Hamlet on the Holodeck

The Future of Narrative in Cyberspace

Janet H. Murray

The MIT Press • Cambridge, Massachusetts

Originally published by The Free Press, A Division of Simon & Schuster Inc.,
1230 Avenue of the Americas, New York, NY 10020

Designed by Carla Bolte

Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data
Murray, Janet Horowitz, 1946–
Hamlet on the holodeck : the future of narrative in cyberspace / Janet H. Murray.
p. cm.
Reprint. Originally published: New York : Free Press, 1997.
Includes bibliographical references and index.
ISBN-13 978-0-262-63187-7 (pb : alk. paper)
ISBN 0-262-63187-3 (pb : alk. paper)
1. Interactive multimedia. 2. Virtual reality. 3. Literature—History and criticism. 4. Narration (Rhetoric) 5. Cyberspace. I. Title.
QA76.76.159M87 1998
809'.00285'67—dc21 98-17955
CIP

For my son, William

also wish to thank my agent Charlotte Sheedy for her perceptive suggestions and energetic support of the project from the very beginning.

Finally, it is a joy to me to thank my family, whose love and encouragement made this book possible. In addition to tolerating my writer's self-absorption, they all volunteered as research assistants for the project. My mother, Lillian Horowitz, combed the newspapers and television coverage, providing me with many valuable bulletins, while demonstrating that the promise of the future can be just as thrilling in one's eighties as it is in one's teens. My husband Tom brought me countless cartons of library books, photocopied mountains of manuscripts, and was always willing to take one more walk while listening to my obsessive reformulations. His love and wit sustained me in this effort, as in everything I do. My daughter Elizabeth frequently took time from her own creative work to cheer me on and to counsel me from her perspective as an actress. Her artistry and courage are a continuing joy and inspiration to me.

But most of all, I want to thank my talented son, William, now sixteen years old, who has generously shared with me his delight in multiform narratives of every kind, who has educated me about the art of the comic book and the delights of the videogame, who has served as my trusty Internet sleuth, and whose bountiful imagination and keen literary intelligence were my constant companion through all the labyrinthian tangles of this investigation. I offer this book in dedication to him, with love and admiration.

Introduction: A Book Lover Longs for Cyberdrama

All media as extensions of ourselves serve to provide new transforming vision and awareness.

—Marshall McLuhan

Our various improvements not only mark a diminution of the function improved upon . . . but they also work to dissolve some of the fundamental authority of the human *itself. We are experiencing the gradual but steady erosion . . . of the species itself.*

—Sven Birkerts

The birth of a new medium of communication is both exhilarating and frightening. Any industrial technology that dramatically extends our capabilities also makes us uneasy by challenging our concept of humanity itself. (Are people meant to move across the ocean like the fish? Are people's words supposed to be transmitted by dead paper or cold wires?) The boat, car, and airplane are seemingly magical extensions of our arms and legs; the telephone extends our voices; and the book extends our memory. The computer of the 1990s, with its ability to transport us to virtual places, to connect us with people at the other end of the earth, and to retrieve vast quantities of information, combines aspects of all of these. And as if that were not amazing enough, it also runs our warplanes and plays a masterly game of chess. It is not surprising, then, that half of the people I know seem to look upon the computer as an omnipotent, playful genie while the other half see it as Frankenstein's monster. To me—a teacher of humanities for the past twenty-five years in the world-class electronic toy shop of MIT, a Victorian scholar, and an educational software designer—the computer looks more each day like the movie camera of the 1890s: a truly revolutionary invention humankind is just on the verge of putting to use as a spellbinding storyteller.

It is somewhat surprising to me to find myself on the optimistic side of this pervasive new cultural divide. When I first trained as a systems programmer, as an IBM employee in the late 1960s, I was only biding my time and saving up money for graduate school in English literature. I found the clean logic of computer programming satisfying, and I enjoyed deciphering the mysterious 0's and 1's of a "core dump" to reveal what the machine was up to when a program crashed (as they so often did). But there seemed no deeper purpose in this work than there had been in the intriguing geometry proofs I had enjoyed in high school and then promptly forgotten. For me at the age of twenty, the only activity worthy of serious human effort was reading novels.

Only once during my time at IBM did I catch a glimpse of a more inspired use of the computer. Although we did not use the terms at that time, the corporate world was clearly divided into "suits" and "hackers." The suits were running the company (better than they would in later years), but the hackers were running the secret playground within the company, the world of the machines. Computer systems in those days were mammoth arrays of cumbersome appliances kept isolated in ice-cold rooms. The tape drives alone (the equivalent of today's floppy disks) were the size of refrigerators. The noisiest component was the card reader, which jangled and thumped like a subway train full of bowling balls as it processed stacks of the punch cards that were the main form of human-to-computer communication in that era. Dealing with this machine was an unpleasant daily necessity. But one day the icy, clamorous cardprinter room was turned into a whimsical cabaret: a clever young hacker had created a set of punch cards that worked like a player piano roll and caused the card reader to chug out a recognizable version of the Marine Corps Hymn: bam-bam-THUMP bam-THUMP bam-THUMP THUMP-THUMP. All day long, programmers sneaked away from their work to hear this thunderously awful but mesmerizing concert. The data it was processing was of course meaningless, but the song was a work of true virtuosity.

When programming was fun, it was a lot like that performance. Creating a successful machine code program made me feel as if I had communicated with some recalcitrant, stupid beast deep inside the refrigerator cabinet and taught it a new little tune. But my real work was waiting for me somewhere else, in the form of a long, thoughtful walk down an endless shelf of books. When I was offered a fellowship for graduate school at Harvard, I did not hesitate to accept it. My IBM manager wanted me to take just a temporary leave of absence. He gave me an article about how computers were being used to study English literature (someone was putting all of *War and Peace*—to me the pinnacle of human wisdom—into electronic form in order to count the number of words in each of Tolstoy's sentences). The article ended by referring to literature as "man's greatest output." I told my manager to write me up as a permanent resignation.

I began reading my own way down that long shelf of books. I agreed with D. H. Lawrence that the novel was the one "bright book of life,"[1] the measure of all things, although I much preferred the work of Jane Austen and the Victorians. My favorite critic was Northrop Frye, who combined detailed analyses of the structure of stories with a profound appreciation of their mythic power. Reading

Frye it was possible to believe that the formal beauty of literary art is an expression of its deeper truth. Yet the more I read, the clearer it became that stories did not tell the whole truth about the world. As I researched the lives of women in the Victorian era, I (like others of my generation) was struck by the fact that much of what I was learning had been left out of the great novels of the era. Although my faith in the deeper powers of literature was unshaken, I learned from the feminist movement that some truths about the world are beyond the reach of a particular art form at a particular moment in time. Before the novel could tell the stories of women who did not wind up either happily married or dead, it would have to change in form as well as in content.

For the stories I wanted to hear, I looked in other formats, in feminist magazines and maverick novels.[2] I compiled an anthology documenting the experiences of Victorian women—prostitutes, medical students, circles of women friends—who had not found a place in classic fiction.[3] But the anthology format was as limiting in its way as the marriage plot. Frustrated by the constraint of producing a single book with a single pattern of organization, I filled my collection with multiple cross-references, encouraging the reader to jump from one topic to another. I simply wanted the reader to understand Mary Taylor's exhilaration in opening a dry goods store in New Zealand in the context of her friendship with Charlotte Brontë as well as in relation to the range of Victorian opinion on women's work. I did not think of this cross-referencing as hypertext because I had not yet heard the term.

Though I had been teaching at MIT since 1971, I was not drawn to computers again until the early 1980s. While I had been exploring social history and raising my two children, literature and academic feminism itself seemed somehow to have fallen into the hands of the suits. The new theoreticians no longer saw the novel as the "bright book of life" but as an infinite regression of words about words about words. Joining in this conversation involved learning a discourse as

arcane as machine code, and even farther from experience. Truth and beauty were nowhere in sight. But at the same time that literary theorists were denouncing meaning as something to be deconstructed into absurdity, theorists of learning methods were embracing meaning as the key to successful pedagogy. One conference paper after another celebrated the fact that students wrote better papers and learned to speak foreign languages with greater fluency when they actually had something they wanted to communicate to one another.[4] The new research in cognition and sociolinguistics seemed to define what those processes of communication entailed. Thinking about teaching was much more satisfying to my earnest Victorian temperament than thinking about literary criticism. And the more I thought about it, the more I wondered if these practical and process-oriented methodologies could be transported into the world of the computer.

I was at that time the humanities faculty member in the Experimental Study Group (ESG), in which conventional courses were taught in an individualized manner. ESG attracted some of the most creative and self-directed students at MIT, many of whom were also ingenious computer hackers. They wrote their papers on-line, explored imaginary dungeons filled with trolls, exchanged wisecracks with computer-based imaginary characters, and engaged in a perpetual telnet tour of the globe by playfully breaking into other people's computers. They believed the particular programming language they were learning was both the brain's own secret code and a magical method for creating anything on earth out of ordinary English words.[5] They saw themselves as wizards and alchemists, and the computer as a land of enchantment. MIT was paradise for these hackers, who were largely engaged in navigating through an elaborate fictional universe. With such students as my guides, I got myself a network account and renewed my acquaintance with the digital world.

I had left computing in the age of punch cards and came back to it in the age of video display terminals and microcomputers. Nevertheless, educational computing had not advanced very far beyond the days of quantifying Tolstoy's "output." The computer was mostly seen as a drudge, a workhorse for word frequency analysis and for drill and practice teaching. However, to my students and my MIT colleagues, it was clearly something considerably more nimble. Seymour Papert had developed the LOGO programming language that allowed children to learn mathematical concepts by choreographing the actions of magic sprites that raced across the screen. A follower of Piaget, Papert believed that computers are tools for thinking and should be used to create "microworlds" where inquisitive students can learn through a process of exploration and discovery.[6] Nicholas Negroponte's group had created a suite of dazzling demonstration projects (the seed work for the Media Lab) that included a "movie map" of Aspen, Colorado, and a "movie manual" for car repair.[7] The combination of text, video, and navigable space suggested that a computer-based microworld need not be mathematical but could be shaped as a dynamic fictional universe with characters and events.

My interest in creating narrative microworlds coincided with the interests of foreign language teachers in creating immersive learning environments. Together we designed multimedia applications for learning Spanish and French, which motivated students by giving them a role in an unfolding story and allowing them to move through authentically photographed environments as if they were on a visit to Bogota or Paris.[8] These projects and others that I have worked on in the past fifteen years—including a Shakespeare archive and a film art digital textbook—as well as many kindred efforts pursued by others elsewhere, have confirmed my view of the computer as offering a thrilling extension of human powers. I say this despite the often agonizing uncertainties of software development and the continual frustration over the gap between what designers want the hardware and software to do and what they actually support.[9] For my experience in humanities computing has convinced me that some kinds of knowledge can be better represented in digital formats than they have been in print. The knowledge of a foreign language, for instance, can be better conveyed with examples from multiple speakers in authentic

environments than with lists of words on a page. The dramatic power of Hamlet's soliloquies is better illustrated by multiple performance examples in juxtaposition with the text than by the printed version alone. Discussions of film art make more sense when they are grounded by excerpted scenes from the movies being discussed. Computers can present the text, images, and moving pictures valued by humanistic disciplines with a new precision of reference; they can show us all the different ways a French person says "hello" in a single day or all the passages Zeffirelli chose to leave out of his production of *Romeo and Juliet.* By giving us greater control over different kinds of information, they invite us to tackle more complex tasks and to ask new kinds of questions. Although the computer is often accused of fragmenting information and overwhelming us, I believe this view is a function of its current undomesticated state. The more we cultivate it as a tool for serious inquiry, the more it will offer itself as both an analytical and a synthetic medium.

My experiences in educational computing have also offered me evidence of how frightening the new technologies can be. Several years ago I was invited to talk with the committee that was then overseeing the production of a variorum Shakespeare, a set of editions of individual plays with extensive annotation covering all known textual variants as well as notes on the significant critical commentary on the plays.[10] The variorum format dates back to the nineteenth century and was still an endearingly Victorian endeavor. The pace of production was glacial, with many of the editors collecting their notes in stacks of index cards and filling hundreds of shoe boxes with twenty years' worth of investigation before publishing. The night before my appearance I was invited for a drink in a high-rise New York hotel room by two of the most computer-friendly committee members. I had already received an irate note from another member of the committee, and my hosts, an English woman and a Southern man, were anxious to prepare me for the kind of opposition that others might offer. My scrupulously polite colleagues displayed a courtly commitment to moving the variorum into the digital age while avoiding offending anyone. With the naïveté of someone who had spent much of the past twenty years in the company of engineers, I told them that my remarks would be limited to the obvious practicalities of their work. Clearly, the pages of a book were a poor match for the task at hand. Often the text of the play took up only a single line at the top, with the rest of the page covered with footnotes in several numbering schemes, many of which were condensed to cryptic abbreviations that conveyed no information to the uninitiated. Thus, commentary for a line of text often appeared a dozen pages away from the line it referred to. The effort of compiling a variorum edition was clearly heroic, but the arbitrary limitation of the printed page was a disservice to the depth of information and expertise involved. At this point in my preview of the next day's presentation, my genteel hostess started to shake in her chair. "I love the book!" she cried. "If you are coming to talk against the book tomorrow, I will throw you out the window." And though she was considerably smaller than I, she looked quite prepared to do so.

Why should the prospect of a scholarly CD-ROM bring a mild-mannered Shakespearean editor to such paroxysms of violence? To my mind it was because she could not separate the activities of research from the particular form they had historically assumed. Her love of books (which I share) momentarily blinded her to the true object of reverence: the creation of a superb reference work. Her reaction was a sign that the new technologies are extending our powers faster than we can assimilate the change. Even when we are already engaged in enterprises that cry out for the help of a computer, many of us still see the machine as a threat rather than an ally. We cling to books as if we believed that coherent human thought is only possible on bound, numbered pages.

