Research proposal (3/2003)

Information Visualisation in CASE Environments

Enhancing the Graphical Information System Models with Advanced Visualisation Techniques

Jouni Huotari

jouni.huotari@jypoly.fi
Jyväskylä Polytechnic / School of Information Technology

ABSTRACT

Our research focuses on supporting understanding of complex information spaces, such as design repositories maintained by CASE tools. Due to the growing size and complexity of modern information systems, critical design information is often distributed via multiple diagrams at many different levels of abstraction. This slows search performance and results in errors that later cause omissions and inconsistencies in the final designs. We examine how advanced visualisation techniques help understanding these complex relationships between information elements and thus reduce cognitive overload. We implement an application through which a user can maintain a sense of location and direction by getting assistance what information exists within an information space and how information elements connect to each other. The ideas are applicable also in other domains where the information space contains hierarchical and / or network structure.

Keywords: information visualisation, information interfaces and presentation, human information processing, diagrammatic representation, spatial ability, visual search

Supervisor: Professor Kalle Lyytinen
Department of Information Systems
Case Western Reserve University
10900 Euclid, 44106 Cleveland, USA

CONTENTS

1. INTRODUCTION AND OVERVIEW TO THE RESEARCH

Introduction

During the last decades, we have seen a rapid growth in the use and importance of information technology (IT). IT has many implications in the society. Although the aim of using IT is to help people to more productive work, there are many examples of failures. The problems experienced with these systems originate from several factors (Ewusi-Mensah, 1997), but basically, we can identify one common factor: erroneous human nature. Consequently, when people make information systems for other people to use, there is always a possibility of an error to occur. The increasing demand for high-quality systems that should be delivered to market in a fast pace often results over-the-budget, poorly documented systems that are difficult to maintain.

The rapid growth of electronic information has resulted an increasing difficulty to understand what information actually means and how it relates to other information. These complex relationships between information elements exist also in information system development (ISD) and computer-aided systems/software engineering (CASE) environments. They consist of many interrelated diagrams and documents ranging from requirement specifications to documentations of codes.

Developing information systems is a complex process. Information systems development is "a change process taken with respect to an object system in an environment by a development group using tools and an organised collection of methods to produce a target system" (Lyytinen, 1987). This definition takes into consideration technical, conceptual, and organisational aspects. Consequently, the concept of a system is wide, including e.g. code, a database, or a method. Even an organisation itself can be seen as a system.

An ISD method defines the way of carrying out the change process. There is no unified view what constitutes a method. We define a method as an organised collection of concepts, techniques, beliefs, values, and normative principles (Hirschheim et al., 1995). A central part of a method is the technique, which explains the procedure how to perform a task to accomplish a desired state (Welke, 1983). One of the key issues is to address the problem how to visualise effectively those outcomes that form the output from different modelling techniques.

The automation of the ISD process has long been an ultimate goal in software development companies. Even though the problems in ISD are often non-technical – the commitment of the stakeholders as an example (Iivari, 1994)

– properly used CASE tools, CASE shells (Bubenko, 1988), or metaCASE tools (Alderson, 1991), can increase productivity, improve system's quality, shorten the development lifecycle, and result in reusable methods, method components, and code. CASE, on the large, includes any computer support from designing specifications to automatic generation of database and code. A CASE tool typically supports only on part of the development process, whereas a CASE environment[1] is aimed to support a large part of the ISD process. Standardising the ISD process may solve many of the design time problems, but it does not help people to understand those descriptions (specifications, design decisions etc.) that emerge during the ISD process (Shu, 1988).

Systems development is, fundamentally, a problem solving activity (Vessey and Glass, 1998). We can identify four problem areas where research and tools should be extended: the representational, conceptual, methodological, and implementation (Kelly and Smolander, 1996). Current commercial CASE (and CAME) tools and environments fall short in many respects (Kelly, 1997):

- Lack of different representational paradigms (diagrams, matrices, tables, etc.)
- Only simplistic checking and visualisation of the contents of the repository
- Only partial graphical support for the (meta)modelling process
- Difficulty to see how the modifications of the metamodel influence the corresponding model
- Lack of support for finding and seeing reused and reusable components
- Poor horizontal and vertical method integration
- Insufficient integration of different tools, such as project management tools, from different levels and phases of ISD

In addition to rigorous modelling, CASE tools should support soft aspects of software development, such as creativity and problem solving (Jarzabek and Huang, 1998). The most fundamental reason for the poor visualisation support is the lacking of features that enable the linking of and further showing concrete links between different information elements. For example, the (meta) data model must be made more powerful to enable the representation of all the kinds of information necessary (Kelly, 1997). The associated properties of the design information, such as creator of an object, creation date, or motivation for creation / modification could be of value when deciding further actions for an object. This metadata could be used also as categorising and further visualising that information.

We propose several solutions to address the problems mentioned above. To begin with, we need guidelines and standards, which tell what information to include, where to store that information, how it is done, how it is

---

[1] CASE environments has many different names, such as Integrated Project Support Environment (IPSE) and Software Engineering Environment (SEE).

achievable, and who does what and when. People need to be motivated to follow those guidelines. We should therefore explain why this is important and teach how to do it. Then, we need also technical solutions, such as an efficient repository, which integrates parts of the design documents to one manageable entity. Lastly, we need an intelligent browser, which supports finding and visualising information. In our research, we are concentrating mainly on visualisation aspects.

Modelling is a central part of the ISD. We build models for several reasons: to communicate the desired structure and behaviour of a system, to visualise and control its architecture, and to better understand the system we are building (Booch et al., 1998). In addition, computer supported modelling often helps simplification, reuse, and risk management.

Successful development of IS requires effective communication between all stakeholders (Figure 1). To help communication and to achieve a comprehensive understanding of the desired target, developers should produce meaningful and understandable representations (IS specifications). As the complexity of systems increases, so does the importance of good modelling and visualisation techniques. Grasping the idea behind an existing IS from its documentation is often time-consuming, partly because of the complexity to interpret how different parts of the documents relate to each other. Therefore, more efficient ways to visualise the design information are needed.

