John W. Tukey (1915-2000)

“The best single device for suggesting, and at times answering, questions beyond those originally posed is the graphical display.”
Exploratory Analysis of Relational Data
Main Idea

Formal framework for extracting, manipulating, and drawing graphs implied by a relational database.
Node and Edge Tables

<table>
<thead>
<tr>
<th>paper_ID</th>
<th>title</th>
<th>year_binned</th>
<th>authors</th>
<th>cite_count</th>
</tr>
</thead>
<tbody>
<tr>
<td>806819</td>
<td>A reflectance mode</td>
<td>1980-1985</td>
<td>Robert L. Cook, Ke</td>
<td>23</td>
</tr>
<tr>
<td>801255</td>
<td>Light reflection func</td>
<td>1980-1985</td>
<td>James F. Blinn</td>
<td>24</td>
</tr>
<tr>
<td>806820</td>
<td>Vectorized procedure</td>
<td>1980-1985</td>
<td>Nelson Max</td>
<td>7</td>
</tr>
</tbody>
</table>

Citations

<table>
<thead>
<tr>
<th>citers</th>
<th>cites</th>
</tr>
</thead>
<tbody>
<tr>
<td>807414</td>
<td>563871</td>
</tr>
<tr>
<td>807456</td>
<td>807361</td>
</tr>
<tr>
<td>807507</td>
<td>807360</td>
</tr>
</tbody>
</table>

Graph Interpretation

Node Table
- Nid

Node Size

Node Color

Node Label

Edge Table
- N2
- N1

The rendering

The hemi-cub

Distributed

Ray tracing

Rendering algorithm

Light reflection

Shade trees

Simulation

A simple model

Ray tracing

3-D transfer

Surface rendering

Models of light
node and edge tables

\[ \begin{array}{ccc}
  a_1 & a_2 & a_3 \\
  \vdots & \vdots & \vdots \\
  a_1 & a_2 & a_3 \\
  \vdots & \vdots & \vdots \\
  \vdots & \vdots & \vdots \\
  \vdots & \vdots & \vdots \\
  \vdots & \vdots & \vdots \\
  \end{array} \]

graph interpretations

\[ \begin{array}{ccc}
  a_1 & a_2 \\
  \vdots & \vdots \\
  a_1 & a_2 \\
  \vdots & \vdots \\
  \vdots & \vdots \\
  \vdots & \vdots \\
  \vdots & \vdots \\
  \end{array} \]

graph visualizations

operators

implemented with relational algebra
defined semantically for graphs
**edge table**

<table>
<thead>
<tr>
<th>N1</th>
<th>N2</th>
<th>citation date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sarah</td>
<td>John</td>
<td>2001</td>
</tr>
<tr>
<td>John</td>
<td>Pam</td>
<td>2000</td>
</tr>
<tr>
<td>George</td>
<td>Pam</td>
<td>1997</td>
</tr>
<tr>
<td>Mary</td>
<td>Pam</td>
<td>1998</td>
</tr>
<tr>
<td>Mary</td>
<td>Pam</td>
<td>1999</td>
</tr>
<tr>
<td>Carlos</td>
<td>Pam</td>
<td>2000</td>
</tr>
<tr>
<td>Carlos</td>
<td>Fred</td>
<td>2000</td>
</tr>
<tr>
<td>Harry</td>
<td>Carlos</td>
<td>2002</td>
</tr>
<tr>
<td>Harry</td>
<td>Carlos</td>
<td>2004</td>
</tr>
</tbody>
</table>

**node table**

<table>
<thead>
<tr>
<th>N_id</th>
<th>name</th>
<th>salary</th>
<th>school</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>160K</td>
<td>Yale</td>
<td></td>
</tr>
<tr>
<td>Mary</td>
<td>70K</td>
<td>MIT</td>
<td></td>
</tr>
<tr>
<td>Sarah</td>
<td>80K</td>
<td>Yale</td>
<td></td>
</tr>
<tr>
<td>Carlos</td>
<td>95K</td>
<td>Harvard</td>
<td></td>
</tr>
<tr>
<td>Fred</td>
<td>65K</td>
<td>Duke</td>
<td></td>
</tr>
<tr>
<td>George</td>
<td>80K</td>
<td>MIT</td>
<td></td>
</tr>
<tr>
<td>Pam</td>
<td>120K</td>
<td>MIT</td>
<td></td>
</tr>
<tr>
<td>Harry</td>
<td>55K</td>
<td>Duke</td>
<td></td>
</tr>
</tbody>
</table>

**graph interpretation**

- T_E: citations table
- N1: author1
- N2: author2
- T_N: authors table
- N_id: name

\[ f_{n}: \text{node color} \rightarrow \text{school} \]
\[ f_{n}: \text{node size} \rightarrow \text{salary} \]

**graph visualization**

- Aggregate the graph by the school attribute results in a new graph interpretation, and corresponding visualization.
Network Intrusion Detection

