NAVIGATION IN HYPERTEXT: A CRITICAL REVIEW OF THE CONCEPT

Andrew DILLON, Cliff McKNIGHT and John RICHARDSON

HUSAT Research Institute, Elms Grove, Loughborough, Leics. LE11 1RG

This item is not the definitive copy. Please use the following citation when referencing this material: Dillon, A., Richardson, J. and McKnight, C. (1990) Navigation in Hypertext: a critical review of the concept. In D.Diaper, D.Gilmore, G.Cockton and B.Shackel (eds.) Human-Computer Interaction-INTERACT'90. North Holland: Amsterdam, 587-592.

With the advent of hypertext it has become widely accepted that the departure from the so-called “linear” structure of paper increases the likelihood of readers or users becoming lost. In this paper we will discuss this aspect of hypertext in terms of its validity, the lessons to be learned from the psychology of navigation and the applicability of the navigation metaphor to the hypertext domain.

1. IS NAVIGATION A PROBLEM?

There is a striking consensus among many of the “experts” in the field that navigation is the single greatest difficulty for users of hypertext. Frequent reference is made to “getting lost in hyperspace” (e.g. Conklin 1987, McAleese 1989), and Hammond and Allinson (1989) speak for many when they say:

“Experience with using hypertext systems has revealed a number of problems for users..... First, users get lost... Second, users may find it difficult to gain an overview of the material... Third, even if users know specific information is present they may have difficulty finding it” (p294).

In the following section we will discuss what is known about the psychology of navigation in physical environments and show how this might have relevance to the ‘virtual’ worlds of information space.

2. THE PSYCHOLOGY OF NAVIGATION

2.1. Schemata and Models of Generic Environments

Individuals possess schemata or models of the physical environment in which they find themselves. This is acquired from experience and affords a basic orienting frame of reference for navigatory purposes. Thus, we soon acquire schemata of towns and cities so that we know what to expect when we find ourselves in one: busy roads, numerous buildings, shopping, residential and industrial areas, many people, churches, pubs, etc. According to Downs and Stea (1977) such frames of reference exist at all levels of scale from looking at the world in terms of east and west or First and Third Worlds, to national distinctions between north and south, urban and rural and so on down to local entities like buildings and neighbourhoods.

Such frames of reference also guide our responses to the environment in terms of how we should behave. Therefore we soon realise that to interact effectively with an urban environment (e.g., to get from A to B) there are probably a variety of information sources available to us such as maps, street-signs, landmarks, tourist information facilities and so forth. In this sense the frame of reference is identical to the concept of script (Schank and Abelson 1977).

While schemata are effective orienting guides, in themselves they are limited. They do not reflect specific instances of any one environment and provide no knowledge of what exists outside of our field of vision. Yet humans have such knowledge of places with which they are familiar. So what is this detailed knowledge that we acquire of our environment and how does it emerge?

2.2. The Acquisition of Cognitive Maps

Current theories of navigation vary and it is no longer the province of psychologists alone. Geographers, anthropologists and urban planners all show an interest (see for example Downs and Stea, 1974). However, Tolman's (1948) paper on cognitive maps is frequently cited as seminal. Tolman postulated the existence of a cognitive map, internalised in the human mind which is the analog to the physical lay-out of the environment. In dismissing much of the then popular behaviouristic school of psychology, Tolman argues that information impinging on the brain is:

"worked over and elaborated....into a tentative cognitive like map of the environment indicating routes and paths and environmental relationships..."

Recent experimental work takes the notion of some form of mental representation of the environment for granted, concerning itself more with how such maps are formed and manipulated. Many theorists agree that the acquisition of navigational knowledge proceeds through several developmental stages from the initial identification of landmarks in the environment to a fully formed mental map. One such developmental model has been discussed by Anderson (1980) and Wickens (1984) and is briefly described here.

According to this model, in the first instance we represent knowledge in terms of highly salient visual *landmarks* in the environment such as buildings, statues, etc. Thus we recognise our position in terms relative to these landmarks, e.g., our destination is near building X or if we see statue Y then we must be near the railway station and so forth. This knowledge provides us with the skeletal framework on which we build our cognitive map.

