SAND2009-0057P

Human Factors Literature Review

Travis Bauer
Courtney Dornburg
Laura Matzen

The literature review included 200 papers

- Goal of Literature Review
  - To educate the team members
  - To stimulate discussion and generate ideas on
    - Application evaluation
    - User interviews and knowledge elicitation
  - To look for gaps in the current methodologies
- Summary of Paper Types
  - Evaluations of specific pieces of software
  - Studies of and proposals for standard evaluation methodologies
  - Systematization of different aspects of evaluation
  - Studies of analysts' workflow and environment

Talk Outline

General InfoViz Issues
- Design Issues
- Evaluation Issues

Task Types
- Knowledge Tasks
- Analytic Functions

Methodologies
- Qualitative
- Quantitative
- Tools

General Information Visualization Issues

- Issues for Design Methodologies
  - Faulty Assumptions (Ackoff88)
    - Management's most critical information need is for more relevant information
    - If managers have all the information they need, they will perform better
  - Societal Responsibilities of Developers (Beusmans89, Shneiderman91)
  - Dealing with Uncertainty (Amar04, Ellis06)

General Information Visualization Issues

- Issues for Evaluation
  - Diversity of data sets (Ellis06)
  - Complexity of visualizations (Ellis06)
  - **IDEA:** Radical diversity between old and new software
  - Variability among analysts and their tasks
    - Makes choosing an appropriate subject pool difficult. Makes generalization difficult. Some software might help some people and hinder others (Greitzer05)
  - Need for validation and methodology (Craft05)
  - Creation rather than adherence to standards (Craft05)
  - Domain specificity of representations (Amar04)

Tasks Aided by Information Visualization

- Knowledge Tasks
  - High level goals related to the problem as opposed to the data.
- Analytic Functions
  - Things one does to the data in order to accomplish Knowledge Tasks

Knowledge Tasks

- Exposing Uncertainty (Amar04)
- Concretize Relationships (Amar04)
- Formulate Cause and Effect (Amar04, Teoh03)
- Determination of Domain Parameters (Amar04)
- Multivariate Explanation (Amar04)
- Confirm Hypotheses (Amar04)
- Answering questions you didn't know you had (Plaisant04, Teoh03)
- Looking at the same data from different perspectives over a long time period (Plaisant04)
- Discovery of Patterns (Teoh03)

Analytic Functions

- Retrieve Value (Amar05, Craft05, Shneiderman96, Valiati06, Ardito06)
- Filter (Amar05, Craft05, Shneiderman96, Ardito06)
- Compute Derived Value (Amar05)
- Find Extremum (Amar05)
- Sort (Amar05, Valiati06)
- Characterize Distribution (Amar05, Valiati06)
- Find Anomalies (Amar05)
- Cluster (Amar05, Valiati06)
- Correlate (Amar05, Valiati06)
- Overview (Craft05, Shneiderman96, Ardito06)
- Zoom (Craft05, Shneiderman96, Ardito06)
- **IDEA:** Pan or Scroll
- Relate (Craft05, Shneiderman96)
- History of the analytic process (Craft05, Shneiderman96, Ardito06)
- Extract important findings (Craft05, Shneiderman96, Ardito 06)
- Identify/Find (Valiati06)
- Distinguish (Valiati06)

Evaluation Methodologies

- Quantitative
- Qualitative
- Tools

Evaluation Methodologies

Quantitative

- Simulation/Human Performance Modeling (Baines04)
  - EPIC – predicting performance in multi-modal high-performance tasks (Kieras97)
- Average Response Time (Nijim05)
- Average Throughput (Nijim05, Andrews06)
- Controlled experiments comparing design elements (Plaisant04)
- Controlled experiments comparing two or more tools (Plaisant04, Lam08)
- Eye Tracking (Huang08, Korner04, Shen08)
  - **IDEA:** Which aspects of the display draw attention? Are those the most important?
  - **IDEA:** How long does it take an analyst to spot the relevant information in a display?
  - **IDEA:** How long do users look at a piece of information before deciding what to do with it?
- Event-related potentials (Verwey96)

Evaluation Methodologies

Quantitative (cont.)

- Measurements of Cognitive Load (Huang08, Verwey96)
  - Primary and/or secondary task performance
  - Subjective assessments of workload
  - Physiological measures
- Size Metrics (Bertini06)
- Visual Effectiveness Metrics (Bertini06)
- Feature Preservation Metrics (Bertini06)

- **IDEA:** How long does it take to complete a task or subtask?
- **IDEA:** How many documents can an analyst find/sort/read in a given amount of time?
- **IDEA:** How many iterations of a search are needed to find all of the relevant information?
- **IDEA:** How much relevant information is missed?
- **IDEA:** How many times are the same actions repeated?

Evaluation Methodologies

Qualitative

- Heuristic Evaluation (Frokjaer08, Zuk06, Nielsen93)
- Cognitive Walk Through
  - **IDEA**: Augment existing evaluations of Analyst's work flow with more formal methods.
- Think aloud testing (Frokjaer08, Andrews06, Wharton94)
- Case studies of tools in realistic settings (Plaisant04)
- MILC (Multi-dimensional in-depth long-term case studies) (Valiati06, Lam08, Andrews06, Shneiderman06)
- Process Centric Evaluation (McNee08)
- Grounded Evaluation – evaluation in context (Isenberg08)
- Log file evaluation (Rester06)
- Focus Group (Reser06, Mazza06)
- Maintaining concentration (Shneiderman05)
- Metaphors of Thinking (Frokjaer08, Johnson91)

Methodologies for Evaluation

- Tools
  - Benchmark Data sets and Tasks (Plaisant04)
    - **IDEA**: benchmark performance on some data set with and without NGC tools.
  - Toolkits and development tools (Plaisant04)
  - Threat Stream Generator (Whiting08, Whiting06)

Possible NGC HF Research Direction

- Measure cognitive load using secondary task performance
  - Advantages
    - Can be used as a general tool across many different kinds of visualizations
    - Indicates how easy or difficult it is to use the software
  - If more cognitive resources are available while using the new software, those resources can be used for more in-depth processing of the data.
    - Errors should be less likely and gaining new insights into the data should be more likely

Summary Slide

- Goal of Literature Review
  - To educate the team members
    - **We read and discussed about 200 papers**
  - To stimulate discussion and generate ideas on
    - Application evaluation
      - **Examined different types of evaluations**
    - User interviews and knowledge elicitation
      - **Discussed some methodologies for this**
      - **Read some papers that were about IC specifically**
  - To look for gaps in the current methodologies
    - **These are the next steps – we are already discussing possibilities**

Time for Discussion

- How do we know if the NGC tools . . .
  - save analysts time?
  - reduce effort?
  - help to minimize errors?
  - enable analyses that weren't possible before?
  - change the analysts' work flow?
  - account for differences in individual analysts' styles
- How do we know if the NGC team's evaluation . . .
  - has a large enough sample size
  - does a fair comparison between old and new ways of doing things.