Division of Research
School of Business Administration

March 1988

USER INDEPENDENCE IN HYPERTEXT

Working Paper #555

Manfred Kochen
Moonkee Min
Christopher Westland
The University of Michigan

USER INDEPENDENCE IN HYPERTEXT

Manfred Kochen, Moonkee Min, Christopher Westland
The University of Michigan

Abstract: The concept of independence, as understood in database systems, is extended to hypertext systems. Experience in the classroom use of an adaptive encyclopedia system is presented to clarify the trade-off between cognitive coupling to individual users (accessibility) and the greater functionality stemming from user-independence. This system is viewed as a prototype hypertext model that is analyzed, and for which optimal degrees of independence under various conditions are derived. A new concept of an adaptive, intelligent, cognitive view-filter or intellectual zoom-lens is analyzed in the context of how persons with diverse cognitive styles can best use and update a general hypertext system.

1. Introduction

In the evolution of storage and access to knowledge, hypertext can be viewed a conceptual step up from database technology. While database applications address the need for fixed access to well-defined datasets, such as corporate financial or production-related data, hypertext allows for variable resolution access to more loosely structured knowledge where procedural, declarative and categorizing information are not well distinguished. It is now implemented in various ways (Conklin, 1986) with two-dimensional displays of limited resolution. More imaginative future extensions of the man-machine interface technology (Adams, 1982) have yet to be considered in this context.

Independence is a central concept in database technology. Various programs can access the same sets of data, and the logical organization of the data is not affected when new data is added or when different means of physical access are used. Basically, independence means that a user can obtain from the database useful information in response to a question expressed in the form of an algorithm (implemented as a program or command sequence) without knowledge of the way data is stored physically, how it is accessed, or even how the hardware and software are connected. Independence means accessibility to information for a wide variety of programmed and non-programmed algorithms, at the same time that a wide variety of cognitive styles in the formulation of inquiries, instructions, and learning modes by its users can be accommodated.

An early hypertext-like personal system for storing, retrieving and processing concepts, called AMNIPS, was constructed at the IBM Research Center in 1960-3. It was intended to be the first in a sequence of systems leading toward what is today called hypertext. Determining the degree of user-independence was a central issue. In this paper we explicate this concept in a formal model of contemporary hypertext technology, and enlighten our results from field experience with the second system in the sequence that started with AMNIPS.

The stimulus for the design of AMNIPS (Adaptive Man-machine Non Arithmetic Information Process System) came from a paper by Mary Stevens

(1958). Its architecture was that of a network or graph of nodes connected by links. The nodes corresponded to stored symbols or symbol strings that could refer to other symbols. Symbols served as names for objects, persons, places, times, sentences or paragraphs and, most importantly, concepts. The links connecting the nodes also corresponded to symbols. These denoted predicates in an applied predicate calculus corresponding primarily to two-place relations.

A second generation after AMNIPS was developed in 1965-70 and used for several years to guide the learning of mathematics. It is described in Section 2. AMNIPS was a personal relational "database" with inferential capability. It was intended to extend its user's private memory. As such, it did not have the defining property of a database: independence of the user.

In the second generation after AMNIPS, some of the knowledge and organization of mathematical concepts was intended to be shared by all users. It thus had a greater degree of user independence. To the extent that the personal meanings attached to certain mathematical concepts were shared by various individuals, to that extent those concepts could be used for public communication. They should also correspond to the interpretations of past authors who presented their findings in the published mathematical literature. Where possible, then, the conventionalized, consensual or canonical definitions of key concepts were made to correspond to the common core concepts. Yet, each individual still had his own internal representation of a concept, and that is reflected in the personalized part of the system. Hence, it was not completely user-independent.

The long-range vision for the later generations beyond AMNIPS was an electronic version of H. G. Wells' "World Encyclopedia" (Kochen, 1967). Such a system, called WISE, is continually updated from a variety of the most competent sources of specialized expertise. It continually screens, evaluates, integrates, synthesizes and reorganizes its knowledge base so that what is known at any time can be effectively and easily brought to bear on the day-to-day problems of ordinary people.

The second-generation version described in the latter part of Section 2 is called WISE-Mathematics. To be used <u>easily</u> by the average would-be user, the system has to be cognitively coupled with his idiosyncratic ways of thinking, remembering, expressing, interpreting and organizing. To be used <u>effectively</u> by an individual, it has to be able to change his preparedness for action, his intentions, his capabilities, his commitments in a way that is consistent with what he values or with values implicit in a value system that he consciously accepts. The system cannot both be very effective and very easily used by diverse users.

A central issue then becomes how to trade the extent to which the system can be tailored for a good fit to each individual user and the extent to which it stores and organizes user-independent knowledge that is of objective value. Results of deriving an optimal set of links and nodes are presented in Section 3.

In Section 4 we introduce and analyze the idea that a user can adjust the mask or filter that presents him with an individualized view of the hypertext. The adaptivity (the A of AMNIPS) lies not only in the plasticity and flexibility with which it can accommodate to the needs and use patterns of individuals, and to the changing corpus of knowledge, but to the self-improving features of the system itself. As a network that coordinates many knowledge processors (persons as well as their technological supports), many knowledge domains and many symbolic constructs, it has the properties of a cooperative, self-organizing system that, like the human brain and many of the systems analyzed with the help of synergetics (Haken, 1983), should exhibit interesting emergent properties and the informational equivalents of synergies.