The next stage of development is the acquisition of *route* knowledge which is characterised by the ability to navigate from point A to point B, using whatever landmark knowledge we have acquired to make decisions about when to turn left or right. With such knowledge we can provide others with effective route guidance, e.g., “Turn left at the traffic lights and continue on that road until you see the Bull’s Head public house on your left and take the next right there...” and so forth. Though possessing route knowledge, a person may still not really know much about his environment. A route might be non-optimal or even totally wasteful.

The third stage involves the acquisition of *survey* knowledge. This is the fully developed cognitive map that Tolman (1948) described. It allows us to give directions or plan journeys along routes we have not directly travelled as well as describe relative locations of landmarks within an environment. It allows us to know the general direction of places, e.g., “westward” or “over there” rather than “left of the main road” or “to the right of the church”. In other words it is based on a world frame of reference rather than an ego-centred one.

It is not clear if each individual develops through all stages in such a logical sequence. Obviously landmark knowledge on its own is of little use for complex navigation, and both route and survey knowledge emerge from it as a means of coping with the complexity of the environment. However, it does not necessarily follow that once enough route knowledge is acquired it is replaced by survey knowledge. Experimental investigations have demonstrated that each is optimally suited for different kinds of tasks. For example, route knowledge is better for orientation tasks than survey knowledge, the latter being better for estimating distance or object localisation on a map (Thorndyke and Hayes Roth 1982, Wetherell 1979). Route knowledge is cognitively simpler than survey knowledge but suffers the drawback of being virtually useless once a wrong step is taken (Wickens 1984). Route knowledge, because of its predominantly verbal form, might suit individuals with higher verbal than spatial abilities, while the opposite would be the case for survey knowledge.

While such theoretical work on navigation is primarily concerned with travels through physical space such as cities and buildings it does offer a perspective that might prove insightful to the design of hypertext systems, where navigation is conceptualised as occurring through an information space. Variations in navigational knowledge might account for many of the opposing views expressed on the validity of navigation problems.

3. NAVIGATION APPLIED TO ELECTRONIC DOCUMENTS

3.1. Schemata and Models

The concept of a schema for an electronic information space is less clear-cut than those for physical environments or paper documents. Electronic documents have a far shorter history than paper and the level of awareness of technology among the general public is relatively primitive compared to that of paper. Exposure to information technology will almost certainly change this state of affairs but even among the contemporary computer literate it is unlikely that the type of generic schematic structures that exist for paper documents have electronic equivalents of sufficient generality.1

Obviously computing technology's short history is one of the reasons but it is also the case that the media's underlying structures do not have equivalent transparency. With paper, once the basic *modus operandi* of reading are acquired (e.g., page turning, footnote identification, index usage and so forth) they retain utility for other texts produced by other publishers, other authors and for other domains. With computers, manipulation of information can differ from application to application within the same computer, from computer to computer and from this year to last year's model. Thus using electronic information is often likely to involve the employment of schemata for systems in general (i.e., how to operate them) in a way that is not essential for paper-based information.

The qualitative differences between the schemata for paper and electronic documents can easily be appreciated by considering what you can tell about either at first glance. A paper text is extremely informative. When we open a hypertext document however we do not have the same amount of information available to us. We are likely to be faced with a welcoming screen which might give us a rough idea of the contents (i.e., subject matter) and information about the authors/developers of the document but little else.

Performing the hypertext equivalent of opening up the text or turning the page offers no assurance that expectations will be met. Many hypertext documents offer unique structures (intentionally or otherwise) and their overall sizes are often impossible to assess in a meaningful manner (McKnight *et al*. 1989). At their current stage of development it is likely that users/readers familiar with hypertext will have a schema that includes such attributes as linked nodes of information, non serial structures, and perhaps, potential navigational difficulties! The manipulation facilities and access mechanisms available in hypertext will probably occupy a more prominent role in their schema for hypertext documents than they will for readers' schemata of paper texts. As yet, empirical evidence for such schemata is lacking.

