The role of direct manipulation of visualizations in the development and use of multi-level knowledge models

Randy Goebel, Wei Shi, Yuzuru Tanaka
Department of Computing Science, University of Alberta, Edmonton, Alberta T6G 2E8, Canada,
MemeMedia Laboratory Hokkaido University N 13 W 8, Kita-ku, Sapporo 060- 8628
rgoebel@ualberta.ca, shiwei@meme.hokudai.ac.jp, tanaka@meme.hokudai.jp

Abstract

The proliferation of touch sensitive display screens has created a new generation of human-computer interaction styles which are so natural and common that even the youngest of users now perceive ordinary static media like a glossy magazine as a broken iPad. The volume of users who expect to be able to pinch, grab, twist and manipulate images on screen is rapidly growing; they drive a renewed interest in developing, assessing, and delivering new direct manipulation systems. Our premise is that one can exploit new technologies to develop new repertoires of direct manipulation, but with increasing pressure to provide semantically-coupled direct manipulation methods to experiment with computational information models.

We develop this premise by noting highlights in the evolution of direct manipulation interfaces, and suggest that their selection and deployment can be tailored as visual experiments to debug and extend more complex computational models of information systems and processes. These systems and processes include those of natural systems such as arise in systems biology (e.g., modelling multiple levels of protein structure), but also in "unnatural" systems such as in the identification of hubs and authorities in artificial systems like the World Wide Web (WWW).

The immediate consequence of our premise suggests that the design of direct manipulation tools should proceed with the semantics of the modelled systems in mind, so that each users' manipulations provide a new perspective on the concept of "data mining" of large data sets. This will allow users to not just expose implicit relationships, but to incrementally combine explanatory and exploratory investigation by direct manipulation, to adjust and improve the computational knowledge models that emerge from the underlying data.

Keywords— **visualization, direct manipulation, explanatory inference, exploratory inference, multi-level modelling.**

1 Introduction

A recent Youtube posting helps make an easy case for the intrinsic value of direct manipulation within the scope of human computer interaction [22]. The annotated 1:36 length video clip shows a one year old girl who, after a brief experience with an iPad, expects a glossy magazine to behave in a similar fashion. Perhaps the most intriguing part of the video is how the baby seems to conduct a test of whether her finger is still "working" by pressing it to her leg, and then again trying her finger on the magazine. The finger seems ok, but the magazine is "broken." An earlier posting of the same video was without the annotation in that cited here (e.g., the explicit annotation "Yet my finger is working" suggests an experiment, but that may be a bit of an overstatement). While the earlier posting leaves more to the viewer's interpretation, it is still clear that direct manipulation is easily conditioned and quickly expected to be true of all flat glossy objects.

With this observation as background, our goal is to consider some highlights in the history of direct manipulation, and to promote the premise that increasingly sophisticated direct manipulation actions must *necessarily* be designed to both respect and even help articulate the semantics of the underlying data models of the objects they manipulate. So while grab and pinch might well and naturally adjust the size and orientation of a static image interpreted as a photograph, manipulation actions that select and identify objects in a photo, like a face or a tree, must *necessarily* encode the semantics of what it means to be a face or a tree.

This simple premise is neither naive nor profound. It is not profound, because the current generation of human computer interface techniques have for some time provided direct manipulation techniques that encode the notion of a graphical model that one can inspect and adjust (e.g., consider the variety of techniques used for interaction with virtual museums [15, 20]).

But it is also *not* naive, because the growing volume and complexity of accumulating data requires much more