G-Café

Ling Xiao
Joel Brandt
Nagendra Modadugu

With Prof. Pat Hanrahan
Proposed Approach

Design a system that combines
- Data mining methods
- Visualization techniques

For a system that can intelligently interact with the system administrator.
Approach: Idea

- Incorporate user domain knowledge
- Interactive system
  - Administrator does the thinking
  - Algorithm does the grunt work
Shaded Relief and Contours from Elevation Data

-Mike Cammarano
Motivation for LOD

Figure 24.6  Two representations of Great Britain and Ireland at the scale of about 1:15,000,000. The map on the left (A) is simplified to fit the scale and is suitable for a reference map intended to give the impression of detailed precision. The map on the right (B) is a diagrammatic generalization suitable as a base on which to display thematic data. Note that (B) captures basic shapes, which tend to be masked by the detail in (A).
Motivation for LOD
## Cartographic Generalization

- **Selection**
- **Simplification**
- **Exaggeration**
- **Regularization**
- **Displacement**
- **Aggregation**

(Monmonier 96), [MacEachren 94], [DiBiase 91]
Approach
Shading Model
Motivation

• Standard online route maps difficult to use
Three Generalizations for Route Maps

• Our observations from handdrawn examples:
  • Distortion
    • Road length
    • Turning angle
  • Simplification
    • Road shape

• Generalizations emphasize turning points!
LineDrive: Route Map Design System

Handdrawn route map

LineDrive route map
Results: Bellevue to Seattle
Flow Map Layout

Doantam Phan
Ling Xiao
Ron Yeh
Pat Hanrahan
Terry Winograd

25 October 2005
Related Work

[Tobler 2004]
Carte en représentation des quantités de Vin Français exportées par mer en 1864

[Minard 1864]
Geographic Distortion
Edge Merging
Edge Routing
System Diagram

- Layout Adjustment
- Primary Clustering
- Rooted Clustering
- Spatial Layout
- Edge Routing

(Geometric Distortion)
(Edge Merging)
(Edge Routing)
Visualizing Dataspaces

Mike Cammarano
G-café, 23 Feb. 2006

With: Pat Hanrahan
Alon Halevy
Dan Ramage
Too many bits, in too many places.

Our emphasis is on personal information collections.
Dataspace

- Associations and attributes added by many independent sources.
Reorganize on-demand

- How do I ...
  - say what I want (query)
  - get around (browse)

- How does the computer ...
  - present the content (vis)

Goal: Seamless information space.
Overview

• What’s old:
  – Text search
  – Faceted navigation

• What’s current:
  – Dataspaces and semantic web
  – Semex
  – Other associative browsers

• What are we contributing:
  – Pervasive multiple-selection
  – Browse via sets of associations
  – Fit visualizations with schema resolution
Flamenco
Faceted navigation demo
Visualization-matching

• Like schema-matching

Figure 1: Computer Science Department Ontologies
Visualization-matching

• Fit objects into a particular visualization.
  – If they are missing a needed attribute, search for a similar one nearby.
Visualization matching

• Need: Person->Image
  • Try: Person->HomePage->Image

• Need: File->Author
  • Try: File->ContainedIn->Author
Interactive Gigapixel Displays

fluid interaction on large paper surfaces

Ron Yeh
Jonas Boli
Joel Brandt
Scott Klemmer

G-Café · 16 March 2006
Mission Statement

Design and Develop...

1) visualizations for large, paper-based displays
2) techniques for interacting with them.
Benefits

- Familiarity
- Size & Amount of Data
- Robustness & Permanence
- Physicality
- Resolution
- Flexibility
- Mobility
- Collaboration
Drawbacks of Paper-Based Interfaces

- Static Organization of Content
- SLOW Refresh Rate for Output
- No Computation (e.g., Text Search)
- Lacks Network Connectivity
- Not enough Data Storage
Combine Them!

Semi-Static Output
- Print a Visualization on a Wide-format Printer

Real-Time Output
- Mobile Device
- Nearby Display (LCD or Laptop)
- Projected Overlay

Input
- Digital Pen
Applications
Applications: Maps

- Biology
- Astronomy
- Archaeology
- Architecture
- Tourism
- Subway/Bus
- Ski Resorts 😊
social networks: vizster
Summary & Contributions

Large Interactive Paper Surfaces

Research
• Design Space for Interactive Gigapixel Displays
• Technique for Rapid Prototyping & WOz

Future
• Library for developing interactive visualizations
• Toolkit for new interactions & visualizations
Reorderable Matrices
&
Voting patterns in the US House of Representatives

Flash G
7 Feb 2006
Veterans and Dependants Millenium Education Act (S 1402)
On Motion to Suspend the Rules and Pass, as Amended
Passed (417-0-0-17)
Bipartisan 199-0  216-0
Education
Smith's vote: Yea
Thanks

- Maneesh Agrawala
- Dave Akers
- Joel Brandt
- Joel Crosby
- Pat Hanrahan
- Alon Helevy
- John Gerth
- Doantam Phan
- Barbara Tversky
- Ling Xiao
- Ron Yeh