A Semantic Approach in Modeling and Visualizing WWW Bookmarks

COSTAS PETROU, DIMITRIOS CHARITOS, DRAKOULIS MARTAKOS
Department of Informatics
National and Kapodistrian University of Athens
TYPA Buildings, Ilisia, 15771, Athens
GREECE

Abstract: - Bookmarks are the facilities found in current web browsers, which allow a user to revisit a WWW page. Most WWW browsers provide a hierarchical structure for aiding bookmark organization in searchable subsets. Average users tend to create a large number of bookmarks during multiple browsing sessions and consequently such hierarchies quickly become complex and hard to manage. Moreover, it is questionable whether a user recalls the significance of a single bookmark in the context of his current informational needs, as the majority of WWW browsers only store the title of the relevant page. As an answer to the aforementioned issues, we propose a novel system, which supports the organization of bookmarks and also suggests a 3D spatial metaphor as a visualization technique for navigating within bookmark collections. In specific, a semantic modeling framework for shaping the bookmark "space" and a corresponding 3D visualization for this "space" are suggested. The proposed framework is able to capture both human and system derived meta-information which describes bookmarks. It also incorporates inference capabilities by exploiting fuzzy set theory and fuzzy logic. Additionally, this paper attempts to define the types of meta-information, which could be utilized for describing bookmarks. Finally, a system architecture for acquiring, manipulating and visually rendering these types of meta-information in a 3D spatial context is proposed.

Key-Words: Bookmarks, 3D Visualization, Semantic Modeling, Fuzzy Logic, Expert Systems, Virtual Environments.

1 Introduction

The study of user's navigation in complex informational spaces, such as hypertext, has revealed problems termed as "disorientation" and the "lost in the hyperspace" syndrome. As an answer to such problems, also encountered in the more specific area of the World-Wide Web, browsers like Netscape Navigator© and Microsoft Internet Explorer© have been geared with interface facilities that would allow a user to revisit a node[1]. These facilities are known as "bookmarks" and can be found even in dated closed hypertext systems such as the Symbolics Document Examiner [1]. Jakob Nielsen [2] defines the main purpose of bookmarks as: the "marking" of a hypertext node which characterizes it as "interesting", for the purpose of gaining direct access to it at some later point in time, without having to traverse the path that has been previously traversed to reach it. He also suggests that bookmarks can be seen as objects in their own right, meaning that they can have an existence outside a simple bookmark list. The obvious advantage of such a modeling view is the added flexibility to move bookmarks around and to build different kinds of collections of bookmarks for different purposes. Influenced by Nielsen's viewpoint, this paper proposes a novel approach to bookmark modeling. Moreover, the paper suggests a 3D spatial metaphor as a visualization technique for navigating within bookmark collections. It has been proposed [3] that a set of requirements, should be considered when designing bookmarks organization schemes. We have refined these requirements and concluded the following points:

- The size of a typical bookmark list is significantly smaller than any kind of digital document repository. Therefore, we expect that the system can afford to store and manipulate a rich set of descriptive meta-information regarding bookmarks.

[1] In the context of this paper, the word "node" indicates a WWW HTML page.

- The collection is highly evolutionary, as the user continuously updates it during his browsing sessions. This implies that the re-organization of the collection should follow any shift to user's conceptual space (also noted in [4]).
- The collection retains a subjective, user-oriented profile, rather than an objective and public-oriented one. Therefore, the subjectivity, which is inherent in bookmark categorization, must be captured and represented by the system. Originating from the theory of semantic nets, the idea of modeling bookmarks using graph-based representations has been previously tested (e.g. object-oriented modeling in General Magic's Magic Cap user interface [2]). In the same vein, this paper favors a graph-based modeling approach and further attempts to capture human ambiguity, as will be explained in a subsequent paragraph.
- The frequency of adding bookmarks is expected to be a small proportion of the frequency of visiting nodes during browsing sessions. This allows for analyzing and processing the contents of the node without adding a considerable overhead to the system.

Irrespective of the adopted organizational scheme, the user still has to cope with searching amongst a large number of bookmarks for choosing the most relevant to his current interests. Most WWW browsers organize bookmarks in hierarchies, for facilitating such a search. However, a problem that users are usually faced with is that they end up with large hierarchies, which they cannot fully expand and effectively manipulate at the same time. Furthermore it is not possible to re-arrange bookmarks in a way that reflects users' ad hoc informational requirements. Therefore, we consider as an essential requirement for any system that supports bookmarking, to facilitate navigation within the bookmark "space" as well as to adapt the "space" to users' current informational needs.

This paper proposes a semantic framework for modeling bookmark collections and suggests a corresponding 3D visualization for these collections. The proposed framework is able to capture a rich set of both human and system derived meta-information, regarding bookmarks. The different types of meta-information are classified. The paper also proposes a system architecture for acquiring, manipulating and visually rendering these types of meta-information in a 3D spatial context. A prototype system implements the above framework and some technical aspects of it are also discussed.

2 Semantic framework of WWW bookmarks

2.1 Types of meta-information

We suggest an orthogonal categorization space for bookmark meta-information. The two axis of this space are:

Reliability: denotes that the meta-information remains valid, e.g. time invariant. This ranges from meta-information that is not changing over time to time dependent meta-information.

