Joining Cognitive Capabilities in 3D Media Spaces

Maya Dimitrova, Claude St-Jacques

Abstract: *In design of ergonomic interface it is crucial to make the most of abstract representations to convey various aspects of the semantic content. We are developing interface based on joining users' visual computing skills to the soft computing capability of fuzzy logic to model the proximity relationship of the document corpus onto 3D visual representations. The paper presents the main ideas behind the proposed interface and the simulations to achieve immersion effect in real 3D semantic interaction.*
Key words: *Ergonomic Interface, Soft Computing, Fuzzy Logic, Immersion Effect.*

INTRODUCTION

Abstract representation of semantic information is a new and promising trend in Web search tool design. We are developing interface based on joining users' visual computing skills to the soft computing capability of fuzzy logic to model the proximity relationship of the document corpus onto 3D visual representations. This portion of research is part of the Teleweb project which aims to provide a friendly online access to large multimedia databases – texts, images, soundtracks, videos and animations [1], [2]. The feeding and maintenance of this multi-formatted and distributed media space rests on the voluntary collaboration of multiple people and networked/"Webbed" organizations. All the files are binary data, although each of them has its own structure: HTML, MPEG, QuickTime Flash, etc. Being rich and varied, the media space presents its content in an under-determined, or *seemingly* "chaotic" way that makes the retrieval of media files on a specific topic as problematic as the structuring of the hypermedia itself.

It is very difficult for someone to access optimally the media space via merely an Internet connection, a Web browser, "punching in" one or more words into some search engine. For example, Google returns 4,930,000 hits for "autism" and returns 3,900,000 hits to the word "cognition" and the only way to scan them is by going through the list, one by one. Moreover, the unsolicited hits constitute some form of "noise" that can be reduced in these search engines by adding words to the query. The problem has been addressed, for a long time, in the context of document indexing for libraries by adding keywords that better reflect its content and facilitate its retrieval. However, the quantity and the diversity of the available files on the Internet make these centralized methods of indexing unfeasible because this task is highly specialized and relies on individuals with their own background knowledge and expertise. In regards to this problem we focus on two main aspects in the Teleweb project: the cognitive ergonomics of presenting the whole set of documents instantly and the logical and computational aspects of offering a dynamic and automatic classification process.

To enhance cognitive synthesis in our experiment we have replaced the lists with 3D representations. On the logical and computational side the attempts to use a taxonomic approach with trees, thesauri, etc., or a linguistic approach to a set of primitives and generation rules are put aside to focus on a dynamic and *a posteriori* automatic classification in the soft computing environment of the fuzzy logic [3]. Initially, all hits resulting from a query were mapped onto hilly ground by taking their proximities and groupings into account [1]. This time, we develop Java 3D medium where clusters are mapped onto grids and sites on nodes with the ability to rotate the grid in space and obtain full view from any perspective. With a fuzzy logic classification engine coupled to a 3D representation of its semantic clustering by means of a cybernetic loop, Teleweb will provide an intuitive meaningful interface to the user.

CYBERNETICS OF WEB SEMANTICS

From a cybernetics point of view, we have started from the information given in natural language and build a system in three steps to make the selection of the descriptions available by regrouping them by granularity. Following Wiener [4], who called the part of the system that controls "the loss of meaning from language" "cybernetics of semantics", we explore three perceptual invariants: *resemblance*, *difference* and *data-space*. Invariance is understood as a perceptual mechanism *always* present as a human cognitive process [2].

In the next sections each construction step is described that is used to transcend semantic information into the three regulating principles emerging from the granularity and the similarity of visual units within their perceptual *invariance* in natural language. Section 1 is devoted to the first regularity that can transcend as a perceptual invariant in the semantics of these descriptors, i.e. the degree of resemblance [5] among documents sharing common features. The resemblance is explored by looking at the fuzzy similarity of documents in vector space representation. In section 2 the second perceptual invariant is dealt with that transcends from the semantic information of the visual units. That is, grouping entails the granularity of descriptions for objects in the same group. This is called the difference between the resemblances of what can produce an unbundling amongst associated objects when some semantic distinctions appear in the mind of the user. Fuzzy cluster analysis is used to reproduce the data space in a similar group by their differences. The third regulating principle or perceptual invariant, used for the semantics emerging from the description of the visual units, is a topological representation sensitive to cognitive ergonomics, i.e. the data-space [2]. In section 3 the "cybernetic loop" is illustrated by a 3D representation of these semantics in an immersive environment. The user is immersed into a grid 3D world where each level represents regrouping of visual units of the same topic/style. Exploring the brightly colored "Smarties" in the grid the user discovers by the dynamic action of clicking on them that the description of a unit is more prototypal when it is near the center of the grid level and less when it is at its outskirts. The conclusion section makes comments on this experimental design and the prospective development of the Teleweb project.

