythmical (such as red things or plummeting household appliances) and an object-time pattern and then enforce it. Once more we see a need for lazy evaluation, so the Editor can find red things when it needs them.

[20] Extended flashbacks are already used in a current IF system—Disch's *Amnesia* [11].

A more challenging question is why the Editor would want to use this technique. It seems to us that the purpose of this technique in film is tied with those deep issues of the overall aesthetic effect of the film—not something easily specified for automation. These "deeper" purposes should eventually be incorporated as higher level rules (or meta-rules—rules specifying the slant of rules) in the Editor. The deep creative decisions would be made by a particular IF world's author, who would embed her biases in these rules.

It's a little easier to see how shallow uses of rhythm can be justified and automated. The first example that springs to mind is Anderson's ghost world [4]. The ghost's episodes of throwing things around have meaning: they tell how the ghost feels about the UC (alternatively, the status of the UC in the achievement of his goal of exorcising the ghost). The author may wish to put some rules in the Editor that recognize when the UC is in a situation which features the repetition of similar actions (data pieces?), and activates a rhythm scheme—pick a mood/time map (so that the frequency of thrown objects slows down when the UC says the right things, or so that the frequency of the thrown objects is faster now than it was the last time the ghost was throwing things because the ghost is angrier now) and use it. The Editor reads the ghost's attitude in the in the ghost's mood box, or in the UC's destiny box, or in the overall plot box.

In the above discussion, the timestream in which these events are rhythmically embedded is real time (that is, time as the user experiences it). Changing the relation of this stream to the stream of the reality is a technique (distortion of natural rhythms) discussed later. Note however that if this relation is changed, then the use of this rhythm method implies that the altered timestream will affect the world (the SR), not just the user's experience of it. (If the user is experiencing plummeting household appliances at 0.5 Hz, then these appliances need to plummet at 0.5 Hz real time no matter how fast or slow real time becomes with respect to time in the simulation.)

- *Symbolic camera position.* In a textual medium such as the IF systems we're considering, what is the equivalent of a symbolic camera position? We have concluded that it is a certain tone or descriptive viewpoint. Perhaps a certain officer is highly charismatic. A film might communicate his prowess and leadership by shooting upwards at him—he always appears as an imposing giant. In text, we could convey this feeling by appropriate word choice—his voice is stentorian, and he strides rather than walks.

  To implement such tonal control in a system with current NL processing technology, we cannot just adjust some control variables on a filter through which we send data-pieces regarding this officer. However, if the author can foresee such tone adjustments then she can attach to the character or object involved a tag specifying the customized vocabulary and style appropriate to that character or object. If this character appears frequently then the tag needs to be sufficiently rich in order to avoid having the resulting effect seem rather silly.

- *Conflicting movement.* The purpose of this technique is to convey through physical opposition some deeper relationships. In our plague story, perhaps Our Hero has to fight his way through a crowd of people (the church picnic stampede?) all rushing the other way. Besides being physically necessary, this could be symbolic of Our Hero's relatively unique reaction the plague, or something like that.

  We imagine this is another of the deeper techniques that the author would have set explicit rules for, according to her own creative mind.

  There are several levels to implement this, and which is best is not at all clear. The author could set up the interaction laws in the SR itself to see that this sort of thing happens occasionally. Perhaps we have plot managers, with fingers into the relevant character models and a global plot box, attached to each scene-node. Alternatively, we could try to do something with the Editor to recognize when there is opposing motion going on that could be symbolic, and make sure it is reported. Conceivably we could even use lazy evaluation to ensure that symbolic opposing motion occurs whenever the Editor wants it.

  The Plot Manager approach seems a little more reasonable, as using this technique may require deep manipulation of the SR, not the shallow (or nonexistent) manipulation usually performed by the Editor. Any use of symbolism in deep SR events raises this issue.

- *Distortion of natural rhythms.*

  Suppose Our Hero sees across a crowded square some close associate or friend who is about to be attacked. Our Hero sees the danger but Close Friend is unaware. For some reason the only way Our Hero can warn Close Friend is to run across the square to the Friend. But it's too late.

  In the SR, this happens in just a few seconds. We *could* present it that quickly.

> Across the square you see Close Friend. Behind him stands the high school football coach. The coach holds an ax, and seems ready to swing it at Friend.
>
> SHOUT "LOOK OUT!" TO CLOSE FRIEND.
>
> He cannot hear you. Coach starts swinging the ax.
>
> RUN ACROSS THE SQUARE.
>
> You run.
>
> You reach your friend just as the coach's ax has decapitated him.

We could increase the suspense of this scene (and the user's feeling of impotence in the face of the horror) by slowing down the presentation of the run. Maybe even we'd want to cross cut between the UC running and the coach swinging.

> Across the square you see Close Friend. Behind him stands the high school football coach. The coach holds an ax, and seems ready to swing it at Friend.