The fact that hypertext offers authors the chance to create numerous structures out of the same information is a further source of difficulty for users or readers. Since schemata are generic abstractions representing typicality in entities or events, the increased variance of hypertext implies that any similarities that are perceived must be at a higher level or must be more numerous than the schemata that exist for paper texts.

3.2. Acquiring a Cognitive Map of the Electronic Space

The roots of this issue can be traced back to the literature on users interacting with non-hypertext databases and documents as well as with menu-driven interfaces, where it has been repeatedly shown that users can lose their way in the maze of information (Canter *et al*. 1985). Hagelbarger and Thompson (1983) claim that when users make an incorrect selection at a deep level they tend to return to the start rather than the menu at which they erred. Research by Tombaugh and McEwen (1982) and Lee *et al*. (1984) indicates that the actual to minimum ratio for screens of information accessed in a successful search is 2:1, i.e., users will often access twice as many menu pages as necessary. All of this leads such researchers to conclude that navigation through electronic (but non-hypertext) databases can pose severe navigational problems for users.

In terms of the model of navigational knowledge described above we should not be surprised by such findings. They seem to be classic manifestations of behaviour based on limited knowledge. For example, returning to the start upon making an error at a deep level in the menu suggests the absence of survey type knowledge and a strong reliance on landmarks (e.g., the start screen) to guide navigation. It also lends support to the argument about route knowledge that it becomes useless once a wrong turn is made. Making “journeys” twice as long as necessary is a further example of the type of behaviour expected from people lacking a mental map of an environment and relying on landmark and route knowledge only to find their way.

3.3. Acquiring a Cognitive Map of a Hypertext Document

McKnight *et al*. (1989) looked at navigation in terms of the amount of time spent in the contents and/or index sections of the documents employed using two hypertexts and two linear documents. They found that subjects in both hypertext conditions spent significantly greater proportions of time in the index/contents sections of the documents. They noted that this indicated a style of interaction based on jumping into parts of the text and returning to base for further guidance (a style assumed not particularly optimal for hypertext) and concluded from this that effective navigation was difficult for non-experienced users of a hypertext document.

Once more this is a classic example of using landmarks in the information space as guidance. Subjects in the linear conditions (paper and word processor versions) seemed much happier to browse through the document to find information, highlighting their confidence and familiarity with the structure presented to them. Similar support for the notion of landmarks as a first level of navigational knowledge development are provided by several of the studies which have required subjects to draw or form maps of the information space after exposure to it (e.g., Simpson and McKnight 1989). Typically, subjects can group certain sections together but often have no idea where other parts go or what they are connected to.

Unfortunately it is difficult to chart the development of navigational knowledge beyond this point. Detailed studies of users interacting with hypertext systems beyond single experimental tasks and gaining mastery over a hypertext document are thin on the ground. Edwards and Hardman (1989) claim that they found evidence for the development of survey type navigational knowledge in users exposed to a strictly hierarchical database of 50 screens for a single experimental session lasting, on average, less than 20 minutes. Unfortunately the data is not reported in sufficient detail to critically assess such a claim but it is possible that given the document’s highly organised structure, comparatively small size and the familiarity of the subject area (leisure facilities in Edinburgh) such knowledge might have been observed. Obviously this is an area that needs further empirical work.

4. PROVIDING NAVIGATIONAL INFORMATION: BROWSERS, MAPS AND STRUCTURAL CUES

4.1. Graphical Browsers

A graphical browser is a schematic representation of the structure of the database aimed at providing the user with an easy to understand map of what information is located where. According to Conklin (1987) graphical browsers are a feature of a "somewhat idealized hypertext system", recognising that not all existing systems utilise browsers but suggesting that they are desirable. The idea behind a browser is that the document can be represented graphically in terms of the nodes of information and the links between them, and in some instances, that selecting a node in the browser would cause its information to be displayed.

