Towards A Role of Visualization in Social Modeling

Remco Chang and **Alex Godwin** and **Robert Kosara** and **William Ribarsky**
UNC Charlotte
Charlotte Visualization Center
www.viscenter.uncc.edu
{rchang | rkosara | ribarsky}@uncc.edu, jagodwin@gmail.com

Abstract

The traditional role of visualization in large scale social modeling projects has mostly been relegated to presentation and reporting. While these projects see the potential of communicating information visually, the visualization component has too often been considered as the final stage of a long process, sometimes as a last-minute add-on. The result is that these visualization components are limited in capabilities, and often appear disjointed and forced from the rest of the project. In this paper, we propose that for visualization to be more commonly accepted, it needs to fit the role as an analytical tool on top of being a presentation component. We use our probe-based visualization as an example of how such cross-over can occur, and present some challenges in social modeling that can be addressed using visualization techniques.

Visualization is becoming prevalent in our everyday lives. Popular websites such as Google Maps, the Visual Thesaurus, and Baby Name Voyager have increased the public's awareness of how information can be efficiently communicated visually. In the past few years alone, visualization components are appearing with more frequency in the final products of large scale social modeling projects.

However, a gap still remains between how the visualization community identifies itself and the perception of the role of visualization by the other disciplines. For too often visualization is considered to be the last step of a project, where visualization experts are asked to "make pretty pictures out of the final data" without having participated in the design or implementation of the overall system.

We propose that for visualization to take a more integral part in these systems, it needs to fulfill both the roles of a presentation tool as well as an analytical tool. Not only does visualization need to deliver the "pretty pictures," it also needs to provide the analytical capability that is difficult for non-visual components to deliver. As an example, we describe our probe-based visualization that supports interactive analysis and presentation of an agent-based simulation. We then propose additional challenges in social modeling that can be addressed by integrating a visualization component.

Visualization for Presentation and Analysis

We designed a probe-based visualization as an interface to an agent-based model for simulating the social-economic factors of Afghanistan (Butkiewicz et al. 2008). Although the interface is designed independently of the simulation, it achieves both goals of presentation and analysis. On the presentation level, the interface allows the user to interactively select regions-of-interest and presents pop-up windows containing information visualizations of the specified areas. On the analysis level, these regions-of-interest are connected to the simulation in real time. Local parameters can be modified through the associated pop-up visualizations which allows the user to immediately see the effects (Figure 1).

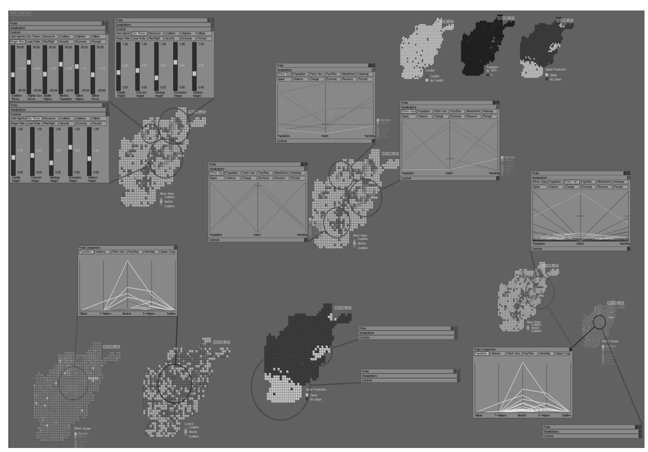

Figure 1: *Our probe-based visualization of an agent-based simulation, used to study the possible effects of development projects and other interventions in Afghanistan.*

Compared to the traditional agent-based simulation interface, the probe-based visualization offers multiple advantages. Without the ability to select localized regions, traditional visualizations for agent-based simulation depict information at a global scale. Similarly, the hundreds of controls for the parameters of the simulation are displayed without context. It is difficult to understand which slider controls what parameter, and at what geospatial locations.

This example illustrates how a visualization can function both as presentation and analysis tools. Through high interactivity, the user can select regions of interest and display information relating to those areas. By directly manipulating these pop-up visualizations, the user can alter the simulation's parameters and experiment with what-if scenarios.

Challenges in Social Modeling

Agent-based modeling and simulation is not the only domain that can benefit from visualizations that function both as presentation and analysis tools. Here we present a few challenges in social modeling that can benefit from visualization techniques.

Analysis with Uncertainty

In problems relating to human activities, uncertainty is often a factor in modeling and simulation. For analysts to effectively make decisions, it is imperative that they understand the uncertainties that are inherent in the data and the solution (Godwin et al. 2008). For a few specific problems, visualization techniques are available for presenting data with different degrees of trustworthiness. For example, fog-of-war (blurring) is a typical technique used to display questionable geospatial data. In information visualization of multivariate data, Xie et al. treat data quality as an additional dimension in the visualization (Xie et al. 2006): uncertainty information is collected with the original data and tracked through the entire program. Any effects of data transformations on quality can thus also be included in the visualization.