2. Early Implementations of Hypertext-like Systems

Figure 1 illustrates a simple use of AMNIPS. It shows a tiny sample of the concepts needed to understand the Pythagorean theorem. The N's are nodes or concepts directly referred to. The L's are links—also concepts—but used to relate N's, and in this case primarily predicates denoting two-place relations. Some of these were symmetric, such as "is adjacent to," (R5) and the same as their inverse. Some were what we called transitive, now said to have hereditary properties, such as "is a special case of," (e.g., R8). (The inverses of predicates--e.g., "comprises" is the inverse of "is a part of:"—are not shown in Figure 1A.)

Generally, the triples Node-Link-Node form sentences. Some of these sentences, such as N6R3N5, "angle has a measure, in degrees, of 90," is a concept with a name of its own, N4, or "right angle," and is treated as a higher-level node. This is indicated by circling N6R3N5 in Figure 1A. Note that in Figure 1B, "right angle" is represented as a construct/figure which is a conjunction of two adjacent perpendicular lines, such as "a" in the configuration N12. One person may think of a ring angle as an angle of measure 90 degrees and another may think of it as the angle between perpendicular, adjacent lines. Both are valid, useful representations, and AMNIPS permitted each user to represent his own view.

Part of AMNIPS had been implemented on a computer by 1963 on an IBM 704. Clearly, the needed technology for properly building a fully operational system in a way that would be recognized as today's hypertext technology was not available at that time. But the concept is the same.

Nevertheless, in the 1970's, we decided to put into actual use a non-electronic version of the second generation system, WISE-Math. This was intended to be more standardized.

While AMNIPS or the first-generation system, was user-dependent, WISE-Math was more independent in two ways. First, it imposed standardized or canonical view of concepts to be shared and passively accepted by all users. Second, it allows users with different views to flexibly and actively relate and adjust their interpretation of concepts to the standardized ones.

The implementation of WISE-Math resembled a programmed text. Had videotex technology such as PRESTEL or TELIDON been available, it would have been used. It was decided to simulate the electronic system with a conventional text and human networking. It was a combined encyclopedia system and learning tool for the major concepts of mathematics. Its intended users were persons who have never learned mathematics, who had been exposed to it

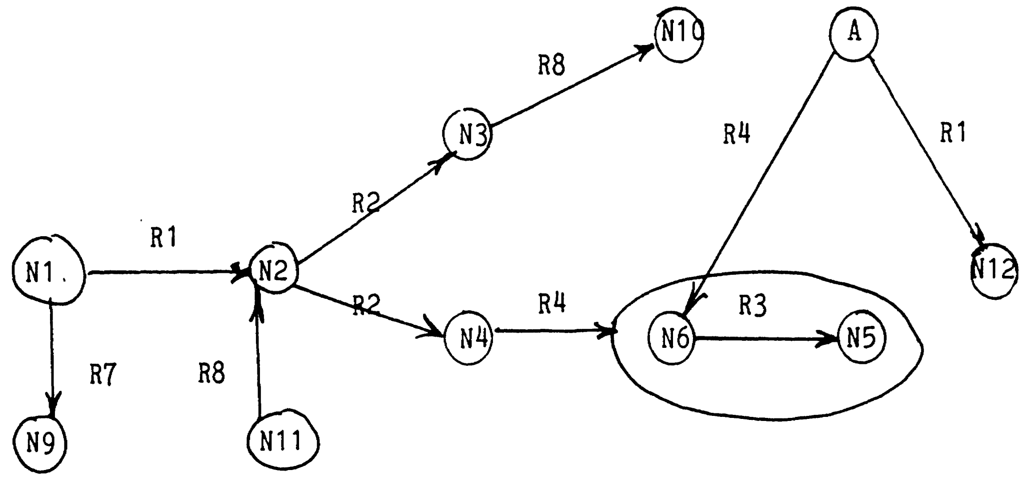


Figure 1A

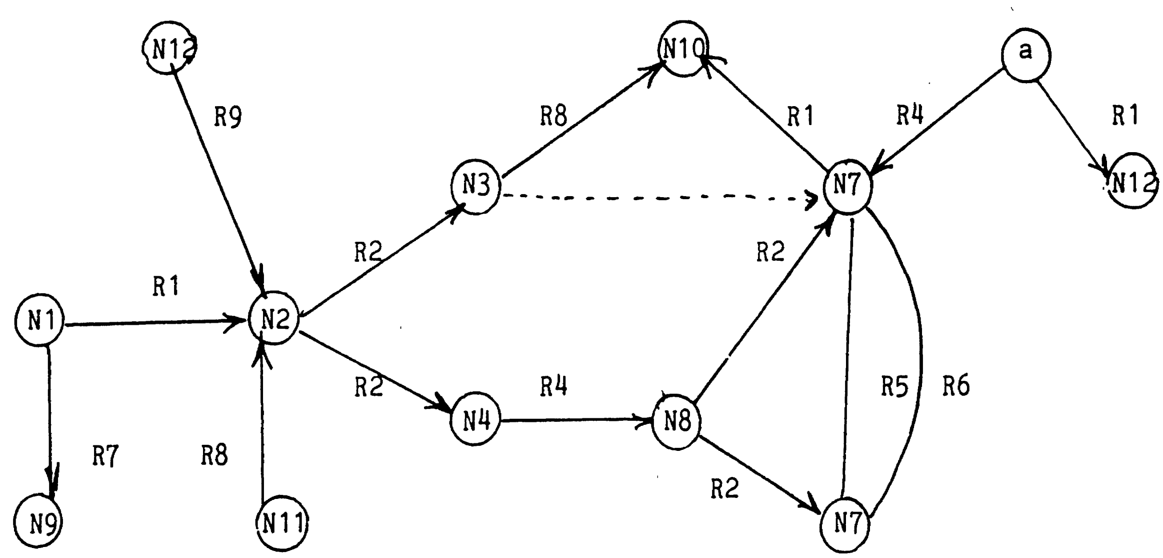
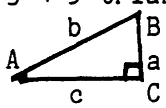