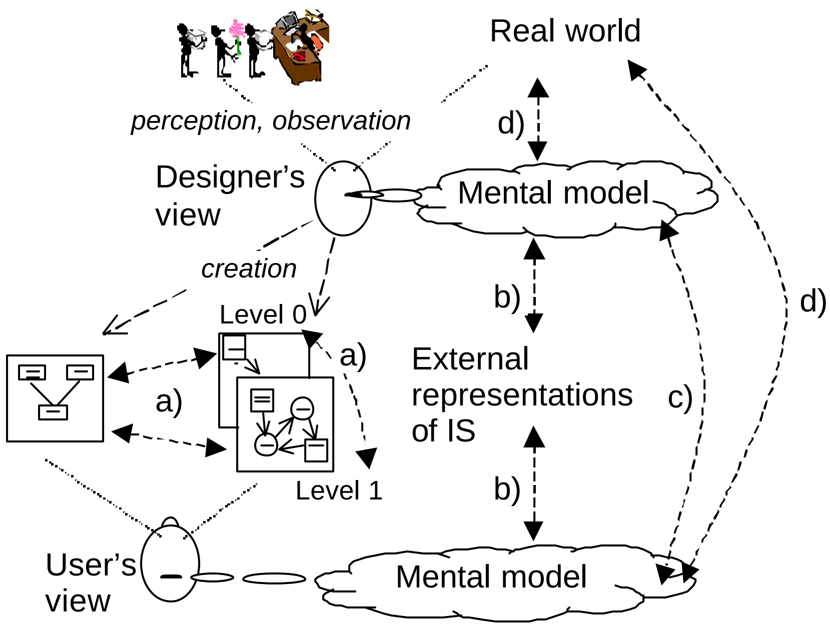

Figure 1. Consistency problems might occur between a) external representations, b) external representations and mental model, c) viewers' mental models, and d) mental model and real world. We concentrate mainly on the consistency problems between external representations (a).

The visualisation and conceptualisation of complex information spaces is considered important in several studies (Card et al., 1999; Chen, 2000; Monarchi and Puhr, 1992; Noik, 1994; Shneiderman, 1997; Spence, 1993; Tufte, 1983; and Tufte, 1990). For instance, through visualisation the structure of information space is explicitly represented, which essentially supports different ways of using the information, such as searching, browsing, interacting with, and thus understanding the complex information.

Information visualisation offers several promising benefits to communicate properties of the information. Moreover, it can reduce cognitive overload by filtering unnecessary information, as well as enhance users' perception and exploration of the content and structure of the information space. Thus, relevant information and overall structure can be shown in a meaningful way. Since advanced visualisation techniques, 3D visualisation, and virtual environments are emerging at a rapid rate, cognitive aspects must be considered before inappropriate designs and practices become common.

Even though the processing power and storage capacity have increased in the orders of magnitude, the screen size has remained quite small. There is simply not enough space on a typical computer screen to visualise the rich information content characteristic of today’s computers and computer networks. Therefore, also big screens and virtual environments could help seeing, exploration, and understanding how a piece of information relates to the bigger context.

A large portion of the brain, approximately 50 %, is devoted to visual processing. People have well-developed languages for visual communication and a capability to see things in two and three dimensions. In addition, people usually learn things better by seeing than by hearing. Vision is therefore argued to be the most important of human senses.

This most important human sense could be supported by advanced visualisation techniques. Now, when computers are capable to provide 3D fluently, even in virtual environments (Ware and Franck, 1996), their use should be carefully studied in order to utilise the full potential. Simultaneously, the characteristics of the internal representations (mental models) must be considered in order to make consistent and rapidly comprehensible depictions.

Our research identifies those representational problems that a information system designers confront. Here, we introduce some suggestions to solve those problems. A special research interest is towards utilising the third dimension (3D). We are taking advantage of 3D capabilities to display information in a way that the users are able to extract important elements from large amounts of information more easily, effectively, and/or rapidly. Now, when technology is capable showing complex representations also in 3D, its' potential should be thoroughly examined.

Our aim is to find suggestions that would support IS designers and other stakeholders in grasping a better the idea behind a complex set of specifications. It is one of the steps towards such CASE tools that would better answer the needs of CASE tool users, particularly by improving the visualisations during an ISD process. The findings do not merely help CASE tool builders to improve the existing CASE tools, but all designers benefit from them. For example, existing information in the Web, in databases, or in multimedia applications can be visualised by using advanced visualisation techniques.

The main objects of this study are information, visualisation, cognition, and ISD. These aspects are studied in many research fields, such as information systems research, cognitive science, and cognitive psychology. This necessitates a multi-faceted, integrating, and a boundary-crossing attitude of research.

In the next section we will define the basic terminology. Then we look at the research problem and the research methodology. The related research is reviewed in Section 4. A short summary of each paper is presented in Section 5, followed by a brief overall conclusion and directions for future research.

Terminology

Here we are going to look into some ontological questions, for example what is information visualisation. Choosing an appropriate definition for the terms is difficult, because different research fields or interest groups differ in their emphasis or point of view. For example, the term "visualisation" can be seen as a process or a product, emphasising either human (mental) or technical aspects. Here we aim at a broad, holistic view of defining basic terminology. Some of the terms are taken from the FRISCO[2] report (Falkenberg et al., 1998), which introduces the most thorough set of definitions of the most fundamental concepts in the information system field.

Data, Information, and Knowledge

What is data, what is information, and what is knowledge? These questions have long exercised the minds of philosophers. The FRISCO report defines the term **data** as any set of representations of knowledge (i.e. meaningful symbolic constructs, such as numbers, characters, or images), expressed in a language. The report denotes **information** as the knowledge increment, brought about by a receiving action in a message transfer, i.e. it is the difference between the conceptions interpreted from a received message and the knowledge before the receiving action. Here the word "message" may have several meanings, depending on the context. We can think of the physical appearance of a message, its syntax and semantics, and its pragmatics in a social context. These semiotic layers are helpful also when thinking terms such as visualisation. In

[2] FRISCO is an acronym for FRamework of Information System COncepts

Table 1 the terms data, information, and knowledge, are positioned on these semiotic layers. We can also identify a counterpart for them at different abstraction levels (Iivari, 1989).

Table 1. The concepts of data, information, and knowledge on different layers.

Main concept	Semiotic layer	Explanation	Abstraction level
Data	Physical	The physical appearance, the media	Technical
	Syntactical	The language, the structure, and the logic used	
Information	Semantical	The meaning and validity of what is expressed	Conceptual
Knowledge	Pragmatic	The intentions, responsibilities, and consequences behind the expressed statements	Organisational

Thinking with these layers or levels, we have been seeing a shift of research effort from technical and conceptual to wider, social context direction Kuutti and Bannon, 1993. In our research, we are focusing mostly on the middle level, while considering also the other levels.

In the same way as abstraction levels, Juhani Iivari (1989) has identified three domains: Organisational, Universe of Discourse (UoD), and Technical. For each domain, three abstractions can be distinguished: structure, function, and behaviour (Iivari, 1989). This apportionment can be used also as selection criteria for visualisation. For example, one could want to see what functional abstractions exist between the organisational and conceptual level. Another example is to change the representational paradigm e.g. from a table to a diagram on one abstraction level. In theory, we could even choose the method according to which the representation changes, e.g. the conceptual schema can be shown as an entity-relationship diagram (ERD) or as a class diagram in UML. In practise, mappings for such method conversions are very difficult to implement. In all, it is difficult to map the real world (RW) system to the information system and to compare representations of the RW system and IS (Wand, 1996).

