How do people create visualizations?

**Chart Typology**
- Pick from a stock of templates
- Easy-to-use but limited expressiveness
- Prohibits novel designs, new data types

**Component Architecture**
- Permits more combinatorial possibilities
- Novel views require new operators, which requires software engineering

Today’s first task is not to invent wholly new [graphical] techniques, though these are needed. Rather we need most vitally to recognize and reorganize the essential of old techniques, to make easy their assembly in new ways, and to modify their external appearances to fit the new opportunities.

Chart Typologies
Excel, Many Eyes, Google Charts

Visual Analysis Languages
Tableau VizQL, ggplot2, HiVE

Component Model Architectures
Improvise, Prefuse, Flare

Graphics APIs
OpenGL, Java2D, GDI+, Processing

Exploratory Data Analysis
Protovis - A Graphical Toolkit for Visualization
**d3.js** Data-Driven Documents

- Specialized mark types
  - Streamlined design
  - Limits expressiveness
  - More overhead (slower)
  - Harder to debug
  - Self-contained model

**Protovis**

- Specify a scene (nouns)
  - Quick for static vis
  - Animation, interaction are more cumbersome

**D3**

- Transform a scene (verbs)
  - More complex model
  - Dynamic data, animation, and interaction natural

---

**Interaction**

[There is an] apparent challenge that computational artifacts pose to the longstanding distinction between the physical and the social, in the special sense of those things that one designs, builds, and uses, on the one hand, and those things with which one communicates, on the other.

"Interaction" – in a sense previously reserved for describing a uniquely interpersonal activity – seems appropriately to characterize what goes on between people and certain machines as well.

Lucy Suchman, *Plans and Situated Actions*
**Interaction** between people and machines requires *mutual intelligibility* or *shared understanding*.

**Gulf of Execution**
The difference between the user's intentions and the allowable actions.

**Gulf of Evaluation**
The amount of effort that the person must exert to interpret the state of the system and to determine how well the expectations and intentions have been met.

---

**Gulfs of Execution & Evaluation**

- **Gulf of Execution**
  - Real world
  - Conceptual model
  - Evaluation

- **Gulf of Evaluation**
  - Real world
  - Conceptual model
  - Evaluation

---

**Gulf of Evaluation**

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.07</td>
<td>0.79</td>
</tr>
<tr>
<td>0.38</td>
<td>0.70</td>
</tr>
<tr>
<td>0.27</td>
<td>0.85</td>
</tr>
<tr>
<td>0.71</td>
<td>0.45</td>
</tr>
<tr>
<td>0.52</td>
<td>0.70</td>
</tr>
<tr>
<td>0.28</td>
<td>0.54</td>
</tr>
<tr>
<td>0.51</td>
<td>0.36</td>
</tr>
<tr>
<td>0.11</td>
<td>0.33</td>
</tr>
<tr>
<td>0.16</td>
<td>0.45</td>
</tr>
</tbody>
</table>
Gulf of Evaluation

Real world:

Conceptual model:
x,y correlated?

Evaluation

ρ = -.29

Gulf of Execution

Real world:

Conceptual model:
Draw a scatterplot

Execution

Move 90 30
Rotate 35
Pen down

Interaction Techniques

Are there “essential” interactive operations for supporting exploratory data visualization?

- View Specification (map data to visual vars)
- Navigation (pan, zoom, scale, rotate)
- Selection / Highlighting
- Filtering
- Sorting
- Extract Data

Interactive Visualization

(Graphics and Graphic Information Processing, Bertin 81)
Basic Pointing Methods

Point Selection
Mouse Hover / Click
Touch / Tap
Select Nearby Element (e.g., Bubble Cursor)
Basic Pointing Methods

Point Selection
- Mouse Hover / Click
- Touch / Tap
- Select Nearby Element (e.g., Bubble Cursor)

Region Selection
- Rubber-band or Lasso
- Area Cursors (“Brushes”)

Highlighting / Brushing

Direct attention to a data subset within a graph [Wills, 95]
**Brushing and Linking**

Select (“brush”) a subset of data
See selected data in other views

The components must be *linked*
by *tuple* (matching data points), or
by *query* (matching range or values)

---

**Baseball Statistics** [from Wills 95]

- how long in majors
- avg assists vs avg putouts (fielding ability)
- distribution of positions played
- select high salaries
- avg career HRs vs avg career hits (batting ability)

---

**Linking Assists to Positions**
GGobi: Brushing

http://www.ggobi.org/

Dynamic Queries

Query and Results

SELECT house FROM palo_alto_homes
WHERE price < 1,000,000 AND bedrooms > 2
ORDER BY price

