Towards the Development of Environments for Designing Visualisation Support for Visual Data Mining

Simeon J. Simoff

Faculty of Information Technology
University of Technology, Sydney
NSW 2007, Australia
simeon@it.uts.edu.au

Abstract. The design of consistent information visualisation models is a key component in the development of visual data mining methods. However, it is a challenging activity to find the methods, techniques and corresponding tools that are suitable for specific visual mining task, or a particular type of data. The comparison of visualisation techniques across different designs is not a trivial problem. This paper discusses the issues connected with the development of consistent approach to formal development, evaluation and comparison of visualisation methods. Proposed formal approach is illustrated with examples of development of a visualisation model for data from the area of team collaboration in virtual environments and evaluation of visualisation schemes for the results of text analysis. The papers concludes with the discussion of the limitations of the proposed approach and future directions of the research and development of proposed approach.

Keywords: data visualisation, visual data mining, metaphor representation, virtual environments, formal methods, affective computing

1. Handcrafting and creativity in the design of visualisation techniques

Visual data mining[1] (Michalski et al., 1999) is an approach to explorative data analysis and knowledge discovery that is built on the extensive use of visual computing (Gross, 1994, Nielson et al., 1997). The basic assumption is that large and normally incomprehensible amounts of data can be reduced to a form that can be understand and interpreted by a human through the use of visualisation techniques based on a particular metaphor or a combination of several metaphors (preferably, but not necessarily preserving the consistency of their combination). The popularity of digital terrain models, based on the geographical framework (Hetzler et al., 1998a, Hetzler et al., 1998b) and CAD-based architectural models of cities, mapping the metaphor of urban design to information visualisation, has demonstrated that multi-dimensional visualisation can provide a superior means for exploring large data-sets, communicating model results to others and sharing the model (Brown, 1998). Some recent developments are extending visual data mining with algorithmic animation techniques (Meisalo et al., 1998), multimedia support (Noirhomme-Fraiture, 2000) and incorporation of virtual reality immersive representations, aiming at involving wider range of human "input" channels in the mining and discovery processes.

The design of visualisation models for visual data mining, in broad sense, is the formal definition of the rules for translation of data into graphics. Generally, the visualisation of large volumes of abstract data, known as 'information visualisation' is closely related but sometimes contrasted, to scientific visualisation, which is concerned with the visualisation of

---

[1] Visual data mining spans from 2D visualisations into the use of a virtual reality systems, for example, see http://www.cs.auc.dk/3DVDM/about.html.

(numerical) data used to form concrete representations (Nielson et al., 1997). The frequently used expression "the art of visualisation" appropriately describes the state of research and development in that field. *Currently, it is a challenging activity for information designers to find out the methods, techniques and corresponding tools available to visualise a particular type of information*.

The *comparison of visualisation techniques across different designs* is not a trivial problem (Chen, 1999). Partially, current situation is explained by the lack of generic criteria to access the value of visualisation models. This is a research challenge, since most people develop their own criteria for what makes a good visual representation. The design of visualisation schemata is an area, dominated by individual points of views, which has resulted in a considerable variety of ad hoc techniques (Chen, 1999a). A recent example is the visualisation of association rules proposed in (Hofmann et al., 2000), where rule predicates are associated with shaded rectangular areas in attempt to convey visually some quantitative, qualitative and even semantic information about the rule patterns.

Another integral part of information visualisation is *the evaluation of how well humans understand visualised information* as part of their cognitive tasks and intellectual activities, the efficiency of information compression and level of cognitive overload. (Crapo et al., 2000) has investigated some aspects of developing visualisation schemata from cognitive point of view. Different cognitive styles in human activities, especially in stimulating insights and creativity can play key role in the success of one or another method. For example, the circle of linked histograms visualisation mechanism of Daisy[2] offers a simple visualisation of relationships of different groupings of data records, which has the potential variety of emergent geometries that can lead to some insights about more complex relationships in the data.

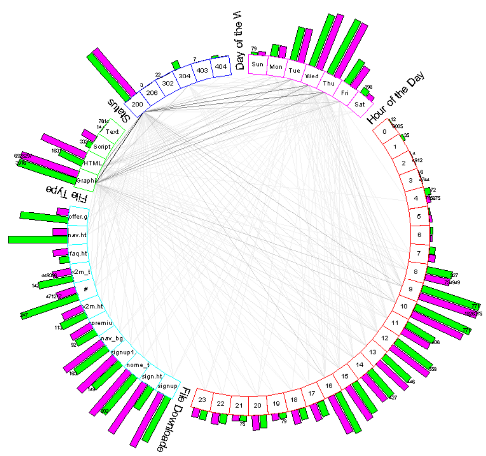


Figure 1. The visualisation schemata of Daisy Analysis visual data mining tool

Consequently, with the increasing attention towards development of interactive visual data mining methods the development of more systematic approach towards the design of visualisation techniques is getting on the "to do" list of the data mining research.

