“Show Me What I Want to See”: Providing a Better Overview by Reducing Visual Complexity

Hanseung Lee
Catherine Plaisant
Ben Shneiderman

University of Maryland
Department of Computer Science

May 23, 2013
What do I have to do Next?

First look at this complex overview

What do I have to do Next?
Contributions

• Goal:
  • Develop and integrate a Visual Complexity Guideline software in EventFlow.

• Contributions:
  • Five different metrics that can measure visual complexity in EventFlow.
  • Interface that suggest (Align, Remove) sets for users to help reduce visual complexity.

• Examples:
  ✓ (Align with Event A)
  ✓ (Remove Event X, Remove Event Y)
  ✓ (Align with Event A, Remove Event X, Remove Event Y)
Five Visual Complexity Measures

- Total Number of Visual Elements
- Number of Visual Elements per Record
- Average Element Height
- Max Element Height
- Max Number of Record

Terms:
- Visual element: colored vertical bar in the aggregated view
- Element height: the height of the vertical bar as a percentage of the display height
Visual Complexity Measures

1) Total Number of Visual Elements

- Total # of Visual Elements = 72
- Total # of Visual Elements = 424
Visual Complexity Measures

2) Number of Visual Elements per Record

Total # of Visual Elements = 424
Number of Visual Elements per Record = 10.06

Total # of Visual Elements = 201
Number of Visual Elements per Record = 22.3
Visual Complexity Measures

3) Average Element Height

\[
\text{Average Element Height} = \frac{5.35}{40} = 0.134
\]

\[
\text{Average Element Height} = \frac{1.45}{40} = 0.036
\]
Visual Complexity Measures

4) Max Element Height

- No Alignment (original)
- Aligned with Green Event: Max Height = 0.44
- Aligned with Blue Event: Max Height = 0.75

😊
Visual Complexity Measures

5) Max Number of Records

- No Alignment (original)
  - Max Record = 4

- Aligned with Green Event
  - Max Record = 4

- Aligned with Blue Event
  - Max Record = 3
Visual Complexity Measures

Max Element Height vs. Max Number of Records

No Alignment (original)  
Aligned with Green Event  
Aligned with Blue Event

Max Height = 0.44  
Max Record = 4

Max Height = 0.75  
Max Record = 3
Visual Complexity Measures

6) Rate of Convergence to Individual Instances

- Find an exponential curve $y = Ae^{-\alpha x}$ with the best fit.
- In a random sequence of events, the probability of two sequences having the same prefix of length $n$ will be proportional to $e^{-n}$.
Summary of Visual Complexity Measures

• Five visual complexity measures have different characteristics.
• Users perceive visual complexity in various ways based on data and tasks.
• Different measures can be selected based on users’ intention.

<table>
<thead>
<tr>
<th>Absolute Values</th>
<th>Total Number of Visual Elements</th>
<th>Max Number of Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Values</td>
<td>Number of Visual Elements per Record</td>
<td>Max Element Height</td>
</tr>
</tbody>
</table>
Users can select the number of actions shown in the main list.
User Interface & Interaction

Five different metrics on visual complexity measurement can be selected.

- number of visual elements
- number of visual elements per record
  - average element height
  - max element height
- max number of records
Users can use the checkbox to select categories of interest.

- Focus categories: used to calculate visual complexity
- Search categories: used to search categories in suggested action sets
User Interface & Interaction

- **Focus categories**: used to calculate visual complexity
- **Search categories**: used to search categories in suggested action sets

---

Users can use the checkbox to select categories of interest.

- Focus categories: used to calculate visual complexity
- Search categories: used to search categories in suggested action sets
User Interface & Interaction

Visual Complexity Guidelines

Control Panel
- # of Actions: Using two actions
- Sort By: Average Height
- Order: Ascending
- Which Side?: Left
- Filter By: Align
- Focus on: click categories below...
- Search for: click categories below...

