Report Number: WUCS-89-02

1989-01-01

The Medical Informatics Group: Ongoing Research

Steven B. Cousins, Mark E. Frisse, Michael G. Kahn, and James C. Beard

Two current research projects within the Medical Informatics Group are described. The first, the Diabetes Data Management Project, has as its major goal the effective analysis, display, and summarization of information relevant to the care of insulin-dependent diabetics. These goals are achieved through the use of quantitative and qualitative modeling techniques, object-oriented graphical display methods, and natural language generation programs. The second research activity, the Hypertext Medical Handbook Project, emphasizes many aspects of electronic publishing and biomedical communication. In particular, the project explores machine-assisted information retrieval by combining user feedback with Bayesian inference networks.

... Read complete abstract on page 2.

The Medical Informatics Group: Ongoing Research

Steven B. Cousins, Mark E. Frisse, Michael G. Kahn, and James C. Beard

Complete Abstract:

Two current research projects within the Medical Informatics Group are described. The first, the Diabetes Data Management Project, has as its major goal the effective analysis, display, and summarization of information relevant to the care of insulin-dependent diabetics. These goals are achieved through the use of quantitative and qualitative modeling techniques, object-oriented graphical display methods, and natural language generation programs. The second research activity, the Hypertext Medical Handbook Project, emphasizes many aspects of electronic publishing and biomedical communication. In particular, the project explores machine-assisted information retrieval by combining user feedback with Bayesian inference networks.

THE MEDICAL INFORMATICS GROUP: ONGOING RESEARCH

Steve B. Cousins, Mark E. Frisse,
Michael G. Kahn and James C. Beard

WUCS-89-02

January 1989

Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
Saint Louis, MO 63130-4899

This work was supported by a major equipment grant from Apple Computer Inc., by grants from the New Medical Foundation, Miles Laboratories, by Training Grant 5 T15 LM07049 from the National Library of Medicine. Dr. Frisse is a Teaching and Research Scholar of the American College of Physicians.

Abstract

Two current research projects within the Medical Informatics Group are described. The first, the Diabetes Data Management Project, has as its major goal the effective analysis, display, and summarization of information relevant to the care of insulin-dependent diabetics. These goals are achieved through the use of quantitative and qualitative modeling techniques, object-oriented graphical display methods, and natural language generation programs. The second research activity, the Hypertext Medical Handbook Project, emphasizes many aspects of electronic publishing and biomedical communication. In particular, the project explores machine-assisted information retrieval by combining user feedback with Bayesian inference networks.

1 Introduction

This paper describes some of the Medical Informatics Group's current research activities.[1] This group is but one of several University groups engaged in the development of computer science theory and the application of engineering techniques to medicine. The activities described herein are neither all-inclusive nor discussed in great detail, but they are representative of the type of endeavors engaging our interest.

We discuss two major research initiatives in this report. The first, the **Diabetes Data Management Project**, has as its major goal the effective analysis, display, and summarization of information relevant to the care of insulin-dependent diabetics. Section 2 describes several major aspects of this research:

- We discuss how issues in diabetes data management guide the design of a general architecture for our diabetes data management software. We observe that effective management of diabetes is also impeded by an inability to summarize data and by inaccurate, qualitative measurements of crucial biological parameters.

- We discuss our early efforts to allow medical practitioners to visualize and manipulate diabetes data crucial to the making of therapetic decisions. We describe the features of the multimedia **TimeLine Browser Toolkit** used in this process. The toolkit uses object oriented programming techniques to incorporate new data types, to annotate data, and to effect the display of data points and temporal intervals.

In section 3 we discuss our second research activity, the **Hypertext Medical Handbook Project**. This project was initiated in 1987 in response to an evolving vision of the role of computers in biomedical communication. The Project emphasizes many aspects of electronic publishing and biomedical communication. In this document:

- We discuss authoring tools designed to facilitate creation and revision of structured medical documents like the Washington University *Manual of Medical Therapeutics*. We report some impressions gained from an informal survey of *Manual* authors and describe how we used our general purpose **Graphical Browser Toolkit** [6] to study the organization of individual *Manual* chapters.

- We discuss our efforts to develop a general theory for information retrieval from hypertext. This theory has as its foundation principles of probabilistic **belief networks**, machine learning, heuristic term-weight propagation techniques, and automatic document classification. The intent of this theoretical work is to

[1] Currently, the Medical Informatics Group is an unofficial designation not to be confused with Washington University's Medical Informatics Training Program. The latter provides stipends and tuition for pre-doctoral and post-doctoral students who desire to pursue graduate training in computer science and who wish to emphasize biomedical problems in their research.

develop a program which allows automatic browsing assistance through reading and feedback.

Although the Diabetes Data Management Project and the Hypertext Medical Handbook Project are directed towards widely different problems in clinical medicine, we find many applied and theoretical issues common to the two efforts. We find, for example, that temporal descriptions of change in state are essential to biomedical description, whether it be for a qualitative physiological model, or for a paragraph in a textbook. We also find that a common set of well-designed software tools has utility across both domains.

2 The Diabetes Data Management Project

Type I diabetes is a widespread disorder characterized by a complete inability to secrete insulin. Insulin is a vital hormone necessary to maintain normal levels of blood glucose. In addition to short-term consequences like "diabetic coma," people with type I diabetes suffer early heart and vascular disease as a result of accelerated atherosclerosis. Although insulin can be given by injection, this therapy approximates human insulin secretion only if accompanied by multiple measurements of blood glucose concentration. In addition, diabetics must adhere to complex guidelines dictating meals, exercise, and insulin doses. It is the feeling of many experts that these heuristics can be applied successfully only if extraordinary record-keeping and educational efforts are an integral part of diabetes management. We believe these issues make diabetes an ideal domain for computer-assisted medical decision support.

Effective diabetes decision-making requires that the physician and patient pay close attention to subtle relationships between clinical signs and symptoms. To aid in this task, the patient frequently maintains a hand-written data logbook that contains information essential to proper diabetes management. Microelectronics have made possible the introduction of small devices to measure accurately blood glucose concentration, and most authorities believe that the capabilities of these devices can be expanded to incorporate measurement of meals, exercise, and other significant daily events. This technological advance will allow for the replacement of the conventional paper logbook with an electronic device which can download its data for microcomputer analysis. This process will in principle allow for earlier detection of clinical problems and for optimization of complex therapeutic decisions.

The presence of additional patient-specific data has opposing cost-benefit effects. On one hand, the physician will have more data readily available to detect subtle clinical conditions and to make fine therapeutic adjustments but on the other hand, the physician also will require significantly more time to sift through this additional information. We seek to develop new computer tools that assist in patient-specific decision-making. Our goal is to maximize the clinical summarization derived from the logbook data and to minimize the time required for the detailed analysis of patient data.

Our diabetes data management system will incorporate a diabetes log database, mathematical models of insulin-glucose metabolism, clinical heuristics for data interpretation, and grammars for production of text summaries (Figure 1). Data about blood glucose measurements, exercise, and meals are entered via a small electronic diabetes diary or "log." These data are maintained in a database. A qualitative/quantitative model of glucose metabolism is combined with clinical heuristics to develop expectations and interpretation of trends. The data from the trends are maintained as an augmented transition network and, via software based on PROSENET [10], presented as text to the health care professional.

As a part of our diabetes research, we are developing a general **TimeLine Toolkit** to manipulate and display data from diabetes logbooks (Figure 2). The toolkit dis-

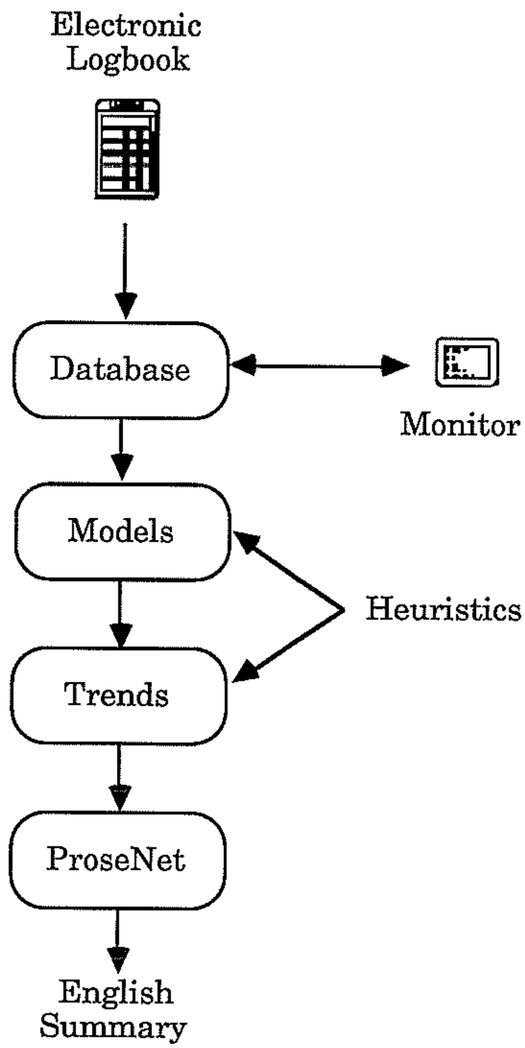


Figure 1: One possible organization for a diabetes data summarization system.

