COMPUTERS AND BIOMEDICAL RESEARCH **25**, 238–263 (1992)

As We May Think: The Concept Space and Medical Hypertext*

JAMES J. CIMINO

Center for Medical Informatics, Columbia University

PETER L. ELKIN

Massachusetts Eye and Ear Infirmary

AND

G. OCTO BARNETT

Laboratory of Computer Science, Massachusetts General Hospital

Received April 18, 1991

Hypertext, a medium for presenting written material in a nonsequential manner, is gaining popularity as a format for medical text. The structure of "traditional" hypertext documents (hyperdocuments) includes author-created links among text segments. This structure poses challenge for those who create and maintain hyperdocuments, while reading them can introduce disorientation and cognitive overload. An alternative model is presented in which text segments are linked to the concepts which they contain and the concepts are linked to each other in a semantic network called the "Concept Space." The concepts and semantic links attempt to approximate potential topics of interest, allowing the reader to browse the hyperdocument in an individualized manner, rather than in an author-designated one. The *concept space* approach offers advantages for both the author and the reader. © 1992 Academic Press, Inc.

INTRODUCTION

The current "information explosion" makes thorough intimacy with all but the most limited topics beyond the capacity of individual humans. The rapid change and sheer volume of knowledge are impediments to accessing appropriate, up-to-date information in a timely manner; the practice of medicine is not immune to these difficulties (*3, 4, 5*). However, this flood of information is not a recent phenomenon.

* Portions of this work were presented at the *First Annual Educational and Research Conference of the American Medical Informatics Association* in 1990 and appear in abstract form in the conference proceedings (*1, 2*).

0010-4809/92 $5.00

In 1945, Vannevar Bush lamented the outmoded methods which were used then (and, for the most part, now) for disseminating knowledge. He anticipated many ways in which advancing technology would compound the problem by increasing the rate of growth in scholarly works, but he also proposed a solution. He described a machine, dubbed "memex," which would use microfilm for information storage and mechanical methods for retrieval. In essence, text and pictures would be linked to each other in ways that deviated from the traditional sequential structure of documents. Users would retrieve one document and then evoke related documents based on links which were included in the initial document. They would also be able to create their own links, forming their own trails of information retrieval which could be retraced at a later date or shared with others. Bush proposed that such trails would provide information in a manner "as we may think" (*6*). Memex was never built, but its functional description is generally credited as the inspiration for a new paradigm of information retrieval, now known as "hypertext" (*7, 8*).

Over the past two decades, the medium of scholarly information has evolved from its exclusively physical, printed form. At first, there were bibliographic retrieval services which provided references, and perhaps abstracts, on-line (*9*). More recently, CD-ROM databases (*10*) and on-line full-text retrieval have become available, providing virtually limitless access to medical information (*11*). In all of these, the format of the text remains essentially unchanged from that of the printed form (except for the omission, in most media, of tables, graphs, and illustrations). During this same era, researchers inspired by Bush's engine began to explore ways of implementing nonsequential retrieval of electronic textual data. In 1963, Engelbart described a model for using the computer to augment man's ability to manipulate symbols used to represent intellectual concepts (*12*) and by 1968 an actual implementation of this model appeared, called NLS, which allowed storage and retrieval of a variety of structured documents and provided the ability to link individual segments (called "statements") between documents. The links could then be traversed in ways unrelated to the original text sequence or hierarchy (*13*). Nelson coined the term "hypertext" to describe documents which could be read by following references from one point in the text to another, and his Xanadu project substantially advanced the mechanisms and applications of hypertext (*14, 15*). For details of these and other projects, the reader is referred to an excellent review by Conklin (*7*). The fruits of hypertext research are evidenced by the hypertext systems which are now commercially available (*16*).

The solution offered by hypertext to the information explosion problem is twofold. First, electronic storage of text permits large libraries to be readily available for retrieval. Second, the links between and within documents provide a means for supplying the reader with access to relevant information. Several innovative hypertext models have already been applied to medical texts (*17, 18*).

Unfortunately, the types of links and indexes for hypertext documents (or hyperdocuments) and the methods by which they are created have received scant attention. Bush anticipated that document trails would be created manually and made commercially available. Nelson also relegated the task to enterprising companies (*19*). But relatively little work has been directed at the standardization of link types and well-defined methods for their creation. (A notable exception is the work of Frisse, in which a hyperdocument was automatically created from "The Washington Manual of Medical Therapeutics," a text which is itself highly structured (*18*).) Difficulties with hypertext links exist, both for authors (*20*) and readers (*14*). This paper analyzes those difficulties and proposes a new model for the organization of hypertext which, like Bush's memex, makes use of a model of human information retrieval. This model is referred to as "Concept Space."

"Traditional" Hypertext

For every hypertext project and product, there is a different set of functional specifications for what constitutes a valid hypertext system. In part, this is due to various applications of hypertext, such as interactive teaching (*21*), cooperative authoring (*22*), and global publishing (*19*), to name a few. This paper considers one particular application of hypertext: information retrieval and browsing. Common commercial applications of hypertext (such as NoteCards, HyperCard, Guide, ToolBook, and Folios) allow the creation of hyperdocuments and navigational tools for browsing them. The underlying paradigm in all of these systems is that text exists in "chunks" (often of a size which will fit on a screen or in a window on a screen) and links between the chunks. The chunks, then, can be thought of as nodes in a graph. One system that makes use of this model is CASPER (*20*). Although it includes much more than static hyperdocuments, CASPER uses a model for information retrieval and browsing as shown in Fig. 1 that illustrates nicely the arrangement of a traditional hyperdocument.

Links between chunks may be unidirectional or bidirectional and may originate and terminate at text chunks or at specific points within the text. The links may be of many different types, each of which expresses some relationship between the chunks. One common type of relationship is the hierarchical link, which reflects the structure familiar in traditional paper documents. Other links can act as footnotes, providing paths to reference material. The textual content of the hyperdocument can be created with the system's word processor or may be imported from other sources. Hierarchical links may be created automatically by the system, while other links are created manually by the author or editor of the document. It is after the document is created that the hypertext features of the system come into play. The user of a browsing system can access information through static paths (explicitly provided by the author) or through dynamic ones (such as text searches) (*23*). The typical features of such a system (defined below) are: hierarchical browsing, index browsing, nonsequential browsing, text searching, filtering, and paths.