1. Document Resemblance

For the experiments the index term "autisme" is used to query the database of the University of Quebec Montreal library. The method used to detect the resemblance between the analytical entries - title and index terms - is borrowed from fuzzy information retrieval [6], where segments of text are represented in vector space. The set of terms given by the title, subjects, and a list of index terms is the vector for each analytical entry. According to the method, informative weight is allotted initially to each term based on its frequency of occurrence in one document with doubled weight if the term is in the title. A matrix (A) is built from the vectors of weighted terms with the elements in each entry in the matrix.

The next step is to find the granularity expressed by the description of a document by searching for "resembling" objects in a vector space. According to [6] this cannot be directly accessible by a single word, but the abstraction or the idea of a concept is retrievable from words co-occurring on a regular basis in natural language. A "fuzzy association of terms" is called the most abstract definition of a concept, their connection expressed by the similarity relation. A pseudo-thesaurus of related terms is built for the matrix of documents by computing their fuzzy similarities with the formula:

$$R(t_i, t_j) = \frac{\sum_k^M \min(t_{ik}, t_{jk})}{\sum_k^M \max(t_{ik}, t_{jk})}. \quad (1)$$

The relation *R* of the fuzzy similarity connecting terms t_i and t_j is given by the *min* value between their weights and among all of their co-occurrences starting from the first document k until the last entry of matrix M, divided by the *max* value of the same kind of co-occurrences across the matrix M. To avoid problems of "mismatch" of user and database designer vocabularies [7], the query is augmented with a pseudo-thesaurus [8]. The augmentation process of a query Q is initiated by the user's choice of terms that are projected automatically onto the set of related terms in the pseudo-thesaurus P, hence the operation of composition of relations *max-min* is used. The augmented query B is given by the composed relation o,

$$B = Q \circ P. \quad (2)$$

If the word "autisme" is used as a query, the obtained from the composition with the pseudo-thesaurus P the augmented query B = [autisme (1), enfant (0.036), aspect (0.039), cognitive (0.004), trouble (0.034)…]. The augmented query B is applied onto the matrix A of documents by composing their relation, once again, to obtain the results (C = $B \circ A$) from the *max-min* composition.

Table 1. Matrix C resulting from the query "autisme"

Doc. #	Analytical Entry	μ_C
11	Psychose, autisme et défaillance cognitive chez l'enfant. Aspect clinique; Autisme; Enfant autistique; …	0.5
270	Cognition et information en société. Aspect social; Connaissance; Politique gouvernementale; …	0.25
60	Élaboration et validation auprès de femmes d'un questionnaire clinique d'éléments cognitifs individuels…	0.2
…	…	…

Table 1 shows a sample of the matrix C obtained from the corpus of 300 documents from the library database. At the end of each analytical entry the membership degree μ is given of the set C of relevant results to the query. By applying an augmented query to the corpus of documents it is evident that it is possible to connect the "cutting" of a subset of a database with the fuzzy similarity between the analytical entries.

2. Differences in Resemblance

An essential ergonomic component of user perception of abstract semantic relations is the level of granularity grounded in natural language and reinstated by the automatic processing of (fuzzy) linguistic features. The granularity of a document can be "grasped" only by its partition according to its features in natural language. Despite the fact that the underlying media space has rather unstructured, "chaotic" nature, the visual regrouping of the documents are produced by the common features shared with other documents. In terms of the fuzzy similarity relation from equation 2, the result of a query has a minimum amount of resemblance and the necessary next step is to make cluster analysis of the documents by looking at the differences between their resemblances. Since the entropic tendency of natural language is the "major opponent to communication" [6], it is necessary to account for the uncertainty in the grouping of the documents. The method used for fuzzy cluster analysis from [10], [11] is based on the following objective function:

$$J_m(U, v_{ik}) = \sum_{w=1}^{j} \sum_{c=1}^{i} (\mu_{kq}(\omega_{jq}))^m (d_{kq})^2 \ . \qquad (3)$$

