

Espen J. Aarseth

CYBERTEXT

*Perspectives on
Ergodic Literature*

© 1997 The Johns Hopkins University Press
All rights reserved. Published 1997
Printed in the United States of America on acid-free paper


4 6 8 9 7 5 3

The Johns Hopkins University Press
2715 North Charles Street
Baltimore, Maryland 21218-4363
www.press.jhu.edu

Library of Congress Cataloging-in-Publication Data will be found
at the end of this book.

A catalog record for this book is available from the British Library.

ISBN 0-8018-5578-0
ISBN 0-8018-5579-9 (pbk.)

Title page illustration: The design is the 811th generation of the staircase
hexomino, which was automatically evolved using Andrew Trevorow's
program LifeLab (with the 3-4 rule) from the initial state .

A World Wide Web site for this book can be found at
<http://www.hf.uijib.no/cybertext/>

It contains links to many of the texts and computer programs discussed,
as well as pointers to other relevant resources.

Literature is a combinatorial game that pursues the possibilities implicit in its own material, independent of the personality of the poet, but it is a game that at a certain point is invested with an unexpected meaning, a meaning that is not patent on the linguistic plane on which we were working but has slipped in from another level, activating something that on that second level is of great concern to the author or his society. The literature machine can perform all the permutations possible on a given material, but the poetic result will be the particular effect of one of these permutations on a man endowed with a consciousness and an unconscious, that is, an empirical and historical man. It will be the shock that occurs only if the writing machine is surrounded by the hidden ghosts of the individual and his society.

I T A L O C A L V I N O

<i>One</i>	Introduction: Ergodic Literature	1
	The Book and the Labyrinth 1 / Some Examples of Ergodic Literature 9 / The Aim of This Study 13 / What Is Cybertext? 17	
<i>Two</i>	Paradigms and Perspectives	24
	Problems in Computer Semiotics 24 / Textuality, Nonlinearity, and Interactivity 41 / Cyborg Aesthetics and the "Work in Movement" 51	
<i>Three</i>	Textonomy: A Typology of Textual Communication	58
	Previous Efforts 59 / Method 60 / The Typology 62 / The Texts 65 / Analysis and Results 67 / Conclusions 73	
<i>Four</i>	No Sense of an Ending: Hypertext Aesthetics	76
	Paradigms of Hypertext 82 / The Sense of a Novel: Michael Joyce's <i>Afternoon</i> 86 / The Rhetoric of Hyperliterature: Aporia and Epiphany 90 / The Poetics of Conflict: Ergodics versus Narration 92 / Transclusions 95	
<i>Five</i>	Intrigue and Discourse in the Adventure Game	97
	A Brief History of the Genre 97 / A Schematic Model of Internal Structure 103 / Some Issues in Adventure Game Criticism 106 / Intrigue, Intrigant, Intriguee 111 / The Autistic Detective Agency: Marc Blank's <i>Deadline</i> 115 / Intrigue and Discourse 124 / The End of Story? 127	

Six

- The Cyborg Author: Problems of Automated Poetics 129
- The Case of Racter 132 / A Typology of Authors in the Machine-Human Continuum 134 / Laurel's Playwright: Seducing the User 136 / From Author Simulacrum to Cybertext 141

Seven

- Songs from the MUD: Multiuser Discourse 142
- Literature in the MUD? 142 / A Historical Perspective on MUDs and Nonlocal Communication 149 / The Aesthetics of Nonlocal Discourse 152 / Netiquette and Discourse 158

Eight

- Ruling the Reader: The Politics of "Interaction" 162
- The Death (and Politics) of the Reader 162 / Democracy in Cybermedia 167 / Levels of Usership 173

Nine

- Conclusions: The Ideology of Influence 178
- Anamorphosis versus Metamorphosis 178 / Toward Theories of Ergodic Literature 182

References 185 / Index 197

This work was made possible by a three-year Ph.D. scholarship from the Norwegian Research Council. I wish to thank the Faculty of Arts, University of Bergen, for their generosity in providing me with basic equipment crucial to my project, when they were under no obligation to do so.

A large number of people in various countries have contributed invaluable help and inspiration during my years of research. By naming none, I hope to include all. I am indebted to my supervisors, Aste Kittang and Richard Holton Pierce, for their critical and inspiring support. I am also very grateful to my colleagues and friends at the Humanities Computing Section, in particular Roald Skarsten.

The Book and the Labyrinth

A few words on the two neoteric terms, *cybertext* and *ergodic*, are in order. *Cybertext* is a neologism derived from Norbert Wiener's book (and discipline) called *Cybernetics*, and subtitled *Control and Communication in the Animal and the Machine* (1948). Wiener laid an important foundation for the development of digital computers, but his scope is not limited to the mechanical world of transistors and, later, of microchips. As the subtitle indicates, Wiener's perspective includes both organic and inorganic systems; that is, any system that contains an information feedback loop. Likewise, the concept of cybertext does not limit itself to the study of computer-driven (or "electronic") textuality; that would be an arbitrary and unhistorical limitation, perhaps comparable to a study of literature that would only acknowledge texts in paper-printed form. While there might be sociological reasons for such a study, we would not be able to claim any understanding of how different forms of literature vary.