Visualisation, Perception, and Representation

The four concepts – perception, conception, representation, and visualisation – are difficult to define separately without referring to one another. They all have many definitions, depending on the research field. The researchers from the computer science field emphasise **visualisation** as a product or a technique. As an example, visualisation is "the visual representation of a domain space using graphics, images, animated sequences, and sound augmentation to present the data, structure, and dynamic behaviour of large, complex data sets that represent systems, events, processes, objects, and concepts" (Williams et al., 1995). Researchers from many other fields view it more as a cognitive process, for example visualisation can be seen as ‘‘mechanisms by which humans perceive, interpret, use, and communicate visual information’’ (McCormick, 1987).

Card et al. (1999) synthesise these “hard” and “soft” views by regarding visualisation as “the use of computer-supported, interactive, visual representations of data to amplify cognition”. In addition, they characterise visualisations as “adjustable mappings from data to visual form to the human perceiver” (p. 17). Note that although visualisation is commonly related to seeing, we can use all other senses, too. For example, blind people use touch, and sound can augment seeing. Visualisation is thus a multi-faceted process, which has several research areas to be examined.

Ware (2000), McCormick et al. (1987), and Tufte (1983) list several advantages of visualisation. It

- Enables to view, compare, and comprehend huge amounts of data
- Allows the perception of emergent properties and thus fosters profound and unexpected insights
- Reveals hidden problems, deficiencies, or errors
- Eases hypothesis formulation and enriches the process of scientific discovery
- Facilitates understanding of both large-scale and small-scale features of the data

Visualisation can thus be invaluable in quality control and perception of patterns linking local features (Ware, 2000). Friedhoff (1990) argues that graphically rendered information is assimilated at a much faster rate. Robertson (1991) emphasises the selection of the appropriate representation, because it affects to an observer's mental model and further subsequent analysis, processing, or decision-making.

As noted before, people have remarkable perceptual abilities for visual information. **Perception** covers all human senses, even though vision is the most important one. Human perception and the interpretation of received information is a complex process. It involves several levels of processing, ranging from low-level sensory mechanisms to higher-level cognitive mechanisms. One fascinating view (Bruner, 1957) considers the perceptual process to be like science itself. According to Bruner, the perceptual process consists of finding clues, formulating a hypothesis about what the clue is, verifying it, and then either accepting the hypothesis or reformulating it and trying again. Bruner also thinks that “all perceptual experience is the end product of a categorisation process”. Moreover, Bruner states that concepts or categories have to be defined before any recognition is possible. In addition, subjects' knowledge and expectations influence how one interprets the perceived stimulus. Here we adopt this constructivist view, although we agree that “direct” pickup of relevant information is also possible (see Gibson, 1979).

Representation is one of the central concepts related to human cognition. Its varied use, however, is confusing. For example, Marr (1982) defines a representation as “a formal system for making explicit certain entities or types of information, together with a specification of how the system does this”. Here the system refers to brain, which processes the received information. In information system models, a representation is typically a labelled graphical symbol showing the values of properties of a given conceptual model component (Kelly, 1997). Representation can contain exact positioning information (coordinates) of a particular object in a diagram. Other examples of the properties of a representation are size, colour, and font information. These representational elements constitute a so-called notation. A representation can contain also some intelligence, e.g. when, where, and how the property values are displayed.

Distinction between conceptual and representational aspects of information is important. This can be seen as a dimension of information in CASE (Smolander et al., 1991), another dimension being type-instance. Kelly (1997) gives a good example of maintaining the same conceptual information and displaying it in several different representations, e.g. as a diagram (possible with different layouts), matrix, or a table. According to this example, "a Class Diagram graph could have two different diagram representations, one stressing the inheritance hierarchy and the other the aggregation hierarchy: the underlying conceptual graph would be the same for both. Most of the conceptual objects would then have representations in both diagrams, whereas the conceptual relationships would mostly have a representation in only one diagram or the other." Tweedie (1997) introduces three aspects to representation: data "behind" the representation (meta data / raw data); forms of interactivity (direct / indirect); and input and output information (is it represented and in what direction, e.g. Input → Output). From these points of view we can infer that a representation contains more or less hidden information.

Data Visualisation, Scientific Visualisation, and Information Visualisation

Defining information visualisation is difficult because of the various, contradictory or confusing uses of the term. One of the reasons is the mixed use of the words data and information, as noted before. Data visualisation or data graphics consists of scientific visualisation and information visualisation (Card et al. 1999, see Figure 2). Scientific visualisation is mainly concerned with phenomena that are based on physical world. The data is collected from the earth, buildings, molecules, or other. Information visualisation, in contrast to scientific visualisation, is concerned with visualisation of large volumes of abstract, non-physical data.

Kamada and Kawai (1991) presented one of the early definitions of information visualisation, which is “translation from textual or internal representations into pictorial representations”. Furthermore, they regard the visualisation process as the translation from textual languages into two- or three-dimensional visual languages and called this process "translation into pictures". More recently, Card et al. (1999) define information visualisation as "the use of computer-supported, interactive, visual representations of *abstract* data to amplify cognition". This broader view is more applicable for our purposes, as will be shown later.

There are several classifications and frameworks for visualisation and visual representations. Functional classifications focus on the intended use and purpose of the graphics, whereas structural categories focus on the form of the image (Lohse et al., 1994). The latter is closer to the previously chosen definition of a representation, because it is intended to correspond to an underlying representation in memory.

The framework suggested by Ben Shneiderman (1999) supports research in visualisation in two dimensions. The first dimension is the data-type of the objects to be represented in the interface (Figure 2), and the second is interface functionality, i.e. tasks that can be executed in the visualisation. Those tasks include overview, zoom, filter, details-on-demand, relate, history, and extract (Table 2). According to his observations, Shneiderman proposes an information-seeking mantra, which is "overview first, zoom and filter, then details-on-demand".

Table 2: The main features of a comprehensive information visualisation tool (extended from Shneiderman, 1997).

Feature	Explanation
Overview	A clearly organised overview map is provided; if 3D is used, it should deal effectively with occlusion
Search strategy	Search using keyword(s) or index, browsing, or agent can be used
Zoom and filter	Zoom in on items of interest and filter out uninteresting items (global context retained)
Details-on-demand	The actual information source can be achieved; logic behind the visualisation is shown or explained
Relate (focus + context)	The relationships among items and the place of selected information is shown (in the overview)
Visual layout and structure	Simple and clear; viewer's attention is not drawn from the essential information; both group of objects and single nodes can be browsed
Navigation aids	Landmarks, history, and backtracking facilities are provided; traversal path can be seen if desired
Extract, customisability	Current state of the visualisation and user options can be saved; dynamic links can be created (hypertext functionality)

