NSF Science of Interaction for Data and Visual Analytics

Workshop report (David Ebert, Kelly Gaither, Niklas Elmqvist)

The ever-growing volume of science and user-generated data in today's society creates a need for valuable, timely analysis tools that provide critical insights from this sea of data through interactive visual analysis for sensemaking. Key findings in visual analytics have demonstrated that interaction and collaboration with data are key components of an integrated computational-human decision making loop. This human-information interaction occurs at many levels: from individual manipulation of data representation, through interactive cognitive discovery combined with automated analysis, and to distributed interactive analysis among groups of individuals.

The outcome of this workshop series was to define a research agenda for the "**Science of Interaction**" to support ubiquitous and collaborative analysis and discovery utilizing new interactive tools. What resulted was the identification and discussion of four distinct areas of consideration, including descriptions of their issues and limitations as well as a call to arms for open research problems. These four areas of research encompass a rich set of issues, capabilities, and needs for next generation interaction in these ever-growing large data environments. Below we discuss each of the research areas in more detail:

Fluid Interaction – The interaction component in a visual analytics tool is the process through which analysts (a) convey their intent to and (b) receive responses from the tool during sensemaking. While all visual analytics tools are by their definition interactive, it is also true that the style, quality, and efficiency of the interaction differ widely across tools. One way to capture this distinction is through the concept of *fluid interaction*, defined as the degree to which the response from the visual analytics system matches the user's intent. However, fluidity cannot be achieved merely by optimizing aspects of an interface, or by providing additional functionality in the interface or system, but is a function of both the people making use of the interfaces and the interface itself. There are several challenges inherent in this research area: (1) taking advantage of new computing platforms, including touch-based, pen-based, and gesture-based interaction as well as mobile, wearable, and ubiquitous computing platforms; (2) enabling fluid collaboration between multiple analytic performers, each with their own expertise, role, and authority; (3) seamlessly integrating (make available, integrate, and provide ready-at-hand) analysis algorithms into a visual analytics interface; (4) identifying cross-cutting interactions common to virtually sensemaking; and (5) evaluating fluidity in an interface for the purpose of improving it.

Interactive Sensemaking in Information-Rich Environments – There has been a great deal of work to help individual analysts explore and understand large quantitative datasets, including human-machine interfaces for visual interaction with data. This work has helped the *information* within that *data* to be more easily extracted, used, and combined with the domain expertise provided by the analysts. As a result, the interaction burdens and technical expertise requirements being placed on analysts, as barriers to fluid interaction and understanding of the data, are being ever lowered. Some of this work has been extended to support distributed collaboration in visual data exploration, e.g. over a Web interface where multiple people can each explore the data, finding and sharing their insights with one another. Interactive sensemaking encompasses issues to enable large-scale, distributed teams to interactively make sense of big multimodal data and complex problems and to operationalize that understanding to coordinate their

actions. The biggest challenge is that of enabling large groups of people to interact with each other and data to do sensemaking on complex problems where no one person can understand the problems, let alone specific questions or answers. We want to understand how to build interaction environments that enable people with different expertise to interact with one another and the data, understand each other's insights, and iteratively build on each other's insights. The second major research challenge in this area is supporting interaction with multimodal data, where each piece and set of data may have different structure, properties, and format and may originate from a wide variety of heterogeneous sources. The third and final challenge is deriving suitable metrics for measuring sensemaking. Having clear and meaningful metrics are essential for making this a science supported by rigorous research practices.

Sensemaking Enhanced by Interactive Computation Guidance – Going beyond the computer-as-tool metaphor and into the computer-as-partner metaphor for sensemaking would enable enhancing the analytical process by balancing active user input with automatic guidance provided by the computer system. The role of this computational guidance is to amplify the cognitive powers of analysts, not to remove any of their responsibilities. Adding such guidance would require developing process models tuned to domain-specific problems. However, balancing active user input with automatic guidance is potentially problematic. We must explore new processes for the system providing guidance for analysts, for example, different guidance processes depending on the tasks, and guidance breadth versus depth. We must also study the process of analysts steering the system or model, both bi-directional and implicit model learning. Finally, we must manage issues of uncertainty and trust and determine methods for accounting for the growth of uncertainty over a lengthy analysis session.

Visual Discourse – Visualizing data of all kinds is necessary, but not sufficient for visual analytics. Data need to be understood and applied for analysts to be able to reach agreement and operationalize analysis in action. Generalizable analytic methods and structures must be found that support argumentation and negotiation of meaning, common ground, and coordinated actions across a range of data types, problems, and human organizations. This entails discourse, sensemaking, and storytelling. An interrelated set of challenges is the development and evaluation of interactive adaptive visual tools that incorporate existing semantic elements and syntactic rules, and invent new ones to support the visual rhetoric of data exploration, hypothesis testing, understanding, description, explanation, decision, interaction, collaboration, and persuasion. Some of the specific topics included in the research agenda for visual discourse include exploration, understanding, collaboration, description, explanation, decision, dissemination, and persuasion about data, underlying concepts, and plans for action.