# An Environment for Aiding Information-Browsing Tasks

**Alan Wexelblat**
MIT Media Lab, E15-305
20 Ames St.
Cambridge, MA 02139 USA
Tel: 1-617-253-9833
E-mail: wex@media.mit.edu

## Introduction

Information-seeking tasks are usually characterized in terms of search and retrieval [1]. In these tasks, users are seen to be looking for some items which are more or less loosely specified. An agent or system to aid such a task works by presenting items (or descriptions of items) to the user until the desired item is found. However, this is not the only possible task for users interacting with a large quantity of information. In particular, users encountering a new information space may simply wish to *browse* that space, to get a sense of what it might offer or what kinds of items could be retrieved in the course of a search.

The abundance of new users and new sites on the World Wide Web (WWW) has lead to the popularization of a "surfing" metaphor for information-seekers on the web. The metaphor lends itself to visualizing a user moving more or less effortlessly over a wide sea of information, twisting and turning as the currents and his interests dictate.

Despite the appeal of the metaphor, there are no real environments for web surfing, or browsing. Most people experience the web as a contextless set of links, each page supplying tens or sometimes hundreds of options for further exploration. Even those sites which categorize their links, do not usually indicate how the links may relate. Users simply go from page to page, following whatever link seems best. This depth-first wandering provides no organization and no means of relating one discovery to another.

The web also supports a number of information-finding engines. However, most of these, such as Yahoo [2] or Lycos [3], are oriented toward directed search tasks, rather than browsing. In this directed search, the user supplies a series of key terms or other information and the system responds with a list of pointers to pages which contain the given terms. While directed-search tools are useful in their own right, they do not really help with browsing.

Even in a situation where a user knows a topic which he wishes to explore, the sheer volume of results returned by the various search engines can be overwhelming. For example, a query to Digital's Alta Vista search engine [4] for the phrase "Babylon 5" returns ten thousand 'hits.'[1] *Babylon 5* is a relatively unpopular television series; the much-more popular show *Star Trek* returns an order of magnitude more hits.

Even if the search engines were able to eliminate duplicate links, the results would not necessarily be of interest to someone who wanted to know what the Web had to offer on the topic of *Babylon 5*. Many of the pages simply contain the phrase because the author of the page watches it, or knows someone who does. The presence of the search term is no guarantee that the page will be about the desired topic.

We need an intermediate solution, one which permits users to get a sense of what an information space has to offer. This intermediate task I refer to as *browsing* — a casual survey of a space done to familiarize oneself with what it might contain.

In this paper I describe what I see as the requirements of a system to support browsing tasks on a large, dynamic, heterogeneous information collection such as some part of the WWW. I then describe the kind of system I am building in an attempt to satisfy those requirements. The system will draw on and extend work from learning interfaces and autonomous agents.

## Characteristics of the Task

The goals of users involved in information-browsing tasks (for simplicity we call them 'browsers' hereafter) are usually hard to specify. Rather than seeking a particular set of things, they are more interested in gaining some sense of familiarity with the possibilities offered by what they are browsing. My model of this activity is much like the browsing we engage in while wandering the aisles of an unfamiliar bookstore or music store. We also may browse the contents of a particular book — looking at the table of contents, at some pictures, perhaps reading a paragraph here or there.

Browsing is characterized by frequent changes of topic, of interest, and by changes in level of detail: when we come across something that interests us, we may explore it in more detail. However, this is an incomplete description of the

1. A 'hit' is a link to a page; multiple hits may link to the same page, so the count is approximate.

task. Browsing is often most easily characterized by what it is not.

### It is not a search for a specific target.

Some information searching tasks take place because the user cannot recall the target by name, but feels he will be able to recognize it ("I'll know it when I see it."). In these tasks a standard approach is to present the user with a series of options and take some measure of how similar the presented items are to what is desired [5]. It is assumed that the presented items will have some features in common with the desired object and that by extracting these features and presenting possible targets which have -- or have more of — these features an eventual match can be found. This is, analogously, often the process one goes through in shopping for "just the right" article of clothing or jewelry.

Browsing does not have this targeted behavior. Although a browser may stop and examine information items in detail, there is no sense in browsing of narrowing down toward a specific target. Items which are examined sequentially in a browse may have few features in common.

