

"Applying Hypertext Concepts to Print," Proceedings of the International Professional Communication Conference, Santa Fe, NM, September-October 1992.

Hypertext is non-linear text envisioned as a collection of independent modules ("nodes") and electronic pathways among these nodes known as "links" or "jumps." The different ways of arranging nodes and links are called "topologies." Hypertext theorists, designers, and empirical researchers have devoted much attention to the various topologies of nodes and links, to the experiences readers have navigating these topologies, and to such related concepts as modularity, pointers, and maps.

Can this effort be applied to print documents? The answer, I believe, is yes. This paper argues that we can identify a "topological perspective" on rhetoric and information design. Furthermore, this perspective can improve our understanding of both traditional and non-traditional print documents as well as electronic documents and can unify our study of print and electronic documents.

In a very useful article about teaching students the differences between creating hypertext and traditional documents, Henrietta Shirk (1991) calls for "a new and separate rhetoric of hypertext" (p. 189). On this one point I must disagree with Professor Shirk, for I am convinced that hypertext concepts are relevant to the study of paper as well as electronic documents. Hypertext concepts offer insights where traditional rhetoric is weak and provide a new perspective on traditional documents. At the same time, hypertext concepts help us understand the various kinds of modular documents that have been marginalized by our long-standing focus on the traditional sequential-hierarchical document model. Most important, hypertext concepts can unify our study of paper and electronic documents. They can, for example, reveal underlying similarities and differences among such diverse documents as Mark Bernstein's pioneering hypertext history *The Election of 1912* (1988); multimedia kiosks; online help facilities; newspapers; encyclopedias; and the full gamut of conventionally designed non-fiction books, technical reports, and other print documents.

The Dominance of the Sequential-Hierarchical Model

The design and logical structure of the printed book have changed little for at least the last few centuries. Authors of non-fiction very typically conceptualize and compose their documents in a linear reading sequence, with one division of the document serving as a prerequisite to the next. The body of the book is divided into chapters and in many cases is further divided into sections and subsections, and these divisions are represented in a table of contents and in a system of headings and running heads.

All of these devices, however, are more useful for previewing what lies ahead than for non-sequential reading, since the body of the text supports non-sequential reading haphazardly, if at all. Open a traditional book to Chapter 5, Section 2, and you may or may not be able to read productively. We have no conventions regarding this. The index supports non-sequential reading by directing the reader to the page or pages where a specific topic is discussed. But, again, there is no assurance whatsoever that the reader will understand the discussion well unless she has already read the book in its entirety or has a strong grasp of the subject matter. With only rare exceptions fiction writers follow the same general model, but dispense with the index and assume strictly linear reading.

Applying Hypertext Concepts to Print

DAVID K. FARKAS
(206) 685-8659

*Department of Technical
Communication
University of Washington
Seattle, WA 98195*

The various kinds of modular print documents are a major departure from the sequential-hierarchical model. One is the encyclopedia, a genre in which modules (in the form of independent articles) are arranged alphabetically and linked by cross references. Another is the collection of independent and yet closely related components such as meditations, poems on a single subject, or the suggestions offered in *50 Simple Things You Can Do to Save the Earth* (1989). Still another is the newspaper, a genre in which independent articles are presented together on the same page to compete, via headlines, for the reader's attention.

Because of the dominance of the traditional model, the communities of rhetoric as well as document design, literary study, composition studies, and technical communication have almost completely ignored modular documents. Furthermore, these same communities, though they have studied the traditional model intensely and in innumerable ways, have been largely blind to the dependency relationships among the units of information making up documents, to multiple pathways through documents, to access points, and to the conventions governing the hierarchical division of traditional documents and the representation of this hierarchy. In short, we have missed something that I will later define as the "topological" perspective on information design.

The Rise of Modular Documents

This topological perspective is now forcing itself upon us; modular documents, in both print and electronic form, are taking on unprecedented importance in human culture. There are at least three aspects to this transition.

One is the recent burgeoning of hypertext and multimedia. These technologies are rapidly maturing, large numbers of hypertext and multimedia presentations are being produced, and these presentations are steadily gaining a wider audience. Although fiction (including cinema) is still a bastion of the fully linear, hyperfiction is being written and studied.

Somewhat beyond the hypertext community, modular documents are also flourishing in the realm of computer documentation. With the transition to online documentation, the sequential-hierarchical user's guide is giving way to online help—most commonly a set of modular help topics richly interconnected via electronic pathways. Online help is bringing the modular document model to almost all computer users.

