The Mind's Eye

An Approach to Understanding Large Complex Information-Bases Through Visual Discourse

Earl Rennison

Bachelor of Science, University of California,
Davis, California 1988

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning
in Partial Fulfillment of the Requirements for the Degree
Master of Science in **Media Arts and Sciences** at the
Massachusetts Institute of Technology

September 1995

Author Earl Rennison
Program in Media Arts and Sciences
August 11, 1995

Certified by William J. Mitchell
Dean, School of Architecture and Planning
Thesis Advisor

Accepted by Stephen A. Benton
Chairperson
Departmental Committee on Graduate Students
Program in Media Arts and Sciences

The Mind's Eye

An Approach to Understanding Large Complex Information-Bases Through Visual Discourse

Earl Rennison

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, on August 11, 1995
in Partial Fulfillment for the Requirements of the
Degree of Master of Science in Media Arts.

Abstract

To develop an effective medium that adapts organically to changing information and responds dynamically to users' changing interests, we must rethink the process of designing and presenting information. Most information systems focus on reducing the information load by filtering information. This thesis presents an alternative approach, called the *Mind's Eye*, that allows people to see all information available in a corpus of information from a high level, while allowing them to quickly dive in for details. The Mind's Eye demonstrates a novel interface approach for visualizing, navigating and accessing information objects in a large body of unstructured information such as on-line news stories and photographs available via Clarinews; electronic mail; articles in a historical information-base; and World Wide Web documents.

To explore issues of design in these dynamic information environments, a prototype system was developed. This system provides mechanisms to analyze the relationships between information objects and build a representation of the underlying structure of the entire corpus of information. This relational structure is used to construct a visual information space with which the user interacts to explore the contents of the information-base. As a user moves through the information space, the system responds by dynamically changing visual attributes or restructuring the space to illustrate the underlying organization and structure of the information. This process creates a *visual discourse* that is much like a conversation that uses natural language as the medium to communicate, except natural language is replaced with a virtual space. The process of visual discourse enables understanding of a large complex information-base by allowing people to see and explore complex relationships between information objects.

Thesis Supervisor: William J. Mitchell
Title: Dean, School of Architecture and Planning

This work was sponsored in part by ARPA, JNIDS, and NYNEX.

The Mind's Eye

An Approach to Understanding Large Complex Information-Bases Through Visual Discourse

Earl Rennison

The following people served as readers for this thesis:

Thesis Reader

Glorianna Davenport
Associate Professor of Media Technology
Program in Media Arts and Sciences

Thesis Reader

Michael Hawley
Assistant Professor of Media Arts and Sciences
Program in Media Arts and Sciences

Thesis Reader

Pattie Maes
Associate Professor of Media Technology
Program in Media Arts and Sciences

Acknowledgments

First, and certainly most importantly, I thank my parents, brothers and sisters for their unending support.

Lisa, Nicolas and Sara, thank you for your undying friendship through the dark days, as well as the days full of sunshine. Each of you (especially Lisa) hold a very special place in my heart. You have made this a very special time in my life.

To the AARDVARKS (Deb, Jeff, Josh, Nicolas, and Sara), thank you for your weekly dose of fun, excitement and encouragement. Without your friendship this would have been an entirely different experience.

I would like to thank my initial advisor, the late Muriel Cooper, for her vision, inspiration and support without which I would not have embarked on this journey. She was my compass that pointed my way to new discoveries. I would also like to thank my thesis advisor, William Mitchell, for his many insights and guidance. Thanks also to my readers, Glorianna Davenport, Michael Hawley, and Pattie Maes, who helped shape both the research and this thesis document.

I would like to thank my colleagues in the Visible Language Workshop– Suguru, Robin, Ishantha, Ron, Dave, Lisa, Jeff, Yin Yin, Louie, and Xiaoyang– for providing a rich and diverse environment. I have learned much from each of you.

I would like to give a special thanks to two members of the VLW: Lisa Strausfeld and Suguru Ishizaki. Lisa, thank you for your inspiration, collaboration, commitment and compassion. I attribute you to the origin of this research, and continued insights along the way. Suguru, I thank you for your countless hours of invaluable discussions, advise, design input, and everything else a student could possibly hope for.

Thanks also to my UROPs–Loren, Amy and Vivek– for your many hours of committed work in constructing tools for annotating objects and building the Millennium database.

Linda, Santina, and Nancy, I thank each of you for bringing order to my chaotic ways.

Table of Contents

Chapter I

Introduction

As traditional forms of media transition to a non-linear digital environment, we need to rethink the nature of the computer as medium and the process of designing and presenting information in this new environment. Is it possible to construct a new medium that adapts to a world of rapidly changing information and users whose interests change moment by moment as they explore these vast information landscapes? If we are to create such a medium, how should the design process change to reflect this new medium? This thesis begins to address these questions.

In this thesis, I have explored a new interactive medium that *responds* dynamically to a user's changes in direction of gaze and movement through a virtual information space that *expresses* structural relationships between information objects. These virtual information spaces contain visual elements whose presentation *represents* the underlying structure of a large complex information-base. While the visual elements form output, they also provide the basis of user input. As a user moves relative to these visual elements, the virtual space interprets this movement by changing the visual properties of the elements or restructuring the space itself to show different relationships. An organized progression through these virtual spaces constitutes a *visual discourse*. This *movement-based interaction* with information breaks the stultifying boundaries of the WIMP (window, icon, menu, pointer) interface and the desktop metaphor to provide a new *experience* with information.

These ideas were explored in four different projects. The first project, Galaxy of News, formed the foundation for this work and focused on categorical organizations of information (see Figure 1.1). The second project, the Millennium Project, explored geographical, temporal, as well as categorical, organizations of historical information covering art, science, music, philosophy, and political events from 1906 to 1918 (see Figure 1.2). The Millennium project formed the basis of most of the work discussed in this thesis. The remaining two projects, Interactive Boston (see Figure 1.3) and InteractiveMTV (see Figure 1.4), explored the use of narrative techniques to dynamically express stories using this new medium.

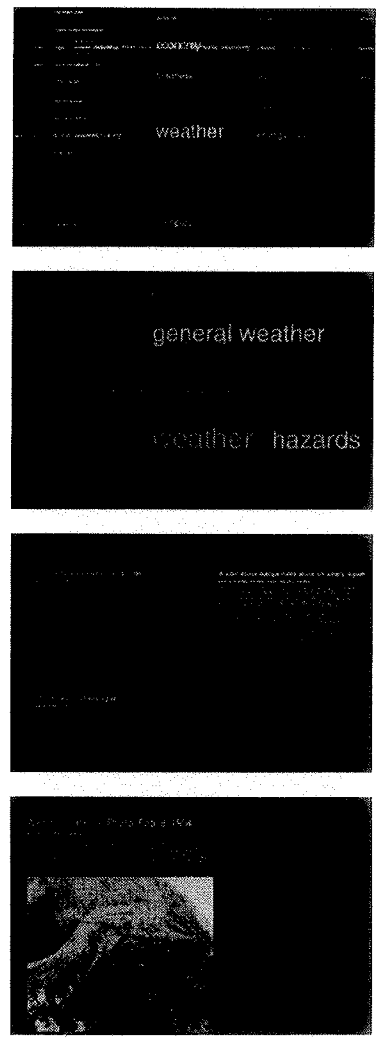


Figure 1.1 *This sequence of images (from top to bottom) shows a progressive zoom into a Galaxy of News. As the user zooms into the space, the system responds to the users changing focus of attention by restructuring the space to show increased level of detail.*

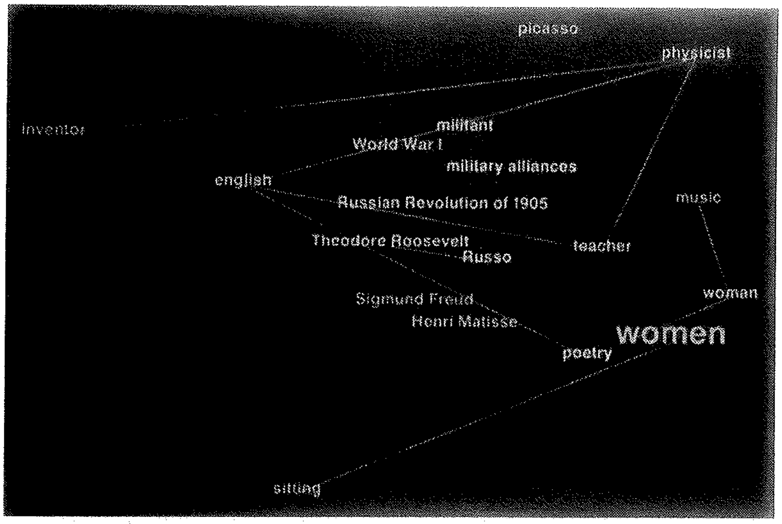


Figure 1.2 Taken from the Millennium Project, this sequence of screen captures shows a user moving through categorical, geographical and geographical-temporal virtual information spaces.

Figure 1.2.a A top-level categorical space that shows the main categories of information in a historical information-base covering the period of 1906 to 1918. This space provides a map to a symbolic landscape that represents the information-base. The interconnecting lines indicate direct relationships between the categories the lines connect. Similar categories are placed in relative proximity.

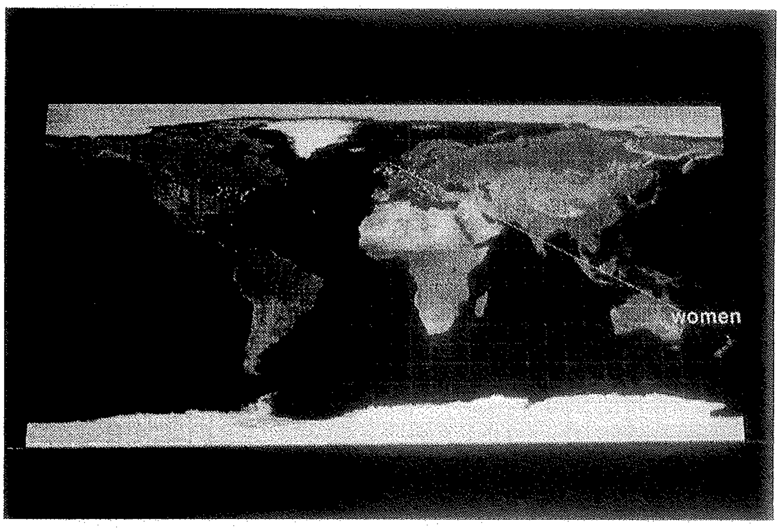


Figure 1.2.b A geographical space showing the layout of information objects placed according to their corresponding geographical position(s). The symbol women *represents a thread that connects information objects that related to the category of* women *(in this case most of the articles discuss the women suffrage movement). We reached this space by moving up to the category* women *shown in Figure 1.2.a above and rotating our position about the symbol. As a result, this space was automatically constructed with the* women *thread connecting related articles.*

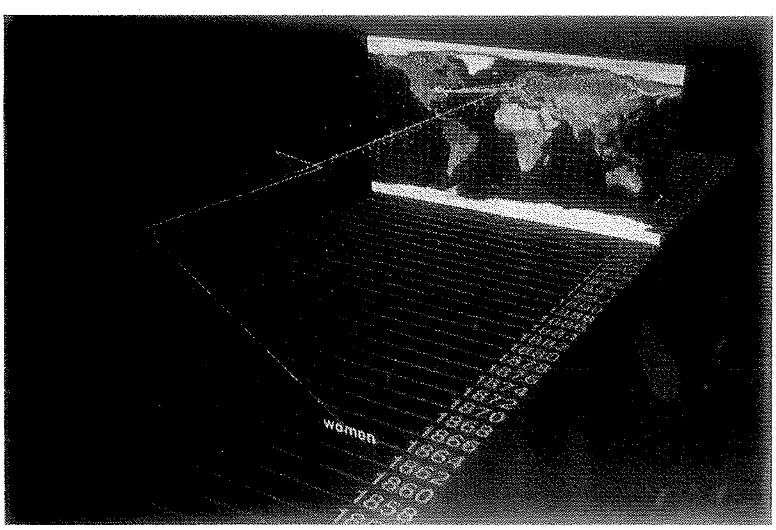


Figure 1.2.c After moving away from and rotating the geographical space shown in Figure 1.2.b, the space transforms itself into a geographical-temporal space showing the time and place relationships between the information objects. In this space, we can also see the temporal relationships between the articles connected by the women *thread.*

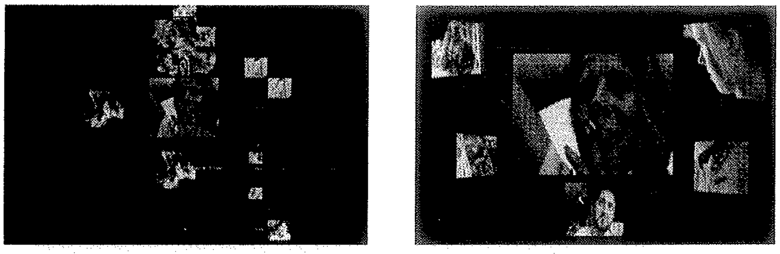


Figure 1.3 Screen captures taken from Interactive Boston showing a story space

Figure 1.4 A sequence of screen captures as a user performs music that has associations attached to the musical components. These associations form queries into an image database. The images is computed using multidimensional scaling techniques. These images were taken from the InteractiveMTV project. (note the similarity of these images to the image shown in Figure 1.2.a.)

The examples shown in Figure 1.1 to Figure 1.4 show selected snapshots of dynamically constructed visual spaces, and give a general feel for the visual imagery experienced when moving through these spaces. Examples of the process of moving between these visual spaces are given in subsequent chapters of this thesis.

In the next section, I describe the general motivations and problems that this thesis addresses.

1.1 Motivation and Problem Statement

While global interconnectivity is solving the problems of information access, it is creating a new problem of finding and assimilating relevant information. With the advent of the internet and technologies like the World Wide Web [Berners-Lee, 92], we will soon have rapid access to all the information produced and published. This technology will afford the opportunity for people to publish as well as consume information. If you or I want to be informed about a topic and information documents on that topic exists, we will have access to it. However, when we go to find information documents of interest, we face the daunting task of finding answers in a sea of information.

The traditional approach to dealing with this problem of information overload is to reduce the information load to a list of only the most relevant information documents. This is the approach taken with information retrieval and filtering systems. Both of these systems deliver a reasonably manageable set of documents to the user.

Information retrieval techniques are effective at reducing the document load if the search parameters are easily quantifiable and where the result of the initial query answers a specific question (see Figure 1.7). In the World Wide Web (WWW) today, information retrieval systems such as InfoSeek and Lycos effectively find specific web pages. For example, a query with the keywords "Media Lab" returns the home page of the Media Lab.

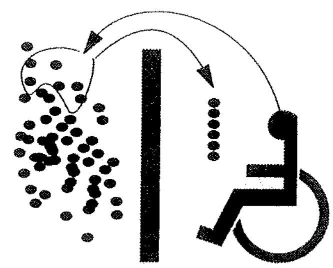


Figure 1.5 *The process of information retrieval where a person specifies a query and the system returns a list of documents.*

There are, however, several significant problems with information retrieval systems. First, they require a user to bring knowledge of what they are looking for (in the form of keywords) in order to access the information. In many cases, it is not easy to clearly state what they are looking for in the form of keywords. And, if their initial query does not return something close to what they are looking for, it is difficult to reformulate the query. Second, and perhaps more significant, the result of an information query is simply a list of relevant information documents. This list says nothing about how the documents relate to one another, nor how they relate to other documents in the information base as a whole. In general, these systems provide no clues to the user about what goes on behind the scene.

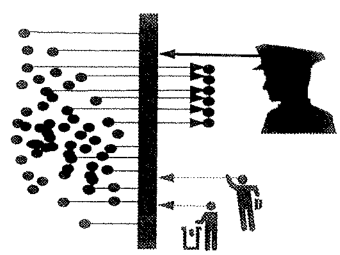


Figure 1.6 *The process of information filtering where "collaborators" specify preferences and the system "recommends" a list of documents.*

Another approach to reducing the information load is to use information filtering techniques such as collaborative filtering techniques like Tapestry [Goldberg, 92], RINGO [Shardanand, 94], Webhound [Lashkari, 95], and MotorMouth [Mentral, 95] (see Figure 1.7). These techniques rely on human evaluation of documents (or music in the case of RINGO) to filter documents based on recommendations of other people. This technique is valuable from the perspective that it facilitates qualitative analysis of information to retrieve relevant and valuable information. These systems do not require that an individual formulate a query using keywords, yet they effectively find documents related to ones previously specified. While these information filtering techniques may successfully reduced the document overload problem, producing a reduced list of documents, the burden still rests upon the reader to make sense of the reduced set of documents. These tools do not reveal to the user how the documents relate to one another, nor about the knowledge contained within them.

Information retrieval and filtering systems usually return results as lists of documents and do not show the relationships between the returned documents, nor to other related documents. In general, users do not have access to the *computational space* used to make decisions on selecting documents, and hence it is difficult for people to understand the process. Ideally, a system should address this issue.

1.2 Approach

My approach does not focus on reducing the load by eliminating information, rather it addresses information overload by visually organizing information to give people a sense of the whole body of information while facilitating rapid visual filtering of the information. What if, rather than using information filtering techniques in isolation, we apply an "information lens" to the information (see Figure 1.7)? In this way an information seeker will understand the process of arriving at a given set of documents or information objects.

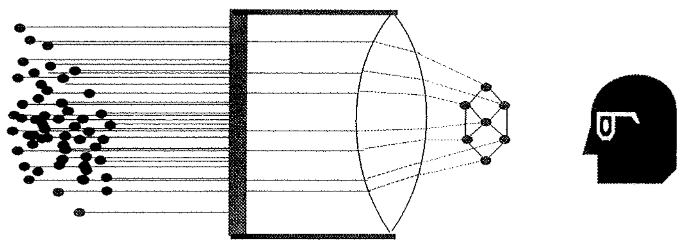


Figure 1.7 The combination of an information filter and an information lens results in a visualization of the structural relationships between retrieved information documents

With an information lens approach, the information documents or objects, like light rays, are passed through an information filter (though the filtering process is optional) followed by a lens that *focuses* the information objects into an image that shows the underlying relationships between the information objects. With this approach, the information seeker can *see* the underlying structures used to select information objects and how this structure relates the objects to one another. From one visual image, we can move to another visual image by *shifting perspective* or *refocusing* on a different set of information. When we move from one information object to another, we see the image in transition, much like we see intermediate images when we refocus the lens of a camera from one object to another. In this way, we engage in a continuous dialog, moving smoothly between images in a symbolic information landscape. This thesis explores this approach.

This approach simulates the notion of the *Mind's Eye* (hence the title of this thesis)–the action of focusing our mental attention on some topic in terms of the direction of our visual gaze at some locus [Lakoff, 89]. Lakoff and Johnson summarize the MINDS-EYE metaphor as follows:

> Via this metaphor, redirecting the mind from one topic to a more important one is understood in terms of pointing from the current locus of visual attention to another, more important visual locus [Lakoff, 89].

This metaphor implies a 3D spatial relationship between entities. For example, our eyes are round objects that rotate in sockets. If we point our eye in another direction, we must rotate our eye around an *x*- or *y*-*axis* or both. The fact that we have two degrees of rotational freedom requires a three dimensional space. Even if we hold rotation constant, changing the in focal length implies the space has depth.

Furthermore, we model the dialog, the process of moving from one image to another, as a *visual discourse*. A visual discourse is much like conversations we have every day that use natural language as the medium to communicate, except natural language is replaced with a virtual space that is populated with conversation elements. This model employs the metaphor DISCOURSE SPACE IS PHYSICAL SPACE; DISCOURSE ELEMENTS ARE ENTITIES [Lakoff, 87]. This approach makes physical those processes that are employed by the mind. Expressions like "Can we go back to our last point?", "I'm lost. Where are we in this conversation?" provide evidence of the spatial nature of discourse. The process of visual discourse is a central theme in this thesis.[1]

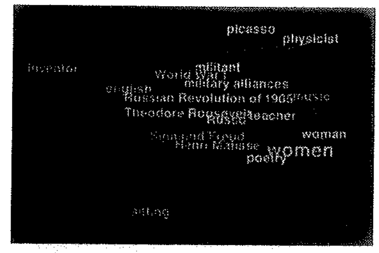


***Figure 1.8** This image shows the relationships between a set of symbols that represent categories of an information-base.*

The process of visual discourse is modeled after concepts of *mental spaces* developed by Fauconnier [Fauconnier, 94]. Mental spaces consist of sets of elements and relationships between those elements. Fauconnier maintains that we communicate meaning through a process of accessing mental spaces through conceptual connections [Fauconnier, 94]. Discourse is represented as a process where 1) categories are used to set up mental spaces, 2) temporary connections to other spaces are established, and 3) new frames are created dynamically as the discourse unfolds. In this process, participants must keep track of the maze of spaces and connections that are built, and this happens through the use of point-of-view and point-of-view-shifts.

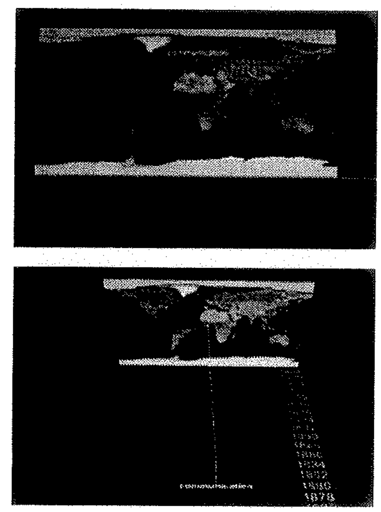


***Figure 1.9** These images show a point-of-view shift from a geographical organization to a temporal organization.*

The visual discourse is the visual analog to mental spaces and discourse. Visual discourse addresses the communication of meaning through a process that contains the following:

- Visualizing relationships between information objects, where the relationships are illustrated by their relative position in virtual space that is phrased by contextual information [Fauconnier, 94; Rennison, 94; Strausfeld, 95][2]. For example, we can create a virtual space with information objects placed in the space so that the distance between information objects maps to the conceptual distance between the objects (see Figure 1.8).

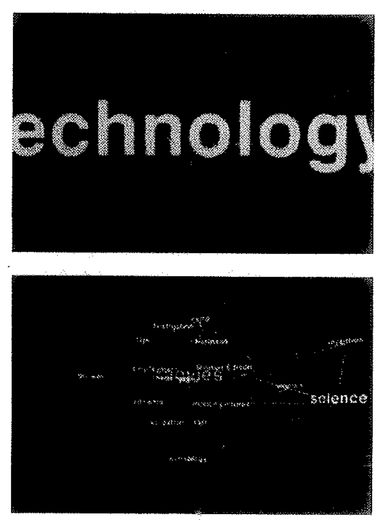


***Figure 1.10** These images show a context -shift as a user shifts from one categorical space to another.*

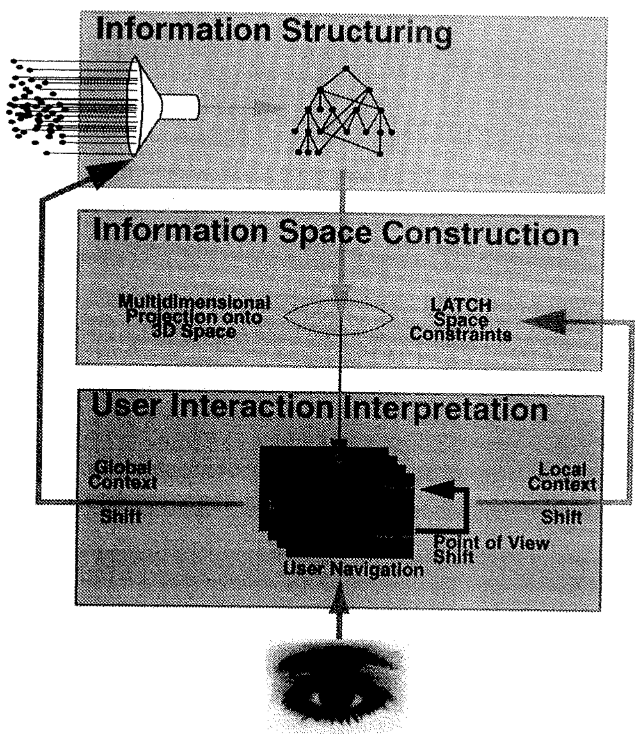
1. I should point out that in the current implementation of visual discourse as discussed in this thesis document, only the computer has the ability to construct spaces. While the user can use the space constructed by the system to formulate input to the computer, the current system does not support the user constructing a space as input. However, I do envision in the future that systems will be capable of supporting a bidirectional space construction and interpretation process.
2. An analogical example of this is a 2D graph where the data points, or in our case information elements, are positioned in a space that is phrased by the labels on the x- and y-axis.

- Dynamic point-of-view shifts, where the information objects are held constant and the types of relationships are changed. For example, we may shift from a space showing information objects positioned on a geographical map, to a space showing the same information objects aligned along a time-line (see Figure 1.8).
- Dynamic context shifts, where a new set of information elements are dynamically established based on the constraints of the new context. For example, we may want to narrow a search from one topic to another more specific topic, as shown in Figure 1.10.

In the next subsection, I describe my computational approach to enabling visual discourse.

Computational Approach

The computational approach to visual discourse, illustrated in Figure 1.11, consists of three main components that work in unison. These components consist of 1) information structure analysis, 2)



Chapter 4

Chapter 5

Chapter 6

Figure 1.11 The Computational Process of Visual Discourse.

information space construction, and 3) user interaction interpretation. The information structure analysis component automatically analyzes a set of information objects to derive conceptual structures that aid us in understanding the relationships between those information objects. These conceptual structures include categorical structures, hierarchical structures, relational structures, radial structures, linear quantity scales, and foreground-background structures. Each of these structures help us understand the relationships between information elements. This component sets the stage for visualizing the information relationships. The information structure analysis process is described in detail in Chapter 4.

The information space construction component builds virtual spaces that show the underlying relationships between information objects. This component is like the information lens. It takes the results from the information structure analysis component and projects them into a visual space.

This projection process leverages on our inherent metaphorical understanding of conceptual structures to project the conceptual structures into virtual space (see Table 3.1 on page 44). Lakoff and Johnson have demonstrated that our conceptual systems stem from our bodily experience in the physical world; we make sense of the conceptual in terms of the physical [Lakoff, 87]. Our conceptual system is grounded in perception, body movement, and experience in physical and social character [Johnson, 92]. Metaphor serves as a mapping from a source domain, our physical experiences, to a target domain, our conceptual understanding. In the process of building virtual spaces, we apply the metaphorical mapping in *reverse*, and map conceptual structures to physical experiences. In this case, the physical experiences are derived from visualizations of and movements through virtual space. The process of constructing virtual spaces is described in Chapter 5.

The interpretation of user interactions in a virtual space, the third component, realizes a visual discourse. The role of the user interaction interpretation component is to interpret a user's movement in a virtual space and respond by moving the point-of-view, shifting the perspective and/or changing the focus of the information lens process described above. As users move within a virtual space, the system responds by reconstructing new spaces to show additional relationships between information. For example, by moving up to a keyword displayed in space, we can automatically form a query for information objects that relate to that keyword (see Table 3.2 on page 46).

As with the construction of the virtual spaces, we utilize metaphorical mappings to interpret user's movements relative to information objects. For example, if the user moves TOWARD an object, the system interprets this action as a desire to see more detail relating to that object. Or, if the user TURNs an object, the system interprets this action as wanting to see information from a different conceptual point-of-view, such as turning from a geographical view of information to a temporal view. The mechanisms for interpreting user interactions and a formal grammar for how they are interpreted is described in detail in Chapter 6.

1.3 Related Work

The traditional approach to information processing and visualization is to have two separate components: one engine to do the analysis and information retrieval or filtering [Deerwester, 90; Goldberg, 92; Lashkari, 95; Mentral, 95; Sulton, 83; Tversky, 77], and a second engine to do the visualization of the analysis results [Feiner, 90; Furnas, 94; Furnas, 89; Perlin, 93; Rao, 94; Robertson, 91; Stone, 94]. These two components are typically decoupled. As a result the visualizations are static in nature. They do not allow people to see the underlying structure from a different point of view quickly and easily. And, when a new point of view or query is generated, there is often a disconnect between the original visualization and the subsequent visualization. Some systems, like Pad [Perlin, 93], allow the user to move through continuous zooms in and out of a hierarchy; however, the underlying hierarchical structure they are visualizing is static and does not allow for the multiplicity of relationships between information.

A fundamental problem faced by these information visualization systems is that, for all intents and purposes, it is impossible to construct a single image that communicates the entire knowledge of the information-base. Even if a very specific question is asked, it is still difficult to generate a single image that communicates the answer because each image represents a single point of view, and that point of view more often than not raises more questions. This, therefore, necessitates the need to generate multiple images to fully communicate the knowledge necessary to answer questions. To address these problems, a system must responds to users changing interests and questions second by second, allowing the user to specify the course of image generation, while the system responds with images.

To respond to user's dynamic queries, we must move beyond the traditional two component approach that decouples analysis from visualization. These two components must be integrated to more effectively communicate knowledge that is multidimensional in nature.

Some systems have attempted to integrate analysis with visualization, such as Starfield Displays [Ahlberg, 94][1]. This system couples a two dimensional visualization of data fields with a data query engine. Users are given sliders to control values that form the basis of queries. This approach provides an effective way to understand a data set by adjusting parameters and seeing the effects it has on the resulting visualization. There are several limiting factors to this approach, however. The first is the fact that only two dimensions can be shown at any given time, whereas the underlying nature of knowledge is that it is multidimensional. Which leads us to the second point, it does not handle unstructured, text-based, information. The underlying database is a relational database that requires all values to fit into a specific field in the database schema. Nonetheless, the system does point out the utility

1. This system is actually a *data* visualizer (as opposed to an *information* visualizer), but we discuss it here because of its merits.

of a dynamic, responsive system in answering questions and generally communicating an understanding of the database.

1.4 Summary of Research Projects

The design and presentation approach taken in this thesis was explored in four primary projects: 1) Galaxy of News, 2) the Millennium Project, 3) Interactive Boston, and 4) InteractiveMTV. Each of these projects and their relation to the thesis work are outlined below.

Galaxy of News

The approach used in this thesis arose out of the Galaxy of News project conducted early in the master's program. The Galaxy of News project addressed many issues relating to visualization of news information. The project resulted in a novel process for constructing dynamic visual environments for exploring information (see Figure 1.12). It suggested new directions and new possibilities that formed the foundation of this thesis. However, it also raised a number of issues that needed to be addressed in subsequent systems and visual designs. These issues, discussed in Chapters 2 and 3, were addressed in this thesis.

Figure 1.12 *Screen shot from Galaxy of News showing a space containing articles on* Weather.

The Millennium Project

The Millennium Project was a follow-on project to Galaxy of News and was conducted jointly with Lisa Strausfeld, another member of the Visual Language Workshop at the Media Lab. This project carried the ideas of visual discourse forward, as well as reinforced the visualization and interaction techniques discovered in Galaxy of News (see Figure 1.13).

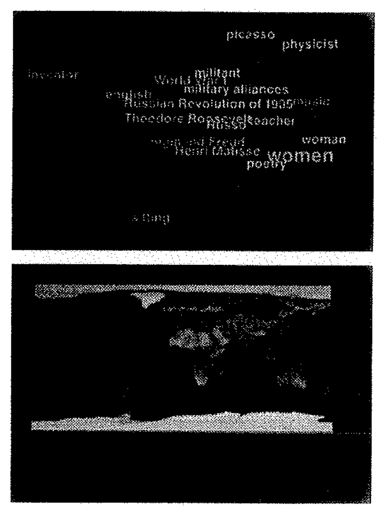


Figure 1.13 *Screen shots from the Millennium Project showing information objects presented in 1) in a categorical space, and 2) a geographical space*

The Millennium Project explored construction of virtual spaces to enable understanding of historical information[1]. We constructed a database of artifacts of philosophy, painting, music, literature, science, and political events of a pivotal time in world history: the years from 1906 to 1918. This database consisted of a set of files that contain information objects that describe events, artifacts, people and ideas pertaining to the years 1906-1918. They are displayed as 3D text objects that sometimes include images, video clips, or sounds. Each of these information objects contain annotations that describe the properties of the information objects. These basic properties include:

1. date
2. location
3. associations (term, object)
4. cause-effect relationships
5. size measurements

1. This project was conducted along with Lisa Strausfeld, another graduate student in the Visible Language Workshop at the Media Lab.

Our objective of this project was to enable understanding of historical information by allowing knowledge seekers to explore and examine information objects in dynamic 3D spaces. These dynamic 3D spaces adapt to the dynamic interests of the knowledge seekers, following the typically nonlinear path of mental understanding.

We chose this project for two reasons: 1) by creating the database, we could control the type of content, how the content was annotated, and how the annotations effected the visualization process, and 2) the database was intended to be a litmus test for the underlying computational approach.

Interactive Boston

Interactive Boston was sub-project to Elastic Boston led by Glorianna Davenport. My portion of the larger Elastic Boston project addressed some of the visual issues of associated with telling stories in a non-linear fashion, and the process in which non-linear stories are authored. The Elastic Boston team created an annotated database of video clips, images, sound bites, animations, and text articles. Given these annotations, I used the techniques developed in this thesis to present the composite story using an abstract virtual 3D space as the stage (see Figure 1.14).

Figure 1.14 *A screen shot from the Interactive Boston project showing a collage of characters.*

InteractiveMTV

InteractiveMTV is a visual environment that responds dynamically to a user's musical performance. This project explored an approach to constructing visual imagery based on annotated musical characteristics. In this environment, the characteristics of the performed music form a query into an image database. The underlying structure of the database in relation to the musical query form the basis for visual layout of retrieved images. The result is a process where as the performer performs music, imagery is generated that may alter the musical choices of the performer, creating a interplay between music and imagery.

Figure 1.15 *A screen shot from the InteractiveMTV showing an image composed while performing music.*

1.5 Summary of Research Results

This thesis developed two main concepts:

- *Movement-based Interface*–Breaking the stultifying boundaries of the desktop metaphor and point and click interfaces, this thesis defined and explored a movement-based interface that is based on movement through a virtual space. A movement-based interface facilitates an intuitive interaction with abstract information because it maps our understanding of interaction in the real world to interaction with abstract concepts.
- *Visual Discourse*–Just as we move through mental spaces and imagery in our mind as we think and converse with other humans, this thesis defined a process for 1) constructing visual imagery, as

expressed through elements placed in a virtual space, that expresses abstract concepts through visual elements, and 2) an intuitive means for moving between images.

More specifically, the results of this thesis are summarized as follows:

- Conceptual Structure to Virtual Space Mapping–Lakoff and Johnson identified seven conceptual structures: categorical structure, hierarchical structure, relational structure, radial structure, linear quantity scale, foreground-background structure [Lakoff, 87]. We defined a mapping from these conceptual structures to virtual space representations based on principles of metaphor. The mapping utilizes our understanding of information in terms of our experiences in the physical world. This mapping included definition of the relationship between conceptual structures, information organizational structures (location, alphabet, time, category, and hierarchy)[Wurman, 89], computational structures, image schemas, and metaphorical mappings. These results are presented in Table 3.1 on page 44.
- *Conceptual Structure Derivation Algorithms*–Three algorithms were developed for deriving structure automatically from an information base, including 1) multiple inheritance categorical classification (combines categorical, hierarchical and relational structures), 2) radial structure, and 3) temporal-relational structure.
- *Visual Design Techniques*–A computational approach to designing virtual information spaces that represent automatically derived conceptual structures was developed and a rationale for applying mathematical processes established. The mathematical processes map structural relationships between information objects to properties of visual elements.
- *Mapping from User Movement to Visual Responses*–A set of user movements and actions in a virtual space were identified and defined, and a mapping between actions and computational operations that effect visual imagery was established. These mappings are based on metaphorical principles of movement in the physical world. These results are presented in Table 3.2 on page 46.
- *Computational Environment for Process Exploration*–A computational environment was developed to explore the process of authoring non-linear dynamic information spaces and illuminate issues involved in this process.

1.6 Organization of Thesis

The remainder of this thesis is organized as follows. Next, I digress slightly in Chapter 2 and discuss the Galaxy of News project. I present Galaxy of News as background both because it was conducted first, and because it provides an example on which we can base discussion. Galaxy of News suggested a number of opportunities, but also raised a number of questions and left several issues unresolved. Chapter 3 lays the foundation for how I went about addressing these questions. In this chapter, I provide a demonstration of the approach that was a result of the Millennium Project. The Millennium Project is then used as an example in subsequent chapters. Chapter 3 also lays down the conceptual framework and the computational approach for the thesis. In the subsequent three chapters (Chapter 4, 5 and 6) I present in detail the results of my investigation into, one, deriving structure (Chapter 4), two, visualizing structural relationships (Chapter 5), and three, a process for interpreting a users movement through a virtual space to automatically formulate queries to the system (Chapter 6). I then briefly present two additional projects, Interactive Boston and InteractiveMTV, that apply and extend the concepts described in the core of this thesis (Chapter 7). And, finally, I conclude with a summary of the results and future directions (Chapter 8).

Chapter 2

Background: Galaxy of News

This chapter describes the Galaxy of News system, a novel approach to visualizing and interacting with news information [Rennison, 94; Rennison, 95a]. Rather than approaching information visualization as a process of creating a static object that we look at and examine, the Galaxy of News visualization approach produces an *experience* with information. In this approach, the user is placed *in* an *immersive* space that is built from the underlying structure of the news information. As the user moves through the space, different elements of the structure are revealed.

Because the approach focuses on creating an immersive experience, the interaction technique is equally important as the visualization approach. The Galaxy of News interface creates a *dialog* between the user and the information space. As the user moves within the space, the system interprets the *movement* and presents more details or more abstraction, depending upon how the user moves. A key aspect of the interaction is the *fluid movement* through the space as the system continuously responds to users movements.

The Galaxy of News work forms the foundation of this thesis. This prototype was developed early in my master's thesis program. The results of this project were so influential on the approach taken in this thesis, I present it here as a foundation for subsequent work.

The remainder of this chapter presents

- the motivation for the project
- an example experience of moving through a news information space
- the visualization and interaction approach
- the design of the interface, and
- related work

This chapter concludes with an evaluation of the results. This evaluation sets the stage for the remainder of this thesis.

2.1 Problem and Motivation

In the Fall of 1994 during a Computers and Graphics Workshop class discussion, the class was discussing news filtering systems and autonomous agents when Muriel Cooper raised the provocative question–"what if instead of filtering information, we could see all the news simultaneously on a single screen?" In many ways, this question posed the primary problem that Galaxy of News sought to address. In pursuing this problem, I also addressed many other issues, and along the way, many additional questions were raised.

The questions and issues addressed by Galaxy of News include the following:

- *Effectively browsing through massively large news spaces*–What support do we need to give people to help them browse through massive information spaces in an intelligible fashion? Can we create a coherent space that naturally facilitates access to related pieces of information?
- *Combine the effective aspects of both browsing and searching/ filtering*–Is there a way that we can combine these two diverse modes of thinking and operation in a seamless and transparent fashion with a single interface?
- *Dynamic level-of-detail selection*–What interaction techniques can we provide to allow people to quickly see information at a level of detail that matches their level of interest? How can we control this process without requiring the user to read a lot of text?
- *Direct interaction with the content*–Can we create an environment that allows people to control the presentation of information they see without having to rely on menus, sliders and buttons that are physically separated from the content? Can we understand where we are in the process of browsing and searching without relying on external maps to show people where they are in an information space?
- *Continuous interaction*–Can we move away from the point and click metaphor and allow more natural types of interactions that are continuous and have intelligible transitions?
- *Visualizing relationships*–Considering the fact that news articles are authored independent of other articles, can we visualize how these articles relate to each other, and find relationships between articles that where not previously known or obvious?
- *Quick relational access*–How do we give the user ability to quickly access related information and understand the relationships between old information and new information?

In addition, I also wanted to see if we could apply concepts of pyramidal encoding of data to the domain of information. For example, a pyramidally encoded JPEG image is represented by layers of data, where the lower layers contain low frequency information and the higher layers contain increasingly higher layers of information. Each layer of information adds increasing detail. The first layer produces a blocky image and subsequent layers increase the resolution until a high quality image is produced. Figure 2.1 illustrates a pyramidally encoded JPEG image, showing the resolution of an image with different layers of detail. I wanted to find a similar structure (and method of presenting that structure) for information. The result would be a way of allowing people to quickly select portions of news information and dive in for more detailed information.

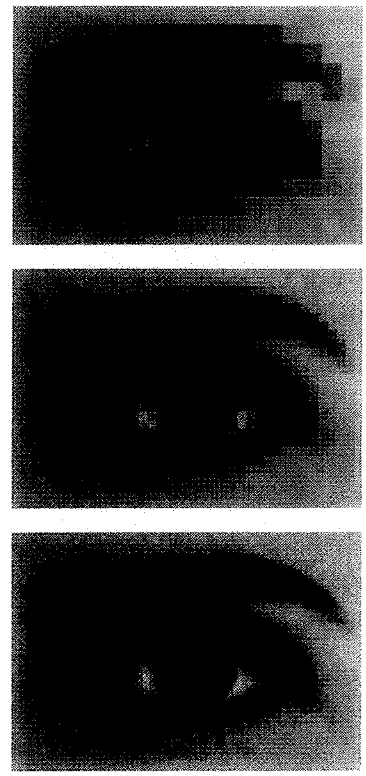


Figure 2.1 *A pyramidally encoded JPEG image. The top image shows the first layer of encoding. The middle image consists of two layers of information. And, the bottom image consists of all the layers.*

2.2 Example News Information Space Experience

The previous section outlined the motivations for develop Galaxy of News. The resulting system met many of these objectives. In this section, I illustrate, in narrative form, how this is achieved, presenting an example experience of moving through a Galaxy of News information space.

Sitting down to the terminal, we initiate the Galaxy of News front-end application. The application queries a relationship database and begins building an information space that represents today's news. The screen turns black as the system presents us empty space.

Using only a three button mouse for navigation, we press the left mouse button as we move forward in the space. Out of the darkness appears a set of words–*Government, Media, Education, Music, Science, Olympics*–organized along a vertical and horizontal grid, and illuminated in an array of colors (see Figure 2.3). These words indicate categories. Blue lines (indicating connections between categories and subcategories) stem out from each word.

Using the mouse as a movement control, we press the left mouse button and slide the cursor to the left side of the screen. The system responds. The space in front of us unfolds as new words under the initial set appear and subcategories emerge. The word *Media* catches our eye, and we continue to move towards it, exposing the words *Newspapers, Books & Publishing,* and *Television* (among others, see Figure 2.2 for the actual hierarchy under *Media*). Radiating red lines appear, pointing off into the space behind the category they connect to. The red lines jump from category to category until we move up to *Media* (see Figure 2.4), where we pause for a moment. These radiating lines show the connection between the category and the articles related to that category, but at this point the articles are not visible yet.

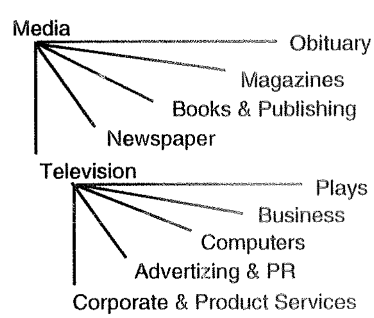


Figure 2.2 *Example of a derived keyword hierarchy*

We continue on our journey, pressing forward into the *Media* space. As we move, the blue lines fade away. The subcategories to *Media* grow larger and become brighter–their colors have a similar hue to *Media*, but slightly more saturated. Simultaneously, *Media* stops growing in size

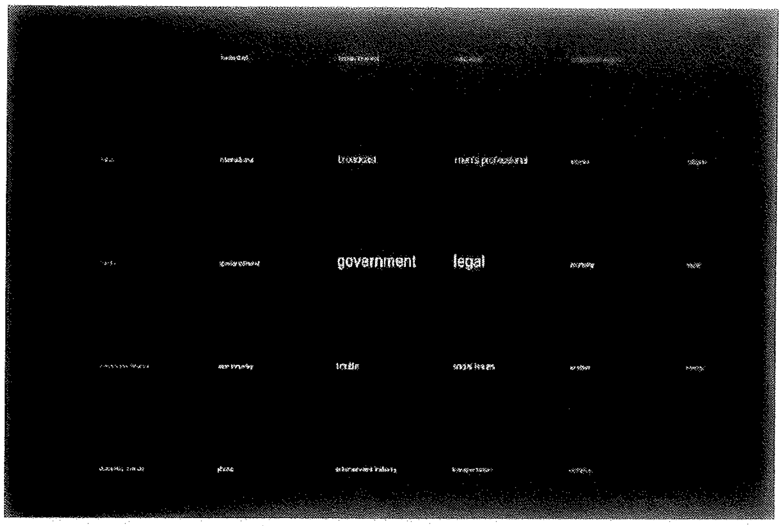


Figure 2.3 Initial view of the news space. It shows a set of root categories of news articles. The blue lines show the parent-child structure of the categories.

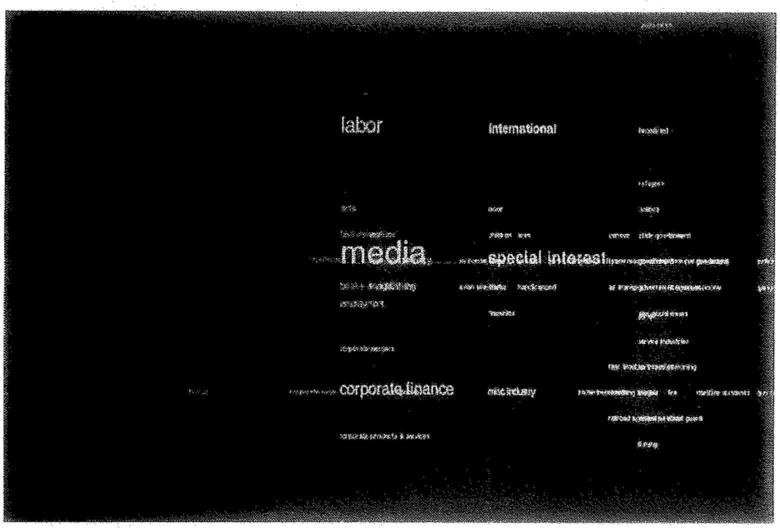


Figure 2.4 The user has begun to zoom into the information space, revealing the subcategories under the "media" cluster.

Figure 2.5 User continues to zoom into the "Media" category as the headlines of the articles associated with Media begin to fade in.

and begins to fade out. Behind *Media*, headlines of the articles begin to fade in (see Figure 2.5).

We hold our position and zoom into the *Media* space. The headlines continue to illuminate until they are no longer transparent (Figure 2.5). We zoom further, experiencing a slight hesitation as the system reaches out across the network to pull in the body of the article, followed by the faint whispers of text under the headline as the body of the article begins to drop in (Figure 2.3). As we continue to zoom into the space the full body of the article drops in (Figure 2.8). In some cases images appear.

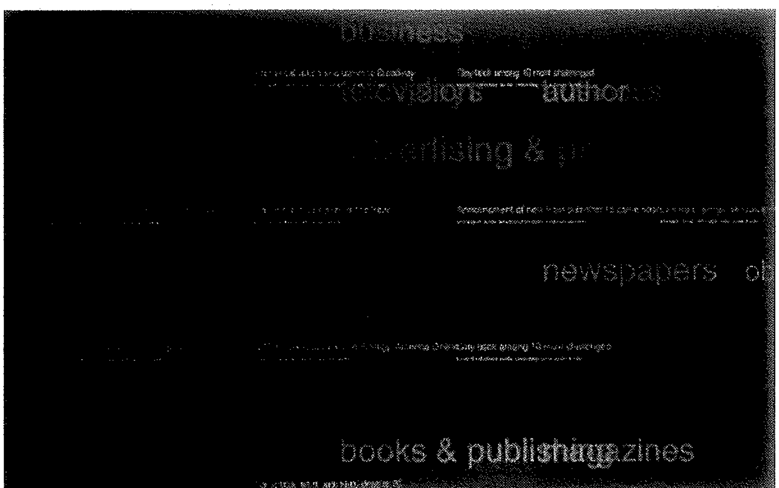


Figure 2.5 Further zooming illuminates the headline completely.

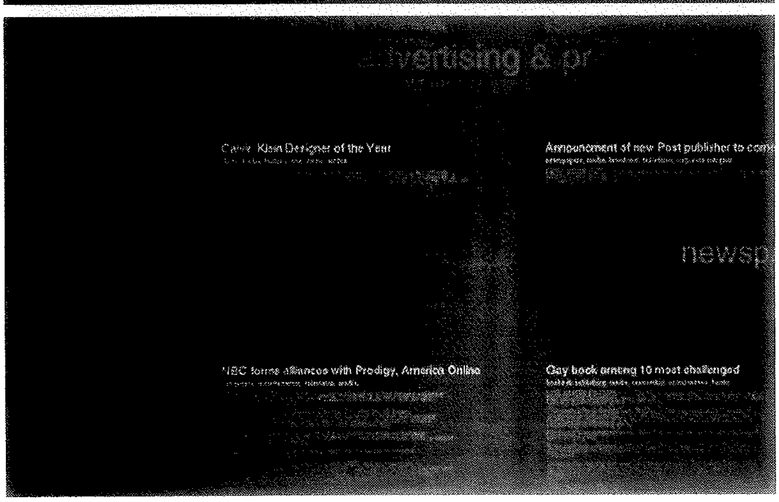


Figure 2.6 As the user zooms closer, the body of the articles begins to drop in.

After briefly browsing through these articles, we decide to reexplore the spaces around *Media*. We back up out of the *Media* space, and the topics *Newspaper, Fashion, Television,* and *Books & Publishing* appear. We move toward the word *Newspaper*. As the word *Newspaper* approaches the center of the screen, the red lines that were radiating out from *Media* jump to *Newspaper*. We have indirectly executed a query to the system, searching for those articles that contain *Media* and *Newspaper*. The red lines show the result of that query, as the selected articles begin to move

Figure 2.7 As the user reaches the back of the space, the full headlines and body of the article are presented.

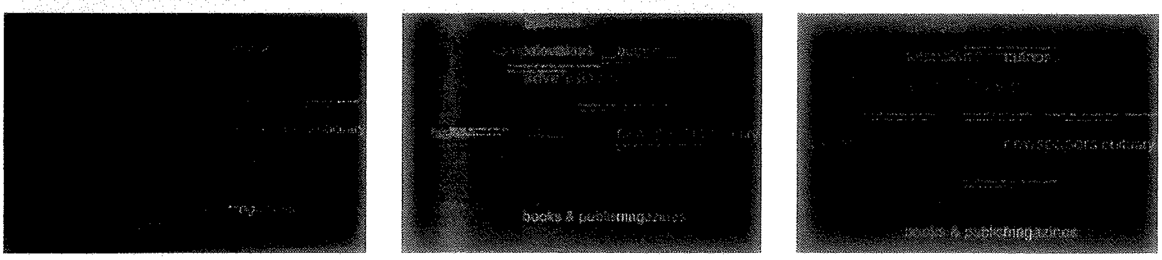


Figure 2.8 *A sequence of images showing the transition from the* Media *space to the* Newspaper *space.*

from their position around *Media* to their respective positions around *Newspaper* (see Figure 2.8). We examine the articles and continue our journey...

2.3 Information Visualization and Interaction

The example above illustrates many of aspects of the experience created through information visualization and interaction in Galaxy of News. Though it is difficult to illustrate with these static images, the experience is fluid and continuous. The system responds dynamically to our movement in the space, presenting information at a readable level of detail for a given position. At every stage of the experience, the system tries to present an appropriate level of abstraction so that the user is not overloaded with information.

The system interprets the users movement relative to objects displayed on the screen. For example, if the user moves close to a word representing a category, a query is automatically formulated. The system dynamically presents the results of this query to the user using animation to show the difference between the old state and the new state resulting from the query. If the user *backs away from* keywords or news articles, the system automatically interprets this action as wanting less detail and more abstraction, responding appropriately.

Visualization and Interaction Characteristics

The Galaxy of New information visualization and interaction approach can be characterized by the following:

- Pyramidal encoding or presentation of news elements to provide progressive refinement of news information
- Visual clustering of news elements based on the content of news articles to provide structured information access
- Abstract three plus dimensional spaces that contain information objects
- Semantic zooming and panning, where zooming is synonymous with searching or filtering, and panning is synonymous with browsing
- Fluid interaction to help understand and maintain the context of the information being presented
- Animation and motion to illustrate relationships between news elements
- Dynamic visual cues to aid in the navigation through an abstract news space
- Dynamic visual presentation of information to present the proper quantity of information at each instance of interaction and to eliminate distracting clutter

These characteristics define an outline for building a structured pyramidal visualization of news, whereby the upper portions of the pyramid consist of general descriptions or abstractions of the lower levels which contain increasing levels of detail. Pyramidal representation offers news readers the ability to progress through a process of glancing, to investigating, to reading details in a fluid and selective manner, while maintaining context of where they are in the process (as illustrated above). Hence, the information is structured such that news readers can gain a good understanding of the full range of news by looking at the top levels of the news information pyramid, and through fluid interaction, gain access to increasing levels of detail.

An important and interesting aspect of the space constructed by the Galaxy of News system is that it is *not* based on any physical metaphors that we encounter on a daily basis such as windows, desks, folders, cabinets, rooms, buildings, streets, books, and so forth. Rather, it is based on abstract conceptual metaphors, e.g. galaxies and solar systems, which we understand, but only on a conceptual level since we do not experience these types of environments in our daily lives. As a result the space is freed from dimensional constraints, and hence, can represent many conceptual dimensions simultaneously. At first one might think that this would be very confusing to a user; however, usage of the system has shown that people have the ability to adapt to this abstract space given that appropriate visual cues are provided to the user.

2.4 Information Space and Interaction Design

The Galaxy of News system explores the separation of news information space design and the authoring of news articles. Rather than creating a final visual product, such as a news paper, designers must design a *process* for creating a final product. In this case the final product is an information space. In this environment, the role of an information space and interaction designer is to specify 1) the types of relationships between news articles and the process of constructing these relationships; 2) rules for constructing a multidimensional spatial layout based on the relationships between news articles, and rules for building constraint networks to dynamically manage the spatial layout; and 3) actions the system takes when the user navigates through the space.

The iterative design process used in the development of the Galaxy of News information space resulted in an information space organized into distinct layered structure. The information space consists of three layers: 1) a hierarchy of keywords that go from general to more specific keywords, 2) headlines of articles, and 3) the body of articles (as illustrated in Figure 2.8). The category structure is derived from the relationships between news articles. An example keyword hierarchy derived automatically from an actual news base is shown in Figure 2.2 above. Because this structure was derived automatically, unlikely relationships were determined, e.g. the relationship between "Obituary" and "Media" which was derived from an article detailing the death of a comic book artist. Also, since the actual structure derived by the algorithm is an acyclic directed graph, keywords may appear in several places in the spatial representation of the information categories. The articles, on the other hand, move dynamically around the space depending upon where the user is within the space.

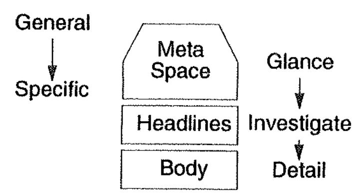


Figure 2.8 Organization of the Galaxy of News information space.

The presentation of the information space is non-linear. The system determines what information is displayed and how it is displayed at each instance the user moves through the space. As the user navigates through the space, the system controls the following parameters:

- Size of the keyword fonts
- Transparency of the keywords (as the user zooms past a keyword, it is kept in front of the user and is faded out over time, which helps with navigation)
- Location of articles within the space, animating the move between locations
- The color of articles as they move between keyword groups
- Line transparencies between
  - Parent and child keywords
  - Keywords and articles, indicating the relationship between the two

- Size of the article headline fonts
- How much of the article body to display, if any, and the transparency of the portion displayed (the body of the articles gradually drops in as the user zooms toward an article)

The designer of the information space must decide how these parameters change as the user moves through the space. The design approach I took in designing the Galaxy of News information space is described in the example presented above.

When I designed Galaxy of News, I decided to fix the position and attributes of some elements in the information space, while varying others. Keyword locations are held constant to maintain a basic sense of structure. The color of the keyword groups remain constant to indicate keyword clusters. These fixed parameters aid the user when navigating through the space by giving the user a sense of where they are within the space without having to provide a global map.

Some elements of the space change dynamically. These elements include: the size of the keyword fonts, the transparency and position of the keyword fonts as the user navigates into the region of a particular keyword, the lines emanating from keywords to articles becoming more transparent as the user navigates close to the articles–also provide implicit navigational aides.

The combined effects of the fixed and dynamic parameters were carefully designed to assist the user in navigating through the space without using explicit navigational aides.

2.5 Related Work

At first glance, Galaxy of News seems similar to cone trees [Robertson, 91]; however, there are several significant differences. First, the hierarchical form is not explicitly presented to the user. The hierarchy is not presented as a static object; it is primarily used to present information to the user at the appropriate time. Second, not all the elements of the hierarchy are visible at a single glance. Rather, only elements that are relevant to the user's present view are shown. This is significant because it allows for an infinitely deep information structure to be presented. Third, the user is able to navigate through the space in an immersive fashion. As the user goes deeper into the structure, the system reveals the substructures.

The process of zooming in the information hierarchy is a form of interactive filtering. A similar approach was explored in the PAD system [Perlin, 93], which provides an infinite two dimensional information plane. One of the main limitations of the PAD approach is that once the space has been constructed, it is rigid as objects have fixed locations on the plane, and hence does not address the multiplicity of relationships between news stories. The Galaxy of News visualization and interaction approach addresses this by dynamically restructuring the space to pull in information relative to a given view. This process is animated to illustrate to the user what the system is doing. In effect, the approach is to construct information worlds within information worlds similar to [Feiner, 90], yet different in that the space is not tied to any dimensions–the space is abstract.

2.6 Discussion and Evaluation

In developing the Galaxy of News system, I considered several alternative approaches. I considered using hypertext and hypermedia concepts and systems to aid in accessing related articles or information in general [Nelson, 81]. This technique has met with some success and has been employed by Netscape and Mosaic as an interface to World Wide Web documents [Andreessen, 93]. However, there are several inherent problems with the hypermedia approach, such as that used with Mosaic.

One of the most significant problems with hypermedia is its "hyper" aspect; the process of jumping to another location in an information space can easily confuse a user. This is primarily a result of the lack of a general, or known, structure of the information available to the user. Unless the author of the hypermedia document clearly presents the structure of the information, the user has no idea what other information is available other than the clues indicated by hot spots or hot text that link one node to another node. Hence, the utility of hypermedia systems are at the mercy of hypermedia content authors. Further, if all of the links between related news articles must be authored by hand, this problem will only expound with the growth of computer connectivity and the amount of news information available.

Lessons Learned

The development of Galaxy of News went through several unsuccessful phases before arriving at the final design. While these approaches were unsuccessful, they did provide some valuable insights into information visualization and interaction, so I present them here.

The design of Galaxy of News evolved through the following stages (each one building on the previous one):

1. *2D simulation of keyword relationships*

   This approach used dynamic clustering of keywords based on a structure that emerged from the contents of news articles. The relationships between keywords were shown by movements relative to other keywords.

   This approach was not effective because the

   - movement was disjoint
   - the movement of the keywords on the screen was far to complex to understand

- the user had no way to declutter the screen, and
- the user could not figure out what was going on because they could not get more specific details.

2. *3D Keywords and Headlines*

   This approach used dynamic clustering of both keywords and article headlines in the same space, and the user could zoom into the space to access more details, reducing the clutter.

   This approach was not effective because it was

   - too difficult to identify clusters
   - too difficult to understand the relative relationships between different elements
   - too complex in general.

3. *3D Spring-Lattice Structure*

   This approach added color coded lines between keywords, between headlines, and between keywords and headlines. The length of the lines between the keywords were defined by the frequency of co-occurrence of two keywords in news articles. A spring simulation was used to show the relative relationships between keywords. When the spring simulation settled in on a resting position, the resulting structure showed the relationships between the keywords and the articles. This in effect simulated the process of Multidimensional Scaling (MDS). Figure 2.9 illustrates several example visualizations taken from this approach.

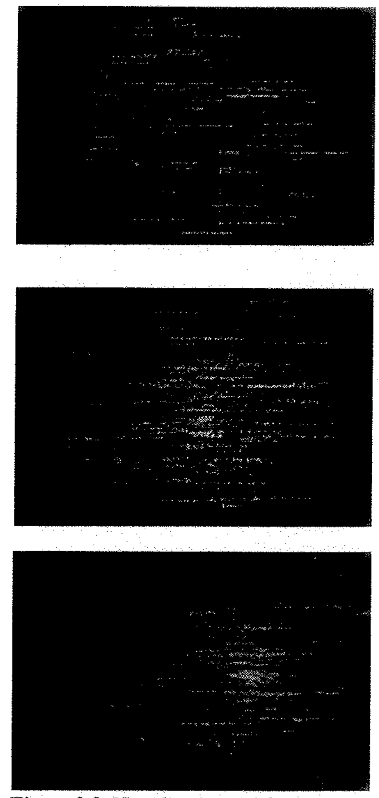


Figure 2.9 Visualizations of the relationships between keywords extracted from news articles. The top image shows the relationship between 20 news articles. The middle image shows the relationship between 75articles. And, the bottom image shows the relationship between 500 articles.

   This approach worked quite effectively when the number of articles was small (as shown in the top image in Figure 2.9), but suffered from several problems as the number of articles grew. For example, when the number of documents rose to a modest 500 articles, the resulting structure was unintelligible. Additionally, the approach suffered from the following problems:

   - difficult to maintain context when zooming
   - can not disambiguate complex relationships
   - ultimately, too many dimensions to understand in one view.

Although these projects failed, they pointed out some essential problems almost all information visualizations systems face and must overcome. First, when we try to visualize large complex information bases, it is not possible to communicate all the details with one image, nor with one structural object, nor with one space. Second, the underlying dimensions of the conceptual space are far to great to compress into a three dimensional space and show then all at one time. Third, without some mechanism to dynamically control the visualization, selecting and disambiguating specific portions of the information structure, we can not discern what the structure means.

In fact, nearly all information visualization systems to date have run into the problems outlined above. Most have not gone past this mark, and as a result are not very effective for visualizing large complex information bases. Most information visualization systems produce a static object represented in a 2D or 3D space. For the reasons described above, this is not sufficient for understanding the complex relationships between information objects.

In developing Galaxy of News, I learned that effective visualization involves a process of presenting a simplified visualization, and then allowing users to change the visualization interactively. The final version of Galaxy of News addressed this problem, giving the user interactive control based on movements through conceptual structures.

Chapter 3

Mind's Eye Approach

Galaxy of News, presented in the previous chapter, formed the foundation for a new approach to visualizing and interacting with a large corpus of information. While this work addressed many issues and suggested possibilities for future human-computer interactions, it also raised a number of questions. The remainder of this thesis addresses these issues and builds towards a new form of human-computer interaction–*visual discourse*.

This chapter is broken down into four parts. First, I present a number of issues that were unresolved with Galaxy of News and discuss the opportunities that the work suggests. Second, I present a conceptual approach to addressing these issues and opportunities. This conceptual framework forms the foundation for the remainder of the thesis. Third, I present an example showing a path way through a historical information-base constructed in the Millennium Project. This example illustrates a Mind's Eye view of a symbolic landscape, highlighting important aspects of the computational approach. And, fourth, I present the computational approach used to implement the conceptual approach and test its validity of this approach. Subsequent chapters provide detailed discussions of both the theoretical and computational approach.

3.1 Problems and Opportunities

While Galaxy of News addressed many problems in visualizing a large corpus of information, a number of problems and limitations persisted. These problems and limitations can be characterized as follows:

- *Visual Form*–The layout of keywords and articles within a given category in a Galaxy of News space has little or no meaning. The grid layout used in Galaxy of News, while providing an organizational form, does not convey any information. Much more information is available to communicate, but we need to find better forms to communicate that information. Can we give form to the organization of the keywords or symbols and articles such that a person can quickly *see* the relationships?

- *Continuous Dialog*–In Galaxy of News the space is fixed and bounded, and hence the level of *depth* of the interaction is limited. Can we create a continuous, progressive dialog that has no limitations on the scope of exploration?
- *Extended Interaction Lexicon*–Only two levels of interaction are supported: 1) zoom in/out 2) move side to side. Given the constraints on movement, we could add a push-through operation, but that is the extent. This limitation arrises from the orthogonal constraints placed on the projection and movement. This constraint was carefully chosen to solved some visual problems. If we free the constraint of orthogonality, we end up with another problem, namely that subgroups do not lie "behind" their respective subcategories. This was the original problem I faced with Galaxy of News and the reason I decided to apply orthogonal constraints to the view. As a result, I needed to rethink the structure of the space. Can we extend the lexicon to support more complex forms of dialog, including point-of-view shifts and context-shifts?
- *Point-of-View*–Only a categorical view is provided in Galaxy of News. We need to extend the structuring techniques to expound more of the underlying structure of the information to give people different ways of accessing information.
- *Immersive Experience*–The Galaxy of News interface does not give the user a sense of changing direction, which is OK when we only have one point-of-view. But, if we add the ability to shift point-of-view or have an endless conversation, we need the ability to change the course of the conversation. Since the projection in Galaxy of News is orthogonal, when the user changes the topic the movement is lateral, rather than axial. As a result, the user does not get a sense that they are *changing the course* of the conversation.

Even with these problems, Galaxy of News suggests the possibility of having an endless visual conversation with the computer, if we can address the problems outlined above. This conversation would use a visual space to communicate, as opposed to the mental spaces used with natural language. Movement through a conceptual space simulates the movement of the *Mind's Eye* as we shift our attention from one conceptual structure to another in our mind.

3.2 Conceptual Framework

In this section I formulate foundations for a new framework and approach to interacting with computers that address the issues outlined in the previous section. First, I present some background information on the general process of understanding with relation to theories of metaphor and embodied cognitive models. Second, I describe the concept of *visual discourse* that forms the foundation of my approach.

Pathways to Understanding and Embodied Cognitive Models

How do we understand information? One postulate put forward by Lakoff and Johnson is that understanding comes from our experience in the physical world by experiencing it: we look, we touch, we move around, and we relate things to our bodies [Johnson, 92]. Information, however, is typically abstract. We cannot literally see it, touch it, move though it, or relate it to our bodies. Yet we are able to somehow structure abstract ideas in such a way that we understand them and can store them in our minds.

Linguistic metaphor theory claims that the way we experience the physical world through our bodies makes its way into the cognitive models we use to structure abstract ideas. Understanding is *embodied* in that our bodily experience directly influences the way we structure thought [Johnson, 92; Strausfeld, 95b]. Metaphor is how we map the concrete cognitive structures or models in our minds to an abstract domain, like information.

The problem of representing the abstract with the concrete is the problem of language. We encounter this problem every time we attempt to express ideas outside the realm of the physical world. In *Metaphors We Live By*, Lakoff and Johnson expose the way language allows us to implicitly (and often sub-consciously) reference our physical and cultural experience in the world to express or understand abstract concepts or ideas. Lakoff and Johnson show that language is based on a conceptual system that is metaphorical in nature. They write: *"The essence of metaphor is understanding and experiencing one kind of thing in terms of another"* [Lakoff, 80].

In her Master's thesis, Strausfeld describes a five-part approach to enabling the understanding of information based on kinesthetic cognitive models [Strausfeld, 95b]. These cognitive models put forward by Strausfeld encapsulate ideas on metaphor theory by Lakoff and Johnson [Lakoff, 80; Lakoff, 87; Johnson, 92]. According to Lakoff, there are four types of cognitive models around which thought is structured.

1. Propositional
2. Image schematic
3. Metaphoric
4. Metonymic

Propositional models specify elements, their properties, and the relations among them (like the *desktop* and its relationship to *file folders*). Image schematic models specify schematic images that operate on a more conceptual level such as *trajectories* or *containers.* Metaphoric models are mappings from a propositional or image-schematic model in a source domain to an analogous structure in a target domain (like the mapping of the desktop model to a graphical user interface). Metonymic models involve the substitution of a part of one of the above model types for the whole, or for another part of the model (e.g. "The White House refused to comment on the issue.")

The first two cognitive models, propositional and image schematic, characterize concrete thought structures while the second two models characterize mental mappings of these concrete structures onto abstract thoughts. The role of my thesis work (as well as that of Strausfeld) is to apply these mappings in reverse. After analyzing and deriving computational structures that represent abstract conceptual structures, we can apply inverse metaphorical mappings to construct visual information spaces that are physical in form, yet representative of their conceptual counterparts.

Visual Discourse

The above embodied cognitive models illustrate how we understand concepts relative to our bodily experience. The dynamic component to these models emerges when we consider the process of discourse used in natural language. One important cognitive model for describing aspects of meaning construction in natural language is the concept of "mental spaces" as developed by Fauconnier [Fauconnier, 94]. From models of natural language discourse, I have correlated and extrapolated a visual and interactive process for exploring information. I call this process "Visual Discourse" and is the foundation of this thesis. I describe in subsequent chapters the computational environment that supports this process.

Mental Spaces and Discourse Representation

The concept of visual discourse is precipitated from Fauconnier's theories on mental spaces. Mental spaces are characterized by the following properties [Fauconnier, 94]:

- Spaces may contain mental entities, such as beliefs, mental images, inference models, etc.
- Spaces may be structured by cognitive models [Lakoff, 87], such as
  - Categories
  - Hierarchical Structure
  - Relational Structure
  - Radial structure
  - Foreground-background structure
  - Linear quantity scales
- Spaces may be related to other spaces by "connectors"
- An entity in one space may be related to entities in other spaces by connectors

- Spaces are extendable, in that additional entities and idealized cognitive models may be added to them in the course of cognitive processing

Fauconnier maintains that we communicate meaning through a process of accessing mental spaces through conceptual connections [Fauconnier, 94]. Examples of this include mappings between source and target domains in conventional metaphor [Fauconnier and Turner, 93; Lakoff, 87; Lakoff and Johnson, 80; Lakoff and Turner, 89]; and discourse involving time, viewpoint and reference [Seuren, 84]. Fauconnier represents discourse as a process where 1) categories are used to set up mental spaces, 2) temporary connections to other spaces are established, and 3) new frames are created dynamically as the discourse unfolds. In this process, participants must keep track of the maze of spaces and connections that are built, and this happens through the use of point-of-view and point-of-view-shifts [Fauconnier, 94].

The Visual Discourse Analog to Natural Language Discourse

Mental spaces and discourse suggest a compelling model for the communication of meaning. As illustrated in Figure 3.1, natural language is used as a medium for building mental spaces as discourse unfolds. In this process, the expert traverses preestablished cognitive models, and the knowledge seeker, listening to the expert, constructs a mental space in his or her mind based on the grammatical structures of natural language elicited by the expert. In this way, the utility of the conversation is measured by the ability of both parties to construct similar mental spaces.

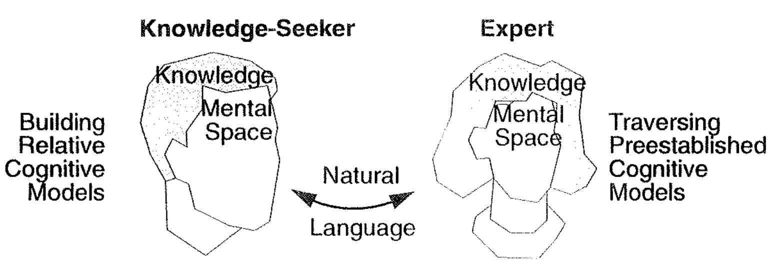


Figure 3.1 Representation of discourse using natural language as the exchange medium.

Our goal is to find the analog of this process, where natural language is replaced with a visual information space and a users movement and interactions in the space form the basis for communication. In this way, visual discourse addresses the communication of meaning through a process of

- *Visualizing relationships between information elements*, where the relationships are illustrated by their relative position in space that has been phrased by contextual information [Fauconnier, 1994; Strausfeld, 1995a][1] (as illustrated in Figure 3.2.a in the example below).

- *Dynamic point-of-view shifts*, where the information elements are held constant and the types of relationships are changed (as illustrated in Figure 3.2.h and Figure 3.2.j), and
- *Dynamic context shifts*, where a new set of information elements are dynamically established based on the constraints of the new context (as illustrated in Figure 3.2.b and Figure 3.2.c).

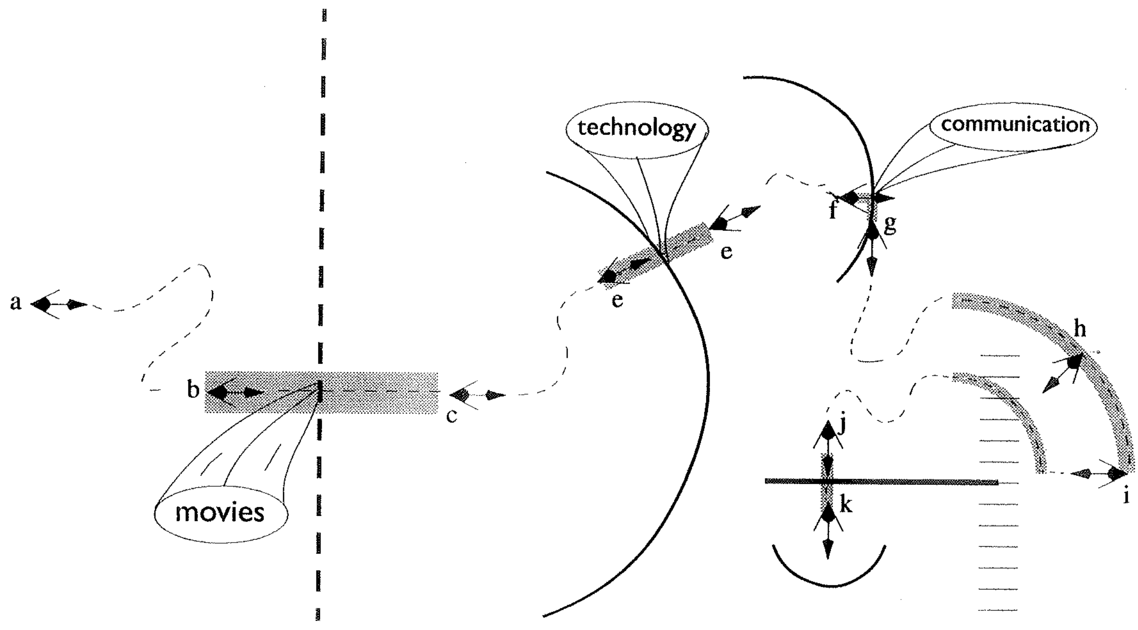
In visual discourse, one participant (in this case the computer) outputs spaces and visual elements in those spaces. In turn, these spaces and visual elements are used as input. For example, if the computer presents the user a set of categorical classifications, each of the symbols that represent the categories can be used as input. Another example is the use of an information object. It can present information on an event, and can also be used to formulate a new query for additional information.

3.3 A Mind's Eye View of a Symbolic Landscape

By addressing the problems outlined above, we can construct dynamic symbolic landscapes that unfold as we move through them, much like the spaces of our mind unfold as our Mind's Eye moves through them. To show what this would be like, let's take an actual tour through the symbolic landscape (constructed automatically by the system that will be described shortly) that represents a historical information-based that contains information objects that describe events, people, and artifacts from the period of 1906 to 1918. This information-base was constructed as part of the Millennium Project[1]. Figure 3.2 presents a map that provides an overview of this tour through a symbolic landscape. Note that each eye/camera icon corresponds to numbered image or image sequence illustrated in Figure 3.2.

Sitting down at a workstation, we initiate one of the Mind's Eye applications, *SymSpace*. SymSpace queries an information-base for the structure of the information contained within. In this case, the information-base contains information on political events, art, philosophy, music, poetry, architecture, design, and science from a pivotal time in our history: the years from 1906 to 1918. The system responds by presenting a symbolic road map of this information (Figure 3.2.a). Some words stand out like landmarks, while others are connected together like roads on a map.

1. An analogical example of this is a 2D graph where the data points, or in our case information elements, are positioned in a space that is phrased by the labels on the x- and y-axis.
1. Amy Schneider, Vivek Palan, Lisa Strausfeld and myself built this database from a variety of sources including Britannica On-line and other web resources.



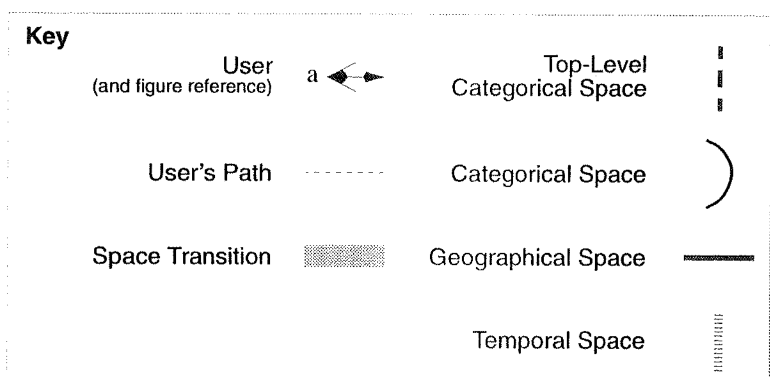


Figure 3.2 A map representation showing a path through symbolic spaces. Each letter next to a camera position corresponds to figures shown on the next few pages. This path through a symbolic landscape shows a) a top-level categorical space, b) a transition to a sub-categorical space, c) the sub-categorical space, d) another transition to a sub-categorical space, e) transition from a categorical space to a geographical space, f) the geographical space, g) transition from a geographical space to a temporal space, h) a temporal space, i) transition from a geographical space to a categorical space, and j) the categorical space.

Using the mouse to navigate, we push forward in the space. We move forward, laterally, up and down with in the space, browsing the options we have before us.

We pick a topic–*movies*–and move up to it. As we approach the word *movies* it highlights, and a set of words appear behind it. We continue to push forward and push into a sub-space formed around *movies* (Figure 3.2.b). As we move into the *movies* sub-space, we notice that the space warps around us and connections between the words that lie behind the word *movies* begin to form. As we press into the space, the connections grow stronger until they form an interesting shape (Figure 3.2.c). We quickly see relationships between topics such as *invention, science, woman* and *technology.*

The topic of *technology* peaks our attention as we wonder what technology was being developed around the turn of the century. To

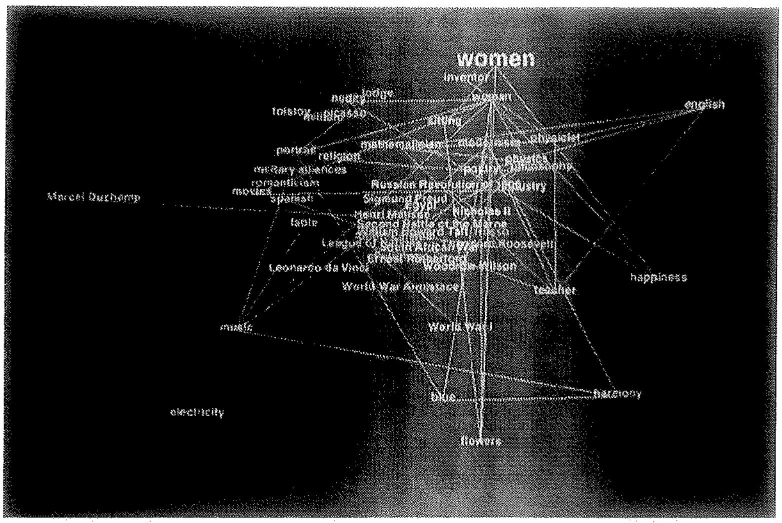


Figure 3.2.a View of the top-level categorical space.

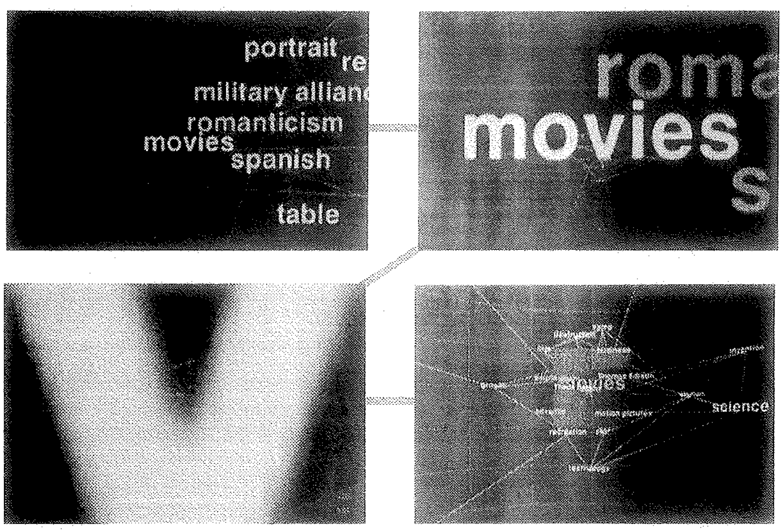


Figure 3.2.b Transition from the top-level space into the movies *sub-categorical space*

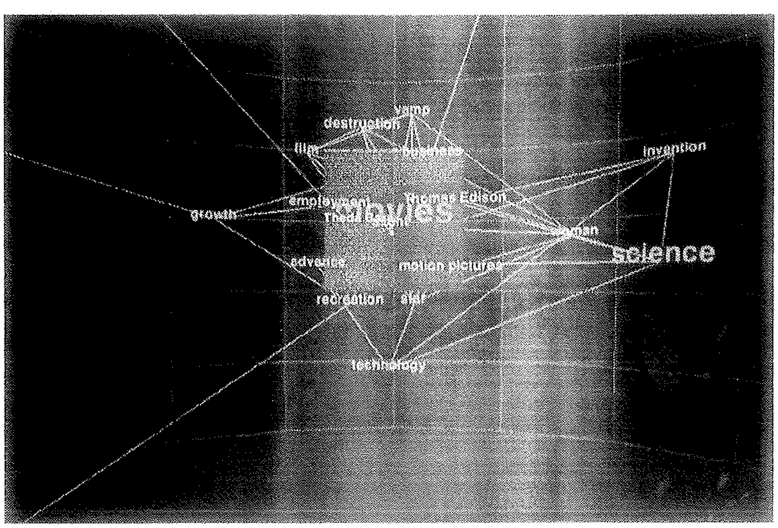


Figure 3.2.c view of the movies *sub-categorical space*

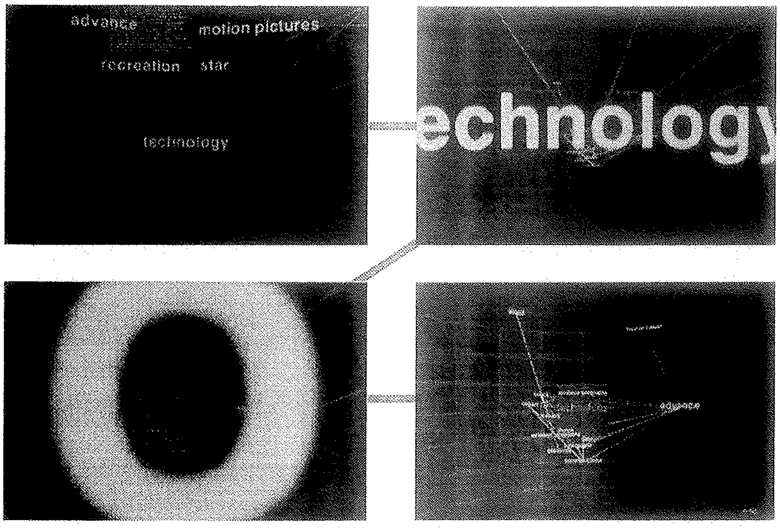


Figure 3.2.d Transition from a categorical space to the technology *sub-categorical space.*

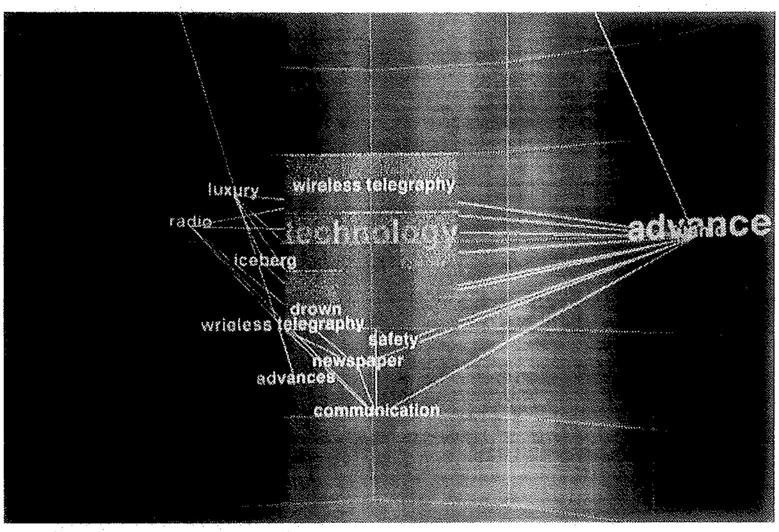


Figure 3.2.e View of the technology *sub-categorical space.*

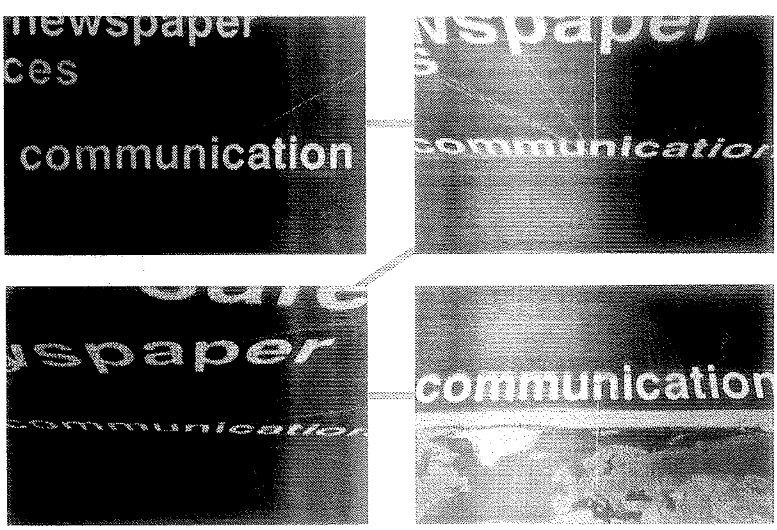


Figure 3.2.f Transition from a categorical space to a geographical space

satisfy our curiosity, we push forward into the space and move up to the word *technology.* As we approach it highlights, and a set of words or symbols appear behind it. We continue to push forward and transition into the *technology* sub-space (Figure 3.2.d). Upon arriving in the *technology* sub-space (Figure 3.2.e), we see a new form relating a set of words together. *Communication, wireless telegraph, radio,* and *newspaper* are among them.

We wonder what the role of communications had on the arts and sciences and wish to investigate this further. We move up to the word *communication.* Again it highlights as we approach it. This time we decide that instead of investigating sub-categories, we would like to see the where events that related to *communication* took place. So, while the *communication* is highlighted, we rotate ourselves about the word. As we do so, a map begins to appear from under the *communication* word (Figure 3.2.f). Stretching out in front of the map is a long line with the word *communication* at the end of it. The line is a thread that ties articles that relate to *communication* together. We investigate the thread by moving up to the word *communication* that now rests above the map (Figure 3.2.g). As we move close to the *communication,* the system takes over, like a roller coaster ride, and we move along the thread examining the articles that relate to *communication.* After examining the articles in the thread, we return to the beginning of the thread.

Curious about the sequence of events that relate to *communication,* we rotate the map on it's side. In doing so, the system responds by transitioning to a geographical and temporal space (Figure 3.2.h). We keep rotating the space until we arrive at a temporal space (Figure 3.2.i). Articles sprinkle the landscape. We move up to then and they unfold before us. Eventually, we transition back to a geographic space. This time we are located above Europe. *Geneva* catches our eye, so we zoom up close to it. We keep moving forward, pushing into the map (Figure 3.2.j). As we push through the map, a symbolic landscape appears before us (Figure 3.2.k), and we continue our exploration...

This example, while arbitrary with respect to content, illustrates what it is move through the space. It illustrates how people can move between spaces and briefly shows how movement relative to words or symbols in the space is the basis for connecting to different spaces. This brief example should give a flavor for the types of dialogs that are possible with a visual discourse.

3.4 Computational Approach

In this subsection I describe a computational approach used to instantiate the types of visual discourse exemplified by the scenario presented above. In the Millennium Project, Strausfeld and myself developed a method of mapping between conceptual structures and virtual spaces. The conceptual structures addressed in this mapping include categorical structures, hierarchical structures, relational

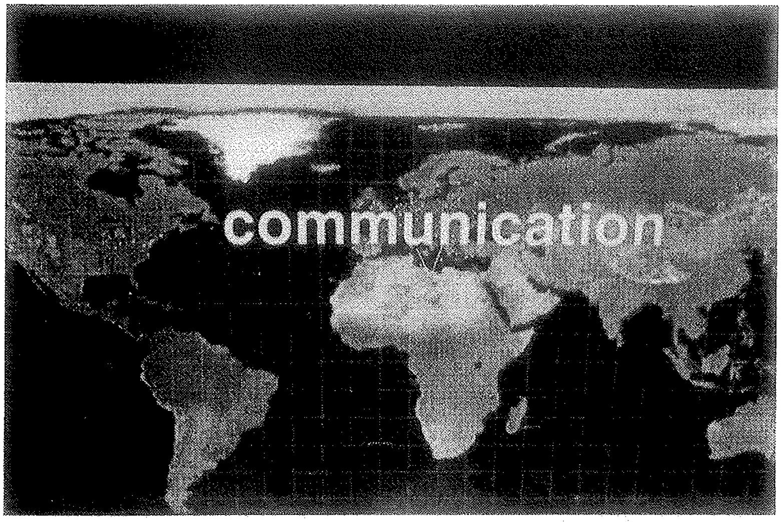


Figure 3.2.g View of a geographical space with the head of a thread indicated by the symbol communication.

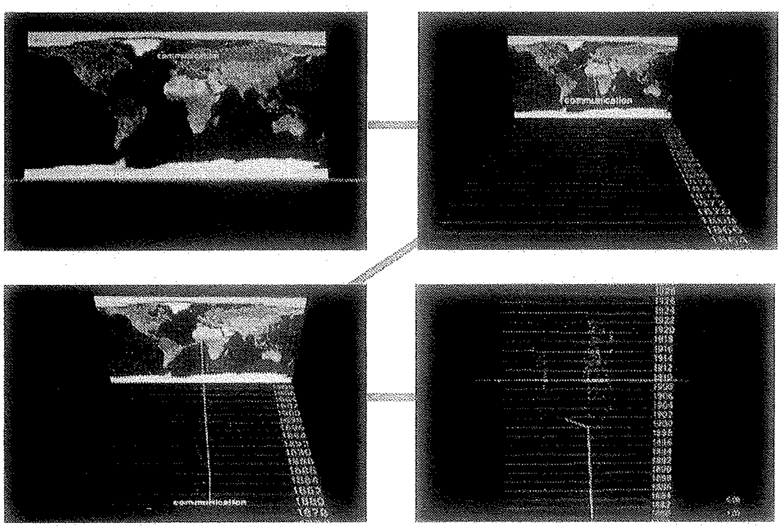


Figure 3.2.h Transition from a geographical space to a temporal space

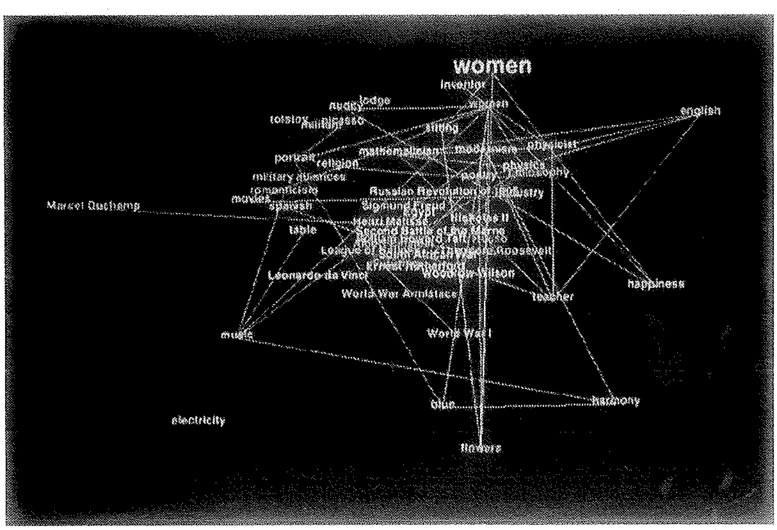


Figure 3.2.i View of a temporal space

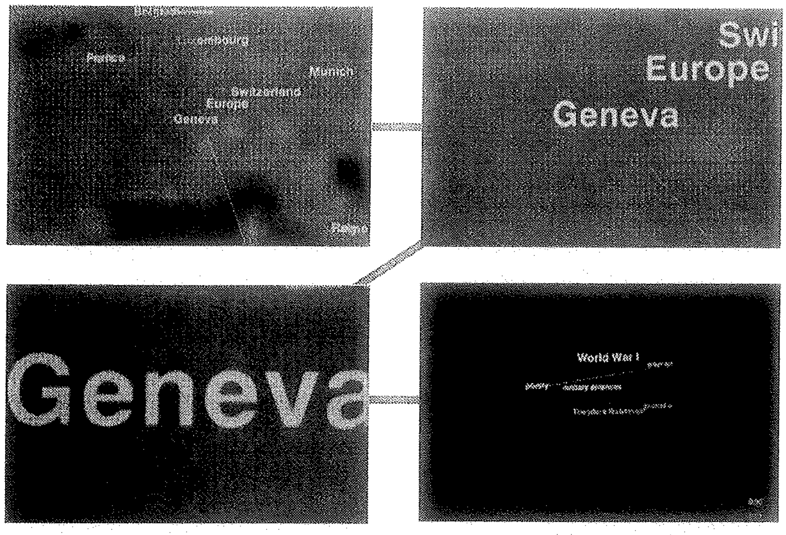


Figure 3.2.j Transition from a geographical space to a categorical space.

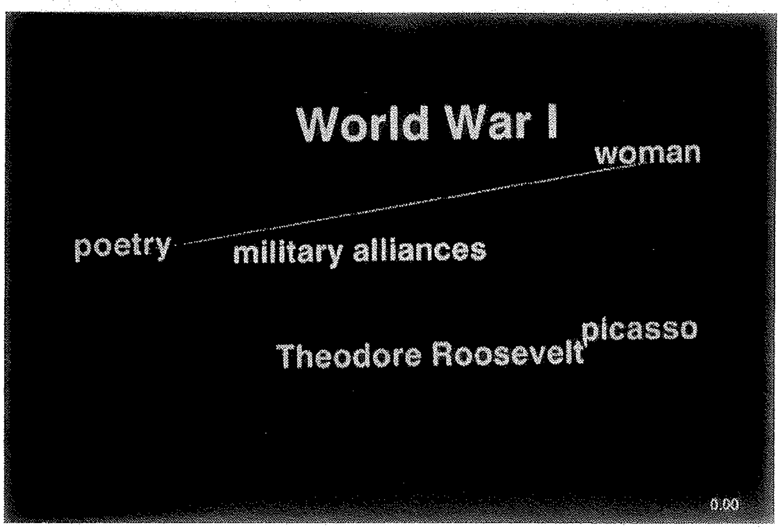


Figure 3.2.k Categorical space after transitioning from a geographical space to a categorical space.

structures, radial structures, linear quantity scales, and foreground-background structures [Lakoff, 87]. Each of these structures help us understand the relationships between information elements.

Table 3.1 presents an overview of a conceptual and computational framework we developed to project information organized in conceptual structures into virtual information spaces. This table shows how conceptual structures relate to the five information organization structures proposed by Richard Saul Wurman [Wurman, 89] are denoted by the LATCH acronym, and include:.

- Location
- Alphabet position
- Time

Table 3.1. Projection from Conceptual Structures into Virtual Information Spaces

Conceptual Structure	Information Organization Structure	Computational Structure	Corresponding Image Schema	Metaphorical Mapping	Virtual Space Mapping
Categorical Structure	Categorical Temporal (i.e. periods)	ARN/Graph	CONTAINER	CONTEXTUAL as INSIDE	Graphical objects as containers that either define space or occupy space (see Figure 3.2.c)
Hierarchical Structure	Hierarchical	Acyclic directed graph	PART-WHOLE UP-DOWN SCALE	GENERAL as WHOLE ABSTRACT as HIGHER IMPORTANT as BIG	Graphical objects as containers inside other containers (see Figure 3.2.c and Figure 3.2.e) Graphical objects scaled relative to importance (see Figure 3.2.a)
Relational Structure	Categorical Temporal (cause-effect) Location (geographical)	ARN TARN LARN	LINK	RELATED as CONNECTED SIMILAR as CLOSE	Graphical objects attached (see Figure 3.2.a) Graphical objects positioned relative to one another using MDS (see Figure 3.2.a, Figure 3.2.c, and Figure 3.2.e)
Radial Structure	Categorical (fuzzy)	Fuzzy cluster graphs	CENTER-PERIPHERY	IMPORTANT as CENTRAL	Spherical, axial and hyperbolic spaces (see Figure 3.2.c and Figure 3.2.e)
Linear Quantity Scales	Hierarchical Alphabetical	Sorted list	UP-DOWN LINEAR ORDER	MORE as UP	Graphical objects viewed sequentially (see Figure 3.2.i)
Foreground-background Structure	Temporal Alphabetical	Sorted list	FRONT-BACK	FUTURE as IN FRONT	Graphical objects viewed sequentially (see Figure 3.2.h)

- Category
- Hierarchy[1].

In addition, Table 3.1 also shows the correspondence between conceptual structures and image schemas (as introduced in Section 3.2), and between conceptual structures and metaphorical mappings. This table presents the computational structures we use to represent the conceptual structures, and shows how we map these conceptual structures to virtual information spaces.

Each of the images shown in Figure 3.2 illustrate aspects of the mapping shown in Table 3.1. For example, Figure 3.2.c shows how a radial and categorical structure is combined to form a visual space. In this image, *movies* is used to define a *container* that delineates the *movies* space. The symbols in this space are objects that indicate subspaces. The symbols are positioned relative to one another using a multidimensional scaling algorithm, and are attached to symbols that are directly related to each other by an article.

I have also developed methods for dynamically presenting the information to the user through "visual discourse", a process that interactively unfolds over time. There are two important aspects of visual discourse: 1) how the conceptual structures are mapped to virtual space (as shown in Table 3.1), and 2) how user interaction is interpreted.

Table 2, *Interpretation of User Interaction,* outlines our[2] approach to interpreting user interactions in a virtual space that is based on metaphorical mappings between a person's experience moving in the physical world and movement in a conceptual or mental space [Rennison, 95b]. It lists the possible user interactions and their effect on the display of objects and contexts as well as the underlying information representation. An information object will display more detailed information up close than it will from far away, for example, and will foreground and background different information from different points of view [Strausfeld, 95a]. The left column lists possible user interactions which consist of movement of self and manipulation of objects. The middle column describes how we interpret user actions based on the cognitive models we outlined in our conceptual framework. The right column describes what changes are made to the current context based on the our interpretation of the user's actions.

1. Wurman refers to hierarchy as the relative size and position of things. This differs from categorical hierarchies.
2. This user interaction interpretation table was developed jointly with Lisa Strausfeld as part of the Millennium Project [Rennison, 95b; Strausfeld, 95b].

Table 3.2. Interpretation of User Interaction

User Interaction		Action Interpretation	Computational Operation
Movement of Self	WITHIN a container	User wants to explore current context	Different views generated within current context
	INTO an object	User wants the object to establish a new context container	NEW SUB-CONTEXT New space constructed Transition performed between old and new spaces
	TOWARDS object(s)	User wants more detail about the object	NEW OBJECT REPRESENTATION by adding detail
	AWAY (BACKING-UP) from object(s)	User wants less detail/ more abstraction	NEW OBJECT REPRESENTATION by removing detail When outside container, NEW CONTEXT, original objects replaced with abstracted representations
	THROUGH an object	User wants next in a sequence	NEW SUB-CONTEXT with next in sequence computed
	OVER, UNDER, AROUND object(s)	User wants to see object(s) from different points of view	Different views generated within current context
Manipulation of Object(s)	Translate: MOVE, PUSH, PULL or DRAG	User wants to see object(s) in different relation to other objects in space	Currently, no operation. (Future work may use object manipulation to generate a new context that user "builds" interactively.)
	Rotate: TURN	User wants to see object(s) from different points of view	CHANGING OBJECT REPRESENTATION based on view angle with respect to context container
	Scale: STRETCH or COMPRESS	If scaling container, user wants to extend or contract context constraints (e.g. time)	NEW CONTEXT generated by extending or contracting constraints mapped to the XY, or Z axes

Key:

NEW CONTEXT: Context generated by establishing new filtering constraints and refiltering the original set of information objects

NEW SUB-CONTEXT: Context generated by adding additional constraints and refiltering the information objects in the current context to generate a new set of information objects

NEW OBJECT REPRESENTATION: Space is restructured by adding or removing related- and/or sub-structures that correspond to the object

The computational process I explored in this thesis consists of the following steps:

1. *Analyzing* the information-base to construct a representation of the relationships between the information objects, namely analyzing the underlying *structure* of the information-base
2. *Presenting* the information relationships in a 3D virtual space that provides a particular contextual view on the information, and

3. *Interpreting* user movements and actions in the 3D virtual space to dynamically query for additional information, and dynamically reconstruct the virtual space to show the relationships between the objects returned from the query.

The relationships between these steps and important subprocesses are illustrated in Figure 3.3. Each of these steps is discussed in detail in the following chapters.

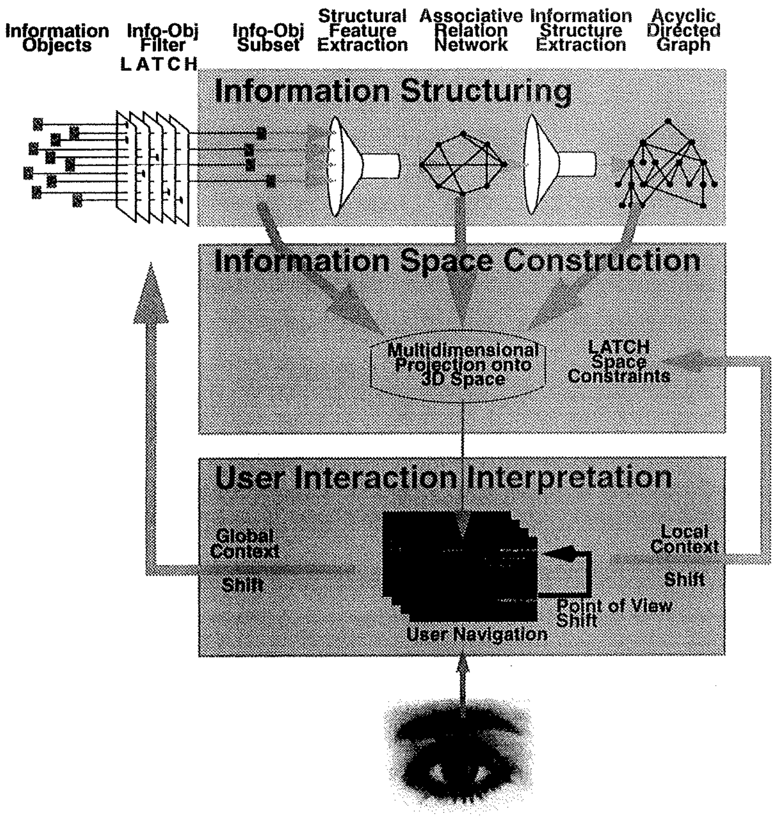


Figure 3.3 The computational process of visual discourse consists of three interrelated subprocesses. The first subprocess, information structuring, filters information objects and derives relational structures from extracted features. the second subprocess, information space construction, consists of projecting computed conceptual structures into a virtual space. The third subprocess, user interaction interpretation, entails reacting to user movements within a virtual space to either aid the understanding of the space or create a new space.

3.5 System Architecture Overview

A prototype system was developed to explore the computational process described above. The architecture of this prototype system is illustrated in Figure 3.4. Each component of this architecture and their relation to the conceptual framework illustrated in Figure 3.4 are described below.

The *information structurer* organizes an unstructured body of information. To perform this operation it builds a structural representation of a selected body of documents by parsing the set of information objects and building relationships between symbols and attributes contained in the objects. How these relationships are constructed is defined by a relationship specification (or procedure) defined by the information space designer. The information structurer interprets the relationship specification, and uses it to build a structure

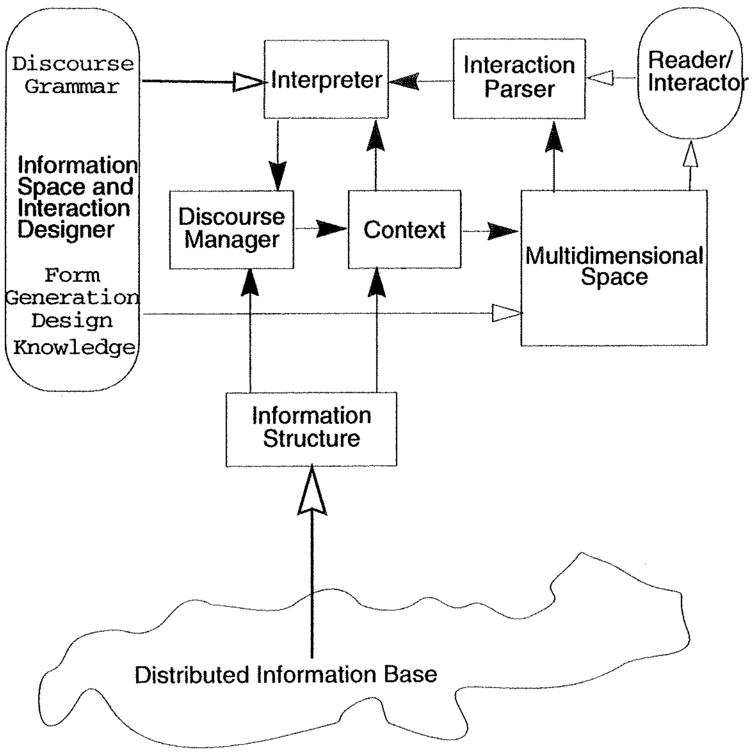


Figure 3.4 System Architecture Overview

representing the relationships between information objects stored in an information-base.

The *interaction parser* takes as input the model of the user's position in space, the user's movement in space, and the current spatial context, and builds a parse-tree based on the discourse grammar provided by an interaction designer. The resulting parse-tree is subsequently passed to the interpreter for interpretation.

The *interpreter* takes as input the parse-tree generated by the parser and outputs a new context, if appropriate. In the process of generating a new context, the interpreter queries the information structurer through the visual discourse manager to derive the information objects and structure of the new context.

The *discourse manager* maintains a history of contexts and their associated spaces, and controls the flow of the visual discourse.

The *multidimensional space* is constructed based on the relationship structures specified by a context for a given point of view. The rules for constructing the spatial layout are defined by an information space designer. In addition, the layout manager controls what and how information is presented at each instance of interaction. Since the space the user navigates through is non-linear, the space must compute a layout each time the user moves within the space. The space utilizes a knowledge-base provided by the information space designer to perform this operation.

The subsequent chapters of this thesis describe in detail the computational process for deriving emergent information structures, building information spaces that illuminate the information structure, and the mechanisms for moving between information spaces. The combination of these three aspects form a visual discourse.

Chapter 4

Deriving Structure

This chapter describes an approach to finding the underlying structuring of an information-base. This structure is intended to help enable people to understand the complex relationships among information objects contained within the information-base. This chapter describes a process for analyzing information objects and deriving structures that convey relationships between the information elements. Information elements in this definition include the original information objects as well as features that are extracted from the information objects (such as symbols or keywords). We use the extracted features to analyze the structure of the information objects (I describe this process below).

4.1 Key Conceptual Structures

Before I embark on describing the process of deriving structure, we must first ask the question–what conceptual structures are effective in communicating the knowledge contained in an information base? What is the role of organizational structures to the process of communicating knowledge? Are certain organization structures more effective than others? To provide a framework for addressing these questions, we explored theories developed in linguistics and cognitive science, including (but not limited to) the work of Lakoff, Johnson, Turner, Fauconnier, Jackendoff, Ortony, and Arnheim. In addition, we also explored techniques used by graphic designers to communicate information visually, including the work of Wurman. The results of this analysis concluded in two classes of structures: one, conceptual structures that correlate to cognitive structures [Lakoff, 87], and two, organizational structures that are used to physically structure and present information [Wurman, 89]. The identified cognitive structures are summarized in [Lakoff, 87] and include

- categories
- hierarchical structures
- relational structures, and
- radial structures.

Richard Saul Wurman in his book *Information Anxiety* identified five information organization structures, for which he defined the acronym LATCH:

- Location
- Alphabet
- Time
- Category
- Hierarchy, or more precisely, a continuum.

Wurman provides numerous examples of how these basic organizations structures are employed[Wurman, 89]. The correlation between the conceptual structures and organizational structures is shown in Table 4.1.

Table 4.1. Mapping from conceptual structures to organizational structures

Conceptual Structure	Information Organization Structure
Categorical Structure	Categorical Temporal
Hierarchical Structure	Categorical Hierarchical
Relational Structure	Categorical Temporal (cause-effect) Location (geographical)
Radial Structures	Categorical
Linear Quantity Scales	Hierarchical Alphabetical
Foreground-background structure	Temporal Alphabetical

4.2 Overview of Computational Process

The objective of this analysis process is to find structures that correlate to cognitive structures such as categories, hierarchical structures, relational structures, and radial structures. These structures will are in turn used in the process of mapping the structural relations onto a visual space that is presented to the user. In addition, these conceptual structures aid the user in navigating through the virtual information spaces, as well as aid in understanding the relationships between information objects describing events and artifacts that span place and time.

As illustrated in Figure 4.1, the conceptual structures are derived through the following process:

1. Filtering the original set of information objects to a reduced subset (via a LATCH filter, optional)
2. Extracting key features from the reduced set of information objects (e.g. keywords)
3. Constructing a computational representation that captures the structural relationships between extracted features and the underlying information objects (e.g. Associative Relation Network)
4. Processing the structural relationship representation to extract computational structures that correspond to conceptual structures (e.g. Acyclic Directed Graph).

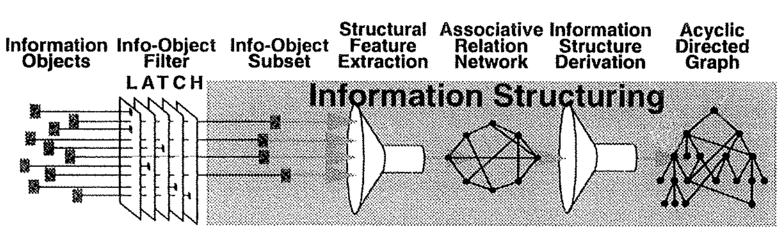
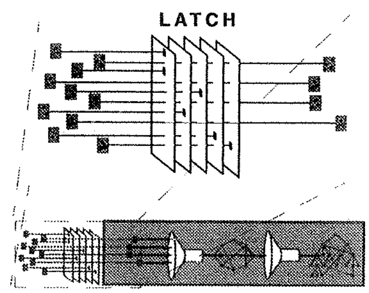


Figure 4.1 Overview of the information structure derivation process and the analysis stages involved

We describe each of these steps in the following subsections.

4.3 Information Object Filtering

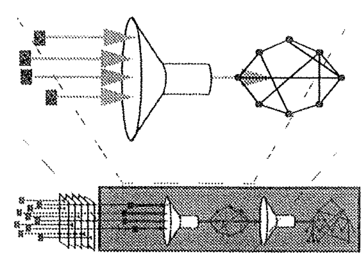


The first step of the structure analysis process is to filter the original set of objects to a reduced set. This essentially establishes the initial or global context for a discourse. This filtering process is based on an initial condition specified by the user. For example, "Let's start will information that pertains to the geographical location of Vienna, Austria, during the period from 1911 to 1912, that fall into the categories of painting and abstraction." This sentence formulates a query or filter that screens information objects to derive a subset of objects. Queries for information objects are either made explicitly, via a text entry mechanism such as a dialog box, or through implicit interaction within an information space. Implicit information queries, which are based on users' movements in the information space, are describe in more detail in Chapter 7.

We have defined a filtering process based on Richard Wurman's five methods for organizing information, as described above. We call our initial filter a "LATCH Filter." In this initial stage, objects are passed through a LATCH filter to establish the initial set of information objects.

It is also important to note that this filtering stage is optional. If the user does not specify an initial condition, the entire database is used as an initial context and the following process continues from there.

4.4 Extracting Key Features



The second stage of the analysis process is to extract key features from the information objects. These features include such information as the dates/duration that an event occurred, location an event occurred, and sets of symbols that describe the information object (refer to Figure 4.2). The symbols in this case refer to elements such as nouns, noun phrases, verbs, and verb phrases that describe the subjects, actions, and objects of the information context. They may also include constructs such as Universal Record Locators (URLs) and names.

In the implementation describe in this thesis, we allow for three levels of feature definition: 1) features extracted from the content or body of the information object, 2) features defined by an object annotator, and 3) features associated with the object by the end-user, or knowledge seeker. Each of these features are treated separately and the user has control over how the system applies them in constructing the information spaces.

The features fall into two categories:

- general properties, and
- structural relations.

```
<ODFile 0.9>
<ObjType people>
<ObjName `alma-mahler.html'>
<Annotator (Lisa Strausfeld, Earl Rennison)>
<Author `'>
<Location (`Vienna, Austria', `New York, New York, USA')>
<Date (`Aug. 31, 1879', `Dec. 11, 1964')>
<Source `Britannica On-line'>
<!-- Association Sets that describe this object -->
<AssociationSet Subjects ((music, art, piano, writer),
  (woman, marriage, wife, divorce, relationships, affairs, love),
  (Mahler Symphony No. 6, Mahler Symphony No. 8, The Tempest
  Wozzeck, And the Bridge Is Love),
  (Gustav Mahler, Oskar Kokoschka, Gustav Klimt, Walter Gropius,
  Franz Werfel, Arnold Schoenberg, Gerhart Hauptmann,
  Enrico Caruso, Alban Berg)) >
<AssociationSet Influenced (Gustav Mahler, Oskar Kokoschka,
  Gustav Klimt, Walter Gropius, Franz Werfel) >
<TITLE>Alma Mahler</TITLE>
<H1> Alma Mahler </H1>
(b. Aug. 31, 1879, Vienna, Austria-Hungary--d. Dec. 11, 1964, New
York, N.Y.,U.S.) <p>
Alma Mahler (also known as Alma Maria Schindler, Alma Gropius,
and Alma Werfel) was wife of Gustav Mahler, known for her
relationships with celebrated men. <p>
The daughter of the painter Emil Schindler, Alma grew up
surrounded by art and artists. She studied art and became friends
with the painter Gustav Klimt, who made several portraits of her.
Her primary interest, however, was in music: she was a gifted
pianist and studied musical composition with Alexander von
Zemlinsky. <p>
In 1902 she married Gustav Mahler, who at first discouraged her
from composing; he is said to have changed his mind after hearing
her songs. Mahler left a musical portrait of her in the first
movement of his Symphony No. 6, and he dedicated Symphony No. 8
to her. After his death in 1911 Alma had an affair with Oskar
Kokoschka, who painted her many times, most notably in "The
Tempest" (1914; "Die Windsbraut"). In 1915 she married the
architect Walter Gropius; they were divorced after World War I.
She married the writer Franz Werfel in 1929. In the late 1930s
the Werfels left Nazi Germany, eventually settling in the United
States. <p>
During her lifetime Alma Mahler became friends with numerous
celebrated artists, including the composer Arnold Schoenberg, the
writer Gerhart Hauptmann, and the singer Enrico Caruso. The
composer Alban Berg dedicated his opera Wozzeck (1921) to her. <p>
Alma Mahler published two collections of Gustav Mahler's letters
as well as her memoirs, And the Bridge Is Love (1958). She also
published a number of songs. <p>
```

***Figure 4.2** Example Information Object File in HTML format.*

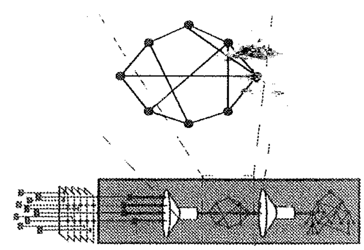
General properties include information such as size, date/time, location, and so forth, and are measured according to an absolute reference[1] such as a time according to the Roman calendar and geographic location. The general properties of the information objects vary according to the type of object. For example, information objects that pertain to artifacts may contain a size of the artifact, date produced, location produced, and who created it. Information objects that pertain to events would not include a size (unless some conceptual size can be specified), the date may be specified as a duration, the location may be specified as a region that may change over time, etc.

1. Absolute reference in this context refers to a reference that is stationary relative to other reference frames, e.g. the death of Christ as a reference frame for time, and the center of the earth as a reference for location.

Structural information consists of sets of symbols that indirectly bind an information object to other related objects. These elements are relative measures or characterizations, and are relative to one another. For example, symbols are defined and classified relative to other symbols.

We extract key symbols and symbol sets from the contents of textual information via one of three techniques. First, we provide a mark-up language that allows authors or annotators to explicitly embed specifications of AssociationSets[1] in the body of an information object description file (as illustrated in Figure 4.2). These AssociationSets can have a hierarchical structure such as the "Subject" AssociationSet illustrated in Figure 4.2. This hierarchical structure is similar to the sentence-paragraph-section-chapter-book type structures that bind words together, but operates on the principles of association as opposed to grammatical structures[2]. Second, we can use automatic text indexing techniques based on symbol frequencies to extract keywords from a text document. And, third, we can use a part-of-speech tagger [Brill, 1992] to identify the nouns, noun phrases, verbs, and so forth.

4.5 Constructing Relationship Representation



Once we have extracted important features from the documents, we use these features to construct a representation that captures the *emergent relationships* between the information objects. A key element of my research is to find emergent structural properties that are not globally or explicitly defined, but rather emerge from the amalgamated properties of the individual objects. Hence, I do not impose a global structure on the information spaces; they are derived automatically from the contents of the information-bases through this bottom-up structuring process.

General Structure

In this thesis, I specifically use associative relations that define co-occurrences of symbols as the basis for our structural representation [Rennison, 94]. In addition, I also use temporal sequence relationships, and geographical and absolute temporal parameters (as specified by the authors of the information objects) to build a representation of the underlying structure. Figure 4.3 illustrates my core representation of the information structures. As described above, each information object can contain a set of dates, a set of locations, and associated sets of symbols (AssociationSets). When these sets of symbols, dates, and locations are

1. AssociationSets are sets of symbols (such as keywords and URLs) that co-occur and are bound together for some structural or grammatical reason, such as a sentence. The symbols in the list can also have weights and counts.
2. Future research would involve analyzing the emergent properties of amalgamated grammatical structures, such as the subject-action-object relationships.

inserted into the core representation they strengthen weights between the symbols, dates and locations.

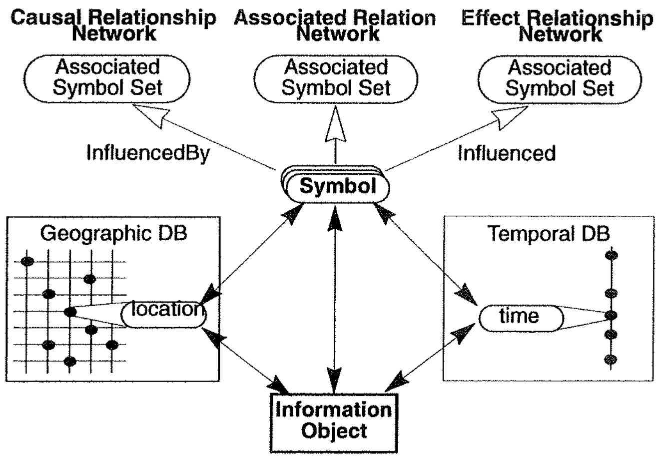


Figure 4.3 Information Relationship Representation

As noted above, features extracted from the information objects fall into two categories: one, absolute values, and two, relative values. The absolute values are maintained in a standard database. The relative measures are maintained in associative relation networks, describe below.

In addition, each symbol in the representation has a reference to all the locations and times that the symbol occurred as defined by an information object. Likewise, each location and time has a reference to associated symbols, and back to the information objects that contain the location or time. The locations are also stored in a geographic database that facilitates quick filtering and searching of either symbols or information objects. Times are stored in a temporal database that facilitates quick filtering and searching for related symbols and information objects.

Associative Symbolic Relationships

The basic structure for representing relationships between information objects is called an Associative Relation Network, or ARN. An ARN captures weighted relationships between co-occurring symbols extracted from information objects. This structure is based on the principals of association of ideas:

> Association of Ideas: In psychology, the conditions under which one idea is able to recall another to consciousness. These conditions may be classified under two general heads, the *law of contiguity*, and the *law of association*. The first states the fact that actions, sensations, emotions, and ideas, which have occurred together, or in close succession, tend to suggest each other when any one of them is afterward presented to the mind. The second indicates that the present actions, sensations, emotions, or ideas tend to recall their like from among previous experiences. On the physical side the principles of

association correspond with the physiological facts of reexcitation of the same nervous centers[1].

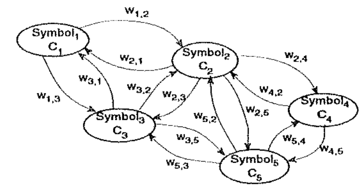


***Figure 4.4** Simplified Associative Relation Network Representation*

An ARN, illustrated in Figure 4.4, maintains weighted relationships between symbols contained in the network, and the relationship between symbols and the documents to which they relate.

An ARN defines an N-dimensional space that contains N^2-N terms (though many of the terms are null, resulting in a sparse space). The basis vectors of the space are defined by symbols extracted from the information objects. Associated with each basis vector (i.e. symbol) is a vector that defines the relationship between itself and all the other basis vectors (i.e. symbols). With an ARN, the information objects reinforce the associative weights between symbols that represent the relationships between information objects. A symbol also forms a link between objects. However, the link that the ARN forms between information objects is not a simple index between information objects. It contains structural information that determines the strength of the relationship between the objects.

An ARN is described as follows: For a given set of documents D, there exists a set of symbols S. The frequency of occurrence for symbol S_i, is defined as

$$c_i = \sum D_x\big|_{S_i \in D_x}$$

where, $D_x\big|_{S_i \in D_x}$ denotes a document containing S_i

The weighted relationship between S_i and S_j in a symmetric network is defined as

$$w_{i,j} = w_{j,i} = \sum D_x\big|_{S_i, S_j \in D_x}$$

where, $D_x\big|_{S_i, S_j \in D_x}$ denotes a document containing both S_i and S_j

With an ARN, the documents reinforce the associative weights between symbols that represent the relationships between documents. And, a symbol forms a link between documents. However, the link that the ARN forms between documents is not a simple index between documents. It contains structural information that determines the strength of the relationship between the documents.

In effect, an ARN is used to learn about the structure of an information base. When an ARN is created it has no previous knowledge about the symbols that are used to construct the network, and as it sweeps through

1. Webster's New Universal Unabridged Dictionary, p. 113.

the information base, it learns the relationships between a set of symbols contained within the database. It has properties of a neural network as well as properties of a semantic network. This gives it the ability to learn as well as maintain semantic relationships.

Temporal Associative Relation Network

The ARN described above is also used to capture relative temporal relationships between information objects, and implicitly the cause and effect relationships between information objects. Our current information object mark-up language allows authors and annotators to specify sets of symbols that the subject of the information object was *influenced by*, and a set of symbols that the subject of information object *influenced* (as shown in Figure 4.3). Each of the *influenced by* symbols are associated with each of the symbols that describe the information object, and these relationships are maintained in a separate ARN that also maintains the temporal distance between associated symbols. We call this extended ARN a Temporal ARN, or TARN. Likewise, each of the *influenced* symbols are associated with each of the symbols that describe the information object. This relationship is maintained in a separate TARN.

A TARN can also capture both *explicit sequences* of events, or *implicit sequences* of events. Explicit sequences of events are those specifically specified by an information object such as a video clip. In a video clip, Subject X may be discussed, then Subject Y, followed by Subject Z. This specifies an explicit sequence of subjects, namely X->Y->Z. An implicit sequence can be specified by placing information objects on a single event timeline and then examining the temporal ordering of subjects. An example of this includes sequences of e-mail messages exchanged over a period of time.

The temporal extension to an ARN for an explicit sequence of subjects is defined as follows: For a given set of documents D, there exists a set of n symbols, S, where the symbols in S are ordered $S_1, S_2, S_3, \ldots, S_n$. The weighted relationship between S_i and S_j, is defined as

$$Aw_{i,j} = \sum D_x\big|_{D_x \supset S_i, S_j}$$

where, $D_x\big|_{D_x \supset S_i, S_j}$ denotes a document containing S_i *followed* by S_j,

and

$$Bw_{i,j} = \sum D_x\big|_{D_x \supset S_i, S_j}$$

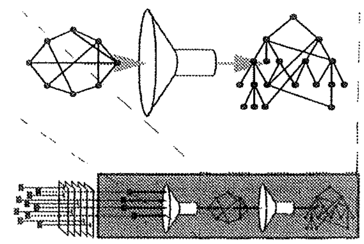
where, $D_x\big|_{D_x \supset S_i, S_j}$ denotes a document containing S_i *proceeded* by S_j

The weighted relationship Aw maintains the relationship between symbols that occur one *after* another. And, the weighted relationship Bw

maintains the relationship between symbols that occur one *before* another. These relationships can be used to learn sequences of subjects (defined by symbols), as well as hierarchical structures of temporal events.

The primary utility of this representation is the ability to compute probability, similarity, and distance measures between symbols and information objects. These measures are used in computing categorical classifications, fuzzy clusters, hierarchical structures and sorted lists as described in Table 4.1. The complex representations described above are dynamically processed to extract these structural relationships that are implicitly maintained by the representation. This process is discussed in the next subsection.

Computing Conceptual Structures



The most important step of the structuring process is deriving computational structures that correspond to conceptual structures and implicitly define structural relationships between information elements. We specifically compute the following computational structures:

- *graph,* where each node in the graph corresponds to a category[1] and linked nodes correspond to related symbolic categories
- *acyclic directed graphs,* where each node in the graph corresponds to a symbolic category and linked nodes correspond to symbolic sub-categories
- *fuzzy cluster graphs,* where each node in the graph corresponds to a symbolic category and linked nodes correspond to related symbolic categories such that the node is the central theme (as in a conceptual radial structure)
- *sorted lists,* where each node represents a place in some linearly ordered sequence or scale
- *circular sorted lists*, where each node's neighbors are the most similar nodes.

The relationship between these computational structures and their corresponding conceptual structures and information organizational structures is shown in Table 4.2. The next three subsections of this chapter describe how these structures are computed.

1. Note that in some cases the nodes may correspond to times, locations or the information objects depending upon the type of conceptual or information structure we are generating.

Table 4.2. Relationship between conceptual, organizational and computational structures

Conceptual Structure	Information Organization Structure	Computational Structure
Categorical Structure	Categorical Temporal	ARN/Graph
Hierarchical Structure	Categorical Hierarchical	Acyclic directed graph
Relational Structure	Categorical Temporal (cause-effect) Location (geographical)	ARN TARN LARN
Radial Structures	Categorical	Fuzzy cluster graphs
Linear Quantity Scales	Hierarchical Alphabetical	Sorted Lists
Foreground-background structure	Temporal Alphabetical	Sorted Lists

4.6 Structure Derivation Algorithms

Basic Measures

Several basic measures are used at various times within the system to perform operations such as classify, sort, spatially position and organize, assign colors, and cluster information objects and their features. These measures are broken down into three classes:

- probabilities
- similarities, and
- distances

We calculate these measures for symbols, symbol sets (or associative sets), and associative relation networks. We describe how each of these measures are calculated below. How these measures are applied is described in subsequent section of this chapter, as well as in the next chapter on visualizing the derived conceptual structures.

Probability

Probabilistic measures are useful for computing classifications of features. The primary measure for computing categorical classifications is the probability measure, $P(a|b)$, or the probability of symbol a occurring given that symbol b has occurred in an information object. The relationship between the set of documents containing symbol a, as denoted by A, and the set of documents containing symbol b, as denoted by B, is illustrated by the Venn-diagram shown in Figure 4.5. P(a|b) is computed as follows:

$$P(a|b) = P(A|B) = P(A \cap B)/P(A) = f_{A \cap B}/f_A \quad \textbf{(EQ 4.1)}$$

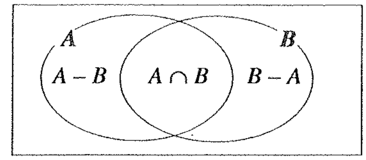


Figure 4.5 *Venn diagram showing relationship between two sets of symbols*

It is also useful to measure the probability of a symbol-set[1], S_A, given a symbol or another symbol-set, S_A. This measure is calculated as follows:

$$P(S_A|S_B) = P(\bigcap A_i \cap \bigcap B_i)/P(\bigcup A_i) \quad \textbf{(EQ 4.2)}$$

Likewise, the probability of an ARN given another ARN (i.e. the probability that one set of information objects will have the certain structure given another structure) can be computed as follows:

$$P(ARN_A|ARN_B) = P(\bigcup ARN_{A_i})/P(\bigcap ARN_{A_i} \cap \bigcap ARN_{B_i}) \quad \textbf{(EQ 4.3)}$$

Similarity

Similarity measures are useful for retrieving similar information objects, filtering objects, and clustering objects. Similarity between information objects and symbols can be computed in one of two ways: 1), using a geometric model (i.e. a vector space model)[Deerwester, 90]; or 2), using a set-theoretical approach[Tversky, 77]. In the set-theoretical approach the probability of co-occurrence is taken as a similarity measure. The geometric model is used frequently in information retrieval applications [sulton, LSI, etc.]; however, for purposes of classification, the set-theoretical approach is more appropriate. In this thesis, we implemented both approaches, applying them selectively based on their merits. We briefly describe the geometric model below.

With the geometric model, symbols are assigned a coordinate axis, and information objects are represented by points in the resulting coordinate space, as illustrated in Figure 4.6. The similarity between two information objects is computed by taking the dot product of the two normalized vectors representing the information objects, namely the normalized projection of one set of features onto another. This works fine for comparing two information objects, but when two symbols are compared, the resulting similarity is zero by definition. In other words, when symbols are used as the basis of a space, sets of symbols can be compared for similarity (e.g. information objects are represented by a set of symbols); however, two symbols can not be compared for similarity. Hence, a vector space model is not sufficient for computing similarities between meta features such as symbols.

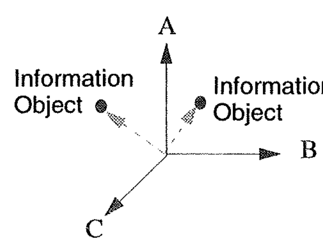


Figure 4.6 *Geometric model for representing information objects in a symbolic coordinate space*

An ARN, on the other hand, does capture similarity information between symbols. In an ARN, each symbol is assigned a coordinate

1. Examples of symbol-sets the set of symbols associated with an information object, or the set of symbols associated with another symbol as represented in an ARN.

axis, as done with traditional geometric models; however, each symbol also contains the following additional information:

1. a basis vector that aligns along the coordinate axis and has a magnitude proportional to the frequency of the symbol in the collection of information objects, and
2. a vector consisting of components from associated symbols, where each component aligns along the associated symbols axis and has magnitude proportional to the frequency of co-occurrence of the basis symbol and the associated symbol.

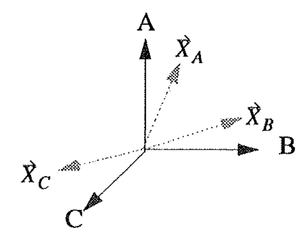


Figure 4.7 Symbolic coordinate space with vectors that relate symbols, assigned to coordinate axis, to other symbols

These properties of an ARN are illustrated in Figure 4.7. The vector $\vec{X}_A$ denotes the vector that relates symbol A with all of the other symbols:

$$\forall (S_i \in A) \qquad \vec{X_A} = \sum w_i \vec{S}_i$$

With this characteristic, we can compute the similarity between symbol A and symbol B as follows:

$$S(A,B) = \vec{X}_A \cdot \vec{X}_B \qquad \textbf{(EQ 4.4)}$$

The similarity between two associated symbol-sets is

$$S(S_A,S_B) = \vec{C}_{S_A} \cdot \vec{C}_{S_B} \qquad \textbf{(EQ 4.5)}$$

where,

$$\vec{C}_{S_A} = \sum \vec{X}_{A_i}$$

The similarity between two ARNs is

$$S(ARN_A,ARN_B) = \vec{C}_{ARN_A} \cdot \vec{C}_{ARN_B} \qquad \textbf{(EQ 4.6)}$$

where,

$$\vec{C}_{ARN} = \sum \vec{X}_i$$

Distance

A distance measure is useful for spatial layout of information objects and symbols. A distance measure is typically considered to have an inverse relationship to a similarity measure, and in many psychological experiments distance has a constant inverse proportional relationship. In this thesis, we compute distance using the same geometric model describe above, and can compute the distance between information objects, symbols, symbol sets, and ARNs. Each of these constructs is represented by a vector, $\vec{V}$, consisting of weighted symbolic components:

$$\forall (S_i \in A) \qquad \vec{V_A} = \sum w_i \vec{S}_i$$

Given this representation, the distance between construct A, represented by $\vec{V_A}$, and construct B,

$$D(A,B) = |\vec{V}_A - \vec{V}_B|$$

$$D(A,B) = \sqrt{\sum_{i=0}^{N} (w_{A,i} - w_{B,i})^2} \qquad N = dimensions \qquad \text{(EQ 4.7)}$$

Categorical Classification

An important conceptual structure for organizing information is a categorical classification of information objects. Objects can be classified into clusters (flat organizations of objects that contain similar features), hierarchical organizations, or acyclic directed graphs. Of these organizational structures, an acyclic directed graph is most desirable because it contains properties of clusters and hierarchical organizations, while also allowing categories to be classified into two different groups (i.e. multiple inheritance). For this reason, we have developed several algorithms to classify symbols, and hence information object from which they were extracted, into an acyclic directed graph.

We use several techniques to compute acyclic directed graphs. These techniques fall into two categories: clustering and probabilistic search and sort algorithms. Within these two categories we use two primary techniques: top-down and bottom-up. The clustering algorithms use similarity and distance measures calculated from an ARN. The probabilistic sorting techniques use probabilities measures computed from an ARN [Rennison, 94]. The probabilistic categorical classification algorithms are described in below.

The information hierarchy resulting from the categorical classification process is used to aid the user in navigating through information structures. This process essentially defines a technique for abstracting and generalizing. As the philosopher William James noted "we acquire knowledge through a process of differentiating characteristics. This process of differentiation is based on finding dissociations between elements" [Arnheim, 69]. This process captures the essence of this objective.

Currently, we compute a fuzzy categorization graph by first computing an acyclic directed graph using a top-down probabilistic approach. Then, we apply a clustering algorithm using each node in the graph as a centroid and searching for all symbols that fall within the range of the symbol, where the range is defined as the farthest distance from the node symbol to a child symbol.

The result of the computational processes described above is a set of computational structures that map to conceptual structures. In the next chapter, we describe how these computational structures are used to construct spaces that reflect the underlying conceptual meaning.

Top-Down Categorical Classification

There are essentially two approaches to categorical classification, *top-down* and *bottom-up*. The top-down approach is used to classify an entire collection of information objects, while the bottom-up approach is used to *abstract from* an information object. Each of these algorithms is described below.

The top-down approach to categorical classification starts from an entire set of information objects, and generates an acyclic directed graph that represents the classification for the entire collection. This approach has several important functions, including 1) providing a way of visualizing an entire collection of documents at selective levels of detail, and 2) providing high level entry points into the information base without requiring prior knowledge of the information base.

The top-down categorical classification algorithm relies heavily on the ARN described above for both the mathematical measures as well as its structural organization that accelerates the classification process. The basic structure of the algorithm consists of two parts:

1. Search through the ARN to find roots to the graph (that is the set of symbols that are the most abstract), and
2. For each root, find the subgraphs or sub classifications.

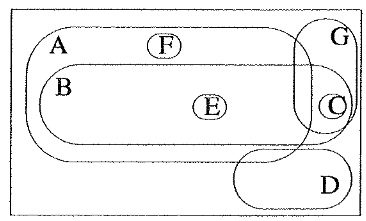


Figure 4.8 *Venn diagram showing relationships between sets of symbols extracted from a collection of information objects*

To further describe this algorithm, let's consider the an example set of documents containing symbols *a, b, c, d, e, f* and *g*, where *A, B, C, D, E, F* and *G* define the corresponding sets of occurrences of the symbols. For example purposes, consider the Venn diagram (shown in Figure 4.8) illustrating the probabilistic relationships between these symbol sets. A key measure for this algorithm is $P(X|Y)$ as computed by EQ 4.1. The probability measures for the symbols corresponding to the symbols in the Venn diagram to the left are given in Table 4.3.

Table 4.3. *Probability of a Symbol X occurring Given Symbol Y*

$P(X \mid Y)$		Y						
		A	B	C	D	E	F	G
X	A	1	<T	-1	<T	<T	<T	<T
	B	>T	1	<T	-1	<T	-1	<T
	C	-1	1	1	-1	-1	-1	1
	D	<T	-1	-1	1	-1	-1	-1
	E	1	1	-1	-1	1	-1	-1
	F	1	-1	-1	-1	-1	1	-1
	G	<T	<T	<T	-1	-1	-1	1

Note: T indicates a probability threshold
-1 value indicates mutual exclusion

Given this set of relationships between symbols, we can compute the categorical classification of these symbols as follows:

- **Step 1**: Find root symbols in an ARN

  For each symbol, S_i, in the ARN

  if $(\forall S_{i,j}, P(S_i|S_{i,j}) \le T)$

  then S_i is a root symbol

  where, $S_{i,j}$ is a symbol associated with S_i, and

  T is a variable threshold that can be used to control the levels of classification

  There is one special case in this algorithm:

  if $((P(A|B) > T)$ and $(P(B|A) > T))$

  then both A and B are judged to be independent

  The result of this step in the given example is illustrated in Figure 4.9.

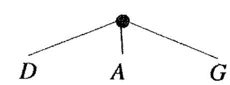


Figure 4.9 Root symbols derived from the symbol sets illustrated in Figure 4.8

- **Step 2**: For each root symbol, S_r, find subclasses

  Step 2.1: Find all symbols that are dependent of symbol S_r

  if $(\forall S_{r,j}, P(S_{r,j}|S_r) > T)$

  then $S_{r,j}$ is a dependent symbol of S_r

  where, $S_{r,j}$ is a symbol associated with S_r, and

  The resulting graph of this step is illustrated in Figure 4.10.

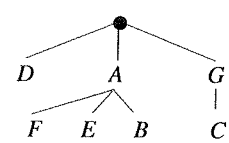


Figure 4.10 Classifying dependent symbols under root symbols

  Step 2.2: Find independent symbols in set of dependent symbols

  if $\left(\forall S_{r,j_y}, P\left(S_{r,j_x} | S_{r,j_y}|_{y \neq x}\right) \le T\right)$

  then S_{r,j_x} is independent of other symbols

  dependent of S_r

  The resulting graph of this step is illustrated in Figure 4.11.

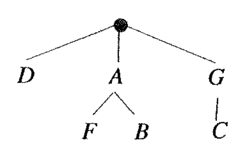


Figure 4.11 Result of removing dependent symbols from subgraph

  Step 2.3: Find subgraph for each independent symbol

  For each independent symbol under the root symbols, Steps 2.1, Step 2.2, and Step 2.3 are repeated until there are no more dependent symbols in the subgraphs.

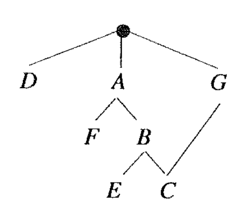


Figure 4.12 Resulting classification of symbols

The result of carrying this process forward in the example provided (in this case it is carried forward one additional level of recursion) is illustrated in Figure 4.12. This example illustrates several important aspects of the algorithm. One, the use of "fuzzy" classification where a variable thresholding level, as defined by *T*, is used to control the breadth of the graph. Depending upon the value of this parameter, symbol *B*, could be classified as either a root symbol, or a subcategory of symbol *A*. In the example, *T* was set such that *B* was classified under *A*. The second important aspect of the example is the classification of symbol *C*. In this example it is classified under both *B* and *G*, indicating multiple inheritance. Multiple inheritance is an important aspect of the resulting acyclic directed graph.

Bottom-Up Categorical Classification

The bottom-up approach to categorical classification starts from a single information object, and incrementally generates an acyclic directed graph that represents *tiers* of information objects. With this algorithm, each *n*th tier represents an abstraction of the set of information objects that are *n* levels of association away. In addition, the bottom-up categorical classification approach represents a radial classification where the original information object is the central theme.

The basic structure of the algorithm is as follows:

- Step 1: Find all information objects directly associated with the initial information object
- Step 2: Build an ARN for the collection of associated information objects
- Step 3: Extract the root groups and subgroups from the ARN (as described in the top-down approach)

These three steps derive a radial structure that is one step removed from a specific information object. We can continue from the information object to derive an abstracted, radial structure where the original information object is the central foci of the structure. To perform this operation the following steps are performed for N steps back from the information object:

- Step N.1: Find all information objects directly associated with the information objects in the N - 1 step back. If the number of information objects in N is the same as in N - 1, we are at the top most level.
- Step N.2: Build an ARN for the collection of associated information objects
- Step N.3:Extract the root groups and subgroups from the ARN (as described in the top-down approach)

4.7 Discussion and Evaluation

In general, I found the classification algorithms capable of finding a structures that represent the *whole* information-base by analyzing associations from *parts* of the information-base, specifically information objects. The tope-down classification algorithm is, in general, effective in finding multiply inherited categorical classifications, as represented by an acyclic directed graph. This does not, however, imply that the algorithm generates structures that generally classify english language, or other languages for that matter. This algorithm is only effective at classifying the content contained with in a given information-base.

The fact that the top-down classification algorithm does not rely on external references such as WordNet is both useful and problematic. If there is an underlying structure contained in the information-base and is reflected in the information objects, the algorithm will find that structure. Problems arise when the content does not have any cohesive structure to begin with. This is especially an issue if the structure is derived from annotations, as opposed to the actually body text of articles. If an annotator does not have a clear structure in his or her head and/or does not express it in terms of association sets, then the algorithms will not find a clear structure. The "garbage in, garbage out" principle applies. This does not, however, mean that annotations must explicitly specify structural relationships. Rather, it means that annotators must think on multiple scales, both abstract and detailed, when combining symbols to form association sets. In this way, abstract symbols are spread out across many information objects, while more specific symbols are isolated in fewer objects. This requirement illustrates the underlying nature of the algorithm–it learns structure through re-enforcement. In general, the way that we abstract general properties is to distinguish between common and more specific characteristics. The result is that annotators must enumerate many features for each information object and those features must range from specific to very general.

Categorization Control

The initial version of the classification algorithm developed for Galaxy of News used a fully inclusive OR not fully inclusive approach to determining roots to an acyclic directed graph. This approach works if 1) the underlying structure is very clearly delineated into inclusive sets (and the information objects directly reflect this structure), and the number of roots is a *manageable* number. What constitutes a manageable number of roots upon the types of elements we are classifying. For example, if we are classifying keywords, a manageable number is less than 50 to 80 keywords or phrases, depending upon the length. the reason for this is that type is only legible at a certain size, and because screen space is limited we can only fit a limited number of keywords on a screen at any one time. If the number of roots is too large, the resulting layout will be incomprehensible to the user.

If the underlying structure is not clearly delimited by inclusive sets, or the information object's features do no reflect an inclusive set structure, the classification algorithm will derive a large number of root groups and a shallow graph. The Millennium Project information-base exemplifies this situation. Analysis of the information-base using fully inclusive set analysis results in nearly 350 root keywords. This clearly exceeds a manageable number of keywords and presents a significant design problem. As a result, it was necessary to develop a partial set inclusive algorithm, as was presented in this chapter.

To address partial set inclusive classifications, I utilized a thresholding value to control the level of classification. While this thresholding value introduces a level of arbitrariness to the algorithm, its has a useful characteristics in that the breadth and depth of the graph can be controlled. Figure 4.10 illustrates the results of changing the threshold value. If the threshold is decreased to below 0.5, the number of root groups found is a manageable number (on the order of 80). As a result, the depth of the graph is much deeper. This controllable aspect couples well with the visual design algorithms discussed in the next chapter.

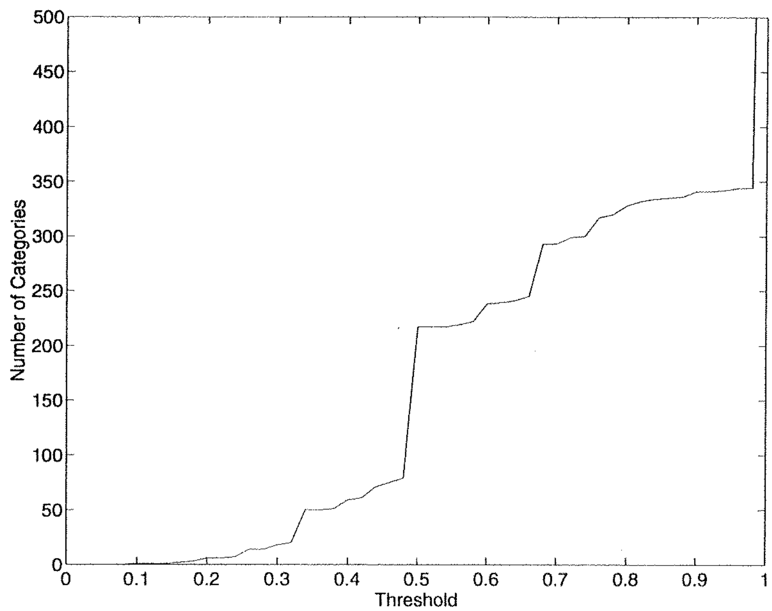


Figure 4.13 *This graph shows the relationship between the classification threshold value and the number of root categories in a top-down categorical classification algorithm described in this chapter.*

Threshold Characteristics

A more careful analysis of the graph in Figure 4.10 illustrates an interesting aspect of the Millennium Project information-base. The graph illustrates clear breaks at 0.5, 0.67, 0.33, and to a lesser degree at 0.25 and 0.75. After considering these characteristics, it became clear that this graph illustrated how symbols or keywords were associated together. For example, the clear break at 0.5 indicates that there were many cases of symbol associations that take on the form:

(...,A,...,B,...)
(...,A,...,C,...)
(...,C,...,B,...)

The other distinguishable break points illustrate similar patterns.

Chapter 5

Visualizing Relationships

In this chapter we discuss the process of building visual spaces that represent relationships between information objects. The main objective of the space building process is to build form that represents the underlying conceptual structures of an information base derived through the process discussed in the previous chapter, and to bring to bear as many visual cues that illuminate the underlying knowledge as possible. In this chapter we describe the process of mapping conceptual structures to virtual spaces, and formulate a theory for this mapping based on the principles of metaphor.

During the course of this thesis research, several questions were addressed:

- What mappings are effective in communicating the underlying structure of the information base?
- What visual cues are effective, and under what conditions are they effective?
- Are there visual cues that hinder understanding of the structure?
- Why do certain cues and mappings work better than others?

In addressing these questions, we hope to illuminate some principles for designing virtual spaces that communicate information structures. However, it is not the intent that these principles be applied as prescriptive measures, but rather as elements of a thought process future designers of these spaces may employ. Having said this, the objective of this part of the thesis is to explore *processes* for constructing virtual spaces that reflect the complex structure of information. The express intent of this thesis is to encode design knowledge into an automatic design system. My approach to accomplishing this goal is to develop mathematical models for space design. Derivation of these mathematical models directly reflect concepts of constructing symbolic landscapes.

5.1 Space Building Overview

The presentation aspect of the visual discourse process consists of projecting the multi-dimensional structures, as derived in the previous chapter, into a three dimensional visualization (as shown in Figure 5.1).

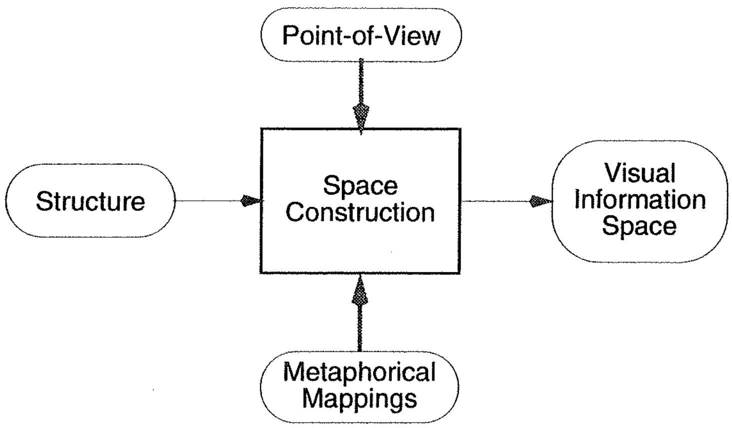


Figure 5.1 *Process of constructing a visual information space.*

Because of the high dimensionality of the underlying structure (a direct correlation to the number of features extracted from the information objects), it is not possible, or at least not meaningfully intelligible, to project the entire underlying space into a 3D representation directly. In other words, it is not possible to create an *object*, entirely visible on the screen, that represents the underlying structure. As discussed in Chapter 2, the structure is too complex. An interesting phenomena occurs, however, when the boundaries of an object exceed the edge of our peripheral vision–the object forms a *space*, and we perceive ourselves as being *contained* within that space [Strausfeld, 95b].

Hence, the process of creating visualizations is two fold: one, we directly create visual objects that represent information (such as images or text) or connections between objects (such as lines and planes), and two, we indirectly create spaces from collections of visual objects. If the collection of objects extend beyond the edge of our vision (in this case the edge of the screen) we get a sense that we reside inside of a space. It is not necessary to be fully enclosed within the space, only that the extent of the space suggests enclosure. And, for all intents and purposes, this occurs when our entire view is filled with an object or collection of objects.

Given that we can not display all information simultaneously, we need to build layers of visual information and allow for intuitive interaction to access additional detail. To this end, it is the role of a 3D space builder to automatically construct information contexts from a list of information objects and a list of extracted features (such as keywords) which are also displayed as graphical objects. An information context is displayed as an enclosure that contains the set of information and feature objects.

At each stage of the projection or design process we need to make decisions on how to layout and present information. Summarily, the decision process involves

- placing
- orienting
- scaling
- coloring, and
- setting the transparency

of visual elements. Our main objective is to use these values to express the underlying structure of the information. So, the main question follows: how do we assign values to these variables?

To address this question we have explored the use of metaphor theory as a basis for mapping or projecting conceptual structures into virtual spaces. The projection strategy for mapping from conceptual structures to virtual spaces is illustrated in Table 3.1 on page 44. These metaphorical mappings form the basis of our development of principles described later in the chapter.

Another task of our 3D space builder is the generation of *transitional spaces*. Transitional spaces allow gradual transitions from one context container to another. A transition between a context that is contained inside another (i.e. the information object list of the new context is a subset of the old information object list) is experienced like a power-of-ten shift or an infinite zoom [Morrison, 94].

Space Construction Process

During the course of this thesis we have defined a model and process for projecting the structural information into a 3D space. The process is dependent upon the type of view, or the conceptual viewpoint, of the information for a given space. Currently, we have parameterized the types of spaces that can be generated according to location, alphabetical position (though the use of this constraint is limited), time (*absolute*, e.g. at time T, and *relative*, e.g. before, after), category, and hierarchy, or as Wurman terms LATCH.[Wurman, 89] These parameters may be specified individually, or by combinations. For example, we can combine geographical and temporal parameters to illustrate the location and time relationships between information elements (which may include combinations of the original information objects, and features extracted from the information objects). Or, a temporal relationship may be combined with a geographical relationship. Specification of these parameters define the *context* in which the information elements are positioned in space. Some particularly meaningful contexts include the following:

- Categorical[1]

- Categorical-Temporal (absolute)
- Categorical-Temporal (relative)[1]
- Categorical-Geographical
- Geographical-Temporal
- Categorical-Geographical-Temporal

We describe the designs of these spaces in the next subsection.

5.2 Visual Space Design

In this subsection we present the designs of several key spaces (as listed above) explored in this thesis. These spaces were designed as part of the Millennium project, a joint project with Lisa Strausfeld. Some of these spaces were designed in part or in whole by Lisa and are noted as such. We describe these spaces here for both completeness and for discussion sake. For further discussion of spaces designed by Lisa, please refer to her Master's thesis [Strausfeld, 95b].

Space Design Specification

To provide some structure in describing the designs of the virtual spaces, we define a set of design specifications for each space. These specifications include the following:

- *Plan view*–showing the top down view of the space with the normal projecting to the left of the page
- *Elevation view*–showing the front perspective, where the normal of the space projects out of the page
- *Foreground elements*–the set of dynamic elements that are the main focus of the space and that change with each instance of the space. For each type of element, the setting or mapping from a conceptual measure or feature to one of the following values is described:
    - position
    - scale
    - orientation
    - color
    - transparency
    - type face

1. Categorical spaces correspond directly to conceptual spaces and conceptual structures.
1. This in effect shows causal relationships.

- *Dynamic contextual (background) elements*–the set of elements that define the context of the foreground elements (e.g. the name of the category) and that change with each instance of the space. For each type of element, the setting or mapping from a conceptual measure or feature as described for the foreground elements
- *Fixed contextual elements*–the set of elements that remain the same for each instance of the space. A brief description of the element is provided
- *Computational derivations*–if layout values for elements are derived using a computational process, the process is described
- *Space dynamics*–a brief description of how the space reacts to a user's movement through the space. These dynamics differ from the dynamics of the visual discourse in that they do not effect a context shift and rather are cues to illuminate various aspects of the existing context

In addition, before we describe the construction of each space, we need to first provide some general context on how a user will arrive in one of the space. How a user arrives in a space is dependent upon the context of the originating space. The specific process of moving from one space to the next is describe in the next chapter; however, the critical point is how the initial context is established.

Following each space design, we discuss the rationale behind key choices in the design process, and where appropriate describe general principles.

5.3 Top-Level Categorical Space Design

A top-level categorical space visually represents the root categories of a collection of information objects and the relationship between these categories. This space is intended to be entered into as an initial context or a gateway into categorical subspaces, mainly because it does not require a user to specify an initial context–the context is created automatically. As with the initial view of Galaxy of News (described in Chapter 2) this space is intended to present an abstract representation of the entire collection of information objects.

This space consists of two classes of foreground objects: one, the root symbols of the acyclic directed graph categorizing the collection of objects, and two, lines connecting the symbols and showing their relationship. Since no context has been established with this space, there are no contextual or fixed background elements.

The image in Figure 5.2 illustrates a top-level categorical space constructed from the Millennium Project information-base. The root categories presented in this space are entry points into information-base. This space provides a high-level road map to the information objects contained in the space. The root symbols in this space are placed in relative proximity to other related symbols.

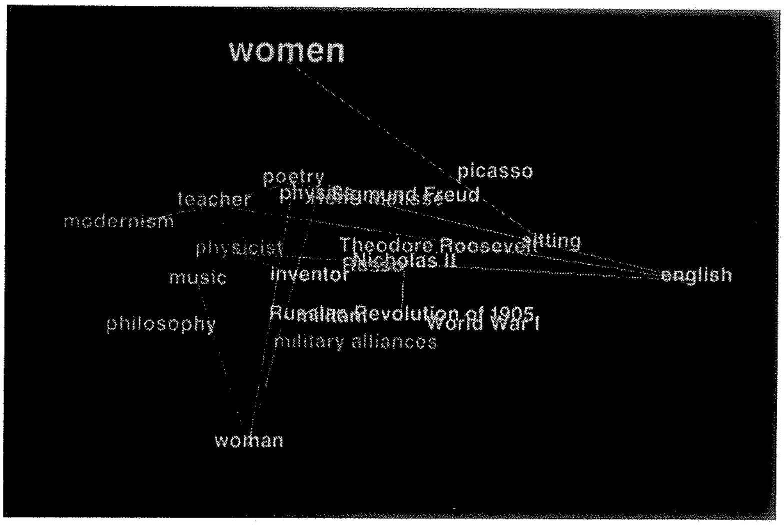


Figure 5.2 An example of a top-level categorical space taken from the Millennium Project.

The specification of if a top-level categorical space is presented in Table 5.1.

Computational Derivations in Space Construction

Construction of this space requires mathematical computation of three variables:

- Symbol position
- Symbol scale, and
- Symbol color.

Symbol Position: The position of the symbols is computed using a multidimensional scaling algorithm that scales the multidimensional relationships between the root symbols into two dimensional *x-y* coordinates. The basis of this algorithm are measures of distance (see EQ 4.7) or similarity (see EQ 4.4) between each root symbol. The distance measure used to compute this space is the euclidean measure given in EQ 4.7:

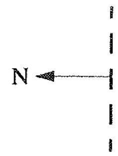
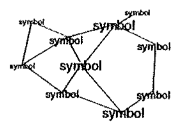
$$\delta_{i,j} = D(S_i, S_j)$$

Given the distance measure $\delta_{i,j}$, we compute the *x-y* location of each symbol by minimizing the following function:

$$F = \sum_i \sum_j [(x_i - x_j)^2 + (y_i - y_j)^2 - \delta_{i,j}^2]^2 \qquad \text{(EQ 5.1)}$$

Or, in other words, we minimize the error between the actual distance between the *x-y* coordinates of two symbols and the conceptual distance as measured by $\delta_{i,j}$. To minimize this function, we explored the

Table 5.1. *Space Specification of a Top-level Categorical Space*

Plan View		
Elevation View		
Foreground Elements		**Mapping/Setting**
Symbols	Position	located on a vertical *x-y* plane *x-y* locations mapped relative to other symbols using MDS (see below)
	Scale	size maps to frequency of occurrence of the symbol in the information object set
	Orientation	fixed horizontal orientation
	Color	maps to "conceptual district"
	Transparency	set to opaque
	Typeface	set to Swiss bold
Lines	Position	vertices connect between two associated symbols
	Scale	set to thickness of 1.0
	Orientation	N/A
	Color	vertex color set to symbol color
	Transparency	dynamic depending upon user position in space
	Typeface	N/A
Contextual Elements		**Mapping/Setting**
None (no context established)		
Fixed Contextual Elements		**Description**
None		

amoeba and Powel methods for minimizing functions and selected the amoeba method as being more efficient for this function[1].

Symbol Scale: The symbol scale is computed based on the frequency of occurence of the symbol in the information object set. The more frequently the symbol occurs, the larger the symbol. The equation for mapping the symbol frequency to a normalized symbol scale is

1. The author would like to thank Vivek Palan, a UROP in the Visible Language Workshop, for investigating the mathematical formulas for multidimensional scaling.

$$s = s_{min} + \left((s_{max} - s_{min}) \times \left(\frac{f_s - f_{min}}{f_{max} - f_{min}} \right) \right) \qquad \textbf{(EQ 5.2)}$$

where, s_{min} is the minimum scale and s_{max} is the maximum scale,

f_s is the frequency of the symbol, and

f_{min} and f_{min} are the minimum and maximum frequencies of symbols in the set of root symbols

Figure 5.3 shows the effect of changing the s_{min} and s_{max} for a given space.

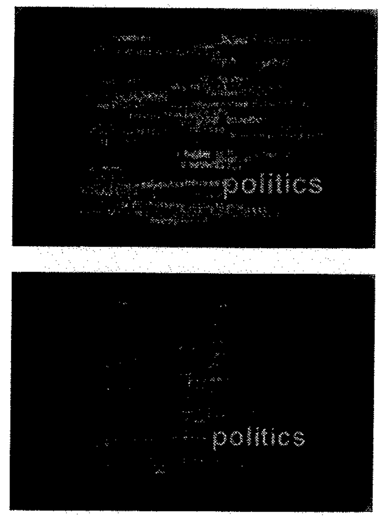


Figure 5.3 *The top space has a min to max scale ratio of 2 to 4, while the bottom space has a min to max ratio of 2 to 8.*

Symbol Color: Our objective in mapping colors to symbols is to create conceptual spaces. In effect, we would like to establish a concept *district*, such that when a user enters the category represented by the symbol and further enters subcategories under that symbol, they will get a sense that they are in a conceptual district. We also intend to use color to provide navigational cues as a user moves through a symbolic landscape.

The objective in assigning color at the root level is to assign a unique color, but that the color be similar to those root symbols that are conceptually close to each other. Further, at the root level we would like to assign a hue value, while keeping saturation and value constant so that they can be used to represent different values in subspaces.

To compute conceptual districts, we scale the distances between root symbols onto the outer edge of a circle (as shown in Figure 5.4), creating a single continuous and circular dimension. The parameter of this dimension is the angle θ, which maps conveniently to hue in the HSV color model. In this model we seek to minimize the function

$$F = \sum_i \sum_j [(\cos\theta_i - \cos\theta_j)^2 + (\sin\theta_i - \sin\theta_j)^2 - \delta_{i,j}^2]^2$$

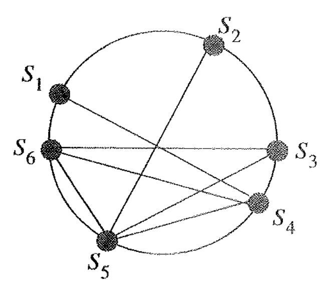


Figure 5.4 *Symbols mapped onto the edge of an HSV color circle where the arc distance between represents the conceptual distance. The symbols color hue maps to the angular placement of the symbol.*

using the amoeba method as described in [reference numerical recipes]. For computational expediency, the above function is reduced to

$$F = \sum_i \sum_j [2(1 - \cos(\theta_i - \theta_j)) - \delta_{i,j}^2]^2 \qquad \textbf{(EQ 5.3)}$$

Note, that the radius of the circle used is 1; hence, we use a distance measure $\delta_{i,j}$ computed as

$$\delta_{i,j} = D(S_i, S_j) = 2(1 - S(S_i, S_j)) \qquad \textbf{(EQ 5.4)}$$

so that symbols that are dissimilar (i.e. similarity of 0) are exactly 2 units away and at the opposite side of the circle. The resulting θ is mapped to the hue of the symbol color.

Space Dynamics

The main variable that changes as the user moves through the space is the transparency of the lines connecting symbols. If the user is far enough away, the lines are opaque. As the user moves closer to the symbols, the lines fade away until they disappear.

Design Discussion

Several issues arose when designing the top-level categorical space. The first was the algorithm used to layout the symbols. My general approach was to use a multidimensional scaling algorithm to place the symbols. This was an intuitive and logical choice. The main issue involved was what distance measure to use. I tried several different approaches. First, I tried using a non-metric method using the similarity measure given in EQ 4.4. This approach tended to place items in a circular form which did not accurately reflect the relative distances between symbols, as shown in Figure 5.5. Second, I tried to use a distance measure of one minus the similarity. This produced different results, but a similar form. Finally, I tried computing a euclidean distance measure of the symbols in a vector space, as given in EQ 4.7, and used a metric scaling technique. The analysis described in Chapter 4 provides for an accurate, computational measure of conceptual distance. This approach formed a space that more closely mapped the conceptual distances to a spatial layout.

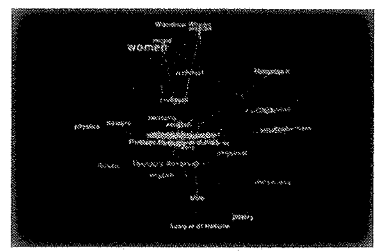


Figure 5.5 *Layout of symbols in a top-level space using a non-metric multidimensional scaling algorithm. Note that the symbols tend to form a circular pattern.*

Another significant issue was the number of root symbols that were derived by the structuring algorithms described in Chapter 4. The threshold value used to derive the roots significantly altered the number of root clusters derived. Figure 5.6 illustrates the results of changing the

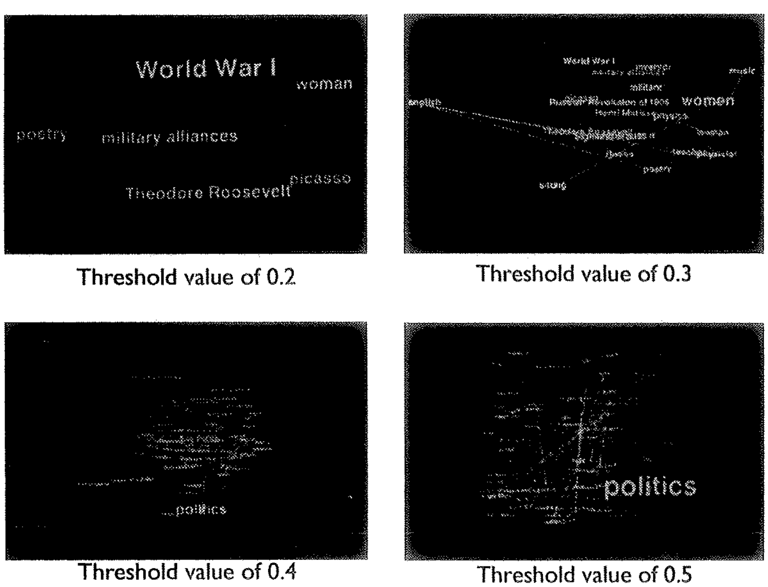


Figure 5.6 *This sequence of images shows the effect of changing the threshold value on the number of root categories*

threshold values. This threshold control is very useful in the design process because the legibility of text limits the number of symbols that can be displayed while still being comprehensible. The number of text symbols that can be displayed in one space depends upon many factors, including the length of the symbols, and the scale at which they are presented. Using the threshold value, the design algorithm can control the number of symbols derived.

5.4 Categorical Space Design

A categorical space represents a category, it's subcategories and the information objects that relate to the category. An important differentiation between a top-level categorical space and a categorical space is that a user may be *outside* or *inside* the space. If the user is outside a categorical space, the user perceives the space as an object. But, as the user moves inside the space, the user perceives the categorical space as a space and that he or she is in that space. This maps to the notion that the user enters into a context. A categorical space is entered into from the top-level categorical space, or recursively through another radial categorical space.

The foreground elements include a set of information objects that contain the symbol that represents the category, a set of subcategories (if they exist), and lines showing the relationships between the subcategories. The subcategories can also be entered into presenting the user a self-similar space.

Figure 5.2 shows an example of a categorical space (in this case the category is *inventor)* constructed from the Millennium Project information-base. This space was entered into from the top-level categorical space shown in Figure 5.2. This space shows a set of information objects that relate to *inventor*, and a set of subcategories that lead objects on related topics. The connected lines indicate the relationships between the related topics and gives form to the information that relates to *inventor*.

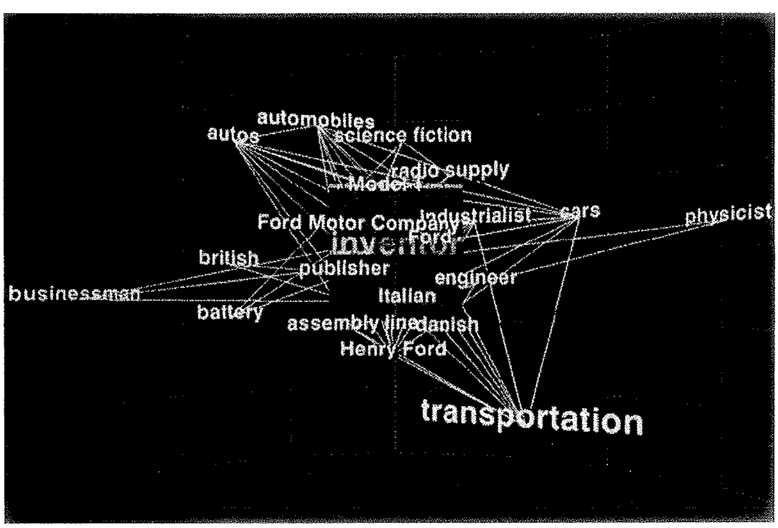


Figure 5.7 An example of inventor *categorical space taken from the Millennium Project.*

The specification of a radial categorical space is presented in Table 5.1.

Computational Derivations in Space Construction

Construction of a categorical space requires mathematical computation for the following variables:

- Subcategory symbol position
- Subcategory symbol scale
- Subcategory symbol orientation
- Subcategory symbol color
- Information object position

Subcategory Symbol Position: The position of the subcategory symbols are computed using a multidimensional scaling algorithm that scales the multidimensional relationships between the root symbols into two dimensional *x-y* coordinates *and* positions the symbols radially from the symbol that represents the category. This algorithm uses two distance measures: one, the inter subcategory symbol distance $\delta_{i,j}$ as computed in EQ 4.7, and two, the distance measure ρ_i between the base category symbol and the subcategory. ρ_i is computed also computed using EQ 4.7:

$$\rho_i = D(S, S_i)$$

Given the distance measures $\delta_{i,j}$ and ρ_i, we compute the *x-y* location of each symbol by minimizing the following function:

$$F = \sum_i \sum_j [(x_i - x_j)^2 + (y_i - y_j)^2 - \delta_{i,j}^{2}]^2 + \sum_i [x_i^2 + y_i^2 + \rho_i]^2 \quad \textbf{(EQ 5.5)}$$

Or, in other words, we minimize 1) the error between the actual distance between the *x-y* coordinates of two subcategory symbols and the conceptual distance as measured by $\delta_{i,j}$, and 2) the error between the actual distance between the base category symbol and the subcategory symbol and the conceptual distance. This function is minimized using the amoeba method.

Subcategory Symbol Scale: The symbol scale is computed based on the frequency of co-occurence of the subcategory symbol and the base category symbol. The more frequently the subcategory symbol co-occurs with the base category symbol, the larger the symbol scale. The equation for mapping the symbol co-occurrence to a normalized symbol scale is

$$s_i = s_{min} + \left((s_{max} - s_{min}) \times \left(\frac{f_{S \cap S_i} - f_{min}}{f_{max} - f_{min}} \right) \right)$$

Table 5.2. *Space Specification of a Categorical Space*

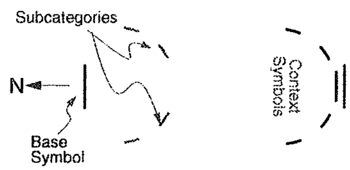
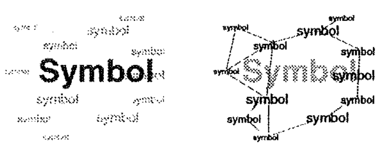
Plan View		 Outside Inside
Elevation View		 Outside Inside
Foreground Elements		**Mapping/Setting**
Subcategory Symbols	Position	mapped as a function of the symbols relationship to the category (base) symbol and the relationship to other symbols in the space
	Scale	mapped to the frequency of co-occurence with the category (base) symbol
	Orientation	normal of the symbol points to the arc center of the space
	Color	mapped as a function of the relationship between the symbols parents in the acyclic directed graph
	Transparency	dynamic depending upon users position in the space
	Typeface	set to Swiss Bold
Lines	Position	vertices connect between two associated symbols
	Scale	set to thickness of 1.0
	Orientation	N/A
	Color	vertex color set to symbol color
	Transparency	dynamic depending upon user position in space; only visible when inside space
	Typeface	N/A

Table 5.2. Space Specification of a Categorical Space

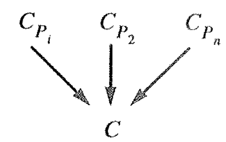
Information Objects	Position	place on an *x-y* grid at the back of the space *x-y* position mapped to relationship between other information objects
	Scale	set to thickness of 1.0
	Orientation	normal set to point in the -*z* direction
	Color	dependent upon object type
	Transparency	dynamic depending upon user position in space; only visible when inside space
	Typeface	dependent upon object type
Contextual Elements		**Mapping/Setting**
Category Symbol	Position	base location set by super-context position along normal changes dynamically as user moves into the space
	Scale	set by super context
	Orientation	normal set by the super context
	Color	set by the super context
	Transparency	dynamic depending upon users position in the space
	Typeface	set to Swiss Bold
Historical Context Symbols	Position	positioned sequentially on the ground plane starting with the oldest context
	Scale	scale increases in step from oldest
	Orientation	normal of the symbol points in the +*y* direction
	Color	set from original context
	Transparency	transparent when outside the space dynamically increases opacity as user enters space
	Typeface	set to Swiss Bold
Fixed Contextual Elements		**Description**
background grid		mostly transparent defines boundary of the space

where, s_{min} is the minimum scale and s_{max} is the maximum scale,

$f_{S \cap S_i}$ is the frequency of co-occurrence, and

f_{min} and f_{min} are the minimum and maximum frequencies of co-occurrence in the set of subcategory symbols

Subcategory Symbol Color: The subcategory symbol color is computed as a weighted average of the colors of the categories under which the



symbol in question lies. This is computed as follows: let weight of the parent category (S_i), w_i, be computed as follows

$$w_i = \frac{P(S|S_i)}{w_{mag}}$$

where,

$$w_{mag} = \sqrt{\sum_i (P(S|S_i))^2}$$

Then, the hue of the subcategory symbol is computed as follows

$$hue = \frac{\sum_i w_i hue_i}{i}$$

And, the saturation is set as follows

$$sat = P(S|S_i)$$

Probability measures are used in computing the color values because the technique for generating the acyclic directed graph uses probabilistic measures.

Information Object Positions: The information objects are also positioned using a multidimensional scaling algorithm. This is achieved by minimizing EQ 5.1, where the distance measure $\delta_{i,j}$ is a measurement of the conceptual distance between two information objects. Since information objects are characterized by a set of associated symbols (this is the basis from which the associations between symbols are derived in the first place), the conceptual distance between two information objects is computed by taking the weighted euclidean distance measure, as given in EQ 4.7, between the set of symbols contained in each information objects:

$$\delta_{i,j} = D(I,J) = \sqrt{\sum_{i=0}^{N} (w_{I,i} - w_{J,i})^2} \qquad N = dimensions$$

An example layout is illustrated in Figure 5.8

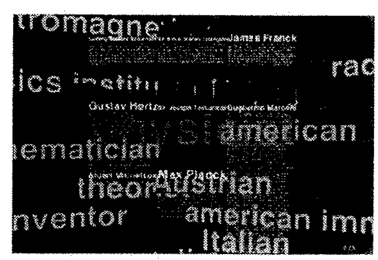


Figure 5.8 *Example layout of information objects.*

Space Dynamics

The categorical space changes dynamically in many ways as user moves into and within the space. The following variables change dynamically:

- *Subcategory Symbol Transparency*–From outside the space, the subcategory symbols are transparent. As the user moves into the space the symbols gradually fade in until the user is fully inside the space
- *Line Transparency*-From outside the space, the subcategory connection lines are transparent. As the user moves into the space, the line fade in until the user is in the center of the space. As the user

moves closer to the subcategory symbols lines fade away (since the pattern of interconnectivity is not visible from a close-up view anyway)

- *Information Object Transparency*– From outside the space, the information objects are transparent. The information objects fade as the user enters the space and moves forward. They are fully opaque when the user reaches the center of the space
- *Category Symbol*–As the user approaches and begins to move into the category space corresponding to the base category symbol, the position of the symbol moves back along it's normal, until the symbol reaches the back of the space, where it stops
- *Category Symbol*–The transparency of the base category symbol fades as it is pushed by the user to the back of the space
- *Historical Context Symbols*–These symbols are transparent when the user is outside the space and opaque when the user enters the space.

5.5 Geographical and Temporal Space Designs

As a user navigates through the symbolic landscapes illustrated in the previous two sections, they may wish to see information contained within the given context displayed using a geographical or temporal organization. At any point in the symbolic spaces, a user can move up to an object and pull up a geographical or temporal space. In this section, I present some examples of geographic and temporal spaces developed during the Millennium Project. These spaces were designed jointly with Lisa Strausfeld and are discussed in her Master's thesis [Strausfeld, 95b]. The images shown in Figure 3.2 illustrate the designs of these spaces.

5.6 Evaluation and Discussion

In this chapter I presented the Mind's Eye approach to design of information spaces. These new Mind's Eye designs represent an advancement over those constructed in Galaxy of News in the following ways:

1. The layout of symbols on each level have a spatial meaning. The conceptual distances between symbols is represented spatially. As a result, as a users moves into the space, they can quickly see related symbols or information objects without having to search around the space to find conceptually related information.
2. The shape of the spaces give form to the content of the information-base. The lines that connect related symbols give form to the underlying structure of the information base. At the top level this form serves as a road map to the space as a whole. Within a subcategory, these lines indicate the underlying structure of the information within the given context. Examples of some of these forms are shown in Figure 5.9. These forms

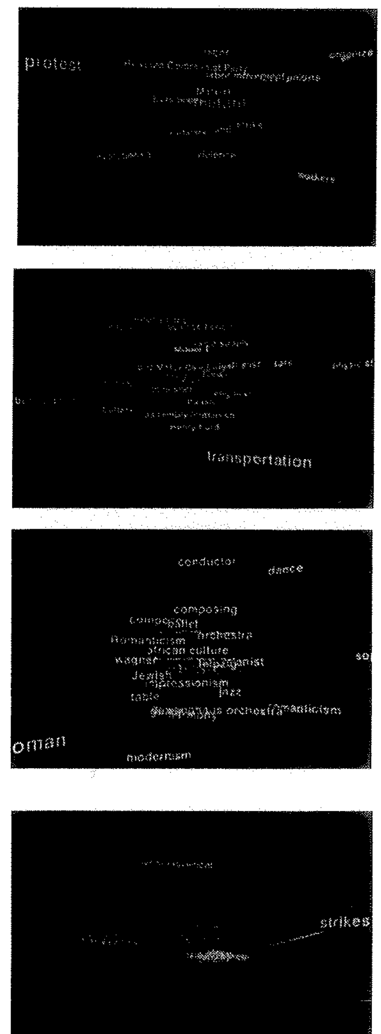


***Figure 5.9** Examples of different forms that represent different segments of an information-base.*

directly reflect the content of the information-base and change as the information objects change. This aspect reflects the organic nature of these spaces.

3. The shape of the spaces give a sense of direction. The 3D curved shape of the subcategory spaces physically gives the user the sense that they are changing the direction when they move into a subcategory. This aspect makes literal the conceptual aspects of changing direction in a conversation as evidenced by metaphor theory. This is not without its problems. The movement to a 3D space from the 2 1/2D space used in Galaxy of News raised new questions of physical navigation. These issues are addressed below.

Navigation Issues

The designs of the spaces described in this chapter raised some new questions of navigation and movement. In Galaxy of News, I constrained the camera movement to look only in the -*z* direction and constrained the movement to be either forward or laterally. I did not allow the user to look from side to side. But, this does not give the user the sense of a change in direction. By building true 3D spaces, as shown in this chapter, we have to give the user the ability to move and look around in this 3D environment.

Initially, we built a 3D navigation camera (as part of the Millennium Project) that had six degrees of freedom. This allowed people to "fly" around the information space much like a plane or bird would fly. This seemed natural at first, but these spaces are abstract information spaces and giving the user six degrees of freedom actually hindered the understanding of the space. People spent most of their time trying to physically move around the space, struggling to keep themselves upright and pointed in the correct direction.

Not only was the physical movement difficult it actually had an interesting psychological effect. It turned these abstract conceptual spaces into physical spaces. And, as a result, the text in the space became objects as opposed to abstractions such as the text on this page represent. Imagine at this point that you paid more attention to the letter forms and shapes than to the words and concepts I am communicating to you at this point. This was the effect it created.

As a result, it quickly became apparent that it was important to restrict the degrees of freedom. But, how do you restrict the range of movement? After further experimentation, it became clear that the degrees of freedom are intimately coupled with the form of the space. For example, the subcategory spaces described above are essentially cylindrical. As a result, we need to give the user the freedom to rotate about the *y*-axis; however, rotation about the *x*- or *z*-axis is not necessary and adds to confusion. Further, the user can move forward and backward, up and down, but not side to side. A sideways movement could result in looking at the text at an angle. However, if we are in a top-level space, we restrict rotation entirely (the space is flat), but allow sideways movement.

Because the spaces change in scale, the movement speed of the camera was also an issue. The scale of the spaces from a category to a subcategory is two orders of magnitude (i.e. about 100 times smaller)[1]. If the camera moved at the same speed, it would quickly blow through the space. So, their needs to be come communication between the space the user resides in and the navigation camera.

Discussion

In this chapter, I present mainly presented the design and construction of categorical spaces, though I presented some views of geographical and temporal spaces. I do not go into detail on geographical nor temporal space construction, nor analysis in Chapter 4. The reason for this is that when we consider these measures and organizations from an absolute perspective, their relationships are reasonably well understood. A geographic relationship is a 2D relationship and a temporal relationship is a linear relationship.

The relative characteristics of categorical classification, however, are very highly dimensional, and hence, are not easily understood nor conveyed. I addressed this issue in this thesis, as it was a central theme. This does not, however, mean that geographical and temporal organizations of information are not useful. Quite the opposite is true. Since these relationships are well understood, they are effective techniques for expressing relationships between information objects.

My approach to utilizing geographical and temporal organizational relationships is to bring them up at a time when the user deems them useful. Hence, I address this issue (as will be described in Chapter 6) by giving users a movement-based method for bringing up these different points of view in a given context. I describe in this thesis a way for a user to shift between different points of view relative to a given object or information space.

Further, I did not specifically address in this document combinatorial aspects of geographical, temporal and categorical analysis and space construction. This detail is beyond the scope of this thesis. However, the algorithms and process involved are combinatorial in nature and leverage off of the building blocks discussed in this chapter and Chapter 4. For example, if a designer of an information space wants to show information objects that relate to a location on a map yet there are too many objects to show in the area around the location, the designer can

1. The reason why these spaces are smaller is because their size is relative to the symbol that represents a subcategory. Further, movement into a subcategory constitutes moving into a smaller search space conceptually.

apply the top-down categorical classification algorithm on the set of information object directly associated with the geographic location, and layout the root symbols around the location. Or, the designer can choose to apply an MDS algorithm on the information objects to place them around the location.

This chapter has illustrated through some select examples how the structural analysis building blocks described in Chapter 4 can be combined with spatial analysis building blocks described in this chapter. I have shown how these building blocks can be combined to solve visual design problems. This approach is not intended to be a prescriptive measure, rather the provide examples of a general approach to dynamic design and non-linear presentation.

Chapter 6

Interpreting Interaction

This chapter addresses how users interactively navigate *between, into* and *out of* subspaces and related or connected spaces as a user builds or changes context. The previous two chapters discussed how structures are derived from an information base and how those information structures that represent relationships between information objects are projected into information spaces that illuminate those underlying structures. The previous chapter also discussed how the constructed spaces reacted dynamically to the users movement within the space, selectively changing visual cues (enhancing some elements while de-emphasizing others) to aid his or her understanding of the space. This discussion was, however, only addressed the systems response within a single space. This chapter addresses how a user can dynamically build new spaces based on movement within the current space. This movement between connected spaces is the basis of visual discourse.

This chapter lays the conceptual foundation for how movements through a space can be interpreted. This is followed by a presentation of the system that implements and realizes those concepts.

6.1 Overview

An important aspect of meaning communication, and hence understanding information-bases, is the dynamic process of shifting point-of-view and shifting context. As Fauconnier delineates, a central theme in meaning construction is access through conceptual connections that define mappings between source and target domains [Fauconnier, 94]. An instantiation of this concept is the continuous process of building a context, exploring, extending the context, exploring, building a new context, and so forth.

In our computational environment we define a context to be a set of information objects and the relationships between those objects. A context is represented and presented to a user as a container, presented as a space, and a set of contained objects, where spatial representation of the container defines the relationship between the objects [Rennison, 95b].

A context shift is defined as either global or local. Establishing a global context implies *filtering* or re-filtering the original information objects into a working subset of information objects. For example, we may wish to establish a global context to be all objects "in the geographic area of 'France' during the period of 1911 to 1912." Local context shifts imply a change in *conceptual viewpoint* on the subset of information objects and *illustrate* a new set of relationships between the context of objects. For example, we can shift between a `categorical` view, to a `categorical-temporal` view, to a `geographical-temporal` view.

Key questions that arise from this process include:

- How does a user indicate a context shifts?
- How are these context shifts executed?
- How does a user's movement in a space translate to changes in context?
- How do we distinguish movement of one's self in a virtual space verses movement of objects in a space?

Table 3.2 on page 46, *Interpretation of User Interaction*, outlines the approach I explored to address these questions. This approach is based on metaphorical mappings between a person's experience moving in the physical world and movement in a conceptual or mental space [Rennison, 95b]. It lists the possible user interactions and their effect on the display of objects and contexts as well as the underlying information representation. An information object will display more detailed information up close than it will from far away, for example, and will foreground and background different information from different points of view [Strausfeld, 95a].

6.2 Visual Discourse Grammar

The metaphorical interpretations of user interactions described in Table 3.2 provide a grounding of interpretation based on experiences in the physical world. While this provides some structure in local contexts, it does not necessarily provide structure the discourse as a whole. We would like to additionally lay a grammatical structure to the discourse. To this end, we have defined a formal grammar for interpreting user interactions. The model of a structured visual discourse can be formally specified by the grammar (in BNF format) shown in Table 6.1.

The notation for the discourse grammar is glossed as follows:

- A *situation* consists of a *user* positioned in a *context* consisting of objects positioned in a 3+ dimensional space, where the placement of objects are based on relationships and structural meaning
- A *context* consists of a 3+ dimensional *space*, a set of *placed-objects*, and the *perspective* from which they are viewed

Table 6.1. Visual Discourse Grammar

<discourse>	::=	<situation> <transition>*
<situation>	::=	<user> <context>
<transition>	::=	<situation> <context-shift> <situation>
<context-shift>	::=	<user> <movement-action> <context>
<user>	::=	location direction-vector
<movement-action>	::=	investigate | enter |
		look-through | push-through |
		look-back | enter-back |
		look-forward | enter-forward |
		look-down | enter-down |
		look-up | enter-up
<context>	::=	space <placed-object>+ <perspective>
<placed-object>	::=	<info-object> location <attributes>*
<perspective>	::=	geographical |
		categorical |
		temporal |
		categorical-temporal |
		categorical-geographical |
		categorical-geographical-temporal
<info-object>	::=	<symbol> | image | video | Article | map
<symbol>	::=	keyword | symbol-image

- A *transition* consists of a change from one *situation* to another that is executed by a *context-shift*
- A *context-shift* consists of a *user* moving between two *space* that defines *contexts*, where the *movement* specifies the change of *context* or *perspective*
- A *user* consists of a user's *location* in *space* and a *direction-vector.* of where the user is gazing
- A *movement-action* are the core of the discourse and consists of queries to the underlying structure of information, such as
  - *investigate*–present more detail without changing context
  - *enter*–create a new context specifying more details on selected information
  - *backup-abstract*–create an more abstracted context relative to the current context
  - *look-through*–present a a more general view, with reference to a central theme (i.e. an abstraction based on a specific object), without changing context

- *enter-through*–create a new context specifying a more general view relative to a specific object
- *look-back*–present a view looking back in time without changing context
- *enter-back*–create a new context specifying a view back in time (e.g. show the events that led up to the current event)
- *look-forward*–present a view looking forward in time without changing context
- *enter-forward*–create a new context specifying a view forward in time (e.g. show the events that were caused by the current events)
- *look-down*–present a geographic view without changing context
- *enter-down*–create a new context specifying a geographic perspective (e.g. show the events that led up to the current event)
- *look-up*–present a categorical view without changing context
- *enter-up*–create a new context specifying a categorical perspective (e.g. show the events that led up to the current event)

- An *placed-object* consists of an *info-object* positioned in the information space
- An *info-object* consists of either a *symbol, image, video, Article,* or a *map*
- A *perspective* consist of one of the four ways of looking at information, namely geographical, categorical, temporal, and hierarchical

This grammar defines explicitly and formally the boundary of the visual discourse, and provides a formal method for specifying the range of visual discourses. The key element of this grammar is *transitions* that consist primarily of *context-shifts.* Context-shift requests take on the similar form of the subject-verb-object relationship in the English language, and ground the discourse in a similar linguistic structure. Refinements of the above formal grammar define how the system responds to the user as they move through a space relative to the objects that define a context. This grammar forms an implicit rule structure for the visual discourse.

6.3 Context-Shifts and Spatial Connections

In this thesis, we have defined a context-shift to be the process of connecting one space to another, where each space illustrates a different set of relationships between information objects. The key question is: how does a user execute a transition from one space to another? In this thesis, I explored movement relative to objects in a space as a means of navigating between these spaces. In this way, movement is used to automatically connect spaces.

In this thesis, I investigated the following context-shifts:

1. categorical -> sub-categorical
2. categorical -> geographical
3. categorical -> temporal-categorical
4. geographical -> categorical
5. geographical <-> temporal

where A -> B indicates unidirectional context-shift from context A to context B, and A <-> B indicates a bidirectional context-shift between A and B, such that a user can freely move back and forth between space A and B.

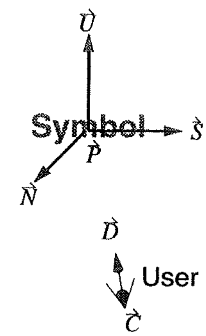


***Figure 6.1** Definition of an object reference frame and a user's reference frame.*

The key question I addressed is: how is a user's movement mapped to one of the above context-shifts? My approach is based on the relationship between the general reference frame of an object (or space) and a model of the user. As illustrated in Figure 6.1, an object is modeled by a *position,* $\vec{P}$, a *normal vector,* $\vec{N}$, an *up vector,* $\vec{U}$, a *side vector,* $\vec{S}$, which is the cross product of the up vector and the normal vector. The user is modeled with a *position,* $\vec{C}$, and a *direction of gaze,* $\vec{D}$. Given these models of objects (or spaces) and a user, the conditions for executing a transition from one context to another for each of the transitions listed above are describe here.

- *categorical -> sub-categorical*–If the user is within an activation area, is looking at the front-face of the object (i.e. $\vec{D} \bullet \vec{N} \approx -1$), and moves toward the object, we begin to execute a context transition to a sub-categorical space.
- *categorical -> geographical*–If the user is within an activation area, is looks *down* at the object (i.e. $\vec{D} \bullet \vec{U} \approx -1$), and moves toward the object, we begin to execute the context transition to a geographical space.
- *categorical -> temporal-categorical*–If the user is within an activation area, is looks *back* at the object (i.e. $\vec{D} \bullet \vec{S} \approx -1$), and moves toward the object, we begin to execute the context transition to a categorical-temporal space that looks back in time.

  If the user is within an activation area, is looks *forward* at the object (i.e. $\vec{D} \bullet \vec{S} \approx 1$), and moves toward the object, we begin to execute the context transition to a categorical-temporal space that looks forward in time.
- *geographical -> categorical*–If the user is within an activation area, is looking at the front-face of a location symbol (i.e. $\vec{D} \bullet \vec{N} \approx -1$), and moves toward the location symbol, we begin to execute the context transition to a categorical space. An example of a location symbol is the name of a country, state, or city.

- *geographical <-> temporal*–If the user rotates the space on it's side(i.e. $\vec{D} \bullet \vec{N} \approx -1$), we begin to execute the context transition between the two types of spaces.

6.4 Example Context-Shifts

The example given in Section 3.3 illustrate three different types of context shifts that are delineated by a *movement-action*: 1) categorical to subcategorical spaces, 2) categorical to geographical, and 3) geographical to temporal-geographical. I present here a simplified example for each of the five context-shifts delineated above to illustrate clearly the types of actions that trigger a transition. These examples are given in Figure 6.2.

Figure 6.2 A set of simplified examples showing how a users movement relative to an object will cause a transition from one type of space to another.

1. Categorical to sub-categorical
2. Categorical to geographical
3. Categorical to temporal categorical
4. Geographical to categorical
5. Geographical to temporal

6.5 Computational Approach

The visual discourse process is directly supported by the architecture illustrated in Figure 6.3. In this architecture, the *visual discourse*

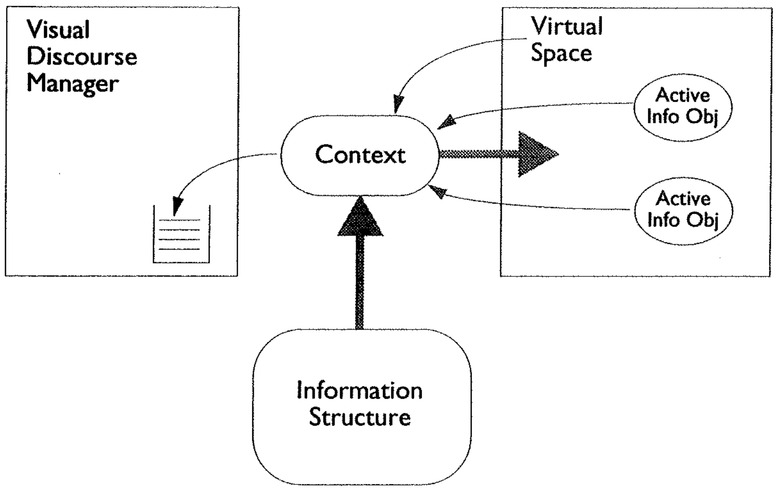


Figure 6.3 *Visual discourse architecture. The wide gray lines indicate flow of relationship information as a virtual space is constructed to illustrate a conceptual structure. The thin lines indicate context-shift activation feedback. The visual discourse manager maintains a context stack.*

manager creates and manages *contexts.* When a context is created, the context queries the *information structure* for the appropriate relationships between information objects. The context uses this information to create a *virtual space* that contains *active graphical objects.*

As the user moves through this virtual space, each of the active graphical objects, as well as the virtual space, independently *sense* the position and orientation of the user relative to it's internal reference frame, as illustrated in, and take an appropriate action. Each of the

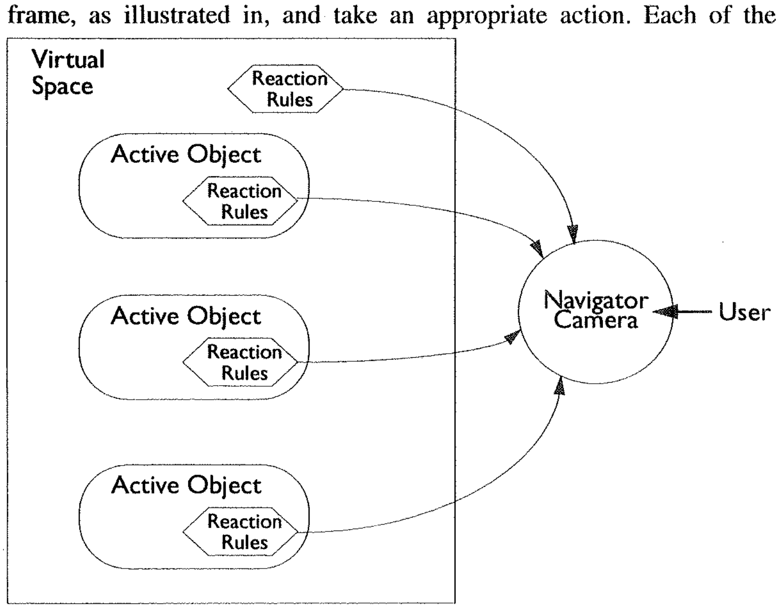


Figure 6.4 *Active graphical objects and virtual spaces act as autonomous agents, sensing the user 's movements and comparing the sensed values against a set of* reaction rules. *If a context-shift condition is met, the active object notifies the visual dossiers manager via the context, which in turn executes a context-shift.*

active graphical objects compare the sensed parameters against a set of *reaction rules* that are specified as part of the discourse grammar. When a reaction rule files, changes its internal state, and as a result it may either 1) modify it's visual parameters, 2) modify the visual parameters of the local virtual space, or 3) activate a context-shift.

When a context-shift is activated, the active graphical object notifies the current context and in turn the visual discourse manager. The context-shift is executed by having the current context create a sub-context, pushing the old context on a stack (or popping it off if the user is backing out of a context), creating a virtual space from the new context, and activating the new virtual space and its graphical objects. When the new virtual space is activated, it begins sensing the user's movements in the virtual space.

6.6 Evaluation and Discussion

In developing the visual discourse grammar that defines how the system responds to user movements and the architecture described above, a number of issues were addressed. These issues are briefly discussed below:

1. *Global or local reference frames*–In arriving at the above architecture and approach I tried a number of different ways to measure and respond to user movements. One of the main issues I faced was whether or not to use a global or local reference frame. I tried using a global reference frame for computational expediency. Using a local reference frame requires that the camera's position and orientation (the parameters that model the user) be transformed into local coordinates space which requires inversion of the modeling matrix. Using a global reference frame does not require this inversion and transformation of the camera parameters for each active graphical object. However, because a context-shift can result in a change of scale, orientation, and position of the space the user resides in, it is difficult to maintain the orientation, scale and position of the active object in world coordinates. And, though it requires additional computation, it is much easier to define a set of reaction rules based on local coordinates. This simplicity is well worth the computational costs.
2. *Position and orientation measurements*–In developing the system, I set up a general scheme to compute *sensing* measurements, including the following measurements: *camera distance, distance from normal, distance to normal intersect point, angle from normal, angle from up, angle from side,* and *camera position* in local coordinates. Initially, I used the orientation measurements to determine if the user was looking at an active object. However, since many of the objects are not square, the measurements were insufficient for determining if the user was actually looking at an object. The reason for this is that the active object is modeled as a point, and the measurements only computed radial measurements. And, in many cases, as with

text, the active objects are long and rectangular. This problem is resolved, however, by using the camera position to determine if the user is in an active area, and then use orientation measurements.

3. *Transitions*-Perhaps one of the most significant issues addressed in this thesis is the issue of transitions. How transitions between spaces are executed is critical. If the transitions are very abrupt, the result is significant disorientation when entering a new space. And, as a result, users get a sense that the space is disorganized and quickly get lost. This interesting enough is often the experience felt with hypermedia environments. And, in fact, this was also the result of a study conducted by Kevin Lynch to investigate how people construct mental images of cities.In his analysis of people's mental constructions of cities, Lynch identified five major properties of cities: paths, edges, nodes, districts, and landmarks. After interviewing a number of people about their understanding of a city, Lynch discovered, among many other things, that if the edges are very abrupt between districts people had a general sense that the city was disorganized [Lynch, 60]. This same characteristic also appears to apply to virtual spaces. In evaluating the new spaces I constructed, I realized that where the transitions were very abrupt, it was very easy to get lost in the web of virtual spaces. Even the sense of movement did not help disambiguate what was going on.

To provide a specific example of the issues faced in developing transitions. In one rendition of the categorical spaces, when a user executed a transition to a sub-categorical space by "pushing through" a symbol that represented the sub-category. As the user pushes into the symbol, the symbol gets very large until it fills the screen with a solid color. When the user pops out the other side, he enters the subspace and the transition is completed, as shown in Figure 6.5. At this point, the

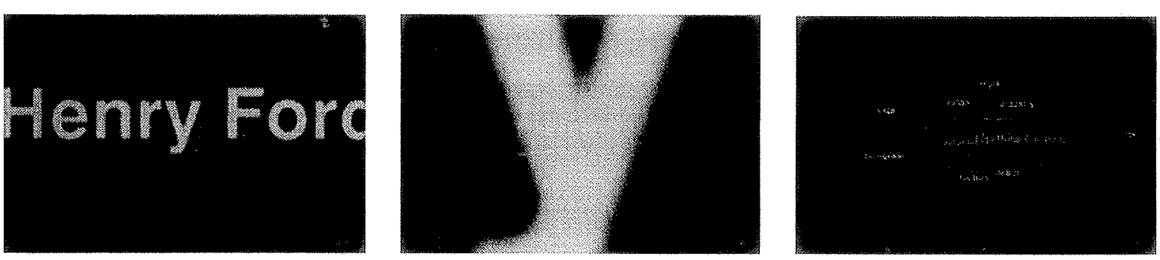


Figure 6.5 *Transition into a categorical subspace where the user "pushes through" a symbol to arrive in the sub-categorical space. This results in disorientation and a general sense that the space is disorganized.*

user had to figure out where they were.

This problem can be alleviated, though, through the use of *smooth transitions*. This greatly helped maintain orientation and a general sense of organization[1]. However, I also found that animations between spaces must be directly linked to a users movements. If the objects move autonomously, it is difficult to distinguish between user movement and object movement. This adds to confusion. But, if contextual elements respond directly to a users movements, the transitions were fluid and effective in helping to maintain a sense of orientation and position in the global structure.

1. Interestingly enough, this was effectively demonstrated by Galaxy of News which had this characteristic, though I did not know it's significance at the time. In hindsight, I realize that this was one of the most significant aspects of Galaxy of News.

Chapter 7

Applications and Extensions

This chapter addresses the application of the processes described in this thesis to storytelling. Two projects, Interactive Boston and InteractiveMTV, explored how, if at all, the techniques describe in this thesis can be applied and extended to tell stories. The Interactive Boston project explored non-linear visual presentation issues associated with interactively presenting the story about the history and development of the multibillion dollar Boston Artery project using video clips, images, sound bites, animations, and text articles. The InteractiveMTV project explored an approach to constructing visual imagery in conjunction with a musical performance. While Interactive Boston explored issues of sequencing temporal media, InteractiveMTV explored issues of temporal contexts and contextual history.

7.1 Interactive Boston

Stories has traditionally been expressed with linear media where the author has control over how content is presented to the reader or viewer. The non-linear, interactive aspects of the computer as a storytelling medium presents a problem because the reader has direct input on the direction of the story. How does this non-linear aspect change the way a story is authored and presented to a reader, while still maintaining the intent of the story teller?

Two of the significant differences between the issues raised in the main portion of this thesis and interactive storytelling are the temporal nature of the output medium, and issues of ordering and sequence. Up to this point, most of the discussion in this thesis addressed presentation of static media such as text and images. Video and sound, on the other hand, have a temporal presentation component. A key aspect of storytelling is the sequence in which video and audio clips, images, and text are presented to the reader. Hence, a key issue of interactive storytelling is how is a sequence is automatically generated. Problems also arise when several video clips are presented simultaneously, and especially when several audio samples are presented simultaneously. While the cocktail party effect (playing many audio clips simultaneously) may be useful in some situations, this is not generally considered an acceptable way of presenting a set of audio clips.

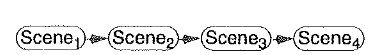
Interactive Storytelling Model

To preface the issues of sequencing addressed in the Interactive Boston project, let's consider a model for interactive storytelling. Storytelling, in general, involves the development of *characters* along a *theme* in a setting or *place.* The process of telling a story involves *time, plot,* and *point of view.* A primary aspect of storytelling concerns the fact that character, place, theme, time, plot and point of view each represent different relationships to idea of moving a story forward, or flashing back or reviewing. With chronology a story forward chains through events. With character and setting, a story can chain forward, backward, laterally, or in some combination. The plot of the story is how particular events are chained in a sequence. Events disclose situations and changes to situations. Each of these elements are placed in question when we consider the interactive nature of the computer as a medium for expressing stories.

In developing a model for interactive storytelling, we face an interesting challenge of providing a common model for both local interaction and global interaction. By local interaction I refer to the presentation of elements of the story within a specific context. And, by global interaction I refer to a larger sequence of contexts, where the story structure in each context build upon each other to form the whole story. In the local context, each interaction must advance communication of ideas to be expressed within that context. At some point, there must be a transition between one local context and another. How is this transition established? And, how is the order of contexts established and advanced?

In a book, a chapter or section establishes a certain context to explain or present a sequence of events that develop characters and/or themes. The linear sequence of chapters establishes the global context and connections between the chapters. Similarly, shots and shot sequences form a scene, and sequences of scenes form the story as expressed through video or film. However, computers offer us the ability to break from the linear structure of books and film. They afford us interaction. And, with interaction comes the dynamic multiplicity of sequences and orderings. But, there must be some underlying structure if we are to communicate the story effectively; otherwise, the interaction will simply lead to chaos.

In traditional dramatic models, local interaction is synonymous with a *scene* in a movie or play. In the global structure of a narrative, each scene advances the narrative, one after another in a sequential fashion, until the whole story is told. The sequential structure in many cases is needed to ensure that certain ideas are established before others are presented.



In many cases the structure of the narrative is not strictly sequential. In more complex narratives, the *plot*, or sequence of events or ideas, are multidimensional (see Figure 7.1). However, the presentation medium,

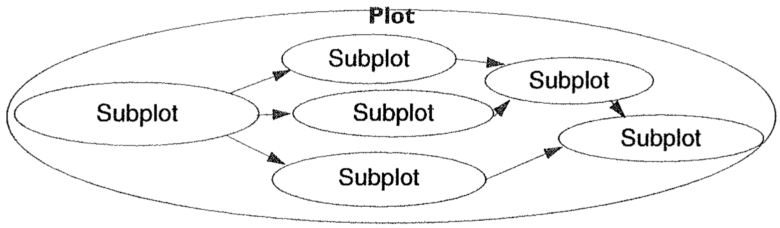


Figure 7.1 Plot structure of a drama– subplots feeding into subsequent subplots, forming a narrative.

whether it is film or a play, still remains sequential. As a result, the director must weave a sequence of scenes through the plot structure, revealing portions of the plot as the narrative advances (see Figure 7.2).

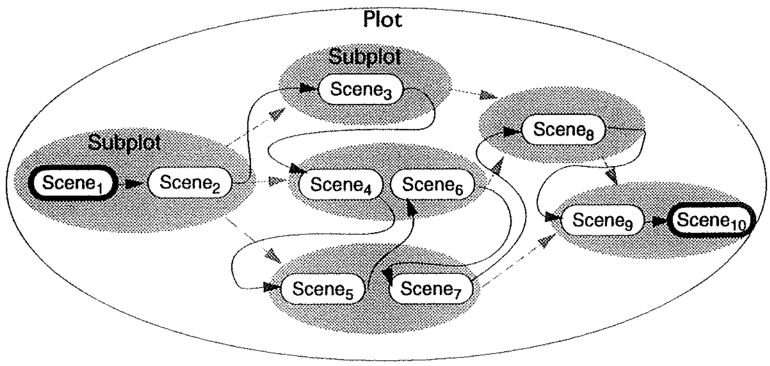


Figure 7.2 Plot structure of a drama, and it's relationship to a sequence of scenes that weave through the plot structure revealing the story.

With film and theater the sequence of scenes are carefully designed ahead of time, and in their final presentation the sequences are fixed. In the computer environment we are not restricted to these limitations. And, in fact, any number of sequences may be followed by a free-willed interactor. As a result, we need to extend these models beyond their current definitions and allow for interactor initiated transitions from one scene or context to another.

A key issue in defining an interactive storytelling model is establishing transitions between one element of a plot to another. To create an engaging environment, transitions should leave the user with an undiminished feeling of free will. The whole purpose of the environment is to give the *active* user an *experience* that conveys meaning and understanding, providing the user with a sense of immersion and presence. Giving the user the ability to effect the outcome of the story is also desirable.

A limitations of the movement between plot elements shown in Figure 7.2 is that they indicate singular transitions, almost like hyperlinks in a hypermedia model. These types of transitions disrupt the flow of events and can hinder the clear development of ideas and concepts. Alternatively, we would like to develop a model that allows for fluid transitions between different stages of the plot. To achieve this goal we must rethink the problem from a different perspective.

In defining an interactive drama model we need to consider the interplay between the space the user interacts within and the underlying structure of the message. In essence, the spaces described in this thesis are an intermediary between the structures of the story and the interactor exploring those ideas. The spaces and the interactions within the spaces are the avenues for communicating ideas between the computer and human interactors. And hence, a model for interactive drama should reflect this relationship.

The model illustrated in Figure 7.3 represents an abstract relationship between spaces, elements or *actors* contained within those spaces, and the transitions between those spaces and actors. Each circle indicates a

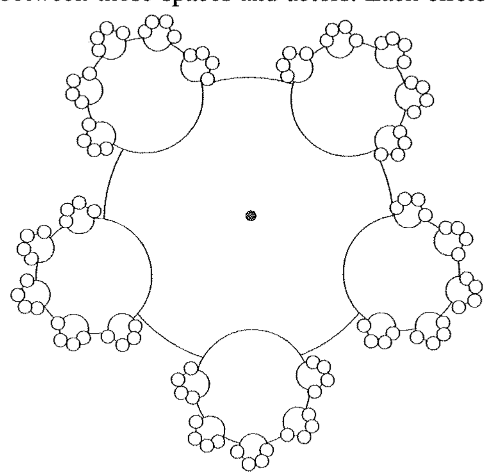


***Figure 7.3** Model for interactive story. Each circle represents a subplot. The edge between adjacent circles indicates a transition from one subplot to another, until one of the outer edges of the circles is reached, thus ending the performance.*

subplot, and is represented by a story structure such as associative relationships between video clips. Each subplot contains a space that sets the stage for interaction, and a set of actors to interact with in that space. Transitions are made between the each subplot on the edge of one subplot and another. This transition is made by moving from one subplot space to another subplot space. Actors common to both spaces form the basis for transitions from one subplot to another. In this model, the user has the freedom to explore each space at his or her own free will. When the user reaches a point on the edge of the space, they can push forward and enter the next space, and hence, the next subplot.

Interactive Boston Approach

To explore the issues of interactive storytelling outlined above, we applied the space building techniques described in this thesis to tell stories. In addition, we extended the space construction process by adding the ability to dynamically form sequences of video clips, images, and sound clips once a user entered into a space. In this model, the space represents the structure of the story. When a story is told, an interactor is presented with a set of objects, organized in a space, that provide entry ways into the story. As the interactor moves up to the objects, a space containing additional objects folds before them. If they enter the space, a sequence of images, video clips, and/or sound clips are presented. If the interactor was intrigued by any of the images or video clips, he or she can move up the object and a new space will unfold before him or her. The interactor can press forward or back up and take an alternative path through the story space. The interactor can wonder through the space until all the material is exhausted. Example screen captures of this experience are shown in Figure 7.4.

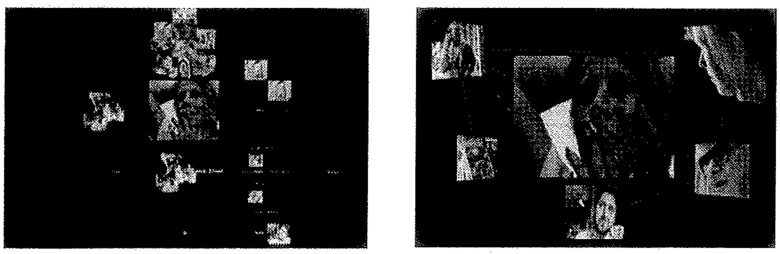


Figure 7.4 Screen captures taken from Interactive Boston showing a story space

Generating Sequences

An important aspect of the Interactive Boston approach was the process of selecting sequences. The approach I explored for generating sequences was the point–counter-point model. When the user moves into a space that represents a sub-plot, the system generates a sequence of counter-punctual clips.

To generate a sequence of counter-punctual media objects, I developed an algorithm that is based on pro and con points of view. The basic process involves the following:

1. The storyteller annotates a set of media objects with annotations that include
   - *Subject Sequence*–A sequential set of associations that describe the topics the media object
   - *Pro Subjects*–A set of subjects that the content of the media object supports
   - *Con Subjects*–A set of subjects that the content of the media object does not support

   Each of these media objects are stored in a database.
2. The database is parsed and the annotations extracted. The *subject sequence* association sets are used to construct a temporal ARN that is used to establish the global story structure. An initial context of the story is established by analyzing the TARN and extracting the root subjects.
3. The initial context is presented to the user as a space with the media objects positioned in the space. As the user moves into this initial context, a sequence of video clips, images and audio clips is generated. These sequences are generated by creating a multidimensional space from the *pro* and *con* annotations. To do this we create a weighted vector space, where the symbols used by the annotator

form the basis of this conceptual space. The media objects are represented by a vector composed of positively weighted *pro* symbols, and negatively weighted *con* symbols, i.e.

$$\vec{V}_o = \sum_i \vec{P}_i + \sum_j -\vec{C}_j$$

4. From the vector space constructed in Step 3, we can generate a sequence of media objects by selecting one of the objects as the start of the sequence, and then for each subsequent object, finding the object that is the furthest away in the conceptual space. For example, if we collapse the multidimensional space down into a single circular and continuous dimension (as shown in Figure 7.5), we can build a sequence by selecting objects that are opposite one another. Objects directly opposite one another take a pro and con stand point. The resulting sequence is a point–counter-point sequence.

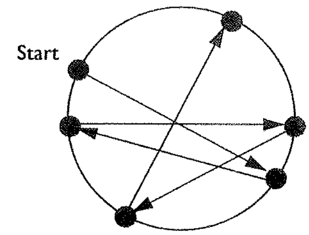


***Figure 7.5** Media objects placed on a circle. The arc distance from one object to another represents the conceptual distance between the objects. Objects on the opposite side of the circle* oppose *each other conceptually.*

Evaluation, Discussion, and Future Directions

In general, Interactive Boston was rather effective at showing relationships and connections between characters, places, and events. An interactor can wonder through a space and be presented with sequences of connected and interrelated images, audio and video clips. This was not surprising as it was the objective of this thesis in general.

However, in evaluating the end result, this prototype highlighted many deficiencies in this approach as a medium for telling stories, at least in it's current form. I present these deficiencies here as directions for future research. Some of the deficiencies include the following:

Perhaps the most limiting aspect of my Interactive Boston approach is that the structure of the story is static. The structure remains constant throughout the duration of the presentation. And, it does not change until new material is added. The problem with this is that it does not show how the plot structure unfolds over time. It does not show the effect certain events have on the development of character and theme. For this medium to more effectively convey a story, the system would need to be modified to handle changing structures as a plot unfolds.

Because interaction is a primary driving factor for how the story unfolds, the system would need to be extended such that a user's interactions would effect the state of the story. This would fundamentally change the way the story is authored. The author would have to define the initial states of the story and define how a user's movement within the space effected how connections are established and how the state of the story changes. The system would have to maintain a model of the story that captures the history of a users movement through the story.

In extending the system, another element must also be added. The current structure does not have explicit notions of characters. The concepts of subjects can be substituted for themes, and the system has knowledge of location and time, but the system does not explicitly allow authors to specify relationships between characters and theme, place and time. The main issue with adding notions of characters is that characters undertake *actions*. And, actions effect the state of the story.

At first glance, this may seem like adding support for characters and dynamic models would require a major rework of the system; however, this is not necessarily the case. Because of the way the system was architected, this could potentially be accomplished in the following minor extensions. One, we could extend the information representation to support the concept of a character as a fundamental element, similarly to the way the system represents information objects, symbols, locations, and dates as fundamental representation elements (see Figure 4.3 on page 55).

Two, we could extend the model of an information object to support the notion of a *character action.* Rather than specifically defining a character, a story author would define a set of information objects that describe actions that a character undertakes. An action description should be representative of the action a character takes in a video clip, for example. In addition, this information object would describe the effect the action has on the state of the story model, and the effect the presentation of the video or audio clip has on the current context.

Third, we could extend the model of an *active object* to support a model of a *character–character-action.* A character–character-action would be represented by 1) the physical representation of the character, such as a still image of the character, 2) the rules that would cause the action to take place (e.g. the user moves close to the physical representation), 3) the presentation to the user when the action is undertaken (e.g. presenting a video clip), and 4) the effects to the state of the story model. The effects to the state could be as simple as adding the properties of the associated information object to the underlying representation. Or, they could be more sophisticated by allowing the author to specify different actions for each information object.

As for the underlying representation, the approach taken in the prototype Interactive Boston system suggests that the representation used in this thesis would be sufficient for representing the current state of the story (with the addition of the character representation described above). To initiate the state, the author would specify which information objects initiate the story. These objects would be parsed, inserted into the information structure representation, and the resulting initial structure derived (as described in Chapter 4). Then, each time a character-action is executed, the structure would be updated.

The result of these extensions would potentially be an environment that supports an unfolding plot and story structure that dynamically adapts to a user as he or she drives the story forward.

7.2 InteractiveMTV

The InteractiveMTV project was an exploration in constructing dynamic imagery in response to musical performances. To this end, as a musician performs music, the system responds by producing imagery that reflects the characteristics of the music being played, as illustrated in Figure 7.2. This project was an experiment in the interplay between music and imagery, where the presentation of one effects the other, and visa versa.

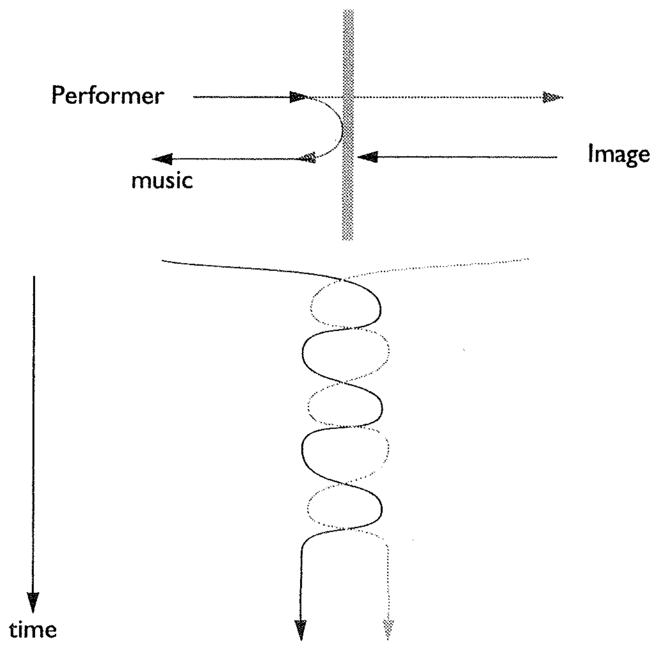


Figure 7.6 Model for an interactive music and imagery jam session. As a performer performs music, the system responds by producing imagery that reflects the musical characteristics.

The significant differences between InteractiveMTV and the main aspects of this thesis and Interactive Boston include the following:

- *Navigation*–The user does not directly navigate through the image space. Navigation is indirectly defined by the musical performance.
- *Query formulation*–Queries are formulated based on the characteristics of the music.
- *Context History*–The context history (the history of queries and points of view) has a temporal element. Queries are accumulated as the interactor performs music. And, the accumulated queries are degraded over time, leaving the most recent query as the most salient and dominating factor.

The last factor is perhaps the most distinguishing element of InteractiveMTV. As a user performs music, he or she can emphasize certain musical characteristics or motifs, and the longer the performer emphasizes them, the stronger the query. As the query gets stronger the imagery focuses in on that subject or theme. As a result, the performer can build to a climax and then release, moving off to another theme.

The system basically works as follows. At constant intervals, a a query is automatically formulated based on music the performer is playing. This query is used to retrieve relevant images from the image database. The returned images are then positioned using the MDS algorithms described in Chapter 5. As the query changes over time, the images retrieved and their relationship between each other change reflecting the changes in music qualities. Examples of screen captures are illustrated in Figure 7.7.

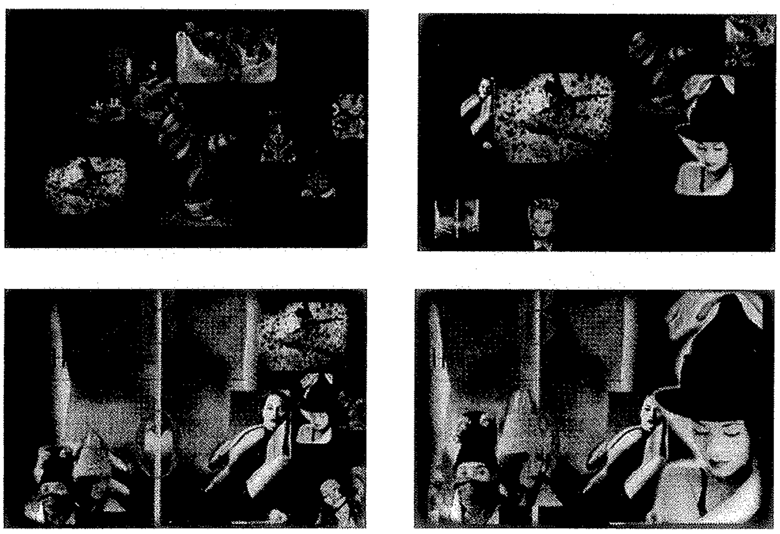


Figure 7.7 A sequence of screen captures taken from a jam session with Interactive MTV. A user performs music that has associations attached to the musical components. These associations form queries into an image database. The images is computed using multidimensional scaling techniques.

Chapter 8

Conclusions

In this thesis I have sought to breaking the stultifying boundaries of the desktop metaphor and point and click interfaces. To this end, I have defined and explored a movement-based interface that facilitates intuitive interaction with abstract information. These interactions are intuitive because they map our understanding of interaction in the real world to interaction with abstract concepts.

Based on this concept of movement through a virtual space as input to the computer, I have also demonstrated the concept of a *visual discourse*. Just as we move through mental spaces and imagery in our mind as we think and converse with other humans, this thesis defined a process for 1) constructing visual imagery that illustrate abstract concepts through visual elements, and 2) an intuitive means for moving between images.

More specifically, the results of this thesis are summarized as follows:

- *Conceptual Structure to Virtual Space Mapping*–Lakoff and Johnson identified seven conceptual structures: categorical structure, hierarchical structure, relational structure, radial structure, linear quantity scale, foreground-background structure [Lakoff, 87]. We defined a mapping from these conceptual structures to virtual space representations based on principles of metaphor. The mapping utilizes our understanding of information in terms our experiences in the physical world. This mapping included definition of the relationship between conceptual structures, information organizational structures (location, alphabet, time, category, and hierarchy)[Wurman, 89], computational structures, image schemas, and metaphorical mappings. These results are presented in Table 3.1 on page 44.
- *Conceptual Structure Derivation Algorithms*–Three algorithms were developed for deriving structure automatically from an information base: 1) multiple inheritance categorical classification (combines categorical, hierarchical and relational structures), 2) radial structure, and 3) temporal-relational structure.
- *Visual Design Techniques*–A computational approach to designing virtual information spaces that represent automatically derived conceptual structures was developed and a rationale for applying mathe-

matical processes established. The mathematical processes map structural relationships between information objects to properties of visual elements.

- *Mapping from User Movement to Visual Responses*–A set of user movements and actions in a virtual space were identified and defined, and a mapping established to computational operations that effect visual imagery as expressed through the virtual space. These mappings are based on metaphorical principles of movement in the physical world.
- *Computational Environment for Process Exploration*–A computational environment was developed to explore the process of authoring non-linear dynamic information spaces and illuminate issues involved in this process.

8.1 Future Directions

In developing the prototypes described in this thesis, I choose work with an information-base that could controlled. This decision was made to explore how annotations could be used by authors to build relationships between information objects. And, since we build the information-base ourselves, we knew the content and could judge whether or not the system had correlated the information in a way that we desired. In this way, I was also experimenting with the process of authoring multidimensional information environments. Having done this, I would now like to apply this system to large complex information-bases that I have not authored as an avenue for exploration.

3D Web Browser

The most obvious extension of this work is to build a 3D World Wide Web browser. This tool would not be intended to explore 3D models, but rather as a tool for understanding the complexities of the information available via the World Wide Web. This is perhaps the most challenging and fruitful domain for this work because it represents a fully distributed information source, where meaning and understanding is gained through an amalgamation of many information objects, authored by many people.

A main issue in applying the approach described in this thesis to a 3D web browser is that it requires that each article be parsed and key information extracted. With an expanding information base such as the information in the World Wide Web (WWW) the quantity of information is extremely large, and will only continue to grow. As a result, it is time consuming and perhaps impossible to parse through all the information contained in the WWW. Further it is a waste of resources (both computing and network bandwidth) to parse through information that is not desired by the user. Hence, rather than doing a linear search through the information base, what we would like to have is a system that probes, searches, explores and narrows in on information that is relevant to a particular users desires. This suggests a need for a special information seeking agent environment that would search through the WWW to extract the structural relationships between information objects contained within the web. The derived structural representation would then be presented to the user via an abstracted via information space described in this thesis.

Reactive and Expressive Information Objects

As described in Chapter 6, information objects, including text symbols, react to the users as they navigate through an information space. For example, if a user moves up to a symbol in a categorical space and continues to push towards it, the user transitions to the subcategory space indicated by the symbol. When this occurs the user effects a query for additional information. For each of the active objects in the scene, the reaction rules had to be hand coded. The system was designed such that these rules could easily be extended, but a language interpreter could be developed, based on the discourse grammar, that would allow people who author information objects to attach behaviors to the objects. These behaviors could be expressive in nature. For example, when a user moved up to an object, it could respond by performing an animation. Or, the behaviors could alter or reconstruct the space.

Extensible and Dynamic Space Design

As I noted in Chapter 5, the spaces designed in this thesis were intended to illustrate the coupling of analysis with presentation. Each of these designs were custom designed. The prototype system was architected such that these spaces could be extended and new spaces added, but ideally we would like the space design process to be much like designing books or documents. What if instead of having one or two categorical spaces, we had fifty or a hundred or thousands, each one expressing a specific type of relationship between elements. To do this, we need to build a set of extendible building blocks and an interpretive language for connecting them together. This approach is being explored by Ishizaki [Ishizaki, 95] and Weitzman [Weitzman, 95].

Space Building as Input

The rules associated with the discourse grammar specified in Chapter 6 are limited. The interactions that these rules afford a whole new language of interaction based both on movement and on manipulation, but only a small subset of these interactions based on movement were explored. What if the user was able to construct a symbolic space that represented a query, and the computer was able to interpret this space and respond by extending the space to show how entities in the space relate to each other. In turn, the computer's output, in the form of objects in a space, could be reformulated as input. In this way, the visual discourse would be bidirectional and could be much more conversation like.

References

[Ahlberg, 94] Ahlberg, C. and B. Shneiderman, Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays, *Proceedings of Human Factors and Computer Systems CHI,* 1994. Boston, Massachusetts.

[Allport, 95] Allport, D., E. Rennison, L. Strausfeld, Issues of Gestural Navigation in Abstract Information Spaces, *Proceedings of CHI 95, Conference Companion.*Denver, Colorado.

[Andreessen, 93] Andreessen, M., NCSA Mosaic Technical Summary. 1993, National Center for Supercomputing Applications.

[Arnheim, 54] Arnheim, R.. *Art and Visual Thinking*. London: University of California Press, 1954.

[Arnheim, 69] Arnheim, R. *Visual Thinking*. 1969, University of California Press, London.

[Berners-Lee, 92] Barners-Lee, T. Cailliau, R., Groff, J-F., and Pollermann, B., "World-Wide Web: The Information Universe," *Electronic Networking: Research, Applications, and Policy.* Vol. 1. No. 2, Meckler, CT, 1992.

[Bates, 93] Bates, J., Virtual Reality, Art, and Entertainment, *Presence: The Journal of Teleoperators and Virtual Environments, MIT Press* (1993)

[Brachman, 85] Brachman, R. J., *On the Epistemological Status of Semantic Networks,* in *Readings in Knowledge-Representation,* R.J.B.a.H.J. Levesque, Editor. 1985, Morgan Kaufmann Publishers, Inc.: Los Altos, CA. p. 191-215.

[Brill, 1992] Brill, E.. A Simple Rule-Based Part of Speech Tagger. *Proceedings of Third Conference on Applied Natural Language Processing.* 1992. Trento, Italy: ACL.

[Brill, 1993] Brill, E., *A Corpus-Based Approach to Language Learning*. 1993, University of Pennsylvania:

[Chang, 1993] Chang, B-W., and D. Ungar. Animation: From Cartoons to the User Interface. *Proceedings of UIST.* 1993. Atlanta, Georgia:

[Colby, 92] Colby, G., and L. Scholl. Transparency and Blur as Selective Cues for Complex Information. *Proceedings of SPIE.* 1992.

[Cooper, 94] Cooper, M., et.al. Information Landscapes. MIT Technical Note. April, 1994.

[Cutting, 91] Cutting, D., J. Kupiec, J. Pedersen, and P. Sibun. A Practical Part-of-Speech Tagger. *Proceedings of Third International Conference on Applied Natural Language Processing*. 1991.

[Cutting, 92] Cutting, D. D. Karger, J. Pedersen, J. Tukey. Scatter/Gather: A Cluster-based Approach to Browsing Large Document Collections. *Proceedings for the International Conference on Information Retrieval SIGIR 92.* Denmark, 1992.

[Deerwester, 90] Deerwester, S., S. Dumais, G. Furnas, T. Landauer, R. Harshman, Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science.* April, 1990.

[Fauconnier, 94] Fauconnier, G. *Mental Spaces: Aspects of Meaning Construction in Natural Language.* Cambridge University Press. Cambridge, England, 1994.

[Strausfeld, 95a] Feiner, S., and C. Beshers. Worlds within Worlds: Metaphors for Exploring n-Dimensional Virtual Worlds. *Proceedings of UIST.* 1990. Snowbird, Utah: ACM.

[Feiner, 93] Feiner, S. and McKeown, K. Automating the Generation of Coordinated Multimedia Explanations, *Intelligent Multimedia Interfaces,* Maybury, M. Ed, AAAI Press/MIT Press, 1993. Cambridge, MA.

[Furnas, 89] Furnas, G. Generalized Fisheye Views. *Proceedings of Human Factors and Computer Systems, CHI.* 1989.

[Furnas, 94] Furnas, G. and Zacks, J. Multitrees: Enriching and Reusing Hierarchical Structure. *Proceedings of Human Factors and Computer Systems CHI, 1994.* Boston, Massachusetts.

[Goldberg, 92] Goldberg, D., Nichols, D., Oki, B., and Terry, D., "Using Collaborative Filtering to Weave an Information Tapestry," *Communications of the ACM,* Vol. 35. No. 12. December, 1992.

[Holtzman, 94] Holtzman, S. *Digital Montras.* MIT Press, 1994. Cambridge, Massachusetts.

[Ishizaki, 95] Ishizaki, S. *Typographic Performance: Design Solutions as Emergent Behaviors of Active Agents.* PhD Thesis in progress. MIT, 1995.

[Jackendoff, 84] Jackendoff, R. Semantics and Cognition. MIT Press, 1983. Cambridge, Massachusetts.

[Johnson, 92] Johnson, M. *The Body in the Mind.* Chicago: The University of Chicago Press, 1992.

[Karp, 90] Karp, P. and Feiner, S. Issues in the Automated Generation of Animated Presentations, *Proceedings of Graphics Interfaces*, Halifax, Canada. 1990.

[Kelso, 93] Kelso, M. T., P. Weyhrauch, and J. Bates. Dramatic Presence. *Presence: The Journal of Teleoperators and Virtual Environments,* MIT Press (1993)

[Lakoff, 80] Lakoff, G., and M. Turner. *Metaphors We Live By.* University of Chicago Press, 1980. Chicago, Illinous.

[Lakoff, 87] Lakoff, G. *Woman, Fire, and Dangerous Things: What Categories Reveal about the Mind.* University of Chicago Press, 1987. Chicago, Illinous.

[Lakoff, 89] Lakoff, G, and Turner, M. *More than Cool Reason: A Field Guide to Poetic Metaphor.* University of Chicago Press, 1989. Chicago, Illinous.

[Lashkari, 95] Lashkari, Y. Z. *Feature Guided Automated Gollaborative Filtering,* MS Thesis, Program in Media Arts and Sciences, MIT, September, 1995.

[Lee, 93b] Lee, T. B., Protocol for the Retrieval and Manipulation of Textual and Hypermedia Information. CERN Technical Report, 1993.

[Loeb, 92] Loeb, S., Delivering Interactive Multimedia Documents over Networks. *IEEE Communications Magazine*, May 1992.

[Lynch, 60] Lynch, K. *The Image Of The City.* Cambridge, MA: MIT Press, 1960.

[Mentral, 95] Mentral, M. *MotorMouth: A Generic Engine for Large-Scale, Real-Time Automated Collaborative*

Filtering, MS Thesis, Program in Media Arts and Sciences, MIT, May, 1995.

[Minsky, 88] Minsky, M. *Society of Mind.* Simon and Schuster, 1988. New York.

[Morrison, 94] Morrison, P. & P., and the Office of Charles and Ray Eames. *Powers of Ten: About the Relative Size of Things in the Universe.* New York: Scientific American Library, 1994.

[Nelson, 81] Nelson, T., *Literary Machines.* Swarthmore, PA, 1981.

[Perlin, 93] Perlin, K., and D. Fox. Pad: An Alternative Approach to the Computer Interface. *Proceedings of Computer Graphics.* 1993.

[Rao, 94] Rao, R. and Card, S. The Table Lens: Merging Graphical and Symbolic Representations in a Interactive Focus+Context Visualization for Tabular Information. *Proceedings of Human Factors and Computer Systems CHI* 1994. Boston, Massachusetts.

[Rennison, 94] Rennison, E. Galaxy of News: An Approach to Visualizing and Understanding Expansive News Landscapes. Proceedings of *User Interface Software Technology* , Marina Del Ray, CA. November, 1994.

[Rennison, 94a] Rennison, E. Using Genetic Algorithms to Explore Large Distributed Information Bases. MIT Technical Note, April, 1994.

[Rennison, 95a] Rennison, E. Personalized Galaxies of Information. *Conference Companion to CHI 95.* Denver, CO. May, 1995.

[Rennison, 95b] Rennison, E., and L. Strausfeld. The Millennium Project: Constructing a Dynamic 3+D Virtual Environment for Exploring Geographically, Temporally and Categorically Organized Historical Information. Proceedings of the *Conference on Spatial Information Theory.* Vienna, Austria. September, 1995.

[Robertson, 91] Robertson, George G., Jock D. Mackinlay, and Stuart K. Card. Cone Trees: Animated 3D Visualizations of Hierarchical Information. *Proceedings of UIST.* 1991. Hilton Head, South Carolina: ACM.

[Seligmann, 91] Seligmann, D. and Feiner, S. Automated

Generation of Intent-Based 3D Illustrations. *Proceedings of the ACM SIGGRAPH 91.* Las Vegas, NV, 1991.

[Seuren, 84] Seuren, P. *Discourse Semantics.* Oxford University Press, 1984. Oxford.

[Shardanand, 94] Shardanand, U. *Social Information Filtering for Music Recomendation,* SM Thesis, Department of EECS, MIT, September, 1994.

[Shneiderman, 92] Shneiderman, B., *Designing the User Interface: Strategies for Effective Human-Computer Interaction: Second Edition,* Addison-Wesley Publishing Co, 1992. Reading, MA.

[Small, 94] Small, David, Suguru Ishizaki, and Muriel Cooper. Typographic Space. Companion to *Proceedings of Human Factors and Computer Systems CHI. 1994.* Boston, Massachusetts.

[Stone, 94] Stone, M. Fishkin, K. and Bier, E. The Movable Filter as a User Interface Tool. *Proceedings of Human Factors and Computer Systems CHI 1994.* Boston, Massachusetts.

[Strausfeld, 95a] Strausfeld, L. Financial Viewpoints. *Proceedings of CHI 95, Conference Companion.* Denver, Colorado.

[Strausfeld, 95b] Strausfeld, L. *Embodying Virtual Space to Enable Understanding of Information.* MS Thesis. Massachusetts Institute of Technology. Cambridge, MA. August, 1995.

[Sulton, 83] Sulton, G., and McGill, M.J., *Introduction to Modern Information Retrieval.* McGraw-Hill, 1983.

[Tversky, 77] Tversky, A. Features of Similarity. *Psychological Review.* Vol. 84. No. 4. July, 1977.

[Weitzman, 95] Weitzman, L. *The Architecture of Information: Interpretation and presentatin of information in dynamic environments.* PhD Thesis. MIT, 1995.

[Wurman, 89] Wurman, R. S. *Information Anxiety: What to do when information doesn't tell you what you need to know.* Bantam Book, 1989. New York.