History Tools for Collaborative Visualization

Ali Sarvghad, Narges Mahyar, and Melanie Tory

Abstract—In the context of collaborative data visualization and analysis, history tools can play an important role. We present a compilation that characterizes users' probable objectives when using history tools for collaborative work, as well as operations commonly performed on histories. We further characterize user objectives according to the likely time/space setting in which they would be used, and whether they are likely to be used by individuals, groups, or both. We conclude by compiling a list of design and implementation challenges and research questions that need to be discussed and investigated in order to make history tools adequately support collaborative visualization activities.

Index Terms—History tool, collaboraton,visualization, analysis.

INTRODUCTION

In this paper, we present a preliminary list of history operations and users' most common objectives in the context of collaborative data visualization and analysis. We then identify a set of key research challenges that will need to be addressed in order to make history tools effective for collaborative visualization tasks. Though the lists of history operations and objectives presented in this paper are detailed, they are neither final nor complete. They are preliminary lists assembled to trigger discussions and raise questions regarding collaborative use of histories in data visualization and analysis.

Many researchers have mentioned advantages of history tools and their importance for collaborative data visualization and analysis [5][6][10][12][14]. However, to date, visualization histories have been designed only for individual use, not communal use. Histories for group use will demand a new set of functionalities and design considerations.

Several applications provide a general-purpose undo/redo tool but this simplest form of history reuse is inadequate for participants of a collaborative data visualization and analysis task. They need to use history items, individually or collectively, to coordinate their work, try a different course of visualization and analysis, recover from a system crash, train naive users, and so on. Scientific workflow management tools such as Vistrails [1] and Taverna [16] capture very detailed information about scientific workflows. This information consists of data, created visualizations, and their manipulation. Though these systems maintain rich historical (provenance) information, they are designed primarily for expert users who are able to understand and manipulate complex workflows for creating visualizations from scientific data. More importantly, these systems have not been designed with collaborative work in mind.

History items can be browsed [5][12], searched [5][12], edited [5][12], filtered [5] and exported [5] for different purposes such as analysis, decision-making, validation and correction. As the name "history tool" suggests, users can revisit and reuse historical items. This reuse involves enacting specific operations to achieve specific objectives. In the following sections we will point out what we expect to be the most common operations and objectives performed and intended by history tool users.

- *Ali Sarvghad is with the University of Victoria, E-Mail: asarv@cs.uvic.ca.*
- *Narges Mahyar is with the University of Victoria., E-Mail: nmahyar@cs.uvic.ca.*
- *Melanie Tory is with the University of Victoria, E-Mail: mtory@cs.uvic.ca.*

1 MOST COMMON OPERATIONS ON HISTORY REPOSITORIES

Heer et al. [5] list a number of operations that a history tool should support. We built our list of operations largely based on their work, but we make some alterations. We expect the most common operations on history repositories to be:

- Browse
- Search
- Filter
- Edit
- Delete
- Export

We consider an *editing* operation to be changing the content of a history item, such as adding metadata, and we consider *deleting* history items to be independent from editing. We also consider *searching* and *filtering* as two different operations. Other researchers also point out the importance of *browsing* [3][12], *searching* [12][13] and *editing* [4] operations for history tools and some other researchers [4][8] mention the necessity of a tool to *export* and communicate history.

There is no one to one dependency between operations and user objectives. In other words, an operation, solely or in conjunction with other operations, can be performed to achieve a number of objectives. For example, searching and filtering both are required to accomplish analysis and validation objectives.

2 MOST COMMON OBJECTIVES

Based on a literature survey and our own experience, we expect history operations would be mainly used to achieve the following objectives:

- **Analysis** [3][8]: Users can traverse a history item repository and revisit different data visualizations to investigate data. Products of this analysis can vary from making a decision to verifying a hypothesis. We define analysis as investigating data with a specific goal in mind, in contrast to exploration.

- **Validation** [5][8][10][14]: Correctness and admissibility of decisions/findings or appropriateness of a single visualization can be examined by using history items. For instance, analysts may review visualizations created in the course of an analysis process to double-check that their findings are correct, or they may revisit a particular visualization to ensure that it is the result of correct mapping and filtering of data. This might be more helpful when users' collaboration style changes over time such as autonomous collaboration. Participants may need to corroborate the outcomes on individual works that will be concatenated later.