### It is not necessarily a terminable task.

Most information search tasks have a specifiable end condition, even if it is only the user identifying a desired target after a long search. System efficiency may be measured by the rapidity with which the search is terminated. Browsing, by contrast, is a procedure more governed by the level of the user's interest. As long as interest is maintained, the browsing may continue indefinitely. At the very least, it is not possible to specify in advance what termination conditions might hold. Browsing systems may be said to be performing better the longer users spend in the system. Efficiency may be measured by the user's level of comprehension after using the system to browse the information space. In addition, we are working to develop some in-system measures of efficiency, as described below.

### It is not necessarily linear.

Most information search tasks involve a process of generating candidates, testing them against some criteria, and then accepting or rejecting each candidate in turn. Browsing, by contrast, is highly non-linear and often characterized by serendipitous discovery. These discoveries are made by a process of associations within the browser's brain which is still not well understood in cognitive science. Something catches the browser's eye or ear and is followed up as potentially interesting; it may remind us of something, or it may simply catch our fancy for a while. Often the thing which sends us off on a tangent is not per se interesting in and of itself, but it suggests other things that may be of more direct interest. As part of this serendipitous discovery process, browsers often need to be aware of possibilities beyond the one directly in front of them and of the relationship of the other items to the one being considered. That is, if I see something which makes me think of something else, I need a relatively easy way to access that other thing or find out if there's something like what I have thought of.

## Characteristics of the Solution

Given the problem characteristics described above, certain aspects of the solution seem obvious. In particular, we want to provide the user with the opportunity to browse through parts of the data space without being overwhelmed by the possibilities. Spatial organizations have been used for many information retrievals tasks, both at the interface level (as in the various tools that make up the Xerox PARC Information Visualizer[6]) and at the environment level (as in a number of virtual reality systems, for example [7]). The goal of spatial layouts of information is usually to present the user with as much information as possible without overloading her or creating a "lost in hyperspace" feeling.

In our domain, spatial layouts are interesting for another reason: narrative — specifically story-telling — can provide a rational, understandable basis for organizing information-finding tasks. In particular, what I call the "travel story" [8] gives us a way to help users comprehend where they are and what they might do next. The travel story is a simplified version of what Joseph Campbell called the "hero's journey." [9] This archetypal tale is told in many cultures in many forms and has many specific cues; the travel story is simplified in that it lacks the mythic cues of the hero's journey but retains the simple structure of leaving home, exploration and encounter, and return.

The key feature of a travel narrative is that it provides a metaphor that contextualizes the amorphous activity of browsing. Travel stories are simple narratives of the form "first we did this, then we went there, then we saw that" and so on. This ordering is the temporal form of a simple linear spatial ordering (such as "this thing was on the left, this other thing was next to it, and that thing was on the right"). In some sense the story structure simply exchanges space for time in organizing information. However, it seems likely that story structures are somehow innate to the way we organize information in our brains [13]. The simple form of a travel story, with its organizing principles of travel, explore, and return, helps people remember the stories for retelling and will help people remember and navigate an information space, if a story of travel through and exploration of the space can be created.

This reasoning suggests that we might organize information for browsing in the form of a spatial structure which users can browse by moving through. In effect, a "scientific visualization" such as is sometimes seen in virtual reality systems [10]. However, unlike a typical scientific visualization which tries to present as much information on-screen at one time as possible in order to facilitate getting an overall or holistic picture, our structured space should support limited serendipitous discovery and avoid information overload. This suggests using a space which is subdivided into rooms or areas. Each area will be used to visualize a small amount of related information; links between related information groups can be easily represented as doors between areas.

In contrast to previous efforts in this area (such as Xerox PARC's Rooms system [11] and follow-ons) we do not intend to construct a static set of precomputed and pre-arranged rooms and connecting doors. Instead, we envision our space as a dynamic environment, supporting the browsing task. In particular, the system keeps track of the browser's interests and reactions as she moves through the space and constructs the visualization ahead of the browsing process in response to what the browser has shown interest in earlier on the journey. Specifically, we can create the rooms in a personalized and contextualized way, giving us a different set of rooms for each user.

The basic idea is to have two different information spaces to meet the needs of two very different tasks. In one space — the original link-based space — information is created and linked together as the information authors think appropriate for publication and archival or scholarly purposes; in the case of the WWW these are the web-page authors, who place content and links on their pages as they see fit. In the other space — the personalized set of rooms — information is organized to facilitate browsing, and constructed dynamically to account for evolving user preferences and interests.

