Collaboratively Analyzing Large Data Sets using Multitouch Surfaces

Judith M. Brown
Dept. of Psychology
Carleton University
Ottawa, Canada
mmjbrown@gmail.com

Robert Biddle
Institute of Cognitive Science
Carleton University
Ottawa, Canada
robert_biddle@carleton.ca

Stevenson Gossage
School of Computer Science
Carleton University
Ottawa, Canada
estemenson@gmail.com

Jeff Wilson
School of Computer Science
Carleton University
Ottawa, Canada
rjp.wilson@gmail.com

Steven Greenspan
CA Labs
CA Technologies
Islandia, New York
Steven.Greenspan@ca.com

CSCW'12, February 11–15, 2012, Seattle, Washington, USA.
ACM 978-1-4503-0556-3/12/02.

Abstract
The analysis of large data sets is increasingly collaborative, multidisciplinary and even distributed. There are many advantages including numerous checks and balances on the results. But, even in highly distributed analytic tasks, it would be useful to concurrently support co-located analyses by groups or teams. Large multi-touch surfaces present opportunities for presenting large amounts of data. They also naturally support the co-analysis of this data by co-located, multi-disciplinary teams of analysts or groups of people. We present a field study of multi-disciplinary collaboration that exposes some physical and behavioral aspects of multi-disciplinary collaboration. Then we describe our current work on supporting the co-analysis of large data sets in data centres where we are addressing psychological issues pertaining to user interface design.

Author Keywords
Collaboration, multi-touch surfaces, large data sets

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

General Terms
Design, Human Factors

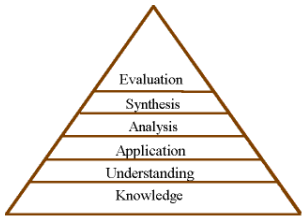

Figure 1: A Model: This model of learning from Benjamin Bloom helped to guide the design of a learning game. The model ensured a variety of learning was supported in the game such as knowledge-based learning and analytic skills. Mental models also guide the work of analysts.

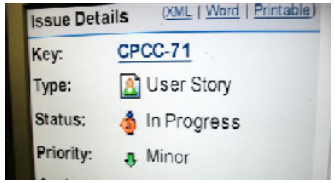

Figure 2: A Plan: Work organized around stories. Large analysis tasks are enabled by plans.

Introduction

Data analysis is a common task in many domains, including science and engineering, and in many workplaces ranging from business environments to data centres. Increasingly due to larger data sets and the requirement for more complex analyses, this activity is becoming a multi-disciplinary collaborative endeavor. Occasionally such an analysis can be distributed across the globe. We address the issue of co-analysis by teams and our research helps to generate more understanding and ideas that help with the design of tools for this new form of analysis activity. We briefly describe our research on multidisciplinary collaborative analysis (and design) in the software domain and the insights that shed some understanding on the physical and behavioral aspects of analysis tasks. We then report on our research on the issues around the use of large multi-touch surface technologies in data centres to analyze large amounts of data.

Research on Multidisciplinary Collaboration

We conducted an empirical study of multidisciplinary face-to-face collaboration in 8 experienced software teams that were analyzing the requirements for (and designing) challenging interfaces for software products. We focused on interactions between two members of the team: 1) the software developer, whose background is usually computer science, and 2) the interaction designer, whose background is varied.

For each of the teams, we collected interview and demographic data from one designer, one developer and a team leaders. At four of the organizations we also observed, videoed, and took photos. We used a socio-cultural framework [1] to focus data collection and grounded theory [2], a well-established qualitative analysis method, to analyze our data [3]. The end result of our analysis was a categorization scheme of the collaborative contexts and artifacts that supported collaborative multidisciplinary events and a thick description of collaborative behaviors [3].

We found that analysis and design work in the software domain was a combination of individual and collaborative work. The collaborative work served to clarify product and project aims, to plan out future work, and to shift and realign the individual activities of designers and developers, whose work was highly intertwined. We found that multi-disciplinary collaborative work of a similar nature occurred in three contexts. 1) In scheduled collaborations in prearranged locations that often involved most of the team members. 2) In impromptu collaborations involving several colleagues in convenient team rooms, hallways or offices. 3) In work-related desk-side chats. The proportion of each type of work in each workplace varied with the organizational culture and the physical aspects of the workplace setting. We also found that artifact use (such as those depicted in Figure 1 and 2) was ubiquitous in collaborative work. Some of the artifacts directly depicted what was being analyzed (the context of use of the future software product, the end users, their tasks and so on), others were outcomes of the analysis (such as sketches of interfaces), but there were many other supporting artefacts that aided analysis work such as models and plans. Overall, artefacts served to reveal, direct and constrain the analysis and design work. Extrapolating from our context, some artifacts that may be found in analysis work of large data sets are summared in Figure 3

Our field study shows that to support collaborative behaviors, systems that support multi-disciplinary collaborative work must be highly flexible.

Artifact Category	Description
Data	Input stream.
Visualization	Output of the analysis task.
Idea	Suggestions for moving the task ahead.
Question	A potentially challenging inquiry.
Exemplar	Another visualization that is like the visualization being created in some way.
Model	A mental model of knowledge, a thing, or a process or a representation of such.
Plan	A statement about intended future work.
Stipulation	A comprehensive description of something, e.g., The data or purpose of the analysis.