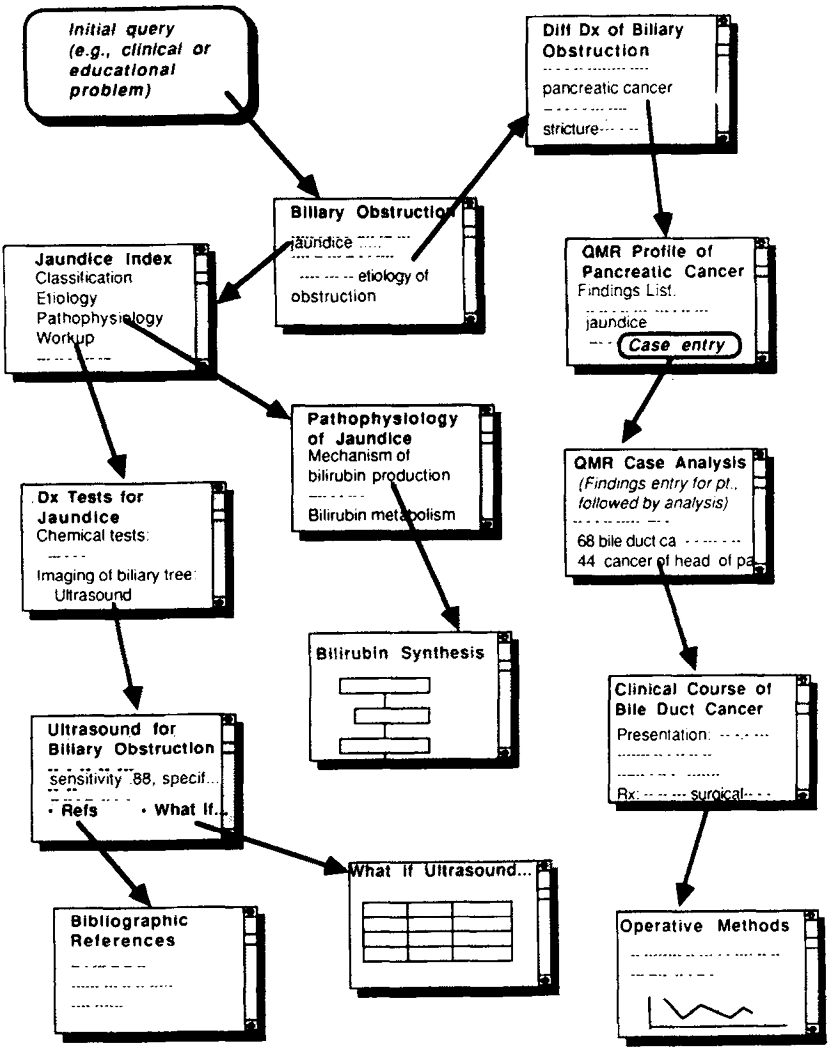

FIG. 1. Sample relationships between components of CASPER. This diagram shows examples of hotspots, buttons, and an index which are typical of those found in a traditional hypertext document. The CASPER system provides additional features which build upon the traditional model (from Greenes *et al.* (*20*)).

Hierarchical Browsing

Most hypertext systems permit (and some require) a hierarchical document structure (library, book, chapter, subchapter, etc.). When hypertext systems do not provide for this structure, the authors will often construct it, using links (*14*). This hierarchical structure provides a retrieval and browsing method familiar to all readers of conventional literature: selection of a book or journal, perusing the table of contents for a chapter or article of interest, and perhaps selecting a section of the chapter (e.g., the introduction) or an article (such as the abstract) to read. For example, a reader wishing to read about syncope might select a "book" on cardiovascular diseases from the hyperdocument and locate the chapter on syncope. The reader would then find a subchapter of interest or might read the entire chapter in the traditional sequential manner by traversing the hierarchical structure of the chapter in a depth-first manner. Note that, in order to reliably find information through hierarchical browsing, the reader's task is easier if he or she already has some knowledge of the subject (e.g., by knowing enough not to select a book on psychology or gastroenterology).

Index Browsing

The index used by hypertext systems is completely analogous to that found in conventional textbooks or, for that matter, the index of any database: a list of potentially interesting terms with pointers to text where each term is discussed. A reader simply turns to the index, locates the desired term ("jaundice," for example, as shown on the left-hand side of Fig. 1) and is presented with a list of links to text; selecting a link causes the associated text to be displayed. The index is created by the hypertext author and is static.

Nonsequential Link Browsing

The essential part of all hypertext systems is the ability to traverse nonsequential links which the author includes with, or the editor adds to, the text. They generally appear in two forms, which are typically referred to as "hotspots" and "buttons." Hotspots are highlighted words or phrases appearing within the text itself, while buttons are labeled icons that appear on the same screen as text but are spatially distinct from it. Beyond their placement in relation to the text, hotspots and buttons serve the same fundamental purpose. By selecting a hotspot or button (usually through some pointing device, such as a mouse), the reader will obtain additional information which the author has linked to the current text. For example, while reading a block of text about the differential diagnosis of biliary obstruction (such as in the upper right-hand corner of Fig. 1), the phrase "pancreatic cancer" appears as a highlighted term. By selecting it, the reader is shown a "findings list" for this disease. At the bottom of the list is a button, labeled "Case entry" which provides a link to another screen (in this case, an interactive decision support system). Hot spots and buttons are typically created and linked to related text in a manual fashion by the hypertext author or editor. Note that, for the author to provide relevant information, he or she must anticipate not only *what* the reader might be interested in but *why*.

Text Searching

Besides static, author-created links, some hypertext systems provide for free text or keyword searching. The reader enters a term of interest and the system generates a set of links to documents where the term appears, either in the text

1....precise clinical definition of **syncope** is sudden...
2....must be distinguished from true **syncope** on the basis...
3....arrhythmias commonly associated with **syncope** are...
4....it is not associated with **syncope**. The manifestations...
5....in vasovagal **syncope**, the precipitating factor...
6....drugs which cause **syncope** include vasodilating agents...
7....differential diagnosis of **syncope** includes a wide variety...
8....use of this drug is discussed in the **syncope** section...
9....most common causes of **syncope** in the elderly are...
10...treatment of **syncope** is directed at the underlying...

FIG. 2. Filtering. The results of a hypothetical search for the word "syncope" in a hyperdocument are shown as a list of "filtered" references, with only the text containing the search term being shown. The reader exercises further filtering by examining only those references which, from the snippets shown, appear to be most worthwhile. For example, a reader interested in causes of syncope would be unlikely to select references 1 or 2, and no reader would be likely to select reference 4 or 8.

or in an author-defined keyword list. For example, if a system has no index, the reader might search for "syncope" in the text. With this type of retrieval, the reader would need to be aware that the search term might appear in the text in many forms (such as "syncopal," "fainting," "loss of consciousness," etc.). Some systems permit wild-card and Boolean searches to improve the sensitivity and specificity of the retrievals (*24*). With such a system the reader might request all text which includes any word beginning with "syncop" or containing the word "fainting" or containing the word pair "loss" and "consciousness." Although the precise format for such a query will differ from system to system, a typical form might be: "SYNCOP$ or FAINTING or (LOSS and CONSCIOUSNESS)."

Filtering

By its very nature, hypertext can often present the reader with a large and potentially overwhelming set of links. This may be due to thoroughness on the part of the author-editor, but it can be compounded by hierarchical links, indexes, and text searches. A commonly employed method for assisting the reader in narrowing the furnished choices to those which are relevant to the reader's interest is to show a portion of the text which will be found by traversing the link. In the case of text searches, the context of the search term is displayed. If the search for syncope identified 10 references in the hyperdocument, the reader might select the best ones by use of a display such as is shown in Fig. 2. Because only the text which actually contains the term is displayed and because it allows the reader to screen candidate text for appropriateness, this technique is known as "filtering" (*7*).

Paths

Most hypertext systems allow for more than the simple display of additional information. Readers can traverse the nonsequential links so that the additional text becomes the new center of attention. Each chunk of text can proffer its own hot spots and buttons which the reader may then choose to traverse in an iterative or recursive manner. Figure 1 illustrates some sequences that a user might choose through a small subset of the CASPER system. Many hypertext systems maintain a history of the user's traversals in what is known as a path or web (7). As might be expected, the path can quickly grow long and meandering, taking the reader far from the original topic of interest. Since it may be possible to reach a particular block of text through multiple links and then traverse multiple links to other blocks of text, the potential paths are innumerable. While the individual links are static, the aggregation of links that form a user's path may not be anticipated or intended by the author.

Problems with Editing "Traditional" Hypertext

Part of the appeal of hypertext is that it can provide rapid access to large amounts of information. For a system to accomplish this, large amounts of text must be written for, or large numbers of text resources must be loaded into, the hypertext environment. Since multiple authors, styles, and formats are typically involved, editing and organizing the material into a coherent collection can be laborious. Providing hierarchical links is usually not difficult, since the source documents are typically structured hierarchically; however, problems can occur in building the index, creating the nonhierarchical links, and maintaining those links.

Building the Index

A good index to any substantial document (such as a text book) is an extremely valuable adjunct to the text itself. When properly created, it provides a good controlled vocabulary of the topics discussed in the document and includes relevant synonyms. It must also provide references to locations in the document where the topics are discussed (even if not explicitly mentioned), while omitting references to locations where the topics are mentioned explicitly but not discussed. Creating an index for a single piece of text is difficult enough; the problem is compounded when an index is needed which unifies several different documents, since synonymy must be recognized. For example, if one document discusses "syncope" while another describes "fainting," it is possible that a common index might include both terms without taking their synonymy into account. Thus, when wishing to find text on this disorder, a reader would need to find both index entries in order to retrieve all the relevant material. Hypertext authoring systems often provide an indexing facility, but the result is usually little more than a concordance of the unique words appearing in the text.