The objective function J_m is given by the optimization of the membership degrees to the fuzzy sets U in respect to the prototypal vectors v_{ik} in such a way that the sum ($\sum$) of the square distance errors $(d_{kq})^2$ to the clusters $c = \{1,...,i\}$ and the sum ($\sum$) of the membership degrees μ_{kq} are optimal. The partitioning of the data-space is achieved by calculating iteratively the Euclidean distance between the vectors of documents and the prototypal vectors for each set of the document groups. This distance is given by the following formula:

$$d_{kq} = d(\omega_{jq}, v_{ik}) = \|\omega_{jq} - v_{ik}\| = \left[\sum_{i=1}^{c} (\omega_{jq} - v_{ik}))^2 \right]^{1/2}, \qquad (4)$$

where $q = \{q_1,...,q_n\}$ is the index of each of the descriptors in the analytical entry ω_i and k = $\{k_1,...,k_m\}$ is the index for each of the prototypal descriptors of the prototype v_i of a cluster. The optimization of the objective function is obtained by reducing the distance (4) to its smallest value (e.g. threshold value) after several cycles or after some fixed number of iterations. Details of membership to a cluster calculation and repositioning of prototypes after each cycle are given in [9], [10].

The partitioning on the data space made by the query "autisme" regroups similar documents into 5 clusters following their granularity. Considering that the subject of our corpus is "cognition" and the topic is "autisme", the automatic fuzzy cluster analysis of the differences between resembling documents generates five clusters of under-determined subtopics: 1) The social aspect of cognition; 2) Cognitive disorder; 3) Children's knowledge and thinking; 4) Cognitive process of learning; 5) Children's mental process of thinking. By applying the granularity assumption, this rule of the "cybernetics of semantics" governs the control of the "loss" of what is meant by difference of resembling documents.

3. Sensitive Representation

So far, the representation of documents is multi-dimensional, whereas the topology principle of representation requires three dimensions. In [10] dimensionality reduction from n to 2 dimensions is achieved by borrowing a strategy from fuzzy controller's theory, where the rules of operation of the system are expressed in natural language. A fuzzy controller rule can be seen as a function with several premises entailing only one conclusion: $f(x_1, x_2,..., x_n) = y$. The method leads to the reduction of a multidimensional rule in a function with one argument by taking the minimum value of the premises: $y = \min|f(x_1, x_2,..., x_n|$ [11]. Given that our visual representation is a regrouping of several multidimensional vectors representing a certain number of documents in one cluster, then the conclusion y is obtained by combining the functions of all dimensions:

$$y = \max[\min f_1(x_1, x_2,..., x_n),..., \min f_m(x_1, x_2,..., x_n)] \, . \qquad (5)$$

The conclusion y is derived from the maximum value of the minima calculated for each function. For the 3D representation to infer the rule that governs the "granularity" of connecting objects in the media space is similar to finding the rule of operation for a system in a given state. To infer automatically classification rules from linguistic knowledge NEFCLASS-J is used [12], [13]. It employs a fuzzy neural feed-forward network with the first layer formed by the input variables $X = \{x_1, x_2, ..., x_n\}$. The second layer stands for the fuzzy rules and the third are the output classes obtained by the classifier. NEFCLASS is trained via its supervised learning algorithm to adapt iteratively the fuzzy sets to the already known clusters. The network is fed with the input variables from matrix C. The number of rule nodes is fixed to the number of prototypes discovered by the cluster analysis and each output unit is trained on the classes of documents. In this way the x value of the function $f(x)$ reduced to one argument is obtained.

The topology of a sensitive visual representation of a media space relies on mapping of quantitative information, obtained from the fuzzy c-mean algorithm [14] and the fuzzy classifier [13] via the method of dimensionality reduction from [10]. The z coordinates of the grid visualization are given by the membership degree of a document to a cluster, which is the height above the ground level. To compute the coordinates of each document a bell-shaped function is used:

$$\mu_C(x) = e^{-\left(\frac{\|x_i - v_k\|}{\sigma_k}\right)^{\alpha}}, \qquad (6)$$

where $\|x_i - v_k\|$ is the Euclidean distance between x and the prototype v in the same dimension and σ and α are the parameters of the standard deviation and the mean of the curve, respectively [10]. The curve shape, generated by equation (6), serves as a template for the visualization interface (Figure 1).