However, in social modeling and simulation there are fewer techniques that effectively convey the uncertainty within the data. One exception is the use of lattices for depicting words with multiple meanings in different contexts. The Visual Thesaurus is a good example of this type of visualization. Collins et al. (Collins, Carpendale, and Penn 2007) apply the same technique to display multiple possible translations of the same sentence in a foreign language. Similarly, Albrecht et al. (J.S. Albrecht 2008) also utilize node-link structures for translating Chinese to English that takes into account the context of the translated words. Although both Collins et al. and Albrecht et al. apply similar techniques for visualizing the uncertainty within a single translator, it is not difficult to extend the same technique to "merge" translations of the same sentence from multiple translators. This type of visualization intuitively conveys the ambiguous nature of translations and offers the user a simple display to understand the potential multiple meanings behind the same sentence.

Heterogeneous Processes

Data used in analyzing social behaviors is often complex and uncoordinated. For example, to understand the cause and effect of civil strife (Wang et al. 2008), one must at least consider social, economic, political, and religious factors. Collecting data for these aspects is difficult enough, but to weave them together to form a coherent picture is even more challenging. Worse yet, unifying these data could introduce conflicts within them, or sometimes remove parts of the data that are of value.

Instead of coordinating heterogeneous data into a single dataset, we propose that a visualization approach to managing heterogenous processes could be more efficient and cost-effective. In this approach, individual processes developed to analyze each type of data is run independently of the other processes. The visualization then serves as a management tool for coordinating the necessary process to execute and organizing the resulting knowledge artifacts. Similar to the Scalable Reasoning System by Pike et al. (Pike, May, and Turner 2007), these artifacts can be treated as hypothesis, evidence, etc. in the final analysis process to fully understand the causes and effects of events.

Wang et al. (Wang et al. 2009) apply this concept to visually analyzing bridge conditions. In their system, the condition of a bridge is summarized through three different sources of data - the bridge engineer's report, remote imagery, and 3D laser scans (LiDAR). These three sources of data are interpreted and analyzed using disparate technologies, but each analysis result is integrated into a cohesive visual analytical system. However, the analysis results of the three sources of data often contradict each other, and it is through the use of the visual analytical system that the bridge analysts can identify the reasons behind the contradictions and make informed decisions by forming a complete understanding of all aspects of the problem.

Visualizing Search Spaces

In an experiment involving financial analysts using a visual analytical tool for fraud detection in wire transactions, Stukes et al. find that analysts have a tendency to "re-discover" the same suspicious activities multiple times (Stukes et al. 2009). One reason is that the analysts' memory fades over time. Another is that analysts often take different strategies to find suspicious behaviors, but the strategies sometimes lead to the same suspicious transactions. When the analysts apply different strategies, it is impossible for them to foresee which activities these strategies will reveal. By the time they realize that they have arrived at the same discoveries, valuable time has already been wasted.

Similarly, Jeong et al. (Jeong et al. 2008) demonstrate that there is a great deal of knowledge embedded in an analyst's interactions with an interactive visual analytical system. Through the use of appropriate tools and methodologies, much of the analysts' knowledge and the reasoning processes used to perform the analysis are in fact recoverable only through examining the analyst's interactions (Dou et al. 2009). Together with the findings of Stukes et al., the discovery of Jeong et al. give weight to the need of a visualization that allows analysts to see and understand the interactions and analysis that have been performed.

Visualizing the analysts' search space and interactions has two important functions. During the investigation, the analyst can see the path of their analysis and project if they are investigating the same transactions or accounts. It can also be used as a communication and reporting tool. By sharing the visualized search space, the analyst can quickly show fellow analysts the areas that have been examined. More importantly, this visualization will demonstrate the paths that the analyst undertakes to identify the fraudulent activities. Together, this simple visualization functions as both presentation and analysis tools to understanding the problem domain.

Social Network Analysis

An increasingly popular trend in online user interaction is the creation and maintenance of relationships through social networking websites. Friendster, MySpace, and Facebook are just a few examples of sites boasting membership in the millions from around the globe. Appropriate analysis of online social networks has applications in advertising, trend-spotting, market prediction, and defense.

One of the most frequent goals of social network analysis (SNA) is the identification of communities. While some social networking communities, such as Facebook, do allow users to place themselves into groups, others do not. Communities must therefore be detected through analysis. Visualization tools for the presentation and analysis of social networks, such as Heer and Boyd's Vizster (Heer and Boyd 2005), allow analysts to detect community structure when none is explicitly provided by users through visual exploration of the social network in a node-link diagram.

Perer and Shneiderman's SocialAction (Perer and Shneiderman 2006) takes another approach. SocialAction provides integration of visual exploration and statistical methods to discover community structure as well as rank the importance of individuals to the connectivity of the social model. Systematic exploration is tightly coupled to the visualization of the social network, and as a result, the model and its representation evolve as the user interacts with the data.