Figure 1B

N1 : hypotenuse
N2 : right triangle
N3 : triangle
N4 : right angle
N5 : 90
N6 : angle
N7 : side/leg/line
N8 : construct/figure
N9 : pythagorean theorem
N10: polygon
N11: 3-4-5 triangle
N12:



R1 is part of:
R2 is a conjunction of:
R3 has a measure, in degrees, of:
R4 is a name for:/of:
R5 is adjacent to:
R6 is perpendicular to:
R7 is used in:
R8 is a special case of:
R9 is a pictorial representation of:

but disliked or feared it, or who have learned it long ago and forgotten it. Its major functions were to: (1) offer such users an integrated perspective of the major ideas and achievements of mathematics based on careful screening, evaluation, and above all, synthesis, of mathematical concepts that evolved throughout its history up to the day of use; (2) to couple the organization and presentation of this knowledge with the experiences, interpretations, patterns of thinking of individual users in a custom-tailored manner so that they can readily internalize and comprehend it and its utility in their daily coping with matters of importance, as well as to enjoy it as a monument to the human intellect, comparable to other peaks of humanistic culture in art, music, literature; (3) to reverse their negative effect and orient them to mathematics as a constructive attitude toward confident problem-solving.

WISE-Mathematics will be described using the current language of hypertext. Like AMNIPS, it consisted of a knowledge-base of nodes and links between them. The nodes corresponded to "objects," such as pages or paragraphs in some text, human tutors, physical models such as hexaflexagons, games such as EQUATIONS, computer workstations, special graphics such as videotapes, charts, etc. Had videotex or windows been available then, the information in the nodes or some representation of the"objects" they denote would have been displayed as the content of windows on VDTs. Paper was used to simulate the use of windows, with the use of pointers on a page to simulate linking, zooming, etc. (The idea of intellectual zooming was elaborated in Kochen [1972] in this context.)

The links were directives to users for shifting their attention. They were activated by them. The nodes served as choice points as well as sources of information on which to base the choice. Learning was viewed as pursuing a path in this network. A "class" of learners was viewed to be a set of learners who pursued the same paths at about the same pace. Interaction among members of a class was facilitated by computer conferencing (CONFER), face-to-face meetings with tutors and instructors, and conventional mail or messaging on paper. A course was viewed as the process of pursuing a path. It featured individualized learning that offered group interaction where appropriate.

Fear of failure in a competitive atmosphere and in a context where there are mainly right or wrong answers is a major cause of the negative attitude to mathematics. This system avoided such conditions. Students in this "course" came from all parts and levels of the University of Michigan: freshmen, graduate students in urban planning and in library science, undergraduates majoring in premed, engineering, business, science, music, literature, etc. Each student set his or her own learning objective for their "course." If they took it for credit, the duration of the course would be a semester, with a contract for a regular schedule of interactions with the system. At the end of this term, they would present the "cognitive map" of mathematics that they created. They would exhibit evidence of changed affect and attitude. They would demonstrate their ability to bring their new knowledge to bear on solving problems that mattered to them. It was intended, however, that the path in the network they pursued would not end when the contract ended, but that it would continue for life. Indeed, follow-up with some alumni showed this to be the case for several; they were to learn as well as to continually update a limited personal version of the system by themselves for their own use, and their enjoyment in doing this was to be their incentive.

The nodes were of the following kind:

(a) Text with graphics. First encounter with a topic, whether selected by them or as part of a recommended sequence, describes at an "entry-level," a concept such as that of "triangle" or "function." "Entry-level" is user-specific; it is the point at which a user first enters the encyclopedia. Thus, someone who had exposure to calculus might encounter a definition of a function as an ordered set of pairs, in which no two pairs with the same first element can have two different second elements. Should this be too abstract or use concepts with which the learner is not completely comfortable, he can immediately and simultaneously look at an entry for "set," and for "ordered," etc.

Secondly, it offers the learner a means of testing his comprehension and affect, usually by asking him a question or giving him an instruction. The preceding item will have made it possible for him to learn as little as he needs to know to answer the question.

Thirdly, it offers the learner a choice about what to do next, generally on the basis of his response to the question or instruction. The question was intended to be answered correctly by about half the respondents, and whose who answered it correctly were directed to a node with a question at a more advanced level, following the system's answer. Those who answered it incorrectly were also given the system's answer, and directed to a question at a more elementary level.

Fourthly, it answers any questions he may have to help him make the wisest and most knowledgeable choice.

(b) Stand-by tutors. Since text, whether prerecorded or generated on demand as needed, still requires a programmer to anticipate a great deal, especially for answering learners' questions, we used human experts (seniors in mathematics honors, who took this for credit and who also learned a great deal).

(c) Multi-media, such as video and audio tapes, mixes of media and live presentations on campus (or elsewhere for learners at remote locations). The learners and tutors serve to an extent as scouts for such materials— as well as for the text/graphics in item a--and evaluated them as well. These recommendations would be used to continually replace and add items to the knowledge base, on the assumption (one of the many hypotheses tested) that students in the same class and/or of the same type would be served by an item in a manner similar to the way it served the recommender.

(d) Work stations, in the form of computer terminals or facilities for special exercises or activities with instructions for their use. This included drill for students who wished to acquire a specified degree of accuracy and speed in executing routine, formal algorithms for, say, differentiation or integration (usually to pass an advanced placement exam) in calculus.