The concept of cybertext focuses on the mechanical organization of the text, by positing the intricacies of the medium as an integral part of the literary exchange. However, it also centers attention on the consumer, or user, of the text, as a more integrated figure than even reader-response theorists would claim. The performance of their reader takes place all in his head, while the user of cybertext also performs in an extraoemetic sense. During the cybertextual process, the user will have effectuated a semiotic sequence, and this selective movement is a work of physical construction that the various concepts of "reading" do not account for. This phenomenon I call *ergodic*, using a term appropriated from physics that derives from the Greek words *ergon* and *hodos*, meaning "work" and "path." In ergodic literature, nontrivial effort is required to allow the reader to traverse the text. If ergodic literature is to make sense as a concept, there must also be nonergodic literature, where the effort to traverse the text is trivial, with no extraoemetic responsibilities placed on

2 Cybertext

the reader except (for example) eye movement and the periodic or arbitrary turning of pages.

Whenever I have had the opportunity to present the perspective of ergodic literature and cybertext to a fresh audience of literary critics and theorists, I have almost invariably been challenged on the same issues: that these texts (hypertexts, adventure games, etc.) aren't essentially different from other literary texts, because (1) all literature is to some extent indeterminate, nonlinear, and different for every reading, (2) the reader has to make choices in order to make sense of the text, and finally (3) a text cannot really be nonlinear because the reader can read it only one sequence at a time, anyway.

Typically, these objections came from persons who, while well versed in literary theory, had no firsthand experience of the hypertexts, adventure games, or multi-user dungeons I was talking about. At first, therefore, I thought this was simply a didactical problem: if only I could present examples of my material more clearly, everything would become indisputable. After all, can a person who has never seen a movie be expected to understand the unique characteristics of that medium? A text such as the *I Ching* is not meant to be read from beginning to end but entails a very different and highly specialized ritual of perusal, and the text in a multi-user dungeon is without either beginning or end, an endless labyrinthine plateau of textual bliss for the community that builds it. But no matter how hard I try to describe these texts to you, the reader, their essential difference will remain a mystery until they are experienced firsthand.

In my campaign for the study of cybertextuality I soon realized that my terminology was a potential source of confusion. Particularly problematic was the word *nonlinear*. For some it was a common literary concept used to describe narratives that lacked or subverted a straightforward story line; for others, paradoxically, the word could not describe my material, since the act of reading must take place sequentially, word for word.

This aporia never ceased to puzzle me. There was obviously an epistemological conflict. Part of the problem is easily resolved: hypertexts, adventure games, and so forth are not texts the way the

average literary work is a text. In what way, then, are they texts? They produce verbal structures, for aesthetic effect. This makes them similar to other literary phenomena. But they are also something more, and it is this added paraverbal dimension that is so hard to see. A cybertext is a machine for the production of variety of expression. Since literary theorists are trained to uncover literary ambivalence in texts with linear expression, they evidently mistook texts with variable expression for texts with ambiguous meaning. When confronted with a forking text such as a hypertext, they claimed that all texts are produced as a linear sequence during reading, so where was my problem?

The problem was that, while they focused on what was being read, I focused on what was being read *from*. This distinction is inconspicuous in a linear expression text, since when you read from *War and Peace*, you believe you are reading *War and Peace*. In drama, the relationship between a play and its (varying) performance is a hierarchical and explicit one; it makes trivial sense to distinguish between the two. In a cybertext, however, the distinction is crucial—and rather different; when you read from a cybertext, you are constantly reminded of inaccessible strategies and paths not taken, voices not heard. Each decision will make some parts of the text more, and others less, accessible, and you may never know the exact results of your choices; that is, exactly what you missed. This is very different from the ambiguities of a linear text. And inaccessibility, it must be noted, does not imply ambiguity but, rather, an absence of possibility—an aporia.

So why is this so difficult to see? Why is the variable expression of the nonlinear text so easily mistaken for the semantic ambiguity of the linear text? The answer, or at least one answer, can be found in a certain rhetorical model used by literary theory. I refer to the idea of a narrative text as a labyrinth, a game, or an imaginary world, in which the reader can explore at will, get lost, discover secret paths, play around, follow the rules, and so on. The problem with these powerful metaphors, when they begin to affect the critic's perspective and judgment, is that they enable a systematic misrepresentation of the relationship between narrative text and reader; a spatiodynamic fallacy where the narrative is not perceived as a pre-

4 Cybertext

sentation of a world but rather as that world itself. In other words, there is a short circuit between signifier and signified, a suspension of *différance* that projects an objective level beyond the text, a primary metaphysical structure that generates both textual sign and our understanding of it, rather than the other way around.

A reader, however strongly engaged in the unfolding of a narrative, is powerless. Like a spectator at a soccer game, he may speculate, conjecture, extrapolate, even shout abuse, but he is not a player. Like a passenger on a train, he can study and interpret the shifting landscape, he may rest his eyes wherever he pleases, even release the emergency brake and step off, but he is not free to move the tracks in a different direction. He cannot have the player's pleasure of influence: "Let's see what happens when I do *this*." The reader's pleasure is the pleasure of the voyeur. Safe, but impotent.

The cybertext reader, on the other hand, is not safe, and therefore, it can be argued, she is not a reader. The cybertext puts its would-be reader at risk: the risk of rejection. The effort and energy demanded by the cybertext of its reader raise the stakes of interpretation to those of intervention. Trying to know a cybertext is an investment of personal improvisation that can result in either intimacy or failure. The tensions at work in a cybertext, while not incompatible with those of narrative desire, are also something more: a struggle not merely for interpretative insight but also for narrative control: "I want this text to tell *my* story; the story that *could* not be without me." In some cases this is literally true. In other cases, perhaps most, the sense of individual outcome is illusory, but nevertheless the aspect of coercion and manipulation is real.