Data about blood glucose measurements, exercise, and meals are entered via a small electronic diabetes diary or "log." These data are maintained in a database and displayed via the TimeLine ToolKit. A qualitative/quantitative model of glucose metabolism is combined with clinical heuristics to develop expectations and interpretation of trends. The data from the trends are maintained as an augmented transition network and, via PROSENET software, presented as text to the health care professional.

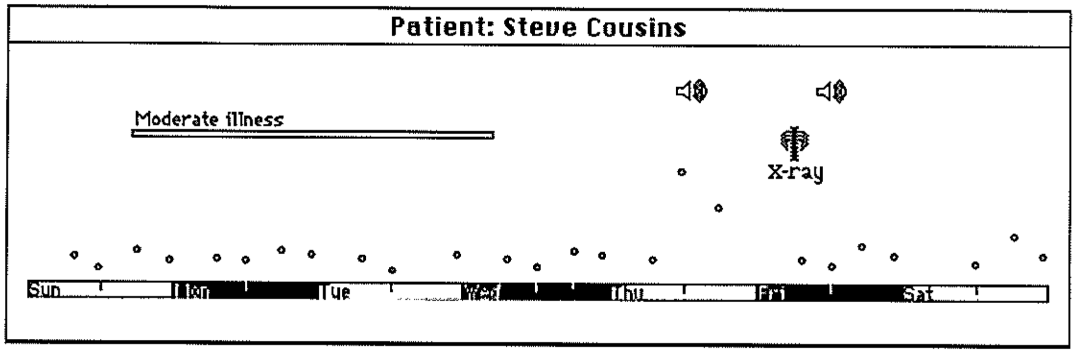


Figure 2: One week of medical data.

The circles represent individual laboratory measurements. The x-ray icon notes the availability of a digitized radiograph obtained on the date denoted by the time line. The speaker icons denote audio messages. These can be commentaries on other data, requests, or even sounds recorded from examination of the heart. This interface might be useful in the design of medical voice mail systems.

plays arbitrary time-tagged data structures and allows the naive user to query and manipulate data intervals or individual points. Data displayed on the browser are dynamic. They can be "queried" to reveal detail or, in the case of graphics or sound, to initiate display or playing respectively (Figure 4). Other operations are similar to those of our graphical browser (Figure 3). The user can "zoom" or "roam" through the data. One can also "stack" segments of timelines or "restructure" into a calendar-like format. Finally, one can filter specific data types, and, with permission, add new data or modify existing data. Some of these operations are still in primitive developmental stages.

Our research in time-dependent medical reasoning promises to yield both applied systems (like the diabetes advisor) and sound theoretical contributions to computer science research. We also believe that our findings and our toolkit will have great use as a general tool for scientists and clinicians involved with exploratory analysis of time-oriented data.

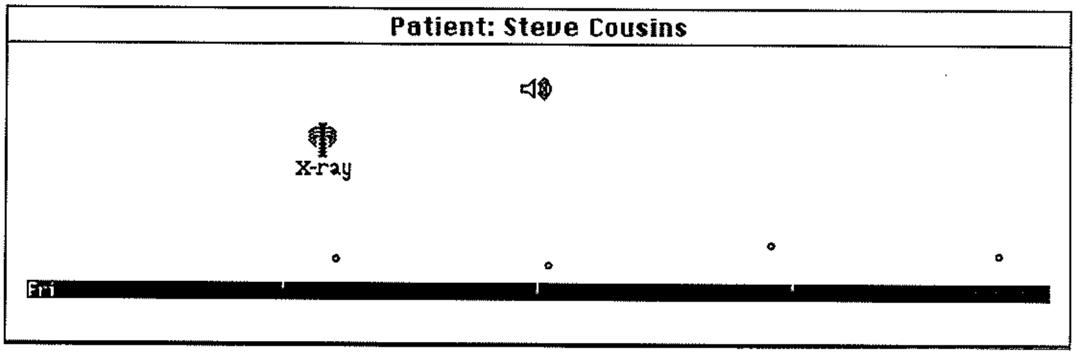


Figure 3: A "zoom" operation applied to the data in the previous figure.

This operation is useful when multiple data have been obtained within a short period of time.

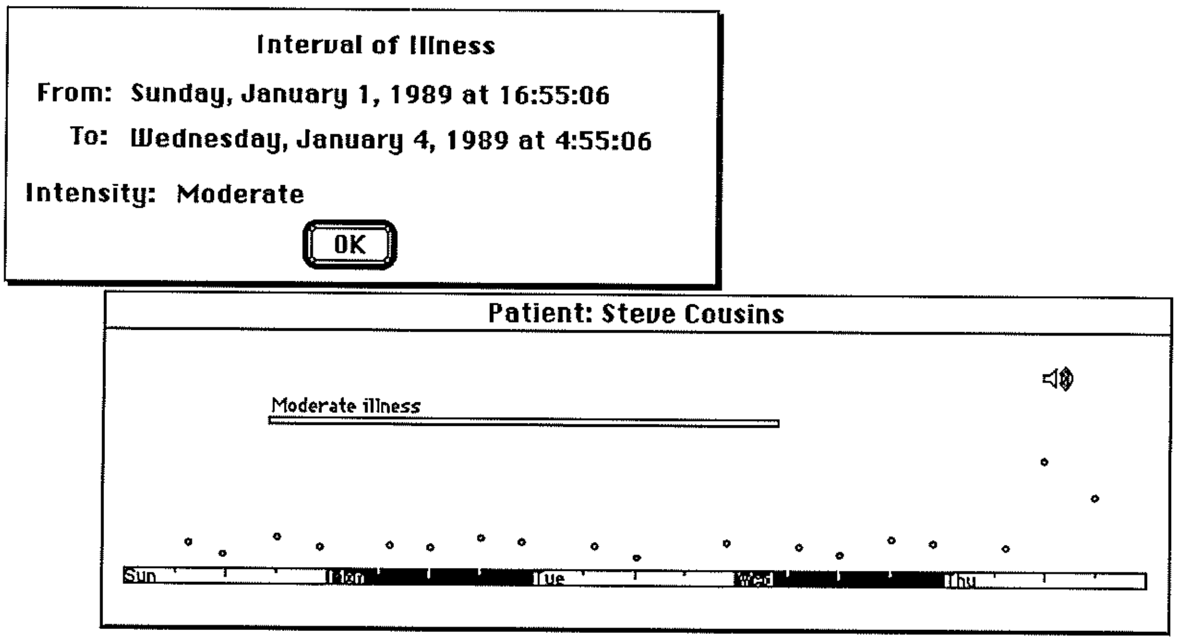


Figure 4: Querying a data point.

Double-clicking on a conventional data point leads to the presentation of a dialog box providing other details about the datum. These details can include numerical data, time stamps, comments, and relationships to other data elements.

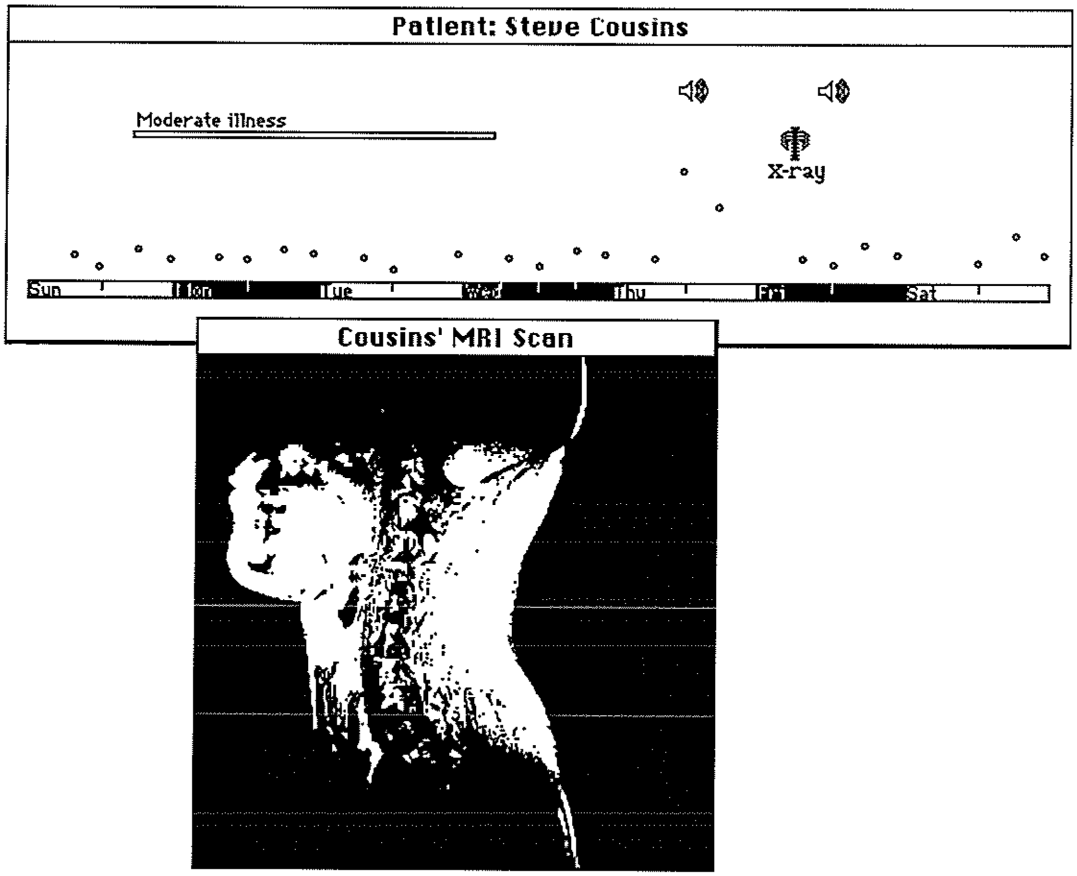


Figure 5: Querying a radiograph point.