1.1 The special needs of collaborative and immersive data mining

The success of a visual data mining method depends on the development of an adequate computational model of selected metaphor. This is especially important in the context of communicating and sharing of discovered information, and in the context of the emerging methods of computer-mediated collaborative data mining. This new visual data mining methodology is based on the assumption that individuals usually may respond with

[2] http://www.daisy2000.com/

different interpretations of the same information visualisation (Snowdon et al., 1995). A central issue is the communicative role of abstract information visualisation components in collaborative environments for visual data mining. In fact, "miners" can become part of the visualisation. For example, in virtual worlds this can happen through their avatars[3]. In a virtual world, a collaborative perspective is inevitable, thus a shared understanding of the underlying mapping between the semantic space and the visualisation scheme becomes a necessary condition in the development of these environments. The results of CMCDM are heavily influenced by the adequate formalisation of the metaphors that are used to construct the virtual environment, i.e. the visualisation schemata can influence the behavior of collaborators, the way they interact with each other, the way that they reflect on the changes and manipulations of visualisations, and, consequently, their creativity in the discovery process.

1.2 Challenges towards the development of environments for design of visualisations

Currently, it is a challenging task for designers of visual data mining environments to find the strategies, methods and corresponding tools to visualise a particular type of information. Mapping characteristics of data into a visual representation in virtual worlds is one promising way to make the discovery of encoded relations in this data possible. The model of semantically organised place for visual data exploration can be useful for the development of computer support for visual information querying and retrieval (Del Bimbo, 1999, Gong, 1998) in collaborative information filtering. The development of a representational and computational model of selected metaphor(s) for data visualisation will assist the design of virtual environments, dedicated to visual data exploration. This paper argues for the development of more formal approach towards the design and evaluation of visualisation techniques.

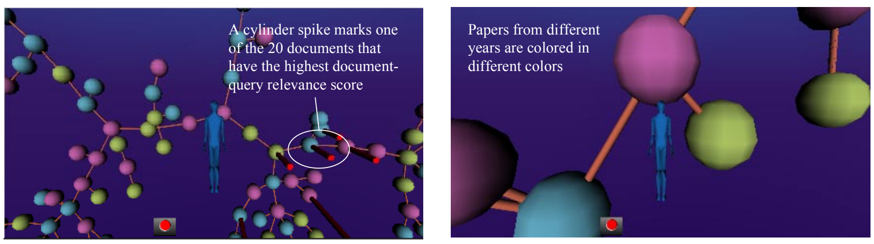
2. Background to a formal approach for the design of visualisation techniques

The formal approach presented in this paper is based on the concept of *semantic visualisation* defined as a visualisation method, which establishes and preserves the semantic link between form and function in the context of the visualisation metaphor. Establishing a connection between form and functionality is not a trivial part of the visualisation design process. In a similar way, selecting the appropriate form for representing data graphically, whether the data consists of numbers or text, is not a straightforward procedure as numbers and text descriptions do not have a natural visual representation. On the other hand, how data are represented visually has a powerful effect on how the structure and hidden semantics in the data is perceived and understood. An example of a virtual world, which attempts to visualise an abstract semantic space, is shown in Figure 2. The visualisation of the semantic space of the domain of human-computer interaction is automatically constructed from a collection of papers from three consecutive ACM CHI conference proceedings (Chen, 1999). The overall landscape is designed according to the theory of cognitive maps. In such a world, there is a variety of possibilities for data exploration. For example, topic areas of papers are represented by coloured spheres. If a cluster of spheres includes every colour but one, this suggests that the particular topic area, represented by the missing coloured sphere, has not been addressed by the papers during that year. However, without the background knowledge of the semantics of coloured spheres, selected information visualisation scheme does not provide cues for understanding and interpreting the environment landscape. It is not clear how the metaphor of a "landscape" has been formalised and represented, what are the elements of the landscape. Associatively, this visualisation is closer with the visualisation of molecular

[3] 3D representations of people in virtual worlds. Avatar is an ancient Sanskrit term meaning ‘a god’s embodiment on the Earth’ (Damer, 1998).

structures. Consequently, will an expert in molecular chemistry be more efficient in discovery specific relations between the visualised entities, based on her/his knowledge of possible associations between particular link configurations and molecular properties.

Semantic visualisation is considered in the context of two derivatives of visualisation - *visibilisation* and *visistraction* (Choras and Steinmann, 1995). Visibilisation is visualisation focusing on the presentation and interpretation which complies with rigorous mapping from physical reality. By contrast, visistraction is the visualisation of abstract concepts and phenomena, which do not have a direct physical interpretation or analogy. Visibilisation has the potential to bring key insights, by emphasising aspects that were unseen before. The dynamic visualisation of the heat transfer during the design of the heat-dissipating tiles cover of the underside of the space-shuttle is an early example of the application of visibilisation (Gore, 1981). Visistracton can give a graphic depiction of intuition regarding objects and relationships. The 4D simulation of data flow is an example of visistraction, which provides insights impossible without it. In a case-base reasoning system, visistraction techniques can be used to trace the change of relationships between different concepts with the addition of new cases.



a. b.