SHOUT "LOOK OUT!" TO CLOSE FRIEND.

He cannot hear you. Coach starts swinging the ax.

RUN ACROSS THE SQUARE.

You run:

Step. *(these details are fed out slowly and regularly)*

Step.

Step.

The ax moves closer.

Step.

Step.

The ax head is only a foot from your friend's neck.

Step.

Step.

You're nearly there. The ax breaks the skin.

Step.

Step.

You arrive. Your friend's freshly severed head plops down in front of you.

To implement this, the Editor just slows down the flow of time and magnifies details (ie, the hierarchy of data-pieces for this scene is grown far deeper than usual). Possibly the Editor needs to do something to block interaction (or at least make sure the user can't do anything that doesn't fit the slow motion context).

Slowing down the presentation would free up the computation power normally used to maintain all the out-of-sight scenes, so we could do some garbage collection now. What if we speed up time? Then it's possible that out-of-sight scene maintenance will be starved (if the UC is just rocketing around then scenes don't need to be modeled as richly, true, but adjustment in that direction is still limited).

Once more we see the need for a complete hierarchy of clocks and phases and ticks.

- *Subliminals.* The purpose of using nearly subliminal flashes is to suggest quickly some mood or idea without having to embed it in the linear progression of the SR.

  Perhaps Our Hero is afraid he'll become infected and go around doing horrible things. Occasionally a window can pop up briefly in which this fear is visualized.

  You stand smiling holding the butcher knife. The cuff of your flannel shirt is soaked with blood.

It's easy to see how rules can be set up to lead the Editor to this technique in order to express some mood or idea. The problem is that the domain of things that can be flashed is large and unwieldy. Unlike standard flashback fodder, these items can't be enumerated by just scanning the UC box for traumatic memories; the items and their activation conditions will need to be hardwired.

We could try to go directly for subliminal influence by flashing messages on the screen for, say, a sixtieth of a second.

> *People getting shot right and left...*
> flash: You're next.[21]

Should subliminals actually work, this approach would heighten the user's tension and fear through direct subconscious suggestion.

- *It's only IF.* The film examples provided by Huss and Silverstein (*Tom Jones*) and in conversation by Joel Bloch (Woody Allen) all feature a bit of camp. Use of this conceit heightens the sense of camp and allows direct communication of characters' thoughts.

  Can this technique be used outside of a camp context? We don't think so (but we're sure someone will think of a way).

  What would this technique look like in IF? A direct adaptation would have characters stepping out of context and addressing the user directly (perhaps because the Editor was told by a plot demon that it's important for the user to learn the information this character can reveal). A more creative approach would poke fun at the limitations of IF: having characters discuss their rules of interaction (this behavior of course would be programmed into the rules of interaction), or having the narrating voice joke with the user about having to go do some disk fetches (perhaps to mask some lengthy searches).

  We suspect that this technique will only be useful in a small number of IF worlds (after all, overuse would destroy its effectiveness).

Infeasible Techniques

These are techniques for which we cannot conceive adaptations for IF, or whose IF implementations would be too difficult or unwieldy. As with all limits, however, we encourage the reader not to accept these judgments but rather to regard them as challenges to the reader's creativity and imagination.

- *Voiceovers.* Joel Bloch explained to us the theory behind narrators in films—no narrator, omniscient external narrator, character as narrator. This is an interesting issue to consider in relation to

[21] This example was suggested by Peter Weyhrauch.

interactive fiction. Narration is crucial to such text-based systems—whose voice is it that keeps addressing the user?

Things such as commentary by a character within the story fall under such techniques as "it's only IF." But something we have not considered is attaching a personality, with quirks and attitudes, to the narrating voice. Having the narration occur from the subjective viewpoint of a character in the story opens up some interesting avenues for communication, provided that the user is sufficiently certain of the underlying truth to perceive the subjectivity.

Such an approach to narration requires both effective non-narrative communication and powerful "live" personality simulation and NL generation, all of which makes us label this technique as interesting but quite beyond the scope of what we expect to achieve with OZ.

- *Strange camera angles.* The purpose behind strange camera angles (or viewpoints) to heighten a sense of unease, of disorder.

> Through the beakers, Jones' face is distorted and gray beard tinted yellow.
>
> A rippling figure in a tinted yellow suit appears behind him. Jones' yellow arm raises a wavy test-tube. Shots ring out.
>
> Dark yellow stains appear on the back of Jones' light yellow lab coat. The body falls.

We could report the stampede at the church picnic from ground view: dust, towering figures, falling potato salad meteors.

These two examples are entirely visual—what textual techniques (that wouldn't be too hard to build into a natural language generator) would qualify as the equivalent of "strange camera angles?"