In the realm of popular reading, modular books are becoming ever more prevalent. Typical titles include: *50 Simple Things You Can Do To Save the Earth*, *All I Really Need to Know I Learned in Kindergarten* (1986), and *Elvis: His Life from A to Z* (1990). Richard Saul Wurman's controversial *Information Anxiety* (1989) is a modular book that explicitly challenges the sequential-hierarchical model. The popularity of modular books has much more to do with changes in our reading habits than the influence of hypertext but parallels changes in the world of electronic information. In the realm of scholarly book publishing, collections of separate contributions make up an ever-larger portion of the whole, while full-length monographs are becoming a less favored vehicle. In these collections, as in special issues of journals, it is common for the individual contributions to cross reference each other and be linked implicitly to one another by means of an editor's introduction, thus creating a kind of print hypertext.

One can argue that the shift to modular documents, in all or in some of its aspects, is harmful. But whether or not this is true, the rise of modular documents means that we need intellectual tools to better understand the structure of these documents.

Contributions of Hypertext Theory

Hypertext is a broad area that extends in certain directions which pertain exclusively to computer science or to electronic information. I am, however, arguing that significant portions of hypertext theory are highly relevant to print as well as online documents, to the concerns of traditional rhetoric, and even to fiction, most obviously fictions composed of interrelated modules such as *The Canterbury Tales* or Edgar Lee Master's *Spoon River Anthology*. Below I explain four fundamental hypertext concepts and point out their relevance to rhetoric and information design. Then, I offer a very brief "mini-analysis" of a non-fiction print document in an attempt to demonstrate the kinds of insights that hypertext-derived concepts can bring forth.

Modularity and Dependency. As we have seen, the dependency relationships among the chapters, sections, and other units of traditional non-fiction documents are haphazard and unpredictable. Furthermore, while many individual writers have no doubt grappled with this issue, it has not much concerned rhetoricians. In creating hypertext, however, the issues of modularity and dependency relationships are inevitably a central concern, for in hypertext the connections among the units of information must be explicitly established, requiring the designer to consider the success that different classes of readers will have as they navigate along the pathways the designer is building.

My own approach to the concept of modularity is to define several levels. In its most rigorous definition, modularity requires that any clearly bounded component (whether a "node" or "chapter" or "section") be fully understandable to any member of the intended audience regardless of when this individual accesses the component. Thus, an encyclopedia article on thermometers is meant to be fully understandable to the general reader, whether or not this reader has first consulted the article on mercury. A less rigorous definition allows that some topics may be only partially understandable if a reader from the unsophisticated part of the intended audience begins with a relatively specific or advanced component. A still more relaxed definition stipulates that the unsophisticated reader who jumps in "over her head" may not understand the component at all, but is at least directed to a well-chosen set of prerequisite components. Many help systems meet only this definition of modularity. Finally, if a document contains nodes or sections or chapters that can leave a member of the intended audience utterly bewildered and undirected, these components cannot be termed modular and the document cannot be termed truly modular at the level of granularity represented by these non-modular components.

Some may not look at modularity quite as I do, but this general kind of thinking necessarily governs the design of hypertext and, furthermore, can lead to better-designed sequential-hierarchical documents. Often, it will prove desirable for authors to plan out multiple reading pathways and to make them explicit to readers, whether in a preface or elsewhere. Textbooks do this regularly, because textbook publishers understand that teachers often assign only a subset of the chapters and have their own ideas about the sequence in which

topics should be taught. These prefaces, then, tell teachers which chapters the publisher deems essential and non-essential and which chapters must be read in sequence to be properly understood. Software manuals and especially third-party computer books often do much the same thing because it is recognized that users have diverse backgrounds and agendas and hence different information needs.

Pointers. Pointers are devices that direct or transport the reader from one part of a document to another. Tables of contents and indexes are collections of external pointers directing the reader from the front and back matter, respectively, into the body of the document. Cross references are internal pointers from one part of the body to another. So are ad-hoc forward and backward references such as "This problem will be discussed in more detail in Chapter 7." It seems fair to say that in strongly linear documents, there is a single implicit pointer extending from the first page to the last. Electronic documents, of course, have corresponding kinds of pointers. Hypertext jumps, for example, are the electronic equivalent of print cross references. In electronic documents, however, pointers are more than suggestions; they are the actual electronic pathways that readers must take to access a new node.

In traditional rhetoric pointers are beneath notice. But the number, type, placement, and distribution of pointers, especially when all the pointers in a document are considered from a systems perspective, constitute a significant design issue not only in electronic documents but in modular and sequential-hierarchical print documents as well.

Maps. Maps are representations of the content and structure of a document. A table of contents, in addition to being a set of pointers, is one kind of map. So are a book's system of headings and its running heads. Like pointers, maps have been largely ignored by rhetoricians. In contrast, maps are a major emphasis in hypertext, partly so because computer technology makes possible a wide range of design options including dynamic, interactive maps far more flexible and responsive to the reader's needs than print maps can ever be.