1550-6037/13 $26.00 © 2013 IEEE
DOI 10.1109/IV.2013.95

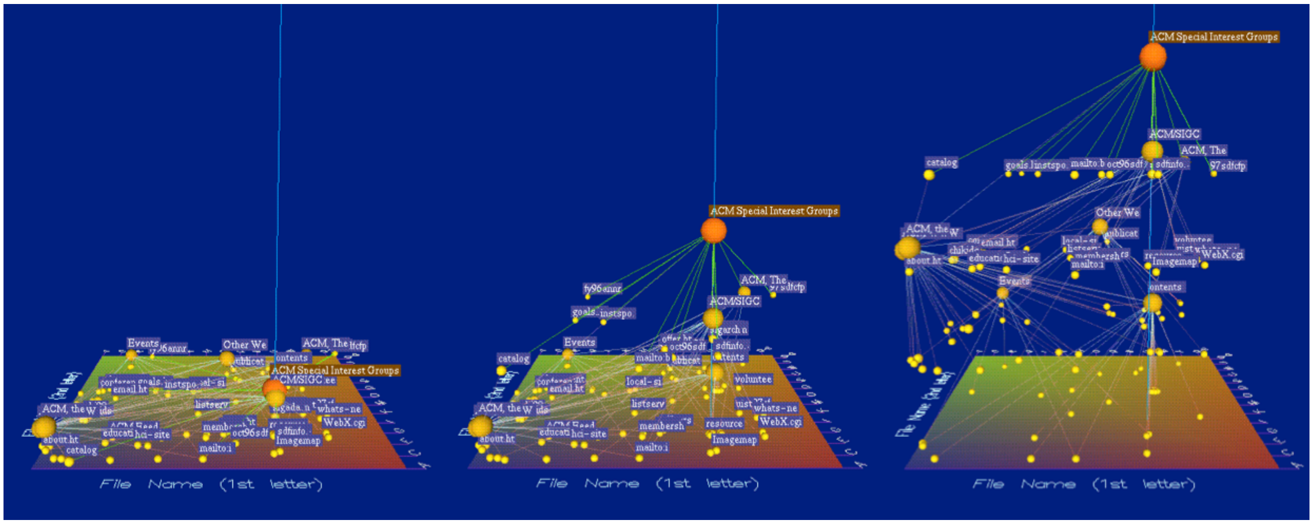

Figure 1: The lift up operation of Natto in three frames.

sophisticated knowledge models, and therefore similarly more sophisticated direct manipulation techniques. Despite the work on classifying a repertoire of interaction techniques into informal semantic categories (e.g., [25], it remains the case that nearly every example of direct manipulation systems are individually engineered for a relatively brittle scope of domain specific application. Alternatively, the example of the Foldit protein folding game is an early illustration, where a relatively sophisticated protein dynamics model underlies the use of a touch screen manipulation of example proteins. The knowledge model for proteins constrains the game players repertoire of actions, with the intent of discovering new unconsidered possible protein foldings [5].

To further push the motivation for the need for increased semantic sophistication in the design of direct manipulation of complex knowledge models, consider the anticipated data capture rates of ≈864,000Gb per day of the radio astronomy square kilometre array project [19]. Direct manipulation interfaces for that project will have to encode the semantics of manipulations like "adjust cosmic magnetism" or "deflect the magnetic field around a black hole." While similar in spirit to using existing touch interfaces to grab faces and trees (rather than regions on a flat image), it is clear that the intrinsic value of direct manipulation will have to achieve semantic sophistication commensurate with the complexity of the computational models being manipulated.

Note that it is *not* a premise of our position that no good or useful work has been done in systems which support the coupling of direct manipulation of visualization. Rather it *is* fundamental that the accumulation of data is so rapid, that only machine learning techniques will enable the construction of the kinds of multi-scale models that support coherent direct manipulation (e.g., see [26]).

The rest of this brief paper is organized as follows. Section 2 provides some highlights in the evolution of direct manipulation interfaces, and suggests that much of the development of direct manipulation interfaces has been driven by semantic-based intent. The core of the argument for semantics-based direct manipulation is in Sections 3 and 4, where we first provide some examples where the semantics of both natural and unnatural systems required semantic manipulation techniques to confirm the explanatory power of interaction. Then Section 4 provides some insight into the challenges arising when direct manipulation provides the basis for exposing anomalies in existing knowledge models, and so provides exploratory power that can debug those models. We conclude with Section 5, with the briefest of summaries and our anticipation of much work to be done.