Using two actions

Using a single action

Using three actions

Possible Previous Action List
Possible Next Action List
User Interface & Interaction

Visual Complexity Guidelines

Control Panel
- # of Actions: Using two actions
- Sort By: Average Height
- Order: Ascending, Descending
- Which Side?: Left, Right
- Filter By: Align, Only Remove
- Focus on: click categories below...
- Search for: click categories below...

Actions: R: Remove, A: Align

Main Action List

Possible Previous Action List

Possible Next Action List

Using two actions
- R: ▲ R: ▲
- R: ▲ A: ▲
- R: ▲ A: ▲
- R: ▲ A: ▲
- A: ▲ R: ▲
- A: ▲ R: ▲
- A: ▲ A: ▲
- A: ▲ A: ▲
- A: ▲ A: ▲
- A: ▲ A: ▲
- A: ▲ A: ▲

Using a single action
- R: ▲
- A: ▲

Using three actions
- R: ▲ R: ▲ A: ▲
- R: ▲ R: ▲ A: ▲
- R: ▲ A: ▲
- R: ▲ A: ▲
- R: ▲ A: ▲
- R: ▲ A: ▲
- R: ▲ A: ▲
- R: ▲ A: ▲
- R: ▲ A: ▲
- R: ▲ A: ▲
Case Studies

Three Case Studies

Professor data set

Diabetes data set

Basketball data set
Case Study 1. Professor Data Set

- Professor Data Set
  - Total # of Records = 40
  - Total # of Events = 767
  - # of Point Event Categories = 7
  - # of Interval Event Categories = 3
Case Study 1. Professor Data Set

• Observation 1. Understanding the Data More Quickly

✓ Taking faster actions, which are placed at the top first 2~3 on the list
✓ Can easily capture the average time periods between events
✓ Found outliers, who became an assistant professor before earning a Master’s degree.
Case Study 1. Professor Data Set

- Observation 2. Obtaining More Information on Alignment

✓ After the “Associate Professor” Event
✓ There are few number of publications
✓ Some professor have no publications at all
Case Study 2. Diabetes Data Set

- Diabetes Data Set
  - Total # of Records = 97
  - Total # of Events = 3575
  - # of Point Event Categories = 7
  - # of Interval Event Categories = 2
Case Study 2. Diabetes Data Set

Let’s align this with a certain event.

Can you guess which alignment will give the fewest visual elements on the left side?
Case Study 2. Diabetes Data Set

- Observation 1. Obtaining More Information on Alignment

- Around 17% of patients did not have a previous event before “diabetes”

- A large portion of the patients had several “visit: clinic” event before “diabetes”
Case Study 2. Diabetes Data Set

- Observation 2. Finding Evidences of Subsequences

- When aligned with “VISIT: IP/OBS WITH ULCER”, the overview has the maximum height on its right side.

- “diabetes” event is the highest visual event right after the “visit: w/ulcer” event

- “visit: ip/obs with ulcer” → “diabetes” is the most frequent subsequence event
Case Study 3. Basketball Data Set

- Basketball Data Set
  - Total # of Records = 74
  - Total # of Events = 477
  - # of Point Event Categories = 9
  - # of Interval Event Categories = 2
Case Study 3. Basketball Data Set

• Observation 1. Obtaining More Information on Alignment
Demo
Visual Analytics Process

I did some analysis using these categories.
OK.
Now I understand the data. I just want to add two more categories in this visualization.
Visual Analytics Process

Let’s add two categories back to this aggregated view!
Visual Analytics Process – Current Aggregation

What just happened?
New Aggregation: An Incremental Aggregation

Now… let’s add the two categories back in while preserving the existing visualization

Much better !!!
Conclusions and Future Work

• Introduce a novel visual complexity guideline interface
  • Guiding users to reduce visual clutter in the aggregated overview
  • Supporting users to find a clearer aggregated overview

• Conduct simple case studies
  • Understanding the data more quickly
  • Obtaining more clues on Alignment, Remove, and Subsequences

• Future work:
  • Conducting more user studies to evaluate metrics
  • Improving the aggregation method
Thank you for your attention!
hanseung@cs.umd.edu

Questions, comments, or feedbacks?

Special thanks to Megan Monroe, Rongjian Lan, and Tak Yeon Lee