It is not difficult to see why this might be useful. Like a map of a physical environment it shows the user what the overall information space is like, how it is linked together and consequently offers a means of moving from one information node to another. Indeed, Monk *et al*. (1988) have shown that even a static, non-interactive graphical representation is useful. However, for richly interconnected material or documents of a reasonable size and complexity, it is not possible to include everything in a single browser without the problem of presenting 'visual spaghetti' to the user. In such cases it is necessary to represent the structure in terms of levels of browsers, and at this point there is a danger that the user gets lost in the navigational support system!

Some simple variations in the form of maps or browsers have been investigated empirically. Studies by Simpson (1989) requiring users to locate information in hypertexts have experimentally manipulated several variables relating to structural cues and position indicators. In one experiment she found that a hierarchical contents list was superior to an alphabetic index and concluded that users are able to use cues from the structural representation to form maps of the document. In a second study she reported that users provided with a graphical contents list showing the relationship between various parts of the text performed better than users who only had access to a textual list. Making the contents lists interactive (i.e., selectable by pointing) also increased navigational efficiency. In general, Simpson found that as accuracy of performance increased so did subjects' ability to construct accurate post-task maps of the information space.

4.2. The Provision of Metaphors

A second area of research in the domain of navigational support concerns that of metaphor provision. A metaphor provides a way of conceptualising an object or environment and in the information technology domain is frequently discussed as a means for aiding novices' comprehension of a system or application. The most common metaphor in use is the desk-top metaphor familiar to users of the Apple Macintosh amongst others. Prior to this metaphor, the word processor was often conceptualised by first-time users as a typewriter.2

The logic behind metaphors is that they enable users to draw on existing world knowledge to act on the electronic domain. As Carroll and Thomas (1982) point out:

"If people employ metaphors in learning about computing systems, the designers of those systems should anticipate and support likely metaphorical constructions to increase the ease of learning and using the system."

However, rather than anticipate likely metaphorical constructions, the general approach in the domain of hypertext has been to provide a metaphor and hope (or examine the extent to which) the user can employ it. As the term 'navigation' suggests, the most commonly provided metaphor is that of travel.

Hammond and Allinson (1987) report on a study in which two different forms of the travel metaphor were employed: "go-it alone" travel, and the "guided tour". These two forms were intended to represent different loci of control over movement through the document, the first being largely user-controlled and the second being largely system-controlled. Additionally a map of the local part of the information structure was available from every screen. Hammond and Allinson stress the importance of integrating the metaphor in the design of the system, which they did, and not surprisingly they found that users were able to employ it with little difficulty.

Of course, one could simply make the electronic book look as similar to the paper book as possible. This is the approach offered by people such as Benest (1989) with his book emulator and as such seems to offer a simple conceptual aid to novice users. Two pages are displayed at a time and relative position within the text can be assessed by the thickness of pages either side which are splayed out rather like a paper document would be. Page turning can be done with a single mouse press which results in two new pages appearing or by holding the mouse button down and simulating "flicking" through the text. The layout of typical books can also be supported by such a system, thereby exploiting the schematic representations possessed by experienced readers.

If that was all such a system offered it would be unlikely to succeed. It would just be a second-rate book. However, according to Benest, his book emulator provides added-value that exploits the technology underlying it. For example, although references in the text are listed fully at the back of the book, they can be individually accessed by pointing at them when they occur on screen. Page numbers in contents and index sections are also selectable, thereby offering immediate access to particular portions of the text. Such advantages are typical of most hypertext applications. In his own words:

"the book presentation, with all the engrained (*sic*) expectations that it arouses and the simplicity with which it may be navigated, is both visually appealing and less disruptive during information acquisition, than the older ‘new medium demands a new approach’ techniques that have so far been adopted.”

This may be true but at the time of writing no supporting evidence has been presented and in the absence of empirical data one should view all claims about hypertext with caution.