Creating Nonhierarchical Links

The nonhierarchical link is the essence of hypertext but its creation is labor-intensive, since it is done manually (*20*). The author/editor must decide which points in the text are appropriate hot spots and predict the reader's intent when selecting a hotspot, so that the appropriate text can be linked. If the author/editor anticipates a potential need for additional information which cannot be linked with hotspots (because the text does not include an explicit reference), buttons can be added to complement the text and provide pathways to appropriate text.

Maintaining Nonhierarchical Links

Maintaining links is also nontrivial since, as the text is updated, some links may be "orphaned" as referenced text is deleted, while other links may need to be reevaluated to determine if more appropriate text has been added. If all of the text which refers to a particular piece of text is deleted, that piece then can become stranded and inaccessible to the reader. Similarly, if all of the text to which a particular piece of text refers is deleted, that piece becomes a dead end.

Problems with Reading "Traditional" Hypertext

The application of computers to large tasks often produces equally large side effects. In the case of hypertext, easy access to huge volumes of text can easily overwhelm the reader who cannot keep track of the individual pieces of text that are encountered or the links between them. The difficulties associated with reading hyperdocuments have been described as "Cognitive Overhead" and "Disorientation" (*7*).

Cognitive Overhead

While hypertext is intended to help alleviate the difficulty of the "information explosion" it can actually exacerbate it by providing access to more information than the reader can manage. Filtering can help to some extent, but a filter which reduces choices by 90% may be inadequate when there are a 1000 alternatives to begin with. If the reader attempts to pursue multiple paths to collect the desired information, it can become difficult to keep track of what already has been read and what should be read next. This has been called the problem of "Cognitive Overhead" (*7*).

For example, if the reader wishes to decide what medical procedure would be best for diagnosing the cause of syncope, 10 blocks of text might be identified, each of which discuss two causes of syncope. The reader would then need to pursue each of these 20 causes (some of them possibly redundant) to find diagnostic procedures for each. Having collected all of this information, the reader now needs to digest it in order to select the single best procedure. Besides the medical thought processes being performed, the reader must also perform some tasks related to deciding what links to traverse and remembering those which were ultimately worth traversing. This requires a certain amount of mental bookkeeping which can become sufficiently distracting so as to adversely effect the ability to perform the original task (selecting a diagnostic procedure). To a large extent, this overhead is due to the current information explosion; medical research provides the knowledge that lets multiple causes be considered and offers numerous diagnostic choices.

Hypertext systems can help by providing access to all this information, but without additional tools to manage the information, a reader may be overwhelmed. Hypertext links in some systems give an indication of their function (by using colors, fonts, symbols, or some combination thereof). Keeping track of the meanings of different types of links adds to the cognitive overhead required to read the hyperdocument.

Even when such visual cues are present, the reader may be seeking information in a way not anticipated by the author and, therefore, not explicitly addressed by existing links. For example, when encountering a cause for syncope in the reading, the text might include links to information about the diagnosis or treatment of the underlying condition. However, depending on the reader's level of understanding, information might be desired regarding how the condition causes syncope or simply a definition of the condition. If no links are present which appear to address the desired topic explicitly, there is no way to be sure which links will provide the desired information without actually traversing them. The result is that the reader can become frustrated by pursuing links which are irrelevant (to the reader) or by being unable to find relevant ones. Futile casting about can contribute to the problem of cognitive overhead and can easily become a deterrent to exploring additional paths—an effect which is antithetical to the purpose of hypertext.

Disorientation

Disorientation occurs when the reader creates a complex path, or sets of paths, through the traversal of various links (7). Since, by definition, the document is not meant to be read in a particular sequence, the reader is free to create these paths. In a large hyperdocument, the number of possible paths is enormous. When the path becomes long and tortuous, the reader is in danger of becoming disoriented. It is easy to become diverted from one related topic to another until it becomes impossible to recall not only how a location was reached, but why, or where to go next.

Another situation by which a reader can become disoriented occurs when a particular piece of text is desired but is not found. This might occur because the reader cannot recall the necessary path, or it might be due to the "you can't get there from here" phenomenon which can occur when the author has simply not provided the desired path. In Fig. 3, for example, a reader might arrive at the text shown at upper left. In this text, the word "dyspepsia" is highlighted. Selecting this hotspot brings the user to a piece of text which provides the

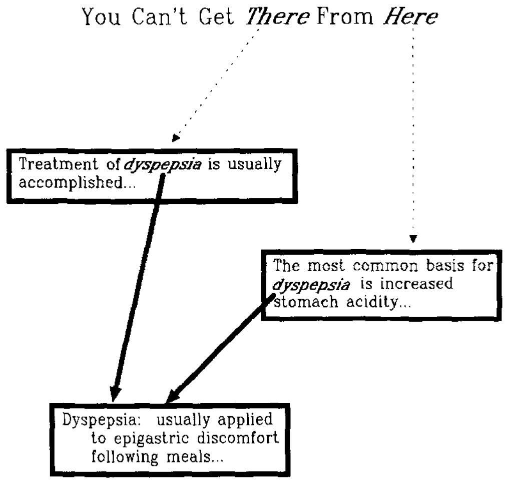

FIG. 3. "You can't get there from here." One form of disorientation, which can occur if no path has been created to allow the reader to reach a desired section of the text. In this case, a reader viewing a paragraph about the cause of dyspepsia (in the upper right-hand block of text) may already know the definition of the term and wish to read about its treatment. Although relevant text exists, no path has been provided for reaching it.

definition of "dyspepsia." Similarly, a reader arriving at the text shown at right encounters a "dyspepsia" hotspot in text describing the cause of this condition; this hotspot also provides access to the definition. There is, however, no author-provided path for the reader who, on encountering the text describing the cause of dyspepsia, wishes to read about its treatment. This prevention of the reader from finding a desired path can be classified as a type of disorientation.

ONE MODEL FOR SOLUTIONS—THE CONCEPT SPACE

The difficulties described above are not unique to hypertext; indeed, they are also encountered by authors and readers of conventional text. But the power and speed of the computer, which accentuates the positive aspects of nonsequential reading, serves to amplify the problems as well. This problem became apparent while undertaking research on the National Library of Medicine's Unified Medical Language System project (UMLS), when manipulation of large numbers of related literature citations and medical textbook chapters became unwieldy.

As various approaches were developed to deal with this problem, techniques were recruited from unrelated UMLS work on semantic networks, being carried out in collaboration with other UMLS contractors (for example, see (*25*, *26*)). It became clear that there were many similarities between the relationships among the literature citations and those found in the semantic network of medical concepts. Experiments were carried out in which the two sets of information (medical literature and the semantic network) were brought together into a single representational scheme. The result was a model for hypertext that appeared to provide some solutions to the problems found with traditional hypertext. A hyperdocument was authored using this model and a prototype browser was constructed (*1, 2*).

An understanding of the model requires special definitions for two terms: "concepts" and "topics." A *concept* is defined here as *any self-contained, atomic term within a field of discourse*. Concepts correspond to noun phrases composed of one or more words. Single words are not necessarily concepts. For example, "yellow" is a self-contained concept in art, while it is not in medicine. In medicine, "yellow" appears as part of a concept name, such as "yellow fever." Concepts are combined to convey complex meanings, such as "yellow fever is transmitted by mosquitoes."

A *topic* is defined here as *a medical subject of potential interest that concerns one or more concepts in a particular context or relationship*. It is a basic hypothesis for the concept space hypertext model that a reader seeking information in the text is often interested in a particular concept (such as "dyspepsia" or "syncope") as it appears in some context (such as "treatment of dyspepsia" or "cause of syncope"). These definitions for concepts and topics are in keeping with those of other workers studying the information content of medical literature (*27*).