2 A Few Highlights in the History of Direct Manipulation

Direct manipulation of data representations, whether the representation is visual or not, is an old idea. Even software pan and tilt controls within graphics systems are a form of direct manipulation of visualizations, as are the kinds of display manipulation controls used in modern modelling software (e.g., [21]) or as the basis for scheduling software [6]. Similarly, the idea of computational steering within resource intensive scientific simulation is now commonly providing user interaction to help reduce resource load and more quickly find desired simulation results (e.g., see [14]). It is clear that the intended manipulations are developed for understanding the content and pro-

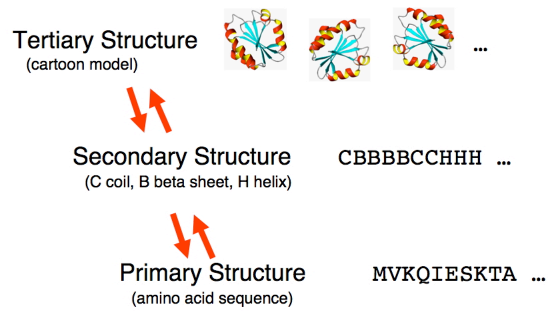

Figure 2: An abstraction of three levels of protein representation

cesses of the underlying data. Indeed, extensive work on classifying a repertoire of interaction categories has made a fundamental contribution to understanding both what is possible and what is difficult about direct manipulation of visualizations of data (e.g., [25]).

Much of the development of direct manipulation has arisen from demand for technical innovation in order to replace other methods (e.g., a mouse). But this demand has not only been motivated only by "pushing" technology, but by the "pull" desire to augment human intellect [8], or to improve what we could label the "semantic coupling" between a human and a control system (e.g., [3]).

Now the most common direct manipulation technologies — touch sensitive screens — offer little in the way of support for the semantics of visualization, beyond what we might describe as syntactic stretch-shrink-rotate operations that are meaningful only with respect to visualizations considered as flat two dimensional images. However, direct manipulation of visualizations is a feasible conceptual approach to the so-called "top 10 unsolved problems of visualization," which are largely about how humans can confirm their interpretation of a broad variety of information visualization schemes: in a phrase, to confirm the semantics of any visualization [4]. The salient point is that any repertoire of direct manipulation actions must be constrained by the semantics of the underlying information, not merely by the technological rendering device. Note, for example, that direct manipulation in visually immersive 3D environments requires an explicit coupling between the third dimension of depth and some attribute of the underlying data (e.g., [1]).

To provide just a glimpse of the trajectory of using direct manipulation to broach the visualization semantics problem, consider a simple first step from the relatively shallow touch screen manipulation of images to the "lift-up" tool of Natto [18]. In this domain of WWW pages and their hyperlink connections, Figure 1 shows a three frame static image sequence to depict a direct manipulation operation that helps a user understand web page connectivity. In the online dynamic interaction with Natto, the representation of webpage connectivity is presented in a third dimension, so that when a single page (node) is grabbed and pulled, the depth of connectivity with other webpages becomes obvious.

There are several important points here. First, the lift-up operation helps expose the connectivity amongst web pages, so that a user gets a sense of the connection density amongst pages. One learns to expect that grabbing and pulling a page will reveal some idea of the connection density for that page. Second, and perhaps more importantly, the user *expects* the connectivity shown to be accurate, so that one is *exploring* an accurate presentation of the connectivity in a graphical way, rather than finding and counting hyperlink annotations in an HTML document. And thirdly, to anticipate the transition from *explanation* to *exploration*, the user has very little chance of finding errors or bugs in the visual presentation of the connectivity, because the volume of pages and links in the actual HTML space are too numerous to consider.

So it is this kind of obvious inference that is facilitated by direct manipulation. But the accuracy of the inferences drawn from that direct manipulation depends on the precision of the relationship between the visual domain and the knowledge model framework for the data domain. And the semantics of those relationships for the user are only accurate when the direct manipulation actions respect those relationships.

3 Direction Manipulation and Model Extending

In the case of building or extending scientific models of large volumes of data, the challenge is to create general principles based on observing specific relations amongst base data. So, in the classical inductive scenario, observing that the nth of a sequence of white swans is actually black, one confirms the property that not all swans are white. Observing a swan's colour can he considered as the simplest possible direct "manipulation" of the data to make some inference. Consequently, any direct manipulation action on a world-wide swan colour database would have to respect that principle, and not allow a contravening manipulation.