In essence, this second space represents an externalized form of the cognitive model a user builds of a particular information set. As the browser repeatedly encounters the information, her model grows and the space grows with it. Rooms and connections which were made in past travels through the space stay in place to facilitate learning landmarks and remembering past encounters. However, any time the browser steps outside the bounds of what has been encountered before (even if it is only to delve into more detail on a previously-seen topic) the space must grow to accommodate the new information being encountered.

## Assisting rather than Automating

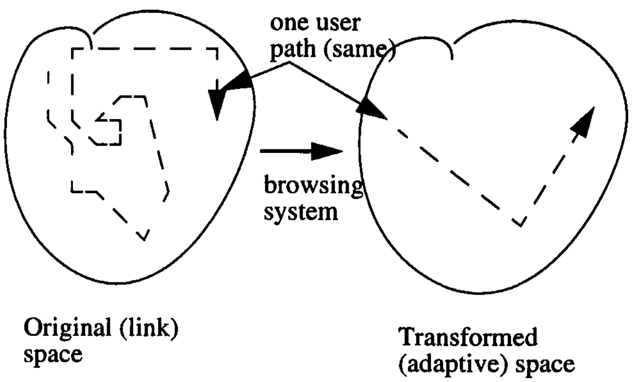
Another difference from conventional approaches which this task seems to require is that we try not to automate the basic browsing task. Rather, we envision systems that assist users. While browsing under one's own control is usually pleasurable activity, it is hard to see how browsing could be automated — except perhaps as a computational version of the neighbor's vacation slide show.

The first sort of assistance given by the system is the dynamism of the space, constructed for the browser through passive observation. The space assists browsing by automatically filtering and grouping information into related and easily-grasped chunks and providing gentle transitions between the chunks.

This allows us to think about new kinds of measurements of efficiency, based on what the user does in the system, rather than testing knowledge outside the system. In particular, we are investigating using measures of backtracking as a way to gauge how well the system is performing.

The idea, in its simplest form, is that if the system is correctly understanding the user's evolving preferences, then it should put the item most desired up for next viewing by the browser. If the desired item is not presented, the browser will need to backtrack and take a different path. Of course, this is a theoretical ideal, and would probably never be approached in reality.

Mathematically speaking, we can look at this as a function which has the effect of translating a random walk into an approximation of a straight line:




That is, the browsing system computes a function which transforms the space on the left into the space on the right, while satisfying the property that the user's path (which appears to be a random walk in the original space) as closely as possible approximates a straight line in the transformed space.

Of course, there is no general solution for this problem (which is roughly NP hard). However, we are not proposing a *prescriptive* solution; rather, this is a *descriptive* attribute of the solution, if it is implemented properly.

The second form of assisting we are implementing into the system is that of passive learning. Rather than requiring browsers to attempt to formalize or regularize their interests and desires, the underlying algorithms of the space operate "passively" in the sense that they work only with information "given off" by the browsers as they interact with the space during their task. In specific, inputs to the learning algorithm consist of external observables such as the amount of time spent viewing each piece of data, the sequence of items viewed, and the trail of links between clusters (the story being told). These inputs are related by the system to the underlying data, represented in a domain-specific knowledge base. The learning algorithm manipulates the underlying knowledge to extract the next appropriate clusters for the browser to see.

The importance of a passive learning algorithm is magnified by the type of task we are considering. Browsing is an inherently contemplative task, often involving lengthy associational chains of interests and serendipitous discoveries. Requiring the browser to instruct the system actively in order to proceed would interrupt these chains of thought and throw the person out of the task. This could only be counterproductive.[2]

The final facet of our solution strategy involves making available to the browser an agent guide. This agent serves much the same role as a museum tour guide or a knowledgeable sales clerk in a bookstore. The guide agent will be equipped with a domain-specific knowledge base about the subject matter being browsed. Thus the agent will be able to "understand" what the browser is looking at and will be able to provide more in-depth information if desired. It is still an open research question what information should be directly visualized and what should be accessible through the agent guide.

In general, all visualizations have this problem: they hide or deemphasize certain things in order to reveal or highlight certain others [1]. Generally speaking, most visualization techniques give the user ways to find out additional information about something which is visualized — revealing detail, as it were.

The guide agent is important in this process for two reasons. First, it gives the person a natural way to find out about things that are not directly visible, by asking questions such as "What's behind that door?" or "Can I find out about *Babylon 5*?" If the user's query matches terms in the agent's knowledge base, the agent can formulate a conversational answer. If the query is about location, the agent can show maps, or add informative labels to the thing being queried. Being able to ask about and interact with areas of the information space not directly visible will aid users in making leaps of intuition or changes in interest which the incremental learning algorithms cannot predict.