Figure 2. An example of virtual world visualising abstract semantic structure

Both kinds of semantic visualisation play important role in visual data mining. However, semantic visualisation remains a hand-crafted methodology, where each case is considered separately. This paper presents a consistent approach to semantic visualisation based on a cognitive model of metaphors, metaphor formalisation and evaluation. We illustrate the application of this approach with examples from visistraction of communication and collaboration data. Further, the paper is organised as follows. Section 3 presents the Form-Semantics-Function framework for construction and evaluation of visualisation techniques, Section 4 presents an example of the application of the framework towards the construction of a visualisation technique for identifying patterns of team collaboration, Section 5 presents an example of the application of the framework for evaluation and comparison of two visualisation models. The paper concludes with the issues for building visualization models that support collaborative data mining.

3. Form-Semantics-Function: a formal approach towards constructing and evaluating visualisation techniques.

The Form-Semantic-Function (FSF) approach includes the following steps: *metaphor analysis*; *metaphor formalisation*; and *metaphor evaluation*. Through the use of metaphor, people express the concepts in one domain in terms of another domain (Lakoff and Johnson, 1980, Lakoff, 1993). The closest analogy is VIRGILIO (L'Abbate and Hemmje, 1998), where the authors proposed a formal approach for constructing metaphors for visual information retrieval. The FSF framework develops further the formal approach towards constructing and evaluating visualisation techniques, approaching the metaphor in an innovative way.

3.1 Metaphor analysis

During metaphor analysis, the content of the metaphor is established. In the use of metaphor in cognitive linguistics, the terms *source* and *target*[4] refer to the conceptual spaces connected by the metaphor. The target is the conceptual space that is being described, and the source is the space that is being used to describe the target. In this mapping the structure of the source domain is projected onto the target domain in a way that is consistent with inherent target domain structure (Lakoff, 1993, Turner, 1994). In the context of semantic visualisation, the consistent use of metaphor is expected to bring an understanding of a relatively abstract and unstructured domain in terms of more concrete and structured visual elements through the visualisation schemata.

An extension of the source-target mapping, proposed by (Turner and Fauconnier, 1995) includes the notion of generic space and blend space. Generic space contains the skeletal structure that applies to both source and target spaces. The blend space often includes structure not projected to it from either space, namely emergent structure on its own. The ideas and inspirations developed in the blend space can lead to modification of the initial input spaces and change the knowledge about those spaces, i.e. to change and evolve the metaphor. The process is called conceptual blending - it is the essence in the development of semantic visualisation techniques.

In presented approach, the form-semantics-function categorisation of the objects being visualised, is combined with the (Turner and Fauconnier, 1995) model. The form of an object can express the semantics of that object, that is, the form can communicate implicit meaning understood through our experiences with that form. From the form in the source space we can connect to a function in the target space via the semantics of the form. The model of the metaphor analysis is shown in Figure 3.

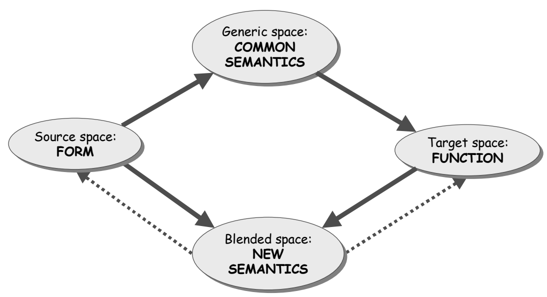


Figure 3. A model for metaphor analysis for constructing semantic visualisation schemata

The term *visual form* refers to the geometry, color, texture, brightness, contrast and other visual attributes that characterise and influence the visual perception of an object. Thus, the source space is the space of 2D and 3D shapes and the attributes of their visual representation. "Functions" (the generalisations of patterns, discovered in data) are described using concepts from the subject domain. Therefore the target space includes such concepts associated with the domain functions. This space is constructed from the domain vocabulary. The actual transfer of semantics has two components - the common semantics, which is carried by notions that are valid in both domains and what is referred as new semantics - the blend, which establishes the unique characteristics revealed by the correspondence between

[4] In the research literature the target is variously referred to as the primary system or the topic, and the source is often called the secondary system or the vehicle.

the form metaphor and functional characteristics of that form. The schema illustrates how metaphorical inferences produce parallel knowledge structures.

3.2 Metaphor formalisation

The common perception of the word "formalisation" is connected with the derivation of some formulas and equations that describe the phenomenon in analytical form. In this case, formalisation is used to describe a series of steps that ensure the correctness of the development of the representation of the metaphor. Metaphor formalisation in the design of semantic visualisation schemes includes the following basic steps:

- *Identification of the source and target spaces of the metaphor* - the class of forms and the class of features or functions that these forms will represent;
- *Conceptual decomposition of the source and target spaces* produces the set of concepts that describe both sides of the metaphor mapping. As a rule, metaphorical mappings do not occur isolated from one another. They are sometimes organized in hierarchical structures, in which `lower' mappings in the hierarchy inherit the structures of the `higher' mappings. In other words, this means that visualisation schemes, which use metaphor are expected to preserve the hierarchical structures of the data that they display. In visistraction, these are the geometric characteristics of the forms from the source space, and other form attributes like colours, line thickness, shading, etc. and the set of functions and features in the target space associated with these attributes and variations;
- *Identifying the dimensions of the metaphor* along which the metaphor operates. These dimensions constitute the common semantics. In visistraction this can be for instance key properties of the form, like symmetry and balance with respect to the center of gravity, that transfer semantics to the corresponding functional elements in the target domain;
- *Establishing semantic links, relations and transformations* between the concepts in both spaces, creating a resemblance between the forms in the source domain and the functions in the target domain.