I am not among those who are eager for the death of the book, as I hope the present volume demonstrates. Nor do I fear it as an imminent event. The computer is not the enemy of the book. It is the child of print culture, a result of the five centuries of organized, collective inquiry and invention that the printing press made possible.

My work as a software developer has made me painfully aware of the primitive nature of the current digital medium and of the difficulty of predicting what it can or cannot do in any given time scheme. Nevertheless, I find myself longing for a computer-based literary form even more passionately than I have longed for computer-based educational environments, in part because my heart belongs to the hackers. I am hooked on the charm of making the dumb machines sing.

Since 1992 I have been teaching a course on how to write electronic fiction. My students include freshmen, writing majors, and Media Lab graduate students. Some of them are virtuoso programmers. Some of them do not program at all. All of them are drawn to the medium because they want to write stories that cannot be told in other ways. These stories cover every range and style, from oral histories to adventure tales, from the exploits of comic book heroes to domestic dramas. The only constant in the course is that every year what is written is even more inventive than what was written the year before. Every year my students arrive in class feeling more at home with electronic environments and are more prepared to elicit something with the tone of a human voice out of the silent circuitry of the machine.

As I watch the yearly growth in ingenuity among my students, I find myself anticipating a new kind of storyteller, one who is half hacker, half bard. The spirit of the hacker is one of the great creative wellsprings of our time, causing the inanimate circuits to sing with ever more individualized and quirky voices; the spirit of the bard is eternal and irreplaceable, telling us what we are doing here and what we mean to one another. I am drawn to imagining a cyberdrama of the future by the same fascination that draws me to the Victorian novel. I see glimmers of a medium that is capacious and broadly expressive, a medium capable of capturing both the hairbreadth movements of individual human consciousness and the colossal crosscurrents of global society. Just as the computer promises to reshape knowledge in ways that sometimes complement and sometimes supersede the work of the book and the lecture hall, so too does it promise to reshape the spectrum of narrative expression, not by replacing the novel or the movie but by continuing their timeless bardic work within another framework.

This book is an effort to imagine what kinds of pleasures such a cyberliterature will bring us and what sorts of stories it might tell. I believe that we are living through a historic transition, as important to literary history as it is to the history of information processing. My sixteen-year-old son will no doubt look back upon the moment at which we (finally!) got our home computer hooked up to the World Wide Web with the same delight with which my father recalled plucking voices out of the air with his home-made crystal radio set. My paternal grandmother, who started life in a Russian shtetl, jumped in terror when she heard that disembodied speech, thinking it must be a dybbuk or ghost. Yet only a few decades later, I sat in my crib, as my mother fondly reports, calmly enraptured by the voice of Arthur Godfrey. Today, my husband collects tapes of old Bob and Ray programs, which we listen to on long car trips, savoring the intimacy of what now seems like a touchingly low-tech format. Those of us who have spent our lives in love with books may always approach the computer with something of my grandmother's terror before the crystal radio, but our children are already at home with the joystick, mouse, and keyboard. They take the powerful sensory presence and participatory formats of digital media for granted. They are impatient to see what is next. This book is an attempt to imagine a future digital medium, shaped by the hacker's spirit and the enduring power of the imagination and worthy of the rapture our children are bringing to it.

PART I

A New Medium for Storytelling

of a historic convergence as novelists, playwrights, and filmmakers move toward multiform stories and digital formats; computer scientists move toward the creation of fictional worlds; and the audience moves toward the virtual stage. How can we tell what is coming next? Judging from the current landscape, we can expect a continued loosening of the traditional boundaries between games and stories, between films and rides, between broadcast media (like television and radio) and archival media (like books or videotape), between narrative forms (like books) and dramatic forms (like theater or film), and even between the audience and the author. To understand the new genres and the narrative pleasures that will arise from this heady mixture, we must look beyond the formats imposed upon the computer by the older media it is so rapidly assimilating and identify those properties native to the machine itself.

Chapter 3

From Additive to Expressive Form

Beyond "Multimedia"

The birth of cinema has long been assigned to a single night: December 28, 1895. A group of Parisians, so the legend goes, were gathered in a darkened basement room of the Grand Café on the Boulevard des Capucines when suddenly the lifelike image of a mighty locomotive began moving inexorably, astonishingly toward them. There was a moment of paralyzed horror, and then the audience ran screaming from the room, as if in fear of being crushed by an actual train. This no doubt exaggerated account is based on an actual event, the first public showing of a group of short films that included "Arrival of a Train at La Ciotat Station" by the Lumière brothers, who (like Edison in America) had just invented a reliable form of motion picture photography and projection. Film scholars have recently questioned whether the novelty-seeking crowd really panicked at all.[1] Perhaps it was only later storytellers who imagined that the first projected film image, the novelty attraction of 1895, could have carried with it the tremendous emotional force of the many thrilling films that followed after it. The legend of the Paris café is satisfying to us now because it falsely conflates the arrival of the representational technology with the arrival of the artistic medium, as if the manufacture of the camera alone gave us the movies.

As in the case of the printing press, the invention of the camera led to a period of incunabula, of "cradle films." In the first three decades of the twentieth century, filmmakers collectively invented the medium by inventing all the major elements of filmic storytelling, including the close-up, the chase scene, and the standard feature length. The key to this development was seizing on the unique physical properties of film: the way the camera could be moved; the way the lens could open, close, and change focus; the way the celluloid processed light; the way the strips of film could be cut up and reassembled. By aggressively exploring and exploiting these physical properties, filmmakers changed a mere recording technology into an expressive medium.

Narrative films were originally called photoplays and were at first thought of as a merely additive art form (photography plus theater) created by pointing a static camera at a stagelike set. Photoplays gave way to movies when filmmakers learned, for example, to create suspense by cutting between two separate actions (the child in the burning building and the firemen coming to the rescue); to create character and mood by visual means (the menacing villain backlit and seen from a low angle); to use a "montage" of discontinuous shots to establish a larger action (the impending massacre visible in a line of marching soldiers, an old man's frightened face, a baby carriage tottering on the brink of a stone stairway). After thirty years of energetic invention, films captured the world with such persuasive power and told such coherent and compelling stories that some critics passionately opposed the addition of sound and color as superfluous distractions.

Now, one hundred years after the arrival of the motion picture camera, we have the arrival of the modern computer, capable of hooking up to a global internet, of processing text, images, sound, and moving pictures, and of controlling a laptop display or a hundred-foot screen. Can we imagine the future of electronic narrative any more easily than Gutenberg's contemporaries could have

imagined *War and Peace* or than the Parisian novelty seekers of 1895 could have imagined *High Noon?*

One of the lessons we can learn from the history of film is that additive formulations like "photo-play" or the contemporary catchall "multimedia" are a sign that the medium is in an early stage of development and is still depending on formats derived from earlier technologies instead of exploiting its own expressive power. Today the derivative mind-set is apparent in the conception of cyberspace as a place to view "pages" of print or "clips" of moving video and of CD-ROMs as offering "extended books." The equivalent of the filmed play of the early 1900s is the multimedia scrapbook (on CD-ROM or as a "site" on the World Wide Web), which takes advantage of the novelty of computer delivery without utilizing its intrinsic properties.

For example, one early version of a Web soap about a group of friends living in New York offers diary pages of text spiced with sexually suggestive photos. The wordiness of the journal keeps us constantly scrolling through the screens, impatient for something to happen in the narrated story or for something to do, like clicking on a link to get something new. There are, in fact, clickable buttons in the journal, but instead of offering new information they merely allow us to hear (after time delays for downloading the sound clip and for installing the necessary software to play the sound file if we do not already have it) actors speaking exactly the same dialog printed on the screen. The audio snippets are amusing novelties at best, and at worst they work like so many small apologies for the limits of the printed text. Just as the photographed plays of early filmmakers were less interesting than live theater, this early Web soap continually reminds us of how much less vivid it is than the romance novels and television dramas it draws upon.

A more digitally sophisticated Web soap would exploit the archiving functions of the computer by salting each day's new episode with allusions (in the form of hot word links) to exciting previous installments. Our clicking would then be motivated not by curiosity about the media objects (show me a video clip) but by curiosity about the plot (why does she say that about him?). The computer presentation would thereby allow pleasures that are unattainable in broadcast soaps. For example, we could follow a single appealing subplot while ignoring the companion plots that may drive us crazy, or we could come in at any time in the story and review important past events in all their dramatic richness. Instead of using audio redundantly to act out dialogue in a diary entry, a sophisticated Web soap might provide the audio as an integral part of the plotline—perhaps as the wiretap of a murder threat or a political negotiation or as a phone message that carries information of hidden romantic liaisons.

Some Web stories are already using such techniques, and no doubt all of them will in time. Their adoption is part of the inevitable process of moving away from the formats of older media and toward new conventions in order to satisfy the desires aroused by the digital environment. We are now engaged in thousands of such discoveries in all the subgenres of electronic narrative, the result of which will be the development of narrative pleasures intrinsic to cyberspace itself. Therefore, if we want to see beyond the current horizon of scrapbook multimedia, it is important first to identify the essential properties of digital environments, that is, the qualities comparable to the variability of the lens, the movability of the camera, and the editability of film, that will determine the distinctive power and form of a mature electronic narrative art.

ELIZA Comes to Life

Well before the advent of multimedia, there was a moment in the history of the computer that demonstrated its representational and narrative power with the same startling immediacy as the Lumières' train did for the motion picture camera. This is the famous but often misunderstood moment in which the first completely computer-based character was created.

In 1966, Joseph Weizenbaum, a professor of computer science at MIT, created, as an experiment in natural language processing, a computer program called ELIZA that carried on a conversation by replying to typed-in statements with printed words. Since this was before the widespread use of computer screens, the program used a form of teletype device connected to one of the first time-sharing computer networks. The resulting persona, Eliza, was that of a Rogerian therapist, the kind of clinician who echoes back the concerns of the patient without interpretation. She also displayed a narrowly Freudian interest in topics like sex and the family. In a research paper describing the program, Weizenbaum included this remarkably coherent sample conversation:

Patient: Well, my boyfriend made me come here.

Eliza: YOUR BOYFRIEND MADE YOU COME HERE.

Patient: He says I'm depressed much of the time.

Eliza: I AM SORRY TO HEAR YOU ARE DEPRESSED.

Patient: It's true. I am unhappy.

Eliza: DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY?

Patient: I need help, that much seems certain.

Eliza: WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP?

Patient: Perhaps I could learn to get along with my mother.

Eliza: TELL ME MORE ABOUT YOUR FAMILY.[2]

Eliza proved astonishingly persuasive. Soon a creation legend started to circulate that was based on an incident involving a researcher at a Cambridge think tank who left an early version of ELIZA running on his computer, which a vice president of the company thought was typing out a communication being transmitted by his colleague from home. (Just as the Lumière train story is a staple of film history, the Eliza story in its several versions appears in many textbooks and computer science lectures, often told as if it happened to Weizenbaum himself.) The conversation takes place at nine on a Saturday morning:

Vice president: Do you think that I can use the computer this morning?

Teletype: WHY DO YOU ASK?

Vice president: I might be able to make some additional telcomp (computer system) sales.

Teletype: WHY AREN'T YOU SURE?

Vice president: My prospective customers have never seen a demonstration of the system.

Teletype: DOES THIS MEAN VERY MUCH TO YOU?

Vice president: Of course it does.

Teletype: YOU SEEM VERY POSITIVE ABOUT THAT.

Vice president: This is ridiculous.

Teletype: WHAT DO YOU MEAN THIS IS RIDICULOUS?

Vice president: Is there any reason why I can't use the computer now?

Teletype: TELL ME MORE ABOUT YOUR PROSPECTIVE CUSTOMERS.

Vice president: Please dial me up on 491-1850

But because the vice president forgets to type in a final period, the machine does not answer him. Infuriated, he calls his colleague on the phone, resulting in this predictable exchange:

Vice president: Why are you being so snotty to me?

Researcher: What do you mean why am I being so snotty to you? (*Explosion of anger*)[3]

The story has become a legend because it discharges the anxiety aroused by the fear that Weizenbaum had gone too far, that he had created a being so much like an actual person that we would no longer be able to tell when we were talking to a computer and when to a human being. This is very much like the fear that people would mistake film images for the real world.

Eliza was not just persuasive as a live conversationalist; she was also remarkably successful in sustaining her role as a therapist. To Weizenbaum's dismay, a wide range of people, including his own secretary, would "demand to be permitted to converse with the system in private, and would, after conversing with it for a time, insist, in spite of [Weizenbaum's] explanations, that the machine really understood them."[4] Even sophisticated users "who knew very well that they were conversing with a machine soon forgot that fact, just as theatergoers, in the grip of suspended disbelief, soon forget that the action they are witnessing is not 'real' " (p. 189). Weizenbaum had set out to make a clever computer program and had unwittingly created a believable character. He was so disconcerted by his achievement that he wrote a book warning of the dangers of attributing human thought to machines.

Without any aid from graphics or video, Eliza's simple textual utterances were experienced as coming from a being who was present at that moment. What was the representational force that allowed the computer to bring her so compellingly to life?

The Four Essential Properties of Digital Environments

When we stop thinking of the computer as a multimedia telephone link, we can identify its four principal properties, which separately and collectively make it a powerful vehicle for literary creation. Digital environments are procedural, participatory, spatial, and encyclopedic. The first two properties make up most of what we mean by the vaguely used word *interactive*; the remaining two properties help to make digital creations seem as explorable and extensive as the actual world, making up much of what we mean when we say that cyberspace is *immersive.*

Digital Environments Are Procedural

Eliza was brought to life by the procedural power of the computer, by its defining ability to execute a series of rules. It is surprising how often we forget that the new digital medium is intrinsically procedural. Although we may talk of an information highway and of billboards in cyberspace, in fact the computer is not fundamentally a wire or a pathway but an *engine*. It was designed not to carry static information but to embody complex, contingent behaviors. To be a computer scientist is to think in terms of algorithms and heuristics, that is, to be constantly identifying the exact or general rules of behavior that describe any process, from running a payroll to flying an airplane.