Issues

1. For programmers
2. Rigid syntax
3. Only shows exact matches
4. Too few or too many hits
5. No hint on how to reformulate the query
6. Slow question-answer loop
7. Results returned as table
### Direct Manipulation

1. Visual representation of objects and actions
2. Rapid, incremental and reversible actions
3. Selection by pointing (not typing)
4. Immediate and continuous display of results
Alphaslider

Title: Moonstruck

[Ahlberg and Shneiderman '94]

Attribute Explorer [Spence and Tweedie '98]

- Video Clip

Zipdecode [Fry '04]

[http://benfry.com/zipdecode/]

[Ahlberg and Shneiderman '94]
Name Voyager

http://www.babynamewizard.com/voyager

TimeSearcher [Hochheiser & Shneiderman 02]

Based on Wattenberg’s [2001] idea for sketch-based queries of time-series data.

3D dynamic queries [Akers et al. 04]

**Pros and Cons**

**Pros**
- Controls useful for both novices and experts
- Quick way to explore data

**Cons**
- Simple queries
- Lots of controls
- Amount of data shown limited by screen space
- Who would use these kinds of tools?

**Trellis Display**
[Becker, Cleveland, and Shyu 96]

**Sorting**

**Condition variables**
- location, year

**Panel variables**
- type, yield

**Trellis Display**
[Becker, Cleveland, and Shyu 96]
Alphabetical ordering

Main-effects ordering

Graph Viewer

Graph Viewer
Assignment 2: Visual Data Analysis

Use visualization software (Tableau) to form and answer questions

First steps:
- Step 1: Pick a domain
- Step 2: Pose questions
- Step 3: Find Data
- Iterate

Create visualizations:
- Interact with data
- Refine your questions

Make wiki notebook:
- Keep record of your analysis
- Prepare a final graphic and caption

Due by 1:00pm on Tuesday, October 16

Assign 3: Interactive Visualization

Create an interactive visualization application. Choose a data domain and select an appropriate visualization technique.

1. Choose a data set and storyboard your interface
2. Implement the interface using tools of your choice
3. Submit your application and produce a final write-up

You should work in groups of 2.
Due by end of day on Tuesday, October 30
Assign 3: Project Partners

For A3, you should work in groups of 2.
If you do not have a partner, you should
1) Add your name + interests to the ProjectPartners wiki page
2) Stay after class to meet potential partners

Assign 3 Tips

1) Start now. It will take longer than you think.
2) Keep it simple. (Eschew the kitchen sink.) Choose the minimal set of interactions that enables users to explore and generate interesting insights. Keep the design clean.
3) Promote engagement. How do your chosen interactions reveal interesting observations?

Generalized Selection

Visual Queries

Model selections as declarative queries.

(-118.371 ≤ lon AND lon ≤ -118.164) AND (33.915 ≤ lat AND lat ≤ 34.089)
**Visual Queries**

Model selections as *declarative queries* over the domain of visualized data.

Applicable to *dynamic, time-varying data*

Retarget selection *across visual encodings*

Perform operations on *query structure*

**Generalized Selection**

Point to an example and define an abstraction based on one or more properties [Clark, Brennan]

"Blue like this"
"The same shape as that"

Abstraction may occur over multiple levels

"Select items like this one."

This is not a sentence.
**Generalized Selection**

Provide *generalization mechanisms* that enable users to expand a selection query along *chosen dimensions* of interest.

Expand selections via **query relaxation**

**Query Builder**

- **Click:** Select Items
  - *(id = ‘China’)*

- **Drag:** Select Range
  - *(2000 < gni AND gni < 10000) AND (.1 < internet AND internet < .2)*

**Legend:**
- Select Attributes
  - *(region = ‘The Americas’)*
Query Relaxation

Generalize an input query to create an expanded selection, according to:

1. A semantic structure describing the data
2. A traversal policy for that structure
**Relaxation using Hierarchies**

Relax using *abstraction hierarchies* of the data. Traverse in direction of increasing generality.

**Examples**
- *A Priori*: Calendar, Categories, Geography
- *Data-Driven*: Nearest-Neighbor, Clustering

---

**Relaxation using Attributes**

If no explicit semantic structure is available, treat data itself as a “flat” hierarchy.

Select all items with matching values along the attributes chosen for relaxation.

---

**Relaxation of Networks**

Diagram showing a network with nodes and edges, illustrating the concept of relaxation in networks.
Lesson

Consider how the structure and/or semantics of the data might be leveraged to aid analysis.

One extension: look beyond data features to incorporate perceptual features of the display.

Summary

**Most visualizations are interactive**
Even passive media elicit interactions

**Good visualizations are task dependent**
Pick the right interaction technique
Consider the semantics of the data domain

**Fundamental interaction techniques**
Selection / Annotation, Sorting, Navigation, Brushing & Linking, Dynamic Queries