But as the world swan colour data count accumulates, a manipulation could be used to observe trends in swan colour, and not just check each one, but suggest new principles to help structure the data. For example, one might notice, if provided with a manipulation that check swan colours on a Google map mashup, that many more swans from a certain place are black — voilà, a new structuring principle for the swan colour data base. And a new contribution to the model.

So there is some value and role of semantically defined direct manipulation especially as related to the explanatory aspect of data mining, or in deeper understanding of complex models and the data that represent them. In fact, for even more complex systems, semantic coupling is even more vital.

This is easy to see when one considers the modeling challenges of modern systems biology (e.g., [12, 11, 7]). In those cases there is the challenge of multi-level or multi-scale modeling for both structure and process. This is easier to understand by consider a simple multi-scale model of protein structure.

For example, Figure 2, depicts three levels of representation of a protein: from bottom up, a sequence of amino acids at the base (i.e., MVKQIESKTA ...), an interpretation of that sequence as a secondary structure of α-helices, β-sheets, and random coils, and finally as possible 3D tertiary structures drawn as so-called "cartoon models."

If we were to provide simple direct manipulation methods for any level of this three-tier protein model, we would expect that manipulation to respect the structure at the other two. For example, if we provided a selection tool for any sequence of "Bs" in the secondary structure (cf. "brushing" methods [24]), then we would expect the appropriate selection of the correct subset of amino acids at the primary structure level, and similarly the appropriate selection of cartoon model segments at the tertiary structure level.

To make the last point about semantically-coupled direct manipulation used for extending complex computational models, consider the FoldIt protein folding game of Seth et al.[27, 5]. Figure 3 shows one panel of the game, which displays a 3D representation of a complex protein. A user can select any part of the displayed protein and drag that point to create a new configuration. If the configuration is not legal, the user is informed, and prompted to try an alternative.

In using direct manipulation to consider alternative configurations of proteins, the user is constrained by what must be assumed to be accurate models of the proteins being manipulated, so that the resulting manipulations produced only legal protein re-configurations. From a logical viewpoint, the manipulation actions can be considered as a kind of user-guided deductive inference: the space of possible consequences is large, but each manipulation or change can be considered as an inference constrained by the logical properties of the protein knowledge model. As in conventional data-mining, the constraints on the underlying data must be sufficiently accurate so that the intuitive use of direct manipulation actions provide visual changes that are predictable, based on a user's expectations.

Note that the kind of direct manipulation we consider is a kind of technologically-evolved version of a scholarly tradition of diagrammatic reasoning, with early origins in Liebniz' pictograms and then later with the logical reasoning on diagrams introduced by Pierce [23]. In the later development of the logical basis for reasoning with diagrams (e.g., [2, 17]), the focus is on general reasoning methods rather than on the capture of knowledge-specific multi-scale domain models. But it is clear that insight from general diagrammatic reasoning will provide at least a preliminary scaffolding for a formalization of general constraint reasoning on multi-scale models.

The direct manipulation embodied in the domain of proteins within Foldit, the general reasoning is a kind of "crowd-sourcing" to investigate the possible configuration space, and gather new hypotheses about possible protein folding configurations. This is highly useful in extending the overall knowledge model because the number of configurations is simply too vast to be searched directly, even by machine.

4 Direction Manipulation and Model Debugging

Section 3 sketched the idea of how direct manipulation can be considered as providing an explanatory mechanism, to manipulate the visualizations of complex models and provide feasible extensions of those models. In the protein example, the manipulations had to be constrained by accurate models of the proteins being manipulated, so that the resulting manipulations produced only legal protein re-configurations. From a logical viewpoint, the manipulation actions can be considered as a kind of user-guided *deductive inference*: the space of possible consequences is large, but each manipulation can be considered as an inference constrained by the logical properties of the protein knowledge model. Thus new and interesting protein foldings can be produced by FoldIt.