How to implement this? We could just glibly say that the Editor, knowing that it needs to make the user uneasy or something, decides to present the data pieces from an angle *Foo*. We need to twist these pieces to reflect *Foo* before generating natural language from them. Should the scene-nodes be required to spit out data pieces not only relative to some incoming edge, but also to arbitrary *Foo*? That's not too satisfying. Should we require some semantic structure processor that takes as input a set of structures and some *Foo* and produces the twisted pieces required? That's more satisfying, but is it feasible?

- *Iris (exclusionary).* As we mentioned earlier, an alternative use of irising is to introduce a feeling of tunnel vision. Maybe we want the user to feel like he's stumbling around blind in some dangerous situation, so rather than hearing about the various stampeding people surrounding him, the user is told about his shoes and what they're standing on. This description is interrupted frequently so the user can be told he's just been bumped into.

We don't see at all how to automate the process of selecting which irrelevant detail on which to iris.

- *Imagery.* The purpose of imagery is clear enough—allude (visually or otherwise) to something in a standard library of images, so that the meaning of a given scene is enhanced by the meanings attached the standard library image. User processing of this connection is frequently on a nearly subconscious level.

  When we're dealing a non-visual medium like IF it is very difficult to perform image allusion. The closest thing we can think of is literary allusion.

  Our hero is going to speak to Mrs. Buckman. To help communicate the proximity of normal life with disaster, we want to have her say something very proper and ordinary to contrast with the bizarre things that have been going on. We borrow from a similar situation in Dylan Thomas' *A Child's Christmas in Wales.*

  > You walk up to Mrs. Buckman's porch. She smiles sweetly and asks "would you like something to read?" She says the right thing, always.

  How to implement this kind of allusion? Presumably, the Editor has enough information to decide that it wants to communicate *foo* and it wants to do that through some given set of data pieces. It is tempting to proceed by saying that a library of standard images will exist and the Editor can scan the library to see if there's a standard image which fits one of the data-pieces and carries meaning that contributes significantly to *foo*. In fact, that seems the only way to automate this sort of thing. But building up such a library—which would have to be customized for each SR—seems quite ambitious.

  An even more ambitious approach is to embed the reference not just in the images but also in the text style (one thinks of the power this technique gives to the "Oxen of the Sun" chapter in *Ulysses*). Sadly, we suspect that rendering this anything more than a dream requires greater dexterity than current NL processing will provide.

- *Momentary excursions into subjective reality.* The purpose of this technique is to communicate fantasies and emotions of the UC by direct experience.

  To continue the example used for nearly subliminal flashes, the UC fears that he might himself be infected. The UC walks into healthy Libby's kitchen. A new window pops up, presenting how the UC picks up a butcher knife, smiles sweetly, and kills. The window then vanishes.

  Such fantasies are deeply tied to the SR—they're not something the Editor can just superimpose. Thus, the critical question to ask is not "do the other characters know this fantasy is happening?"

but rather "how can we employ the character and reality structures and interaction tools to create this fantasy without destroying the the objective reality state in the SR?" The real characters cannot be aware, but there must be twins who provide not only awareness but the complete fabric for an alternate reality.

This view suggests that these alternate realities should be modeled by alternate SR's—by spawning a sub-reality. The interaction rules and overall plot box would all depend on the UC character box in the main SR. If this is done well enough, fantasy sequences could be embedded seamlessly into the user's experience—we wouldn't have to open a window and announce "attention shoppers: this is a fantasy!"

This mechanism, where the Editor calls for an excursion into an alternative reality, would work for a wide class of alternative realities—not just UC fantasies. But what is the scope of memories of the alternative realities? How do we maintain continuity? What happens if we try to introduce multiple users?

Implementation of subjective realities is indeed difficult, but we feel it deserves much exploration. As IF is unique in its potential for interactive dream-logic episodes, we feel that exploiting this potential is essential.

- *Contrast between shot and subject matter.* The purpose of this technique is to suggest that things may not be what they seem.

  The only way we can see to implement this would be to change language style and mood. Given the primitive state of our current natural language generators, we don't think this approach is feasible.

4 Conclusions

We have presented a collection of ideas on narrative for IF so that Bromden-style worlds may be included in the standard paradigm. We did this by examining some elementary film techniques and sketching possibilities for architecture to implement these techniques.

Much work remains to be done. For one thing, we need to refine our sketches so they actually become architecture specifications. What exactly is required for the data-piece machinery we talk about? What do the character-models need to model? What kind of AI architectures can be used for the Editor? Another set of issues are the distribution questions raised by localizing interactions in the scene-graph. How are we going to work out the time-cycle hierarchy?