The study of cybertexts reveals the misprision of the spatio-dynamic metaphors of narrative theory, because ergodic literature incarnates these models in a way linear text narratives do not. This may be hard to understand for the traditional literary critic who cannot perceive the difference between metaphorical structure and logical structure, but it is essential. The cybertext reader is a player, a gambler; the cybertext is a game-world or world-game; it is possible to explore, get lost, and discover secret paths in these texts, not metaphorically, but through the topological structures of the textual machinery. This is not a difference between games and literature but

rather between games and narratives. To claim that there is no difference between games and narratives is to ignore essential qualities of both categories. And yet, as this study tries to show, the difference is not clear-cut, and there is significant overlap between the two.

It is also essential to recognize that *cybertext* is used here to describe a broad textual media category. It is not in itself a literary genre of any kind. Cybertexts share a principle of calculated production, but beyond that there is no obvious unity of aesthetics, thematics, literary history, or even material technology. Cybertext is a perspective I use to describe and explore the communicational strategies of dynamic texts. To look for traditions, literary genres, and common aesthetics, we must inspect the texts at a much more local level, and I suggest one way to partition the field in chapters 4 through 7, each chapter dealing with a subgroup of ergodic textuality.

Even if the cybertexts are not narrative texts but other forms of literature governed by a different set of rules, they retain to a lesser or greater extent some aspects of narrative. Most display some forms of narrative behavior, just as can be found in other non-narrative literary genres. The idea of pure literary forms or discrete genres is not pursued here. Instead, a perspective of complementary generic traits is used to describe the various types as synthetic, composite genres. Perhaps, by studying cybertexts and trying to discover this alterity of narrative, we may also get some small new clues as to what narrative is.

It seems to me that the cybertexts fit the game-world-labyrinth terminology in a way that exposes its deficiencies when used on narrative texts. But how has the spatiodynamic misrepresentation of narrative originated? And was it always inappropriate? An important clue to this question can be found in the historical idea of the labyrinth. Our present idea of the labyrinth is the Borgesian structure of "forking paths," the bewildering chaos of passages that lead in many directions but never directly to our desired goal. But there is also another kind, or paradigm, of labyrinths. Penelope Reed Doob, in her excellent discussion of physical and metaphorical labyrinths of classical antiquity and the Middle Ages (1990), distinguishes between two kinds of labyrinthine structure: the unicursal,

where there is only one path, winding and turning, usually toward a center; and the multicursal, where the maze wanderer faces a series of critical choices, or *bivia*.

Umberto Eco (1984, 80) claims that there are three types of labyrinth: the linear, the maze, and the net (or rhizome; cf. Deleuze and Guattari 1987). The first two correspond to Doob's unicursal and multicursal, respectively. To include the net seems inappropriate, since this structure has very different qualities from the other two. Especially as the net's "every point can be connected with every other point" (Eco 1984, 81); this is exactly the opposite of the fundamental inaccessibility of the other models. Amazingly, Eco also claims that the labyrinth of Crete was linear and that Theseus "had no choices to make: he could not but reach the center, and from the center, the way out. . . . In this kind of labyrinth the Ariadne thread is useless, since one *cannot* get lost" (80). It is hard to believe that Eco is speaking of the labyrinth where Theseus, famously, was the first to find the way out, and only because of Ariadne's thread. This was the same complex labyrinth where even its maker, Daedalus, was lost. Doob (1990, 17-38), on the other hand, citing Pliny, Virgil, Ovid, and others, shows that the literary tradition describes the *Domus daedali* as a multicursal labyrinth.

As Doob demonstrates, the labyrinth as a sign of complex artistry, inextricability, and difficult process was an important metaphor and motif in classical and medieval literature, philosophy, rhetoric, and visual design. Paradoxically, while the labyrinth depicted in visual art from prehistoric times is always unicursal, the literary maze (with the Cretan myth as the chief example) is usually multicursal. The multicursal motif did not appear in art until the Renaissance, but as Doob shows, the two paradigms coexisted peacefully as the same concept at least since Virgil (70-19 B.C.). In Doob's view, what to us seem to be contradictory models were subsumed in a single category, signifying a complex design, artistic order and chaos (depending on point of view), inextricability or impenetrability, and the difficult progress from confusion to perception. Both models share these essential qualities of the labyrinth, and apparently there was no great need to distinguish between the two.

In the Renaissance, however, the idea of the labyrinth, both in

literature and visual art, was reduced to the multicursal paradigm that we recognize today. Consequently, the old metaphor of the text as labyrinth, which in medieval poetics could signify both a difficult, winding, but potentially rewarding linear process and a spatial, artistically complex, and confusing artifact, was restricted to the latter sense. Therefore, I find it reasonable to assume that the image of the text as a labyrinth has undergone an ideological transformation, from a harmonic duality where the figurative likeness of the narrative text as unicursal coexisted with a tropology of multicursal aspects, such as repetition, interlaced narrative threads, prolepsis, and so forth. When the unicursal paradigm faded, however, the multicursal paradigm came to dominate the figure, devolving the rich ambiguity of the classical and medieval labyrinth into the less ambiguous Renaissance model of pure multicursality.

Since we now regard *labyrinthine* and *linear* as incompatible terms, and since the labyrinth no longer denotes linear progress and teleology but only their opposites, its status as a model of narrative text has become inapt for most narratives. For a typical example of this misnomer, consider the following, from a discussion of postmodernist writing: "We shall never be able to unravel the plots of John Fowles's *The Magus* (1966), Alain Robbe-Grillet's *Le Voyeur* (1955) or Thomas Pynchon's *The Crying of Lot 49* (1966), for they are *labyrinths without exits*" (Lodge 1977, 266; last italics mine). Here, the image of the labyrinth has become severely distorted. A labyrinth without exit is a labyrinth without entrance; in other words, not a labyrinth at all.