But what role can semantically-coupled manipulation techniques play when the underlying knowledge model is incomplete or in error? In this case, we claim that the next

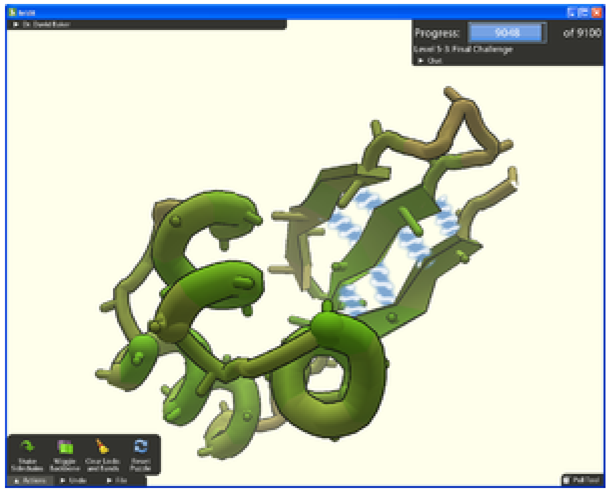

Figure 3: The Foldit direct manipulation of proteins game

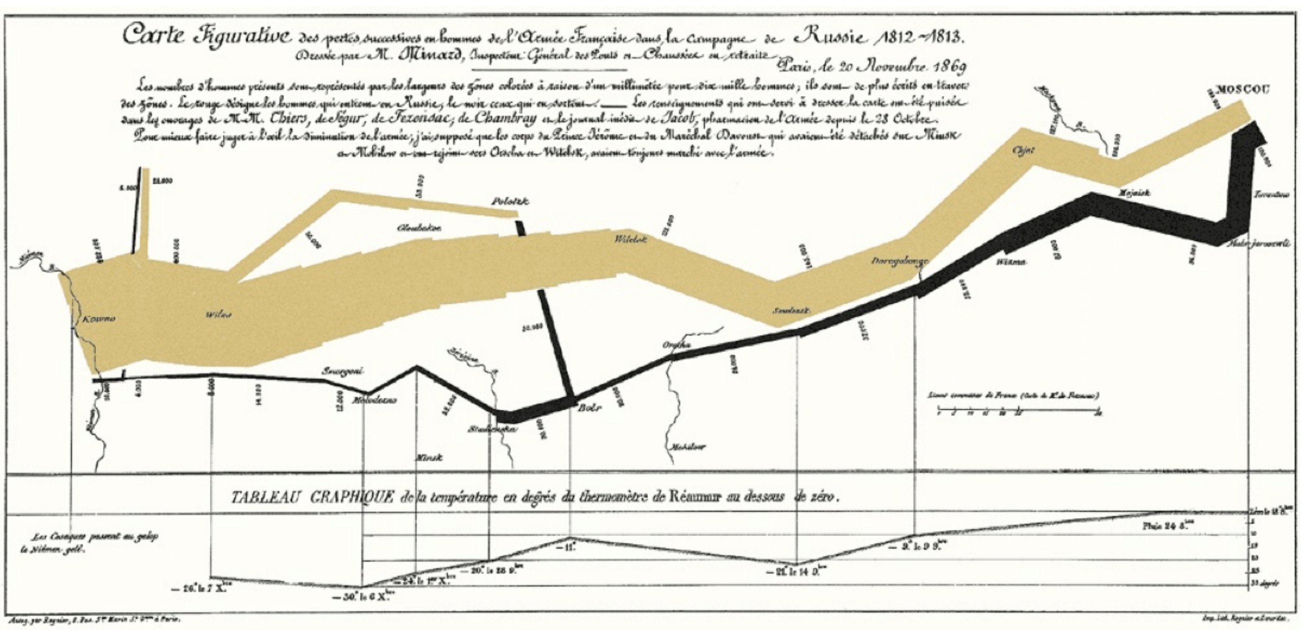

Figure 4: Minard's 1869 graphical presentation of Napoleon's Russian Campaign.