The premise underlying all hyperdocuments is that a reader is likely to encounter some information which prompts a nonsequential traversal throughout the document. Traditional hypertext does not facilitate topic-driven inquiries. Keyword searches promote concept-driven inquiries (i.e., find everything there is about a concept) which may produce results that are irrelevant to a user's query (poor precision). At the same time, keyword searches may have poor recall because of missed synonyms (i.e., find everything which *mentions* the concept but not necessarily everything which *discusses* it). Author-defined hotspots and buttons provide topic-specific links; however, as discussed above, the actual topic may not be clear without traversing the link and the links provided may not address the reader's need. The model described here provides a framework which facilitates authoring of hyperdocuments structured to promote topic-driven inquiry.

The Concept Space

An important feature of the model is that concepts are represented as terms in a controlled vocabulary. Each concept is related to other concepts in various ways. A disease concept, for example, might be related to one drug concept that is an etiology and another drug concept that is a treatment. If the relationship expresses the context in which the terms are related, then the combination of the terms and their relationship can serve as a topic. For each concept, a frame structure exists with slots corresponding to different kinds of relationships (etiology, treatment, etc.). When a relationship exists between two terms, the references to each term appear in the appropriate slot of the other term's concept frame. The frame for the above disease concept would have an "etiology" slot which contains a reference to the first drug and a "treatment" slot which

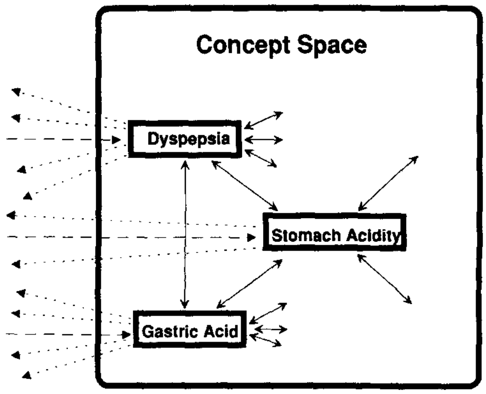

FIG. 4. The concept space. The concept space consists of a collection of "concepts" which are nodes in a semantic network. Three nodes are shown here, with semantic links between them and to other nodes not shown (solid arrows). The reader enters the concept space from the text space (see Fig. 5) via pointers to concepts (broken lines). Once in the concept space, the reader may browse the semantic links and eventually return to the text space via another pointer (dotted lines).

contains a reference to the second drug. Together, the concepts and their meaning-specific interrelationships form a semantic network (*28*), referred to as the *concept space* (see Fig. 4).

The textual component of the hyperdocument exists in a separate structure called the *text space* (see Fig. 5), in which blocks of text are related to each other in hierarchical ways (including footnotes and references), but are not

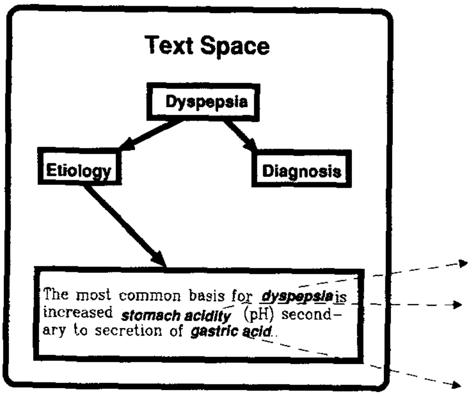

FIG. 5. The text space. In the text space, sections of text are linked to each other through hierarchical links only, which replicate the traditional book–chapter–subchapter–paragraph format of printed matter. This is in marked contrast to the structure of traditional hypertext, as shown in Fig. 1. Hotspots exist in the text (shown here as highlighted terms) which provide links to concepts in the concept space (broken lines). Readers may reach a piece of text through hierarchical browsing or via links from concepts in the concept space (not shown).

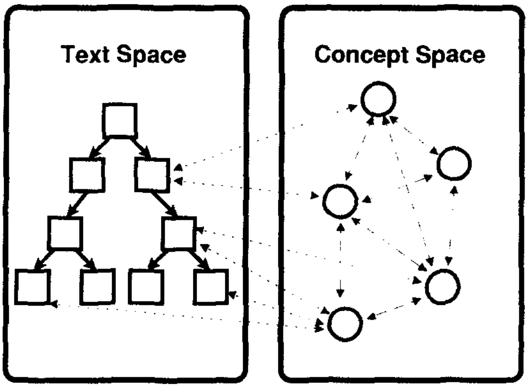

FIG. 6. Links between and within text space and concept space. An overview of the model showing the relationship between the two components of the hyperdocument. Note that the hierarchical relationships are retained in the text while the "nonsequential" links, typically found in a traditional hyperdocument, have been removed to the concept space.

directly linked in conceptual ways (such as hotspots and buttons). Instead, each block of text is related to the relevant frames in the concept space which are, in turn, linked to all relevant blocks of text and to other relevant concepts (see Fig. 6). By this scheme, the move from one block of text to another involves an additional step, but a number of desirable operations are now feasible.

First, the use of a controlled vocabulary eliminates the need for concern about synonymy when retrieving text. If, for example, "Syncope" is a controlled concept, then text which discusses "syncope," "loss of consciousness," or "fainting" will all be linked to the concept frame for syncope.

A second advantage to this arrangement is that the concept frame acts as a "Grand Central Station" for all text which relates to the concept. The reader may then choose from any of the available paths, rather than some preordained one. Furthermore, the "you can't get there from here" problem is eliminated, since you can now get to any related text by traversing exactly two links (one from the text to the concept frame and one back to the second piece of text). A reader who wishes to read everything about a particular concept can easily do so by simply exploring each link from the concept frame to the text space.

Concept Frame Structure

An important part of the model is the way in which links are organized within the concept frame. In a large hyperdocument, the number of pieces of text which refer to the concept can be considerable. To help the reader find the relevant text, three techniques are applied.

First, the medical concepts appearing in the text segments are used to organize the links from the concept frame to the text. It would be extremely unusual for a piece of text to contain only one concept. The other concepts which appear in the same text can serve as clues to the reader as to the relevance of the text. For example, when reading about causes of syncope, starting at the concept frame "Syncope," the reader might be especially interested in reading text which also contains the concepts "Ventricular Fibrillation" or "Epilepsy." In effect, all searches which are Boolean "ands" of syncope and all other concepts have been precompiled. A by-product of this process is that the list of concepts which appear in the same text as "Syncope" can be used to provide the interconcept links which form the concept space.

The second technique is the organization of interconcept links. If a concept is related to many other concepts through a variety of relationships, those other concepts can be grouped by class or by relationship. In the "Syncope" frame, for example, all diseases could be grouped together as "Related Diseases." Alternatively, related terms could be organized by their semantic relationships to the reference concept. For example, related concepts can be grouped under the heading "Causes of Syncope," which might include not only diseases, but other causes, such as chemicals and procedures. It is then a simple matter (if the concepts have been classified previously) to organize the concepts into specific semantic relationships, such as "Diseases which cause Syncope," "Chemicals which cause Syncope," etc. By defining such relationships, the interconcept links needed to form the semantic net are provided. An example of a concept frame with classification of concept links and organization of text links by co-occurring concepts is shown in Fig. 7.

The third technique, one that is commonly used in traditional hypertext systems, is filtering. If, as in Fig. 7, a reader were to look for all the text which dealt with "ventricular fibrillation" and "syncope," there might be 10 pieces of text. Rather than see each in sequence, the reader can be shown a single line from each text block which shows the context of the concepts in the text, much as appears in Fig. 2. To further clarify the relevance of the text, not only the context of the concept in the text is shown but the context of the text in the hyperdocument as well. For example, if the text were a section of a textbook, the section name, chapter name, and book title along with the actual line of text could be shown (see Fig. 8).

Authoring Hypertext for the Concept Space

The division of the hyperdocument into a concept space and text space changes the way in which it is authored. While the text itself must still be written, the process of link creation is simplified to the point where much of it can be automated. There exist four types of links, each created in a different way: links from one piece of text to another, links from text to a concept frame, links from one concept frame to another, and links from concept frames back to text.