2. Unfortunately, there seems to be no substantive work in cognitive science on precisely how the mental processes of browsing work in the human brain.

The alternative would be to have a secondary, contextualizing view (such as a map) to which the user could refer. However, switching views (particularly between first- and third-person views) often presents jarring discontinuities in the interface. In addition, the use of a map would require that the space be mappable in advance, which is contrary to our intention of having the space grow as the user's cognitive model of the information grows. Finally, a map implies a fixed set of connections between any two locations, and we would like our browsing space to present many different paths between two locations, depending on how the browser thinks of the two items being connected. This multiplicity of paths is unlikely to lend itself to a simple map-like (two-dimensional) representation.

Second, the presence of a conversational agent allows us to avoid overloading the visual display with detailed information. For example, if the user is looking at a picture of a character in the *Babylon 5* television show, she might ask "Tell me more." The agent could then respond with a prerecorded dialog about the picture being viewed. While the agent is speaking the text, the user is free to continue browsing. Again, we maintain the natural flow of the browsing task.

The alternative, displaying the information as text to be read on the screen, would require the person to cease browsing in order to read the block of text. Additionally, the displayed text is likely to obscure objects of interest, since both the text to be read and objects of maximum interest must be near the foveal/central area of the display.

In general, autonomous agents are an assistive technology by virtue of being programmed to take over the "dirty work" — the boring, repetitive or uninteresting parts of a task [12]. In our domain, this means accessing what may be lengthy expository blocks of text information. We extend beyond this simple use of agents, though. The guide will also assist the browsing task by engaging other human senses in the system; although agents are not necessary for this kind of multimodal interaction, they provide a natural focus for things such as speech, which otherwise must issue from thin air.

Specifically, the guide will communicate verbally with the browser -- using text-to-speech and speech recognition. This allows the browser to remain mobile — she can move around freely while listening to information, rather than being stuck in place reading blocks of text. Voice commands will allow the user to skip through, terminate, or ask for more information (similar to what was done by Arons with his SpeechSkimmer hardware [14]).

## Conclusion

In this paper I have outlined an approach to solving the problems posed by information-browsing tasks. In general, the solution advocates employing assistive learning and adaptive, personalized agent technologies to create a second information space customized to the task, rather than requiring the user to operate in the native information space, which is not structured for browsing. The environment for browsing must give the user a helpful set of possibilities without either overwhelming her or distracting her from the task.

### References


[1] Furnas, George. "Generalized Fisheye Views," *Proceedings of CHI'86*, ACM Press, 1986.

[2] URL: http://www.yahoo.com/

[3] URL: http://www.lycos.com/

[4] URL: http://www.altavista.digital.com/

[5] Cypher, Allen (ed.) *Watch What I Do — Programming by Demonstration*, MIT Press, Cambridge, MA 1993.

[6] Robertson, George, Stuart Card & Jock Mackinlay. "Information Visualization using 3D Interactive Animation," *Communications of the ACM*, 36(4), 1993.

[7] Stoakley, Richard, Matthew Conway & Randy Pausch. "Virtual Reality on a WIM: Interactive Worlds in Miniature," *Proceedings of CHI'95*, ACM Press, 1995.

[8] Wexelblat, Alan. "Telling the Travel Story," *AAAI Symposium on Interactive Story Systems*, Stanford University, AAAI Press, 1995.

[9] Campbell, Joseph. *The Hero With A Thousand Faces*, Princeton University Press, 1949.

[10] Erickson, Thomas. "Artificial Realities as Data Visualization Environments," in A. Wexelblat (ed.) *Virtual Reality: Applications and Explorations*, Academic Press Professional, 1993.

[11] Henderson, D. Austin and Stuart Card. "Rooms: The Use of Multiple Virtual Workspaces to Reduce Space Contention in a Window-Based Graphical User Interface," *ACM Transactions on Graphics*, 5(3), ACM Press, 1986.

[12] Maes, Pattie. "Intelligent Software," *Scientific American*, September, 1995.

[13] Schank, Roger. *Tell Me A Story*, Charles Scribner's Sons, 1990.

[14] Arons, Barry. "SpeechSkimmer: Interactively Skimming Recorded Speech," *Proceedings of UIST'93*, ACM Press, 1993.