3.3 Metaphor evaluation

In spite of the large number of papers describing the use of the metaphor in the design of computer interfaces and virtual environments, there is a lack of formal evaluation methods. In the FSF framework metaphor evaluation is tailored following the (Anderson et al., 1994) model, illustrated in Figure 4.

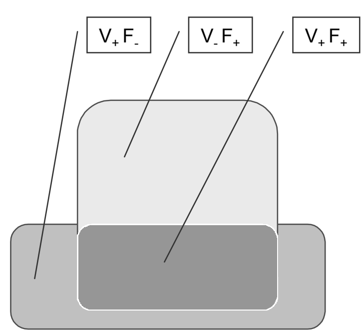


Figure 4. Model for evaluating metaphor mapping (based on (Anderson et al., 1994))

The "V" and "F" are labels for visualisation and function features, respectively. The "VF" label with indices denotes numbers of features, namely:

- V_+F_+ - function features that are mapped to the visualisation schema;

- $V_{-}F_{+}$ - function features that are not supported by the visualisation schema;
- $V_{+}F_{-}$ - features in the visualisation schema, not mapped to the functional features.

The ratio $\frac{V_{-}F_{+}}{V_{+}F_{+}}$ provides an estimate of the quality of the metaphor used for the visualisation - the smaller the better.

4. Application of the Form-Semantics-Function approach

The elements of the Form-Semantics-Function approach are illustrated in the examples of constructing visualisation schema for visual data mining and simple evaluation of two text data visualisation techniques. The first example illustrates the metaphor analysis and formalisation stages for the creation of visualisation form and mapping it to the functional features. In the second example the evaluation of the two visualisation schemata is based on the same set of functional features.

4.1 Constructing visualisation schema for visual data mining

The goal of the development of this visualisation schema is to produce a simple but appealing visual representation of asynchronous collaboration data so that researchers will be able to identify patterns in team collaboration. Asynchronous communication is an intrinsic part of computer-mediated teamwork. Among the various models and tools supporting this communication mode (Maher et al., 2000), perhaps the most popular in teamwork are bulletin (discussion) boards. These boards support multi-thread discussion, automatically archiving communication content. One way to identify patterns in team collaboration is via content analysis of team communications. However, it is difficult to automate such analysis, therefore, especially in large scale projects, monitoring and analysis of collaboration data can become a cumbersome task.

In the research in virtual design studios (Maher et al., 1997, Maher et al., 2000) there have been identified two extremes (labeled as "Problem comprehension" and "Problem division") in team collaboration, illustrated in Figure 5. In "Problem comprehension" collaborative mode the resultant project output - a product, solution to a problem, etc., is a product of a continued attempt to construct and maintain a shared conception and understanding of the problem. In other words each of the participants is developing own view over the whole problem and the shared conception is established during the collaborative process via intensive information exchange.

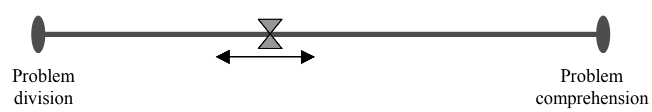
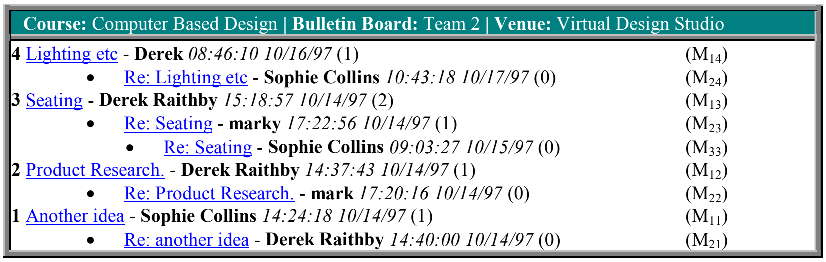


Figure 5. Two extremes in team collaboration

In "Problem division" mode the problem is divided among the participants in a way where each person is responsible for a particular portion of the investigation of the problem. Thus, it does not necessarily require the creation of a single shared conception and understanding of the problem. The two modes of collaboration are two extreme cases. In general, the real case depends on the complexity of the problem.

A key assumption in mining and analysis of collaboration data is that this two extreme styles should be some how reflected in the communication of the teams. Thus, different patterns in team communication on the bulletin board will reflect different collaboration modes. Figure 6 illustrates a fragment of a team bulletin board.