We have paid a price for ignoring print maps. Lack of interest has contributed much to the blind acceptance of a hoary assumption

| | |
|--------------------------------|-----|
| 5. Chemical Treatment | 123 |
| Special Hazards to Personnel | 134 |
| 6. Biological Treatment | 138 |

Map of a print document that violates traditional "tree" structure. This table of contents reveals that Chapter 5 has only a single subordinate section.

| |
|------------------------------|
| Formatting |
| <u>Formatting Characters</u> |
| <u>Formatting Paragraphs</u> |
| <u>Formatting Sections</u> |
| Outliner |
| <u>Using the Outliner</u> |

Map of an electronic document (the Contents screen of an online help system) that violates traditional "tree" structure. The heading "Outliner" has only a single subordinate hypertext jump.

(passed on from one generation of writing teachers to the next) about the structure of documents. Must the table of contents (or an outline of a document)—and hence its actual structure—always be a perfect “tree” in which each chapter or section is divided into two or more parts? Are arrangements such as those shown in the left column inherently dysfunctional? Note that the well-established convention of using sidebar “mini-articles” in magazines allows the writer to add a single subdivision to the main article. There are, in fact, various other respects in which a fresh hypertext-driven look at maps can improve our understanding and design of print documents. I would suggest, for instance, that the inclusion of both a high-level table of contents (“Contents at a Glance”) and a detailed table of contents—an increasingly prevalent design feature in the print world—is partly a response to the more flexible maps people are seeing in various kinds of electronic documents.

Information Topologies. Hypertext derives from computer science. Underlying hypertext, however, is the branch of mathematics called graph theory. Graph theory offers a system of rigorously defined topologies (or arrangements) for anything that can be conceived of as a network of nodes and links. For example, a network of electric-power generating plants and transmission lines can be described topologically. Topologies can also describe the arrangement of nodes and links in a hypertext or multimedia presentation and, in print documents, can describe the various arrangements of document components established by both the explicit pointers and the productive reading pathways.

William Horton (1990) has devised a non-rigorous, but highly practical taxonomy of the topologies commonly used by designers of online information. These include the web, the partial web, the sequence with shortcuts, the sequence with backtracking, the grid, and the hierarchy. Clearly, these topologies and other applications of graph theory can be applied to print documents as well. For example, because several different “sequence” topologies are used in multi-path children’s books (such as the very popular Choose Your Own Adventure series), the topological perspective is an excellent means of categorizing and analyzing the construction of these books.

The Topological Perspective and Wurman’s Information Anxiety

The four hypertext-derived concepts described above form at least the core of an analytic approach to both print and electronic documents. Because the concept of information topologies is the broadest of the four, we may aptly speak of a “topological perspective” on rhetoric and information design.

Let us see just what a very brief topological analysis of a work of print non-fiction might yield. Are there insights here? Do we see things that would not be uncovered by more traditional forms of analysis?

Richard Saul Wurman’s *Information Anxiety* has the physical appearance of a traditional sequential-hierarchical book. It is composed of chapters and sections and has a table of contents and running heads. Although the page design is often flamboyant, from the topological perspective its various maps are conventional.

But Wurman, departing from the traditional document model, insists that not only his individual chapters but the sections within them are completely independent and so meet the strictest definition of modularity. He says, “You can open to any chapter and read forward

or backward" (5). Analysis, however, shows that while the chapters are, in fact, strictly modular, significant context is lost in many instances if sections within a chapter are read in random order. Thus, at the level of granularity represented by the sections, the book is only modestly modular.

The book also features a large number of supplementary elements—quotations, statistics, brief narratives, observations, etc.—located in the margins. These elements are themselves strictly modular and while they are not true pointers, they bear an associative, web-like relationship to the main flow of text and to each other.

Conclusion

The topological perspective has major limitations; it is a narrow rhetoric that ignores style, affect, and even semantics, except that it asks whether readers can create meaning as they navigate from one part of a document to another and defines the structure of documents in terms of both navigable pathways and meaningful reading pathways. Put differently, the topological perspective makes clear that *Information Anxiety* is an unusual book and that it is not quite what Wurman says it is, but offers only hints as to whether it is a good book.

This topological perspective, as we have seen, is simply a subset of hypertext theory applied to print as well as to electronic information. Thus to accept the topological perspective is to accept the integration, not the separation, of hypertext concepts and the traditional rhetoric of print.

BERNSTEIN, Mark

The Election of 1912 (content application software)
Cambridge, MA: Eastgate Systems 1988

THE EARTHWORKS GROUP

50 Simple Things You Can Do to Save the Earth
Berkeley, CA: Earthworks Press 1989

FULGHUM, Robert

All I Really Need to Know I Learned in Kindergarten: Uncommon Thoughts on Common Things
New York: Ivy Books 1986

HORTON, William K.

Designing and Writing Online Documentation: Help Files to Hypertext
New York: Wiley 1990

SHIRK, Henrietta. R.

"Hyper" Rhetoric: Reflections on Teaching Hypertext
The Technical Writing Teacher, 18: 189-201 1991

WORTH, Fred L. and TAMERIUS, Steve

Elvis: His Life from A to Z
Chicago: Contemporary Books 1990