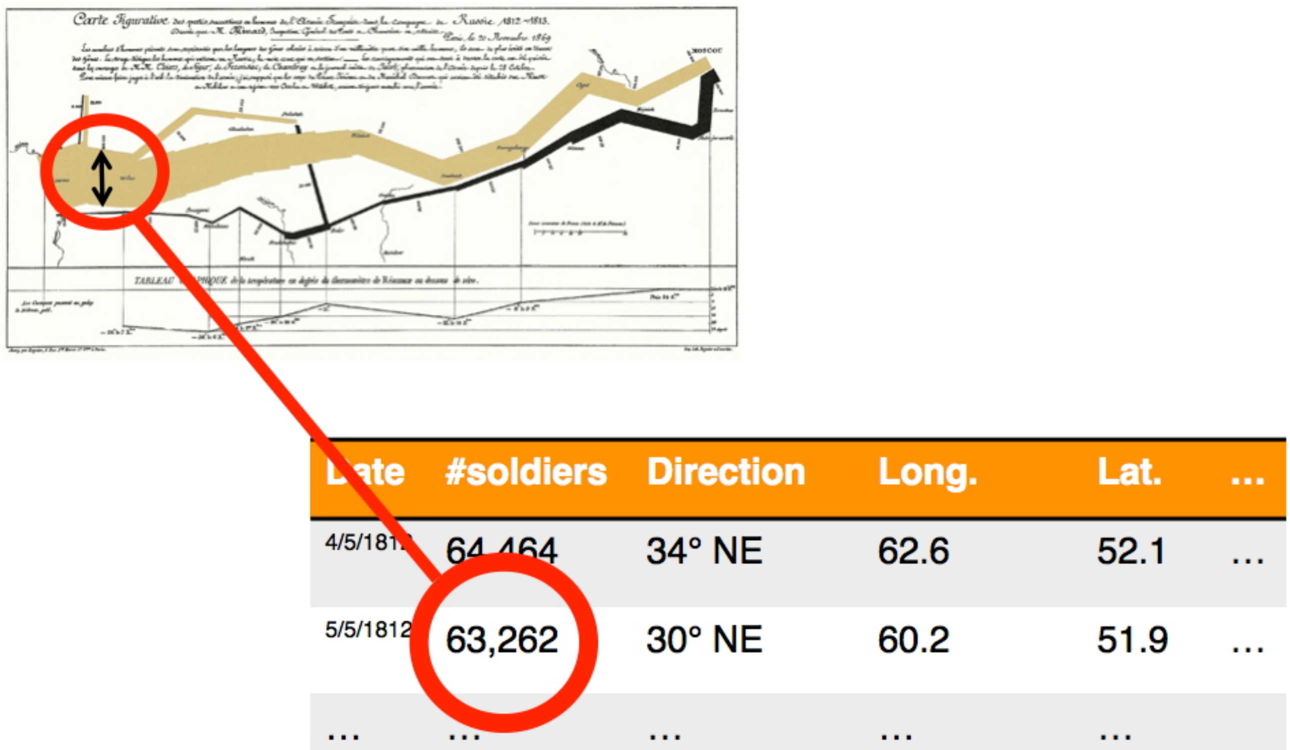

Figure 5: Hypothesized relation underlying Minard's chart

aspect of semantically-coupled manipulation can be developed as an exploratory data mining tool, to reveal missing and incorrect information in the knowledge model.

To provide a simple, but not too simple illustration of this idea, we use another two level knowledge model with the specific example of a famous visualization of Napoleon's Russia campaign, drawn by Minard in 1896 [13, 9]. There are three main kinds of information conveyed in Minard's chart: 1) the direction of Napoleon's campaign as portrayed by the direction or orientation of the line on the map, 2) the size of the army as portrayed by the width of the oriented line (before and after retreat), and 3) date and temperature portrayed along the x-axis at the bottom.

To demonstrate how even naive direct manipulation can help debug an underlying model, we first provide a plausible relational table that Minard might have consulted in drawing his chart, as depicted in Figure 5. The relational table has attributes of date, army size, longitude, and latitude, so that we could imagine the correlation illustrated by the overlayed line with circles, to show a connection between the width of the Minard chart line on one day, and the corresponding size of the army on the same day.

Now consider a possible manipulation based on a select and squeeze operation arising from the original iPad baby scenario. It's plausible enough to consider squeezing the chart line at the circled point in the visualization, but the question is what happens in the relational table? Should the size of the army change, for example, and by how much? Should that change propagate, so that the width of the chart line moving forward in dates is also reduced? It is at least clear that a naive manipulation of this sort creates an enormous volume of questions about the impact on the underlying relation.