Even in highly subversive narratives, such as the novels of Samuel Beckett or Italo Calvino's *If on a Winter's Night a Traveler* . . . (1993), the reader is faced, topologically, with a unicursal maze. Yet there are some novels for which the post-Renaissance model is perfectly valid, for instance Julio Cortázar's *Rayuela* (1966), in which the topology is multicursal. In yet others, such as Vladimir Nabokov's *Pale Fire* (1962), it may be described as both unicursal and multicursal.

The footnote is a typical example of a structure that can be seen as both uni- and multicursal. It creates a bivium, or choice of expansion, but should we decide to take this path (reading the foot-

8 Cybertext

note), the footnote itself returns us to the main track immediately afterward. Perhaps a footnoted text can be described as multicursal on the micro level and unicursal on the macro level. Nabokov's *Pale Fire*, however, leaves the mode of cursality up to the reader; consisting of a foreword, a 999-line poem, a long commentary of notes addressing individual lines (but really telling the commentator's story), and an index, it can be read either unicursally, straight through, or multicursally, by jumping between the comments and the poem. Brian McHale (1987, 18–19) sees it as a limit-text between modernism and postmodernism; it is also a limit-text between uni- and multicursality.

That some texts are hard to define topologically should not surprise us, as it is exactly this aspect of their own ontology they set out to destabilize (cf. McHale 1987, chap. 12). Neither should it discourage us, since the existence of borderline cases and ambiguous structures in no way invalidates the usefulness of categories such as narrative and game or unicursality and multicursality.

The problem is not, finally, that literary critics use words like *labyrinth*, *game*, and *world* as metaphors in their analyses of unicursal works but that this rhetoric seems to have blinded them to the existence of multicursal literary structures and to the possibility that the concept of labyrinth (in their post-Renaissance rendition) might have more analytic accuracy in connection with texts that function as game-worlds or labyrinths in a literal sense. However, this is not the place to criticize in detail the ontological problems resulting from a possible flaw in the terminology of narrative theory. Such an issue deserves at least a separate study, one not focused on the texts that are our primary concern here. Instead, this might be the place for suggesting the reinstatement of the old dual meaning of *labyrinth*, so that both unicursal and multicursal texts might be examined within the same theoretical framework. With such a theory we might be able to see both how, in Jorge Luis Borges's words, "the book and the labyrinth [are] one and the same" (Borges 1974, 88), and how the many types of literary labyrinths are different from each other. It may surprise some readers to find me still using the word *book*, but a number of the cybertexts we shall discuss are indeed books—printed, bound, and sold in the most traditional

fashion. As we shall see, the codex format is one of the most flexible and powerful information tools yet invented, with a capacity for change that is probably not exhausted yet, and I (for one) do not expect it to go out of style any time soon.

Some Examples of Ergodic Literature

At this point it is probably best to live the discussion with some examples of the literature I am primarily addressing. The exposition made here is mostly for elucidation purposes and must not be mistaken for an attempt to produce an exhaustive historical inventory of ergodic literature (see, instead, Vuillemin 1990). Rather than seeking a catalogue of every known instance of ergodicity, I have focused on diversity. As Roland Barthes (1977, 81) maintains in his study of narrative, it is utopian to examine every specimen of a genre; a deductive method, leading to a "hypothetical model of description," should be applied instead. Thus there may well exist major ergodic genres or texts that I have failed to include, but since this is a theoretical rather than an encyclopedic study, the future appearance of any hitherto unknown forms will invalidate my theories only if they fail to comply with my general model of ergodic forms.

Since writing always has been a spatial activity, it is reasonable to assume that ergodic textuality has been practiced as long as linear writing. For instance, the wall inscriptions of the temples in ancient Egypt were often connected two-dimensionally (on one wall) or three-dimensionally (from wall to wall and from room to room), and this layout allowed a nonlinear arrangement of the religious text in accordance with the symbolic architectural layout of the temple (Gundlach 1985).

Possibly the best-known example of cybertext in antiquity is the Chinese text of oracular wisdom, the *I Ching* (Wilhelm 1989). Also known as the *Book of Changes*, the existing text is from around the time of the Western Chou dynasty (1122–770 b.c.) and was written by several authors. The *I Ching* system also inspired G. W. von Leibniz, who developed the binary mathematics used by today's digital computers (Eber 1979). The *I Ching* is made up of sixty-four symbols, or hexagrams, which are the binary combinations of six whole or broken ("changing") lines ($64 = 2^6$). A hexagram (such as

no. 49, ≡ *Ko/Revolution*) contains a main text and six small ones, one for each line. By manipulating three coins or forty-nine yarrow stalks according to a randomizing principle, the texts of two hexagrams are combined, producing one out of 4,096 possible texts. This contains the answer to a question the user has written down in advance (e.g., "How much rice should I plant this year?").

Much simpler examples of nonlinear texts are some of Guillaume Apollinaire's "calligrammes" from early in this century (Apollinaire 1966). The words of these poems are spread out in several directions to form a picture on the page, with no clear sequence in which to be read. A play from the thirties, *Night of January 16th* by Ayn Rand (1936), is about a trial where members of the audience are picked to be the jury. The play has two endings, depending on the jury's verdict. In the early 1960s, Marc Saporta (1962) published *Composition No. 1, Roman*, a novel with pages like a deck of cards, to be shuffled and read in any sequence. It is written in such a way that any combination will appear fluid. (See also Bolter 1991, 140-42.)