Double-clicking on a radiograph icon leads to display of its associated bitmap image. In the future, we plan to add an "annotation" button to the x-ray display.

3 The Hypertext Manual Project

At about the time Vannevar Bush published his 1945 Memex article, the Department of Internal Medicine at Washington University began disseminating a small book detailing therapeutic techniques useful to trainees and practitioners of medicine. This book, the *Manual of Medical Therapeutics,* received immediate acclaim for its presentation of concise, problem-oriented medical information arranged in hierarchical (outline) format. When publication of the *Manual* was transferred to Little, Brown and Company in the mid-1960's, the format was enhanced by the strategic use of bold-face fonts and by the distribution of the text as a spiral-bound, pocket sized handbook. Since that time the *Manual* has remained one of medicine's most popular books.

The *Manual* presents many interesting challenges to the hypertext researcher. First, the medical community's rapidly changing opinions about disease treatment suggests the need for a **revision** cycle more frequent than the current three-year practice. Second, geographic and institutional differences suggest the need for local **modification** and **personalization** to suit the needs of diverse user groups. Third, the information retrieval tactics most frequently employed by the *Manual's* readers are **index-based lookup** and **browsing**–the book is not read sequentially. Finally, the editors are under continuous pressure to incorporate more medical information into the *Manual.* Any further increase in **size** would impact adversely on the *Manual's* **portability.** These constraints, dictated by the physical characteristics of conventional paper media, appear to limit the utility of the book.

The publishing industry shares with the hypertext researcher the awareness of hypertext's potential for overcoming many of the shortcomings of the traditional book.[2] But it should be noted that the motivations of our laboratory differ from those suggested by the publishing industry's forthcoming products. The publisher must be concerned with market share and short-term financial return. Conventional wisdom holds that this can best be achieved by embracing widely accepted software standards and production techniques. On the other hand, research groups such as ours have the luxury of adopting a more risky, long-term outlook. We envision more ambitious products because we believe that the need to store and transmit radiographic images in a rapid and cost-effective manner will make more powerful high-resolution graphics workstations an inevitable part of routine medical practice in the 1990's. Because these workstations will be present wherever medical problem-solving takes place, we believe that they will be an ideal vehicle on which to present medical information that currently is transmitted through medical textbooks and journals. We currently are using the Apple Macintosh II™ to develop medical hypertext in anticipation of the introduction of more powerful medical practice workstations.

[2]For example, Little, Brown and Company recently announced a CD-ROM Library called "MaXX." This library is composed of the Washington University *Manual,* a medical dictionary, and approximately ten other medical books from the Little, Brown "Spiral" series. One of us (M.E.F) serves on the editorial board for this effort.

There are two components of this project: producing a manual in a form that lends itself to hypertext - authoring, and finding ways to effectively get at information contained in the hypertext - retrieval.

3.1 The Manual Authorship Project

The Macintosh SE computers we received from Apple Computer were distributed to a number of physicians engaged in authoring and revising chapters from the upcoming edition of the *Manual of Medical Therapeutics*. Modems, printers, MacWrite (pre-Claris version), and MORE (version 1.1c) were made available to each author with a Macintosh. Although the computers were not delivered in time to conduct a formal study on authorship, we were able to gather meaningful qualitative impressions about the role played by current commercial authoring software packages. We observed some interesting trends in selection of software. Most authors immediately switched to the Macintosh (and many PC users ultimately purchased Macintosh for their offices and laboratories), but very few were satisfied with MacWrite. Most users preferred to use Microsoft Word (version 3.0.1), even though it was not provided with the computers. We believe this was due to "IBM compatibility," the sensation of more word processing power, and the presence of a dictionary spelling-checker. Few made use of the telecommunications capabilities, preferring instead to use the telephone and hand-carried 3.5 inch disks for information sharing.

We also were able to draw some conclusions about *how* people used the software. Some of the more mature authors began using the outline format of MORE, but most switched prematurely to Word. In almost all cases, the editors had to create a second outline while critiquing first drafts, but in no case did authors begin again with MORE. It appears that the format of MORE was alien and that the authors were frustrated by their inability to transfer the documents freely between Word and MORE. Also, most authors were unaware of Word's outlining capabilities.

We believe that most of these observations represent a minor indictment against the organizational and writing skills of our authors, but it is nevertheless quite clear that better software education and a more flexible and powerful groupware product will be needed for the formal study we intend to pursue during the 1990 revision of the *Manual*.

Perhaps our most interesting observations concern the software's adverse impacts on the *Manual's* editors. The clarity afforded by the "WYSIWYG" character of the Macintosh seemed to contribute to a premature emphasis on document physical appearance over manuscript content. Even if a submission was to be completely rewritten, the editors seemed to spend an inordinate amount of time changing fonts, styles, enumerations, reference numbers, and formats. In our opinion, much time would have been saved if the fine points of appearance were added either automatically or only when the manuscripts were in near-final form.

As a prelude to future studies, we adapted a locally created graphical browser for use as a visual editorial tool. By parsing hierarchically organized drafts and displaying them as trees, we immediately were able to note interesting visual characteristics of individual chapters. Early drafts from some of the poorer chapters led to trees that were very broad and shallow (Figure 6). After significant editorial revision, the resulting trees were far deeper and less broad (Figure 7). Visually, the diagrams are more appealing. Certainly the editors must feel that the chapters are better "reads" as well!

We suggest the need for a tool that combines the outline facilities of MORE with the power of our graphical browser and the dictionary/thesaurus capabilities of recent text-editing tools. We also suggest the need for tools to exploit the structural organization of a text rather than the appearance (Our browser adaptation is a primitive example). A visual equivalent to SGML or LaTeX commands seems to be an appropriate foundation. We favor SGML because of the contribution this standard can make to post-authorship editorial revision, information retrieval, and text display. We hope to have these tools available in time for the next revision (1990).

3.2 Theoretical Aspects of Hypertext Information Retrieval

We believe that one of our most significant contributions may be in the area of information retrieval from hypertext documents. Our goal is to develop a sound hypertext information retrieval vehicle which allows automatic browsing assistance through incorporation of information gained through user feedback. At the present time, our research is primarily theoretical, but we believe that our findings will have significant practical implications for the design and use of general purpose hypertext systems.

3.2.1 Two Components of the Hypertext Information Space

The hypertext information space is composed of two logical components, the index space and the document space (Figure 8). The index space can be "flat," "hierarchical," or "networked." The document space can also be "flat," "hierarchical," or "networked." Flat components imply strong independence both within and between individual terms and individual documents. This independence allows rapid computation. Examples of flat spaces include inverted indexes, keyword search, and HyperCard's signature files. Hierarchical components imply conditional dependencies between parent and child nodes. Computation is more complex in these spaces, but remains polynomial in the worst case. The National Library of Medicine's MeSH headings system approximates a hierarchy, but computation is performed as if it were a flat index. Networked components imply complex dependencies among index terms and individual documents. Networked document spaces are usually navigated through browsing. Connectionist models essentially combine a networked index space with a networked document space. Networked and connectionist models remain highly experimental. Computational complexity currently precludes their routine use.