Course: Computer Based Design | **Bulletin Board:** Team 2 | **Venue:** Virtual Design Studio

4 Lighting etc - **Derek** *08:46:10 10/16/97* (1) (M_{14})
- Re: Lighting etc - **Sophie Collins** *10:43:18 10/17/97* (0) (M_{24})

3 Seating - **Derek Raithby** *15:18:57 10/14/97* (2) (M_{13})
- Re: Seating - **marky** *17:22:56 10/14/97* (1) (M_{23})
  - Re: Seating - **Sophie Collins** *09:03:27 10/15/97* (0) (M_{33})

2 Product Research. - **Derek Raithby** *14:37:43 10/14/97* (1) (M_{12})
- Re: Product Research. - **mark** *17:20:16 10/14/97* (0) (M_{22})

1 Another idea - **Sophie Collins** *14:24:18 10/14/97* (1) (M_{11})
- Re: another idea - **Derek Raithby** *14:40:00 10/14/97* (0) (M_{21})

Figure 6. Bulletin board fragment with task-related messages, presented as indention graph

The messages on the board are grouped in threads. (Berthold et al., 1997, Berthold et al., 1998) propose a threefold split of the thread structure of e-mail messages in discussion archives in order to explore the interactive threads. It included (i) reference-depth: how many references were found in a sequence before this message; (ii) reference-width: how many references were found, which referred to this message; and (iii) reference-height: how many references were found in a sequence after this message. In addition to the threefold split, (Sudweeks and Simoff, 2000) included the time variable explicitly. Figure 7 shows the formal representation of the bulletin board fragment in Figure 6.

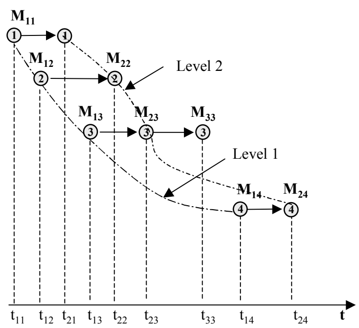


Figure 7. Formal representation of the thread structure in the fragment presented in Figure 6.

4.1.1 Metaphor analysis

Figure 8 shows the Form-Semantics-Function mapping at the metaphor formalisation stage as a particular case of the (Turner and Fauconnier, 1995) model applied to the visualisation of communication utterances data. The source space in this case is the space of 2D geometric shapes, rectangles in particular. The target space includes the concepts associated with the functions that are found in the analysis of a collaborative design session.

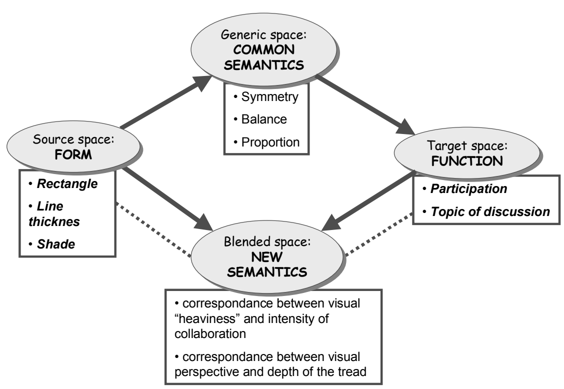


Figure 8. The Form-Semantics-Function mapping for the source space of nested rectangles and the target space of bulletin board discussion threads.

4.1.2 *Metaphor formalisation*

Below is a brief illustration of the metaphor formalisation in this example.

- *Identification of the source and target spaces of the metaphor* - rectangles are the forms that will be used and the class of features or functions that these forms will represent are the messages on a bulletin board;
- *Conceptual decomposition of the source and target spaces* leads to the notion of nested rectangles, whose centers of gravity coincide, with possible variation of the thickness of their contour lines and the background color. Each rectangle corresponds to a message within a thread. Rectangle that corresponds to a message at a level (n + 1) is placed within the rectangle that corresponds to a message at level n. Messages at the same level are indicated by a one step increase of the thickness of the contour line of the corresponding rectangle. Thus, a group of nested rectangles can represent several threads in a bulletin board discussion;
- *Identifying the dimensions of the metaphor* - visual balance and the "depth" or "perspective" of the nested rectangles are the dimension of the metaphor, transferring via the visual appearance the semantics of different communication styles;
- *Establishing semantic links, relations and transformations* - this is connected with the identification of typical form configurations that correspond to typical patterns of collaboration. For example, Figure 9 illustrates two different fragments A and B (each of one thread). Figure 10 illustrates the visualisation of this fragments according to the developed visualisation schema.

The visualisation schema has been used extensively in communication analysis. Figure 11 illustrates communication patterns corresponding to different collaboration styles. An additional content analysis of communication confirmed the correct identification of collaboration patterns.