While many of the links resembled "go to," they were really of the form "If-then-else," and "While-do," or "Next-do." Embedded in the network was a citation net, with the links pointing learners to articles or passages in articles or texts that delved into a topic in greater depth, and through which they could pursue a path right up to the leading edge, to the latest publication or person active on research in that specialty, if he wished. Moreover, he could simultaneously follow several paths if he wished to try to combine concepts.

The features of the system can be summarized as follows:

1. It is hierarchical in the sense that learners could zoom into the prerequisites of a concept or theorem or into its subsequent uses; also in the sense that it exploits a consensual hierarchy among mathematical concepts, such as "a function is a special case of a relation," "a triangle is a particular polygon."

2. It is graph-based, in that the system provided for horizontal links, such as the association between the algebraic, the geometric and the analytical views of just about every concept (e.g., derivative of x^2 as 2x, as steepness of the curve, as $\lim_{\Delta x \Rightarrow 0} ((x+\Delta x)^2 - x^2)/\Delta x)$.) The network of concepts forms clusters that are most helpful to learners in forming personal cognitive maps by iterated chunking.

3. There are many types of links. Some of the links are user-defined for constructing their personal maps. Some of these are just in the user's mind, such as the association of "perpendicular" with a particular T-square the learner has seen. Other user-defined links could be stored in a NoteCard-like personal, external memory aid. They may contribute these links to the system. To the extent that such links are not idiosyncratic but sharable with others learners of the same type, to that extent they can be added to the knowledge base.

4. Many links in a chain can be concatenated to form a new object. Even in AMNIPS (Figure 1), the object names R4 is the chain linking N6 and N5 by link R3. (Wise-Mathematics was not implemented at such a detailed level, though a subsequent AI program to simulate a learner (Kochen and Resnick, 1987) was implemented in KEE.)

5. As explained above, there are many versions of both nodes and links.

6. Nodes are identified by names, which could be meaningless symbol strings for economy of storage and retrieval and clutter control, or meaningful with numerous associational links.

7. The content of the knowledge base is "explorable." This is the most advanced of the three kinds of domains of discourse to which computer systems have been applied (Kochen, 1975): mechanizable (used for automatic question-asking); indexable (used with conventional databases and document retrieval); explorable (where comprehension and discovery is to be facilitated, as in the present case). As such, the knowledge base or encyclopedia can be searched not only for key words and text fragments but by any of the handles, clues or associations that direct learners to analogies, similarities and leads.

8. In principle, any procedure that can be implemented as a program can be attached to a node (or to a link that can also appear as a node). Such procedures can be invoked when the user follows certain instructions or initiates some actions.

9. Another idea that was thought through but not implemented was that of generating certain aspects of the nodes, such as automatic formation of questions or instructions to be issued to the learner. For example, from the generic question-type, "It takes A x hours to V, and B y hours to V; how long does it take them to V, working together," an infinitude of specific questions can be generated (e.g., by substituting personal names for A, B; numbers for x, y; "build fence F" or "dig ditch O" for V).

10. Of course, many learners can use the system simultaneously. When they enter recommendations or feedback, that can also be done simultaneously but requires communication and coordination among them. Editing requires even more coordination and concurrency control.

11. Attributes and their values can easily be associated with nodes and links. In AMNIPS, the attributes were special kinds of links and their values, special kinds of nodes.

12. Of these eleven features, 1-7 are properties of the system as implemented and field tested in 1970-78. All these features were also implemented in a modest way on computers in the 1963 AMNIPS system. Features 8, 10, 11 were implemented in more contemporary object-oriented and other systems (Kochen and Resnick, 1987, in KEE), (Kochen et al., 1986) but in different contexts than that of the integrated WISE-Mathematics. Hypertext systems have some of these features (Conklin, 1987). Because this system has most of the features, there is no doubt that it, too, qualifies to be included in the concept of hypermedia systems, even though it was not automated. Though "hypertext" is implicitly understood today to require the use of computer technology with windows, this requirement is not stated explicitly, and the idea of hypertext should not depend on its implementation.

What has been learned form the years of field experience with this early version of simulated hypermedia? How well did it meet its three stated objectives?

1. It offered users a perspective of mathematics that integrated set theory (from the axiomatics of naive set theory to fuzzy set theory), algebra, logic/automata/information theory, analysis, and geometry/topology around seminal concepts of "function" and "mapping." The core directory to these concepts comprised 600 pages of text (with graphics), collections, and references to several thousand pages of supplementary text, dozens of videotapes, dozens of games, dozens of physical models, about 30 sets of special exercises and computer packages and dozens of live tutors. The cognitive maps drawn by students at the end showed that they had internalized in an organized way a significant set of mathematical concepts and a rich set of relations among them and their uses, and at several levels of aggregation.

2. That they could use it to solve problems of importance to them was demonstrated by original papers they wrote and presented; these also showed their changed comprehension of certain pinnacles in mathematics (e.g., the prime number theorem, the Pythagorean theorem, the central limit theorem) as comparable in beauty to peaks in art or music. There are instances of pre-med or pre-business students with mediocre records whom this system stimulated to change sufficiently to be accepted in the medical or business schools of their choice. Others were helped in obtaining and performing well in jobs of their choice. A comparison between students who used this system and comparable students in the standard calculus sequence showed significantly superior performance for students in this system.