It is interesting for two reasons that Benest dismisses the ‘new medium demands a new approach’ philosophy of most hypertext theorists. Firstly, there is a good case to be made for book-type emulations according to the arguments put forward above about schematic representations. As outlined earlier, such representations facilitate usage by providing orientation or frames of reference for naïve users. Secondly, the new approach which rejects such emulations has largely been responsible for the adoption of the concept of navigation through hypertext.

In response to the first issue it is worth noting that Benest’s approach is, to our way of thinking, correct up to a point. We ourselves have been developing a hypertext journal database and have decided that, on the basis of some of our studies on usage styles and models of academic articles (see McKnight *et al*. 1990), emulating the structure of the journal as it exists in paper is good design. However, we are less concerned with emulation as much as retention of useful structures. This does not extend as far as mimicking page-turning or providing splayed images of the pages underlying either opened leaf. Furthermore, while we advocate the approach of identifying relevant schematic structures for texts we would not expect all types to retain such detailed aspects of their paper versions in hypertext. There seems little need, for example, to emulate the book form to this degree for a hypertext telephone directory. Benest does not seem to draw the line however between texts that might usefully exploit such emulations and those that would not, or state what he would expect unique hypertext documents to emulate.

In response to the second point, it is worth asking whether there is an alternative to navigation as a metaphor. Hammond and Allinson (1987) argue that there are two relevant dimensions for understanding the information which metaphors convey: *scope* and *level* of description. A metaphor’s scope refers to the number of concepts that the metaphor relates to. A metaphor of broad scope in the domain of HCI is the desk-top metaphor. Here, many of the concepts a user deals with when working on the system can be easily dealt with cognitively in terms of physical desk-top manipulations. The typewriter metaphor frequently invoked for explaining word processors is far more limited in scope. It offers a basic orientation to using word processors (i.e., you can use them to create print quality documents) but is severely limited beyond that since word processors do not behave like typewriters in many instances.

The metaphor’s level of description refers to the type of knowledge it is intended to convey. This may be very high level information such as how to think about the task and its completion, or very low, such as how to think about particular command syntax in order to best remember it. Few, if any, metaphors convey information at all levels but this does not prevent them being useful to users. In fact, few users ever expect metaphors to offer full scope and levels of description.

According to Hammond and Allinson, the navigation metaphor is useful in the hypertext domain and when users are offered “guided tours” through an information space they do not expect physical manifestations of the metaphor to apply literally but might rely primarily on semantic mappings between metaphor and system much more heavily. There are numerous rich mappings that can be made between the navigation metaphor and hypertext and thus it seems sensible to use it.

Benest’s book emulation is also a metaphor for using the system and in some instances would offer a broad scope and many levels of description between the paper text and the hypertext. The fact that we can talk about navigation and book metaphors in the one system shows that mixed metaphors are even possible and (though awaiting confirmatory evidence) probably workable in some instances.

It is hard to see any other metaphors being employed in this domain. Navigation is firmly entrenched as a metaphor for discussing hypertext use and book comparisons are unavoidable in a technology aimed at supporting many of the tasks performed with paper documents. Whether there are other metaphors that can be usefully employed is debatable. Limited metaphors for explaining computer use to the novice user are bound to exist and where such users find themselves working with hypertext new metaphors might find their way into the domain. But for now at least it seems that navigation and book emulation are here to stay.

5. CONCLUSION

The concept of navigation is a meaningful one in the hypertext domain in the sense that we can view user actions as movement through electronic space. Research in the psychology of navigation in physical environments has some relevance but needs further empirical investigation to identify the extent to which it may map directly onto users of electronic documents. Limitations in scope and level of application need to be made explicit. The expression of navigation difficulties is rarely supported with clear evidence, however, and the need for sound empirical work here should not be underestimated. The psychological model of navigation knowledge could prove a useful research tool in these circumstances.

FOOTNOTES

1. It is worth noting that, in part, this might be because the electronic document is usually only a stage in the production of a paper one. Few pure electronic texts exist, thus any unique forms have yet to emerge.

2. The history of technological progress is littered with such metaphors e.g., the car as the “horseless carriage”, the first typefaces were imitations of script and so forth.