A rather well-known example is Raymond Queneau's *Cent Mille Millions de Poèmes* (a hundred thousand billion poems; see Queneau 1961), which is a sonnet machine book of 10×14 lines, capable of producing 10^{14} sonnets. Several novels have been identified as ergodic over the years: B. S. Johnson's *The Urfortunates* (1969), Milorad Pavić's *Landscape Painted With Tea* (1990), and many others. The variety and ingenuity of devices used in these texts demonstrate that paper can hold its own against the computer as a technology of ergodic texts.

However, after the invention of digital computing in the middle of the twentieth century, it soon became clear that a new textual technology had arrived, potentially more flexible and powerful than any preceding medium. Digital systems for information storage and retrieval, popularly known as databases, signified new ways of using textual material. The database is in principle similar to the filing cabinet but with a level of automation and speed that made radically different textual practices possible. On the physical level, the surface of reading was divorced from the stored information. For the first time, this breaks down concepts such as "the text itself" into two independent technological levels: the interface and the storage

medium. On the social level, huge texts could be browsed, searched, and updated by several people at once, and from different places on the globe, operations that only superficially seem to resemble what we used to call "reading" and "writing." Armed with a good search engine and a digital library, any college dropout can pass for a learned scholar, quoting the classics without having read any of them.

Several new textual genres have emerged with digital computing and automation. Computer programs, complex lists of formal instructions written in specially designed, artificial languages, can be seen as a new type of the rhetorical figure *apostrophe*, the addressing of inanimate or abstract objects, with the magical difference that it actually provokes a response. Short, simple programs are often linear, but longer programs generally consist of collections of interdependent fragments, with repeating loops, cross-references, and discontinuous "jumps" back and forth between sections. Given the seminatural vocabulary of some modern programming languages, it is not uncommon for programmers to write poems in them, often with the constraint that the "poegrms" (or whatever) must make sense to the machine as well.¹

Programs are normally written with two kinds of receivers in mind: the machines and other programmers. This gives rise to a double standard of aesthetics, often in conflict: efficiency and clarity. Since speed is a major quality in computer aesthetics, an unreadable program might perform much faster than a comprehensible one. The poetics of computer program writing is constantly evolving, and through paradigms such as object orientation it inspires practical philosophies and provides hermeneutic models for organizing and understanding the world, both directly (through programmed systems) and indirectly (through the worldviews of computer engineers).

Through the artificial intelligence research of the sixties, programs emerged that one could "talk" to. The best known of these is

1. For an example of this type of poetry, not to be confused with computer-generated poetry, see Sharon Hopkins' poem "Listen" (Hopkins 1995), written in the computer-programming language Perl.

Eliza, made in 1963 by an MIT computer scientist, Joseph Weizenbaum. Eliza could imitate a Rogerian psychoanalyst, and through a simple pattern-matching algorithm, it used the information given by its human "clients" to make them believe that it somehow "understood" their situations. Another seminal program, and one of the key texts in this study, is the role-playing game *Adventure*, by William Crowther and Don Woods, released on the U.S. research network ARPANET, the precursor of the Internet, in April 1976.² As the microcomputer home market exploded around 1980, *Adventure* was made available on nearly every type of machine and became the first in a short-lived, but influential, textual computer game genre, which ended its commercial life when the graphic adventure games took over in the late eighties.

In the seventies, some artificial intelligence researchers focused on making systems that could analyze and write stories. A well-known project was James Meehan's program Tale-spin, which could construct simple animal fables of the *Æsop* type. Primarily, the researchers were not trying to achieve literary quality, and the stories that were produced typically testify to this lack of ambition. However, some of the "failures" produced by Tale-spin make strikingly original prose, succeeding where the successes failed. A later system, the commercial dialogue program Racter, created by William Chamberlain (1984), is even supposed to have written a book, *The Policeman's Beard Is Half Constructed*, but as it turns out, the book was co-written (at least) by Chamberlain (see Barger 1993 and chapter 6, below). Although the output of these generators are linear stories or poems, the systems themselves are clearly ergodic textual machines, with unlimited possibility for variation.

Another type of digital ergodic text was conceived by the American Ted Nelson around 1965 (Nelson 1965; see also Nelson 1987). Nelson called this *hypertext*, a strategy for organizing textual fragments in an intuitive and informal way, with "links" between related sections of a text or between related parts of different texts in

the same retrieval system. Hypertext has gained in popularity in the last decade, after personal computer programs such as Hypercard were made available and educators started to take an interest in its pedagogical potential. At the same time, literary authors started to experiment with hypertext and have received considerable attention from literary circles. Hyperfictions such as Michael Joyce's *Afternoon: A Story* (1990) engage a modernist poetics to subvert traditional storytelling and present a literary labyrinth for the reader to explore.

In 1980, inspired by William Crowther and Don Woods' *Adventure* (1976), two English programmers at the University of Essex, Roy Trubshaw and Richard Bartle, constructed an adventure game that several players could play at once (see Bartle and Trubshaw 1980; Bartle 1984). They called their invention *Multi-User Dungeon* (*MUD*), also known as *MUD1*, and soon participants from many parts of the world phoned in from their modems to the Essex computer to participate in the new social reality. The first *MUDs* were oriented toward game playing and puzzle solving, but later *MUDs*, such as James Aspnes's 1989 *TinyMUD*, allowed users to build their own textual objects and landscapes, and soon the users came to regard themselves as participants in a community, rather than a game, with communication rather than competition as the main social activity. As literature (although not as textual media), *MUDs* are very different from anything else, with their streams of continuing text and their collective, often anonymous readership and writership. Life in the *MUD* is literary, relying on purely textual strategies, and it therefore provides a unique laboratory for the study of textual self-expression and self-creation, themes that are far from marginal in the practice of literary theory.