Source: denotes the supplier of meta-information, ranging from system- to human-provided. The intermediate values denote semi-automatically defined meta-information.

A visualization of this space is presented in Fig.1.

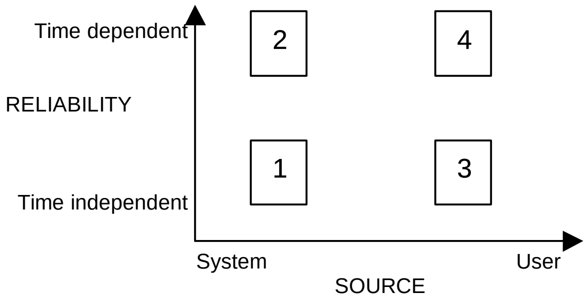

Fig. 1: The proposed meta-information space

Accordingly, the four limits of the space denoted by the boxes one to four, may be described as follows:

1. The meta-information is automatically extracted by the system and always remains valid.
2. The meta-information is automatically extracted by the system and its validity degrades during time.
3. The meta-information is user provided and time-invariant.
4. The meta-information is user provided and its validity degrades during time.

Thereafter, we use these four limits to characterize meta-information as Type-1 through to Type-4. The proposed framework is capable of handling the following types of meta-information:

Type-1

- Last time that the user has visited a node (in date/time).

- Mean-time spent during previous visits to the node (in min).
- Mean-time for transferring the contents of the node (in sec).
- Availability of the host (requests satisfied/ total requests).

Type-3

- A semantic model comprising Concepts, Relationships and Weights.

Between Type-1 and Type-3

- A set of inference Rules, which augments the semantic model with deductive capabilities.

Between Type-1 and Type-2

- A keyword index for the terms found in nodes.
- Size of the node in terms of bytes.
- "Profile" of the node: what type of content the bookmark has (dominating text, images, tables).

In our prototype, the keyword index creation begins with a pre-processing phase, which removes stopwords. Consequently, selected parts of the HTML page are indexed: title, meta tags, headings, emphasis, strong etc. The approach of selectively indexing parts of the node is favored due to its low storage requirements and also due to the fact that in our framework, indexing plays a secondary role in bookmark characterization.

It has to be stressed that the more the meta-information is concentrated nearby the start of the axis, the more system-provided and time-insensitive it is; thus it is "easier" to obtain and it is expected to provide a more "precise" description of the actual nodes. The supported meta-information, has been carefully selected to conform to this argument, without underestimating user-provided meta-information, which would be impossible for the system to automatically elaborate.

2.2 A Semantic model of bookmarks

The proposed model supports a user-centered and domain-depended approach to adaptive bookmark modeling and is based on semantic nets. In the field of human-derived knowledge representation, semantic nets have been widely used, especially for representing aspects of natural language [5][6]. The proposed Semantic model is effectively a kind of user defined semantic net. Apart from nodes and arcs that serve as knowledge representation elements, the model is enhanced with deductive elements. Consequently, the model ensures not only the representation of human knowledge but also the approximate reasoning triggered by this knowledge. Fuzzy set theory and fuzzy logic have been employed for implementing this last feature. A simplified Entity-Relationship (E-R) diagram for the proposed model is shown in Fig. 2.

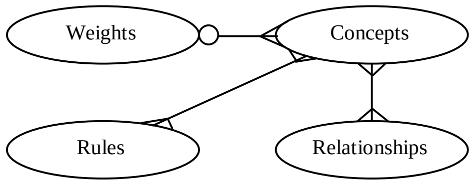

Fig. 2: E-R diagram of the proposed Semantic model

On the basis of the proposed semantic model and the meta-information which is automatically extracted by the system, a bookmark is defined as Table 1 explains:

Attribute	Domain	Comments
ID	*number*	Unique value for every bookmark
Title	*text*	Node's title
Reference	*text*	Node's URL
Size	*number*	Node's size in bytes
LastVisit	*date/time*	Information about last visit to node
SMeanTime	*time*	Mean time in node
TMeanTime	*time*	Mean time to transfer node
Availability	*number*	requests satisfied/ total requests
Keywords	[*keywordID*]	List of keyword references
Profile	*number*	Encoded profile for the node, such as text, text with images etc.
Semantics	[(*ConceptID/ value*)]	List of pairs (ConceptID/value)

Table 1: Bookmark definition

The next paragraphs discuss the four building elements of the model, namely Concepts, Relationships, Rules and Weights:

Concepts provide user defined semantics for the bookmark. These semantics can be assigned with various gradients, which express the notion that the bookmark is partially described by the Concept. The use of fuzzy set theory was favored for modeling Concepts, where each Concept corresponds to a fuzzy set. As noted in [7], "*fuzzy set theory is a marvelous tool for modeling the kind of uncertainty associated with vagueness, with imprecision, and/or with a lack of information … by its very nature, natural language is vague and imprecise … fuzzy logic seems to be most successful in processes when human reasoning, human perception, or human decision making are inextricably involved*". Jun Yan et al. [8] also propose that systems which require human/expert involvement and rely heavily on human or expert experience benefit more from the use of fuzzy modeling. User provided lexical characterizations of bookmarks, are mostly subjective and subject to user's perception of the node's contents. Having to deal with such characterizations, we found that a semantic modeling of Concepts based on and exploiting fuzzy set theory and fuzzy logic may constantly model human knowledge and provide for approximate reasoning (for a further discussion, see [9][10]).