Links from Text to Text

Unlike the usual forms of hypertext, the model does not include nonsequential links between pieces of text. Text–text links are used only to represent the traditional sequences found in printed text: hierarchical links, footnotes, and

Concept Frame: **Syncope**	
Related Terms: **7**	Text References: **30**
Causes of Syncope: **4**	Text References: **20**
Diseases which Cause Syncope: **2**	Text References: **12**
Ventricular Fibrillation	Text References: **8**
Pulmonary Embolism	Text References: **4**
Chemicals which Cause Syncope: **2**	Text References: **8**
Nitroglycerin	Text References: **3**
Propranolol	Text References: **5**
Treatment of Syncope: **3**	Text References: **10**
Procedures which Treat Syncope: **1**	Text References: **4**
Cardiac Pacemaker	Text References: **4**
Chemicals which Treat Syncope: **2**	Text References: **6**
Propranolol	Text References: **1**
Quinidine Sulfate	Text References: **5**

FIG. 7. Concept frame structure. One possible organization for a concept frame is to group associated terms according to semantic relationships. These can be further subdivided to allow the reader to select general or specific references. For example, a reader can chose to look at all 20 references that deal with causes of syncope, at the 12 which deal with diseases that cause syncope, or with the 8 that deal with ventricular fibrillation (see Fig. 8).

references. Producing documents in which these kinds of links are explicit is commonly achieved with word processors and outline processors and should present no great obstacle to the creation of hyperdocuments. Hierarchical text links have been created automatically for a medical therapy handbook (*18*) and, in fact, standards for representing these links already exist (*29*).

Links from Text to Concept Frame

Indexing narrative text with a controlled vocabulary is a nontrivial task; however, the concepts which appear explicitly (i.e., mentioned ver batim) in text can be found with a high degree of reliability using automated techniques (*30, 31*). However, not all concepts which are mentioned explicitly in a text are discussed in that text. In fact, it is quite possible that, through automated techniques, a large number of concepts can be found, none of which are relevant to the subject matter of the text. On the other hand, a concept might be discussed

The Textbook of Cardiology
Chapter 2 - **Syncope**
A. Causes of **Syncope**
Arrhythmias
...another common cause of **syncope** is **ventricular fibrillation**...

B. Treatment of **Syncope**
...when caused by **fibrillation of the left ventricle**, the treatment is...
...another popular treatment for **ventricular fibrillation**...

Journal of Cardiovascular Clinical Medicine Reports
Volume 10, Issue 5
Review of 2000 cases of **Syncope**
Abstract
...in 40% of patients, **ventricular fibrillation** was the suspected...

Methods
...the diagnosis of **ventricular fibrillation (V.Fib.)** was confirmed by...
...a presumptive diagnosis of **V.Fib.** was made in cases where no clear...

Results
definitive diagnosis of **V.Fib.** as the cause of **syncope**...

Editorial: **Syncope** Revisited
...interestingly, **ventricular fibrillation** as a cause of **syncope**...

FIG. 8. Filtering with the use of hierarchical information. If the reader chooses to see the eight references to ventricular fibrillation, as shown in the concept frame in Fig. 7, a filtered display can be presented that shows the relevant text (with the terms highlighted, including recognized synonyms). The context of the text in the hyperdocument is shown through inclusion of hierarchical information. Redundant information can be eliminated; for example, when two articles from the same issue of a journal are included, the title of the journal is included only once.

but not mentioned explicitly, making automated recognition of the concept difficult. This presents two problems: what to do with extraneous references (to concepts which are mentioned but not discussed) and how to assure that appropriate references (to concepts which are discussed but not mentioned) exist and can be found.

The first problem is of little consequence, since extraneous links could appear in the hyperdocument as highlighted terms and could be ignored by the reader. However, it is quite possible that what might normally be considered extraneous by most authors and readers might prove to be of some use to a reader with an unusual query. Since it is impossible to anticipate all of the readers' needs and since the "extraneous" links do not interfere with those who are not interested in them, there is no need to attempt to eliminate them.

The problem of how to assure that the links to concept frames for a given piece of text include ones to concepts which are of central importance to the subject matter is a bit more difficult, especially since it is possible for the concept to be discussed but not explicitly mentioned. One way this issue is addressed is by choosing a vocabulary which is well suited to the indexing of literature in the domain. In case of medical text, the National Library of Medicine's Medical Subject Headings (MeSH) (*32*) is generally felt to provide good coverage of the concepts discussed in medical literature, although not necessarily all medical terms (*33*). It includes 17,000 "Subject Headings" used for indexing, as well as over 70,000 "Entry Terms" which act as synonyms to the subject headings. Through the use of MeSH, it should be possible to identify the major concepts discussed in a given piece of text. If the concept is not stated explicitly, tools are available which can help an author find the closest MeSH concept to which the text may be linked (*34–36*).

For example, a piece of text might be indexed automatically to identify MeSH terms (*30*) which are then displayed as highlights in the text. An author could then scan the text to look for nonhighlighted text phrases which represent additional medical concepts which have been mentioned but not identified by the indexing algorithm. Through interaction with a MeSH lookup facility, the author can attempt to identify the MeSH term which best represents the concept. If one is found, the phrase is linked to the term and becomes a hotspot. If none is found, a new term could be added to the controlled vocabulary, a new *concept frame* is added to the concept space, and the phrase is linked to the new concept. Once all of the text has been scanned for explicitly mentioned concepts, the author re-examines the text, this time looking for concepts which are discussed but not explicitly referenced by any of the hotspots. The author can rectify the deficiency by either recasting the text so that an explicit reference to the concept is made (which becomes a hotspot) or by adding a button to the text which provides the link to the desired concept frame. This manual review should be less labor-intensive than a similarly thorough editing of traditional hypertext for two reasons: first, the author needs only to identify concepts in the text, not anticipate which related topics are of interest to the reader; and, second, once the author identifies a concept term in the controlled vocabulary, the task is complete. There is no need to locate relevant text to link to a hotspot or button.

Links from Concept Frame to Concept Frame

One method for providing links between concept frames has been described above: listing all concepts which co-occur with the reference concept in any of the text in the hyperdocument. For instance, if "pacemaker" appears in the same section of text as "syncope," a link between the two corresponding concept frames can be added to the concept space. The collection of links accrued in this way will not be complete (for example, if the hyperdocument does not mention all of the therapies for syncope in the same text as the term "syncope"). However, it will certainly be long enough to require some organizational structure so that the reader will not be overwhelmed. The structure used in the model is that of semantic relationships between types of concepts. For example, the links from a disease concept to medical procedure concepts might be organized into those procedures which cause the disease, diagnose the disease, treat the disease, or prevent the disease. Multiple semantic links may exist between two concepts (consider, for example, that cardiac pacing can cause, diagnose, treat, and prevent cardiac arrhythmias).

Organizing, and adding to, the list of links from one concept frame to another is a potentially tedious chore. Fortunately, there are tools which can assist the author in this task. First of all, it is possible to know from the controlled vocabulary used what kinds of semantic relationships are possible between two concepts (for example, between a disease and a procedure), thus limiting the number of choices which must be considered. The Unified Medical Language System, for example, provides a large number of such relationships between medical concepts of different semantic types (*37*). In addition, the text space itself can serve as reference material, since it can provide reading material for the author containing the co-occurrence of the two concepts. Finally, it is possible to apply pattern-recognition rules in an automated fashion to literature citations and extract potentially useful semantic relations between medical concepts (*38*).