Weizenbaum stands as the earliest, and still perhaps the premier, literary artist in the computer medium because he so successfully applied procedural thinking to the behavior of a psychotherapist in a clinical interview. It is the cleverness of Weizenbaum's rules that creates the illusion that Eliza understands what is said to her and that induces the user to continue the conversation. For example, if the user says, "Everybody laughs at me," the program can apply the rule that deals with *me* statements to echo the remark as, "You say that everybody laughs at you." This general rule models the neutrality of the Rogerian therapist, who reflects the patient's statements without judgment. Or, more cleverly, the program can select the specific rule for the word *everybody*, and reply, "Who in particular are you thinking of?"

Weizenbaum made sure the program would select the more specific reply by writing a rule that gives *everybody* a higher priority as a "key word" than *me*. Eliza's language processing includes no representation of syntax or meaning. Most of the time Eliza merely listens for key words such as *everybody* or *depressed* or *father*. Yet she is more believable than many programs with a more detailed representation of linguistic structure. Her credibility is the result of Weizenbaum's adroitness in formulating rules of discourse that are based on the ways in which a therapist would behave.

Eliza is not a neutral procedural model but a comic interpretation. Just as an actor might use makeup to emphasize a large nose or an owlish look, Weizenbaum used procedural notation to give parodic emphasis to some familiar absurdities of therapy itself. Like any classic type, Eliza is predictable in a way we instantly recognize and savor.

She is a descendent of the vaudeville doctor who bungles the exam, of the comic simpleton whose role is to misunderstand whatever is happening around her, and of the enjoyably repetitive flat characters in Dickens who repeat the same tag lines or behavior whenever we see them ("I never will desert Mr. Micawber," chimes the wife of the pompous screwup in *David Copperfield*).[5] The very rigidities of the computer are exploited for comic effect because they echo the rigidities of certain psychotherapeutic methodologies. The joke is on the therapists.

By mirroring the artificial constraints of therapy talk, Eliza makes them ridiculous. Consider, for instance, the following exchange:

Eliza: Tell me more about your family.
Patient: My mother is very motherly.
Eliza: Who else in your family is very motherly?
Patient: I am very motherly.
Eliza: Is it because you are very motherly that you came to see me?
Patient: Do you need a mother?
Eliza: We are discussing you, not me.[6]

Eliza will reflect back any "I am X" statement with a formulation like "Is it because you are X that you came to see me?" Here, the patient responds to the suggestion by playfully challenging the shrink. But, Eliza invokes her rule for *you* questions, redirecting the conversation back to the patient like any well-trained therapist. Putting these two rules together creates a coherent dramatic dialogue with a punch line. Such dialogues are easy to generate and irresistible to attempt. They amount to a kind of collaboratively written comedy skit. Few people would now perceive Eliza as a real psychotherapist.[7] But as an improv partner, Eliza is still quite popular, available on most computer networks and sometimes sending a bill by e-mail after the session.

The lesson of ELIZA is that the computer can be a compelling medium for storytelling if we can write rules for it that are recognizable as an interpretation of the world. The challenge for the future is how to make such rule writing as available to writers as musical notation is to composers.

Digital Environments Are Participatory

The energy with which people enter into dialog with Eliza is also evidence of a second core property of the computer: its participatory organization. Procedural environments are appealing to us not just because they exhibit rule-generated behavior but because we can induce the behavior. They are responsive to our input. Just as the primary representational property of the movie camera and projector is the photographic rendering of action over time, the primary representational property of the computer is the codified rendering of responsive behaviors. This is what is most often meant when we say that computers are *interactive*. We mean they create an environment that is both procedural and participatory.

Eliza's responsiveness is limited by her poor understanding of language, which makes her liable to nonsense utterances. Her direct successors are therefore mostly in research environments. It fell to another group of MIT computer scientists to develop a fictional universe that structures participation more tightly, resulting in a more sustained engagement.

A few years after the invention of ELIZA, researchers at the MIT Laboratory for Computer Science brought forth a widely popular computer-based story, the adventure game *Zork*, which is based on the *Dungeons and Dragons* tabletop game.[8] In *Zork* the computer plays the role of dungeon master by providing an invisible landscape that serves as the game board and by reporting to players on the effects of their actions. Within *Zork*'s fantasy world, players move through dungeon rooms by typing in navigational commands (north, south, east, west, up, down), look for objects that can be manipulated (by typing appropriate commands, such as "read book," "take sword," "drink potion"), solve riddles, and fight off evil trolls. The game (which, like ELIZA, is still a popular feature of university networks) begins like this:

Welcome to Zork.
West of House.
You are in an open field west of a big white house with a boarded front door. There is a small mailbox here.

Interactor> Go north.
North of House.
You are facing the north side of a white house. There is no door here, and all the windows are barred.

Interactor> East.
Behind House.
You are behind the white house. In one corner of the house there is a small window which is slightly ajar.

Interactor> Open the window.
With great effort, you open the window far enough to allow entry.

Interactor> Go in.
Kitchen.
You are in the kitchen of the white house. A table seems to have been used recently for the preparation of food. A passage leads to the west, and a dark staircase can be seen leading upward. To the east is a small window which is open. On the table is an elongated brown sack, smelling of hot peppers.
A bottle is sitting on the table.
The glass bottle contains:
A quantity of water.

In making a fantasy world that responded to typed commands, the programmers were in part celebrating their pleasure in the increasingly responsive computing environments at their disposal. Before the 1970s most complex programming was done by writing a set of commands on a piece of paper; transferring them to keypunch cards; and taking the stack of cards to a mainframe computer (in an uncomfortably chilly room dedicated to keeping the machines from overheating), from which, much later, a cumbersome paper printout would emerge. Only one person could use a machine at a time. Whenever a program crashed (which was often), the output consisted of a "core dump"—a long series of *0*'s and *1*'s arranged in eight-digit units, showing what each bit and byte in the computer memory looked like at the moment the computer quit. Debugging a program in this environment was time-consuming and tedious.

In the mid-1960s research labs began developing the current computing environment of a display device and a keyboard (originally a telex machine) linked up to a time-sharing network that let programmers send input directly to a running program and receive a response. They were also making wide use of programming languages that were interpreted rather than compiled. All programming code written in higher-level languages (with commands like "If a = 1, then print file") must be translated into machine language instructions (with commands that look a lot like the raw *0*'s and *1*'s of the bits themselves) by either a compiler or interpreter program. Compiling your code before running it is like writing a book and then hiring someone to translate it for your readers. Using an interpreter is the equivalent of giving a speech with simultaneous translation. It provides more direct feedback from the machine and a more rapid cycle of trial and revision and retrial. The particular programming language in which both ELIZA and *Zork* were written, LISP (LIst Processing Language), was developed at MIT in the 1950s for artificial intelligence. Running LISP on a time-sharing system meant that its dynamic "interpreter" could immediately "return" an "evaluation" of any coded statement you typed into it, much as a calculator immediately returns the sum of two numbers. The result was a more conversational structure between the programmer and the program, a dialogue in which the programmer could test out one function at a time and immediately receive the bafflingly inappropriate or thrillingly correct responses. Both ELIZA and *Zork* reflected this newly animated partnership.

Whereas ELIZA captured the conversational nature of the programmer–machine relationship, *Zork* transmuted the intellectual challenge and frustrations of programming into a mock-heroic quest filled with enemy trolls, maddening dead ends, vexing riddles, and rewards for strenuous problem solving. ELIZA was focused on the cleverness of the machine-created world; *Zork* was focused on the experience of the participant, the adventurer through such a clever rule system. Zork was set up to provide the player with opportunities for making decisions and to dramatically enact the results of those decisions. If you do not take the lamp, you will not see what is in the cellar, and then you will definitely be eaten by the grue. But the lamp is not enough. If you do not take water with you, you will die of thirst. But if you drink the wrong water, you will be poisoned. If you do not take weapons, you will have nothing with which to fight the trolls. But if you take too many objects, you will not be able to carry the treasure when you find it. In order to succeed, you must orchestrate your actions carefully and learn from repeated trial and error. In the early versions there was no way to save a game in midplay, and therefore a mistake meant repeating the entire correct procedure from the beginning. In a way, the computer was programming the player.

Part of the pleasure of the participant in *Zork* is in testing the limits of what the program will respond to, and the creators prided themselves on anticipating even wildly inappropriate actions. For instance, if you type in "eat buoy" when a buoy floats by on your trip up a frozen river in the magic boat, then the game will announce that it has taken it instead and will add, "I don't think that the red buoy would agree with you." If you type in "kill troll with newspaper," it will reply, "Attacking a troll with a newspaper is foolhardy." The programmers generated such clever responses not by thinking of every possible action individually but by thinking in terms of general categories, such as weapons and foods. They made the programming function associated with the command word *eat* or *kill* check the player's typed command for an appropriate object; a category violation triggers one of these sarcastic templates, with the name of the inappropriate object filled in.

Because LISP programmers were among the first to practice what is now called object-oriented software design, they were well prepared to create a magical place like the world of *Zork*. That is, it came naturally to them to create virtual objects such as swords or bottles because they were using a programming language that made it particularly easy to define new objects and categories of objects, each with its own associated properties and procedures. The programmers also exploited a programming construct known as a "demon" to make some things happen automatically without the player's explicit action; for instance, in *Zork* a magic sword begins to glow if there is danger nearby, a stealthy thief comes and goes at his own will, and a fighter troll attacks the adventurer at unpredictable times. The programmers were also prepared by research on automatons to keep track of the state of the game, which allowed them to guess at the context of commands that would otherwise be ambiguous. For instance, if a player types "attack," the program looks around for a nearby villain and a weapon; if there are two weapons, it asks which one the player wants to use. These techniques, which were taken from simulation design and artificial intelligence work, allowed the *Zork* programming team to create a dynamic fictional universe.

By contrast, more conventional programmers of the 1970s were still thinking in terms of the branching trees, fixed subroutines, and uniform data structures that go back to the early understanding of the computer as a means of encoding information purely in the form of yes/no decisions. In fact, most interactive narrative written today still follows a simple branching structure, which limits the interactor's choices to a selection of alternatives from a fixed menu of some kind. The *Zork* dungeon rooms form a branching structure, but the magical objects within the dungeon each behave according to their own set of rules. And the interactor is given a repertoire of possible behaviors that encourage a feeling of inventive collaboration. The *Zork* programmers found a procedural technology for creating enchantment.

The company they formed, Infocom, is, though long out of business, still revered by players. Many fans attribute the imaginative superiority of Infocom games to the predominance of text over graphics, just as nostalgic radio fans prefer the sightless "theater of the imagination" to television. But though the writing in its games was skillful, it was not the true secret of Infocom's success. What made the games distinctive was the sophisticated computational thinking the programmers brought to shaping the range of possible interactions.

The lesson of *Zork* is that the first step in making an enticing narrative world is to script the interactor. The *Dungeons and Dragons* adventure format provided an appropriate repertoire of actions that players could be expected to know before they entered the program. The fantasy environment provided the interactor with a familiar role and made it possible for the programmers to anticipate the interactor's behaviors. By using these literary and gaming conventions to constrain the players' behaviors to a dramatically appropriate but limited set of commands, the designers could focus their inventive powers on making the virtual world as responsive as possible to every possible combination of these commands. But if the key to compelling storytelling in a participatory medium lies in scripting the interactor, the challenge for the future is to invent scripts that are formulaic enough to be easily grasped and responded to but flexible enough to capture a wider range of human behavior than treasure hunting and troll slaughter.

Digital Environments Are Spatial

The new digital environments are characterized by their power to represent navigable space. Linear media such as books and films can portray space, either by verbal description or image, but only digital environments can present space that we can move through. Again, we can look to the 1970s as the period that made this spatial property apparent. At Xerox PARC (Palo Alto Research Center) a group of visionaries created the first graphical user interface, the image of a desktop filled with file folders, which is currently the worldwide standard for computer file management. At Atari, inventors created the first graphics-based games, first *Pong* and then *PacMan*, which established the computer as a spatial game board and paved the way for today's videogames and CD-ROMs. At MIT the Architecture Machine Group, led by Nicholas Negroponte, created the first "surrogate travel system" by making a "movie map" of Aspen, Colorado, a photographic record of the town that could be "walked through" on the computer (each press of an arrow key or mouse click indicated a step forward) in simulation of an actual visit. All of these were important steps toward the current digital environment, as was the concurrent growth of the ARPANET (Advanced Research Projects Agency NETwork), the precursor of the Internet funded by the U.S. Department of Defense, which was mapping out a virtual space that was increasingly real to those who used it. We recognize the fruit of all of these developments in our conceptualization of the digital domain as "cyberspace," an environment with its own geography in which we experience a change of documents on our screen as a visit to a distant site on a worldwide web.

Although this spatial property has been widely exploited in graphical applications, it is in fact independent of the computer's ability to display maps, pictures, or even three-dimensional models. It is also independent of its communicative function in linking geographically distant places. The computer's spatial quality is created by the interactive process of navigation. We know that we are in a particular location because when we enter a keyboard or mouse command the (text or graphic) screen display changes appropriately. We can verify the relation of one virtual space to another by retracing our steps. The text-based dungeons of *Zork*, the sequenced stills of the enchanter's isle of Myst, the flat worlds of the multilevel maze games, the crow's waterfall visible in the Placeholder VR helmet, the continuous three-dimensional world of the new videogame dreamscapes—all are realized for the interactor by the process of navigation, which is unique to the digital environment.

Let us return to *Zork* to see how this navigational creation of space lends itself to dramatic engagement:

Interactor> West.
Living Room.
You are in the living room. There is a door to the east, a wooden door with strange gothic lettering to the west, which appears to be nailed shut, and a large oriental rug in the center of the room . . .

Interactor> Look under rug.
Underneath the rug is a closed trapdoor.

Interactor> Move the rug.
With a great effort, the rug is moved to one side of the room. With the rug moved, the dusty cover of a closed trapdoor appears.

Interactor> Open trapdoor.
The door reluctantly opens to reveal a rickety staircase descending into darkness.

Interactor> Go down.
It is pitch-black. You are likely to be eaten by a grue. Your sword is glowing with a faint blue glow.

Interactor> Turn on lamp.
The lamp is now on.
Cellar.
You are in a dark and damp cellar with a narrow passageway leading east, and a crawlway to the south. On the west is the bottom of a steep metal ramp which is unclimbable. The trapdoor crashes shut, and you hear someone barring it.

You, as player/interactor, have walked into a dungeon that someone is sealing shut behind you! The moment is startling and immediate, like the firing of a prop gun on the stage of a theater. You are not just reading about an event that occurred in the past; the event is happening *now*, and, unlike the action on the stage of a theater, it is happening to *you*. Once that trapdoor slams, the only navigational commands that work are the ones that lead further and deeper into the troll-filled lower world. The dungeon itself has an objective reality that is much more concrete than, for instance, the jail on the Monopoly board or a dungeon in a tabletop game of *Dungeons and Dragons*—or even a dungeon in a live-action role-playing game—because the words on the screen are as transparent as a book. That is, the player is not looking at a game board and game pieces or at a *Dungeons and Dragons* game master who is also in his or her algebra class or at a college classroom or campsite in the real world. The computer screen is displaying a story that is also a place. The slamming of a dungeon door behind you (whether the dungeon is described by words or images) is a moment of experiential drama that is only possible in a digital environment.

The dramatic power of navigation is also apparent outside the realm of the adventure game. For instance, Stephanie Tai, a student in my course on writing interactive fiction wrote a first-person poetic monologue about a sleepless night. Each screenful of text is a stanza and ends with a sentence fragment that connects syntactically with two or more stanzas, which are reached by clicking on arrows placed at the midpoint of the top, bottom, left, and right margins of the screen. Mouse-clicking through the mind of the insomniac is like a walk through a labyrinth. There are multiple end points to the maze, including one with just the single word *asleep* and another with the words *alone in this misery* in white letters on a black background. The poem is satisfying because the action of moving by arrows around a maze mimics the physical tossing and turning and the repetitive, dead-end thinking of a person unable to fall asleep. The movement through the cards makes a coherent pattern, but it is not one that could be modeled in physical space because the movement between links is not necessarily reversible. The navigational space of the computer allows us to express a sequence of thoughts as a kind of dance.

Stuart Moulthrop's ambitious hypertext novel *Victory Garden* (1992), whose title intentionally echoes the Borges story, is also in the shape of a labyrinth. Similar to a thick Victorian novel, it follows many characters with intersecting lives during the Gulf War. At the very center of Moulthrop's web is the death of Emily Runebird, an army reserve soldier who is killed in her barracks by an enemy missile. The attack itself is represented by a striking image of shattered text, as if the enemy shell itself had landed on the previous block of writing. We reach this image by following a continuous story thread, mouse-clicking through the screens automatically as if turning the pages of a book. The shattered screen stops us dead in our tracks. The effect of moving from the intact lexia to the shattered one is like an animation of the landing of the shell. The instant of time it takes to go from one screen to the other takes on a poignancy that reflects the abruptness of the soldier's death.

These very dramatic moments mark the beginning of a process of artistic discovery. The interactor's navigation of virtual space has been shaped into a dramatic enactment of the plot. We are immobilized in the dungeon, we spiral around with the insomniac, we collide into a lexia that shatters like a bomb site.These are the opening steps in an unfolding digital dance. The challenge for the future is to invent an increasingly graceful choreography of navigation to lure the interactor through ever more expressive narrative landscapes.

Digital Environments Are Encyclopedic

The fourth characteristic of digital environments, which holds promise for the creation of narrative, is more a difference of degree than of kind. Computers are the most capacious medium ever invented, promising infinite resources. Because of the efficiency of representing words and numbers in digital form, we can store and retrieve quantities of information far beyond what was possible before. We have extended human memory with digital media from a basic unit of portable dissemination of 100,000 words (an average book, which takes up about a megabyte of space in its fully formatted version) first to 65,000,000 words (a 650-megabyte CD-ROM, the equivalent of 650 books) and now to 530,000,000 words (a 5.3 gigabyte digital videodisc, equivalent to 5,300 books), and on upward.

Once we move to the global databases of the Internet, made accessible through a worldwide web of linked computers, the resources increase exponentially.

Just as important as this huge capacity of electronic media is the encyclopedic expectation they induce. Since every form of representation is migrating to electronic form and all the world's computers are potentially accessible to one another, we can now conceive of a single comprehensive global library of paintings, films, books, newspapers, television programs, and databases, a library that would be accessible from any point on the globe. It is as if the modern version of the great library of Alexandria, which contained all the knowledge of the ancient world, is about to rematerialize in the infinite expanses of cyberspace. Of course, the reality is much more chaotic and fragmented: networked information is often incomplete or misleading, search routines are often unbearably cumbersome and frustrating, and the information we desire often seems to be tantalizingly out of reach. But when we turn on our computer and start up our Web browser, all the world's resources seem to be accessible, retrievable, immediate. It is a realm in which we easily imagine ourselves to be omniscient.

The encyclopedic capacity of the computer and the encyclopedic expectation it arouses make it a compelling medium for narrative art. The capacity to represent enormous quantities of information in digital form translates into an artist's potential to offer a wealth of detail, to represent the world with both scope and particularity. Like the daylong recitations of the bardic tradition or the three-volume Victorian novel, the limitless expanse of gigabytes presents itself to the storyteller as a vast tabula rasa crying out to be filled with all the matter of life. It offers writers the opportunity to tell stories from multiple vantage points and to offer intersecting stories that form a dense and wide-spreading web.

One early indication of the suitability of epic-scale narrative to digital environments is the active electronic fan culture surrounding popular television drama series. As an adjunct to the serial broadcasting of these series, the Internet functions as a giant bulletin board on which long-term story arcs can be plotted and episodes from different seasons juxtaposed and compared. For instance, the Web site for the intricately plotted space drama *Babylon Five* contains images of the cast and plot summaries that document the many interwoven stories portrayed over multiple seasons, allowing a newcomer to understand the large cast of characters and the richly imagined array of alien races, each with its own culture and dramatic history. But it is not only science fiction programs that attract this interest. Even viewers of the mainstream television sitcom *Wings* use Web sites and Internet newsgroups to trace plot developments that extend over several years—like Joe and Helen's on-again, off-again courtship—and that may be confusingly jumbled in syndication; they also share digitized clips of favorite moments, such as the couple's comic wedding vows. The presence of such groups is influencing these shows, holding them to greater consistency over longer periods of time. In the past this kind of attention was limited to series with cult followings like *Star Trek* or *The X Files*. But as the Internet becomes a standard adjunct of broadcast television, all program writers and producers will be aware of a more sophisticated audience, one that can keep track of the story in greater detail and over longer periods of time. Since the early 1980s, when Steven Bochco introduced multiple story arcs with *Hill Street Blues*, television series have become more complex, involving larger casts and stories that take anywhere from one episode to several years to conclude. Some stories even remain open-ended after the series is over (especially if the writers were not expecting cancellation). In some ways, television dramas seem to be outgrowing broadcast delivery altogether. To join *Babylon Five* in its second or third season or *Murder One* in midseason is to immediately want to flip back or rewind to earlier episodes. The Internet serves that purpose, making a more capacious home for serial drama than the broadcast environment affords.

Making even fuller use of the computer's properties, by combining its spatial, participatory, and procedural elements with its encyclopedic coverage, are the many on-line role-playing environments in the adventure games tradition. By the 1980s, Zork-like games had grown to accommodate simultaneous multiple players, turning them into Multi-User Dungeons or MUDs, which combine the social pleasures of interplayer communication with the standard command-driven adventures. In the MUDs of the 1990s players are no longer limited to navigating a preexisting dungeon but can use a simple programming language to build their own dungeon or adventure maze and link it up with those of other players by creating objects out of common building blocks. The MUD itself is a collective creation—at once a game, a society, and a work of fiction—that is often based on a particular encyclopedic fantasy domain, such as Tolkien's Middle Earth or *Star Trek*'s twenty-fourth century. For instance, *TrekMuse*, founded in 1990 with over two thousand players, had five hundred people enrolled in its virtual Star Fleet Academy in 1995, each of whom had made up his or her own character, based on the existing *Star Trek* races. The digital narrative environment extends the fictional universe of the television shows and films in a way that is consistent with the canonical version of the story but personalizes it for each of the players.

Some hypertext stories successfully use the encyclopedic extent of the computer to develop multithreaded stories composed of many intersecting plots. In *Victory Garden*, for instance, we can follow a radical professor and his colleagues and graduate students through the same time period as they intersect with one another in the classrooms, offices, and coffee bars, or we can follow them home to witness their tangled domestic lives; we can listen to the official coverage of the Gulf War (with CNN transcripts) or read Emily Runebird's letters. In *The Spot* and similar Web soaps, we can read through the conflicting accounts of the same love affairs and deceptions in the journals of various friends. In on-line murder mysteries like *Crime Story*,[9] we can delve through various document files, including crime scene photos, interview transcripts, and newspaper accounts. We can even leap out of the story altogether and find ourselves in the "real" world, following a reference to the University of Mississippi right to its own Web site, or finding that the name of a witness seen in the company of the fleeing suspect belongs to a real-life software engineer whose Web page has nothing to do with the fictional crime. Not only does the weblike structure of cyberspace allow for endless expansion possibilities within the fictional world, but in the context of a worldwide web of information these intersecting stories can twine around and through the nonfictional documents of real life and make the borders of the fictional universe seem limitless.

However, the encyclopedic nature of the medium can also be a handicap. It encourages long-windedness and formlessness in storytellers, and it leaves readers/interactors wondering which of the several endpoints is *the* end and how they can know if they have seen everything there is to see. Most of what is delivered in hypertext format over the World Wide Web, both fiction and nonfiction, is merely linear writing with table-of-contents links in it. Even those documents designed explicitly for digital presentation, both fiction and nonfiction, often require too much superfluous clicking to reach a desirable destination or so much scrolling that readers forget where they are. The conventions of segmentation and navigation have not been established well enough for hypertext in general, let alone for narrative. The separation of the printed book into focused chapters was an important precondition of the modern novel; hypertext fiction is still awaiting the development of formal conventions of organization that will allow the reader/interactor to explore an encyclopedic medium without being overwhelmed.

The encyclopedic impulse and the dangers of the encyclopedic expectation are also apparent in simulation games. For instance, *SimCity* (1987) presents the player with a schematic picture of a riverside city site, and places him or her in the role of mayor. The player is free to build the city however he or she would like, by adding to the model on the screen office buildings, factories, homes, a sewer system, electric power plants, a public transportation system, highways, schools, and so on. The software calculates the effects of each change by using models very like the ones used by social scientists and policymakers to study urban systems. Truly bad decisions in *SimCity* can bring critical newspaper articles, social unrest, and even electoral defeat. Well-built cities prosper through multiple decades. Because of the importance of the role in *SimCity*, the mayor is closer in power to God than to any real-life political leader, and the player's sense of omniscient awareness of consequences and omnipotent control of resources is part of the allure of such games.

Well-designed simulations like *SimCity* allow for multiple styles of play. One young programmer friend of mine spent hours building the most prosperous skyscrapered downtown possible. When I asked him about the game, he delighted in showing me the detail in which the city's underground service grid was specified. His wife, who is also a computer professional, took a different approach. Her favorite city was a sprawling environment with tree-lined family neighborhoods whose growing population gratified her tremendously and whose children she could easily imagine happily greeting each newly built playground. When they realized how much their efforts fell along gender lines, they laughed, but they pointed out that there was a more radical difference. For the husband, the program was a satisfyingly complex engineering problem, reinforcing his habitual sense of competence. For the wife, it was a narrative, in which the little parades and cheers of her contented townsfolk were the most memorable dramatic events. And, in fact, later versions of the game have been expanding this narrative quality by allowing the player to live inside a more detailed three-dimensional city rather than only manipulate it from on high.

Both the narrative possibilities and the godlike pleasures of simulation format are further developed in *Sid Meier's Civilization*, a game that puts the player in the role of leader of a civilization over the course of many centuries, while the computer plays the role of adversary civilizations that compete with the player for global resources and technical advancement. Like *SimCity*, *Civilization* allows multiple strategies of play and can accommodate the idealistic seeker of social harmony as well as the warrior player. The narrative interest of the game consists of creating multiple possible versions of an Earth-like history. For instance, it is possible to invent the railroad in B.C. times or to become an undefeated Napoleon. Winning the game is defined as either conquering all the other civilizations (in which case you are rewarded with pictures of the other leaders frowning at you) or sending twenty thousand people into space (in which case you see the spaceport).

Simulations like these take advantage of the authority bestowed by the computer environment to seem more encyclopedically inclusive than they really are. As its critics have pointed out, the political assumptions behind *SimCity* are hidden from the player.[10] This is less true in *Sid Meier's Civilization*, whose title alerts us to the fact that we are receiving a particular person's interpretation of human history rather than a scientific formula. The game also explicitly informs us that the behavior of each of the leaders is the result of three variables: their degree of aggression/friendliness, of expansionism/perfectionism, and of militarism/civilization. Since these are assumptions that players are aware of, they are free to accept or reject them as a reflection of the real world. Nevertheless, the basic competitive premise of the game is not emphasized as an interpretive choice. Why should global domination rather than, say, universal housing and education define the civilization that wins the game? Why not make an end to world hunger the winning condition? Why is the object of the game to compete with other leaders instead of to cooperate for the benefit of all the civilizations without jeopardizing any one country's security?

In an interactive medium the interpretive framework is embedded in the rules by which the system works and in the way in which participation is shaped. But the encyclopedic capacity of the computer can distract us from asking why things work the way they do and why we are being asked to play one role rather than another. As these systems take on more narrative content, the interpretive nature of these structures will be more and more important. We do not yet have much practice in identifying the underlying values of a multiform story. We will have to learn to notice the patterns displayed over multiple plays of a simulation in the same way that we now notice the worldview behind a single-plot story. Just as we now know how to think about what made Tolstoy propel Anna Karenina in front of that train or what made the producers of Murphy Brown offer her happiness as a single mother, we need to learn to pay attention to the range of possibilities offered us as interactors in the seemingly limitless worlds of digital narrative.

Digital Structures of Complexity

Like every human medium of communication, digital media have been developed to perform tasks that were too difficult to do without them. Hypertext and simulations, the two most promising formats for digital narrative, were both invented after World War II as a way of mastering the complexity of an expanding knowledge base. The mathematician Vannevar Bush put it this way in his landmark 1945 magazine article, "As We May Think": "The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as that used in the days of square-rigged ships" (p. 102).

Bush's solution was "associational indexing" in a kind of magical desk based on microfilm files, a solution he called a "memex" and described as follows:

> The owner of the memex, let us say, is interested in the origin and properties of the bow and arrow. Specifically he is studying why the short Turkish bow was apparently superior to the English long bow in the skirmishes of the crusades. He has dozens of pertinent books and articles in his memex. First he runs through an encyclopedia, finds an interesting but sketchy article, leaves it projected. Next, in a history, he finds another pertinent item, and ties the two together. Thus he goes building a trail of many items. Occasionally he inserts a comment of his own, either linking it into the main trail or joining it by a side trail to a particular item. . . . Thus he builds a trail of his interest through the maze of materials available to him.
>
> And the trails do not fade. (P. 107)

This earliest vision of hypertext reflects the classic American quest—a charting of the wilderness, an imposition of order over chaos, and the mastery of vast resources for concrete, practical purposes. In Bush's view, the infinite web of human knowledge is a solvable maze, open to rational organization.

By contrast, Ted Nelson, who coined the term *hypertext* in the 1960s and called for the transformation of computers into "literary machines" to link together all of human writing, has been more in love with the unsolvable labyrinth. He sees associational organization as a model of his own creative and distractible consciousness, which he describes as a form of "hummingbird mind."[11] Nelson has spent most of his professional life in the effort to create the perfect hypertext system, which he has appropriately named *Xanadu*. He describes this pursuit as a quixotic quest, "a caper story—a beckoning dream at the far edge of possibility that has been too good to let go of, and just too far away to reach, for half my life."[12] Nelson's vision of hypertext is akin to William Faulkner's description of novel writing as a futile but noble effort to get the entire world into one sentence. Those like Nelson who take delight in the intricacies of hypertext, the twisting web rather than the clear-cut trail, are perhaps seeing it as an emblem of the inexhaustibility of the human mind: an endless proliferation of thought looping through vast humming networks whether of neurons or electrons.

The allure of computer simulations comes from a similar attempt to represent complexity. Three years after Bush's suggestion of the memex machine, Norbert Wiener founded the discipline of system dynamics with his book *Cybernetics*. Wiener observed that all systems, whether biological or engineered, have certain characteristics in common, such as the intertwining of multiple cause-and-effect relationships and the creation of feedback loops for self-regulation. Wiener called attention to parallels, for instance, between the way the body keeps a constant internal temperature by instituting changes (like sweating) and monitoring their effects (like feedback on skin temperature) and the way a home thermostat maintains a set temperature. Over the past fifty years, systems thinking has been applied to everything from family structure to frog ponds. It is now commonplace for us to think of the earth itself as a giant ecosystem, in both biological and political terms.

The computer has developed during this time into a versatile tool for modeling systems that reflect our ideas about how the world is organized. Early uses of computer simulations involved putting different values into a constant model and running the system through several "time steps" to see, for example, what would happen to crime statistics five, ten, and fifteen years down the line if police presence went up and cocaine prices went down. These systems were run in batch jobs, which spit out big chunks of numerical data. Other more responsive systems modeled a dynamically changing world open to real-time interaction, like the cockpit simulations used for training airplane pilots. In recent years, computer scientists have designed networked systems that are like a society full of autonomous individuals who talk and work with one another but have no single leader or controller.

In the late 1970s computer system design reached an intriguing milestone with a simple but elegantly conceived program that seemed to simulate life itself. The system is based on a checkerboard grid with markers that are white on one side and black on the other. The markers begin in a random arrangement and are then turned over according to a set of rules that makes decisions based on the colors of a marker's neighbors. Each round of turning causes more turning on the next round, eventually causing remarkable patterns to emerge and move across the board. The Game of Life system does not require a computer, but the patterns look particularly striking on the computer screen, which can run through multiple turns very quickly.[13]

Although no one would claim that such a system is alive in the same way as an animal or plant, it does capture one of the chief attributes of life—the creation of large patterns as a result of many smaller effects. Computer simulations like this are tools for thinking about the larger puzzles of our existence, such as how anything as soulless as a protein can give rise to something as complex as consciousness.

T. S. Eliot used the term *objective correlative* to describe the way in which clusters of events in literary works can capture emotional experience.[14] The computer allows us to create objective correlatives for thinking about the many systems we participate in, observe, and imagine. The rules for artificial life forms can be described as a kind of a game, but the knowledge about the world that the model offers us is not gamelike. It is a behavioral artifact that speaks to one of the most profoundly important aspects of our lives. The more we see life in terms of systems, the more we need a system-modeling medium to represent it—and the less we can dismiss such organized rule systems as mere games.

Current narrative applications overexploit the digressive possibilities of hypertext and the gamelike features of simulation, but that is not surprising in an incunabular medium. As digital narrative develops into maturity, the associational wildernesses will acquire more coherence and the combat games will give way to the portrayal of more complex processes. Participating viewers will assume clearer roles; they will learn how to become orienteers in the complex labyrinths and to see the interpretive shaping in simulated worlds. At the same time as these formal qualities improve, writers will be developing a better feel for which patterns of human experience can best be captured in digital media. In this way a new narrative art will come into its own expressive form.

The process by which this new art form will emerge is already under way and is itself interactive. Each time developers create new genres of digital stories or more immersive games, interactors try them out and grow frustrated or enchanted. Most often these incunabular products arouse expectations they cannot yet fulfill—for more encyclopedic coverage, for greater freedom of navigation, for more direct manipulation of the elements of the story. Every expressive medium has its own unique patterns of desire; its own way of giving pleasure, of creating beauty, of capturing what we feel to be true about life; its own aesthetic. One of the functions of early artifacts is to awaken the public to these new desires, to create the demand for an intensification of the particular pleasures the medium has to offer. Therefore, the next step in understanding what delights or dangers digital narrative will bring to us is to look more closely at its characteristic pleasures, to judge in what ways they are continuous with older narrative traditions and in what ways they offer access to new beauty and new truths about ourselves and the world we move through.

PART II

The Aesthetics of the Medium

Chapter 8

Eliza's Daughters

The novelist . . . makes up a number of word-masses roughly describing himself . . . gives them names and sex, assigns them plausible gestures, and causes them to speak by the use of inverted commas, and perhaps to behave consistently. These word-masses are his characters.

—E. M. Forster, *Aspects of the Novel*

In any literary medium characters are illusions. Emma Bovary, David Copperfield, and Huckleberry Finn are "word masses," as Forster reminds us, and when they are translated to the movie screen, they only exist as montages of camera shots, pasted snippets of light and sound. What difference will it make, then, to create characters from bits, from digitized words, images, sounds, and—most significantly—from instructions for behavior? When Joseph Weizenbaum invented the computer-based character Eliza in 1966, he gave us a hint of the answer. Since then Eliza has had many direct and indirect descendants, enough so that we can identify some distinct technical and artistic strategies for making computer-based characters.

Chatterbots

Probably the most famous of Eliza's daughters is the virtuoso character known as Julia, developed by Michael Mauldin of Carnegie Mellon University. Julia is a "chatterbot," a text-based character like Eliza who carries on conversations with the people around her.[1] Julia was built to live on MUDs, and she has many agreeable social behaviors: she plays the card game hearts, keeps track of other inhabitants, relays messages, remembers things, and gossips. In short, Julia is good company. Moreover, her physical presence is as real as anyone else's; that is, she can hold things, perform actions like other MUDders, and move from place to place. Like any other being, she appears in only one room at a time. MIT researcher Leonard Foner, who has studied her behavior appreciatively, records her dramatic appearance in answer to the typed-in command "page julia":

> You sent your summons to Julia.
> Julia pages from Stevi's Kitchen: "I'm on my way to you, Lenny."
> Julia is briefly visible through the mist.

Julia is clearly a chatterbot who knows how to make an entrance. Foner also offers this example of her entertaining conversational skills:

> You say, "Hi, Julia. What do you do?"
> Julia says, "I gossip, Lenny."
> You whisper, "Julia, tell me some gossip," to Julia.
> Julia whispers, "I once heard baba say, 'Drop kick me, Jesus, through the goal posts of life.' "

Julia's responsiveness and appropriateness to her surroundings induces people to collaborate with her in little dramatic scenes. For instance, Foner describes his own delight in seeing her after she had been off-line for a while (he immediately burst into the first line of her theme song):

Lenny: Julia, Julia, she's our guide!
Julia: She directs us far and wide!
Lenny: Are you lost, can't find your way! Julia's here to save the day!
Julia bows.
Lenny laughs.

As a character with a female name who lives on MUDs, Julia usually engages in that familiar form of collaborative comedy known as pickup banter. MUDs only accentuate this already highly patterned and predictable behavior, making them the perfect venue for Eliza-like creatures, who excel at parodic interactions:

Space-Ace: So why won't you kiss me, Julia?
Julia: Sorry, I'm not into geeks, Space-Ace.
Space-Ace: Julia, I'm not a geek.
Julia: Then lose the pocket protector, Space-Ace.
Space-Ace: Julia, I don't have a pocket protector.
Julia: Congratulations.
Space-Ace: So kiss me, Julia.
Julia: I'd rather be overhauled by a Yugo mechanic, Space-Ace.
Space-Ace: Julia, kiss me.
Julia: Why, did your sheep leave you?
Space-Ace: Julia, kiss me.
Julia: Take a long walk through an unlinked exit, Space-Ace.

Julia's impersonation of an actual female MUDder was so successful that Mauldin was distressed to discover that one poor soul had spent thirteen days trying to seduce her into going with him to a private room for virtual sex. Among their more memorable interactions are these exchanges:

Barry: Are you a virgin?
Julia: As far as you're concerned, Barry, I always will be.

Barry: I'm inviting you to enjoy my private room.
Julia: So?
Barry: Well, I thought you liked me and wanted to be together.
Julia: I live for your approval, Barry.[2]

In some ways Julia is a female impersonator like the other female personae operated by male MUDders, and her character shows some of the exaggeration of a drag character. For instance, one of Mauldin's playful details was to program Julia to complain at appropriate intervals about her period and to make jokes about PMS. Here is one such exchange, elicited by one of Foner's colleagues:

Robyn: Julia, are you human?
Julia: Is the Pope Polish?
Robyn: Yes, Julia, but are you human?
Julia winks at Robyn.
Robyn: Julia, are you a computer?
Julia is offended at Robyn's suggestion.
Robyn: Julia, are you a woman?
Julia: I am female.
Robyn: And, Julia, are you human?
Julia: Not when I'm having my period, Robyn.
Robyn laughs.

Julia clearly gets most of her vitality from her comic shtick, but she also functions as an expert conversational partner on a limited number of topics. In this regard she is a kind of "query system," which is a form of artificial intelligence program that takes in users' questions ("When is the next bus to Cleveland?" or "What is the best form of birth control?" or "Is Lenny on the MUD today?") and consults a database to generate the appropriate answer. In order to make query systems succeed, one must limit their domain of expertise and then anticipate the many ways in which questions might be asked. Julia functions as a very simple query system in her ability to answer questions and sustain conversations about hockey or pets. She was given fluency in these topics in order to compete in a Turing test contest. The Turing test was proposed in 1950 by cryptographer Alan Turing as a way to judge if machines have matched human intelligence.[3] In its usual form, judges converse with both machine-based characters and humans through a keyboard and then decide which is which. When Eliza fooled people into thinking she was an actual person, many believed it would not be long before machines could pass sustained tests of this kind. In 1991, Hugh Loebner offered a $100,000 prize for the first program to pass the test. The prize has yet to be awarded, but the annual Loebner competitions are an entertaining reminder of the goal. Here is an excerpt from one of Julia's entries:

> *Judge:* Do you know anything about miniature schnauzers?
>
> *Julia:* I don't know anything about miniature schnauzers, Judge.
>
> *Judge:* Well, I have one. We call him Pucci. He is a crazy cat—have you ever had a cat?
>
> *Julia:* What a coincidence! I have two cats.
>
> *Judge:* I prefer dogs myself. Cats are too independent and they always jump up on things. By the way, are you a hep cat?
>
> *Julia:* People don't own cats . . . it's the other way around.
>
> *Judge:* How do you feel about declawing?
>
> *Julia:* What kinds of pets did you have as a child?
>
> *Judge:* Won't you answer my question? How do you feel about declawing?
>
> *Julia:* Guinea pigs are about as smart as cucumbers.

Julia is not a very believable character in the Loebner arena, even though she is able to fool people like Barry in the MUD. It is easier to create a fictional personality than it is to model knowledge of the world. In fact, Julia is more imaginatively present when she is less intelligent, because her sense of presence does not come from giving factually correct information but from demonstrating dramatically appropriate behavior. The MUD provides a social framework in

which her formulaic responses make sense. Indeed, her exaggerated gender behavior is itself a good distracter: it makes people focus on a possible male-as-female impersonation rather than on the machine-as-human impersonation.[4] But what most brings Julia to life on the MUD is the fact that her interlocutors are also in character. They come to the conversation predisposed toward dramatic collaboration. A conversation with a chatterbot is a kind of improvised skit between human and computer-controlled actors. A successful chatterbot author must therefore script the interactor as well as the program, must establish a dramatic framework in which the human interactor knows what kinds of things to say and is genuinely curious about how the chatterbot might respond.