- **Memory aid**: The limitation of humans' short-term memory is a known fact, and a history tool can act as external memory aid [12]. Data analysts can add important notes, observations, calculations et cetera to history items for future referral.
- **Correction/Recovery**: If data analysts find their current visualization undesirable for any reason, they can perform a selective undo/redo [3][5][8][11][13][5]. It is also possible to continue a visualization and analysis process from the last point in the history repository after a system failure.
- **Exploration**: Exploration involves investigating data without a specific goal in mind. Having a repository of history items enables data analysts to try different courses of visual analysis by revisiting a history item and trying a different possible path. "Insight often comes from comparing the results of multiple visualizations that are created during the data exploration process" [2].
- **Reporting** [5][12]: A history repository, wholly or partially, can be sent to peers or upper management as a progress report, indication of the amount of work done, or formal report of findings.
- **Presentation** [5]: History items can be summarized and presented in a meeting situation. Presentation is similar to reporting, but typically occurs synchronously, as shown in Table 1.
- **Coordination** [4][8][11][12][14]: History items can help collaborators coordinate their effort by increasing awareness in situations such as autonomous collaborative work or remote synchronous/asynchronous situations. Also, viewing another users' history can bring a person up-to-speed on the work done so far.
- **Training** [12]: Novice data analysts can learn from experts by reviewing the history of visualizations created and decisions made.

It is quite possible that users have a combination of objectives when working with history items. For instance, users might review visualizations created in the course of an analysis process to both ensure their validity (i.e. correctness/admissibility) as well as make a decision.

Research is an interesting additional objective offered by rich history tools. Researchers can survey users' behaviour or assess a system's usability by observing the history of analysts' actions [5]. We do not include it in Table 1 because it is not performed directly by visualization users; nonetheless, it is worth mentioning.

3 Effects Of Time/Place Setting

Table 1 predicts the most likely time/place settings in which each objective might occur. As shown in the table, most of the objectives are likely to occur in all of the different time/place settings. However, we suspect that history records may need to be more explicitly displayed for synchronous distributed work in order to help users maintain awareness of others' activities. Additionally, using histories in asynchronous work may require different functionality than synchronous work. For instance, when sharing a history with another user who will take over the work later, a person may want to highlight particularly important findings to ensure they are noticed, or remove an unsuccessful path of analysis and replace it with a simple note to say that investigating that direction was not fruitful.

Table 1: Objectives' most likely time/place setting. ST = same time, DT = different time, SP = same place, DP = different place

	ST, SP	ST, DP	DT, SP	DT, DP
Analysis	√	√	√	√
Validation	√	√	√	√
Memory aid	√	√	√	√
Correction/Recovery	√	√	---	---
Exploration	√	√	√	√
Reporting	---	---	√	√
Presentation	√	√	---	---
Coordination	√	√	√	√
Training	√	√	√	√

4 Individual vs. Collaborative Use of Histories

Reporting, collaborating, coordinating and training are inherently collaborative objectives and require engagement of more than one person; the rest of the objectives could apply to both individuals and collaborating users. Though individuals and groups share most of the objectives, design of a history tool might need to be quite different to support the activities of a group as compared to one person. To adequately support group activities, we anticipate that history tools may need to provide the following functionality:

- Representation of *who* was responsible for each action recorded in the history.
- Both individual and shared histories. This will hopefully prevent users from being overwhelmed with history items from all members of the group. In addition, privacy control may be needed so that some items can be kept private.
- Additional awareness mechanisms, such as an indication that another user has worked on a similar chart or has looked at the same data. This might be similar to awareness mechanisms previously used in collaborative document search [9].
- Extensive editing, highlighting, and annotation capabilities. These will help users to communicate what they have done, or convert a history into a series of visual items and descriptions suitable for a report, presentation, or tutorial.
- Ability to export elements of a history to a document or presentation format for further manipulation.

5 Design Challenges/Questions

There are some important issues to be considered in designing and developing history tools. These issues need to be resolved before history tools can effectively support collaboration:

What content should a history item contain? Researchers have suggested and examined a variety of probable contents such as user interactions (or commands) [3][15][17], software states [5], a combination of commands/states [13] and states plus users' augmented information [5]. User information (which user was responsible for each action) may also be needed for collaborative objectives such as coordination. However, it is still unclear exactly which content is needed to support different collaborative tasks (e.g. training vs. shared analysis) and collaboration styles (e.g. loosely coupled to closely coupled work).

What data structures should be used? Histories can rapidly grow in size and need appropriate data structures and scaling tools [5].

How should a history be represented? Selecting the form that best suits users depends partly on form of the content [5]. For instance, a repository of executed commands can be represented as list of textual commands, a history consisting of a number of graphs can be represented as a comic strip [7], and for hybrid content of commands/states, text and graphics can be used jointly [13]. The ideal representation will also depend on the task, display and input hardware, and setting. For instance, a history that can support distributed awareness during joint analysis may look very different from a history that can support co-located training.

How can we support fluid interaction with histories? Especially for co-located collaboration, where interactive touch surfaces may be used, new mechanisms may need to be developed for interaction with histories.

What are underlying data challenges? It is important to pay attention to the underlying data. Volatile or streaming data add additional challenges for history tools [5]. Moreover, we might need to closely survey different data (e.g. business data and scientific data) to understand their effect on content and representation of history repositories and functionalities they should provide to facilitate collaborative work.

What features of a history tool are needed to support different collaborative activities? Can a single architecture support all of the different time/place settings and user objectives?

