Lecture 4: Frameworks/Models
Information Visualization
CPSC 533C, Fall 2007
Tamara Munzner
UBC Computer Science
19 September 2007

Visualization Big Picture

Mapping
- input
  - data semantics
  - use domain knowledge
- output
  - visual encoding
    - visual/graphical/visual/retinal
  - use human perception
  - processing
    - algorithms
    - handle computational constraints

Data Types
- continuous (quantitative)
  - 10 inches, 17 inches, 23 inches
- ordered (ordinal)
  - small, medium, large
days: Sun, Mon, Tue, ...

Channel Ranking Varies by Data Type
- spatial position best for all types

Mackinlay, Card
- data variables
  - 1D, 2D, 3D, 4D, 5D, etc
- data types
  - nominal, ordered, quantitative
- marks
  - point, line, area, surface, volume
  - geometric primitives
- retinal properties
  - size, brightness, color, texture, orientation, shape
  - parameters that control the appearance of geometric primitives
  - separate channels of information flowing from retina to brain

Shneiderman’s Data+Tasks Taxonomy
- data
  - 1D, 2D, 3D, temporal, rd, trees, networks
- text and documents (Hanrahan)
- tasks
  - overview, zoom, filter, details-on-demand,
    - relate, history, extract
- data alone not enough
  - what do you need to do?
- mantra: overview first, zoom and filter, details on demand

Further Readings
The Structure of the Information Visualization Design Space Stuart Card and Jock Mackinlay, Proc. InfoVis 97
[graphics.stanford.edu/papers/polaris]

Papers Covered
The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations Ben Shneiderman, Proc. 1996 IEEE Visual Languages, also Maryland HCDL, TR 96-13 [citeseer.ist.psu.edu/shneiderman6eyes.html]

Frameworks
- Mackinlay/Card/(Bertin)
  - Data Types, Marks, Retinal Attributes (incl Position)
- Shneiderman, Amar/Eagan/Starke
  - Data, Tasks
- Tory/Moeller, Hanrahan
  - Data/Conceptual Models
- Stolte/Tang/Hanrahan, (Wilkinson)
  - Table Algebra ↔ Visual Interface
  - Value
### Control Room Example

Which location has the highest power surge for the given time period? (extreme y-dimension)

A fault occurred at the beginning of this recording, and resulted in a temporary power surge. Which location is affected the earliest? (extreme x-dimension)

Which location has the most number of power surges? (extreme count)

### Data Models vs. Conceptual Models

- **Data model:** mathematical abstraction
  - set with operations
  - e.g. integers or floats with +, *

- **Conceptual model:** mental construction
  - includes semantics, support data
  - e.g. navigating through city using landmarks

### Models Example

- **from data model**
  - 17, 25, -4, 28.6
  - (floats)
  - using conceptual model
  - (temperature)

### Time

- **2D+T vs. 3D**
  - same or different? depends on POV
  - time as input data?
  - time as visual encoding?

### Average

- **same**
  - time just one kind of abstract input dimension
  - different
    - input semantics
    - visual encoding: spatial position vs. temporal change

- **processings might be different**
  - e.g. interpolate differently across timesteps than across spatial position

### Combinatorics of Encodings

- **challenge**
  - pick the best encoding from exponential number of possibilities \( (n + 1)^8 \)

- **Principle of Consistency**
  - properties of the image should match properties of data

- **Principle of Importance Ordering**
  - encode most important information in most effective way

### Automatic Design

- **Mackinlay, APT**
  - Roth et al, Sage/Visage

- **select visualization automatically given data**
  - vs. Polaris: user drag and drop exploration

- **limited set of data, encodings**
  - scatterplots, bar charts, etc

- **holy grail**
  - entire parameter space

---

**Models Example**

- **from data model**
  - 17, 25, -4, 28.6
  - (floats)

- **using conceptual model**
  - (temperature)

**to data type**

- burned vs. not burned (N)
- hot, warm, cold (O)
- continuous to 4 sig figures (Q)

**using task**

- making toast
- classifying showers
- finding anomalies in local weather patterns

---

**Polaris:**

- Polaris: Circles, Profit/State:Month
- Polaris: Circles, State/Product:Month
- Polaris: Gantt Bar, Country/Time
- Polaris: Circles, Lat/Long

---

**Fields Create Tables and Graphs**

- **Ordinal fields:** interpret field as sequence that partitions table into rows and columns:
  - Quarter = (Qtr1), (Qtr2), (Qtr3), (Qtr4)

- **Quantitative fields:** treat field as single element sequence and encode as axes:
  - Profit = (Profit)

---

**Combinatorics of Encodings**

- **challenge**
  - pick the best encoding from exponential number of possibilities \( (n + 1)^8 \)

- **Principle of Consistency**
  - properties of the image should match properties of data

- **Principle of Importance Ordering**
  - encode most important information in most effective way

---

**Automatic Design**

- **Mackinlay, APT**
- **Roth et al, Sage/Visage**

- **select visualization automatically given data**
  - vs. Polaris: user drag and drop exploration

- **limited set of data, encodings**
  - scatterplots, bar charts, etc

- **holy grail**
  - entire parameter space
### Mackinlay’s Criteria
- **Expressiveness**
  - Set of facts expressible in visual language if sentences (visualizations) in language express all facts in data, and **only** facts in data.
- **Consider the failure cases...**

### Cannot Express the Facts
- A $1 \Rightarrow N$ relation cannot be expressed in a single horizontal dot plot because multiple tuples are mapped to the same position

### Expresses Facts Not in the Data
- Length interpreted as quantitative value
- Thus length says something untrue about nominal data

### Arguments
- **“avoid interaction”** dictum controversial
  - part of power of computer-based methods
  - but can degenerate into human-powered search
- presentation/exposition vs. exploration
- art vs. science vs. technology

### Summary
- **formal approach to picture specification**
  - declare the picture you want to see
  - compile query, analysis, and rendering commands needed to make the pictures
  - automatically generate presentations by searching over the space of designs
- Bertin’s vision still not complete
  - formalize data model
  - formalize the specifications
  - experimentally test perceptual assumptions
- much more research to be done...