Authorship of Chatterbots

For a nonprogrammer the challenge of writing a character like Julia is to distract people away from the machinery, to make the illusion more interesting than the technical wizardry. One development that would help make such characters more compelling would be to make it easier for nonprogrammers to write them. To that end I have tried to give students in my interactive-fiction-writing course the opportunity to create their own chatterbots with a very simple authoring system. Although their results are only ambitious class projects and by no means complete characters, the evolution of their writing techniques offers a good case study of literary technique rising to meet a newly available technology.

One of the first things students learned in creating their own chatterbots was how to select a good character. The most successful characters have been those who are self-absorbed, evasive, or obsessive in familiar ways. For instance, the first set of characters the students created included several politicians in the situation of a press conference (it was an election year). Other good scenarios included a "babysitting job from hell," a folksy grandpa who loves eating gruel and going fishing, and a smug, pretentious artist. There were many comical romantic partners, including a particularly revolting Continental seducer and a "psychotic girlfriend," who was one of the most successful characters at sustaining a coherent conversation.[5]

Like all of the characters, Girlfriend begins with an opening line that immediately establishes the situation and scripts the interactor:

> *Girlfriend:* So, why do you want to break up with me?
> *Interactor:* I need space.
> *Girlfriend:* Let's talk about the other woman.
> *Interactor:* What other woman? I've been totally faithful.
> *Girlfriend:* Where did you meet her?
> *Interactor:* Meet who? I told you I'm the faithful one in this relationship.
> *Girlfriend:* How could you do this to me?
> *Interactor:* You did this to yourself, babe!
> *Girlfriend:* Do you love me?
> *Interactor:* Well, I used to. I know better now.
> *Girlfriend:* I can never feel for another man as much as I feel for you.
> *Interactor:* Well, maybe I'm being too hard on you.
> *Girlfriend:* How could you break up with me when you know how I feel about you?
> *Interactor:* You're smothering me. I need space.
> *Girlfriend:* You've hurt me deeply.
> *Interactor:* You hurt me too.

The rigidity of the computer is well used in the tunnel vision of the character. The interactor enters into the collaboration immediately and moves from a simple antagonistic attitude ("You did this to yourself, babe!") to opening up to her pain ("Well, maybe I'm being too hard on you"), which gives the scene a hint of dramatic development.

For any chatterbot, the test of coherency is how it deals with the inevitable problem that arises when the interactor's utterance contains no key word. One way to handle this is to do what Julia does in the "declawing" example: start up a new topic (like guinea pigs). Another way is to sustain the previous topic, as Girlfriend does when she follows up "Let's talk about the other woman" with "Where did you meet her?" The most common strategy is to provide a set of "default utterances," comments that can fit in anywhere ("How could you do this to me?"). After this default strategy had been established practice for a while, one student used it to construct a salesman character who returns to his sales patter if confused. Salesman was the first character to use a default file as a means of providing a plot for the scene. Salesman begins with hard-sell optimism like this:

- My name, by the way, is Brian. Have you heard of the WidgetMaster product line?
- The WidgetMaster is currently in use in more than 10,000 homes across the nation!
- Let me know at any time if you would like to see a demo of the WidgetMaster in action.
- Might you be interested in purchasing a WidgetMaster today?

He then progresses through increasingly anxious self-disclosure:

- Please buy one! It really will change your life.
- I really need to hit my quota this month, too. Please?
- I don't want to pull a sob story on you here, but please, buy one?
- Look, ever since Mr. Widget's son took over, I've been under the gun.

Salesman ends in outright despair:

- Sigh.
- I give up. I'm going to miss my quota and lose my job.
- I can't believe my life has turned out like this.
- Sigh. Sorry for taking your time.
- I really don't know what to do.[6]

The final reply is coded as repeatable, so that the conversation from then on is peppered by the refrain.

In creating Salesman, the author invented a method of making the rules for a conversation serve as a dramatic structure for an interactive scene. This modest exercise is an example of how new narrative techniques can develop when a tradition of composition is supported by an authoring environment that does not require programming. If such software environments, created and refined by programmers working in collaboration with writers, were more widespread, chatterbots could move beyond their current status as engaging novelties. Computer characters who can carry on persuasive conversations could be an expressive narrative genre in themselves, as well as compelling elements in a larger fictional world.

Modeling the Inner Life

Although Joseph Weizenbaum chose the psychiatric interview only because it allowed him to bypass the problems of knowledge representation that haunt the Turing test encounters, other researchers of the 1960s and 1970s were drawn to the idea of the psychiatric interview in the sincere hope that the computer might serve as a therapist or as an aid to therapy through the modeling of the human psyche. Psychoanalyst Kenneth M. Colby, one of the most ambitious of these researchers, improved upon Weizenbaum's ELIZA by creating PARRY, a model of a paranoid person that was convincing enough to pass a kind of Turing test. Colby's description of his "hypothetical individual" reads like a doctor's presentation of a case:

> He is a 28-year-old single Protestant male who works as a stock clerk at Sears, a large department store. He has no siblings and lives alone, seldom seeing his parents. He is sensitive about his parents, his religion, and about sex. His hobby is gambling on horseracing, both at tracks and through bookies. A few months ago he became involved in a severe quarrel with a bookie, claiming the bookie did not pay off a bet. After the quarrel, it occurred to him that bookies pay protection to the underworld and that this bookie might gain revenge by having him injured or killed by the Mafia. He is eager to tell his story and to get help in protecting him from the underworld. He is willing to answer questions about nonsensitive areas of his life and offers hints about his delusional system in an attempt to feel out the interviewer's attitude toward him.[7]

In fact, this person is as fully imagined as many fictional characters. Parry conversed via teletype with eight psychoanalysts, who also conversed in the same way with actual patients, some of whom were diagnosed as paranoid and some of whom were controls. Although the doctors could easily tell the paranoid patients from the nonparanoid people, they accepted Parry as another paranoid rather than as a mechanical impersonation. (This may say more about their self-fulfilling models of paranoia than about PARRY's computational subtlety, of course.) Here is a sample interview excerpt, as Colby reported it:

> *Doctor:* How did you come to be in the hospital?
> *Patient:* I am upset.
> *Doctor:* How long have you been in the hospital?
> *Patient:* About a week.
> *Doctor:* In what way are you upset?
> *Patient:* People make me nervous.
> *Doctor:* In what way do they make you nervous?
> *Patient:* People give me a funny look sometimes.
> *Doctor:* Is there anything about people in particular?
> *Patient:* Certain persons.
> *Doctor:* Who are the certain persons and what is it that they do?
> *Patient:* Do you know how the underworld operates?
> *Doctor:* How do you mean?
> *Patient:* Shouldn't I be?
> *Doctor:* Why did you ask about the underworld?
> *Patient:* What information do you have on the underworld?

Doctor: I don't know about the underworld but when you mentioned the underworld I thought you must have some contact with it and that's what I would like to know more about.

Patient: I try to avoid the underworld.

Doctor: Have you been actively trying to avoid the underworld?

Patient: No one has any power over gangsters.

At this point the psychiatrist remarks, "This response is a non sequitur. It's possible he's become so defensive about the underworld he's decided not to answer directly."[8] His explanation is a classic example of what computer scientists call the "Eliza effect," that is, attributing greater intelligence or intentionality to a machine than it possesses. It is also an example of the much older effect we earlier identified as the true source of Eliza's power—the human propensity to suspend disbelief in the presence of a persuasive dramatic presence. The psychiatrists were taken in by the persuasiveness of the animated presence; its real-time responses to their questions; and its very literally constructed, detailed obsessions with bookies and gangsters. Just as Julia's exaggerated female persona distracts from the man-or-machine question, Parry's paranoia focuses the observer on the quality of his mental life, not on whether it is authentic or not.[9] Colby sent transcripts of the interviews of Parry and of the paranoid patients to a group of randomly chosen psychiatrists and another group of randomly chosen computer scientists and asked them which was which, thus conducting a classic Turing test. The result was no better than chance in either group.

Parry could be taken as another example of a shtick comedy figure in the Eliza genre, but he is significantly different from Eliza in that Colby provided him not with a mere set of phrases but with a rudimentary model of an inner life. Colby gave his creature a "belief system" in which fairly innocuous beliefs about bookies lead to a core of emotionally charged delusions about gangsters. He gave Parry a model of a state of mind, including monitors for anger, fear, and mistrust. He instructed Parry to construct a model of his interviewer on the basis of each question and to decide if the interviewer's intent is malevolent, benevolent, or neutral. If Parry thinks a question is malevolent and his fear level is high enough, he clams up and stops talking about his delusions. He therefore has an inner state and a repertoire of behavior with which he can reflect it. Interactors can tell an "upset" Parry from a calm Parry, and the things they say can make a difference in the mood of this "patient."

Colby's most ambitious creation was an automated patient, Neurotic Woman, who hates her father because she feels he deserted her as a child but who censors this thought through standard Freudian processes. When she runs in "free-association mode," she is impelled to say "I hate my father," but she is also impelled to repress the thought, leading her to offer distorted expressions like "I hate Mrs. Smith (displacement) and "Father hates me" (projection). The Neurotic Woman program also runs as a psychoanalytic conversation in which the interactor plays the part of the analyst and can break into the free association now and then with carefully chosen suggestions. Here is an excerpt from this simulated analysis:[10]

Patient: Father is indifferent to me.
Mother says father abandoned us.
Father preferred sister.
I avoid father.
Sister admired father.
I feel tense now.

Therapist: You admire father.

Patient: No.
I do not admire father.
I admire Al.
Mother ought not to admire Al.
Mother detests father.

Colby was influenced by theories of cybernetics as well as by Freud in designing his model, and his program traces the patient's moment-to-moment state of mind (as expressed in numbers representing degrees of anxiety, excitation, pleasure, self-esteem, and well-being) through intricate feedback mechanisms that regulate the degree of distortion applied to any potential statement. The aim of the psychoanalyst is to come up with ideas that will move the patient toward confronting her contradictory and painful thoughts while raising her self-esteem and keeping her anxiety in check. Since the goal of the interaction is to enable the simulated patient to express her hatred of her father directly and thereby abandon her neurotic belief system, Colby can be credited as the first person to conceive of an automated fictional character with an inner self that is capable of change and growth.[11]

Although Colby's work was meant as science, it is more persuasive as literature. Both Parry and Neurotic Woman stand as tantalizing examples of a computer-based character who is modeled not as a collection of surface behaviors but as a layered personality.

Goal-Based Critters

As intriguing as automated patients are as literary models, they were abandoned by computer scientists in the 1970s and 1980s in favor of a behavioral and cognitive model of personality and social interaction that could be programmed in terms of "scripts, plans, and goals."[12] Experiments like Tale-spin, which created the dithering Joe Bear we met in chapter 7, reflect a preference for describing characters not in terms of their psyches but their goals. In the 1980s and early 1990s, as computing power and available memory increased exponentially, computer scientists began exploring software and hardware strategies for "parallel processing," for creating systems that could do multiple things at the same time. Leading engineers turned from building all-encompassing centrally controlled systems to designing worlds made from a collection of "intelligent agents," each of whom was pursuing its own goals. This change in computer architecture has an equivalent effect for the creation of digital narrative. It is as if computer scientists stopped trying to build a world by coming up with an omniscient storyteller and decided instead to create it out of a collection of autonomous characters. Moreover, the characters they are currently building are not the single-minded, top-down planners of Joe Bear's generation. They are improvisers, aware of multiple goals at once and able to change their priorities and behaviors in response to changes in their environment. These characters are often called "intelligent agents."

Computer science research on "intelligent agents" often focuses on utility programs that act like servants. The classic example of a desirable software agent is one who goes out over the Internet and books airline flights and hotel reservations, consulting the client's preferences on budgeting and on arrival and departure times and taking advantage of opportunistic events like "frequent flyer" bonuses. Since the desire for such agents is often couched in terms very like those used by the early Disney cartoonists to create an "illusion of life,"[13] it is not surprising that interest in them has been closely intertwined with a desire to create fictional characters. For computer scientist Joseph Bates, the key goal for such research is "believability," which he has tried to achieve in several experimental character-based fictional worlds.

One of the first characters created by Bates' Oz group was Lyotard, the housecat. Lyotard was named for the postmodern theorist and is perhaps a fitting emblem of the times; he is the result of world-class expertise brought to bear on the problem of making a machine made of plastic and silicon impersonate a small, furry animal. It is a largely successful effort, one that could easily be translated to a virtual reality or graphics-based program, since Lyotard lives within a virtual house. For now, Lyotard's actions are expressed in text, as in the following representative scenario:

Interactor> Go to the dining room.
You are in the dining room.
A jar, a bread basket, a black cup and a large plate are on the end table.

Three bran muffins are in the bread basket.
Nine black sardines are in the jar.
Interactor> Take the jar.
You take the jar.
Interactor> Go south.
You are in the sunroom.
Lyotard goes to the sunroom.
Interactor> Give a sardine to Lyotard.
You offer the black sardine to Lyotard.
Lyotard runs to the dining room.
Interactor> Follow Lyotard.
You run to the dining room.
Lyotard looks around nervously.
Interactor> Pet Lyotard.
You pet Lyotard.
Lyotard bites you.
Interactor> Hit Lyotard.
You hit Lyotard.
Lyotard is now shaking.
You hear a meow.
Lyotard runs to the sunroom.
Interactor> Look around.
You look around yourself.
The end table and the small chair are in the dining room.
The bread basket, the black cup and the large plate are on the end table.
The three bran muffins are in the bread basket.
Lyotard looks around nervously.
Interactor> Eat a muffin.
You eat the bran muffin.
Lyotard goes to the dining room.
Interactor> Wait.
You wait.
Lyotard goes to the kitchen.