As mentioned, we only drew on some basic film techniques. What techniques from other areas can

we import? How can we discover techniques unique to IF? In this paper we have taken the model of a flat, realistic synthetic reality and added a dimension of (for want of a better term) affective significance to it. This model still carries an artifact of the theater paradigm—there is one single realistic SR, with maybe some dream-sequence sub-SR's built in. What if we eliminated this artifact? Imagine a hierarchy of intertwined SR's (similar to the tangled levels of reality in the film *The Singing Detective*) or an SR that is itself a spiritual landscape based solely on affective significance and not on physics or narration (as seems to be the case with the realities in the novels of William S. Burroughs).

The theory of narrative we've developed still revolves around a user-driven adventure game format. Many researchers (such as Laurel [20] and several people here at CMU) are also trying to move away from this model. The desire to break away from this format was our initial motivation; yet the ideas we've suggested only deal with the two most general of the four issues we saw this desire raise. (The techniques we've suggested may allow the user to be Bromden, but certainly not Hamlet.) So in addition to exploring the territory we've staked out in this paper, we need to press further into the frontier.

This is only the beginning.

5 Acknowledgments

The authors are grateful for the discussion and suggestions provided by the rest of the CMU Interactive Fiction group: Penny Anderson, Joel Bloch, Merrick Furst, Bryan Loyall, Maggie McCaffrey, Balas Natarajan, Marc Ringuette, and Peter Weyhrauch.

6 References

1. Agre, Philip and David Chapman. "Pengi: An Implementation of a Theory of Activity." *Proceedings AAAI-87 Sixth National Conference on AI.*

2. Agre, Philip and David Chapman. "What are Plans for?" Cambridge, Massachusetts: MIT Artificial Intelligence Laboratory AI Memo 1050, September 1988.

3. Anderson, Lindsay and David Sherwin. *If... a film.* New York: Simon and Schuster, 1969.

4. Anderson, Penny. "Exercising the OZ Character Mechanism." Pittsburgh: CMU CSD (internal), 2 November 1988.

5. Bates, Joseph. "An Analysis of *Deadline*." Pittsburgh: CMU IF Seminar Notes, Spring 1988.

6. Bates, Joseph. "The OZ Character Simulator." Pittsburgh: CMU IF Seminar Notes, Spring 1988.

7. Bates, Joseph. "Some Simple Minds in OZ." Pittsburgh: CMU IF Seminar Notes, Spring 1988.

8. Bates, Joseph. "Synthetic Realities." Pittsburgh: CMU CSD (draft), 1987.

9. Bates, Joseph. "Worlds as an Expressive Form." Pittsburgh: CMU IF Seminar Notes, Spring 1988.

10. *Deadline*. Cambridge, Massachusetts: Infocom, 1982.

11. Disch, Thomas. *Amnesia*. San Mateo, California: Electronic Arts, 1986.

12. Eisenstein, Sergei. (tr. Jay Leyda.) *Film Form* and *The Film Sense*. Cleveland: Meridian Books, 1967.

13. Gagne, Paul R. *The Zombies that Ate Pittsburgh*. New York: Dodd, Mead, and Company, 1987.

14. Hobbs, Jerry R. "Coherence and Co-Reference." *Cognitive Science*, 1 (1979), 67-82.

15. Huss, Roy and Norman Silverstein. *The Film Experience*. New York: Dell Publishing, 1968.

16. Joyce, James. *A Portrait of the Artist as a Young Man*. New York: The Viking Press, 1976.

17. Joyce, James. *Ulysses*. New York: Vintage Books, 1961.

18. Kenner, Hugh. "The *Portrait* in Perspective." in T.E. Connolly (ed.) *Joyce's Portrait: Criticism and Critiques*. New York: Appleton-Century-Crofts, 1962.

19. Kesey, Ken. *One Flew Over the Cuckoo's Nest*. New York: Signet, 1962.

20. Laurel, Brenda K. "Reassessing Interactivity." *The Journal of Computer Game Design*, 1:13, pp. 7-10.

21. Laurel, Brenda K. *Towards the Design of a Computer-Based Interactive Fantasy System*. Ann Arbor: University Microfilms, 1986.

22. McIver, Tom. "Backward Masking, and Other Backward Thoughts about Music." *The Skeptical Inquirer*, 13:1, pp. 50-63.

23. Miall, David S. "Affect and Narrative: A Model of Response to Stories." *Poetics*, 17, pp. 259-272.

24. Porter, M. Gilbert. *The Art of Grit: Ken Kesey's Fiction*. Columbia: University of Missouri Press, 1982.

25. Rheingold, Howard. *Tools for Thought: The People and Ideas behind the Next Computer Revolution*. New York: Simon and Schuster, 1985.

26. Schank, R.C. "Conceptual Dependency Theory" in R.C. Schank (ed.) *Conceptual Information Processing*. North-Holland, 1975.

27. Sussex, Elizabeth. *Lindsay Anderson*. London: Studio Vista, 1969.

28. Tanner, Stephen. *Ken Kesey*. Boston: Twayne Publishers, 1983.