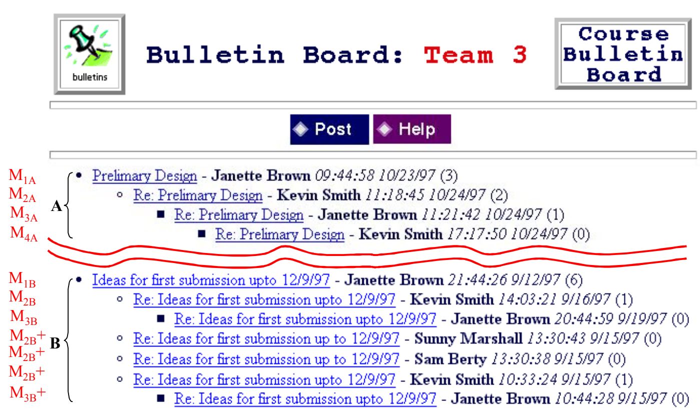


Figure 9. Bulletin board fragment with task-related messages, presented as indention graph

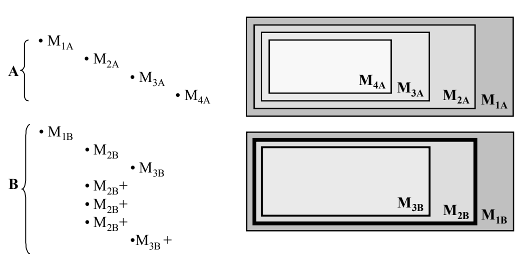
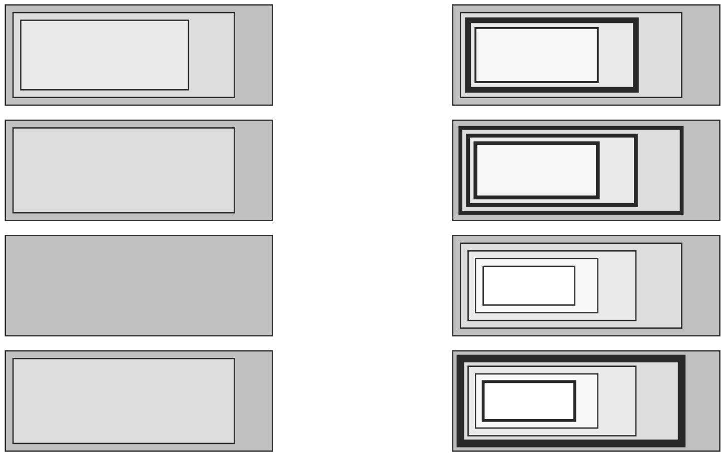


Figure 10. Visualisation of fragments A and B in Figure 9.



a. Communication patterns, typical for "problem division" collaboration

b. Communication patterns, typical for "problem comprehension" collaboration

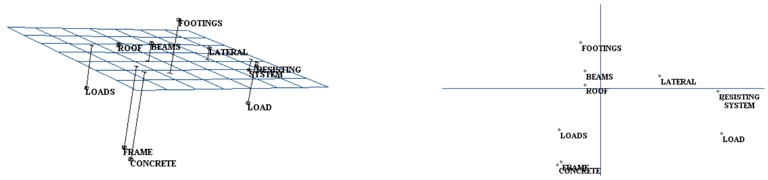
Figure 11. Communication patterns, corresponding to different collaboration styles.

4.2 Evaluation and comparison of two visualisation schemata

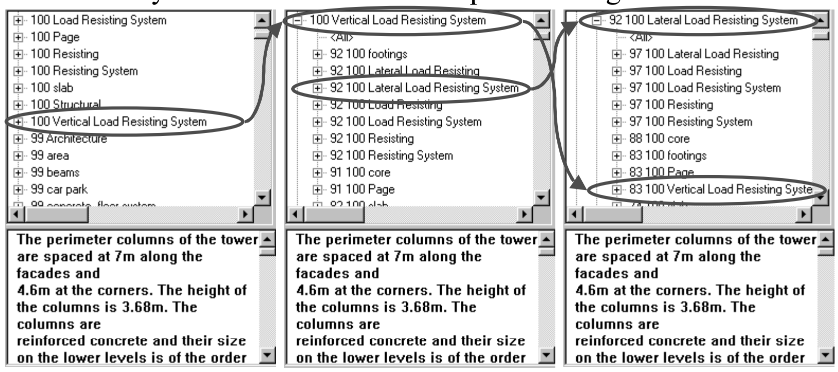
We illustrate the idea by evaluating examples of semantic visualisation of textual data and objects in virtual environments. The role of visistraction in concept relationship analyis is to assist the discovery of the relationship between concepts, as reflected in the available text data. The analysis uses word frequencies, their co-occurence and other statistics, and cluster analysis procedures. We investigate two visual metaphors - "Euclidian space" and "Tree", which provide a mapping from the numerical statistics and cluster analysis data into the target space of concepts and relations between them. The visualisation features for both metaphors and the function features of the target space are shown in Table 1. Examples of the two visualisation metaphors are shown in Figure 12.

Table 1. Visualisation and function features

Visualisation features of Euclidian space metaphor	**Visualisation features of tree metaphor**	**Function features**
- point - alphanumeric single-word point labels - axes - plane - color - line segment	- nodes - alphanumeric multi-word node labels - signs "+" and "-" - branches - numeric labels for branches	- simple/complex concept - subject key word - hierarchical relationship - context link - link strength - synonymy - hyponymy



a. 10 most frequent words in the description of lateral load resisiting system in one of the wide-span building cases[5]



b. Part of a semantic net[6] for the term "load resisting system".