Links from Concept Frame to Text

To a large extent, the problem of providing links from concept frames to text is less a matter of finding links and more a matter of discarding extraneous ones. The links from text to concept frames can be used in a bidirectional manner and the process of identifying these links, as described above, provides the set of all potentially useful links which can be used to bring the reader from the concept frame back into the text space. At the heart of the problem is the fact that even off-hand remarks in the text which have nothing to do with medicine can be interpreted as medical concepts. The preceding sentence, for example, has three such concepts: "heart," "hand," and "medicine," although it discusses none of these. When the numbers of such links are small, filtering (such as that described above) can make the trivial references immediately obvious to the reader. When the number becomes large, other schemes can be used to weight the relevance of each reference automatically (*39*).

Reading in the Concept Space

Reading hyperdocuments constructed with the concept space model is significantly different than reading traditional hypertext. The nonsequential traversals require an extra step (via a concept frame) to move from one piece of text to another. The move through the concept frame also requires that the reader make a decision about which path to travel, rather than simply accept a path predetermined by the author. For example, when reading about syncope, a reader who wishes to know about diseases which cause syncope moves to the "Syncope" concept frame and looks under the heading "Diseases Causing Syncope." A number of diseases are displayed (corresponding to the inter-concept links in the frame slot), such as "Ventricular Fibrillation." The reader now has a choice: reading text which discusses a selected concept (e.g., ventricular fibrillation) in the context of syncope, or moving to the concept frame for the new concept. The reader can continue the traversal from text to concept, from concept to concept and from concept back to text. This sequence is

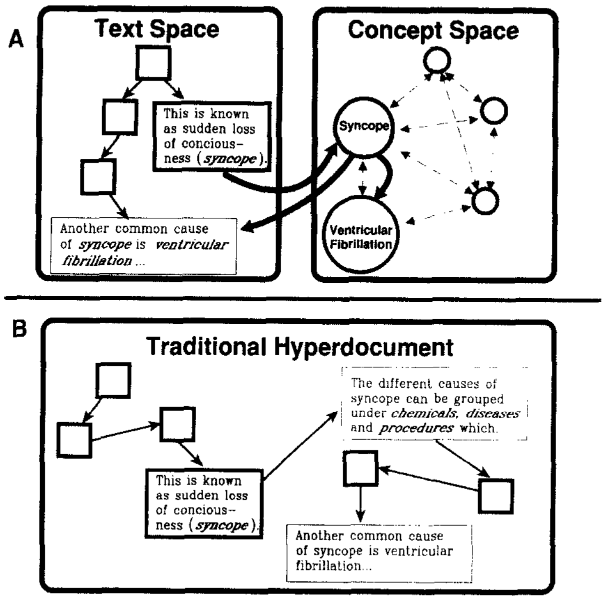

FIG. 9. Comparison of browsing two models for hyperdocuments. In A, the reader encounters the hotspot "syncope" and, upon selecting it, is shown the syncope concept frame (see Fig. 7), which provides links to 30 points in the hyperdocument and to 7 other concept frames. If the reader is interested in the causes of syncope, the concept frame shows that there are 20 relevant places in the text and 4 relevant concepts (2 chemicals and 2 diseases). The reader can look at a filtered list of text references for a particular cause (as in Fig. 8) and can select one or chose to view the concept frame for the etiologic term (ventricular fibrillation, in this case). In B, the reader encounters the same hotspot in a traditional hyperdocument. In this case, the reader may have no clue as to the nature of the link without traversing it. The text at the other end may happen to discuss the topic in which the reader is interested (e.g., causes of syncope), but it is unlikely to provide as close a match to the reader's query as would be available from the "Grand Central Station" offered by a concept frame.

depicted in Fig. 9A and can be compared to the sequence available in a traditional hyperdocument, as shown in Fig. 9B. The extra step required by the imposition of the concept space on the hyperdocument offers some solutions for the problems encountered while reading traditional hypertext.

Cognitive Overhead

Cognitive overhead is reduced in a number of ways by the notion of the concept frame. First, by serving as the focal point of all information about a concept (related concepts and referent text), the reader need not try to recall all the references that might need to be pursued. Instead, it is a simple matter to explore them, one at a time (if desired), with the concept frame as home base. It is easier for the reader to keep track of what has and has not been read when it is collected together in a list.

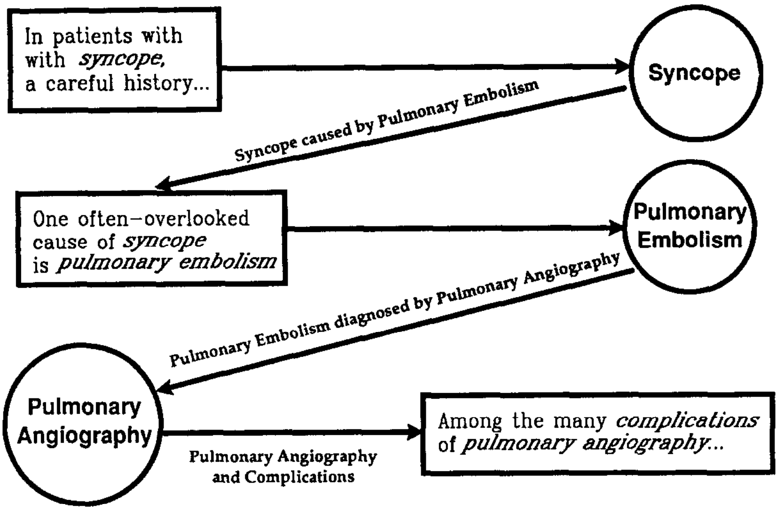

FIG. 10. Tracing a path through a concept space hyperdocument. Each node in the above graph corresponds to a portion of a hypothetical hyperdocument visited by a hypothetical reader. Each box corresponds to a text segment in the text space and each circle corresponds to a concept frame in the concept space. The arrows are labeled with the relationship that was traversed by the reader. Paths such as the one shown here are available because of the structure of the hyperdocument. It is hypothesized that a path such as this can help reduce disorientation and cognitive overhead by helping the reader answer the questions "How did I get here" and "Why did I go here?"

Second, the concept frame allows for the provision of an additional way to filter the text choices presented to the reader. Now, instead of being confronted with all possible choices of text which contain "syncope" in the above example, the reader implicitly narrows the choices down incrementally to text which contains other diseases, specifically those which cause syncope, and even more specifically to ventricular fibrillation (as in Fig. 9A).

Third, the links reflect semantic relationships between the text and concepts they join. Since the user is given information about the nature of the link, there is less likelihood of exploring irrelevant links.

Fourth, the format of the concept frame itself can reduce cognitive overhead since each concept is presented in the same manner. With a familiar structure, the reader can quickly learn which kinds of links are relevant and which are not. This kind of consistency of linking is not possible when a document is created by multiple authors creating ad hoc links from text to text.

Disorientation

Disorientation can be reduced in the concept space model because the system has information pertaining to the reasons why the reader traversed a particular link. Since each link has some semantic meaning attached to it, the reader's path can be depicted with more than a simple replay of what was chosen: it can also indicate why. For example, Fig. 10 shows the kind of path which can be presented to a reader of a document which is structured under the concept space model. A reader who loses track of the original thread of a query should be able to pick it back up again by reviewing such a path.

The kind of disorientation which can occur with the "you can't get there from here" situation shown in Fig. 3 simply cannot occur with the concept space model. The "Grand Central Station" function of the concept frame assures that a direct route exists to each appropriate piece of text. That route can easily be found by looking in the concept frame and can then be reached in a single step, rather than wandering through multiple pieces of linked text.

A Prototype Concept Space Hyperdocument

The model described above was developed after experimentation with medical literature citations and a small semantic network of medical concepts. Using the model, a prototype hyperdocument was constructed which used MeSH as its controlled vocabulary of concepts and a collection of Medline Abstracts and a cardiology textbook chapter as its corpus of text. The text–text links, text–concept links, and interconcept links were all generated using automated techniques. The text–text links were created from the hierarchical structure of the textbook (chapter, subchapter, section, subsection) and of the citations (title, abstract). The text–concept links were created through automated methods for finding MeSH terms in narrative text (*30*). The interconcept links were based on a semantic network created from analysis of Medline citations (*38*).