The Aim of This Study

It is a common belief that the rapid evolution in the field of digital technology from the middle of the twentieth century to the present has (among other equally astounding results) brought on radically new ways of writing and reading. This view, stimulated by the increasing personal experience with computer technology among the academic masses, can be observed even in literary studies, which

2. Personal correspondence with Woods, by E-mail, dated September 29, 1993. I am grateful for his illuminating reply and for the fabulous computer network that makes the Homers of digital literary history still available to researchers.

since 1984 have increasingly attempted to capture and construct computer-mediated texts as objects of literary criticism. The present study can be located both inside and outside of this research. In addition to an analysis—and to some extent a construction—of the perceived objects by means of literary theory, this is a study of the problems of such construction and, hence, a critical study of the strategies used by literary researchers to expand their empirical field in this direction. Especially, I wish to challenge the recurrent practice of applying the theories of literary criticism to a new empirical field, seemingly without any reassessment of the terms and concepts involved. This lack of self-reflection places the research in direct danger of turning the vocabulary of literary theory into a set of unfocused metaphors, rendered useless by a translation that is not perceived as such by its very translators. Thus the interpretations and misinterpretations of the digital media by literary theorists is a recurrent theme of this book.

A related but reverse problem is the tendency to describe the new text media as radically different from the old, with attributes solely determined by the material technology of the medium. In these analyses, technical innovation is presented as a cause of social improvement and political and intellectual liberation, a historical move away from the old repressive media. This kind of technological determinism (the belief that technology is an autonomous force that causes social change) has been refuted eloquently by Langdon Winner (1986), James W. Carey (1988), and others but continues, nevertheless, to dominate the discussion. In the context of literature, this has led to claims that digital technology enables readers to become authors, or at least blurs the (supposedly political) distinction between the two, and that the reader is allowed to create his or her own "story" by "interacting" with "the computer." The ideological forces surrounding new technology produce a rhetoric of novelty, differentiation, and freedom that works to obscure the more profound structural kinships between superficially heterogeneous media. Even the inspiring and perceptive essays of Richard Lanham (1993) are suffused by this binary rhetoric and, ultimately, dominated by politics at the expense of analysis.

Whether concepts such as "computer literature" or "electronic

textuality" deserve to be defended theoretically is by no means obvious, and they will not be given axiomatic status in this book. The idea that "the computer" is in itself capable of producing social and historical change is a strangely ahistorical and anthropomorphic misconception, yet it is as popular within literary-cultural studies as it is in the science fiction texts they sometimes study. Often, in fact, science fiction portrays the technology with an irony that the critical studies lack (see, e.g., William Gibson's short story, "Burning Chrome," in Gibson 1986).

Most literary theories take their object medium as a given, in spite of the blatant historical differences between, for instance, oral and written literature. The written, or rather the printed, text has been the privileged form, and the potentially disruptive effects of media transitions have seldom been an issue, unlike semantic transitions such as language translation or intertextual practices. At this point, in the age of the dual ontology of everyday textuality (screen or paper), this ideological blindness is no longer possible, and so we have to ask an old question in a new context: What is a text? In a limited space such as this, it is impossible to recapture the arguments of previous discussions of this question. And since the empirical basis for this study is different from the one assumed in these discussions, the arguments would be of limited value. In the context of this study, the question of the text becomes a question of verbal media and their functional differences (what role does a medium play?), and only subsequently a question of semantics, influence, otherness, mental events, intentionality, and so forth. These philosophical problems have not left us, but they belong to a different level of textuality. In order to deal with these issues responsibly, we must first construct a map of the new area in which we want to study them, a *textonomy* (the study of textual media) to provide the playing ground of *textology* (the study of textual meaning).

The production of new maps, however, is also a construction of "newness," whose political consequences we cannot hope to escape. The field of literary study is in a state of permanent civil war with regard to what constitutes its valid objects. What right have we to report this war to foreign continents? Even if important insights can be gained from the study of extraliterary phenomena with the

instruments of literary theory (cautiously used), it does not follow that these phenomena are literature and should be judged with literary criteria or that the field of literature should be expanded to include them. In my view, there is nothing to be gained from this sort of theoretical imperialism, but much to lose: discussions of the "literariness" of this or that verbal medium are ever in danger of deteriorating into a battle of apologetic claims and chauvinistic counterclaims. When much energy is spent on showing that *P* is a perfectly deserving type of *Q*, the more fundamental question of what *P* is will often be neglected. These nonproductive (and non-academic) campaigns in favor of marginal media or aesthetic forms of expression are pathetic signs of a larger problem, however: they illustrate only too well the partial and conservative state of the human sciences, in which nothing can be studied that is not already within a field, in which the type rather than the individual qualities of an object determines its value as an accepted member of some canon or other. Where humanistic study used to be genre chauvinistic, it is now medium chauvinistic, organized into empirical fields (literature, art history, theater, mass communication) with not enough concern for general or intermediary perspectives. This "empirical" partitioning is of course unempirical in consequence, since it excludes empirical material that does not belong to the sanctioned sectors. Also, the limited view privileged by this sort of specializing tends to produce apologetics disguised as criticism, in an age where the "inherent" quality of literature (or any other previously dominant mode of discourse) is no longer self-evident.