Relationships represent user defined abstract relationships between Concepts. They are uni-directional, one-to-one and named. The direction defines the way the relationship should be interpreted, e.g. which Concept is the subject and which Concept is the object. The name provides a lexical description of the relationship, e.g. a human-interpreted indication of its meaning.

Rules are Multiple Input Single Output (MISO) Fuzzy Logic Controllers (FLCs). The input is Concept fuzzy values, while the output is a dynamically computed Concept value. A set of Concept values is dynamically computed upon rule's triggering. The dependency graph of the input/output Concepts of the Rules is restricted to be a Directed Acyclic Graph.

Weights: Fuzzy variables. Every Concept is connected with a Weight that provides its domain of definition.

By defining instances of the aforementioned elements, the user constructs a semantic net in the form of a Schema. A user may create more than one Schemata, each of them representing a specific knowledge domain.

To illustrate the notion of the Schema, let us consider the following scenario: a user searches the WWW for information about conferences. While viewing the contents of each node, he decides whether the node refers to conference or not. If a relevant to a conference node has been found, he scans for information about conference topics, dates, travel and registration costs, etc. Following the acquisition and understanding of this type of information, he combines it with his background knowledge, leading to subjective conclusions such as: "the conference is highly focused in hypermedia" or "the conference has a high registration cost".

In such a scenario, the user could have defined a Schema for storing his conclusions and characterizations about the nodes he has visited. An example of such a Schema is presented in Fig.3.

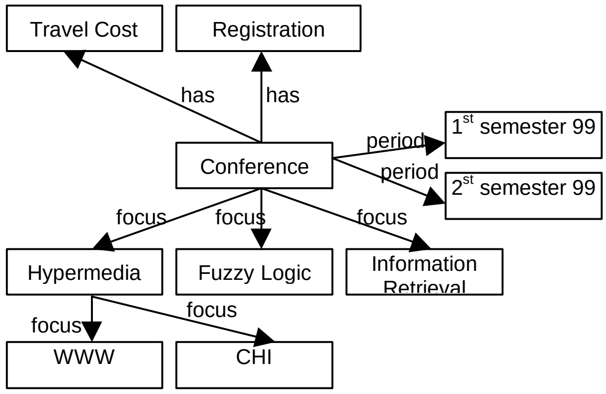

Fig. 3: A Schema example

This Schema definition assures the storage of propositions like "the conference has a registration cost", but also takes into account such lexical descriptions as "high cost". This is accomplished by making use of Weight elements. For example, we may consider a Weight instance named "cost", which provides lexical descriptions for the "Travel Cost" concept, as shown in Fig.4. Such a Weight instance is a fuzzy variable that maps discrete cost values (0 USD – 2000 USD) to lexical descriptions (Low, Medium, High, Very High).

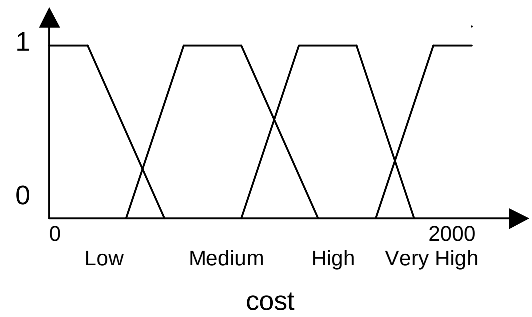

Fig. 4: Weight definition for "cost"

Suppose that given the previously described schema, the user creates a new Concept named "Worth Visit". He would like the system to automatically assign Concept values to newly added bookmarks under the assumption that *if* the bookmark is described as "Conference", it has a low registration cost, it focuses on fuzzy logic, and it is scheduled for the end of the 1st semester *then* the more worthly it is to participate in. This rationale is illustrated in Fig. 5, where "RULE1" infers the value of its single output by being fed from its three inputs.

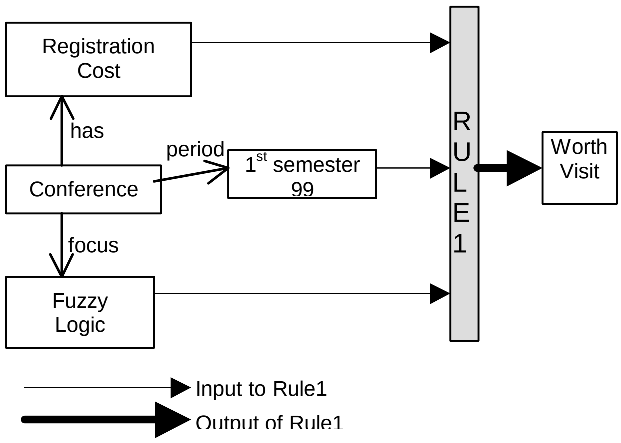

Fig. 5: A graphical example of a Rule

In our framework, we support the definition of Rules like "RULE1", using a format that imitates natural language and provides the rule-base definition of a MISO FLC. More than one sub-rules compose a Rule instance. For example, the following definition could be considered:

RULE1
SubRule1
    IF
    Conference.has.RegistrationCost = "LOW" AND
    Conference.focus.FuzzyLogic = "VERY" AND
    Conference.period.1stSemester = "BEGINNING"
    THEN
    WorthVisit = "VERY"
SubRule2
    IF
    Conference.has.RegistrationCost = "MEDIUM" AND
    Conference.focus.FuzzyLogic = "ENOUGH" AND
    Conference.period.1stSemester = "MIDDLE"
    THEN
    WorthVisit = "ENOUGH"

The advantages that arise from the introduction of Rules are summarized as follows:

- Rules ensure the assignment of Concept values to bookmarks without the interference of the user. This assignment adds both qualitative and quantitative meta-information: bookmarks automatically obtain new Concept descriptions along with the corresponding weight values.
- Due to the exploitation of MISO FLCs, the Rules are capable of firing even with incomplete input values. In turn this retains the workability of the Rules even if the user has only assigned a small number of input Concept values.
- The Rules provide dynamic meta-information that does not have to be permanently stored, thus reducing the total storage requirements.
- Any modification to a Rule does not alter already stored meta-information (apart from the Rule definition itself) and it is immediately activated upon Rule's triggering.

2.3 Assigning Concepts to bookmarks

Nodes characterization is a complex task, mostly subjective and therefore can not be fully automated. A frequently employed method for the automatic extraction of characterized groups (the equivalent to our Concepts), is clustering. It has to be noted that automatic clustering based on keyword vectors, which is mainly used in such approaches, can not be sufficient since:

- An intentional repetition of words might be employed by the author to bias indexing.
- The context of a word may drastically alter its semantics.
- Reader's background knowledge fills the gap between what he actually reads and what he comprehends. Therefore, some Concept might characterize a node, even if the lexical description of the Concept does not appear in the content of the node.

While taking into account the limitations of automatic clustering, this paper proposes three different scenarios for the assignment of a Concept to a bookmark. The selection of the appropriate scenario is based on a priority ordering which is explicitly defined in the next paragraphs:

Manual assignment: The user selects a Concept and inputs the weight. This type of assignment results in the creation of static (permanently stored) Concept values.

Automatic assignment: The system assigns the bookmark to the best matching Concepts, taking into account the cluster center of each Concept. This center calculation is based on distance metrics and keyword vectors. This type of assignment also results in the creation of static Concept values. As indicated in Table 1, for every bookmark, the contents of the corresponding node are keyword indexed and the relevant information is stored in the "Keywords" attribute of the bookmark. Consequently, every Concept is treated as a cluster containing all the bookmarks that it has been assigned to. In this study, we use the well known in the field of Information Retrieval TFxIDF model, to construct term vectors. The TFxIDF model has been used before [11][12] for the resemblance computation of node contents. Each bookmark is represented by a vector of T terms with corresponding term weights. The weight of the k-th term T_k of the i-th bookmark B_i is determined as equation (1) depicts:

$$W_{ik} = \frac{tf_{ik} \bullet \log\left(\frac{N}{n_k}\right)}{\sqrt{\sum_{j=1}^{t} (tf_{ik})^2 \bullet \log\left(\frac{N}{n_j}\right)^2}} \qquad (1)$$

where tf_{ik} is the frequency of term T_k in B_i, N is the total number of bookmarks, and finally n_k represents the number of documents containing term T_k. For efficiency, we use only the first 100 most significant (in terms of occurrences) terms to build the term vectors, and we denote this maximum as MAX_T. The exact process of vector creation is performed for the newly added bookmark. Then each Concept is considered to be a predefined cluster and the center of gravity vector *v* is calculated, applying the following equation (2) for the j-th element of the gravity vector (parameter j ranges from 1 to MAX_T):

$$v_j = \frac{\sum_{k=1}^{N} W_{kj} \bullet C_k{}^w}{\sum_{k=1}^{N} C_k{}^w} \qquad (2)$$

where W_{kj} is the weight for the term T_i of the B_k bookmark, C_k is the Concept value for the B_k bookmark and *w* is a user defined degree of fuzziness ($0<w<1$).

The next step includes the computation of the weight for each Concept that will be assigned to the newly added bookmark. Since the bookmark is required to be as near as possible to the center, we use the Euclidean distance as a measure of similarity between the bookmark and the Concept vectors. Equation (3) computes this distance:

$$D = \sqrt{\sum_{j=1}^{MAX_T} \left(v_j - W_j\right)^2} \qquad (3)$$

Finally the Concept value *C* for the newly added bookmark is computed as equation (4) shows:

$$C = 1 - D \qquad (4)$$

In order to avoid the assignment of many Concepts with low values, a threshold is set to 0.5 for *C* to be accepted.

Dynamic assignment: The Concept values are retrieved and the Rules that apply then fire. A set of dynamic Concept values is generated. This process might be repeated for the newly added dynamic Concepts, with one limitation: each Rule is restricted not to use its output Concept as input, either directly or indirectly.