3. The change of attitude was most dramatic. Because of the nonsequential nature of this system, because of its self-paced, individualized, positively reinforcing nature that still combined the discipline and conventional edifice of accepted, accumulated mathematics, it "clicked" with self-selected students. It "spoke" directly to them, and they could communicate with it as well as with one another through it. They encountered, internalized, rediscovered, used "powerful" ideas (Papert, 1982) and programmed the system at their initiative rather than being programmed by it. They realized their amplified power in using the system, in controlling it. They realized its potential as a tool to help them in lifelong learning, as an aid to their processing of concepts, control of attention, use of analogy, ability to synthesize and, above all, to formulate and solve problems, even problems that at first seemed inappropriately vague or impossible.

   As a growing, multi-level encyclopedia of some mathematics, it was user-independent. As an individualized learning tool, it was tailored to the user. The degree to which the system fits in a personalized way could be varied, within constraints, by the user. The system as a whole was "adaptive" in that it incorporated sharable nodes and links from the users, and modified itself to become more responsive to needs and learning difficulties of its users on the one hand, and to the constantly developing body of mathematics on the other. It was a "man-machine" in that the users were closely matched to the system and part of it. It was non-arithmetic in that the processing transformed symbolic structures.

3. Degrees of User Independence

The idea of user-independence in this context is that the system can accommodate the most varied clientele with a suitable balance of effectiveness and usability. This means, for example, that the utility and beauty of the Pythagorean theorem should be comprehensible, and accessible to as many persons as possible. It means that the way the system organizes the concepts that are required to understand it does not depend on whether a user is a young novice, a person with math-phobia or an advanced learner.

Existing database systems are used mainly with a limited variety of completely programmed algorithms (Simon, 1981). This limits the flexibility of access to database information. The major contribution of independence in hypertext is the cost effective accessibility of information for a wide variety of programmed and nonprogrammed algorithms. To realize this flexibility, though, hypertext must be able to understand and respond to inquiries couched in a wide variety of cognitive styles and arising from a wide variety of educational backgrounds: it should have a degree of user-independence.

By the degrees of independence we mean the number of persons who can use it to meet their information needs at a level of effort they are willing to expend. The key hypertext design variable we are studying here is the number of link-types, with adaptivity as a design variable discussed in the next section. Letting X denote this number, we can clarify the concept of independence as follows. As X increases, uses must expend more effort to express their information need and search for a suitable response. Potential users are distributed according to how much effort of this type they are willing to expend relative to expected returns. Thus, if the effort required, for too large an X, exceeds the mean of that distribution, the number of persons using the system decreases and the degree of independence shrinks. If X is too small, users will not be able to express their information need or obtain what they need, thus incurring an information deficit. Potential users are also distributed by how much information they need. If the information supplied is less than the average need, there will also be a decrease in the number of people using the system and a decline of independence. We can expect an optimal degree of independence that corresponds to an optimal X.

To illustrate how one might obtain such an optimal X, suppose that the effort of using the system incurs a disutility that is proportional to X. and that the information deficit of not being able to obtain answers incurs a disutility that is proportional to 1/X. Let f denote the fraction of the maximum disutility of an information deficit, when X = 1 (instead of 9 as in Figure 1) that measures the disutility of adding one more term to the vocabulary (increasing X by 1). Then, the total disutility is $X^{-1} + fX$, which is minimum when $X = f^{-1/2}$; if f = .0001; for example, then X_{opt} should be 100.

A somewhat more realistic and general model is to assume that the total disutility is $X^{-a} = f \log X$, which is least when $X = (a/f)^{1/a}$; if a = 1 and f = .0001, then X_{opt} should be 10000; if a - 2, X_{opt} should be 141. A refined and more realistic, stochastic model is shown in the appendix.

How to estimate such parameters as a and f must be left for a future study. It is very situation-dependent. It depends on: (1) the domain of discourse; (2) the users' activities or tasks or needs within each domain, (3) the tailorability or extent to which the system can adapt to mesh with users' modes of thought, levels of preparedness, intelligence.

By domain of discourse we mean such areas as law, medicine, physics, history, psychoanalysis, etc. By activities we mean such tasks as are encountered in professional practice, in learning or in researching these fields. Hypertext could probably help a law student learn a general concept by simultaneously displaying to him many special cases, much as a mathematics student was helped by WISE-Math. Hypertext could help a professional lawyer by extending his personal memory of cases, as did AMNIPS for its users; he will wish to keep privately to himself the unusual precedents he found, while making use of publicly known precedents.

In the physical sciences, assistance in generating unusual combinations or patterns of findings, in the manner of exploratory research, could stimulate a researcher who mind is prepared to recognize a novel pattern. The changes of such a usual and productive combination to occur is quite low, however, and the expected costs may outweigh the expected benefits. Hypertext may be as appropriate for scientific research extending large bodies of existing knowledge, e.g., in chemistria, as well as for law because of the much larger number of nodes.

Dependence on domain of discourse can be characterized by generality. How general-purpose is the hypertext system, in the same sense that a programming language has much functionality and can be used to express knowledge and requests for knowledge in a variety of fields, such as law, business, etc. The size of an indexing vocabulary for a field such as medicine is of the order of 5000 ($10^3 - 10^4$), but it would not necessarily double if two related fields, such as business and law, for example, would be combined. Alternatively, we can ask for a given vocabulary how much of the domain can it encompass.

Dependence on the task can be characterized by activity-scope. For a given vocabulary, can only practitioners but not learners or researchers be accommodated?

Dependence on tailorability is characterized by the extent to which the system can be custom-tailored to users with different levels of preparedness, ability/intelligence, motivation and diverse ways of thinking and representing knowledge. The advantages of such a "fish-eye" view in searching a large body of knowledge have been investigated in Furnas [1985].