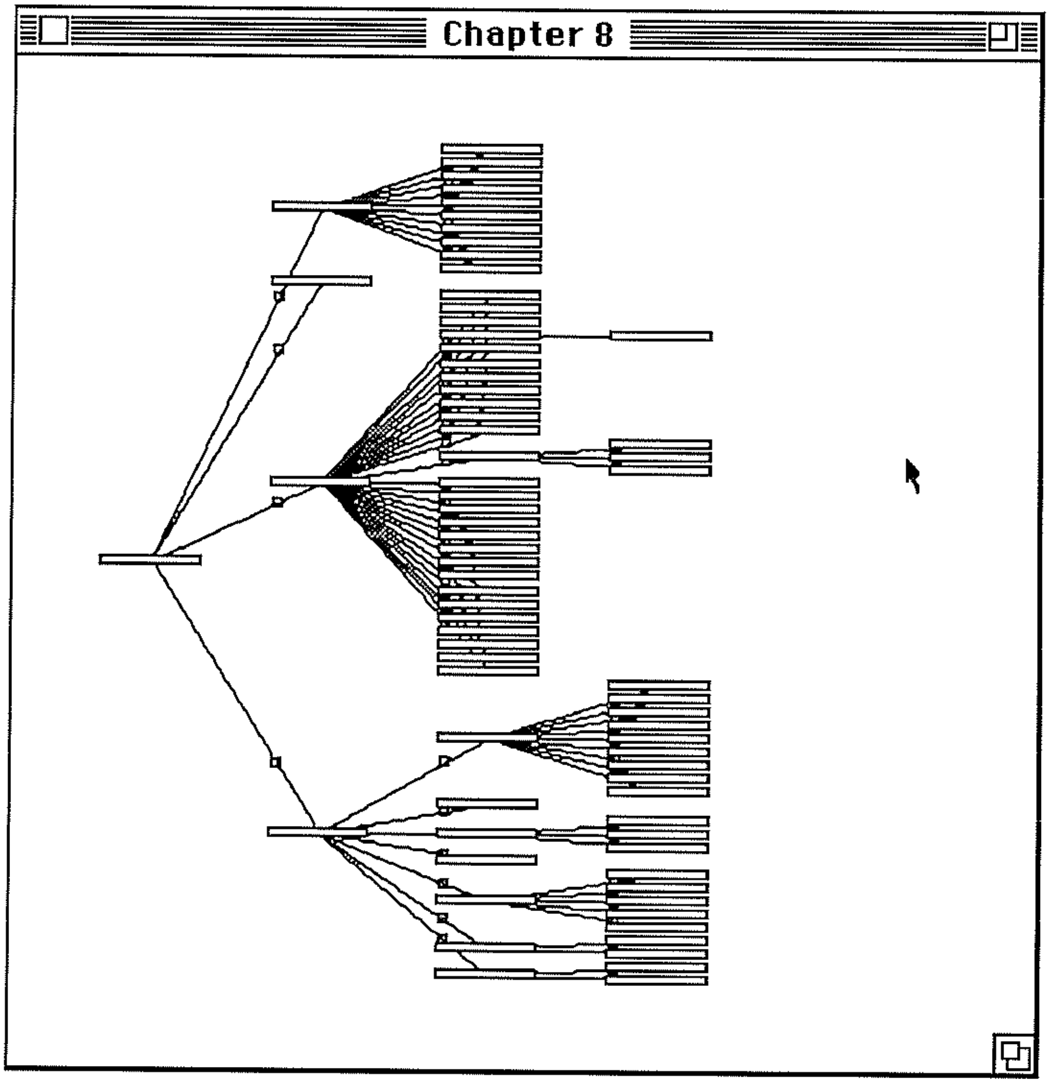


Figure 6: A poorly-organized early draft of a *Manual* chapter.

In the *Manual*, the usual non-leaf-node has approximately six children. In this chapter, one node (about Advanced Cardiac Life Support) has 27 children and only 4 grandchildren.

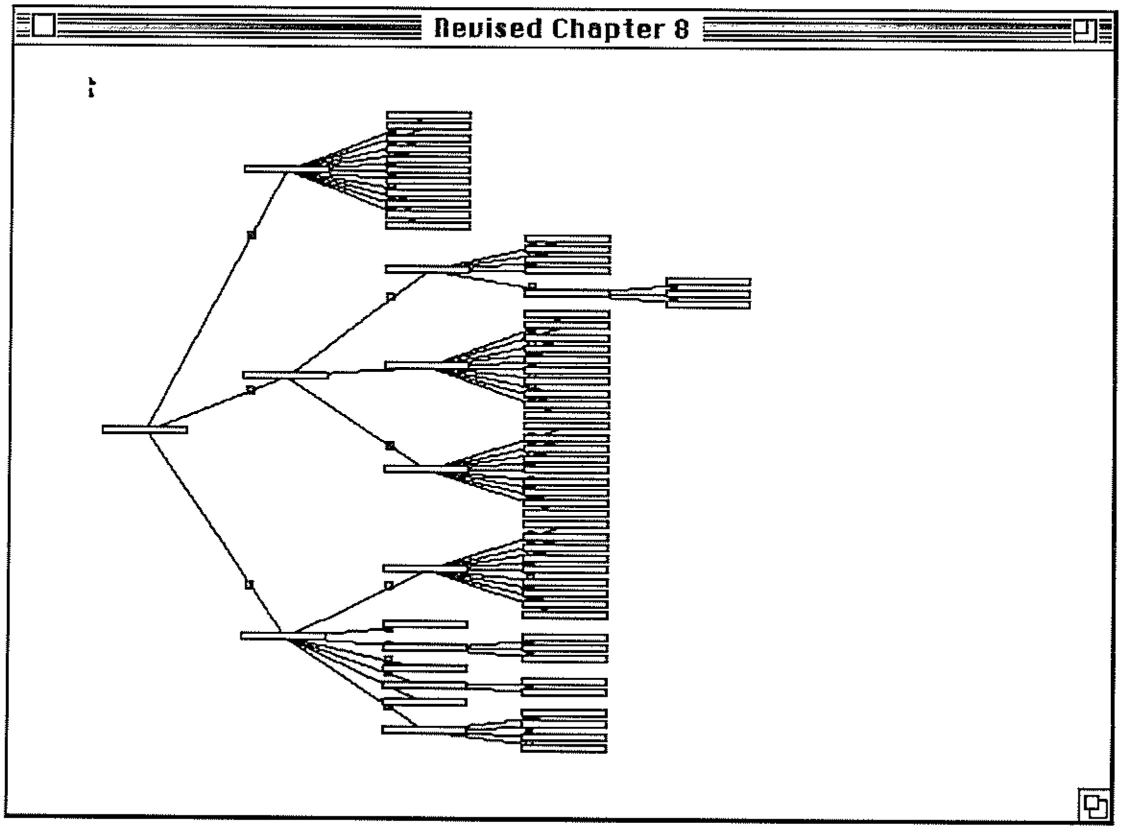


Figure 7: A revision of the manuscript outlined in the previous figure.

The Advanced Cardiac Life Support section has been divided into 3 sub-sections (Primary components, Adjunctive Therapies, and Post-Resuscitation Management). On average, each subsection now has more parts. This reorganization has striking aesthetic effects on the chapter outline.

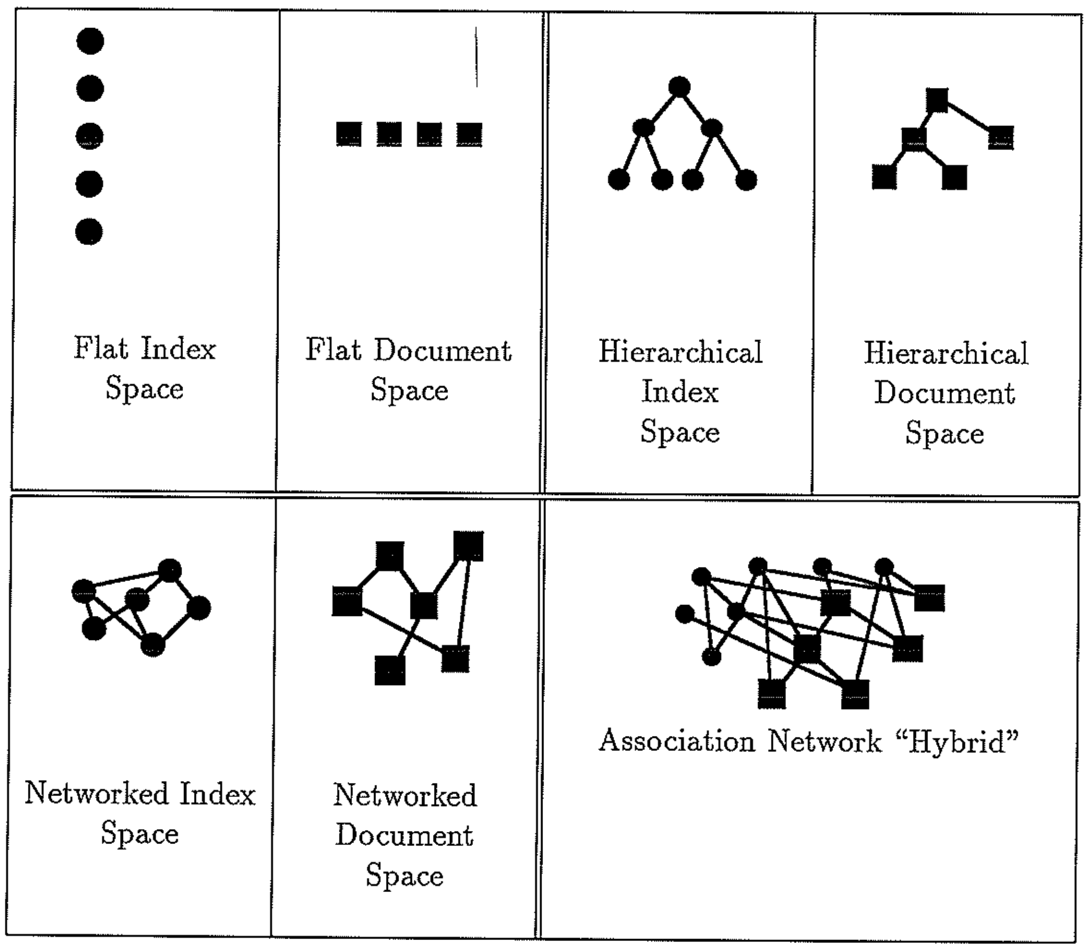


Figure 8: Two Components of the Hypertext Information Space.