In the case of conflicts between Concepts values which might be encountered (i.e. when a user provided value differs from a Rule deducted one for the same Concept), we apply a resolving policy. User's direct choices precede automatic and dynamic assignments, whilst automatic assignments supervene dynamic ones. This policy manifests a certain significance given to the meta-information provider: firstly a user has always a priority in deciding, secondly the user defined rules make the decision, and finally in the absence of the above, the system has to decide.

3 Visualizing the bookmark space

Having presented our approach with reference to bookmark modeling, this paper now addresses the issue of constructing a visualization of the acquired meta-information as an interface, which supports purposeful navigation within bookmark space.

3.1 3D visualization of complex data sets

Graphical representations are a good way of communicating relationships among objects ([13], p.63). Humans have evolved intrinsic skills for detecting visual patterns and for navigating in three-dimensional space. Visualizing information within a graphical, three-dimensional spatial context may exploit such skills. Moreover, it is understood ([14], [15]) that three-dimensional space allows for a more effective arrangement of large sets of informational objects and for visualizing more attributes for each of these objects, than two-dimensional space does. Accordingly, recent advances in VR technology [16] have suggested that interactive visualizations of large information sets in a three-dimensional context may increase the amount of information that people can meaningfully manage.

In the same vein, Fairchild ([13], p.46) suggests that the problem of managing large amounts of complex information may be decomposed to the following three subproblems:

- How to make meaningful visualizations of single objects.

- How to make meaningful visualizations of collections of objects.
- How to allow users to control the selection of the visualizations efficiently.

If these objects are placed into a three-dimensional display, as opposed to a two-dimensional display, the perceived complexity of the information is reduced ([13], p.51). He also goes on to suggest two significant interaction tasks that should be accomplished in such a visualization:

- a user should be able to navigate within the virtual space to dynamically define areas of interest,
- a user should be able to interactively control the positioning of the information objects in space, to move the viewing location dynamically in space and to modify the space according to predefined functions.

Navigating within a complex and extensive informational space, such as WWW, may lead a user into suffering disorientation or "lost in hyperspace" syndrome. For the purpose of decreasing the effect of such phenomena, several attempts have been made for visualizing the hypergraph structure of WWW. Visualizing such complex navigational spaces is a severe task. As relevant research ([3], [11], [14], [15], [17], [18]) suggests, at least three different types of requirements have to be considered:

a) Orientation/ Navigation Requirements: The user must know his current "position" within navigational space, perform more or less drastic "movements" within this space, be presented with an indication of how he reached his current location (intermediate nodes) and finally, have a visual representation of which are the permitted moves that he can proceed with.

b) Decision Support Requirements: The user must be supported in deciding his next move, i.e. which node to visit next. This implies that the user must be provided with meta-information, which describes each available node and may aid him in differentiating amongst these nodes. At the same time, the similarities amongst these nodes have to be presented.

c) Usability Requirements: the user must be able to adjust the space in terms of details shown as well as to "reposition" himself in the space by altering his current "viewpoint". The new "viewpoint" should highlight the visualized part of the space that seems to be of more of interest to the user's current information needs. It is essential that this "repositioning" is smoothly performed for the user to retain orientation with reference to his previous "viewpoint". Finally, the user should be easily acquainted with moving within this space after a short time of training and practice.

It is understood that certain requirements for visualizing large sets of semi-structured information, in general, are still relevant to the visualization of the WWW structure. Relationships amongst WWW nodes though, are explicitly defined by links and such links are being visualized in most cases of WWW structure visualizations (i.e. local maps, fish-eye views, hierarchical views etc.). A set of bookmarks, on the other hand, is a user-specific subset of the overall WWW structure, where each node can be accessed directly and where existing links between nodes become irrelevant. Therefore, links between WWW pages are not relevant when these pages become bookmarks and as a result this paper does not take them into account.

3.2 Related work

This section of the paper considers previous work done on the issue of visualizing information in a 3D spatial context. The relevant literature review has revealed three main approaches.

Firstly, research which generally deals with the issue of visualizing large sets of relatively static data. In some cases this research focuses on visualizing inter-object relationships in detail - the SemNet approach ([19] Fairchild, Poltrock & Furnas, 1988) – or attempts to deal with both inter-object and intra-object relationships - the SGI 3-D Fusion information landscape prototype. Other attempts differ in terms of the way that the visualized data are structured. The Xerox PARC cone tree and perspective wall systems [16] and [20] aim at visualizing hierarchically and linearly structured data respectively. Benford et al. [21] have described systems which support collaborative browsing of information and which correspond to differently structured data sets - well-structured and ordered sets, less well structured without known interrelationships and hypermedia structures.

Secondly, Benford et al. [21] have dealt with data sets which are dynamically re-arranged in 3D space. VR-VIBE in particular supports the positioning of information objects in three-dimensional space according to their relevance to specific concepts called "Points of Interest". Similarly, Fairchild et al. [22] have attempted to arrange objects on the surface of a sphere according to their relationships with a certain object of interest with the VizNet system.

