Visual Language is a Sign System

Image perceived as set of signs

Sender encodes information in these signs

Receiver decodes information from these signs

Frameworks

- Data, Tasks
- Data Types, Marks, Retinal Attributes (incl Position)
- Table Algebra ←→ Visual Interface
- Data/Conceptual Models

Readings


The Structure of the Information Visualization Design Space

Shneiderman, Jack Mackinlay. Proc. InfoVis 97

The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, Ben Shneiderman, Proc. 1996 BEE Visual Languages, also Maryland ICS TR 96-13


Mapping

input
- data semantics
- use domain knowledge

output
- visual encoding
  - visual/graphical/perceptual/retinal
  - channels/attributes/dimensions/variables
  - use human perception

processing
- algorithms
- handle computational constraints
**Bertin: Semiology of Graphics**

- geometric primitives: marks
  - points, lines, areas, volumes
- parameters control mark appearance
- separable channels flowing from retina to brain

[1, y] position
- size
- grayscale
- color
- texture
- orientation
- shape

**Design Space = Visual Metaphors**

![Design Space Diagram]


**Data Types**

- continuous (quantitative)
  - 10 inches, 17 inches, 23 inches
- ordered (ordinal)
  - small, medium, large
  - days: Sun, Mon, Tue, Wed, ...
- categorical (nominal)
  - apples, oranges, bananas

**More Data Types: Stevens**

- subdivides quantitative further:
  - interval: 0 location arbitrary
    - time: seconds, minutes
  - ratio: 0 fixed
    - physical measurements: Kelvin temp

[S.S. Stevens, On the theory of scales of measurements, Science 103:2694-2694, 1944]

**Channel ranking varies by data type**

Spatial position varies best for all types

[1, After, Automating the Design of Graphical Presentations of Numerical Information, ACM TCG'91. 1991]
**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
- separable channels of information flowing from retina to brain

Closest thing to central dogma we've got

---

**Shneiderman's Data+Tasks Taxonomy**

Data
- 1D, 2D, 3D, temporal, nD, trees, networks
- text and documents (Hamran)

Tasks
- Overview, Zoom, Filter, Details-on-demand,
- Relate, History, Extract

data alone not enough: what do you need to do?

[Shneiderman, The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations]

---

**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)

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

---

**Time**

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

same
- time just one kind of abstract input dimension

different
- input semantics
- visual encoding: spatial position vs. temporal change

processing might be different
- e.g. interpolate differently across times than across spatial position

---

**Polaris**

infovis spreadsheet

- table cell
- not just numbers: graphical elements
- wide range of retinal variables and marks
- extends Wilkinson

- table algebra --> interactive interface
- formal language
Mackinlay’s Expressiveness Criteria

Expressiveness

A set of facts is expressible in a visual language if the sentences in that language can express all the relationships in the input data, and only the facts in the data.

Cannot Express the Facts

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

Expresses Facts Not in the Data

A length is interpreted as a quantitative value. Length says something untrue about N data points.

Automatic Design

Mackinlay, APT
Roth et al, SAGE

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

Summary

Formal approach to picture specification
- Declare the picture you want to see
- Compile query, analyze, and rendering commands needed to make the picture
- Automatically generate presentations
- Searching over the space of designs
Berlin’s vision still not complete
  - Formalize data model
  - Formalize the specifications
  - Experimentally test perceptual assumptions
Much more research to be done in this area...
**Value of Vis**

\[ I(t) = V(D,S,t) \]
- Data D transformed by spec S into time-varying image
- \( dK/dt = P(I,K) \)
- Perception P of image by user increases knowledge K

\[ S(t) = S_0 + \text{integral} \ E(k) \]
- Interactive exploration \( E \) changes spec

---

**Cost model**

**costs**
- \( C.I.S.O \) : initial development costs
- \( C.u.S,I \) : initial per-user costs
- \( C.s,S,O \) : initial per-session costs
- \( C.e \) : perception and exploration costs

**benefit**
- \( G = \text{nnM}(\text{deltaK}) \)

**profit**
- \( F = G - C \)
- \( F = \text{nnM}(\text{deltaK}) - C.s - \text{kC}_c - C.j - \text{mC}_u \)

---

**Arguments**

- new methods not better by definition
- vis not good by definition
- must show why automated extraction insufficient
  - e.g., automation not foolproof
- if no clear patterns
  - method limitation?
  - wrong parameters?
  - or truly not there in data?
- inspire new hypotheses vs. verify final truth
  - 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

---

**Credits**

Pat Hanrahan
  (graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding)

Torsten Moeller, Melanie Tory
  - discussions