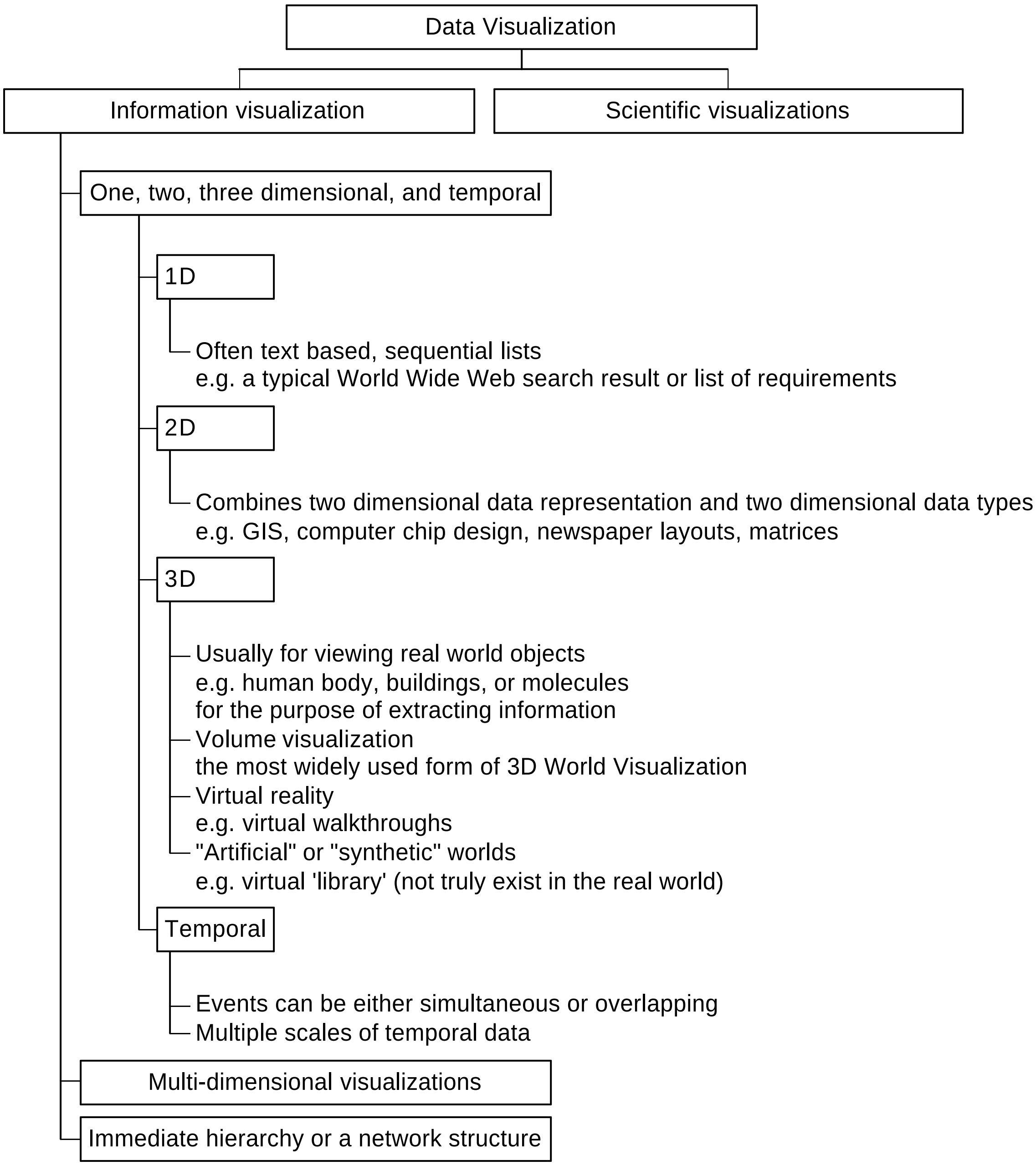

Figure 2. Classification of visualisation (adopted from Shneiderman, 1997).

Research Problem and Methodology

Research Problem

Our research aims towards such CASE environments that would visualise the objects in a design repository effectively. From this, the general research problem of this study is:

How can we improve visualisations of the CASE environments?

Card et al. (1999, p. 23) mention three characteristics for effective mapping from data to visual form: faster to interpret, conveys more distinctions, or leads to fewer errors than other mapping. In our study, we are concentrating on reducing errors, because it is the most important issue during information system development from these three characteristics. Therefore, the previously stated research problem can be reformulated as:

How to utilise visualisation techniques for the CASE environments so that information system models lead to fewer errors than "traditional" ones?

A typical example of a traditional solution is showing diagrams in separate windows. In fact, it is very difficult to show their connections to other design elements. Because current CASE tools lack support for effective visualisations, we are interested in how to improve CASE environments by utilising information visualisation techniques. To answer that question, we need to consider cognitive aspects as well as technical aspects for information visualisation. We are concentrating on a individual designer. Therefore, the following two research questions are asked:

RQ1) What visualisation characteristics should be considered when building support for IS designers?

RQ2) How can advanced visualisation techniques support the user's understanding of complex information spaces? Especially, how can we improve existing CASE tools to facilitate seeing interrelationships between design elements?

Research Methodology

Having identified our research problems, we can now propose a research methodology that directs and describes the way we address those problems. This research applies and refers to many disciplines and research areas. Visualisation, representation, perseptualisation, navigation, and other relevant issues are studied in many fields of science, which all have their own applicable research methods.

Here we are proposing the use of multimethodological research framework (Nunamaker et al., 1991), which has successfully used in many previous studies (e.g. Kelly, 1997, Marttiin, 1998, and Rossi, 1998). The main reason for choosing this framework as a research methodology is motivated by the fact that the research on information system area is still relatively young and thus constructive research is needed. The framework consists of four strategies: observation, theory building, systems development, and experimentation. As can be seen from Figure 3, these strategies may be used in any preferred order.

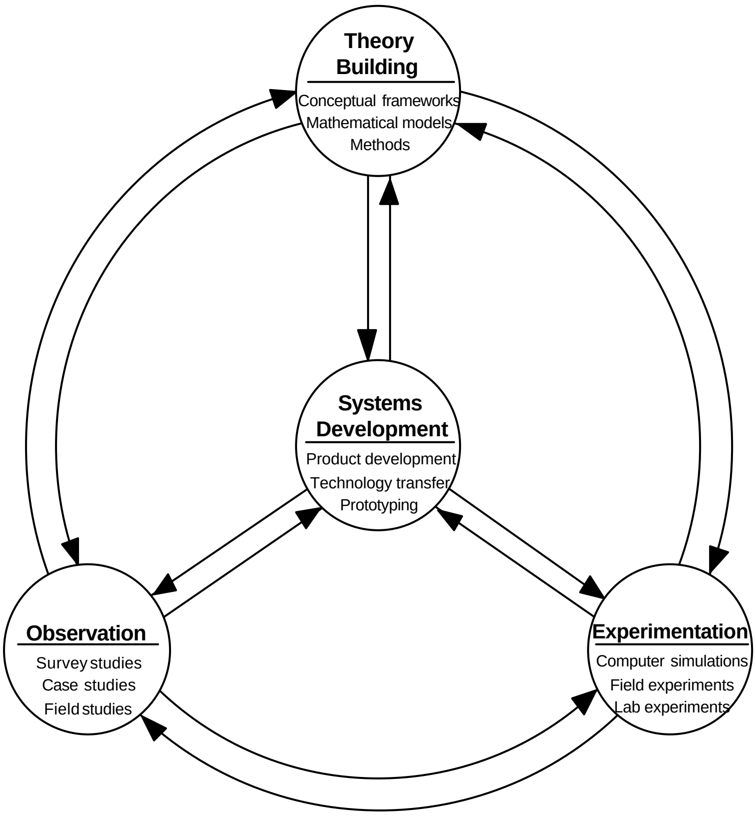

Figure 3. A multi-methodological approach to IS research (Nunamaker et al. 1991, p.94)

Observation includes research methods such as case studies, field studies, and surveys. It is often used when relatively little is known in a research area, to help researches to formulate specific hypotheses to be tested through experimentation, or to arrive generalisations that help focus later research.