Finally, a series of attempts have been made for visualizing hyperstructures of data. Dieberger & Tromp [23] have proposed a model for an "information city" metaphor for a three-dimensional interface to navigating hypertext structures. Andrews [24] has developed the Harmony client for the Hyper-G data model, as a 3D structure map or an interactive three-dimensional visualization of the data structure which may be both hierarchical and hyperlinked. Ingram and Benford [25] have developed a system called LEADS which improves the legibility of information spaces by automatically creating or enhancing certain legibility features within pre-existing visualized information sets. Finally, Snowdon et al. [26] have developed a novel browser called WWW3D, which integrates a representation of all web documents that the user has browsed and history information about the links that the user has visited in the past, within a single 3D display of a part of the web's structure.

Following the relevant literature review, it is hypothesized that it would be preferable to develop a three-dimensional visualization of the rich bookmarks meta-information, for the purpose of supporting user navigation within these bookmarks. To achieve that, an effective way of visualizing each type of meta-information as well as groups of relevant types had to be identified. Two general approaches, in terms of arranging the elements of the visualization in 3D space, were previously detected: static and dynamic arrangement. While acknowledging the difficulties involved in constructing and perceiving dynamically created views, we follow the latter approach, because it provides a user-centric visualization, well adapted to ad-hoc informational needs.

In effect, a three-step procedure is proposed:

1. Firstly, the user poses certain selection criteria based on the semantics of bookmarks, in the form of a selection query.
2. Secondly, the system parses the query and retrieves all bookmarks that partially satisfy the query, constructing an ordered list.
3. Finally, the list order and the meta-information, which describes bookmarks, determine the visual form of each bookmark and the way it is positioned within the 3D space.

The development of the system, based on the VRML 2.0 standard and Java, is in progress. Meanwhile, a prototype of such a system has been developed by making use of Sense8's WorldUp® for simulating the proposed solution.

3.3 Posing selection criteria

A user can pose his preferences defining a SELECT-like query (i.e. a list of Concepts along with their lexical descriptions) that better describes his current interests. 3D space is then shaped in order to "highlight" those bookmarks that better match (e.g. have a greater weight value) for the specified concepts. A visual interface for query formulation is under construction.

A Query example follows:

```
SELECT
Conference.focus.Hypermedia = "very"
AND Hypermedia.focus.WWW =
"very" AND
Conference.has.TravelCost =
"Medium"
```

The Query selects only those bookmarks that have been characterized with the concepts "Conferences", "Hypermedia", "WWW" and "TravelCost". For every bookmark that has been selected, an overall similarity with the selection criteria is computed. This similarity is the result of the computation of the substituted by weight values expression, according to fuzzy logical operators (AND, OR, NOT). A weighted ordered list of bookmarks, which feeds the visualization component for further rendering, is then produced.

Alternatively, the user may select a certain bookmark as representative of his current interests. The minimum distance of all bookmarks that share common Concepts is then computed in order to discover most similar bookmarks. In this case also, a weighted ordered list of bookmarks is produced.

3.4 Description of the 3D visualization

As suggested in 3.1, since a set of bookmarks corresponds to pages which have been manually located within the user-defined Conceptual space, it is no longer significant whether these pages are physically linked or not. As a result, the proposed visualization exploits the meta-information, which has been either explicitly of implicitly derived by user actions and chooses not to render links. In contrast with the information relating to links, which is spontaneous and mostly of significance to the author, meta-information is rich, extensive and user defined. Each web page is represented by an orthogonal parallelepiped box. It is significant to take advantage of perspective as a spatial quality, which may aid users in differentiating between boxes positioned in different depths in relation to the user's viewpoint at each time. Due to difficulties regarding the appreciation of the scale of objects in virtual environments ([28] Charitos, 1998), it is proposed that all boxes should be of the same size; as size degrades so distance from the user's viewpoint increases. [Fig. 6]

By utilizing the ordered list, each bookmark is positioned in 3D space according to its relevance to the specific query. The environment extends over a "floor", which is represented by a horizontal plane at a 0 y-coordinate. The user viewpoint at the moment that the simulation is generated is positioned at the base of a vertical pole – at a (0,0,0) point in 3D space - which extends upwards to the maximum height of the environment. Visualized bookmarks are arranged around the user's viewpoint, their distance from the pole being inversely proportional to the weight of the corresponding list item.

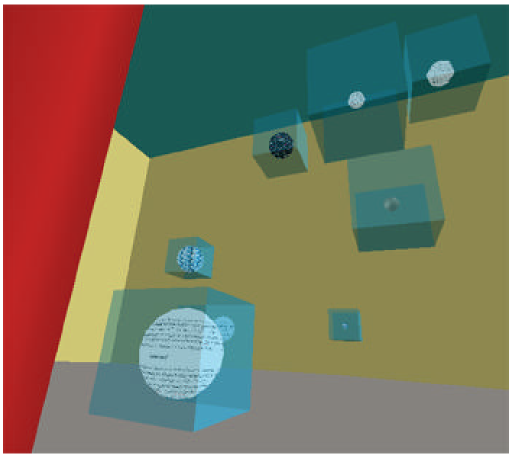

Fig. 6: View of a small selection of bookmarks from the base of the pole

The content of each page is represented by a sphere, which is concentrically placed inside each box. [Fig.7]. The size of the sphere indicates the size of the page file in bytes (Size attribute) and the types of content that the page includes (Profile attribute) are represented by certain images texture-mapped onto the sphere. Meanwhile, the transparency of the box's material corresponds to how recently this page was visited by the user (LastVisit attribute): the more transparent the box is, the more recently the page had been visited. The use of different rates of transparency on the box's material aims at utilizing Type-1 meta-information for visualizing Type-2 meta-information. Intuitively, this means that volatile and time-varied meta-information (size of page, content type) is more clearly visible when the user has recently visited a page (time of last visit).