The result of this process was a hyperdocument consisting of 1921 lines of text with 1625 text–concept links, 1963 Medline citations with 5881 text–concept links, and a semantic network of 885 concepts with 6586 interconcept links. A browsing system was developed which allowed traversal of these links. Formal evaluation of the utility of the concept space model has not been conducted; however, several preliminary findings were made. First, the feasibility of automated hyperdocument construction was demonstrated. Second, MeSH provided good coverage of the terms found in the text and therefore was a good source for text–concept links. Third, while the traversal from text to appropriate concept was easily accomplished, the reverse was not the case. This was due to the problems anticipated above (see "Links from Concept Frame to Text").

Discussion

Textbooks, and their electronic counterparts, are developed to serve many different purposes and are used in many ways by readers. No one model can satisfy all of these; for instance, the model presented here for the structure of hyperdocuments would be unlikely to be useful for programmed text which is used as part of a course syllabus. In a case such as this, the author has a particular path in store for the reader and unrestricted browsing might be inappropriate and nonproductive. The model presented here is more appropriate for the reader who, on becoming interested in a concept while reading, wishes to read more about that concept. While the concepts encountered can be anticipated, the subsequent topic-directed queries they invoke cannot be. The separation of the hyperdocument into a concept space and a text space is an attempt to recognize the cognitive processes which are occurring during the reading of a hyperdocument. This form of representation addresses problems of both the author of the hyperdocument and the reader.

The index of traditional hypertext is replaced with the much richer structure of a semantic network of interrelated medical concepts with links to the text. The text entries listed for each concept are similar to those which would appear in an index, but they can be given a more useful organization because of the semantic network. The text–text links found in conventional hypertext are replaced, in the system, with text–concept–text links. These types of links are easier to create, especially since tools are available for partially or wholly automating the process. These links are trivial to maintain, since the deletion of a piece of text simply requires removing the links between it and the concept, without affecting any of the other links in the document. Thus, no links will be orphaned (that is, have no text at one or both ends) and no text which includes at least one medical concept or is in a text hierarchy will become stranded (unreachable by any means). Finally, the construction of an index using the controlled vocabulary of the concept space makes it a simple matter to merge documents which have been created separately and indexed with the same concept vocabulary.

Construction of traditional hyperdocuments can be tedious, or it can be as simple as the author including ad hoc pointers from one piece of text to another which a reader may find interesting. The concept space approach requires that much more attention be paid to the links that are created. However, the combination of a structured design for a hypertext framework, coupled with the use of a controlled vocabulary, lends itself well to automating the process of hypertext editing and maintenance.

There can be no question that the construction of a high-quality concept space is a nontrivial task. It can, however, be viewed as a knowledge base which is relatively independent of the associated text space. Knowledge from sources outside the text (such as other knowledge bases from expert systems) can be used to enhance the concept space and the concept space can be applied to a variety of hyperdocuments. With the use of a common terminology and structure (such as that of the Unified Medical Language System), concept spaces constructed at disparate sites can be shared between authors. Sharing medical knowledge has already been proposed for other purposes (*40, 41*) and the format for that knowledge may be adaptable for use in a concept space hyperdocument (*42*).

Hypertext models with some similarities to the one presented here have been proposed. Ngyuen and Greenes proposed the EXPLORER-1 model in which information is organized into frames which can be related to each other through semantic links. In addition, they proposed that procedural knowledge could be included in the frames to allow links to be generated automatically (*43*).

Frisse and Cousins described the use of a belief network as a hierarchical index space with each link from an *index term* into their *document space* represented as a probability that a given segment of text is a desirable reference for the term (*44*). Both the concept space described here and the index space of Frisse and Cousins share the notion of index terms in a hierarchical arrangement pointing to text objects. The concept space model adds nonhierarchical semantic links between index terms to assist in browsing the index, while the index space model adds information which addresses the problem of ranking the concept-to-text links. A combination of these two models might alleviate the major weakness of the concept space while providing support for the automated construction of the index space.

The disorientation and cognitive overhead which occur while reading hypertext have been discussed by others, and the attempts to address these problems have usually been directed at providing additional user interface features. The concept space model approaches these problems through a fundamental change in the underlying paradigm of the hypertext structure. This change in paradigm may better represent the cognitive model of the reader. The reader is not interested in medical concepts in isolation, but in their relation to other concepts. Furthermore, the reader can not always be relied upon to be interested in a single, predictable relation or related concept. Reading a traditional hyperdocument is relatively simple, but the results can be dissatisfying because the reader has little ability to describe his or her specific needs. The concept space model offers the reader the opportunity to inform the system about the kind of information being sought. The reader tells the system more than "take me here." By following a semantic relationship, the user conveys implicit information about the query being addressed, in effect saying "take me here because. . . ." The system can then help manage the overhead of where one can go and where one has been. When a reader becomes disoriented, the system can do more than say "you are here"; it can say "you are here because. . . ."

Conclusion

Hypertext is becoming a popular means for organizing textual information for retrieval and browsing. It brings with it, however, a number of difficulties for both the author and reader, which are not solved by computers but are made to occur faster and on a grander scale. By recognizing that readers may be topic-oriented in unanticipated ways and by formalizing concepts and their relationships to text, the hypertext model presented here addresses some of the needs of the reader. Furthermore, the use of a controlled vocabulary, particularly in the medical domain, facilitates the use of existing tools to assist with, simplify, and, in some cases, automate the authoring process.

Acknowledgments

The authors acknowledge fellow UMLS contractors Robert Greenes, Randolph Miller, and David Evans, whose work and collaboration influenced the development of the concept space model. The authors thank Dr. Greenes for permission to use the CASPER diagram and also thank Leslie M. Juceam and George Hripcsak for valuable assistance in preparing the manuscript.