The index space is a data structure containing keyword terms used to access the document space. Each component's data structure can be varied according to the degree to which the designer wishes to impose an organizational framework.

3.2.2 Two Metaphors for Information Retrieval from Hypertext

There are two principal metaphors applied to information retrieval in hypertext. In the first, traditional full-text document retrieval techniques are applied to cards (e.g. Hypercard's "find" command). We call this metaphor the *small-document approach.* This perspective tends to de-emphasize the role of links and emphasize the static structure of a hypertext document. Traditionally, both the index space and the document space are "flat" in systems emphasizing this approach.

Graph traversal (or browsing) is the second metaphor applied to hypertext information retrieval. Browsers reveal hidden cards by facilitating the selection of active labelled link icons located on individual cards. These labelled icons represent the title of a hidden card, a summary of a hidden card's contents, or the semantics of the link between the current card and the hidden card. Although the labels associated with the link icons provide the reader with strong hints about the relevance to his inquiry of their corresponding cards, an accurate assessment of card relevance can be made only after activating the link icon and examining the contents of its corresponding hidden card. No matter how we search for a hidden card, browsing is a "hit or miss" proposition. The true utility of a card to our specific query can only be assessed reliably by direct examination. The browsing document space is networked, but complexity is reduced by restricting the availability of link icons to only those of neighbors immediately adjacent to the card under examination.

Neither the browsing approach nor the graph traversal approach incorporates both a networked index space and a networked document space, suggesting that neither browsing nor pattern-matching techniques will "scale up" adequately to large medical information spaces. If that is in fact the case, it might be relatively easy to "navigate" in a familiar "hyperhandbook," but we will find ourselves lost when we enter an unfamiliar "hyperlibrary."

Our very early experience with the development of a prototype hypertext medical-therapy handbook suggested that our concerns about "scaling up" are warranted even in relatively small application domains. Users of our initial prototype hypertext handbook have great difficulty identifying the small number of useful cards from among the relatively large number of cards retrieved from a typical query. Details of our heuristics and first prototype implementation have been published elsewhere [9].

We currently are using chapters from the upcoming twenty-sixth edition of the Washington University *Manual of Medical Therapeutics* for our prototypes. We use both Apple Computer's HyperCard™ and Coral LISP programs as our primary hypertext development vehicles. We employ a combination of text pre-processing tools developed on a UNIX™ system and a parser (written in LISP) to transfer individual sections of the *Manual* into a hierarchically-structured hypertext data structure [8]. In the HyperCard versions links are "go to" statements. In the LISP version built on the graphical browser, links are "bi-directional" separate objects with programmer-defined properties. Our HyperCard prototypes make extensive use of the "Find" command. Our LISP prototypes incorporate string-search utilities similar to those found in the IRX workbench of the National Library of Medicine.

3.2.3 Operations Performed on the Document Space–Card Weight Propagation

Hypertext systems should respond to queries by providing the user with one or more useful starting points for graphical browsing. We believe that this process can be facilitated in hierarchically structured documents if one recognizes that the utility of any individual card is affected both by its relationship to the query terms (we call this contribution to the total card utility the *intrinsic weight*) and by its relationship to all other cards within the hypertext document (we call this utility contribution the *extrinsic weight*).

We are testing the card weight propagation hypothesis by representing the syntactic hierarchical organization of our document with a specific semantic link type. We use these syntactic relationships to propagate query-generated card weights through the hypertext document. This explicit representation of syntactic relationships among cards allows us to identify parent cards that might not contain any search query terms, but which might nevertheless be ideal starting points because many of their immediate descendant cards contain the query terms.

Our prototypes emphasize the concept of term weight propagation (i.e., a flat index, hierarchical document approach). We calculate the intrinsic card weights by applying a modification of a simple, well-known algorithm promulgated by Gerald Salton. This algorithm assigns term weights to cards as a function of the frequency of occurrence of the term in the entire search space and as a function of the number of cards containing the term. This algorithm assigns a higher weight both to cards containing infrequently used terms and to cards containing several occurrences of a term not found in many other cards.

The intrinsic card weight component is combined with the extrinsic component due to weights of immediate descendant cards through simple summation and division using the following formula:

$$TOTALWEIGHT_i = \sum_j WEIGHT_{i,j} + \frac{1}{y} \sum_d WEIGHT_d$$

Where y is number of immediate descendants of card i containing term j and d is an immediate descendant of card i.

This propagation function is called recursively from the leaf cards to the root card. After card weights are calculated and propagated, the program creates a list of cards which may be good candidates for graphical browsing. The weights of each card are displayed. In some instances cards displayed on the list do not contain any of the actual search terms, but are presented to the user because they have extremely large weights because of their relationships to cards that do contain the search terms.

Although we believe that hypertext information retrieval routines based on card weight propagation will improve our ability to retrieve information from medical handbook chapters, there are two reasons why we believe that card weight propagation methods alone will not provide a satisfactory solution to the more generalized problem of hypertext search. First, these systems are highly dependent upon target document

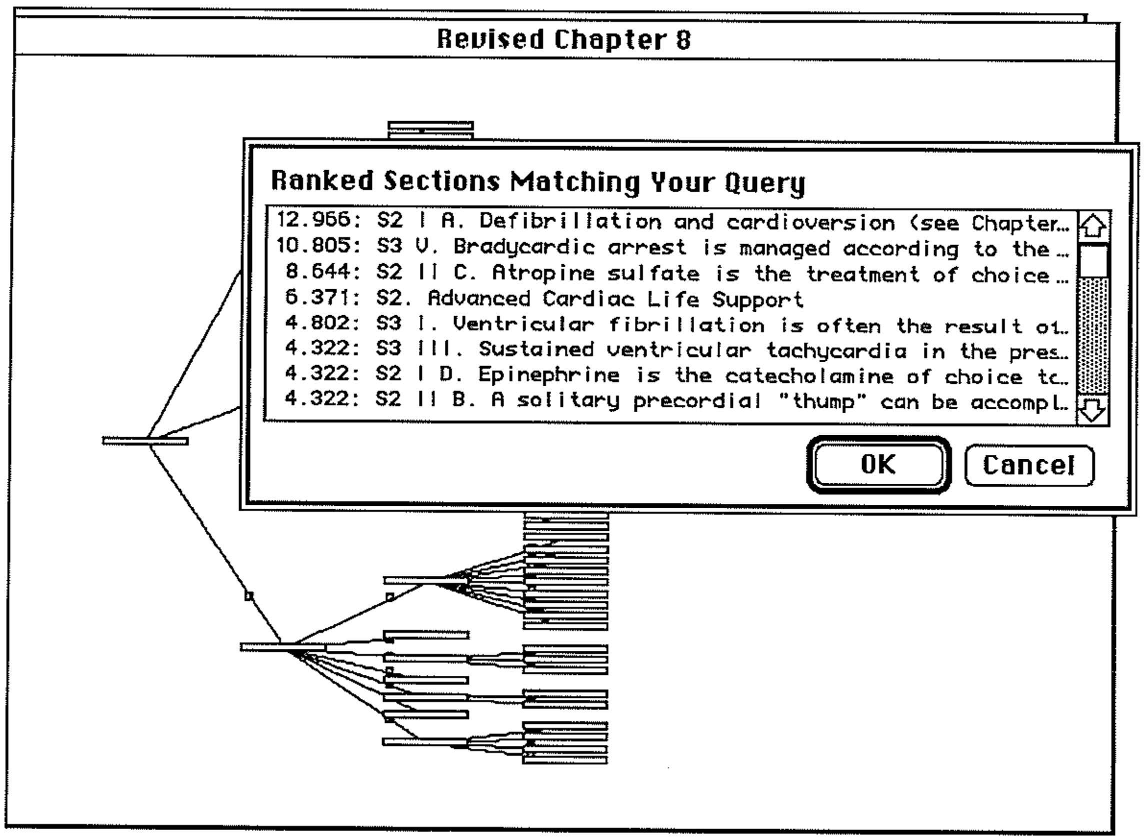


Figure 9: A Card Displaying Browsing.

After searching and card weight propagation, the program produces a window displaying the leading candidates for graphical browsing. Each card is denoted by a weight and a title. Double clicking over any portion of an entry will produce the card identified by the title.

associations between words. As the size of a document decreases (as it does when comparing a scientific article with a single hypertext card), we believe that there will be a lower probability of finding related words within the same document. This will result in lower retrieval sensitivity. Second, these systems do not "learn" from the user's response to card reading. These concerns initiated a study of Bayesian updating and connectionist models for information retrieval. Presently, most of our work concerns the application of Bayesian belief networks to hypertext information retrieval.