REFERENCES

Anderson, J. (1980) *Cognitive Psychology and its Implications*. San Francisco: W.H. Freeman.

Benest, I. D. (1989) A hypertext system with controlled hype. Paper presented at *HYPERTEXT II*, the Second Annual UK conference on Hypertext, York, July.

Billingsley, P. (1982) Navigation through hierarchical menu structures: does it help to have a map? *Proceedings of the Human Factors Society 26th Annual Meeting*.

Canter, D., Rivers, R. and Storrs, G. (1985) Characterising user navigation through complex data structures. *Behaviour and Information Technology*, 4(2), 93-102.

Carroll, J. and Thomas, J. (1982) Metaphor and cognitive representation of computing systems *IEEE Transactions on Systems, Man and Cybernetics,* SMC-12(2), 107-116.

Conklin,J. (1987) Hypertext: an introduction and survey. *Computer,* September, 17-41.

Downs, R. and Stea, D. (1977) *Maps in Minds: Reflections on Cognitive Mapping,* New York: Harper and Row.

Downs, R. and Stea, D. (eds.) (1973) *Image and Environment: Cognitive Mapping and Spatial Behaviour.* London: Edward Arnold.

Edwards, D. and Hardman, L. (1989) “Lost in Hyperspace”: Cognitive Mapping and Navigation in a Hypertext Environment. In R. McAleese (ed.) *Hypertext: Theory into Practice,* Oxford: Intellect.

Hagelbarger, D. and Thompson, R. (1983) Experiments in teleterminal design, *IEEE Spectrum*, 20, 40-45.

Hammond, N. and Allinson L. (1989) Extending hypertext for learning: an investigation of access and guidance tools. In: A. Sutcliffe and L. Macaulay (eds.) *People and Computers V*. Cambridge: Cambridge University Press.

Hammond, N. and Allinson, L. (1987) The travel metaphor as design principle and training aid for navigating around complex systems. In D. Diaper and R. Winder (eds.) *People and Computers III,* Cambridge: Cambridge University Press.

Lee, E., Whalen, T., McEwen, S. and Latrémouille, S. (1984) Optimizing the design of menu pages for information retrieval. *Ergonomics,* 27(10), 1051-1069.

McAleese, R. (1989) Navigation and browsing in Hypertext. In R. McAleese (ed.) *Hypertext:Theory into Practice*. Oxford: Intellect.

McKnight, C., Dillon, A. and Richardson, J. (1989) A comparison of linear and hypertext formats in information retrieval. Paper presented at *HYPERTEXT II*, the Second Annual UK conference on Hypertext, York, July.

Monk, A., Walsh, P. and Dix, A. (1988) A comparison of hypertext, scrolling and folding as mechanisms for program browsing. In D. Jones and R. Winder (eds.) *People and Computers IV*. Cambridge: Cambridge University Press.

Schank, R. and Abelson, R. (1977) *Scripts, Plans, Goals, and Understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Simpson, A. (1989) Navigation in hypertext: design issues. Paper presented at International OnLine Conference '89, London, December.

Simpson, A. and McKnight, C. (1989) Navigation in hypertext: structural cues and mental maps. Paper presented at *HYPERTEXT II*, the Second Annual UK conference on Hypertext, York, July.

Thorndyke, P. and Hayes-Roth, B. (1982) Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, 14, 560-589.

Tolman, E.C. (1948) Cognitive maps in rats and men. *Psychological Review*, 55, 189-208.

Tombaugh, J. and McEwen, S. (1982) Comparison of two information retrieval methods on Videotex: tree structure versus alphabetical directory. *Proceedings of the Conference on Human Factors in Computer Systems*. Gaithersburg, MD: ACM. 106-110.

Wetherell, A. (1979) Short-term memory for verbal and graphic route information. *Proceedings of the Human Factors Society 23rd Annual Meeting*.

Wickens, C. (1984) *Engineering Psychology and Human Performance*. Columbus: Charles Merrill.