**Papers Covered**


Further Readings

The Structure of the Information Visualization Design Space. Stuart Card and Jock Mackinlay, Proc. InfoVis 97. [citeseer.ist.psu.edu/card96structure.html]


The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. Ben Shneiderman. Proc. 1996 IEEE Visual Languages, also Maryland HCIL TR 96-13. [citeseer.ist.psu.edu/shneiderman96eyes.html]

Visualization Big Picture

**Data Types**

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

- **Ordered (ordinal)**
  - apples, oranges, bananas

- **Categorical (nominal)**
  - 10 inches, 17 inches, 23 inches

**Combinatorics of Encodings**

- **Challenge**
  - pick the best encoding from exponential number of possibilities ($n \times 1$)

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

- **Principle of Importance Ordering**
  - handle computational constraints

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

**Bertin’s Semiology of Graphics**

- geometric primitives: marks
- attributes: visual/retinal variables
- parameters control mark appearance
- separable channels flowing from retina to brain
- x-y position
- z
- size
greyscale
color
texture
orientation
shape

**More Data Types: Stevens**

- **Quantitative**
  - 10, 17, 23 inches

- **Ordered (ordinal)**
  - small, medium, large

- **Categorical (nominal)**
  - apples, oranges, bananas

- **Spatial position best for all types**

**Design Space = Visual Metaphors**

**Data Types**

- **Continuous (quantitative)**
  - 10 inches, 17 inches, 23 inches

- **Ordered (ordinal)**
  - small, medium, large

- **Categorical (nominal)**
  - apples, oranges, bananas

- **Spatial position best for all types**

**Channel Ranking Varies by Data Type**

**Lightness**

**Shape**

**Length**

**Angle**

**Slope**

**Area**

**Volume**

**Connection**

**Containment**

**Texture**

**Position**

**Principle of Consistency**

**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...**

**Mackinlay’s Criteria**

**Cannot Express the Facts**

A 1 ↔ N relation cannot be expressed in a single horizontal dot plot because multiple tuples are mapped to the same position

**Mapping**

- input
  - data semantics
  - use domain knowledge
- output
  - visual encoding
  - visual/retinal/perceptual/retinal
  - channels/attributes/dimensions/variables
  - use human perception
- processing
  - algorithms
  - handle computational constraints

**Combinatorics of Encodings**

- **Challenge**
  - pick the best encoding from exponential number of possibilities ($n \times 1$)

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

- **Principle of Importance Ordering**
  - handle computational constraints

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

**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...**

**Lightness**

**Shape**

**Length**

**Angle**

**Slope**

**Area**

**Volume**

**Connection**

**Containment**

**Texture**

**Position**

**Principle of Consistency**

**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...**

**Mackinlay’s Criteria**

**Cannot Express the Facts**

A 1 ↔ N relation cannot be expressed in a single horizontal dot plot because multiple tuples are mapped to the same position

**Visualization Big Picture**

**Data Types**

- **Continuous (quantitative)**
  - 10 inches, 17 inches, 23 inches

- **Ordered (ordinal)**
  - small, medium, large

- **Categorical (nominal)**
  - apples, oranges, bananas

**Combinatorics of Encodings**

- **Challenge**
  - pick the best encoding from exponential number of possibilities ($n \times 1$)

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

- **Principle of Importance Ordering**
  - handle computational constraints

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

**Lightness**

**Shape**

**Length**

**Angle**

**Slope**

**Area**

**Volume**

**Connection**

**Containment**

**Texture**

**Position**
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

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.

- Effectiveness
  - a visualization is more effective than another visualization if information conveyed by one visualization is more readily perceived than information in other.

- subject of the next lecture

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) ⇔

Beyond Data Alone

- bigger picture than just visual encoding decisions

Shneiderman’s data+task taxonomy

- data
  - 1D, 2D, 3D, temporal, nD, trees, networks

- text and documents (Hanrahan)

- tasks
  - preview, 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


Tasks, Amar/Eagan/Stasko Taxonomy

- low-level tasks
  - retrieve value, filter, compute derived value, find extremum, sort, determine range.

- characterize distribution, find anomalies, cluster, correlate

- standardized set for better comparison between papers
  - bottom-up grouping with affinity diagramming

- abstraction from domain task down to low-level task

[Amar, Eagan, and John Stasko. Low-Level Components of Analytic Activity in Information Visualization. Proc. InfoVis 05]

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)


Expresses Facts Not in the Data

- length interpreted as quantitative value
  - thus length says something untrue about nominal data

Mackinlay, APT

[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding]

Design: Designer vs. Automatic vs. User

- designer: studies last time

- automatic: select visualization automatically given data
  - Mackinlay, APT

- limited set of encodings: scatterplots, bar charts...

- Roth et al, Sage/Visage

- holy grail: entire space of infosiv visual encoding
  - nowhere near goal, esp. with relational/graph data

- human-guided: allow user to change encodings

- Polaris: user drag and drop exploration


[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding]


[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding]

[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding]
Threats To Validity: What Can Go Wrong?

- 1

Upstream and Downstream Validation

- humans in the loop for outer three levels
- time: same as one kind of input
- different
- data: semantics (time steps of dynamically changing data)
- input: visual encoding channel of temporal change very different than spatial position change
- processing might be different
- i.e. interpolate differently across timesteps than across spatial position

MatrixExplorer Views

- overviews: matrix, node-link, connected components
- details: matrix, node-link
- controls

Comparing Clusters

- automatic clustering as good starting point
- then manually refine

MatrixExplorer

- domain: social network analysis
- validation:
  - early: participatory design to generate requirements
  - later: qualitative observations of tool use by target users
- techniques:
  - interactively map attributes to visual variables
  - user can change visual encoding on the fly
  - filtering
  - selection
  - sorting by attribute

Automatic Clustering/Reordering

- automatic clustering as good starting point
- then manually refine

Requirements

- use multiple representations
- handle multiple connected components
- provide overviews
- display general dataset info
- use attributes to create multiple views
- display basic and derived attributes
- minimize parameter tuning
- allow manual finetuning of automatic layout
- provide visible reminders of filtered-out data
- support multiple clusterings, including manual
- support outlier discovery
- find where consensus between different clusterings
- aggregate, but provide full detail on demand

MatrixExplorer: a Dual-Representation System to Explore Social Networks.
www.aviz.fr/nhenry/docs/Henry-InfoVis2006.pdf

Comparing Clusters

- relax, check if clusters conserved
- then manually refine

Techniques: Dual Views

- show both matrix and node-link representations

2D+T vs. 3D

- same or different? depends on POV
- input vs. output side
- same
- time as just one kind of abstraction
- different
- input: spatial position
- output: visual encoding channel of temporal change very different than spatial position change
- processing might be different
- e.g. interpolate differently across timesteps than across spatial position

Nested Model

- separating design into levels
- not just the visual encoding level!
- cascading dependencies: outputs from level above are inputs to level below
- domain problem characterization
  - domain problem characterization
  - data/operation abstraction design
  - algorithm design

Pat Hanrahan

graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding