Galaxy of News

An Approach to Visualizing and Understanding Expansive News Landscapes

Earl Rennison
Visible Language Workshop
MIT Media Lab
20 Ames St.
Cambridge, MA 02139
E-mail: rennison@media.mit.edu

ABSTRACT

The Galaxy of News system embodies an approach to visualizing large quantities of independently authored pieces of information, in this case news stories. At the heart of this system is a powerful relationship construction engine that constructs an associative relation network to automatically build implicit links between related articles. To visualize these relationships, and hence the news information space, the Galaxy of News uses pyramidal structuring and visual presentation, semantic zooming and panning, animated visual cues that are dynamically constructed to illustrate relationships between articles, and fluid interaction in a three dimensional information space to browse and search through large databases of news articles. The result is a tool that allows people to quickly gain a broad understanding of a news base by providing an abstracted presentation that covers the entire information base, and through interaction, progressively refines the details of the information space. This research has been generalized into a model for news access and visualization to provide automatic construction of news information spaces and derivation of an interactive news experience.

KEYWORDS: Information visualization, abstracted information spaces, pyramidal information structures, 3D interactive graphics, information space design, information interaction design.

INTRODUCTION

As we enter the information age, concepts concerning news production, distribution, access and visualization are rapidly changing. Until recently, news was authored for relatively static media, such as print, audio/video tape. The information content was distributed in a one-way exchange via newspapers, radio, television, and so forth. And, in many ways existing computer systems and infrastructure are still directed toward these forms of distribution and management. Accessing and understanding news information in this structure is a secondary action and is most often left up to the readers, with little support by the information infrastructure.

To expound the problem, increasing news production resulting from the rise in connectivity has led to increasing complexity of news organization, management, and hence, understanding. As a result, authoring news articles and/or presentations that relate to other information (e.g. articles or presentations that are linked) will become increasingly more difficult to construct in a meaningful way so as to help the reader to understand the full nature of the news or information that is being presented. The current information infrastructure simply cannot handle the exploding scale of news information and its cross correlation. Hence, what is truly needed is an intelligent infrastructure that automatically builds the correlations and relationships between news articles, and automatically construct an environment, based on the information content, that allows readers to dynamically explore the expanding news base and allows them to gain an understanding that is deeper than what they would gain by looking at individual news articles.

To address this problem, we would ideally like to develop news systems that facilitate the following objectives:

- Allow people to explore and effectively browse through massively large news spaces
- Combine the effective aspects of both filtering and browsing and the ability to move between these modes of operation seamlessly with a single interface
- Facilitate ability to understand relationships between independently authored news items
    - To see how articles relate to each other
    - To find relationships between articles that where not previously known or obvious

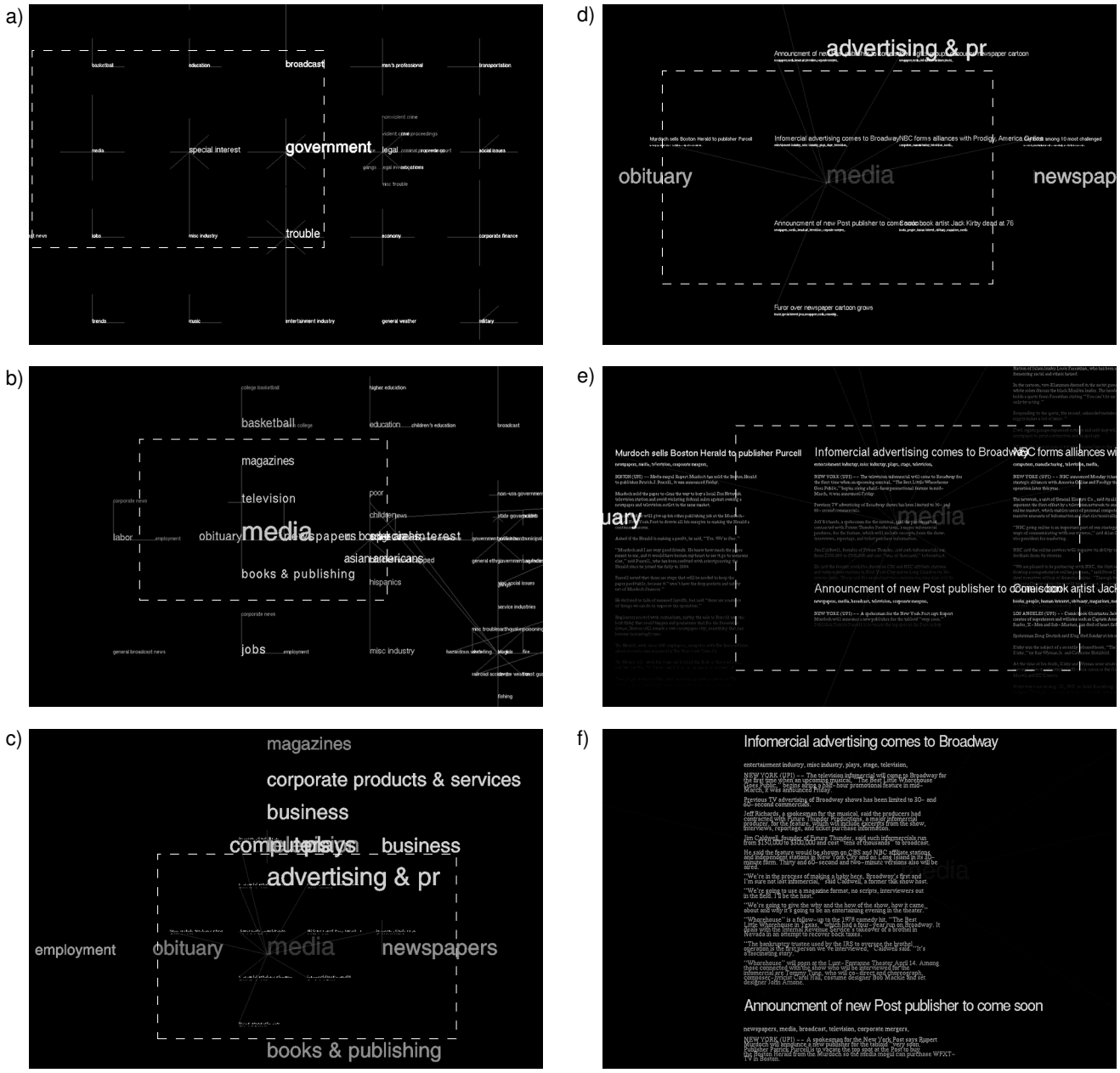

Figure 1: Screen Captures from a Progressive Zoom into a Sample News Information Space
a) Initial, wide-angle view of news space; b) area around Media cluster; c) zoom into Media cluster (headlines begining to show); d) headlines for Media cluster; e) article body beginning to drop in; f) full presentation of article headlines and body.
Dashed rectangles indicate the areas that are zoomed into and shown in the next screen capture.

- Organize disconnected articles into dynamically formed groups, based on the content of the articles, that allow quick access to related information and the ability to quickly understand the relationships

An infrastructure that supports these objectives would automatically build the correlations and relationships between the information elements (hence linking the elements together), and provide coupled mechanisms that dynamically constructs an environment that facilitates interactive navigation and intuitive access to related or correlated information. This is the approach explored with the Galaxy of News System.

EXAMPLE NEWS INFORMATION SPACE

To provide a better understanding of what the Galaxy of News system provides, let's consider an actual scenario currently used to demonstrate the system. In the Visible Language Workshop we receive news feeds from Clarinews. These news feeds are stored in a directory

hierarchy specific to Clarinews. To initiate the visualization and access process, a user specifies all or a portion of the Clarinews database to be processed. The system then parses the content of articles and constructs relationships between the articles. These relationships are then stored in a relationship database. Once the process of creating these relationships is specified, the relationship database can be automatically updated daily.

After setting up the relationship construction process and building some relationships, the user starts a visualization front-end application, specifying what set of relationships to use (and hence, what portion of the database to visualize). When the front-end application starts up, it loads the appropriate relationships, and then constructs an information space that will help the user navigate through the information. Then, using the mouse and two keys (left mouse button to zoom in and right mouse button to zoom out), the user navigates through the information space and the system automatically presents information appropriate for the user's position in the space.

To give you an idea of what is presented to the user, the following example, shown in Figure 1, illustrates a sample path through an actual news space. The image in Figure 1.a. shows the initial view of the information space. In this view, root keywords of news clusters in an abstracted news hierarcy are shown. These root keywords provide an abstracted, representational view of the entire news database on a single screen. These root clusters are derived through an abstraction and generalization process described later in this paper. Also note that the space presented is non-linear and has a simulated fish-eye [9] effect highlighting the elements that are in the center of the space. In Figure 1.b., the user has zoomed toward the "Media" cluster. At this stage, the Galaxy of News application begins to expose the details of this cluster, as well as other clusters in the area. Note for reference that the structure of this cluster is illustrated in Figure 2.

As the user continues to zoom into the “Media” cluster, as shown in Figure 1.c., news articles previously located in other areas of the space are pulled into the background using animation. This process of animation maintains fluid consistency and helps the user understand what the system is doing. Note, at this point the typeface for the word “Media” begins to fade away, yet it is still visible. This gives the user a subtle clue that they have entered the “Media” subspace and helps maintain context as to where he or she is within the space without using a global navigation map.

Further zooming reveals the headlines of the articles relating to the media cluster, shown in Figure 1.c. It is hard to visualize in the screen shots provided, but subtle red lines illustrate to the user the associations of articles to keywords. As the user continues to zoom into the “Media” cluster, the red lines linking keywords and articles gradually fades away, and simultaneously, the bodies of the articles gradually drop in under the article headlines, as shown in Figure 1.e. These gradual changes of the keyword transparencey, red subject-article line transparency, and bodies of the articles provide for a very natural transition from the meta-space provided by the keywords and the detailed space of the news articles. If the user continues to zoom (Figure 1.f.), the full body of the articles is displayed so that it can be read.

An important aspect to point out with these images is that they convey only a limited view of what actually goes on when a user navigates through the space, though they do provide a general sense of the *experience*. A key part of the system is it's fluid movement and display of information, a feeling not conveyed in these images.

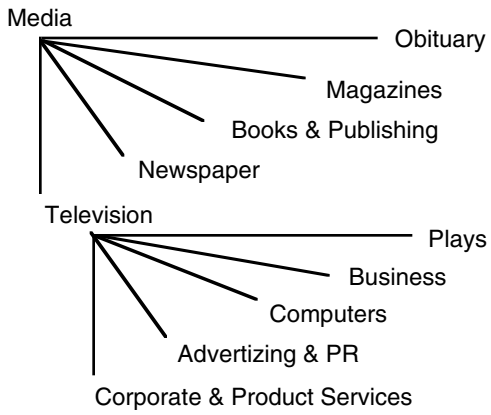

Figure 2: Example Derived Keyword Hierarchy

GALAXY OF NEWS MODEL FOR NEWS ACCESS AND VISUALIZATION

The Galaxy of News system explores a generalized model for news managing, accessing, and visualizing expansive news bases. This model (illustrated in Figure 3) strives to create a structured environment that accommodates automated integration of independently authored articles into self-constructing information spaces that allow users to visualize and access information in a fluid and interactive fashion. The key element of this model is the separation of news authoring, and information space and interaction design into two separate processes. In many ways, this is similar to the process of constructing a daily newspaper, except that the delivery environment is much more dynamic and accommodates direct user interaction.

The model defines three distinguishable perspectives or views: 1) the author's view, 2) the information space and interaction designer's view, and the 3) the reader's view. The author's sole purpose is creating content or information objects. Information space and interaction designers focus on specifying the process of constructing information spaces based on classes of information, as opposed to specific instances of information. And, the reader views the amalgamation of the information space constructed from news content. This approach breaks

information production down into the creation of content, and the creation of ways to "experience" the content (i.e. the combination of visualization and interaction). This split predisposes this model to address information scale.

The model is comprised of four fundamental layers. At the base, or core, of the model is the news information base, and the other three are specifications of what information to visualize and how to visualize that information at each instance of interaction. A description of each of these layers follows.

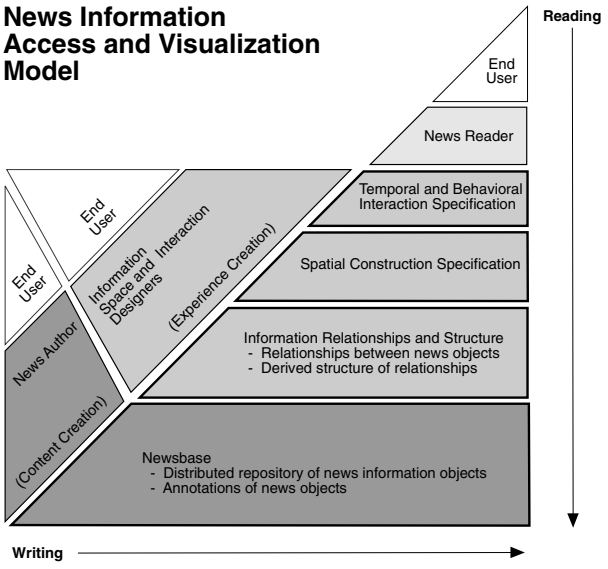

Figure 3: Galaxy of News Model for News Information Access and Visualization

News base

The news base consists of a set of information objects and their annotations. Example news information objects include text articles, photos and captions, graphs, graphic displays, video clips, sound clips, and so forth. In this model, the news base does not consist of raw data; rather only data that has been processed to represent some informational view of the data. Annotations consist of properties or symbolic descriptions of the information objects. They are either supplied by the author or automatically derived from the data, if feasible (e.g., a part of speech tagger can be applied to text articles to extract nouns/noun phrases and verbs [2; 3; 7). News articles provided by wire services provide annotations such as keywords, slugwords, location, time, subject, and so on. This information is used to build relationships between news articles.

Information Relationships

For a given set of information or classes of information, the relationships between instances of information objects can be defined. Examples include relationships between subjects, actions, and/or linguistic objects; between the time events occurred; between temporal ordering of events; and between combinations of these, as well as other properties.

Information relationships also consists of structures derived from other relationships. The model does not define specific techniques for deriving relationships structures, only that the relationships are constructed from the properties or symbolic descriptions extracted from the information objects. The Galaxy of News system defines one approach to constructing relationships and deriving a structure of the relationships (i.e. associative relations between subjects, described later in this paper); however, the model is not limited to that approach. Ultimately, it is the responsibility of the information space designer to specify the relationships, or the process of deriving the relationships, between information. The reason for this is that the spatial construction, and hence the visualization, process is intimately coupled with the information relationships.

Spatial Construction Specification

Using the information relationships described above, a spatial construction specification defines how information is presented graphically and spatially to users. The spaces can consist of two-dimensional, three-dimensional, or n-dimensional layouts, where dimensions above three can be modeled as micro worlds [8]. It is important to note that the information spaces are not defined for specific instances of information objects, rather for classes of information objects and the relationships between these classes. The actual information space constructed is derived from the collection of information objects, resulting in an emergent spatial structure. This is particularly relevant to news because the subjects change dynamically and a static space would not reflect these changes.

Temporal and Behavior Interaction Specification

This level defines the presentation and behavior of information objects during each instance of interaction with the user. The presentation of an appropriate level of detail at each view is fundamentally important. If only a limited subset of the information space can be presented at any moment or view, it is the responsibility of this layer to present the user with dynamic visual cues denoting the structure of the space and where the user is in the space. Temporal behavior can also be used to illustrate the underlying relationships between the information. For this purpose, the information space may be dynamically modified or reconstructed, using animation to illustrate changes [4]. It is the role of the "Information Interaction Designer" to specify these actions.

INFORMATION VISUALIZATION AND ACCESS IN THE GALAXY OF NEWS

One of the primary aspects of the Galaxy of News system is its approach to visualizing and accessing news information; hence the visualization aspects drive the architectural considerations of this system. The visualization approach presented in this paper is the

result of extensive experimental research on visualization of multidimensional information conducted by the author and other members of the Visible Language Workshop at the Media Laboratory [5; 6; 10; 14]. The Galaxy of New system investigates several information access and visualization principles, including:

- 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
- Fluidity of interaction to 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 principles define an outline for building a structured hierarchical representation 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 in Figure 1). 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.

At first glance, these principles are similar to cone trees [13]; however, there are several significant differences. First, the hierarchical form is not explicitly presented to the user. The hierarchy 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 hierarchy to be presented. Third, the user is able to navigate through the hierarchical space in an immersive fashion. As the user goes deeper into the hierarchy, the system reveals the substructures of the hierarchy.

The process of zooming in the information hierarchy is a form of interactive filtering. A similar approach was explored in the PAD system [12], 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 information objects. 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 [8], yet different in that the space is not tied to any dimensions -- the space is abstract.

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.

Information Space and Interaction Design

The Galaxy of News system explores the separation of news information space design and the authoring of autonomous articles in a dynamic environment. To this end, 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 the following features. 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. The keyword hierarchy is derived from the relationships between news articles. An example keyword hierarchy derived automatically from an actual news base is shown in Figure 2 above. Because this hierarchy 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, keywords may be duplicated in the keyword hierarchy, but there is only one copy of each article. An article is dynamically moved around the space depending upon where the user is within the space.

Since presentation of the space is non-linear, the system determines what information is to be 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)

Also, some parameters are held fixed in the information space: the keyword locations in space are kept fixed to maintain a basic sense of structure, and the color of the keyword groups remain constant to indicate keyword clusters. These fixed parameters aid the user when

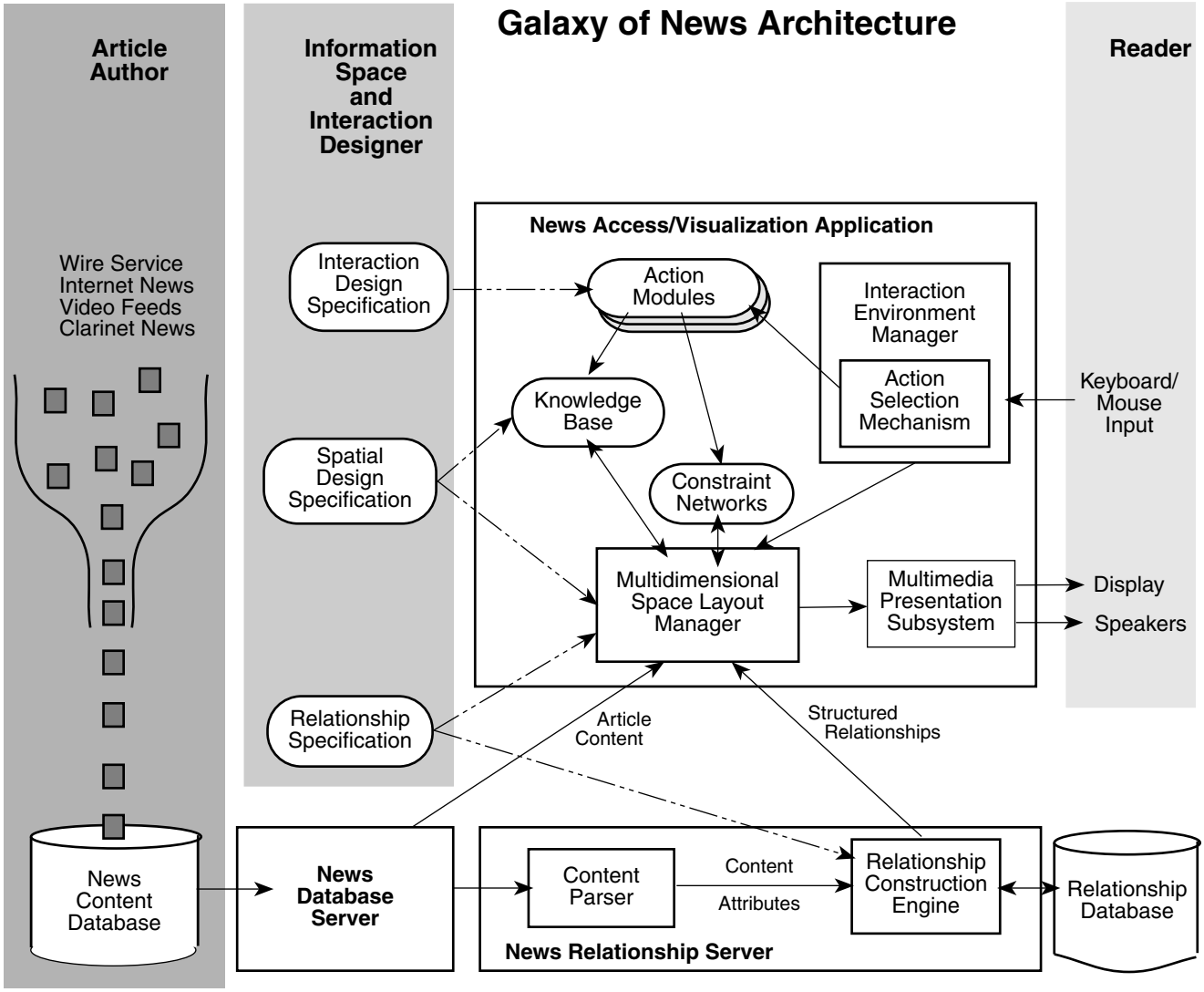

Figure 4: Architectural Overview of the Galaxy of News System

navigating through the space by giving the user a sense of where they are within the space without having to provide something like a global map to aid in navigation. Further, the dynamic elements of the space–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.

GALAXY OF NEWS SYSTEM

The architecture of the Galaxy of News system is illustrated in Figure 4. As shown in this illustration, the architecture emphasizes three important aspects or views of the news production, management, access and visualization process: 1) authoring articles, 2) designing news information spaces for users to explore the news, and 3) reading or exploring the news. The architecture of the system supports the integration of these three views into a single environment.

Information Relationship and Structure Extraction Mechanisms

At the heart of the Galaxy of News system is a mechanism and representation for learning the relationships between news articles. The representation that maintains relationships between news articles is called an Associative Relation Network, or ARN. This representation and mechanisms that drive it use reinforcement techniques to capture the relationships between documents, and symbols extracted from those documents. The relationships between symbols contained in an ARN define the relationships between documents.

An ARN, illustrated in Figure 5, maintains weighted relationships between symbols contained in the network. 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

$$c_i = \sum D_X : \{S_i, ...\}$$

where, $D_X : \{S_i, ...\}$ denotes a document containing S_i

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

$$w_{i,j} = w_{j,i} = \sum D_X : \{S_i, S_j, ...\}$$

where, $D_X : \{S_i, S_j, ...\}$ 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. In effect, an ARN is used to learn about the structure of the news base. Hence, when the system starts, it has no previous knowledge about the symbols that are used to construct the network, and as it sweeps through the database, it learns the relationships between a set of symbols contained within the database.

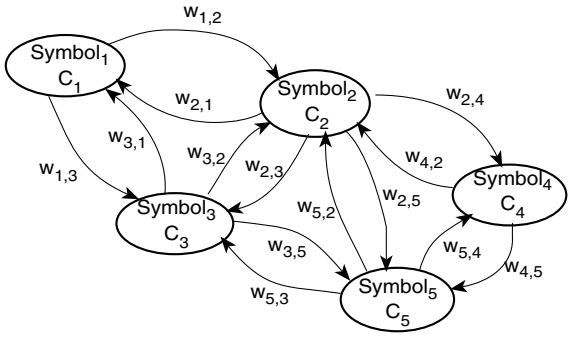

Figure 5: Simplified Associative Relation Network Representation

An ARN forms the basis for constructing an information space to allow people to explore the relationships between documents. Further, this representation is used to construct an abstracted information hierarchy. The information hierarchy is extracted from the ARN by using the following recursive process:

1. Search through the ARN and find all the statistically independent symbols
2. For each independent symbol, find all the symbols statistically dependent on the independent symbol
3. For each set of dependent symbols, find the independent symbols
4. Repeat steps 2 and 3 until all the dependent symbols are independent of one another.

The information hierarchy resulting from this process is used to progressively refine the presentation of information to the user, as described above.

News Information Space Construction

The current implementation uses the following approach to construct the news information space described above. First, news articles are parsed as they are received from their source, i.e. a Clarinet news feed. The parsing process extracts key information, such as keywords, locations, time event(s) occurred, subjects, actions, and so forth. After parsing, the original articles are stored in a database until referenced by the user.

Second, the key symbols are inserted into an ARN to construct relationships between articles. After parsing the articles and constructing the associative relation network, the system searches the network to extract the hierarchical relationships between the symbols, and hence, the articles. This process is described in the previous section.

After sorting the symbols into hierarchies, the peer elements in the hierarchy are sorted spatially so that symbols that are more closely related are placed next to

one another. A relational distance between two symbols is determined by comparing the sets of associated symbols corresponding to the two symbols in question. This relational distance is used to perform a two dimensional sort. A third temporal sort is currently being developed.

Given the hierarchically structured and spatially sorted symbol set, a visual information space is constructed. The current implementation constructs a space that uses *x*- and *y*- dimensions to place the independent or root keyword symbols, and *z*- dimension to place keyword symbols of increasing levels of detail, followed by the articles. A recursive algorithm is used to place keywords within the space. First, the root keywords are placed using a constraint-based grid structure. Then, each set of keywords associated with the root keywords are placed behind their respective root keyword using separate constraint-based grids. The result of this process is a hierarchical grid structure containing the keywords.

During dynamic interaction with the user, the system employs both constraints and a set of heuristics to control the presentation of the hierarchical information to the user. The constraints and heuristics used in the current implementation are described in the previous subsection.

System Architecture Components

The approach used to construct a news information space described in the previous subsection, is split into three main components: a news database server, a news relationship server, and an access and visualization application (illustrated in Figure 4). The relationship server is separated from the front-end application for speed and efficiency reasons. The separation between the news database and the news relationship server is presently a logistical solution, but in the future these two servers may be combined, especially to eliminate the potential problem of database inconsistencies.

The role of the news relationship server is to parse a set of news articles and build relationships between attributes contained in the articles. How these relationships are constructed is defined by a relationship specification (or procedure) defined by the information space designer. A relationship construction engine interprets the relationship specification, and uses it to build a structure representing the relationships between articles stored in a database. The relationship structures, such as the associative relation network described in the previous subsection, are derived by a server separate from the front-end application. These processes are separated because of the time required to build the relationships and allows the relationships to be used multiple times by multiple processes.

The system front-end application consists of two primary components: a multidimensional space layout manager, and an interaction environment manager. These two components are supported by a knowledge-base, constraint networks, and action modules. The knowledge-base is created by the information space designer and specifies rules for constructing a space based on the relationships between articles and meta-data that describes the articles, in this case keywords. The action modules are written by the information space and interaction designer, and define how the system responds to changing situations that are based on user interaction. To effect change in the display of information, the action modules can modify either the knowledge-base or the constraint networks.

Multidimensional Space Layout Manager. The Multi-dimensional Space Layout Manager constructs an abstracted information space based on the relationship structures specified by the 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 layout manager must compute a layout each time the user moves within the space. This continuous spatial regeneration creates a reactive environment. The layout manager utilizes a knowledge-base provided by the information space designer to perform this operation.

Interaction Environment Manager. The Interaction Environment Manager controls the systems response to user interaction. It performs this function by building a model of the user situated in an environment, in this case a three dimensional space. As the user moves within the environment an action selection mechanisms determines what actions, if any, the system should take to respond to the user. In the current implementation, the action selection mechanism uses a simple rule-based approach. When an action is selected and executed, it modifies the state of the knowledge-base and/or the constraint networks. These changes are subsequently interpreted by the layout manager to effect the layout and presentation of the space.

DISCUSSION

In developing the Galaxy of News system , we considered several alternative approaches. We considered using hypertext and hypermedia concepts and systems [11] to aid in accessing related articles or information in general. This technique has met with some success and has been employed by Mosaic[1] as an interface to World Wide Web documents. However, there are several inherent problems with the hypermedia approach, such as 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.

Another approach to this problem is to build news filters. In the past, we had a natural form of news filtering, known as editors, who selected what news we would read, or see (in the case of television). And, we had to either accept their filtering or choose another newspaper or channel. We had no choice but to trust what the newspapers, magazines, and newscasts were telling us. This choice was forced upon us because we had no other means to access information concerning events that happened in distant places and/or at different times. But, with the advent of the Internet and other related technologies, we will soon have direct access to ALL news articles. This clearly presents a dilemma. We cannot possibly read all the news available, and yet can we, or will we, trust news editors and other forms of news filters, such as intelligent autonomous agents. And, if we are aware that we have access to all news information, then trust concerning how information is filtered will be of primary importance. Hence, we took the approach to provide *access* to all the articles instead of using filtering or retrieval techniques.

FUTURE DIRECTIONS

The Galaxy of News system is currently targeted toward a fairly specific class of information–news articles. We are extending the system to handle other types of information objects, such as documents contained in the World Wide Web and images contained in a National Geographic archive. The Galaxy of News system shows promise in providing a tool for understanding the increasing size and complexity of information available via the WWW and other large information bases.

In addition, we are working on extending the system to support more sophisticated relationships between information objects. This work includes extending the system to support not only symbolic relationships, but also parametric relationships. Parametric relationships may be very useful in situations where it is not possible to derive symbolic descriptions of objects, such as with video, audio and images that have not been annotated. And, along with these more sophisticated relationships, we are working on ways to give readers the ability to control how the relationships effect the presentation of information, allowing users to explore deeper and more complex relationships between information objects. In effect, this would give users the ability to switch roles between reader and information space designer. Eventually, we can also see users moving back and forth between author, space designer and reader in a fluid fashion as they read and augment an expanding information space.

Because the space constructed by the current system is based entirely on the contents of a news base, the structure of the space will change along with the news base. This is effective when the user is primarily in an exploratory mode, the type of browsing the Galaxy of News system was originally targeted towards. However, if a user wants to conduct a more directed search with a specific type of information in mind, then it would be useful to have personalized views or structures to assist users with navigation. Hence, we are currently developing a learning algorithm that will monitor a user's interactions and provide the user with the ability to restructure the space so that the placement of information is more familiar to the user.