Theory building includes the development of new ideas and concepts, and construction of conceptual frameworks, new methods, or models (e.g. mathematical models, simulation models, and data models). Theories may be used to suggest research hypotheses, guide the design of experiments, and conduct systematic observations.

Systems development is a central research approach in the Nunamaker's framework. It consists typically of five stages: concept design, architecture construction, prototyping, product development, and technology transfer.

Experimentation includes research methods such as laboratory and field experiments as well as computer simulations. Results from experimentation may be used to refine theories and improve systems.

Application of the Methodology and Contributions

The research was performed as follows. We started from **observation**, where we looked at the state of the practical and research field of visualisation and CASE (paper 1). Parallel to that, we observed user actions and potential problems faced by an information systems designer (papers 2 and 5). Based on our observations, we developed new ideas and suggested how CASE tools could benefit from advanced visualisation techniques (paper 3). As a part of **theory building**, we also built a framework, which was applied when visualisations were evaluated (paper 2).

As the result of the **system development**, we constructed a research prototype (paper 4). It includes some of the suggested visualisation techniques. Finally, we conducted a laboratory **experiment**. There, implementation was examined by giving specific tasks and recording the number of errors (paper 5).

To validate our findings, a new research prototype will be generated and it's applicability to help information seekers will be examined. The results will be presented in the final paper (6) included in the Ph.D dissertation. In addition, user opinions, our conclusions, and suggestions for the future are expressed in all papers.

Table 3. The relation between research questions, papers, and used research approach.

Question	Paper	Main research approach
1) What visualisation characteristics should be considered when building support for IS designers?	1 and 2	Observation; theory building
RQ2) How can advanced visualisation techniques support the user's understanding of complex information spaces? Especially, how can we improve existing CASE tools to facilitate seeing interrelationships between design elements?	2, 3, 4, 5, and 6	Observation, theory building, system development, and experiment

We propose several contributions. Firstly, a synthesis of the main contributions and research methods in the HCI field, especially cognition related concepts, is created (paper 1). Secondly, a framework for evaluation of visualisation techniques is introduced (paper 2). Thirdly, initial findings and suggestions for using advanced visualisation techniques are put forward (paper 3). Then, a research prototype is created (paper 4) and used in a laboratory experiment in order to examine usefulness of proposed solutions (paper 5). Finally, our ideas are tested in a more real setting by using an improved version of a research prototype (paper 6). Overall, the general idea and potential outcome of the research is to give suggestions to improve understanding of stored design information located in repositories. Our aim is to help seeing and understanding the whole behind the details and keep track when navigating within a system specification.

Related Research

Related Fields

The growing number of books and conferences related to data and information visualisation indicate the raising interest towards the subject. There are not, however, many studies with combination of information system area and visualisation area.

Based on the literature, we can distinguish at least four main sources or levels of complexity before understanding of information spaces can happen, namely information space, information retrieval (IR, the ancestor of data mining), information visualisation, and perception (Fig. 4). We will describe each level in the following, focusing on perception and information visualisation.

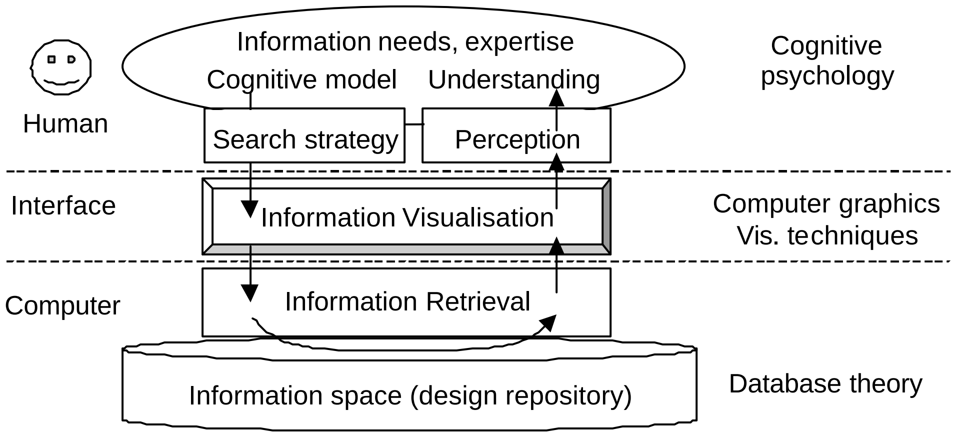

Figure 4: Main levels for information seeking and related research fields.

Information visualisation relates to many fields, from arts to cognitive psychology and neurology, and from databases to computer graphics and further to computer supported cooperative work (CSCW). These areas emphasise different aspects of visualisation: databases include storing and retrieval of data, computer graphics visualising data on the screen, arts proposes aesthetic values, cognitive psychology and neurology dive into inside human brain, and CSCW collaboration aspects. Cognition related research is discussed in more detail in the paper 1.

Research Prototypes and Applications

Although there are a number of research papers and projects, which offer promising solutions, these are rarely commercialised. This implies the difficulty to implement a comprehensive support for information visualisation. Because the research on visualisation techniques and applications are currently advancing rapidly, we do not introduce current solutions here in more detail, we just refer them in appropriate situations. One of our research aims is to select the appropriate solutions and extend their use in CASE.

To begin with, research in metaCASE environments has natural connections to CASE environments. The MetaPHOR group, which I also participated, developed a metaCASE tool called MetaEdit+ and they published over 100 papers in various conferences and journals (see e.g. Lyytinen et al., 1994; Kelly and Rossi, 1997; Tolvanen, 1998). MetaEdit+ can be used also as a CASE tool. Problems and solutions found in metaCASE environments apply often in CASE environments, and vice versa. Therefore, we have reviewed also metaCASE literature extensively.

The second category, which relates to our research, contains hypermedia applications (Garzotto et al., 1996; Oinas-Kukkonen, 1997). These offer solutions for linking parts of (design) documents together. HyperTree (Salampasis et al., 1997) and MICROCOSM (Fountain et al., 1990) are two typical examples.

Research on diagrams or graphs in general has a long history. One of the most acknowledged researchers is Jacques Bertin, who has done extensive work on graphical semiotics and graphic information processing (Bertin, 1983). Mackinley (1986) formalises the Bertin's ideas and proposes the automatic design of graphical presentations of information.

Information system development includes also the generation of code. Software visualisation uses computer graphics and animation to help illustrate and present computer programs, processes, and algorithms (Stasko et al., 1998). As an example, Seesoft (US Patent 5644692) uses focus+context technique for large amounts of source code (Ball and Eick, 1996).

Summary of the Articles

In this Section, I shortly describe the five papers included in the licentiate thesis, the problems addressed, the research methodology used, and results of each. The publication details of the papers and authors are listed for each paper. Note that the order of the papers is logical rather than chronological.

1. Review of HCI Research – Focus on Cognitive Aspects and Used Research Methods

Published in the Proceedings of IRIS 22 "Enterprise Architectures for Virtual Organisations", Vol. 2, Käkölä, T. K. (Ed.), Jyväskylä University Printing House, Keurusselkä, Finland, pp. 435-445, 1999.

Jouni Huotari and Janne Kaipala

This paper investigates the scientific work within the field of Human-Computer Interaction (HCI) with focus on cognitive aspects. Our review analyses and synthesises the main contributions of the field and takes a critical view for some of the influencing methods and theories related to cognition. Moreover, we survey the research methods used in HCI literature dealing with cognition related concepts.

Based on the literature review, HCI is the most distinctive field in IS research that deals with cognitive aspects. We studied the structure of two major IS literature classification systems (ACM and MISQ) and showed the common categories related to HCI. The results indicate that empirical, non-empirical work is almost evenly emphasised, and that most of the empirical research has been experimental. Despite a trend of applying cognitive task analysis and other user-centred system design methods, issues of human cognition and human information processing still need more attention in the IS research.

2. Supporting User's Understanding of Complex Information Spaces by Advanced Visualisation Techniques

Published in the Proceedings of the Eighth Biennial Conference on Artificial Intelligence (STep'98) - Human and Artificial Information Processing, Koikkalainen, P. and Puuronen, S. (Eds.), Finnish Artificial Intelligence Society, Jyväskylä, Finland, pp. 41-50, 1998.

Jouni Huotari

In this paper, we suggest how user’s cognitive capabilities and usability issues should be considered in visual representations. The focus is on supporting user’s understanding of complex information spaces by advanced visualisation techniques.

We conducted a literature review and proposed some suggestions for evaluating novel information visualisation applications and research prototypes. The initial findings are that in recent years the computing power has increased so much that applications using advanced visualisation techniques can be implemented. They support user’s understanding by providing graphical overviews, interaction and efficient search facilities. We show that although there are several novel information visualisation applications and research prototypes available, none of them meets all the suggested evaluation criteria.

3. Towards Advanced Visualisation Techniques in CASE: Initial Findings and Suggestions

Published in the Proceedings of the Seventh International Conference of Information Systems Development (ISD'98), Zupancic, J., Wojtkowski, W., Wojtkowski, W. G. and Wrycza, S. (Eds.), Kluwer Academic / Plenum Publishers, Bled, Slovenia, pp. 742-752, 1998.

Janne Kaipala and Jouni Huotari

We discuss how representations in CASE can be improved using advanced visualisation techniques. While CASE tools allow creating different representations supporting perceptual cues, they have largely ignored the fact that a problem in a design situation can be the amount of irrelevant information. CASE tools also lack overview representations that provide a holistic or focused view on the whole design repository or part of it. We show that advances in research of visualisation can be utilised in solving the above problems. Using a literature review as a research method, we give some suggestions how to improve existing CASE tools to facilitate seeing interrelationships between design elements in a more comprehensible manner.

4. Enhancing Graphical Information System Models with VRML

Published in the Proceedings of the Sixth International Conference on Information Visualisation (IV'02), Williams, A. D. (Ed.), IEEE Computer Society, London, England

Jouni Huotari and Marketta Niemelä

This paper introduces a research prototype, which was developed for a laboratory experiment. Our VRML implementation integrates different types of diagrams in one whole. Three aspects are especially emphasised. Firstly, our solution preserves structure, which can have semantic value. Secondly, it provides both focus and context in order to understand how detailed information relates to other design elements. A third included solution is to enable tracing between diagrams and other design documents. We applied elision technique for decomposition of data flow diagrams (DFD) and added visible lines to link parts of DFD to entity-relation diagrams. In our laboratory experiment, we collected users' subjective opinions and performance in information search tasks. Detailed results are presented in the fifth paper.

5. Improving Graphical Information System Model Use with Elision and Connecting Lines

Submitted to ACM Transactions on Computer Human Interface (TOCHI)

Jouni Huotari, Marketta Niemelä, and Kalle Lyytinen

We applied two visualisation techniques, elision and connecting lines, to resolve the problem of integrating and interpreting design information from different representations. We conducted a laboratory experiment where we compared the impact of used visualisation techniques on search performance (diagram legibility) in designs. We compared paper-based and large screen visualisations while subjects' spatial visualisation ability was measured as a possible covariant. Our results show that visualising relationships explicitly reduces errors in search tasks.

6. (Title will be decided later)

To be submitted to an information visualisation conference (probably IV'04)

Jouni Huotari

Conclusion

Contribution of the Research

The main contribution of the research is to introduce some new visualisation aspects for ISD methods and tools. Our general interest was towards CASE environments, which offer effective visualisations, i.e. they are fast to interpret, rich in distinctions, and decrease the number of errors. We started our studies by examining literature and current applications focusing on cognitive and information visualisation aspects. We found that issues of human cognition and human information processing still need more attention in the IS research (paper 1). We also build evaluation criteria for applications that visualise information (RQ1, paper 2). This criterion includes features such as overview map, support for different search strategies, interaction, focus+context, visual layout and structure, natural metaphor and navigation aids, and customisability. According to our evaluation, current applications do not utilise possible visualisation solutions comprehensively. We also found that current CASE tools and environments still fall short in representing design information, especially showing interrelationships between design elements (papers 3 and 4). We give suggestions how to improve representation in CASE tools and in general (RQ2).

Basing on these observations, we started to build a research prototype as a possible solution in representing multifaceted design information (paper 4). We used this prototype in a laboratory experiment and found out that it really helps people to see connections between different diagrams and between different levels of detail within one diagram (RQ3, papers 4 and 5). Thus, we will still need more studies (paper 6) to find out how to support user's understanding of complex information spaces.

We agree the contingency theoretical view (Katz, 1984), which emphasises selecting the most appropriate actions for specific situations. According to this view, there is no one best visualisation for all situations, and all visualisations are not equally good in all situations (but some visualisations are better than others). Optimal visualisation is produced quickly and cost-efficiently. The external representations must fit / match with the designers' internal representation, and should end up with shared understanding, thus facilitating communication between all stakeholders.