Of course the point is that an appropriate *semantically-coupled* manipulation action would help reveal missing information in this scenario, either by illustrating that the dependencies amongst the relational attributes were missing or incorrectly specified. Referring back to the anticipated direct manipulations of the data and ensuing models of the SKA radio astronomy scenario, having manipulation actions that model the adjustment of semantically sensible actions can be now seen as a kind of modular model building and debugging strategy. In fact, we would argue that scientific model building based on semantic coupling of direct manipulation actions is a necessary component for any scientific discipline that broaches very large data sets (e.g., radio astronomy, genomics, environmental modelling).

5 Conclusions, with much work to be done

Most of the important historical landmarks in the development of methods for human direct manipulation of visualizations of data have been driven by technology. But the growing volume of data and its increasing complexity requires more attention to direct manipulation guided by the semantics of the models that emerge from that data.

Even as brief as this paper is, we have attempted to provide a high level view of how the development of direct manipulation while concurrently arguing that the real value of such manipulation will be in much more complex natural and unnatural knowledge model contexts. We hope it is clear that as the complexity of data and their consequent multi-level models emerge, the idea of semantically-coupled direction manipulation actions are not just tools for use after the models are built, but tools for building and debugging the models too.

The critical role of an integration of direct manipulation with the semantics of multi-scale models is really about providing scientists with tools to build and debug complex scientific models. When simple manipulations on accurate abstractions of high volumes of scientific data produce predictable results, then one expects confidence in more elaborate "abstraction experiments," where novelty in new data can emerge. In fact one could consider the whole exercise as providing experimental abstraction tools to exploit human scientific insight in impossibly large data sets. This of course begs the issue of evaluating "best" or at least preferred repertoires of direct manipulation; this will require an extension of the emerging discipline of visualization evaluation (e.g., [10]), where not just efficacy of visual inference is tested, but also that of direct manipulation adjustments to visualizations.

In other work, we have developed a programming platform that is platform independent, and demonstrates how to design manipulations within a three level representation system based on relational databases, relational views, and their visualization [16]. But there is much left to do, because in general, the development of direct manipulation tools such as those anticipated to change army size in the Minard scenario of Section 4 is really equivalent to the problem of defining non-monotonic reasoning inference operators.

In addition, much of the work in building the multi-scale models for systems biology proceeds by applying a variety of methods from machine learning to induce a variety of scale-free structures from empirical data. So in general, our pursuit of semantic-coupled manipulation will force us to confront both the non-montonic representation and reasoning problem, as well as the induction of multi-scale models from empirical data problem.

That will certainly keep us busy for some time.

Acknowledgements

Special thanks to three reviewers whose careful reading and simple insightful suggestions have contributed both to this brief paper and to our general thinking about semantic coupling of direct manipulation and visualization.

References

[1] Ayman Ammoura, Osmar R. Zaïane, and Randy Goebel. Towards a novel olap interface for distributed data warehouses. In *Data Warehousing and Knowledge Discovery, Third International Conference, DaWaK 2001, Munich, Germany, September 5-7, 2001, Proceedings*, pages 174–185, 2001.

[2] Michael Anderson and Robert McCartney. Diagram processing: Computing with diagrams. *Artificial Intelligence*, 145(1-2):181–226, 2003.

[3] F Beck and B Stumpe. The first capacitve touch screen at CERN. Available at http://cerncourier.com/cws/article/cern/42092, 1973.

[4] Chaomei Chen. Top 10 unsolved information visualization problems. *IEEE Comput. Graph. Appl.*, 25(4):12–16, July 2005.

[5] Seth Cooper, Adrien Treullie, Janos Barbero, Andrew Leaver-Fay, Kathleen Tuite, Firas Khatib, Alex Cho Snyder, Michael Beenen, David Salesin, David Baker, and Zoran Popović. Predicting protein structures with a multiplayer online game. *Nature*, 466:756–760, 2010.

[6] Eugene D. Davis. Odo: A constraint-based scheduler founded on a unified problem solving model. Enterprise Integration Lab, University of Toronto, TR-EIL-94-1, available at http://www.eil.utoronto.ca/scheduling/papers/davis-msc94.pdf, 1994.