As noted earlier in this paper, we are developing a temporal extension to the associative relation network. This will allow us to learn event sequences and event sequence relationships. The temporal ARN can be applied to learn both implicit event sequences (i.e. sequences of events specified over disjoint documents) as well as explicit event sequences such as those found in videos. Once the temporal relationships are learned, they can be used to structure an information space and browse through temporal arrangements of information.

Finally, we are exploring methods for enhancing the action selection mechanism. As the relationships between the information objects grow more complex, the mechanisms for controlling what information is most appropriate to present in a given situation will need to be more sophisticated. We are exploring the utility of behavior-based artificial intelligence techniques to deal with these complexities. Further, we are exploring methods for providing guided tours through information spaces based on a user's previous history of interaction or information that is of interest to a local culture.

CONCLUSIONS

In this paper we have presented results of research conducted in developing the Galaxy of News prototype system. This system allows a news reader to gain a broad understanding of the contents of a news database, and as the user selectively zooms, the news reorganizes itself in relation to the focus of interest and progressively refines the details of the abstracted news information space. The information space is automatically constructed based on relationships derived from the contents of the news stories.

From the lessons learned in developing this system and its projected utility in managing, organizing, visualizing and accessing news information, we have defined a model for news information visualization and access. This model defines four layers that are used to construct an interactive news information space: 1) the news base, 2) derived relationships between the news information objects, 3) spatial construction specifications, and 4) temporal and behavior interaction specifications. The model also defines three different views onto the components used to construct the information space: 1) the reader's view, 2) the information space and interaction designer's view, and 3) the author's view. This model provides a structure for handling the rapidly expanding news base.

The experience of interacting in the abstract information spaces created by the Galaxy of News system clearly indicates the utility of abstracted multidimensional spaces for browsing through news bases. Galaxy of News represents a new approach to addressing problems of information scale and complexity that we will face in the midst of global interconnectivity. We believe the immersive techniques employed by the Galaxy of News system will significantly enhance information visualization, navigation and access.

ACKNOWLEDGMENTS

The author would like to acknowledge the continued support, advice and direction provided by Professor Muriel Cooper, Ron MacNeil, Dave Small and Suguru Ishizaki of the Visible Language Workshop. The author would also like to give a special thanks to Lisa Strausfeld for her insight and ideas which helped formulate and solidify the concepts explored in this research. A special thanks also goes to Maia Engeli, Robin Kullberg, Ishantha Lokuge, Rob Silvers, Jeffry Ventrell, Louis Weitzman, Yin Yin Wong, and Xiaoyang Yang for providing many critiques and suggestions. This work was sponsored by ARPA, JNIDS, NYNEX and Alenia.

REFERENCES

[1] Andreessen, Marc, NCSA Mosaic Technical Summary. 1993, National Center for Supercomputing Applications.

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

[3] Brill, Eric, *A Corpus-Based Approach to Language Learning*. 1993, University of Pennsylvania: PhD Thesis.

[4] Chang, Bay-Wei, and David Ungar. Animation: From Cartoons to the User Interface. *Proceedings of UIST*. 1993. Atlanta, Georgia:

[5] Colby, Grace, and Laura Scholl. Transparency and Blur as Selective Cues for Complex Information. *Proceedings of SPIE*. 1992.

[6] Cooper, Muriel, et.al. Information Landscapes. MIT Technical Note. April, 1994.

[7] Cutting, Doug, Julian Kupiec, Jan Pedersen, and Penelope Sibun. A Practical Part-of-Speech Tagger. *Proceedings of Third International Conference on Applied Natural Language Processing*. 1991.

[8] Feiner, Steven, and Clifford Beshers. Worlds within Worlds: Metaphors for Exploring n-Dimensional Virtual Worlds. *Proceedings of UIST*. 1990. Snowbird, Utah: ACM.

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

[10] Masuishi, Tetsuya, David Small, and Ronald L. MacNeil. 6,000x2,000 Display Prototype. *Proceedings of SPIE/IS&Ts Symposium on Electronic Imaging Science and Technology*. 1992.

[11] Nelson, Ted, *Literary Machines*. 1981, Swarthmore, PA.

[12] Perlin, Ken, and David Fox. Pad: An Alternative Approach to the Computer Interface. *Proceedings of Computer Graphics*. 1993.

[13] 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.

[14] Small, David, Suguru Ishizaki, and Muriel Cooper. Typographic Space. *Proceedings of CHI*. 1994. Boston, Massachusetts: ACM.