Figure 3: An artifact categorization scheme for collaborative work. Some artifacts that are of potential use in co-analysis work.

Representations of data of many types must move easily between personal and shared computers so that data can be displayed on large or small displays as the situation demands. Workspaces must be explicitly designed to accommodate a variety of meeting and display types. Although our software analysts were not analyzing large data sets, this study gives a flavour of analysis activity. It shows that field studies are an essential element of design work because they provide real insight into how individual and collaborative work is intertwined and split, where and how it is carried out, and the actual artifacts that support collaboration, especially those beyond the obvious input and output artifacts of analysis.

Collaborative Data Analysis Using Multi-touch Surfaces

In our current project designing multi-touch applications for data centre analysts we are preparing to conduct field studies. We are also consulting with our three industry partners to find out more about the nature of data centre work and conducting literature reviews to understand the opportunities and limitations of various technologies. One of our industry partners is interested in supporting the collaborative co-analysis of large databases using large multi-touch displays. Our preliminary literature review has shown some promising beginnings on research about using large displays for analysis work. WeSpace, [4, 5] a multi-surface system, was designed for scientists to explore and visualize information and support co-located collaboration. Scientists bring their laptops to a dedicated meeting space and link to a server so that the images on their laptops are shared on a 5-foot, high-resolution display. The scientist's laptops or a multi-touch table act as an input device so that the scientists can 'pin' items (i.e. parts of images) and then 'link' them. The system was trialled with a small team of 4 astrophysicists. The system however lacked many functions required of analysts since viewing and not transforming the data was emphasized. Lark was designed so co-collaborators work on transforming database information on a shared multi-touch tabletop display by scaling, filtering, charting, graphing or colouring various representations of the data [6]. A small lab study with biologists was conducted of this system. However, Lark does not allow analysts to work on parts of the data independently and does not scale well to larger numbers of analysts.

In our work, we are attending to perceptual and communication issues that arise when people and displays

are distributed across spaces, such as data centres. One issue is perceiving the data. High resolution displays help [4, 5] as do tools to magnify portions of the data, and flexible movement or duplication of data between low resolution surfaces, typically projections on walls, to high resolution surfaces, such as tablets [7]. Another issue is that distributed systems of displays require support for deictic language communication through pointing aids. Pointing aids allow analysts to select items on a display or to operate a tool; they also indirectly let other analysts know where an analyst's attention is directed. Nacenta compared different mechanisms for pointing [8] and concluded they have a large impact on task performance. A final issue relates to attention. When data is distributed across large spaces it is useful to know what others are attending to. Pinelle et al. have explored using pointers to indicate awareness [9] and we are exploring using head-and eye-tracking systems with multi-touch displays to indicate attention.

Our system for supporting the analysis of large databases in data centres would involve tablets, smartphones and large high-resolution displays, support for displaying large data sets at resolutions appropriate for human *perception*, easy transfer of information between display screens, tools to support *analysis* work, support for *deictic language* through pointing aids, and mechanisms to indicate where the *attention* of fellow collaborators is directed.

Authors and Acknowledgment

Robert Biddle is co-theme leader for NSERC SurfNet a Digital Surface Software Application Network of researchers, government and industry partners. Brown, Gossage and Wilson are funded through SurfNet. Steven Greenspan is a research staff member at CA Labs, an industry partner.

References

[1] H. Daniels, *Vygotsky and Research*. Routledge, 2008.

[2] A. L. Strauss and J. M. Corbin, *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage, 1998.

[3] J. Brown, G. Lindgaard, and R. Biddle, "Collaborative events and shared artefacts: Agile interaction designers and developers working toward common aims," in *AGILE Conference*, Aug. 2011, pp. 87 –96.

[4] D. Wigdor, H. Jiang, C. Forlines, M. Borkin, and C. Shen, "WeSpace: The design development and deployment of a walk-up and share multi-surface visual collaboration system," in *Proceedings of CHI '09*, 2009, pp. 1237–1246.

[5] H. Jiang, D. Wigdor, C. Forlines, and C. Shen, "System design for the WeSpace: Linking personal devices to a table-centered multi-user, multi-surface environment," in *TABLETOP*, 2008, pp. 97 –104.

[6] M. Tobiasz, P. Isenberg, and S. Carpendale, "Lark: Coordinating co-located collaboration with information visualization," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 15, no. 6, pp. 1065 –1072, 2009.

[7] K. Everitt, C. Shen, K. Ryall, and C. Forlines, "Docubits and containers: Providing e-document micro-mobility in walk-up interactive tabletop environment," in *Proceedings of Interact 2005*, 2005.

[8] M. A. Nacenta, D. Pinelle, D. Stuckel, and C. Gutwin, "The effects of interaction technique on coordination in tabletop groupware," in *Proceedings of Graphics Interface 2007*. ACM, 2007, pp. 191–198.

[9] D. Pinelle, M. Nacenta, C. Gutwin, and T. Stach, "The effects of co-present embodiments on awareness and collaboration in tabletop groupware," in *Proceedings of graphics interface 2008*. Canadian Information Processing Society, 2008, pp. 1–8.