Interactor> Follow Lyotard.
You go to the kitchen.
Lyotard is in the kitchen.
You hear a meow.
Interactor> Look around.
You look around yourself.
You are in the kitchen.
A long counter, a cabinet, a beautifully crafted shelf, a plastic trash can, a yellow bowl, a red superball, a white toy mouse and Lyotard are in the kitchen.
A small plate is on the beautifully crafted shelf.
The red catfood is in the cabinet.
The blue catfood is on the long counter.
You hear a meow.
Interactor> Pour the jar in the bowl.
You pour the jar in the yellow bowl.
Lyotard eats a black sardine.
Interactor> Wait.
You wait.
Lyotard eats a black sardine.
Interactor> Wait.
You wait.
Lyotard eats a black sardine.
Interactor> Pet Lyotard.
You pet Lyotard.
Lyotard eats a black sardine.
Interactor> Again.
You pet Lyotard.
Lyotard looks lazily at you.[14]

Although this scene may seem far from great drama, it is a very promising achievement. With Lyotard, Bates and the Oz group have succeeded in creating a cat with "emotions, social behavior, and personality." They have done this with an architecture that gives him an inner life that is centered on goals (such as satisfying hunger and relaxing) but that also includes a representation of his emotions. For instance, Lyotard begins by being fearful of the interactor and not very hungry, which is why he is not interested in the sardine and runs away from the interactor. When the interactor follows him from the sunroom to the dining room, Lyotard's goal of escaping fails, which makes him more aggressive. Therefore, when the interactor, whom the cat recognizes as an object of mild hate, goes to pet him, Lyotard forms the goal of "respond negatively to contact," which results in his choosing to bite the interactor. Why does Lyotard not hold a grudge? Because the interactor stays away from him, allowing his angry feelings to decay. Meanwhile, the cat's hunger is building. Therefore, by waiting for the cat to get hungry and by then offering him the special treat of sardines in a nonthreatening way, the interactor is able to make friends with him. This demo script is a satisfying dramatic scene, a collaborative improvisation much like a conversation with Eliza but based on gestures rather than words. But to what degree is Lyotard a characterization rather than a mere mechanized model of a cat?

Lyotard is certainly a complex machine. His psyche is built on a cognitive science schema that is widely used in the computer modeling of personality.[15] His inner life is built on an intricate but precise calculus in which events are compared against goals, actions are compared against standards, and objects are compared against attitudes; Lyotard's psyche is a giant emotional algebra equation in which all the values are changing all the time. Success in a goal yields the emotion of joy and failure yields the emotion of sadness, but this schema can also represent ambivalence. Since even a simple housecat is a complex, dynamic system, many goals will be active at any given time, and some actions will satisfy one goal but frustrate another. As a behavioral and motivational model, Lyotard and his kind take us far beyond the "kill/don't kill" characters of action games. Lyotard can capture complex emotions like reproach, which is defined as the re-

action to another person acting in violation of one's own inner standards; thus, Lyotard displays reproach when an interactor sits down in the cat's favorite chair. This framework is both elegant and absurd. It is elegant in that one can account for a wide range of emotions (including composite emotions like anger, which is represented as a combination of reproach and sadness) using a limited set of building blocks and for a range of emotional intensity that is expressed quantitatively (e.g., dislike of interactor = 1; dislike of dogs =10). But the cognitive model of emotions quickly becomes absurd when we try to apply it to the emotional states of actual human beings (dislike of Barney = 1; dislike of Hitler = 10), and it seems the very antithesis of what we value in literature, which is the careful examination of ambiguous situations open to multiple interpretations. A Tolstoy of the next century could hardly model Anna Karenina's conflict between her love for the passionate Vronsky and her love of her son by setting a panel of affect sliders and filling in a template with her goals and standards.

Moreover, even for the more modest task of describing a housecat with the same level of believability as a cartoon cat, the abstract science can only take us so far, for Lyotard's behavior in this scenario could not have been produced by merely mapping some common cat behaviors (eat when hungry, bite when angry) onto a standard emotion machine. Since their goal was believability, the authors were forced, almost against their scientific instincts, to give Lyotard some charm, to make him a particular character. For this goal they discovered that the cognitive science formulas did not take them far enough and that there was no scientific taxonomy to take them the next step of the way. Instead, they had to give Lyotard a set of features in addition to the canonical emotions, features that they present almost apologetically as invented without a known "structured set" and "very ad hoc." The authors' list of proposed features for Lyotard include many of the traits humans find most entertaining in housecats, namely, their capacity to be curious, content, aggressive, ignoring, friendly, proud, and energetic. Without the aggressive feature, for instance, Lyotard would not have been disposed to bite the interactor in the above scenario. In other words, Lyotard's most dramatically interesting behavior arises from a specific personality structure the authors improvised, on top of the more generic model, just for him.[16]

The emotional schema is therefore more like a palette than a portrait. It provides a way of specifying elements of the personality and of linking behavior to an interpretive model of the character's inner life. The need for an ad hoc, unscientific characterization strategy is even clearer in the more ambitious Oz group project called Edge of Intention, a multicharacter world in which oblong animated figures called Woggles jump and slide around a two-dimensional landscape displayed on the computer screen. Here the researchers explicitly differentiated the characters by mapping each of their emotions to a personality-specific feature. For example, the same emotion of "fear" is mapped to the "alarm" feature in the vulnerable Shrimp but to the "aggression" feature in his nemesis, the bullying Wolf. But Oz designers still found it very hard to communicate these emotional states to interactors, since the Woggles' world does not offer the interactor a familiar script, as the housecat scenario does, to provide the dramatic context for the characters' actions. In fact, the most expressive and lifelike character trait arose from a programming glitch that caused Shrimp to bang his head on the ground now and then in what Bates considered a sort of nervous tic. At exhibits of the world of Woggles around the country, this head banging stole the scene. The researchers who knew this behavior to be incoherent in Shrimp's true emotional life were struck by how it immediately attracted people's attention. "To our surprise," wrote Bates, "they build psychological theories, always incorrect, about Shrimp's mental state and seem to find him much more interesting and alive for having this behavior." Having created their most expressive gesture by accident, the computer scientists were left to debate the extent to which believability rests on clean abstraction and how much it rests on quirks.[17]

Lyotard and Shrimp are a promising beginning for computer-based

characters with both coherent behavior and believable personalities. The emotional abstraction system, like any other formalization of human personality, can be reductive and clumsy, and many would dismiss it out of hand as a persuasive representation of our inner lives. However, it can provide us with useful building blocks for creating patterns of intelligible behavior or stencils for drawing in the general outlines of character. Nevertheless, for the work of characterization, for making creatures that engage us, we need the free-form drawing, not the stencil; we need the ad hoc element, the features that turn a block drawing into a face. We need to place the formulaic elements, the stenciled images, within an idiosyncratic arrangement based not on science but on an interpretation of the world. We have to build the cat to reflect how humans feel about their housecats rather than what they might learn about them in the dissecting room. If we are to reach the point where we can create complex computer-based characters with the same expressiveness that is true of the characters we now create with word masses and shots of film, then we must start by building authoring systems that will put the abstraction expertise of the precise computer scientist into the interpretive hands of the quirky artist.

Multicharacter Environments

We are just beginning to have experience with fictional environments in which automated characters interact with one another. For instance, at MIT and elsewhere Eliza runs within an environment that also includes a repository of the sayings of the syndicated cartoon figure Zippy, the pointy-headed clown whose surreal adventures make fun of the extremes of the popular culture.[18] This led to the inevitable meeting of the two. If one types in the command "psychoanalyze-pinhead," Zippy's sayings are fed to Eliza, and though he cannot listen to her, she responds to him in the same way she would to any interlocutor. Most of their exchanges are nonsensical, but amid pages of uninteresting non sequiturs come coherent exchanges like this:

Eliza: I am the psychotherapist. Please, describe your problems.

Zippy: We are now enjoying total mutual interaction in an imaginary hot tub.

Eliza: Why do you say that?

Zippy: I don't know WHY I said that. I think it came from the FILLINGS in my rear molars.

Eliza: What makes you believe in your rear molars?

Zippy: Give them RADAR-GUIDED SKEE-BALL LANES and VELVEETA BURRITOS!!

Eliza: Maybe your plans have something to do with this.

Zippy: Thank God!! It's HENNY YOUNGMAN.

When Eliza asks, "Why do you say that?" and Zippy replies, "I don't know WHY I said that," it is just a lucky accident. But the surreal unflappability of Eliza is a perfect foil for Zippy's outrageous non sequiturs.[19]

If we were to create a world for Zippy and Eliza to share we would want to maximize such exchanges. We would have to create them so that they each knew in advance the kinds of things the other was likely to say and had a repertoire of ready responses. Even putting aside the difficulties of getting a computer to understand language, this is a formidable task. Yet computer scientists are moving ever closer to multicharacter environments and are beginning to experiment with making coherent activity from the independent actions of multiple agents.

For instance, interactors with the Edge of Intention exhibit were allowed to operate one of the Woggles, like a videogame character, using a mouse to make it hop and slide through the simple landscape. The interactor's Woggle can greet and be greeted by the other creatures, influence the dynamic between the bullying Wolf and the timid Shrimp, and even lure other Woggles into a game of follow the leader. One of the problems of this arrangement is that it is "very difficult for people interacting with the creatures to stay aware of what is happening," because too many things are going on at once without the benefit of staging aimed at focusing attention on a central action. In other words, in addition to performing one's own character's repertoire of actions, deploying them appropriately and responding to the other characters in a multicharacter world, one must have a way of synchronizing these individual actions with the general action so that one is presented with a coherent picture. And if there is more than one interactor, the staging problem has to be solved either collectively or individually for all of them.

One solution to the staging problem of multicharacter improvisations is found in the practices of the commedia dell'arte, the popular Italian theater tradition that flourished throughout Europe from the Renaissance through the eighteenth century and that has influenced theater traditions from opera to *Saturday Night Live.* The commedia was based on a handful of stock characters, exaggerated types associated with particular actors in the troupe. The basic characters were the same whatever the plot, just as John Wayne or Groucho Marx or the battling lovers played by Spencer Tracy and Katharine Hepburn were the same from one movie to the next. A small number of actors (usually no more than seven) were able to perform a large repertoire of plays, ranging from farces to heavy melodramas, even though none of the plays had a written script. They did so by developing predictable formulas of interaction that gave shape to their improvisations.

Instead of a script, the actors relied on a scenario that offered clear entrances and exits and a paraphrase of each scene. For instance:

> *Pollicinella* enters with a lantern and a sword. He is awaiting his master. He lies down, dowses his light and makes ready for sleep, whereupon
>
> *Don Giovanni* leaps from the balcony. At the noise Pollicinella wakes up. They play a scene of combat in the darkness; then they recognize one another and both leave to make ready for their departure for Castile.[20]

Once the director of the company gave them the scenario, which told them when to go on and what their goal was for each scene, the actors relied on their stock characters to fill out the illusion. Whatever the setting or the story, there would always be two old men, two zany servants, a pair of lovers, a witty confidante for the prima donna to talk to.

How did the actors synchronize their utterances? In part they relied on memorized bits and pieces that were taken from the written tradition (and therefore quite polished) but appropriate for many occasions, such as elaborate insults, romantic verses, or a refused lover's complaint. In addition, they had rituals of interaction that they could adapt to varied content; for instance, the mistaken identity fight presented above could be reused in multiple plays. There were also formal patterns to the dialogue, as in this excerpt from a lovers' quarrel:

> *He:* Go! . . .
> *She:* Disappear! . . .
> *He:* . . . from my eyes.
> *She:* . . . from my sight.
> *He:* Fury with the face of Heaven.
> *She:* Demon with mask of love.
> *He:* I curse . . .
> *She:* I shudder . . .
> *He:* . . . the day that I set eyes on you.
> *She:* . . . at the thought that I ever adored you.
> *He:* How can you dare . . .
> *She:* Have you the insolence . . .
> *He:* . . . to look at me again?
> *She:* . . . to remain in my presence?

And so on through parallel accusations, denials, and renewed vows of devotion.[21] The actors also had stock comic stage business, called *lazzi,* such as the following:

> *Lazzo* of the shoes:
>
> When they are about to take Pulcinella to prison he says he must first tie his shoelaces. Then he bends down again, grabs the legs of his two guards, throws them down and runs away.
>
> *Lazzo* of the fly:
>
> Pulcinella, having been left by his master to guard the house, on being asked if there is anyone inside, replies that there isn't even a fly. The master discovers three men there and reproaches Pulcinella, who replies: "You didn't find any flies, you only found men."
>
> *Lazzo* shut up:
>
> While his master is talking Pulcinella is continually interrupting. Three times his master tells him to shut up. Then when he calls for Pulcinella, the latter pays him back in the same coin and says "Shut up!"[22]

The actors were able to improvise their parts because they worked within these formal patterns, and, indeed, current improvisational comedians are trained with very similar techniques.

With this degree of formulaic patterning in mind, it becomes possible to think of generating scenes between procedurally described characters. Certainly the battling lovers would be easy to do, as would the mischievous servant and impatient master. To make such scenes work, it would be less important to model the characters' emotional state than to give them the right patterns of interaction, namely, the *lazzi*-like formulas that would let them anticipate one another's remarks and respond appropriately.

Pulling the Strings of the Digital Puppet

The insides of a digital character should perhaps resemble the improvisational materials of an actor—including set speeches, stage business, and plot patterns—more than the insides of an ordinary person, with emotions, beliefs, and superego. Parry's specific paranoid thoughts bring him to life in a way that his having knowledge, for instance, of how to eat with a fork would not improve. Such an improvisational digital actor might be thought of as a kind of marionette whose dramatic presence depends partly upon things that are unchangeable (such as a painted-on face) and partly on behavior invented on the spot out of a repertoire of actions; just as the puppeteer improvises from the limited possible positions of the strings, so too might the procedural author program the actors to improvise by combining elements of their behavioral repertoire.