Figure 12. Visualisation of the results of cluster analysis performed over a text data set[7]

4.2.1 Metaphor evaluation

The first scheme[8] maps the source domain of Euclidean space (coordinates of points in 2D/3D space) to the target domain of word statistics. The blending semantics is that the degree, to which the terms are related to each other, can be perceived visually from the distance between the corresponding data points - the closer the points the tighter is the relationship between the words. The second scheme[9] maps the source domain of the topology of linked nodes to the same target domain of words statistics. This mapping generates one of the possible visualisations of semantic networks. This visualisation includes nodes with single- and multiple-word labels, numeric values of each link between terms and the weight of the term among the other terms in the tree. The results of the comparison between the two metaphors are presented in Table 2 and

Table 3.

The Euclidean space metaphor has a poor performance for visualisation of concept relationships. What is the meaning of such closeness - it is difficult to make a steady judgement about what the relation is and whether we deal with simple (one word) or complex (more than one word) terms. The distance to the surface, proportional to the frequency of the words, can convey the message that a word is a key word. However, there is no feature in the visualisation, which shows context links between words, the strength of this links and other relations between words.

---

[5] This visualisation is used in TerraVision, which is part of the CATPAC system for text analysis by Provalis Research Co.

[6] This is the visualisation of semantic networks in TextAnalyst by Megaputer Intelligence, Inc.

[7] The source text comes from the SAM (Structure and Materials) case library available at http://www.arch.usyd.edu.au/kcdc/caut/

[8] The schema is used in the CATPAC qualitative analysis package by Terra Research Inc.

[9] The schema is used in TextAnalyst by Megaputer Intelligence (see www.megaputer.com).

Table 2. Visualisation support for function features in Euclidean space and tree metaphors

Function features	**Support by the Euclidean space metaphor**	**Support by the Tree metaphor**
Simple/complex concept	-	+
Subject key word	+	+
Hierarchical relationship	-	+
Context link	-	+
Link strength	-	+
Synonymy	-	-
Hyponymy	-	-

Table 3. Comparison of in Euclidean space and tree metaphors

	Euclidean space metaphor	**Tree metaphor**
V_+F_+	1	5
V_-F_+	6	2
$\frac{V_-F_+}{V_+F_+}$	6	0.4

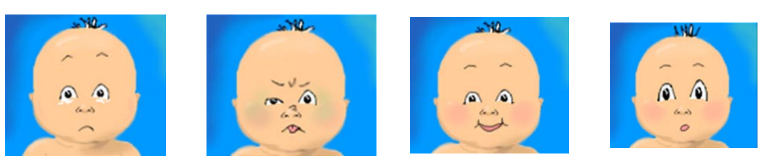
5. Limitations of proposed approach and future directions

The Form-Semantics-Function framework resented in this work is an attempt to develop a formal approach towards the use of metaphors in constructing consistent visualisation schemes. In its current development, it has at least the following limitations:

- to preserve consistency for each change in the visualisation schemata the analysis/formalisation pass has to be conducted again;
- the evaluation part of the framework does not include the analysis of the cognitive overload from the point of information design.

Currently the FSF framework is further developed in the following directions:

- *supporting visualisation schemes based on affective computing models*, which aim at communicating specific emotions for specific patterns in the data. The idea is illustrated by using the emotional expression of a human face to represent four data patterns in Figure 2 (a-d). It is based on the assumption that the observer can perceive any change in the emotional expression of the face, i.e. in one or more of its features, when representing a different pattern in data, accurately.



a. b. c. d.

Figure 13. Emotion-based agents acting as visualisation support in multi-agent data mining environment (the actual faces are adapted from (Ventura, 2000))

- *supporting collaborative visual data mining in virtual worlds*. The different perceptions of a visualisation model in a such environment may increase the gap between individuals as they interact with it in a data exploration session. However, individual differences may lead to a potential variety of discovered patterns and insights in the visualised information across participants.

In this context the future research within the FSF framework will be focused on exploring:

- whether people attach special meanings to abstract visualisation objects;
- what are the design criteria towards visualisation objects, engaged in visual data exploration, that people can effectively construct and communicate knowledge in visual data mining environments;
- what are the necessary cues that should be supported in semantically organised virtual environments.

Acknowledgements

This research has been supported by the University of Technology Sydney and an ATN Research Grant.

References

Anderson, B., Smyth, M., Knott, R. P., Bergan, M., Bergan, J. and Alty, J. L. (1994). Minimising conceptual baggage: Making choices about metaphor. In Cocton, G., Draper, S. and Weir, G. (eds), *People and computers IX* Cambridge University Press, Cambridge, pp. 179-194.

Berthold, M. R., Sudweeks, F., Newton, S. and Coyne, R. (1997). Clustering on the Net: Applying an autoassociative neural network to computer-mediated discussions., *Journal of Computer Mediated Communication,* **2**(4), . Available: http://www.ascusc.org/jcmc/vol2/issue4/bert-hold.html.

Berthold, M. R., Sudweeks, F., Newton, S. and Coyne, R. (1998). It makes sense: Using an autoassociative neural network to explore typicality in computer mediated discussions. In Sudweeks, F., McLaughlin, M. and Rafaeli, S. (eds), *Network and Netplay: Virtual Groups on the Internet* AAAI/MIT Press, Menlo Park, CA, pp. 191-220. Available: http://www.ascusc.org/jcmc/vol2/issue4/bert-hold.html.