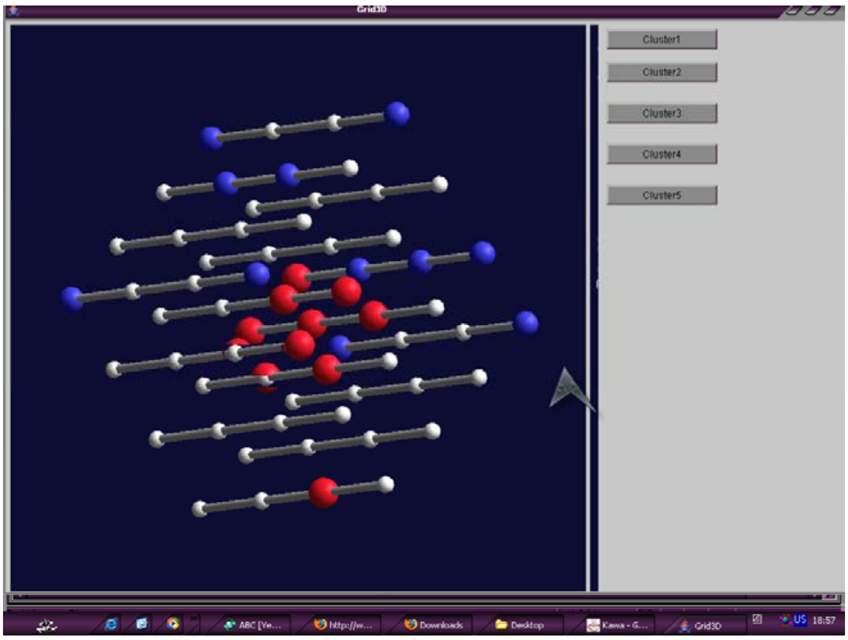

Fig. 1. 3D cluster view of the media space about "autisme"; the "smarties" represent grouping of Urls that can be viewed by mouse-clicking

RELATED AND FUTURE WORK

This research is related to numerous studies of optimising search engine performance, to information retrieval, fuzzy classification and semantic visualization [e.g.

15]. It is also related to studies on automatic Web document classification according to genre and 2D graphical representation [16, 17]. An example of "live" 2D Web visualization via Java-servlet search and classification engine and graphical applet invocation is studied in [17], [18]. The servlet plots along orthogonal dimensions in 2D view the retrieved and classified documents in perceptually salient way, e.g. upper-left boxes visualize brief-technical documents, upper-right boxes – extended-expert, lower coordinates denote popular brief vs. extended documents, respectively. The user can view the document by clicking on the box denoting the Url of the retrieved document. It was experimentally shown that users perceive different document features as independent, so a multifaceted (3D) representation along various *orthogonal* dimensions is ergonomically justified, even enjoyable. With the present research further exploration of "live" ergonomic *abstract* and *immersive* 3D interface is pursued.

Self-organizing maps (SOM) have been extensively applied to Web search and document retrieval and for search result visualization [19]. They combine the features of mathematical simplicity with computational slow-down due to the number of iterations heavily dependent on the perceptual clutter of the immediate visualization tests. Our approach is better in simplifying the representation *before* the visualization.

The problem of mathematical complexity with using NEFCLASS-J is common with many existing tools and techniques [e.g. 20] as well as to many 3D visualization instruments [21]. The data can be sampled by interval of membership to a cluster *before* sending it to NEFCLASS and just give all the information or the data when the user has chosen his area of interest. The duration of NEFCLASS-J computation will be matter of optimisation in future work. The time of computation of few hundreds of vectors having high-dimensional feature-sets already exceeds few seconds. However, the problem is partly overcome by performing an adaptive reduction of the data space by applying an unsupervised dimensionality reduction similar to [22] algorithm of concept indexing to eliminate the most irrelevant dimension of the data space. The similarity between documents being known at the first stage of the computation with the matrix of resembling objects can be added and average each column of this matrix to eliminate the irrelevant dimension of the centroid vector. A lower-dimension subspace is obtained, thus reducing the chance of being trapped in local minima at the clustering step and accelerating NEFCLASS computation. By looking closely at the data and the algorithms *granularity*-driven heuristics are derived. These can be directly fed into present-day Java 3D visualization constructs designed for "*live*" immersive interactive environments [23].

CONCLUSIONS

The paper presented the ongoing research and current results on investigating abstract representations to convey aspects of the semantic content. The main ideas behind the proposed interface are to relate perception and (fuzzy) logic to achieve immersion effect in real 3D semantic interaction between the user and large heterogeneous unstructured databases. A meta-level user-profiling feature "granularity and topology" is investigated in simulations and 3D visualization. The results show that this is a meaningful step towards new cognitive ergonomics on the Web.