Let us use the term "generality" to capture both domain-independence and activity scope. The greater the number of information needs that the system can meet the greater its generality. We suspect that high generality and high tailorability do not co-exist. If that is so, the tradeoff may be captured by the family of indifference curves shown in Figure 2.

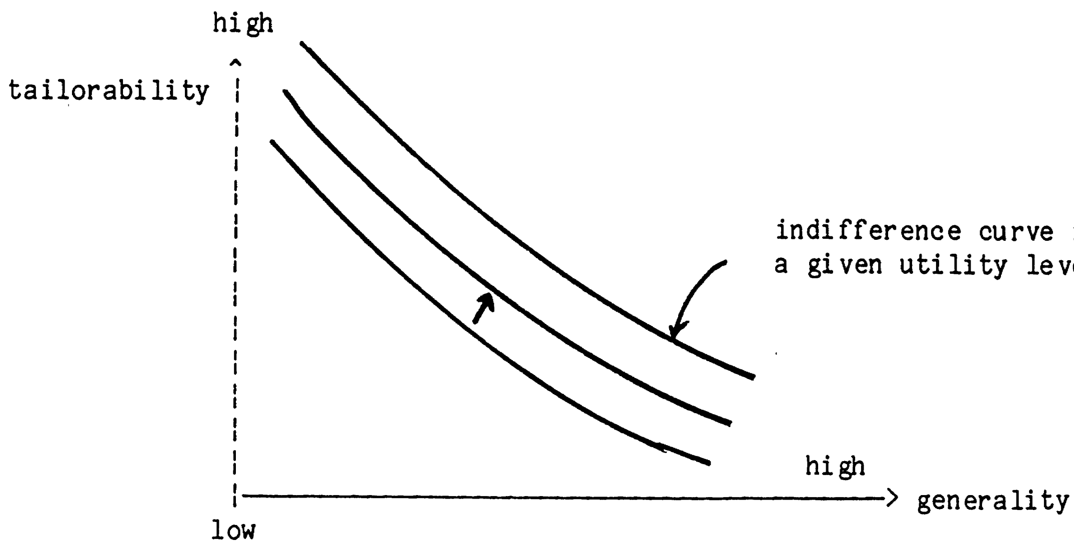


Figure 2

A shift to an indifference curve at a higher level is due by increased adaptivity, explained in Section 4. Some of the hypertext systems could be located in these two dimensions.

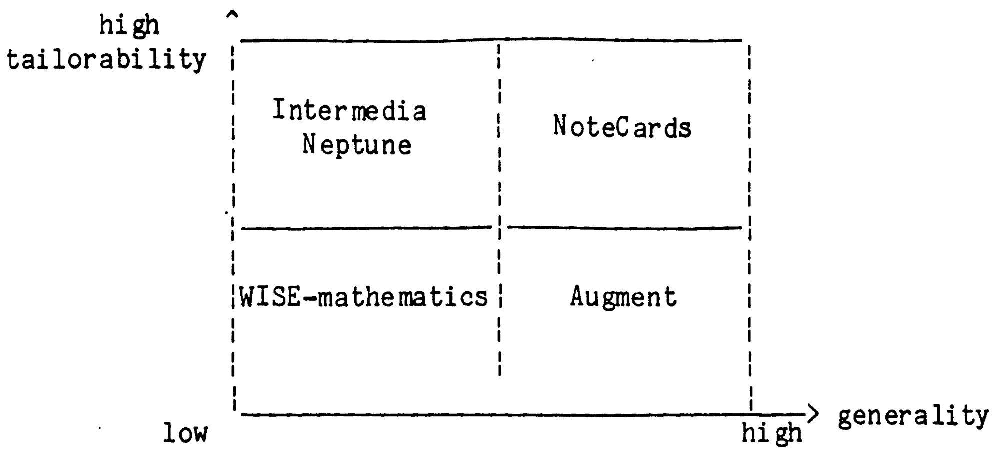


Figure 3

As we see in Figure 3, NoteCards is designed as a general purpose "idea processing" environment (Halasz, Moran and Trigg, 1987). Users can create new types of nodes and links by LISP programming in the case of NoteCards. It can cover any type of task domain. There is some degree of (non-programming) user tailorability in NoteCards as well.

On the other hand, the Intermedia (Yankelovitch, Meyrowitz and van Dam, 1985) and Neptune (Delisle and Schwartz, 1986), and Hypertext software are better in addressing education and software management applications. Augment (Engelbart, 1984) as a general system but differs significantly from NoteCards in being outline-based rather than network based.

4. Adaptivity of Hypertext System

As mentioned in Section 3, the degree of user-independence relies on two factors: (1) number of nodes and links, and (2) adaptivity of the hypertext system to adjust views of hypertext to users with different cognitive styles, tasks, and learning nodes.

The existence of optimum number of nodes and links lead us to the consideration of defining a set of users' dimensions such as task domain.

In this section, we will discuss the ways to adapt hypertext to different users given number of nodes and links. Adaptivity depends on functionality of systems such as mechanisms for zooming, indexing, and filtering of hypertext.