The height at which each visualized bookmark is positioned, in relation to the "floor" (value of the y-coordinate) represents past experience gained regarding the destination of the link, in terms of difficulty with which the contents of a certain web page may be retrieved. In order to transform the selected values of the transfer mean-time (TMeanTime) and host availability (Availability) meta-information to height coordinates, we employ the following equations (5) to (8):

$$T_i = \frac{\sum_{j=1}^{N} t_j}{N} \qquad (5) \qquad\qquad T_M = \frac{T_i}{T_{\max}} \qquad (6)$$

In equations (5) and (6), T_i is the transfer mean-time for the i-th node that has been successfully retrieved N times, and T_M is the normalized mean-time, taking into account the maximum transfer mean-time for all nodes T_{max}. The T_m metric's domain is [0,1]. Analogously, we compute the mean-value of the satisfied URL requests A_M, as equation (7) denotes:

$$A_M = \frac{\#\,\text{Serviced Requests per bookmark}}{\#\,\text{Total Requests per bookmark}} \qquad (7)$$

where the domain of values is also [0,1]. Finally we calculate the height using equation (8):

$$H = T_M \bullet C_T + A_M \bullet C_A \qquad (8)$$

where the constants C_T and C_A express weight for transferability and availability respectively. Given that these constants have been defined, the roof of the space (e.g. maximum height) is the sum of these constants.

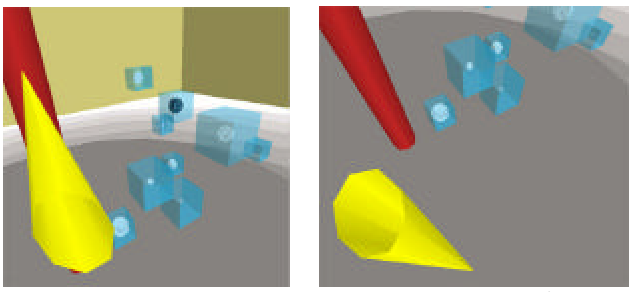

Fig.7: Cylindrical "ray object" emanating from user's viewpoint which indicates the position of the user.

For the purpose of enhancing user's orientation within the 3D space, a "ray object" was designed for indicating the position of the user at all times. This ray emanates from the user's viewpoint and leaves a trace on the central pole as well as on the "floor".

4 Conclusions and future work

The semantic framework described in this paper has several distinct features for modeling and visualizing WWW bookmark spaces. A number of conclusions are drawn based on these features:

- The Semantic model provides for capturing both the explicitly and implicitly obtained semantics of the bookmarked nodes. Due to the fact that it is graph-based, it also provides a great flexibility for constructing simple or more complex models, according to user's preferences and experience.
- The partial assignment of Concept values to bookmarks enriches the expressive capabilities of the model and also ensures the capturing of human ambiguity.
- The framework supports three different types for conceptually characterizing bookmarks: user-provided, based on user's preferences expressed in Rules, and system-provided. To the best of our knowledge, no other attempt has succeeded to support both of these types in a single, flexible and coherent framework.
- A 3D visualization metaphor has been constructed on top of the Semantic model, which efficiently renders the elicited knowledge. The dynamic arrangement of the bookmarks in the 3D space results from user-defined selection criteria, thus providing a navigational "space", well adapted to current informational needs

Initial experimental results have demonstrated that the proposed visualization conforms with the prescribed visualization requirements for WWW bookmark visualization, as these have been stated in this paper. Implementation of a fully functioning visualization is under way and experimental evaluation of this system's effectiveness is scheduled for the immediate future.

References

[1] Walker, J., Document Examiner: Delivery interface for hypertext documents, *ACM Hypertext '97 Conference*, Chaper Hill, NV, USA, November 13-15, 1987, p.p. 307-323.

[2] Nielsen, Jacob, *Multimedia and Hypertext: The Internet and Beyond*, AP Professional, 1995.

[3] Maarek, Y., Shaul, I., Automatically Organizing Bookmarks per Contents, *WWW5 Conference*, Paris, France, May 6-10 1996, p.p. 1321-1333.

[4] McDonald, Sharon, Stevenson, Rosemary, Navigation in hyperspace: An evaluation of the effects of navigational tools and subject matter expertise on browsing and information retrieval in hypertext, Elsevier, *Interacting with Computers*, Vol. 10, 1998, p.p. 129-142.