6 Conclusion

In this paper we compiled a list of operations and objectives related to history tools, and described the importance of such tools for the process of collaborative data visualization and analysis. History tools to support collaborative work are not merely instruments for correcting errors but also provide users with some vital functionality necessary for coordination, training, sharing information, and many other objectives. Designers of software for collaborative work need to take into consideration operations that a history tool must support and objectives that users are most likely to desire. Open research questions include what content to include in histories, how to store histories efficiently, and how histories should be best represented to support different collaborative tasks and situations.

Acknowledgements

This research was supported by grants from SAP Business Objects and the Natural Sciences and Engineering Council of Canada (NSERC). We also thank Kellogg Booth and our colleagues at SAP Business Objects for their many thoughts and suggestions.

References

[1] S.P. Callahan, J. Freire, E. Santos, C.E. Scheidegger, C.T. Silva, and H.T. Vo, "VisTrails: visualization meets data management," *Proc. ACM SIGMOD international conference on Management of data*, Chicago, IL, USA: ACM, pp. 745-747, 2006.

[2] S.P. Callahan, J. Freire, E. Santos, C.E. Scheidegger, C.T. Silva, and H.T. Vo, "Managing the Evolution of Dataflows with VisTrails," *Proc. 22nd International Conference on Data Engineering Workshops, IEEE Computer Society*, p. 71, 2006.

[3] W.K. Edwards, T. Igarashi, A. LaMarca, and E.D. Mynatt, "A temporal model for multi-level undo and redo," *Proc. 13th annual ACM symposium on User interface software and technology*, San Diego, California, United States: ACM, pp. 31-40, 2000.

[4] W.K. Edwards and E.D. Mynatt, "Timewarp: techniques for autonomous collaboration," Proc. SIGCHI conference on Human factors in computing systems, Atlanta, Georgia, United States: ACM, pp. 218-225, 1997.

[5] J. Heer, J. Mackinlay, C. Stolte, and M. Agrawala, "Graphical Histories for Visualization: Supporting Analysis, Communication, and Evaluation," IEEE Transactions on Visualization and Computer Graphics, vol. 14, pp. 1189-1196, 2008.

[6] J. Heer, F.B. Viegas, and M. Wattenberg, "Voyagers and voyeurs: Supporting asynchronous collaborative visualization," Commun. ACM, vol. 52, pp. 87-97, 2009.

[7] R.R. Hightower, L.T. Ring, J.I. Helfman, B.B. Bederson, and J.D. Hollan, "Graphical multiscale Web histories: a study of padprints," Proc. 9th ACM conference on Hypertext and hypermedia: links, objects, time and space---structure in hypermedia systems, Pittsburgh, Pennsylvania, United States: ACM, pp. 58-65, 1998.

[8] J. Heer, F. Ham, S. Carpendale, C. Weaver, and P. Isenberg, "Creation and Collaboration: Engaging New Audiences for Information Visualization," Information Visualization: Human-Centered Issues and Perspectives, Springer-Verlag, pp. 92-133, 2008.

[9] P. Isenberg and Danyel Fisher, "Collaborative Brushing and Linking for Co-located Visual Analytics of Document Collections," Proc. Eurographics / IEEE-VGTC Symposium on Visualization, pp. 1031-1038, 2009.

[10] P. Isenberg, A. Tang, and S. Carpendale, "An exploratory study of visual information analysis," Proc. 26th annual SIGCHI conference on Human factors in computing systems, Florence, Italy: ACM, pp. 1217-1226, 2008.

[11] G. Johnson, "Collaborative Visualization 101," ACM SIGGRAPH - Computer Graphics, vol. 32, May, pp. 8-11. 1998.

[12] A. Komlodi and W.G. Lutters, "Collaborative use of individual search histories," Interact. Comput., vol. 20, pp. 184-198, 2008.

[13] C. Meng, M. Yasue, A. Imamiya, and X. Mao, "Visualizing Histories for Selective Undo and Redo," Proc. 3rd Asian Pacific Computer and Human Interaction, IEEE Computer Society, pp. 459, 1998.

[14] G. Mark and A. Kobsa, "The effects of collaboration and system transparency on CIVE usage: an empirical study and model," Teleoper. Virtual Environ, vol. 14, pp.60-80, 2005.

[15] B.A. Myers and D.S. Kosbie, "Reusable hierarchical command objects," Proc. SIGCHI conference on Human factors in computing systems: common ground, Vancouver, British Columbia, Canada: ACM, pp. 260-267, 1996.

[16] T. Oinn, M. Addis, J. Ferris, D. Marvin, M. Senger, M. Greenwood, T. Carver, K. Glover, M.R. Pocock, A. Wipat, and P. Li, "Taverna: a tool for the composition and enactment of bioinformatics workflows," Bioinformatics, vol. 20, pp. 3045-3054, 2004.

[17] 12-J.S. Vitter, "US&R: A New Framework for Redoing," IEEE Software., vol. 1, pp. 39-52, 1984.