3.2.4 Operations Applied to the Index Space–Belief Networks

We believe that **belief networks** will improve search by retrieving cards indexed under concepts related to those requested in a reader's query. Belief networks represent conditional dependencies between nodes in an acyclic directed graph. The presence of an edge between two nodes implies a conditional dependency between the nodes. Lack of an edge implies that a significant conditional dependency does not exist. Belief in a node is derived from **predictive** support (i.e. the probability that the node is true given that the parent node is true) and **evidential** support (i.e. the probability that the node is true given that a piece of evidence is true). Evidential support can come from inference about other assertions and through direct observation (e.g. "virtual" evidence). In the network, these probabilities can be expressed in terms of the degree of belief in parent node(s), degree of belief in child node(s), degree of belief contributed by "virtual" evidence nodes, and by the conditional probability matrices representing relationships between parent and child nodes. Implementation can be achieved through message passing, and calculations can be performed as an asynchronous process, making inference in some networks (e.g. trees) extremely amenable to parallel computation using multiple processing units.

3.2.5 A Belief Network-Based Index: Design and Implications

There appear to be at least two different approaches to using belief networks to "learn" from hypertext browsing patterns. The first describes a set of externally declared concepts related to one another by probabilistic statements. The second approach ignores explicit declarative representations and instead uses as the framework for probabilistic reasoning the links between hypertext cards and query statements. Our current research emphasizes the former approach. Speculations about the latter approach are discussed in Section 3.2.8.

Our current prototypes employ a hierarchy of medical concepts represented as a directed tree. The most general concept is at the root of the tree, and less general concepts are at the leaves. The semantics of the edges generally are of type "is-a-subdomain," but any similar logical implication can be represented. For example, the root node might be "pets" and two children might be "dog" and "cat." The probability statement represented by the matrix between the "pet" and "dog" node can be crisply explicated as follows: Given that the reader examines cards suggesting the concept of 'pets', the probability matrix denotes the degree of belief of the reader's response to cards about the concept of 'dog' given the readers response to cards suggesting the concept 'pets'. An identity matrix (i.e. 1's along the matrix diagonal) implies that if the reader likes cards suggesting the concept 'pets', she will always like cards suggesting the concept 'dog', and if she does not like cards suggesting the concept 'pets', she will never like cards suggesting the concept 'dog'. A network containing only two nodes connected by an identity matrix is identical to a thesaurus entry table where the parent would be the key term and the child would be a synonym. The network also has the advantage of allowing relative inference between concepts (i.e. "usually similar") and equivalent terms. In a typical belief network, terms would simply be assigned to the same node.

The required probabilistic relationships can be elicited by collecting actual data obtained from a relatively homogeneous group of readers from a collection of indices from medical textbooks (our current approach), or they can be elicited from domain experts. In contrast with association networks, the probabilistic relationships of a belief network are not changed dynamically. Instead, the dynamic computation alter's one's degree of belief (i.e., the value of a node) as new evidence is acquired.

3.2.6 Updating the Belief Network Index

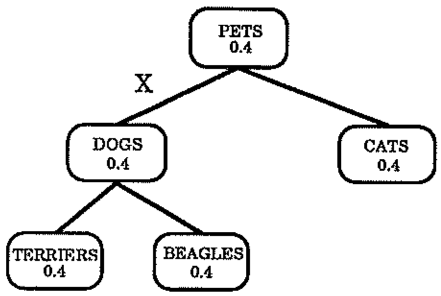
We believe that the true power of incorporating belief network techniques into hypertext information retrieval will come as a result of a network's ability to refine a reader's query automatically, just by "observing" a reader's reaction to cards examined during a traditional browsing session. If a reader does not like a card, there is a significant probability that she will dislike other cards structurally adjacent to the examined card, and there is also a significant probability that she will dislike cards which are classified under the same concepts or keywords but which are not within the structural vicinity (Figure 10). The latter claim can be illuminated by an example. If the reader has just disapproved of a card detailing the life of Franklin Roosevelt's terrier, it seems highly likely that he also will not like cards located in other stacks that also are about terriers. Under the relationships suggested by our concept belief network, it also then seems likely that he will be strongly disinterested in cards about dogs, weakly disinterested in cards about pets, and, depending on other evidence, weakly disinterested in cards about cats (Figure 10).

If our reader then examines several cards about cats and approves of them, the network must increase its degree of belief that the reader is interested in cats. This increase will lead to some increase in belief that the reader is interested in pets (Figure 10). The degree to which each belief changes depends on the prior degree of belief and the strength of the new evidence.

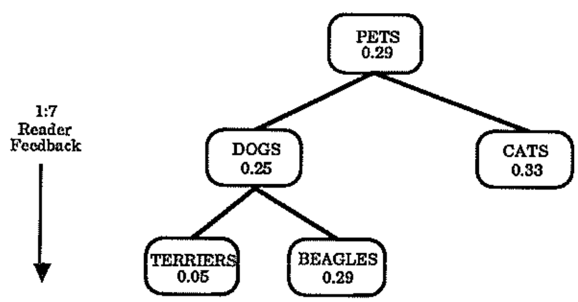
In a belief network, the degree of belief in a given node can be altered by a change in degree of belief of a parent node, a change in degree of belief in a child node, or by new evidence obtained by a user's response to cards indexed under the concept denoted by the node. The updating process is sound, consistent, and can be performed asynchronously. Figure 11 illustrates a belief network index card from an early prototype.

3.2.7 A Typical Iteration

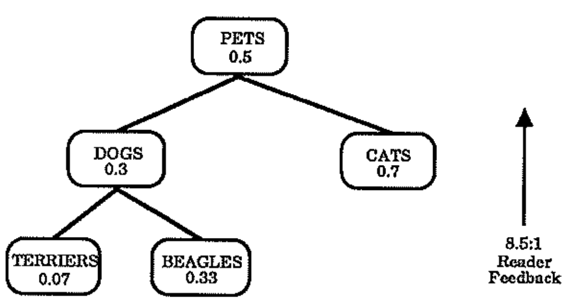
It is apparent that reader feedback can be employed to alter degrees of belief in the network, and that these changed beliefs can be exploited to identify new cards which have a high probability of being useful during the current reading session. Although we are examining many variants of this principle in our research, we currently emphasize hierarchical index and document spaces (Figure 12). We now outline a method that seems to produce reasonably intuitive results.



Both the presence and the absence of links has significance. The presence of a link denotes a significant relationship between the nodes to which it connects. For example, link "X" implies that there is a relationship between a reader's interest in the concept "dogs" and the reader's interest in the concept "terriers." Also, because the only path between "pets" and "terriers" is is through "dogs" (i.e., there is no direct link between "pets" and "terriers"), the semantics of the network imply that if one has complete knowledge about one's belief in "dogs", one can calculate one's belief in "terriers" without knowledge of belief in "pets." In belief network terminology, the node "dogs" blocks the node "terriers" from the node "pets." In this example, the degree of belief is uniform (0.4) throughout the network.



The reader has just expressed strong disapproval of several cards about terriers. Barring other evidence, this leads to a decreased degree of belief in all other nodes in the tree, especially dogs and beagles.



The reader then expresses strong like of several cards about cats. This new evidence increases the degree of belief that the reader is interested in cats, and it also updates the degree of belief that the reader is interested other pets (including, to a mild degree, dogs). The absolute change in degree of belief of the node "pets" will depend on the degree of evidence submitted to "cats" and "terriers," the initial values for the nodes, and the probability matrices represented by the involved edges.

Figure 10: Updating a simple belief network about pets.

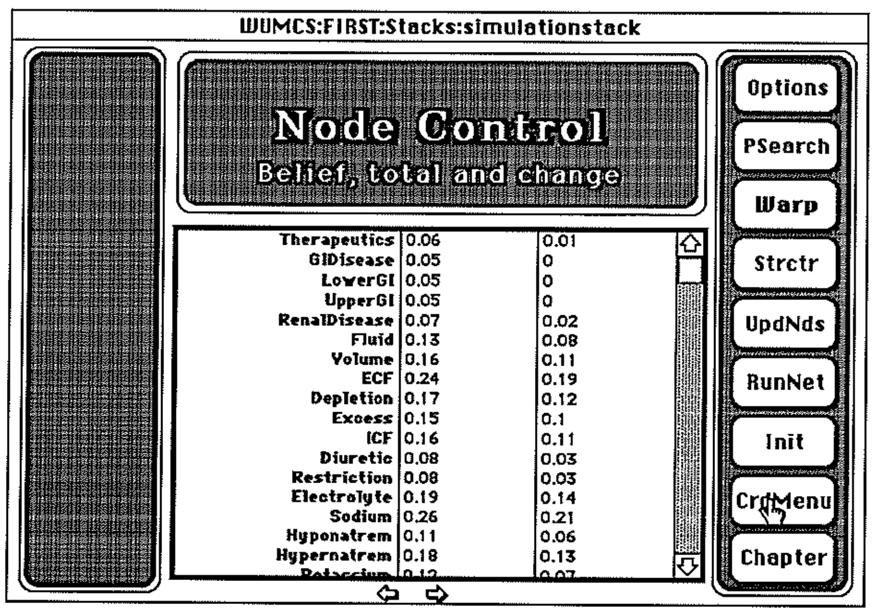


Figure 11: A listing of nodes in a belief network index.