This idea is taken further by Godwin in Time Web (Godwin 2008), a tool used for the generation of similarity networks from activity records of individuals over time. Through the application of a similarity detection algorithm, individuals can be compared to each other to identify those who follow similar trends of behavior (Figure 2). The results of analysis are presented using an evolving representation of the similarity network that allows the user to provide feedback and modify the algorithm. While the algorithm is sufficient to detect similarity between two individuals, this visualization is necessary to develop a greater understanding of the interaction of thousands of individuals or more. Through study of the similarity network, analysts are able to identify shared trends as well as communities who follow the same behavioral patterns. This can yield insight into social interactions between similar individuals.

Conclusion

Visualization has the potential of becoming an integral part of social modeling and simulation. It is up to the researchers in the visualization community to identify methods and tools that can be immediately applied to address the needs of the social scientists. We demonstrate that one promising direction for visualization is to play both the roles of presentation and analysis. By leveraging the capabilities of visualization techniques, we hope that some difficult challenges in social modeling can be efficiently addressed.

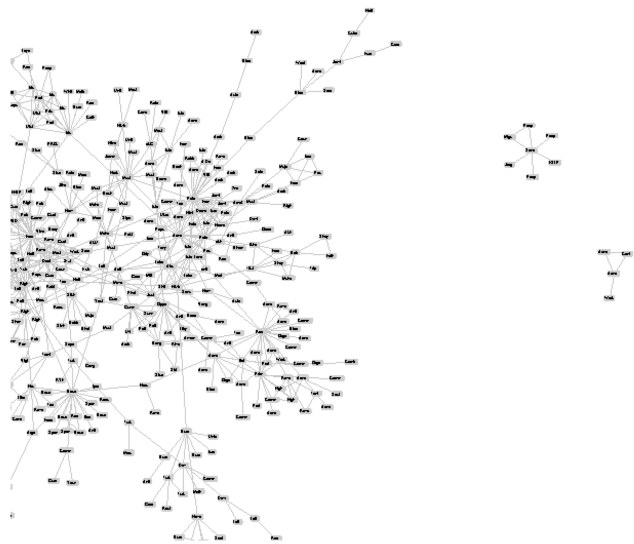

Figure 2: *A similarity network created in Time Web. The source data is a collection of terrorist attacks, which our software grouped by the terrorist organization responsible to uncover similarities between groups. Groups that were found to be very similar were connected in the similarity network.*

References

Butkiewicz, T.; Dou, W.; Wartell, Z.; Ribarsky, W.; and Chang, R. 2008. Multi-focused geospatial analysis using probes. *Visualization and Computer Graphics, IEEE Transactions on*. To appear.

Collins, C.; Carpendale, M. S. T.; and Penn, G. 2007. Visualization of uncertainty in lattices to support decision-making. In *EuroVis*, 51–58.

Dou, W.; Jeong, D. H.; Stukes, F.; Ribarsky, W.; Lipford, H. R.; and Chang, R. 2009. Recovering reasoning process from user interactions. *IEEE Computer Graphics and Applications*. Accepted for Publication.

Godwin, A.; Chang, R.; Kosara, R.; and Ribarsky, W. 2008. Visual analysis of entity relationships in the global terrorism database. volume 6983, 69830G. SPIE.

Godwin, A. 2008. Time web: Comparing unevenly-spaced time sequences using social network analysis of local alignment pairs. Master's thesis, University of North Carolina at Charlotte.

Heer, J., and Boyd, D. 2005. Vizster: Visualizing online social networks. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis)*.

Jeong, D. H.; Dou, W.; Lipford, H. R.; Stukes, F.; Chang, R.; and Ribarsky, W. 2008. Evaluating the relationship between user interaction and financial visual analysis. *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on* 83–90.

J.S. Albrecht, R. Hwa, G. M. 2008. The chinese room - understanding and correcting machine translation. In *IEEE Information Visualization 2008 Poster Compendium*. In Press.

Perer, A., and Shneiderman, B. 2006. Balancing systematic and flexible exploration of social networks. *IEEE Transactions on Visualization and Computer Graphics* 12(5):693–700.

Pike, W.; May, R.; and Turner, A. 2007. Supporting knowledge transfer through decomposable reasoning artifacts. In *40th Annual Hawaii International Conference on System Sciences, 2007. HICSS 2007*, 204c–204c.

Stukes, F.; Dou, W.; Jeong, D. H.; Chang, R.; and Lipford, H. R. 2009. Visual analysis - recall vs. reality. *Information Visualization*. In submission.

Wang, X.; Miller, E.; Smarick, K.; Ribarsky, W.; and Chang, R. 2008. Investigative visual analysis of global terrorism database. *Computer Graphics Forum* 27(3):919–926.

Wang, X.; Dou, W.; Ribarsky, W.; and Chang, R. 2009. Integration of heterogenous processes through visual analytics. SPIE. Accepted for publication.

Xie, Z.; Huang, S.; Ward, M. O.; and Rundensteiner, E. A. 2006. Exploratory visualization of multivariate data with variable quality. *Visual Analytics Science And Technology, 2006 IEEE Symposium On* 183–190.