[7] A. Desgapari and M. Calder. Multi-scale modelling of biological systems in process algebra with multi-way synchronization. In *9th International Conference on Computational Methods in Systems Biology 2011*, pages 195–208, September 21-23 2011.

[8] D.C. Engelbart. Augmenting human intellect: A conceptual framework. Stanford Research Institute Report USA.AFOSR-3233, available at http://www.1962paper.org/web.html, October 1962.

[9] Michael Friendly. Re-visions of minard. http://www.datavis.ca/gallery/re-minard.php, 2012.

[10] Heidi Lam, Enrico Bertini, Petra Isenberg, Catherine Plaisant, and Sheelagh Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Graphics and Visual Computing*, 18(9):1520–1536, 2012.

[11] P. Li, J.O. Dada, D. Jameson, I. Spasic, N. Swainston, K. Carroll, W. Dunn, and F. Khan. Systematic integration of experimental data and models in systems biology. *BMC Systems Biology*, 11(582), 2010.

[12] C. Maus, S. Rybaki, and A.M. Uhrmacher. Rule-based multi-level modeling of cell biological systems. *BMC Systems Biology*, 5(166), 2011.

[13] Charles Joseph Minard. Charles joseph minard, 2012. http://en.wikipedia.org/wiki/Charles_Joseph_Minard.

[14] Jurriaan D. Mulder, Jarke J. van Wijk, and Robert van Liere. A survey of computational steering environments. *Future Generation Comp. Syst.*, 15(1):119–129, 1999.

[15] M. Osorio, P. Figueroa, F. Prieto, P. Boulanger, and E. Londo no. In-situ visualization for cultural heritage sites using novel augmented reality technologies. *Computers and Graphics*, 35:894–903, 2011.

[16] W. Shi, Y. Tanaka, and R. Goebel. A new interactive information visualization framework based on the object-oriented views of querying and visualizing databases, 2012. In progress.

[17] Atsushi Shimojima. Operational constraints in diagrammatic reasoning. In Gerard Allwein and Jon Barwise, editors, *Logical reasoning with diagrams*, pages 27–48. Oxford University Press, Inc., New York, NY, USA, 1996.

[18] Hidekazu Shiozawa and Yutaka Matsushita. WWW visualization giving meanings to interactive manipulations. In *HCI (2)*, pages 791–794, 1997.

[19] SKA. Exploring the universe with the world's largest radio telescope. Square Kilometre Array Consortium, available at http://www.skatelescope.org/the-science/, 2012.

[20] D. Stricker, A. Pagani, and M. Zoellner. In-situ visualization for cultural heritage sites using novel augmented reality technologies. In *I Congreso Internacional de Arqueologła e Informtica Grfica, Patrimonio e Innovacin*, pages 141–146, June 2009.

[21] Nicole Sultanum, Ehud Sharlin, Mario Costa Sousa, Daniel N. Miranda-Filho, and Rob Eastick. Touching the depths: introducing tabletop interaction to reservoir engineering. In Antonio Krger, Johannes Schoening, Daniel Wigdor, and Michael Haller, editors, *ITS*, pages 105–108. ACM, 2010.

[22] UEW. A magazine is a broken ipad. http://youtube/aXV-yaFmQNk, October 2011.

[23] Wikipedia. Diagrammatic reasoning. http://en.wikipedia.org/wiki/Diagrammatic_reasoning, 2013.

[24] Marc Wolter, Bernd Hentschel, Irene Tedjo-Palczynski, and Torsten Kuhlen. A direct manipulation interface for time navigation in scientific visualizations. In *3D User Interfaces, 2009. 3DUI 2009. IEEE Symposium on*, pages 11 –18, march 2009.

[25] Ji Soo Yi, Youn ah Kang, John Stasko, and Julie Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, November 2007.

[26] P. Zikopoulos, D. Deross, K. Parasuraman, T. Deutsch, D. Corrigan, and J. Giles. *Harness the power of big data*. McGraw-Hill, 2005.

[27] David Baker Zoran Popović. Foldit: Solve puzzles for science. http://fold.it, October 2011.