Brown, I. M. (1998). A 3D user interface for visualisation of Web-based data-sets, *Proceedings of the 6th international symposium on Advances in geographic information systems*, , , pp. 100-105.

Chen, C. (1999). *Information Visualisation and Virtual Environments,* Springer-Verlag, London.

Choras, D. N. and Steinmann, H. (1995). *Virtual reality: Practical applications in business and industry,* Prentice Hall, Upper Saddle River, New Jersey.

Crapo, A. W., Waisel, L. B., Wallace, W. A. and Willemain, T. R. (2000). Visualization and the process of modeling: A cognitive-theoretic approach, *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD-2000*, The Association for Computing Machinery, Boston, MA, USA, pp. 218-226.

Damer, B. (1998). *Avatars,* Peachpit Press, an imprint of Addison Wesley Longman.

Del Bimbo, A. (1999). *Visual information retrieval,* Morgan Kaufmann Publishers, Inc., San Francisco, CA.

Gong, Y. (1998). *Intelligent image databases: Towards advanced image retrieval,* Kluwer Academic Publishers, Boston, MA.

Gore, R. (1981). When the space shuttle finally flies, *National Geographic,* **159**, 317-347.

Gross, M. (1994). *Visual computing: The integration of computer graphics, visual perception and imaging,* Springer-Verlag, Heidelberg.

Hetzler, B., Harris, W. M., Havre, S. and Whitney, P. (1998a). Visualising the full spectrum of document relationships, *Proceedings of the Fifth International ISKO Conference:*

Structures and Relations in the Knowledge Organisation, , Lille, France. Available: http://multimedia.pnl.gov:2080/infoviz/isko.pdf.

Hetzler, B., Whitney, P., Martucci, L. and Thomas, J. (1998b). Multi-faceted insight through interoperable visual information analysis paradigms, *Proceedings of IEEE Information Visualization 98*, IEEE, . Available: http://multimedia.pnl.gov:2080/infoviz/ieee98.pdf.

Hofmann, H., Siebes, A. P. J. M. and Wilhelm, A. F. X. (2000). Visualizing association rules with interactive mosaic plots, *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD-2000*, The Association for Computing Machinery, Boston, MA, USA, pp. 227-235.

L'Abbate, M. and Hemmje, M. (1998). VIRGILIO - The metaphor definition tool. (ed.)^(eds), GMD Report, Darmstadt.

Lakoff, G. (1993). The contemorary theory of metaphor. In Ortony, A. (ed.) *Metaphor and thought* Cambridge University Press, Cambridge, pp. 202-251.

Lakoff, G. and Johnson, M. (1980). *Metaphors we live by,* University of Chicago Press, Chicago, IL.

Maher, M. L., J, S. S. and A, C. (1997). Potentials and limitations of Virtual Design Studio, *Interactive Construction On-line,* **January, a1**, . Available: http://www.inconstruction.com.

Maher, M. L., Simoff, S. J. and Cicognani, A. (2000). *Understanding virtual design studios,* Springer-Verlag, London, UK.

Meisalo, V., Sutinen, E., Tarhio, J. and Teraumlsvirta, T. (1998). Combining algorithmic and creative problem solving on the Web, *Proceedings Teleteaching '98/IFIP World Computer Congress 1998*, Austrian Computer Society, , pp. 715-724.

Michalski, R. S., Bratko, I. and Kubat, M. (Eds) (1999) *Machine learning and data mining,* John Wiley and Sons, Chichester, England.

Nielson, G. M., Hagen, H. and Muller, H. (1997). *Scientific visualization : overviews, methodologies, and techniques,* IEEE Computer Society,, Los Alamitos, CA.

Noirhomme-Fraiture, M. (2000). Multimedia Support for Complex Multidimensional Data Mining, *Proceedings of the First International Workshop on Multimedia Data Mining (MDM/KDD'2000), in conjunction with Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD 2000*, ACM Press, Boston, MA, pp. 54-59.

Snowdon, D. N., Greenhalgh, C. M. and Benford, S. D. (1995). What You See is Not What I See: Subjectivity in virtual environments, *Proceedings Framework for Immersive Virtual Environments - FIVE'95*, , London, UK.

Sudweeks, F. and Simoff, S. (2000). Complementary explorative data analysis: The reconciliation of quantitative and qualitative principles. In Jones, S. (ed.) *Doing Internet Research*, Vol. 29-55 Sage Publications, .

Turner, M. (1994). Design for a theory of meaning. In Overton, W. and Palermo, D. (eds), *The Nature and Ontogenesis of Meaning* Lawrence Erlbaum Associates, , pp. 91-107.

Turner, M. and Fauconnier, G. (1995). Conceptual integration and formal expression, *Journal of Metaphor and Symbolic Activity,* **10**(3), 183-204.

Ventura, R.M. (2000). *Emotion-Based Agents*, MSc Thesis, Lisbon Polytechnics, Lisbon.