Strangely, the struggle between the proponents and opponents of "digital literature" deteriorates usually on both sides into material arguments of a peculiar fetishist nature. One side focuses on the exotic hardware of the shiny new technologies, like CD-ROM. Witness especially the computer industry slogan, "information at your fingertips," as if information were somehow a touchable object. The other side focuses on the well-known hardware of the old technology, the "look and feel" of a book, compared to the crude letters on a computer screen. "You can't take it to bed with you" is the sensuous (but no longer true) refrain of the book chauvinists.

Isn't the content of a text more important than these materialistic, almost ergonomic, concerns?

What these strangely irrelevant exuberances reveal, I think, is that beyond the obvious differences of appearance, the real difference between paper texts and computer texts is not very clear. Does a difference even exist? Instead of searching for a structural divide, this study begins with the premise that no such essential difference is presumed. If it exists, it must be described in functional, rather than material or historical, terms. The alternative, to propose an essential difference and then proceed to describe it, does not allow for the possibility that it does not exist and is, therefore, not an option. Whether it exists or not is not of great importance to this thesis, however, as such knowledge would not make much practical difference in the world. The emerging new media technologies are not important in themselves, nor as alternatives to older media, but should be studied for what they can tell us about the principles and evolution of human communication.

My main effort is, therefore, to show what the functional differences and similarities among the various textual media imply about the theories and practices of literature. The exploration is based on the concepts and perspectives of narratology and rhetoric but is not limited to these two disciplines. I argue that existing literary theory is incomplete (but not irrelevant) when it comes to describing some of the phenomena studied here, and I try to show why and where a new theoretical approach is needed. My final aim is to produce a framework for a theory of cybertext or ergodic literature and to identify the key elements for this perspective.

What Is Cybertext?

In the current discussions of "computer literacy," hypertext, "electronic language," and so on, there seems to emerge an explicit distinction between the printed, or paper-based, text and the electronic text, both with singular and remarkably opposing qualities. The arguments for this distinction are sometimes historical, sometimes technological, but eminently political; that is, they don't focus on what these textual genres or modes are but on their assumed

functional difference from each other. Such a strategy is useful for drawing attention to, but less so for the analysis of, the objects thus constructed. It might have been tempting to follow this rhetoric in my investigation of the concept of cybertext and to describe a dichotomy between it and traditional, conventional literature; but the meaning of these concepts is unstable to the point of incoherence, and my construct would therefore probably have reached a similar degree of uselessness.

Cybertext, then, is not a "new," "revolutionary" form of text, with capabilities only made possible through the invention of the digital computer. Neither is it a radical break with old-fashioned textuality, although it would be easy to make it appear so. Cybertext is a *perspective* on all forms of textuality, a way to expand the scope of literary studies to include phenomena that today are perceived as outside of, or marginalized by, the field of literature—or even in opposition to it, for (as I make clear later) purely extraneous reasons. In this study I investigate the literary behavior of certain types of textual phenomena and try to construct a model of textual communication that will accommodate any type of text. This project is not as ambitious as it might sound, since the model is provisional and empirical and subject to future modification should any "falsificatory" evidence (such as an unpredictable object) appear. This pragmatic model is presented in detail in chapter 3.

The rest of this introductory chapter discusses the conceptual foundations and implications of this approach and establishes the terminology applied in the analytical chapters. These chapters (4 through 7) each takes on a main category (or genre) of cybertext roughly corresponding to the results of the analysis in chapter 3: hypertext, the textual adventure game, computer-generated narrative and participatory world-simulation systems, and the social-textual MUDs of the global computer networks. This pragmatic partitioning, which derives from popular convention rather than from my own theoretical model, is motivated by my strong belief that, in such a newly awakened field, theoretical restraint is imperative. Theories of literature have a powerful ability to co-opt new fields and fill theoretical vacuums, and in such a process of colonization, where the "virgin territory" lacks theoretical defense, important

perspectives and insights might be lost or at least overlooked. When we invade foreign ground, the least we can do is to try to learn the native language and study the local customs. Although several studies have already been carried out within most of these subfields, almost none have produced overarching, or universal, perspectives or engaged in a comparative analysis of all the forms of textuality examined here. Therefore, these previous approaches are discussed in their respective chapters rather than in this general introduction.

Because there are strong similarities between new and old types of ergodic literature, "the computer" and "information technology" as such will not be an explaining factor in this study but, instead, part of the field to be explored. This approach frees us from trying to define such vague and unfocused terms as *digital text* or *electronic literature* and allows us to develop a function-oriented perspective, in which the rhetoric of media chauvinisms will have minimal effect on the analysis. To be sure, media are far from neutral, inconsequential carriers of "content," but the essentialist idea of "the computer medium" as a singular structure of well-defined properties of communication is just as untenable and can be based on only a very limited understanding of both computer applications and media theory. Computer technology can sustain many different types of media, with very distinctive characteristics. Such a pluralist perspective will help us avoid the traps of technological determinism and let us see the technology as an ongoing process of, rather than a cause of, human expression. As we shall see, many of the forms of computer-based textuality have more in common with some of the paper media than with each other.

As can be inferred from its etymology, a cybertext must contain some kind of information feedback loop. In one sense, this holds true for any textual situation, granted that the "text" is something more than just marks upon a surface. A reader peruses a string of words, and depending on the reader's subsequent actions, the significance of those words may be changed, if only imperceptibly. The act of rereading is a crucial example: the second time we read a text, it is different, or so it seems. How can we know the text from the reading? Sometimes, a reader may influence the text for other readers, even if all the "marks on the pages" stay the same: a dramatic ex-

ample is the ayatollah Khomeiny's reaction to *The Satanic Verses*. The conventional split between text and reading (between the "intentional object" and the "mental event"), or *signifiant* and *signifié*, is not an impermeable membrane: leaks occur constantly, through various stages of reception such as editing, marketing, translation, criticism, rediscovery, canonization, or banishment.