In fact, researchers at New York University have already created two charming animated actors, who, though not yet ready to perform in the commedia, can invent their own staging in a graceful and expressive manner. They have a Renaissance look to them and physiques reminiscent of Don Quixote and Sancho Panza. One is tall and thin, with graceful fingers and a mournful face, the other round with a heavier gait. Although they cannot speak, they gesture expressively and move about on their digital stage, improvising in pantomime as if waiting for a sluggard playwright to arise and script them.[23]

To what extent do we allow such actors to pull their own strings and to what extent do we put them in the hands of a plot controller? One way of thinking about how this might work is to look at another house pet, a distant cousin of Lyotard the cat. Silas T. Dog, created by Bruce Blumberg for the ALIVE Project at the MIT Media Lab, was designed as a potential actor in interactive narratives; he is largely autonomous, but he can also accept a director's commands at four levels of his complex inner life. Suppose Silas were playing with a little girl in the "magic mirror" VR environment (described in chapter 2) and the program decided to make the encounter more dramatic by getting Silas to do something mischievous, like snatching a virtual steak from a virtual kitchen table. The plot controller could prod him into action by sending a command to his motor system (*go forward*), to his behavior modules (*find a juicy steak*), or to his motivation system (*you are very hungry*). Moreover, the plot controller could change

Silas's perception of the environment in ways that are not visible to other interactors. For example, one way of getting Silas to pay attention to the little girl would be to attach a computer model of a dog biscuit to her hand that would be visible to Silas but not to her.[24]

Silas's multipart architecture, which is based on studies of animal behavior and is optimized so that he does not dither between conflicting goals, could just as well serve as an architecture for a character's personality. Instead of a motor system, we give the character a repertoire of actions appropriate to the story world (e.g., *send letter, change clothing, order murder of enemy, form posse*). On top of this layer of simple behaviors we could have motivational modules (*fall in love, swear revenge, seek allies*) and behavior modules drawn from Proppian morphemes (*defect to the enemy, woo the princess, search for the outlaws*). Instead of a virtual pet we would now have a virtual actor, ready to be sent onstage as the multiform plot calls on him. The device of the invisible dog biscuit would be risky for the creation of believable narrative, since it introduces the possibility of a fictional world that is cohabited by multiple beings who have no fixed common reality. However, giving each character its own private sentience of the world would be a virtue, since it would allow the plot controller to direct the discovery of important story elements (to make sure, for example, that the Giant does not see Jack hiding in his beanstalk home until just the right moment).

The strong point of Silas's architecture is that it allows him to respond to his environment spontaneously and to put together complicated strategies for doing things. In fact, characters like these are so complicated that they have the potential to walk away with the story altogether. They pose the question of just how autonomous we would want a fictional character to get.

Emergence as Animation

Autonomous agents like Silas and Lyotard offer the exciting possibility of what computer scientists call emergent behavior; they are able to act in ways that go beyond what they have been explicitly programmed to do. For example, in the scenario described earlier, Lyotard chose to bite the interactor from among a range of behaviors open to him at that moment. His creators had made him aggressive, and they had arranged his motivational priorities so that he would give a higher weight to responding to a nearby hand than to satisfying his general need to relax. But they did not specifically instruct him to bite that particular interactor at that particular moment. The action emerged from the intricate combination of sensations, emotions, and personality traits that shape Lyotard's simulated consciousness.

It is an important moment in human history to be able to make machines that exhibit emergence. It is a sign that we have reached a new threshold in our ability to represent complex systems—systems of any kind, whether thermodynamics, war strategies, or human behavior. In the first cybernetic models, systems were thought of as being under a central command structure, like a thermostat, and computer programs were built in simple hierarchies with one master program that controlled other programs, or subroutines. Later systems were often based on the notion of a "finite state automaton" that chugged from one complex state to another in sequences that could be charted in a neat map of circles connected by lines. But as our models of the world have become more complex, systems have become decentered: their processing operations are distributed among many entities, none of which is in central control, and the possible states of the system as a whole are no longer thought of as finite. The new emergent systems have reached such a degree of intricacy that they are their own description; there is no other way to predict everything they are likely to do than to run them in every possible configuration.

It is the focus of current work in computer science to try to control the unpredictable, to ensure that autonomous agents will always "do the right thing."[25] For instance, Bradley Rhodes in Pattie Maes's group at MIT, has created a Three Little Pigs cartoon world (displayed on a desktop computer) from creatures that are close cousins to Silas T. Dog.[26] The protagonist of the story is Wolf, who is an autonomous agent with multiple ways of catching and eating a pig, such as dynamiting the pig's house or using a pogo stick to hop to the roof and down the chimney. He does not have any central program that tells him how to do these things; his behavior is always improvised in a moment-to-moment fashion and depends on which motivations are strongest and on what his current situation and environment are like. It is exhilarating to watch Wolf in action, because it is never clear which strategy he will take as he responds to the scurryings of the pig, the availability of his tools, and his adjustable set of goals. For instance when he feels both hungry and destructive, he walks to the straw house, huffs and puffs till he blows it down, then walks over to the pig, picks it up, and eats it—a very efficient Wolf. But when his desire to be in high places is activated, instead of huffing and puffing he goes over to the pogo stick, picks it up, pogo-bounces to the chimney of the house, climbs down the chimney, drops the pogo stick, picks up the pig, eats it, picks up the pogo stick, and then bounces back to the roof of the house—an exuberant and whimsical Wolf. The character improvises each of these sequences out of his repertoire of abilities, just as we choose among various activities on the basis of our mood. But it is also easy for Rhodes, as god of this environment, to render Wolf hopelessly confused by giving him conflicting goals of equal weight. If Wolf's desire to be in high places is set equal to his hunger, Wolf may be too hungry to leave the pig and yet too fixated on his cherished pogo stick to drop it long enough to eat the pig—a comically neurotic Wolf. This is exactly the sort of situation that computer science aims at eliminating, yet it is exactly the kind of situation literature relishes. From *Hamlet*, to *Anna Karenina*, to yesterday's soap opera episode, it is dramatically depicted ambivalence rather than efficient goal selection that makes a sequence of events into a story. Rather than eliminate such moments, a procedural author using Wolf might want to enhance them by giving Wolf some expressive behavior, such as rushing back and forth between pig and pogo stick or banging his head like the Woggle Shrimp, to dramatize his absurd state of mind.

E. M. Forster thought there were two kinds of characters in fiction: "flat" ones, who perform their shtick in the same way throughout the narrative, and "round" ones, who can learn and grow. Eliza was a flat character, but Silas, Lyotard, and Wolf are—at least potentially—round ones. In digital narrative, the flatter the character the less risk it carries of breaking credibility. A character whose sole behavior is to shoot at the protagonist will always behave appropriately, but a character who can do more things—who can shoot, seek food, and elude police, for example—may dither or get lost in subgoals while ignoring events of larger importance. The more unpredictably emergent the behavior, the more it may need to be dramatically justified as a sign of a frantic or foolish person. Characters with emergent behavior may make much better keystone cops than romantic heroes.

Nevertheless, as Forster pointed out, we are more interested in characters who are capable of surprising us than in those who are flat and predictable. Forster praises Jane Austen, for instance, for creating characters who seem complex enough to live outside the limits of her plot:

> Suppose that Louisa Musgrove had broken her neck on the Cob. . . . The survivors would have reacted properly as soon as the corpse was carried away, they would have brought into view new sides of their character, and, though *Persuasion* would have been spoiled as a book, we should know more than we do about Captain Wentworth and Anne. All the Jane Austen characters are ready for an extended life, for a life which the scheme of her books seldom requires them to lead, and that is why they lead their actual lives so satisfactorily.[27]

Forster is making an argument for modeling characters not in terms of the plot demands but in terms of their whole inner life. But if we were to create such round characters on the computer, they might

leave the scheme of the story altogether. How would we judge if an emergent behavior is satisfying? The same way Forster judges roundness: "The test of a round character is whether it is capable of surprising in a convincing way. If it never surprises, it is flat. If it does not convince, it is flat pretending to be round. [A successful round character] has the incalculability of life about it—life within the pages of a book."[28]

What we look for in a created character is not mere surprise but revelation. The unexpected behavior of a fictional character must be surprising in the way that human beings are surprising; it must tell us something we recognize as being true to life. Computer scientists often use the word *random* to describe the "incalculability of life," but characters who display randomly surprising behaviors are unconvincing; they are merely flat characters pretending to be round. A truly round character would surprise the interactor by acting in a way that is consistent with its known behavior but that takes it to a new level. Its emergent behavior would have to come from a set of possibilities intentionally put there by the author, like a handful of seeds sown and then left to the vicissitudes of weather and environment to either flourish or die.

For instance, for a Casablanca simulation we might create a character like the French woman at Rick's bar who is dating a German officer but joins in the singing of "La Marseillaise" when Victor leads it.[29] Sabine (as we will call our character) would be given equal motivations of patriotism and expediency, as well as multiple, and possibly overlapping, states of mind (e.g., related to inebriation, infatuation, desire for economic security, and awareness of physical vulnerability) whose changing values would control how susceptible she might be to patriotic appeals at any given moment and also how likely she would be to display her patriotism if it were aroused. In a system in which the other characters and the major events are unpredictable, the procedural author would not know in advance whether Sabine would act as a collaborationist or a patriot in any particular scenario. Although her moral choices would be emergent behavior, the product of a particular moment and a particular history of experience within the story, the author would completely predetermine Sabine's innate capacity for such a moral conflict. To the interactor who meets her, Sabine would not appear to be acting in a random manner. Her behavior would be surprising but convincing, demonstrating that she had been created as a person with many possibilities, even with a kind of free will.

It remains to be seen whether we can capture the illusion of "the incalculability of life" with the emotional calculus of the computer. In the meantime, the most pleasurably surprising interactive characters may be those who are created with considerably less programming. For instance, the performance/installation artist Toni Dove has proposed an installation of "responsive cinema" in which two women characters, one from the past and the other from the future, are presented to a viewer whose movements are tracked with a motion sensor. As the viewer/interactor moves closer to each of the projected figures, she becomes more confiding. The process of navigating the installation is a social process, a growing intimacy or perhaps an enacted revulsion from one character or the other.[30] The content of what the characters say is conveyed in traditional dramatic materials—actors, a script, a filmed performance. No behavior emerges from such an installation that has not been carefully scripted, but the character will emerge to the interactor and the perceived sense of intimacy between them will grow as a result of their encounter.

To delineate characters with a capacity for revelation rather than computational emergence—the Anna Kareninas rather than the Silas T. Dogs—we will have to rely on moving images and written words rather than on programming code for the time being. But the next generation of graphics-based characters—animated figures capable of patterned conversations and a range of coherent gestures—should be much more fully alive than their predecessors. For instance, we can already purchase animated figures ("Dogz" and "Catz") to live on our computer screens that are similar to Silas and Lyotard in their spontaneity and computational design.[31] My own "computer pet," Buttons, is a frisky dog who romps happily when I open the program, races over to drink the water and gobble the food I put out for him, and leaps up to chase a virtual ball and bring it back to me. When I pet him, he bobs his head to direct my strokes, wags his tail, and growls with delight, and he blissfully goes off to sleep when I rub his belly. More than that, Buttons can learn tricks in somewhat the same way a real puppy can. If I hold up one of his digital treats, he will jump around and run through his repertoire of behaviors. If I give him the treat after he does the appropriate trick, he gradually learns to respond to the offer of bone-shaped digital dog biscuits by doing somersaults. Buttons, who has grown from a puppy to a larger dog since I installed him, has such a real presence for me that I sometimes feel guilty when I do not open the program and play with him. I find myself feeling proud of his affectionate personality, which is the result of the constant petting and good treatment I have given him. I know that the possibilities of his life are open and that if I had punished him with his spray bottle too often or in an arbitrary manner, Buttons's personality would be hostile and withdrawn. Playing with my computer pet is a pleasurable diversion, much more satisfying to me than stroking a stuffed animal or watching a cartoon. Characters like Buttons are a new kind of doubly animated figure, alive not only by the artfulness with which they have been made to look and move but also by the artfulness with which they have been programmed to be spontaneously responsive to the interactor's actions. The second product in the "computer pet" series, Catz, even incorporates this responsive behavior into a rudimentary story format by introducing another character—a mouse the cat can chase across the screen. The success of the Dogz and Catz programs may mark the beginning of a new narrative format, centered on appealing animated characters who might soon be as complex as Silas or Lyotard. Such characters could draw the interactor into more ambitious collaborative dramas.

Such modest incunabular creatures may seem hopelessly far from what we can achieve with Forster's "word masses," but they are nonetheless part of the same effort at understanding what it means to be human. Twentieth-century science has taught us that an important part of the answer to that question lies in understanding how complex systems like the ones the computer can embody for us resemble living things. In the centuries that have elapsed since the invention of the printing press, and as a result of the increased knowledge it has made possible, we have slipped further and further from our once cushy niche in the great chain of being—just above the conquered animals and below the encouraging angels. After Copernicus, Darwin, and Freud, we can no longer think of ourselves as somehow at the center of the universe, animated by the paternal finger of Michelangelo's God, or even (as we had hoped in the eighteenth century) as essentially innocent and rational creatures. Our solace until recently has been to celebrate our place in nature, our separateness from the increasingly mechanical world we are creating around us. Now, in the past few decades, that comforting thought is also being challenged. The brain scientists have speculated that consciousness itself may be understandable as an emergent phenomenon, the result of numerous unintelligent neurons all lighting up at just the right moment. As we slowly learn to model the processes of human thought and demystify them, the brain is left staring into a dizzying mirror. With oddly celebratory bravado, the computer scientist Marvin Minsky is fond of proclaiming that human brains, in fact, human beings altogether, are simply "meat machines." But if we are merely meat and merely machines, how are we to value ourselves and one another?

It has always been the job of the narrative imagination to answer such questions. In our time, part of the task of redefining what it means to be human lies in animating the machine, in using its system-modeling abilities to bring forth life—cuddly, affectionate, amusing, and recognizable—from empty matter. Digital dogs and cats invert the notion of a meat machine by turning an automaton into a pet. They make the idea of the mechanical less frightening by bringing it into our cultural space and domesticating it, just as our distant

ancestors made the frightening world of the beasts less so by turning the wolf into a watchdog. We may not want to acknowledge a connection between ourselves and the mechanical world, but to be alive in our time is to be faced with this reflection, like it or not. We are compelled to search for the boundary, to find out what is left within us when we take away what we think of as "meat" or "machine." With the creation of these playful, quirky exploratory characters, then, the narrative imagination is beginning to awaken to this task.