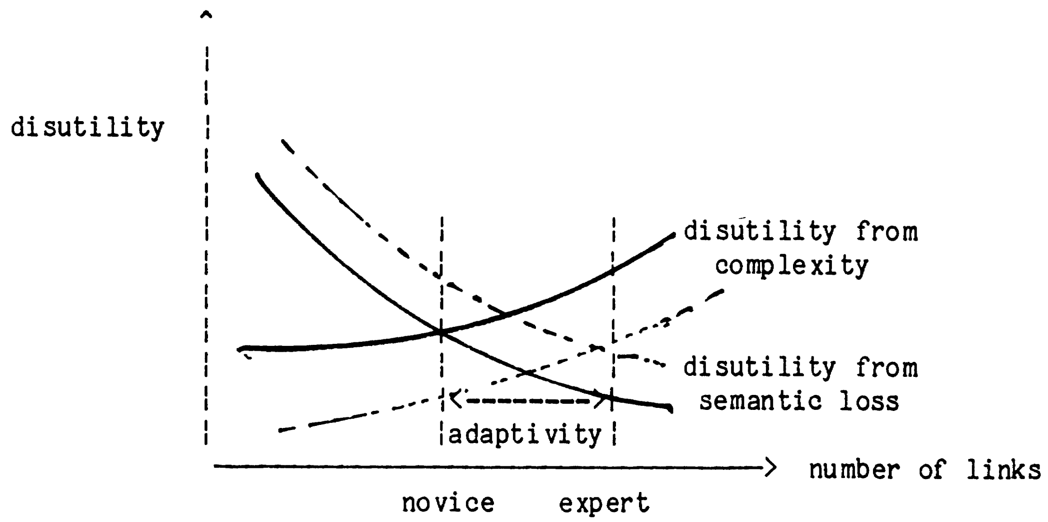
The zoomability is found in second generation of AMNIPS, WISE-Mathematics, described in Section 2. For example, a novice seeking to understand the Pythagorean theorem may need to zoom into more depth into the concept of a "right triangle" or of "perpendicularity" or into what "squaring" means. He should be able to simultaneously see definitions at several levels of, some of which are prerequisite for others.

An index mechanism is found in hypertext system such as NoteCards. The Link Index Mechanism builds a sorted list of nodes by a specified link type or set of link types. This mechanism allows users to see different views of hypertext.

Filtering means aggregation or extraction of hypertext so that the hypertext system provides different number of nodes and links to different users' situations. Figure 4 shows disutility curves of novice and expert users of hypertext systems. A novice user does not experience semantic loss as much as an expert, given the number of links and nodes. Novices may be satisfied with more aggregate types of links, whereas experts may want more links to express their ideas. For example, one link, "larger than," may be enough for novice to express the fact that one quantity is larger than any other. However, an expert may want two links, "a little larger," and "much larger." If an expert is limited to only one link, say "larger than," then his semantic will be bigger than for the notice. Thus, the semantic loss disutility curve of the expert is higher than that of the novice, as shown in Figure 4.

There is a difference between a novice and an expert in complexity disutility curve as well. The novice may not want too many links and concepts that he cannot easily understand. On the other hand, the expert will have heuristics to navigate complex hypertext systems. Thus, complexity disutility curve of expert is higher than that of the novice as shown in Figure 4.

How can we deal with the difference between experts and novices. One way is to provide novices a subset of links and concepts hiding a complex network of nodes and links. This requires another vocabulary directory which helps adjust the number of concepts and links depending on the users' experience.



Novice and Expert

Figure 4

———— for novice
- - - - - - for experts

5. Conclusion

Conditions under which hypertext helps people in generating, organizing and using concepts, ideas, and knowledge can be characterized by the size of the domain of discourse, the diversity of its concepts, the clarity and ambiguity with which they are expressed, by the needs of knowledge users in these fields. For those situations (domains of discourse and uses) for which some form of hypertext is appropriate and effective, the optimal degree of user-independence is greatest if the domain is very consensual, close to most users' experiences and internal maps, as in the case of mathematics for experienced research mathematicians. It is least if the domain is quite incoherent or most users' experiences and internal maps vary greatly, as in the case of learning psychoanalysis.

We argued here that there is a way to define and determine an optimal degree of user independence in various situations and that this is useful because it can be adjusted so as to make hypertext uses more useful and more usable.

REFERENCES

Adams, D., The Hitchhiker's Guide to the Galaxy, New York: Harmony Books, 1982.

Conklin, J., "A Survey of Hypertext," Report of the Software Technology Program, Austin, MCC, October 1986.

Delisle, N. and Schwartz, M., "Neptune: A Hypertext System for CAD Applications," Proceedings of ACM SIGMOD '86, Washington, DC, May 28-30, 132-42, 1986.

Engelbart, D. C., ""Authorship Provisions in Augment," IEEE 1984 COMPCOM Proceedings, 465-72, Spring 1984.

Furnas, G. W., "Generalized Fisheye Views," ACM-CHI86 Proc., April 1986.

Halasz, F. G., Moran, T. P., and Trigg, R. H. "NoteCards in a Nutshell," Proceedings CHI'87, ACM, 1987.

Haken, H. Advanced Synergetics, Berlin: Springer, 1983.

Kochen, M., "Adaptive Mechanisms in Digital 'Concept' Processing," Proceedings Joint Automation Control Conference. New York: IEEE, 1962, 50-59.

Kochen, M., ed., The Growth of Knowledge. New York: Wiley, 1967.

Kochen, M., "WISE: A World Information Synthesis and Encyclopedia," Journal of Documentation, 28, 322-343, 1972.

Kochen, M., and Resnick, P. "PM: A Plausible Mathemachine," Human Systems Management, 7(2), 1987.

Kochen, M., ed., Information for Action. New York: Academic Press, 1975.

Kochen, M., et al., "Computer-Mediated Collaborative Writing," Proceedings National On-line Meeting. New York: Learned, Inc., May 1986.

Kochen, M., and Dreyfuss, G., "An Experiment in Teaching College Mathematics," International Journal of Mathematics, Education in Science and Technology, 3, 1972, 315-328.