The first column lists the index terms, the second column lists the absolute degree of belief in each node, and the third column lists the "adjusted" change in degree of belief caused by recent feedback from hypertext browsing.

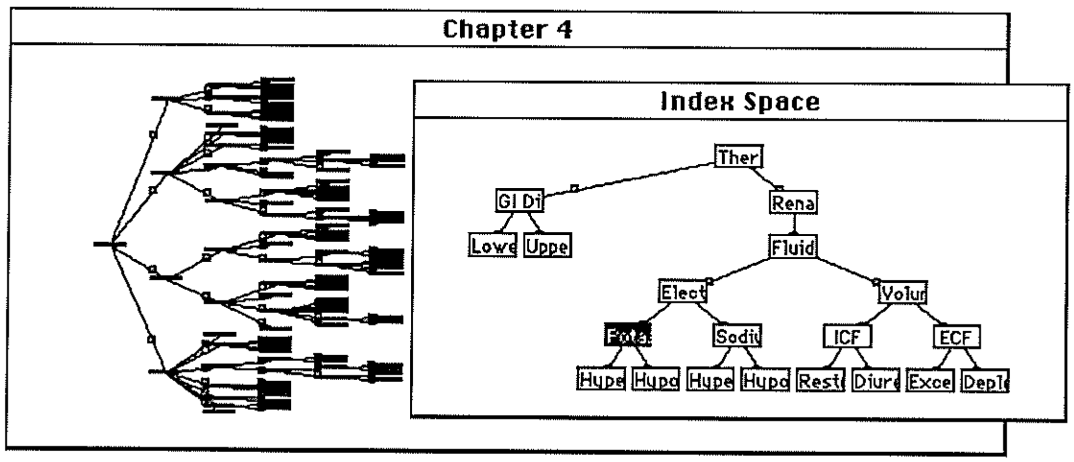


Figure 12: Hierarchical components of a prototype *Manual.*

The left window (entitled "Chapter 4") is a tree from a *Manual* chapter. Each rectangle represents a separate hypertext card. The right window (entitled "Index Space") is a simple belief network denoting relationships among index terms. Feedback obtained from reading cards accessed through the left tree is used to update degrees of belief in index terms represented by nodes in the tree on the right.

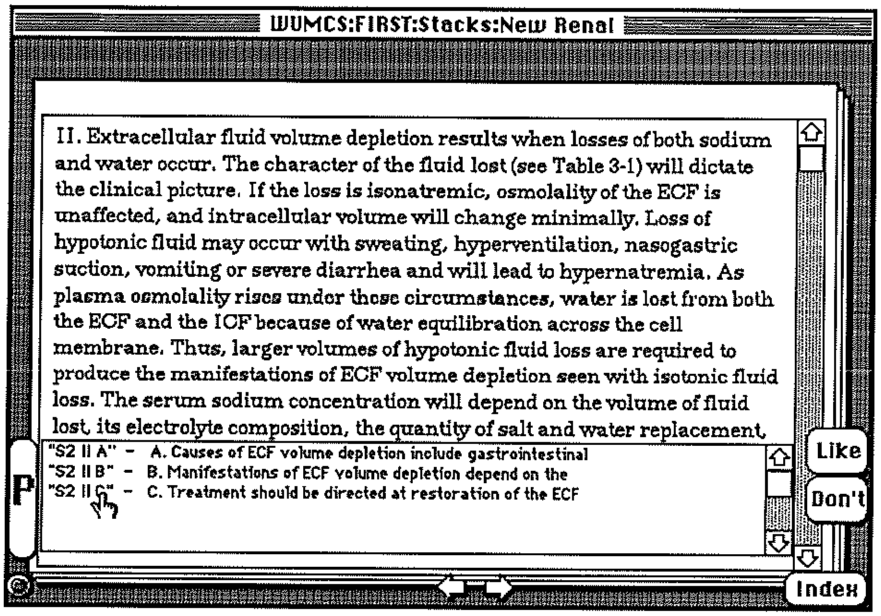


Figure 13: A hypertext medical handbook card from a prototype HyperCard implementation.

Readers can navigate by going to the next card or the previous card (arrows), the parent card (the button labelled "P") or any of the children card (scrolling menu). All links between cards are created dynamically at the time the chapters are parsed from the conventional paper text.

When beginning a session, the reader uses a "fresh" network (either a network in which all nodes have the same degree of belief or a network which has been "tuned" by the usage patterns of colleagues). The reader either initiates a term-weight query (e.g., aortic valve .6 and surgery .3) or she simply begins reading at a convenient place in the hypertext. In the former case, each card is assigned a "belief" proportional to the presence of query terms. In the latter case, a uniform "belief" is assigned to each read card. As the user reads a card, she has the option of responding to what she has read either by selecting a "Like" button or by selecting a "Don't Like" button (Figure 13). She also can also move to another card without selecting either button. If a feedback button is pressed, a pair of keyword term - card weights is generated for each keyword in the card. If the feedback button was "Like," these data structures are added to a global "GoodGuys" list. If the feedback button was "Don't Like," the data are added to a global "Bad Guys" list.

When the user seeks advice on new cards (query reformulation), the program collects the "GoodGuys" and "BadGuys" lists. For each keyword in the list, the corresponding node in the belief network is accessed. Weights from the "GoodGuys" list are added to the "true" field of the feedback odds-likelihood value for the node (i.e.

the virtual evidential support). Weights from the "BadGuys" list are added to the "false" part of the node's feedback odds-likelihood value. This feedback constitutes new evidential support for a concept node. A change in evidential support mandates an update in belief of both the concept node and all nodes related by probability links.

The changes in the network brought about by user feedback change the degree of belief assigned to each node. If the network has been used by many people, the numbers in the "true" and "false" feedback lambda fields are probably large, and any feedback will only change the degree of belief by a small degree. But the question asked in reformulating a query is not "what is the absolute degree of belief in a concept?" but "how has the degree of belief changed as a result of feedback?" When asked for "help" the program "adjusts" the change in degree of belief by assigning a value of zero to the smallest change value, and subtracting this change from each other change value. For each concept node, the program assigns to each card indexed under the concept the node's "adjusted" change in degree of belief . If a card is indexed under more than one concept (usually the case), the change values assigned to a card are combined using Pearl's "noisy OR gate" [11]. This is a monotonicly increasing function that ranges over the probability values between zero to one. Cards are then ranked in descending order and a list of candidate cards is presented to the user. A list of candidate cards from a prototype HyperCard™ implementation is displayed in Figure 14.

3.2.8 Limitations of the Index Hierarchy

One of our major research challenges will be to develop an appropriate index space for use with our medical hypertext retrieval software. We currently are adapting the 15,000–term MeSH classification system for use with our hypertext system, but we believe that the MeSH terminology is both too extensive (i.e. many terms not applicable to our domain) and too coarse (i.e. terms are not sufficiently detailed to distinguish between cards). Alternatives include a semi-automatic structuring of the index terms used in the *Manual* series or adapting a proposed Unified Medical Language standard currently under investigation by the National Library of Medicine.

Early in our investigations, we proposed a belief network structure that combines the index space with the document space. Each card is assigned a to a unique node in the network. The probability statement associated with a card node refers to the degree of belief that the card node in question will be useful to the reader's current line of inquiry. Each keyword is also assigned to a unique node. Keywords are linked to card nodes which contain the keyword. Probability matrices within keyword-card node links are a function of the frequency of each keyword both within a specific document and within the entire document space. Card nodes are connected to one another with card-card links. These links also can be represented by a node if one wishes to ascribe different probabilistic dependencies as a function of the semantics of a link type (Figure 15).

At the current time, we are not assured of the soundness of this configuration, but

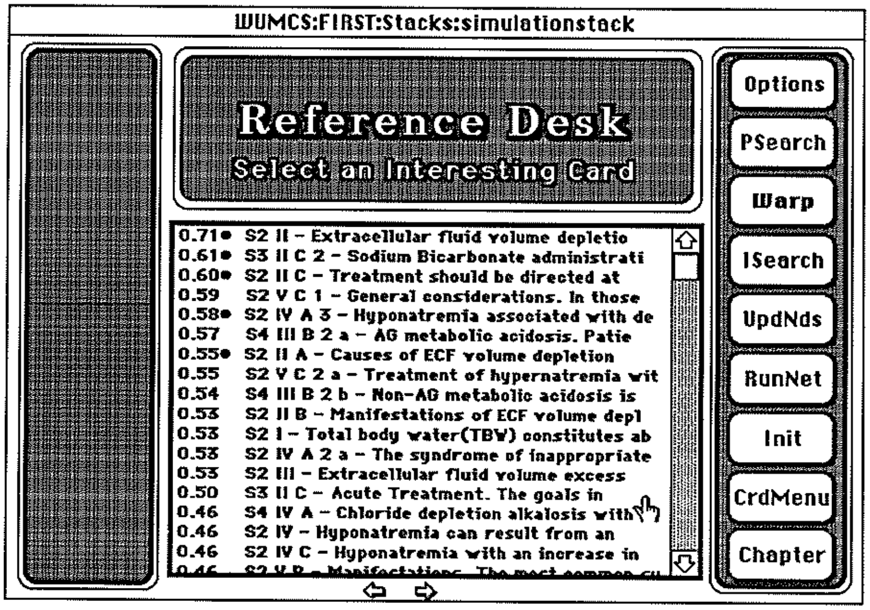


Figure 14: An index of candidate cards for hypertext browsing.