Our research has also some contributions to practice. Based on the statistical analysis and comments of the subjects in our laboratory experiment, our implementations helps seeing interrelationships between different diagrams and between different levels of detail within one diagram. Many novice designers mentioned that they did not understand the meaning of consistency and actual idea of data flow and ER diagrams until they saw those diagrams side by side and the corresponding objects were visualised explicitly. We thus argue that our implementation helps understanding graphical information system models and is applicable in teaching.

Limitations of this Study

We recognise many limitations, such as methodological and technical ones. The chosen research method did not contain qualitative research methods such as action research or interviews. In addition, although thorough understanding of information visualisation requires also understanding the mechanisms of the processes in our bodies, it was beyond our knowledge and resources to measure brain activities or other physiological things. Therefore, some insightful knowledge was not gained.

In our experiment, we were only examining novice designers. We do not know how practical our suggestions are in real use. In addition, our technical solution was a research prototype and as such, we could not include all possible improvements in that one application.

Directions for Further Research

In this study, we have only scratched the surface of information visualisation in CASE environments. There are still many elements to be discovered in relation to human vision and how to transform abstract data into visual-spatial forms (Chen, 2002). In addition, questions related to the concept of information itself need to be answered, such as

- Nature or type of information; what kind of information exists and how to visualise it?
- Properties of information; what visualisation characteristics to attach?
- Structure of information; what visualisation technique is the most appropriate?
- Information retrieval; clustering of information

We have many ideas how to continue our research. One direction is towards better technical support for visualisation, for example to show what parts are related to the selected object (e.g. by blinking). We could utilise more augmented reality, e.g. embedding virtual information in the physical world by using see-through displays, real 3D (where no special glasses would be needed), or by using wearable or mobile user interfaces that would make information available anytime and anywhere.

We see a future, where all system components are reachable (linked together) and visualised in a meaningful way, including visualisation of all documents and design decisions during the design process, and showing the development from the idea to the product. This could be called as visualisation-in-the-large. Lastly, we predict to see ideas from data visualisation to mix with information visualisation, e.g. effective and efficient clustering algorithms for large high-dimensional data sets with high noise level (Keim and Hinneburg, 1999). In all, we are just taking our first steps towards comprehensive information visualisation within CASE environments.

References

Alderson, A. 1991. Meta-CASE Technology, In Software Development Environments and CASE Technology, Proceedings of European Symposium, Königswinter, June 17-19, , Endres, A. and Weber, H. (Eds.), Springer-Verlag, Berlin, pp. 81-91.

Ball, T. A. and Eick, S. G. 1996. Software visualization in the large, IEEE Computer, 29, pp. 33-43.

Bertin, J. 1983. Semiology of Graphics, University of Wisconsin Press, Milwaukee, Wis.

Booch, G., Jacobson, I. and Rumbaugh, J. 1998. The Unified Modeling Language User Guide, Addison-Wesley, Reading Mass.

Bruner, J. S. 1957. On perceptual readiness, Psychological Review, 64, pp. 123-152.

Bubenko, J. A. 1988. A Method Engineering Approach to Information Systems Development, the proceedings of the IFIP WG8.1 Working Conference on Information Systems Development Process, Vol. , Prakash, N., Rolland, C. and Pernici, B. (Eds.), Elsevier Science Publishers, Amsterdam, pp. 167-186.

Card, S. K., Mackinlay, J. D. and Shneiderman, B. 1999. Readings in information visualization: using vision to think, Morgan Kaufmann Publishers, Inc., San Francisco.

Chen, C. 2000. Individual Differences in a Spatial-Semantic Virtual Environment, Journal of the American Society for Information Science, 51, pp. 529-542.

Chen, C. 2002. Information Visualization, Information Visualization, 1, pp. 1-4.

Ewusi-Mensah, K. 1997. Critical Issues in ABANDONED Information Systems Development Projects, Communications of the ACM, 40, pp. 74-80.

Falkenberg, E. D., Hessa, W., Lindgreen, P., Nilsson, B., Oei, J. L. H., Rolland, C., Stamper, R. K., Assche, F. J. M. V., Verrijn-Stuart, A. A. and Voss, K. 1998. A Framework of Information System Concepts: The FRISCO Report (Web edition)., , The IFIP WG 8.1 Task Group, Leiden.

Fountain, A. M., Hall, W., Heath, I. and Davis, H. C. 1990. MICROCOSM: An Open Model for Hypermedia With Dynamic Linking, Hypertext: Concepts, Systems and Applications, Vol. , A. Rizk, N. S. a. J. A. (Ed.), Cambridge University Press, INRIA, France.

Garzotto, F., Mainetti, L. and Paolini, P. 1996. Navigation in Hypermedia Applications: Modelling and Semantics, Journal of Organizational Computing and Electronic Commerce, 6, pp. 211-237.

Gibson, J. J. 1979. An Ecological Approach to Visual Perception, Houghton Mifflin, Boston.

Hirschheim, R., Klein, H. and Lyytinen, K. 1995. Information Systems Development and Data Modeling, Conceptual and Philosophical Foundations, Cambridge University Press,

Iivari, J. 1989. Levels of abstraction as a conceptual framework for an information system, In Information System Concepts: An In-depth Analysis, , Falkenberg, E. D. and Lindgreen, P. (Eds.), Amsterdam North-Holland, pp. 323 - 352.

Iivari, J. 1994. Object-oriented information systems analysis: A comparison of six object-oriented analysis methods, In Methods and Associated Tools for the Information Systems Life Cycle (A-55), , Verrijn-Stuart, A. A. and Olle, T. W. (Eds.), Elsevier Science B.V. (North-Holland), pp. 85-110.

Jarzabek, S. and Huang, R. 1998. The Case for User-Centered CASE Tools, Communications of the ACM, 41, pp. 93-99.

Kamada, T. and Kawai, S. 1991. A General Framework for Visualizing Abstract Objects and Relations, ACM Transactions on Graphics, 10, pp. 1-39.

Katz, R. H. 1984. Transaction Management in the Design Environment, In New Applications of Databases, , Ge, G. G. a. E. (Ed.), Academic Press, London UK.

Keim, D. A. and Hinneburg, A. 1999. Clustering Techniques for Large Data Sets (Tutorial), From the Past to the Future. SIGKDD Conference, Vol. , San Diego.

Kelly, S. 1997. Towards a Comprehensive MetaCASE and CAME Environment: Conceptual, Architectural, Functional and Usability Advances in MetaEdit+, , University of Jyväskylä, Jyväskylä, pp. 267.

Kelly, S. and Rossi, M. 1997. Differences in Method Engineering Performance with Graphical and Matrix Tools: A Preliminary Empirical Study, In Proceedings of 2nd CAiSE/IFIP8.1 International Workshop on Evaluation of Modeling Methods in Systems Analysis and Design, EMMSAD'97, June 16--17, Barcelona, Spain, , Siau, K., Wand, Y. and Parsons, J. (Eds.).