These well-known processes are not entirely trivial, however, because they remind us that a text can never be reduced to a stand-alone sequence of words. There will always be context, convention, contamination; sociohistorical mediation in one form or another. Distinguishing between a text and its readings is not only necessary, it is also quite impossible—an ideal, in other words. On the one hand we need the image of "the text" in order to focus on anything at all; on the other hand we use the metaphor of "reading" to signal that our apprehension of a text will always be partial, that we never quite reach the "text itself," a realization that has led certain critics to question the very existence of such an object (see, for instance, Fish 1980). This hermeneutic movement or desire—perhaps better described as asymptotic than circular—holds true for all kinds of textual communication, but the particular organization of a text can make both the reader's strategic approach and the text's perceived teleology very distinctive, perhaps to the point where interpretation is stretched beyond the cognitive bounds of a singular concept. It is this field of varying textual organization that this study attempts to clarify. The differences in teleological orientation—the different ways in which the reader is invited to "complete" a text—and the texts' various self-manipulating devices are what the concept of cybertext is about. Until these practices are identified and examined, a significant part of the question of interpretation must go unanswered.

The meaning of text used in this study is closer to philological (or observable) work than to the poststructural (or metaphysical) galaxy of signifiers. But though my meaning is related to both of these meanings, it is also radically different from them. Instead of defining text as a chain of signifiers, as linguists and semioticians do, I use the word for a whole range of phenomena, from short poems

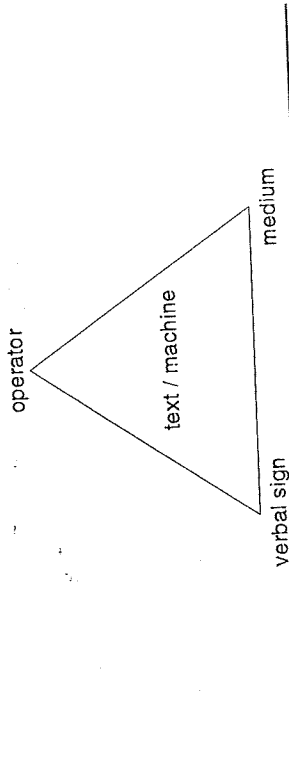


Figure 1.1. The Textual Machine

to complex computer programs and databases. As the *cyber* prefix indicates, the text is seen as a machine—not metaphorically but as a mechanical device for the production and consumption of verbal signs. Just as a film is useless without a projector and a screen, so a text must consist of a material medium as well as a collection of words. The machine, of course, is not complete without a third party, the (human) operator, and it is within this triad that the text takes place. (See figure 1.1.) The boundaries between these three elements are not clear but fluid and transgressive, and each part can be defined only in terms of the other two. Furthermore, the functional possibilities of each element combine with those of the two others to produce a large number of actual text types.

Previous models of textuality have not taken this performative aspect into account and tend to ignore the medium end of the triangle and all that goes with it. In his phenomenology of literature, Roman Ingarden (1973, 305–13) insists that the integrity of the "literary work of art" depends on the "order of sequence" of its parts; without this linear stability the work would not exist. While Ingarden here certainly acknowledges the importance of the objective shape of the text, he also reduces it to a given.

This taken-for-grantedness is hardly strange, since it is only after we have started to notice the "medium" and its recent shifting appearances that we can begin to observe the effect this instability has on the rest of the triangle. As Richard Lanham (1989, 270) observes, literary theorists have for a long time been in the "codex

book business," restricting their observations (but not their arguments) to literature mediated in a certain way. Even within the field of codex literature there is room, as experimentalists from Laurence Sterne to Milorad Pavic have demonstrated, for mediational variation, but these attempts have not, apparently, produced sufficient contrast to provoke a systematic investigation of the aesthetic role of the medium (a notable but much too brief exception being McHale 1987, chap. 12). There is also the fascinating phenomenon known as "Artists' Books," an art movement that originated in the sixties and dedicated to the creation of unique works of art that challenge the presumed properties of the book from within (cf. Strand 1992b and Lyons 1985).

Cybertext, as now should be clear, is the wide range (or perspective) of possible textualities seen as a typology of machines, as various kinds of literary communication systems where the functional differences among the mechanical parts play a defining role in determining the aesthetic process. Each type of text can be positioned in this multidimensional field according to its functional capabilities, as we shall see in chapter 3. As a theoretical perspective, cybertext shifts the focus from the traditional threesome of author/sender, text/message, and reader/receiver to the cybernetic intercourse between the various participant(s) in the textual machine. In doing so, it relocates attention to some traditionally remote parts of the textual galaxy, while leaving most of the luminous clusters in the central areas alone. This should not be seen as a call for a renegotiation of "literary" values, since most of the texts drawn attention to here are not well suited for entry into the competition for literary canonization.

The rules of that game could no doubt change, but the present work is not (consciously, at least) an effort to contribute to the hegemonic worship of "great texts." The reason for this is pragmatic rather than ethical: a search for traditional literary values in texts that are neither intended nor structured as literature will only obscure the unique aspects of these texts and transform a formal investigation into an apologetic crusade. If these texts redefine literature by expanding our notion of it—and I believe that they do—then they must also redefine what is literary, and therefore they

cannot be measured by an old, unmodified aesthetics. I do not believe it is possible to avoid the influence from literary theory's ordinary business, but we should at least try to be aware of its strong magnetic field as we approach the whiter spaces—the current final frontiers—of textuality.