The cards are listed in decreasing order by their "weights" or "degrees of belief that a card will be useful." The first term is the weight, the second describes the cards position in the hypertext hierarchy, and the remaining terms are the initial words in the card. The latter serve as ersatz labels to facilitate browsing. The symbol "•" is used as a place-marker. The presence of this symbol implies that the card already has been examined (a technique modified from Bernstein [3]).

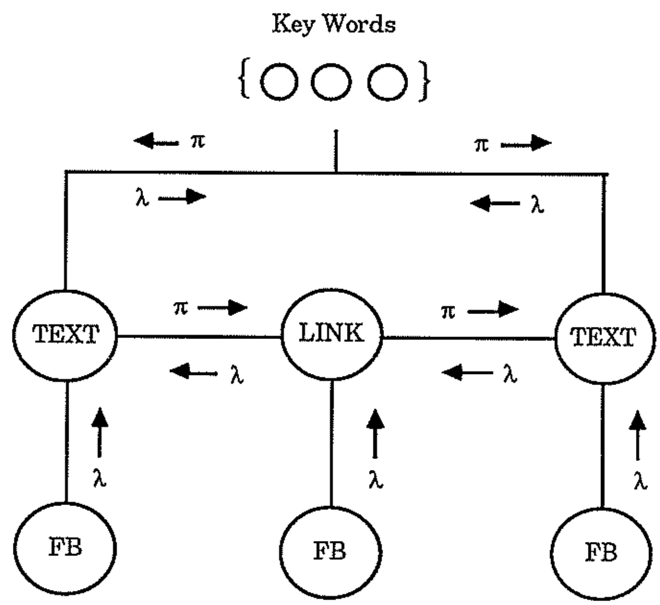


Figure 15: A more complex belief network system with marked similarities to an association network (connectionist) approach. Circles represent cards, links, and user feedback (FB). Degree of belief in links can be set globally or individually.

we are struck by its similarity to work independently done by Belew at UCSD [1]. He uses an association network to represent a similar set of dependencies between keywords and card nodes. Like Belew's work, computational cost is extreme. In fact, we do not believe that this belief network can be computed in our lifetime! We estimate the complexity to be an exponential function of the size of the cutset (i.e., at least 2^{1000}). Pearl [12], Chin and Cooper [5], and Chavez [4] have proposed tractable estimation techniques for networks of this complexity. Pearl's network has poor performance when probabilities are low (as they would be in this case). We plan to investigate further both the Chin and Cooper and the Chavez approach. We will begin by applying their techniques to smaller networks (modifications of our hierarchical scheme) before even considering implementation of this "hybrid" networked configuration.

3.3 Intellectual Challenges Engendered by Hypertext Research

Our research presents us with many exciting intellectual challenges. Our medical authoring projects will require us to develop a formal structural language for the *Manual* and related texts. We believe this language should be implemented by adopting the increasingly popular SGML standard. Our implementation work will require us to review available software and hardware options for delivering both highquality medical images and medical text. We will work closely with the Mallinckrodt Institute's Electronic Radiology Laboratory (ERL) in this effort.

Our data management research will require us to develop a formal language for the display and summarization of temporal data. We expect that this research will carry us deeper into research concerning mathematical models, qualitative models, statistics, rule-based expert systems, and natural language techniques suitable for summarization of medical phenomena.

Our information retrieval research will require the development of a sound, computationally tractable method for updating complex belief networks. We believe the work of Pearl, Chin, Chavez, and Cooper will provide an excellent starting point for our theoretical efforts. To apply our theory to practical issues we will have to develop a better methods for generating index term graphs, better techniques for automatic classification of hypertext cards, and improved approaches for the evaluation of our many prototype information retrieval systems. Ultimately, we will have to explore methods of sufficient performance to allow the introduction to clinical settings of advanced, real-time hypertext systems.

Within the next several years we propose to implement and evaluate an SGML-based medical authoring "groupware" system. As a part of our hypertext authoring research, we also propose to create for the Macintosh a powerful version of Little Brown's "MaXX" medical library. As a part of our data summarization research, we propose to develop a prototype medical data summarization system suitable for commercial development.

Our hypertext research will demand that we develop more powerful parsers for converting conventional documents into hypertext. We also propose to develop more sophisticated automatic classification systems for application to medical hypertexts, and we plan to develop a full-scale medical hypertext prototype suitable for evaluation of advanced information retrieval strategies.

4 Ongoing Research Efforts - Common Elements

The apparent diversity reflected in our research domains belies a number of crucial unifying themes of theoretical and practical significance. We finish this report with a brief description of some of these issues.

From a theoretical viewpoint, each of our projects is concerned with the generic problem of knowledge representation and inference. Our diabetes work, for example, is concerned with the application of qualitative and quantitative models to develop high-level summarizations of low-level temporal entities. Our hypertext research is concerned with the relationships between free-text and "concepts" used by professionals engaged in problem-solving activities. Among the questions we ask are:

- How does one design and evaluate new methods for displaying and manipulating biomedical data?
- How does one know whether or not hypertext information systems confer an advantage over traditional full-text document retrieval systems?
- How does one recognize a "clinically acceptable" biomedical model?
- When are artificial intelligence techniques necessary to the solution of applied biomedical problems?

Our research also uncovers many unifying applied elements. For example, most of our implementations focus on a concept of a high-resolution graphics "workstation". Manipulation of elements displayed in the workstation require the effective application of the object-oriented programming paradigm. Using this paradigm, we create general-purpose programs useful in the design of model-based data summarization programs, electronic medical publishing aids, and hypertext information retrieval engines. As our work evolves, we are generating a number of valuable software "toolkits" easily specialized to meet a wide range of biomedical applications.

We find the intersection between biomedicine and computer science to be an exciting area for basic and applied scientific work. We encourage interested parties to contact the authors for additional information concerning ongoing research activities within the Medical Informatics Group.

References

[1] RK Belew. Adaptive information retrieval: Machine learning in associative networks. PhD thesis, University of Michigan, Ann Arbor, Michigan, 1986.

[2] LM Bernstein and RE Williamson. Testing of a natural language retrieval system for a full text knowledge base. *Journal of the American Society for Information science*, 35:235–247, 1984.

[3] M Bernstein. The bookmark and the compass: Orientation tools for hypertext users. *SIGOIS Bulletin*, 9:34–46, 1988.

[4] RM Chavez and GF Cooper. A fully polynomial randomized approximation scheme for the bayesian inference problem (working paper). Technical report, Stanford University, Stanford California, Fall, 1988.

[5] H Chin and G Cooper. Stochastic simulation of bayesian belief networks. Technical report, Stanford University, Stanford, California, Fall, 1988.

[6] SB Cousins and JA Fingerhut. A graph browser with zoom and roam for Allegro Common Lisp. Technical Report WUCS-88-19, Washington University, St. Louis, MO, October 1988.

[7] JR Cox Jr., SM Moore, GJ Blaine, JB Zimmerman, and GK Wallace. Optimization of trade-offs in error-free image transmission. In *SPIE Conference on Medical Imaging III (in press)*, Newport Beach, California, 1989.

[8] ME Frisse. From text to hypertext. *Byte*, 13, 1988.

[9] ME Frisse. Searching for information in a hypertext medical handbook. *Communications of the ACM*, 31:880 – 886, 1988.

[10] PL Miller. *A Critiquing Approach to Expert Computer Advice: ATTENDING.* Pitman, Boston, MA, 1984.

[11] J Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann, San Mateo, CA, 1988.

[12] J Pearl. Evidential reasoning using stochastic simulation of causal models. Technical report, University of California, Los Angeles, Los Angeles, California, September, 1986.

[13] BR Schatz. Telesophy. Technical Report Bellcore TM-ARH-002487, Bell Communications Research, 445 South Street, Morristown, NJ. 07960-1910, 1985.

[14] S.L.Lauritzen and D Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *J. Royal Statistical Society*, 9:157–224, 1988.

[15] JB Zimmerman, SB Cousins, ME Frisse, KM Hartzell, and MG Kahn. A psychophysical comparison of two methods for adaptive histogram equalization. In *submitted to the SPIE Conference on Medical Imaging III*, Newport Beach, California, 1989.