Kochen, M. "Experiments with Programmed Learning as a New Literary Form," Journal of Chem. Docum., 9(1), 1969, 10-16.

Kochen, M., "Directory Design for Networks of Information and Referral Centers," Library Quarterly, 42(1), 1972, 59-83.

Papert, S., Mindstorms: Children, Computer and Powerful Ideas. New York: Basic, 1982.

Simon, H., The Sciences of the Artificial, Cambridge, MIT Press, 1981.

Stevens, M. E., "A Machine Model for Recall," *Proceedings International Conference on Information Processing*, Paris: UNESCO, 1959, 309-15.

Yankelovic, N., Meyrowitz, N., and van Dam, A., "Reading and Writing the Electronic Book," *Computer*, 15-30, October 1985.

APPENDIX

The following assumptions are made in constructing a sample cost-benefit model of user independence.

m = the number of windows of data
n = the number of links between windows
X = the number of types of links

The more effort a user expends, the less satisfied he will be; let D represent the Disutility operator; then

$$D(\text{effort}) = k_1(m+n)$$

The more linkages that exist between windows, the more information (in the windows) the user will obtain; therefore

$$D(\text{Information Deficit}) = \frac{k_2}{m+n}$$

If each combination of window linakges is especially probable, then the total number of ways of linking a Hypertext network is

$$2^{mX}$$

and the probability of each possible linkage arrangement is

$$Pr(n|X,m) = \frac{\binom{Xm}{n}}{2^{Xm}}$$

Therefore, the expected disutility from user effort expenditure is

$$A = E(D(\text{Effort})) = k_1 \sum_{m=0}^{Xm} (m+n)Pr(n|X,m)$$

and the expected disutility of the Informatijon deficit is

$$B = (D(\text{Information Deficit})) = k_2 \sum_{m=0}^{Xm} \frac{P(n|X,m)}{(m+n)}$$

Although it is difficult to identify the number of users that will be attracted to the system by a given level of utility, we may assume that the market for services saturates relatively quickly, and perhaps this saturation may be model by a Paretian rotation to disutility, i.e.,

$$\text{Total number of users} = k_2(A+B)^{-(1+w)}$$

where $W \in (.1,\ 1.0)$ is some constant.

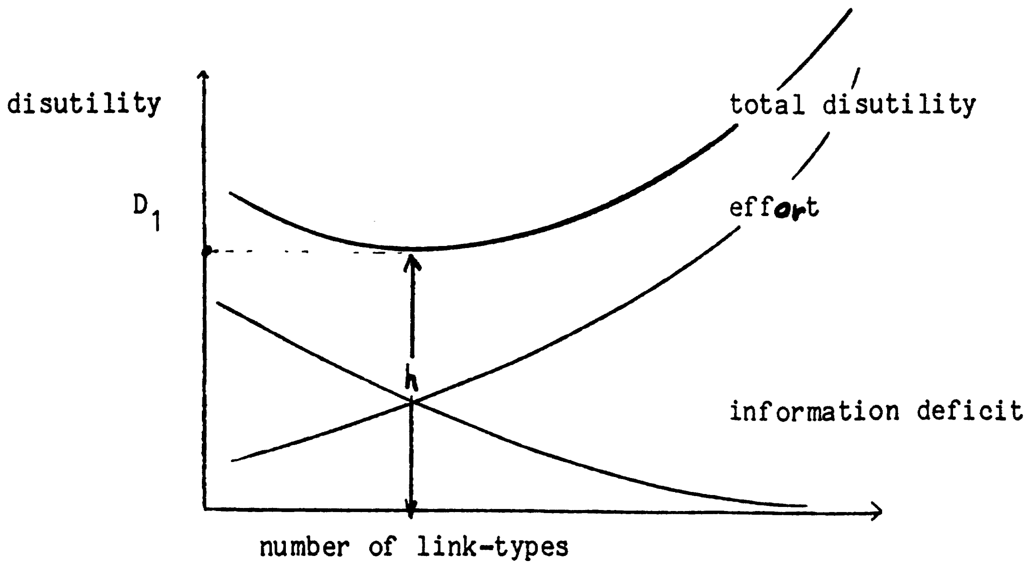


Figure 5

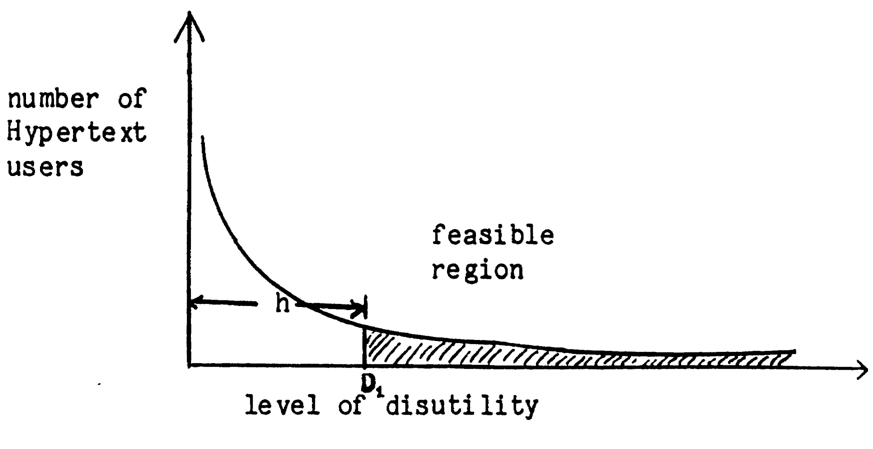


Figure 6