REFERENCES

[1] St-Jacques, C., L.-C. Paquin, M. Lavallée. Dynamic Navigation into 3D Media Spaces. Proc. 2nd IEEE Conference on Intelligent Systems, Varna, Bulgaria, Vol. 2, 2004, 468-473.

[2] St-Jacques, C., L.-C. Paquin. Dynamic Visualization of Information: From Database to Dataspace. Journal of Advanced Computational Intelligence and Intelligent Informatics, 9, (1), 2005, 1-8.

[3] Kosko, B. Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence. Prentice Hall, 1992.

[4] Wiener, N. The Human Use of Human Beings. Boston, MA: Houghton Mifflin, 1950, p.93.

[5] Rosh, E., C.B. Mervis. Family Resemblances: Studies in the Internal Structure of Categories. Cognitive Psychology, 7, 1975, 573-605.

[6] Miyamoto, S. Fuzzy Sets in Information Retrieval and Cluster Analysis. Dordrecht, Netherlands: Kluwer Academic Publishers, 1990.

[7] Kovalski, G. Information Retrieval Systems: Theory and Implementation. Norwell MA: Kluwer Academic Publisher, 1997.

[8] Klir, G.J., B. Yuan. Fuzzy Sets and Fuzzy Logic. Upper Saddle River, NJ: Prentice Hall, 1995.

[9] Bezdek, J.C. Pattern Recognition with Fuzzy Objective Function Algorithms. New York, NY: Plenum Press, 1981.

[10] Höppner, F., F. Klawonn, R. Kruse, T. Runkler. Fuzzy Cluster Analysis. Chichester, England: John Wiley & Son, 1999.

[11] Mandani, E.H., S. Assilian. An Experiment in Linguistic Synthesis with Fuzzy Logic Controller. International Journal of Man-Machine Studies, 7, 1975, 1-13.

[12] Nauck, D., F. Klawonn, R. Kruse. Foundations of Neuro-Fuzzy Systems. Chichester, England: John Wiley & Sons, 1997.

[13] Nauk, U. Design and Implementation of Neuro-Fuzzy Data Analysis Tool in Java. Diplomarbeit, Technische Universität Braunschweig, 1999.

[14] Höppner, F. Fuzzy Clustering Algorithms – A Tool Library: User's manual. 2000.

[15] www.kartoo.com

[16] INFOMINE, http://infomine.ucr.edu/?view=projects/planning/genresid.html, 2004.

[17] Dimitrova, M., N. Kushmerick. Dimensions of Web Genre. Poster at WWW 2003 Conference, Budapest, Hungary, http://www2003.org/cdrom/papers/poster/p143/p143-dimitrova.htm, 2003.

[18] Dimitrova, M., N.Kushmerick, P. Radeva, J.J. Villanueva. User Assessment of a Visual Web Genre Classifier. Proc. 3rd IASTED International Conference on Visualization, Imaging and Image Processing (VIIP), Benalmadena, Spain, 2003, 886-889.

[19] Brittle, J., C. Boldyreff. GENISOM: Self-Organizing Maps Applied in Visualising Large Software Collections. Proc. 2nd International Workshop on Visualizing Software for Understanding and Analysis (VISSOFT), Amsterdam, Netherlands, 2003, 60-61.

[20] Weka Software, http://www.cs.waikato.ac.nz/~ml/weka

[21] Rilling, J., J. Wang, S.P. Mudur. MetaViz – Issues in Software Visualizing Beyond 3D. Proc. 2nd International Workshop on Visualizing Software for Understanding and Analysis (VISSOFT), Amsterdam, Netherlands, 2003, 92-97.

[22] Kaypis, G., E.-H. Han. Concept Indexing: A Fast Dimensionality Reduction Algorithm with Applications to Document Retrieval and Categorization. Proc. 9th International Conference on Information and Knowledge Management (CIKM), 2000.

[23] http://java.sun.com/products/java-media/3D/

ABOUT THE AUTHORS

Research Fellow Maya Dimitrova, PhD, Institute of Control and System Research, Bulgarian Academy of Sciences, Phone: +359 2 979 2054, E-mail: dimitrova@icsr.bas.bg.

Part-Time Professor Claude St-Jacques, University of Ottawa, Member of the Interactive Language Technologies Group, Language Technologies Research Centre, NRC Canada, Phone: +181 9 934 3904, E-mail: Claude.St-Jacques@cnrc-nrc.gc.ca.