[5] Stillings, Neil, et al., *Cognitive Science: An Introduction*, MIT Press, 1995.

[6] Luger, George, et al., *Cognitive Science: The Science of Intelligent Systems*, Academic Press, 1994.

[7] Ross, Timothy, *Fuzzy Logic with Engineering Applications*, McGraw Hill, 1995.

[8] Yan, J., Ryan, M., Power, J., *Using Fuzzy Logic Towards intelligent systems*, Prentice Hall International, 1994.

[9] Petrou, C., Martakos, D., Hadjiefthymiades, S., Hypermedia Linking and Querying: A Fuzzy Based Approach, *IEEE International Conference of Systems, Man, and Cybernetics (SMC'98)*, San Diego CA, USA, October 11-14, 1998, p.p. 1235-1240.

[10] Petrou, C., Martakos, D., Hadjiefthymiades, S., Adding semantics to Hypermedia towards Link's enhancement and Dynamic Linking, *Hypertext-Information Retrieval-Multimedia '97 Conference (HIM'97)*, Dortmund, Germany, 1997, p.p. 175-190.

[11] Chen, Chamomei, Structuring and Visualizing the WWW by Generalized Similarity Analysis, *Hypertext 97 Conference*, Southampton UK, 1997, p.p. 177-186.

[12] Chang, Chia-Hui, Hsu, Ching-Chi, Customizable Multi-engine Search Tool with Clustering, *WWW6 Conference*, Santa Clara, CA, USA, April 7-11, 1997.

[13] Fairchild, Kim M., Information Management Using Virtual Reality-Based Visualizations, *in Virtual Reality Applications and Explorations*, Wexelblat, A. (ed.), Academic Press Professional, London, 1993.

[14] Fowler, R., et al., 3D Visualization of WWW Semantic Content for Browsing and Query Formulation, *WebNet 96 Conference*, San Francisco, CA, USA, October 15-19, 1996, http://aace.virginia.edu/aace/conf/webnet/html/205.htm

[15] Mukherjea, S., Hara, Y., Focus+Context Views of World-Wide Web Nodes, *ACM Hypertext '97 Conference*, Chaper Hill, NV, USA, November 13-15, 1987, p.p. 187-196.

[16] Robertson, G.G., Mackinlay, J.D. and Card, S.K., Cone Trees: Animated 3D Visualizations of Hierarchical Information, in *CHI '91 Proceedings: Human Factors in Computing*

Systems, New York: ACM Press, pp.189-194, 1991.

[17] Carriere, J., Kazman, R., WebQuery: Searching and Visualizing the Web through Connectivity, *WWW6 Conference*, Santa Clara, CA, USA, April 7-11, 1997.

[18] Huang, M., et al., Dynamic Web navigation with information filtering and animated visual display, *Asia Pacific Web Conference (APWeb98)*, Beijing, China, September 27-30, 1998, p.p. 63-71.

[19] Fairchild, K.M., Poltrock, S.E. and Furnas, G.W., SemNet: Three-Dimensional Graphic Representations of Large Knowledge Spaces, in Cognitive Science and Its Applications for Human-Computer Interaction, Fairlawn, NJ: Lawrence Erlbaum Associates, 1988.

[20] Mackinlay, J.D., Robertson G.G. and Card, S.K., The Perspective Wall: Detail and Context Smoothly Integrated, in CHI '91 Proceedings: Human Factors in Computing Systems, New York: ACM Press, pp.173-179, 1991.

[21] Benford, S., Bowers, J., Fahlen, L.E., Greenhalgh, C., Mariani, J. and Rodden, T., Networked Virtual Reality and Cooperative Work, *Presence: Teleoperators and Virtual Environments*, Vol.4, Nr.4, MIT Press, 1995.

[22] Fairchild, K.M., Serra, L., Ng, H., Lee, B.H., and Ang, T.L., Dynamic Fisheye Information Visualizations, 1st British Computer Society Conference on Virtual Reality, 1992.

[23] Dieberger, A., and Tromp, J., The Information City Project – A Virtual Reality User Interface for Navigation in Information Spaces, Proceedings of the Virtual Reality Vienna '93, December 1993. http://www.lcc.gatech.edu/~dieberger/VRV.html

[24] Andrews, K., Visualising Cyberspace: Information Visualisation in the Harmony Internet Browser, Proceedings of the 1st IEEE Symposium on Information Visualisation, Atlanta GA, pp.97-104, Oct. 1995.

[25] Ingram, R. and Benford, S., Improving the Legibility of Virtual Environments, in Proceedings of the 2nd Eurographics Workshop on Virtual Environments, Monte Carlo, 1995.

[26] Snowdon, D., Benford, S., Greenhalgh, C., Ingram, R., Brown, C., Lloyd, D., Fahlen, L. and Stenius, M., A 3D Collaborative Virtual Environment for Web Browsing, in Virtual Reality Universe '97, Westin Santa Clara Hotel, California, USA, April 1997. http://www.crg.cs.nott.ac.uk/research/publications/

[27] Charitos, D., The Architectural Aspect of Designing Space in Virtual Environments, PhD Thesis submitted at the University of Strathclyde, Faculty of Engineering, 1998.