REFERENCES

1. ELKIN, P. L., CIMINO, J. J., AND BARNETT, G. O. The hypertext medical workstation. *In* "Proceedings of the First Annual Educational and Research Conference of the American Medical Informatics Association" (J. A. Mitchell, Ed.), p. 33. Snowbird, UT, June 1990.
2. CIMINO, J. J., ELKIN, P. L., AND BARNETT, G. O. The medical concept space as a model for hypertext. *In* "Proceedings of the First Annual Educational and Research Conference of the American Medical Informatics Association" (J. A. Mitchell, Ed.), p. 41. Snowbird, UT, June 1990.
3. COVELL, D., UMAN, G., AND MANNING, P. Information needs in office practice: Are they being met? *Ann. Intern. Med.* **103**(4), 596 (1985).
4. WILLIAMSON, J. W., GERMAN, P. S., WEISS, R., SKINNER, E. A., AND BOWES, F. Health science information management and continuing education of physicians: A survey of U.S. primary care practitioners and their opinion leaders. *Ann. Intern. Med.* **110**(2), 151 (1989).
5. HUTH, E. J. The underused medical literature. *Ann. Intern. Med.* **110**(2), 99 (1989).
6. BUSH, V. As we may think. *Atl. Mon.* **176**(1), 101 (1945).
7. CONKLIN, J. Hypertext: An introduction and survey. *Computer* **20**(9), 17 (1987).
8. SMITH, K. E. Hypertext—Linking to the future. *Online* **March,** 32 (1988).
9. HAYNES, R. B., MCKIBBON, K. A., WALKER, C. J., MOUSSEAU, J., BAKER, L. M., FITZGERALD, D., GUYATT, G., AND NORMAN, G. R. Computer searching of the medical literature—An evaluation of MEDLINE searching systems. *Ann. Int. Med.* **103**(5), 812 (1985).
10. WOODSMALL, R. (ED.) "Medline on CD-ROM." Learned Info, Medford, NJ, 1989.
11. COLLEN, M. F., AND FLAGLE, C. D. Full-text medical literature retrieval by computer—A pilot test. *J. Am. Med. Assoc.* **254,** 2768 (1985).
12. ENGELBART, D. C. A conceptual framework for the augmentation of man's intellect. *In* "Vistas in Information Handling. Vol. 1. The Augmentation of Man's Intellect by Machine" (P. W. Howerton and D. C. Weeks, Eds.), pp. 1–29. Spartan, Washington, DC, 1963.
13. ENGELBART, D. C., AND ENGLISH, W. K. A research center for augmenting human intellect. *In* "1968 AFIPS Conference Proceedings (Fall Joint Conference)," pp. 395–410. AFIPS, Montvale, NJ, 1968.
14. NELSON, T. H. Getting it out of our system. *In* "Information Retrieval: A Critical Appraisal" (G. Schechter, Ed.), pp. 191–210. Thompson, Washington, DC, 1967.
15. NELSON, T. H. "The Home Computer Revolution," Nelson, South Bend, IN, 1978.
16. Hyper activity: Hypertext products and hypertext research and development. *BYTE* **13**(10), 268 (1988).
17. GREENES, R. A. Knowledge management as an aid to medical decision making and education: The Explorer-1 system. *In* "Proceedings of MEDINFO 86" (R. Salamon, B. Blum, and M. Jorgensen, Eds.), pp. 895–899. Elsevier, Amsterdam, 1986.
18. FRISSE, M. E. Searching for information in a hypertext medical handbook. *Commun. ACM* **31**(7), 880 (1988).
19. NELSON, T. H. "Literary Machines," Chap. 3, p. 5. Swarthmore, PA, 1981.
20. GREENES, R. A., TARABAR, D. B., KRAUSS, M., ANDERSON, G., WOLNIK, W. J., COPE, L., SLOSSER, E., AND HERSH, W. R. Knowledge management as a decision support method: A diagnostic workup strategy application. *Comput. Biomed. Res.* **22**(2), 113 (1989).
21. VAN DAM, A. Hypertext '87 keynote address. *Commun. ACM* **31**(7), 887 (1989).
22. FIDERIO, J. A grand vision. *BYTE* **13**(10), 237 (1988).
23. HERSH, W. R., AND GREENES, R. A. SAPHIRE—An information retrieval system featuring concept matching, automatic indexing, probabalistic retrieval, and hierarchical relationships. *Comput. Biomed. Res.* **23**(5), 410 (1990).
24. ALLEN, D. Text retrieval with a twist. *BYTE* **13**(10), 201 (1988).
25. BARR, C. E., KOMOROWSKI, H. J., PATTISON-GORDON, E., AND GREENES, R. A. Conceptual modeling for the Unified Medical Language System. *In* "Proceedings, 12th Annual Symposium on Computer Applications in Medical Care" (R. A. Greenes, Ed.), pp. 148–151. IEEE Comput. Soc., Washington, DC, 1988.

26. MASARIE, F. E., JR., MILLER, R. A., BOUHADDOU, O., GIUSE, N. B., AND WARNER, H. R. An interlingua for electronic interchange of medical information: Using frames to map between clinical vocabularies. *Comput. Biomed. Res.* **24**(4), 379 (1991).
27. POWSNER, S. M., RIELY, C. A., BARWICK, K. W., MORROW, J. S., AND MILLER, P. L. Automatic bibliographic retrieval based on current topics in hepatology: Hepatopix. *Comput. Biomed. Res.* **22**(6), 552 (1989).
28. WOODS, W. What's in a link: Foundations for semantic networks. *In* "Representation and Understanding: Studies in Cognitive Science" (D. G. Bobrow and A. Collins, Eds.), pp. 35–82. Academic Press, New York, 1975.
29. National Information Standards Organization. "American National Standard for Electronic Manuscript Preparation and Markup." Transaction, New Brunswick, NJ, 1990 (ANSI/NISO Z39.59-1990).
30. ELKIN, P. L., CIMINO, J. J., LOWE, H. J., ARANOW, D. B., PAYNE, T. H., PINCETL, P. S., AND BARNETT, G. O. Mapping to MeSH (The art of trapping MeSH equivalence from within narrative text). *In* "Proceedings, 12th Annual Symposium on Computer Applications in Medical Care" (R. A. Greenes, Ed.), pp. 185–190. IEEE Comput. Soc., Washington, DC, 1988.
31. MCCRAY, A. T. Parsing analyzing and accessing biomedical text. *In* "Proceedings, 52nd Annual Meeting of the American Society for Information Science," pp. 192–197. Washington, DC. 1989.
32. National Library of Medicine. "Medical Subject Headings." Bethesda, MD, 1990 (NTIS NLM-MED-90-01).
33. LINDBERG, D. A. B., AND HUMPHREYS, B. L. The UMLS knowledge sources: Tools for building better user interfaces. *In* "Proceedings, 14th Annual Symposium on Computer Applications in Medical Care" (R. A. Miller, Ed.), pp. 121–125. IEEE Comput. Soc., Washington, DC, 1990.
34. HAYNES, R. B., AND MCKIBBON, K. A. Grateful Med. *MD Comput.* **4**(5), 47 (1987).
35. LOWE, H. J., AND BARNETT, G. O. MicroMeSH: A microcomputer system for searching and exploring the National Library of Medicine's Medical Subject Headings (MeSH) vocabulary. *In* "Proceedings, 11th Annual Symposium on Computer Applications in Medical Care" (W. W. Stead, Ed.), pp. 717–720. IEEE Comput. Soc., Washington, DC, 1990.
36. NELSON, S. J., SHERERTZ, D. D., TUTTLE, M. S., AND ERLBAUM, M. S. Using MetaCard: A Hypercard browser for biomedical knowledge sources. *In* "Proceedings, 14th Annual Symposium on Computer Applications in Medical Care" (R. A. Miller, Ed.), pp. 151–154. IEEE Comput. Soc., Washington, DC, 1990.
37. MCCRAY, A. T., AND HOLE, W. T. The scope and structure of the first version of the UMLS semantic network. *In* "Proceedings, 14th Annual Symposium on Computer Applications in Medical Care" (R. A. Miller, Ed.), pp. 126–130. IEEE Comput. Soc., Washington, DC, 1990.
38. CIMINO, J. J., MALLON, L. J., AND BARNETT, G. O. Automated extraction of medical knowledge from MEDLINE citations. *In* "Proceedings, 12th Annual Symposium on Computer Applications in Medical Care" (R. A. Greenes, Ed.), pp. 180–184. IEEE Comput. Soc., Washington, DC, 1990.
39. SALTON, G., AND MCGILL, M. J. "Introduction to Modern Information Retrieval," p. 63. McGraw–Hill, New York, 1963.
40. CLAYTON, P. D., HRIPCSAK, G., AND PRYOR, T. A. Emerging standards for medical logic. *In* "Proceedings, 14th Annual Symposium on Computer Applications in Medical Care" (R. A. Miller, Ed.), pp. 27–31. IEEE Comput. Soc., Washington, DC, 1990.
41. HRIPCSAK, G., CLAYTON, P. D., PRYOR, T. A., HAUG, P., WIGERTZ, O. B., AND VAN DER LEI, J. The Arden Syntax for medical logic modules. *In* "Proceedings, 14th Annual Symposium on Computer Applications in Medical Care" (R. A. Miller, Ed.), pp. 200–204. IEEE Comput. Soc., Washington, DC, 1990.
42. VAN DER LEI, J., AND MUSEN, M. A. Separation of critiquing knowledge from medical knowledge: Implications for the Arden Syntax. *In* "Proceedings of the IMIA Working Conference on Software Engineering in Medical Informatics" (T. Timmers, Ed.). Amsterdam, 1990.

43. Ngyuen, L. T., and Greenes, R. A. A framework for the use of computed links in the the EXPLORER-1 knowledge management system. *In* "Proceedings of MEDINFO 86" (R. Salamon, B. Blum, and M. Jorgensen, Eds.), pp. 891–894. Elsevier, Amsterdam, 1986.
44. Frisse, M. E., and Cousins, S. B. Information retrieval from hypertext: Update on the Dynamic Medical Handbook Project. *In* "Hypertext '89 Proceedings," pp. 199–212. Assoc. Comput. Mach., ACM Press, New York, 1989.