Kelly, S. and Smolander, K. 1996. Evolution and Issues in MetaCASE, Information and Software Technology, 38, pp. 261-266.

Kuutti, K. and Bannon, L. J. 1993. Searching for unity among diversity: exploring the "interface" concept, Human factors in computing systems, Vol. , Amsterdam, pp. 263-268.

Lohse, G. A., Biolsi, K., Walker, N. and H.H.Rueter. 1994. A Classification of Visual Representations, Communications of the ACM, 37, pp. 36-49.

Lyytinen, K. 1987. A Taxonomic Perspective of Information Systems Development: Theoretical Constructs and Recommendations, In Critical Issues in Information Systems Research, , Jr., R. J. B. and Hirschheim, R. A. (Eds.), John Wiley & Sons Ltd., pp. 3--41.

Lyytinen, K., Kerola, P., Kaipala, J., Kelly, S., Lehto, J., Liu, H., Marttiin, P., Oinas-Kukkonen, H., Pirhonen, J., Rossi, M., Smolander, K., Tahvanainen, V.-P. and Tolvanen, J.-P. 1994. MetaPHOR: Metamodelling, Principles, Hypertext, Objects and Repositories., , Department of Computer Science and Information Systems, University of Jyväskylä, Finland.

Mackinley, J. D. 1986. Automating the Design of Graphical Presentations of Relational Information, ACM Transactions of Graphics, 5, pp. 110-141.

Marr, D. 1982. Vision: A computational investigation into the human representation and processing of visual information, W.H. Freeman, San Francisco.

Marttiin, P. 1998. Customisable Process Modelling Support and Tools for Design Environment, , University of Jyväskylä, Jyväskylä, pp. 238.

McCormick, B. H., DeFanti, T.A. and Brown, M.D. 1987. Visualisation in Scientific Computing - a Synopsis, IEEE Computer Applied Graphics, 7, pp. 61-70.

Monarchi, D. E. and Puhr, G. I. 1992. A Research Typology for Object-Oriented Analysis and Design, Communications of the ACM, 35, pp. 35-47.

Noik, E. G. 1994. A space of presentation emphasis techniques for visualizing graphs, Graphics Interface'94, Banff, Alberta, pp. 225-233.

Nunamaker, J. F., Chen, M. and Purdin, T. D. M. 1991. Systems Development in Information Systems Research, Management Information Systems, 7, pp. 89-106.

Oinas-Kukkonen, H. 1997. Towards Greater Flexibility in Software Design Systems through Hypermedia Functionality, Information & Software Technology, 39, pp. 391-397.

Rossi, M. 1998. Advanced Computer Support for Method Engineering: Implementation of CAME Environment in MetaEdit+, , University of Jyväskylä, Jyväskylä, pp. 215.

Salampasis, M., Tait, J. and Hardy, C. 1997. HyperTree: A Structural Approach to Web

Authoring, Software - Practice and Experience, 27, pp. 1411-1426.

Shneiderman, B. 1997. Designing The User Interface, Third Edition. Strategies for Effective Human-Computer Interaction, Addison Wesley Longman,

Shu, N. C. 1988. Visual programming, Van Nostrand Reinhold Company Inc., New York.

Smolander, K., Lyytinen, K., Tahvanainen, V.-P. and Marttiin, P. 1991. MetaEdit --- A Flexible Graphical Environment for Methodology Modelling, In Advanced Information Systems Engineering, Proceedings of the Third International Conference CAiSE'91, Trondheim, Norway, May 1991, , Andersen, R., jr., J. A. B. and Solvberg, A. (Eds.), Springer-Verlag, Berlin, pp. 168--193.

Spence, R. 1993. A taxonomy of graphical presentation, , Imperial College of Science.

Stasko, J. T., Domingue, J. B., Brown, M. H. and Price, B. A. (Eds.). 1998. Software Visualization, The MIT Press, Cambridge, Ma.

Tolvanen, J.-P. 1998. Incremental Method Engineering with Modeling Tools: Theoretical Principles and Empirical Evidence., , University of Jyväskylä.

Tufte, E. R. 1983. The Visual Display of Quantitative Information, Graphics Press, Cheshire, CT.

Tufte, E. R. 1990. Envisioning Information, Graphics Press, Cheshire, CT.

Tweedie, L. 1997. Characterizing Interactive Externalizations, CHI'97, Vol. , ACM, pp. 375-383.

Wand, Y. 1996. Anchoring Data Quality Dimensions in Ontological Foundations, Communications of the ACM, 39, pp. 86-95.

Ware, C. 2000. Information Visualization: Perception for Design, Morgan Kaufmann Publishers, San Francisco.

Ware, C. and Franck, G. 1996. Evaluating stereo and motion cues for visualizing information nets in three dimensions, ACM Transactions on Graphics, 15, pp. 121-140.

Welke, R. J. 1983. IS/DSS DBMS Support For Information Systems Development, Database Management, Vol. , C. W. Holsapple, A. B. W., Reidel Publishing Company (Ed.), pp. 195-250.

Vessey, I. and Glass, R. 1998. STRONG Vs. Weak Approaches to Systems Development, Communications of the ACM, 41.

Williams, J. G., Sochats, K. M. and Morse, E. 1995. Visualization, In Annual Review of Information Science and Technology (ARIST), 30, pp. 161-207.

2. SCHEDULE FOR THE REMAINING RESEARCH (AIMING AT PHD)

Task#	Duration	Work	Subject
Year 2003: developing a new research prototype			
1.	January - May	1 pwm	Planning; interviews: problems with designer's understanding of a design repository:
2.	March - May	1 pwm	Review of the previous research; focus on the cognitive aspects, used methods, and techniques
3.	April - July	4 pwm	Implementing an application, which supports user's understanding of complex information spaces by advanced visualisation techniques
4.	August	2 pwm	Pilot testing, interviews, observations
5.	June - August	2 pwm	Initial findings and suggestions (when using advanced visualisation techniques in CASE)
6.	September	1 pwm	Technical description
7.	October - December	2 pwm	Reporting of the outcomes
Year 2004: writing the final paper and completing the dissertation			
8.	January - May	1 pwm	Writing the final paper
9.	March - June	2 pwm	Writing the introduction
10.	June	1 pww	Language checking
11.	July	½ pwm	Conference presentation
12.	October-December	1 pwm	Defence

3. FUNDING

For hiring of research assistants, conference fees, language checking, and preparing paper(s), dissertation, and the defence, we need funding as follows:

Year	Purpose	Work	Estimated cost
2003	Research assistants (developing the research prototype etc.)	12½ pwm	10 600 €
2003	Administrative work	½ pwm	600 €
2004	Writing the final paper, conference presentation, and completing the dissertation	3½ pwm	4000 €
2004	Language checking	1 pww	1000 €
2004	Defence	1 pwm	3000 €