Frameworks/Models

Lecture 5 CPSC 533C, Fall 2005

26 Sep 2005

Tamara Munzner

Readings

Chapter 1, Readings in Information Visualization: Using Vision to Think.
Stuart Card, Jock Mackinlay, and Ben Shneiderman, Morgan Kaufmann 1999.

The Structure of the Information Visualization Design Space
Stuart Card and Jock Mackinlay, Proc. InfoVis 97

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

Polaris: A System for Query, Analysis and Visualization of
Multi-dimensional Relational Databases.
Chris Stolte, Diane Tang and Pat Hanrahan, IEEE TVCG 8(1), January 2002.

The Value of Visualization. Jarke van Wijk. Visualization 2005, to
appear.

Frameworks

Shneiderman

- Data, Tasks

Mackinlay/Card/(Bertin)

- Data Types, Marks, Retinal Attributes (incl Position)

Stolte/Tang/Hanrahan, (Wilkinson)

- Table Algebra <--> Visual Interface

Hanrahan, Tory/Moeller

- Data/Conceptual Models

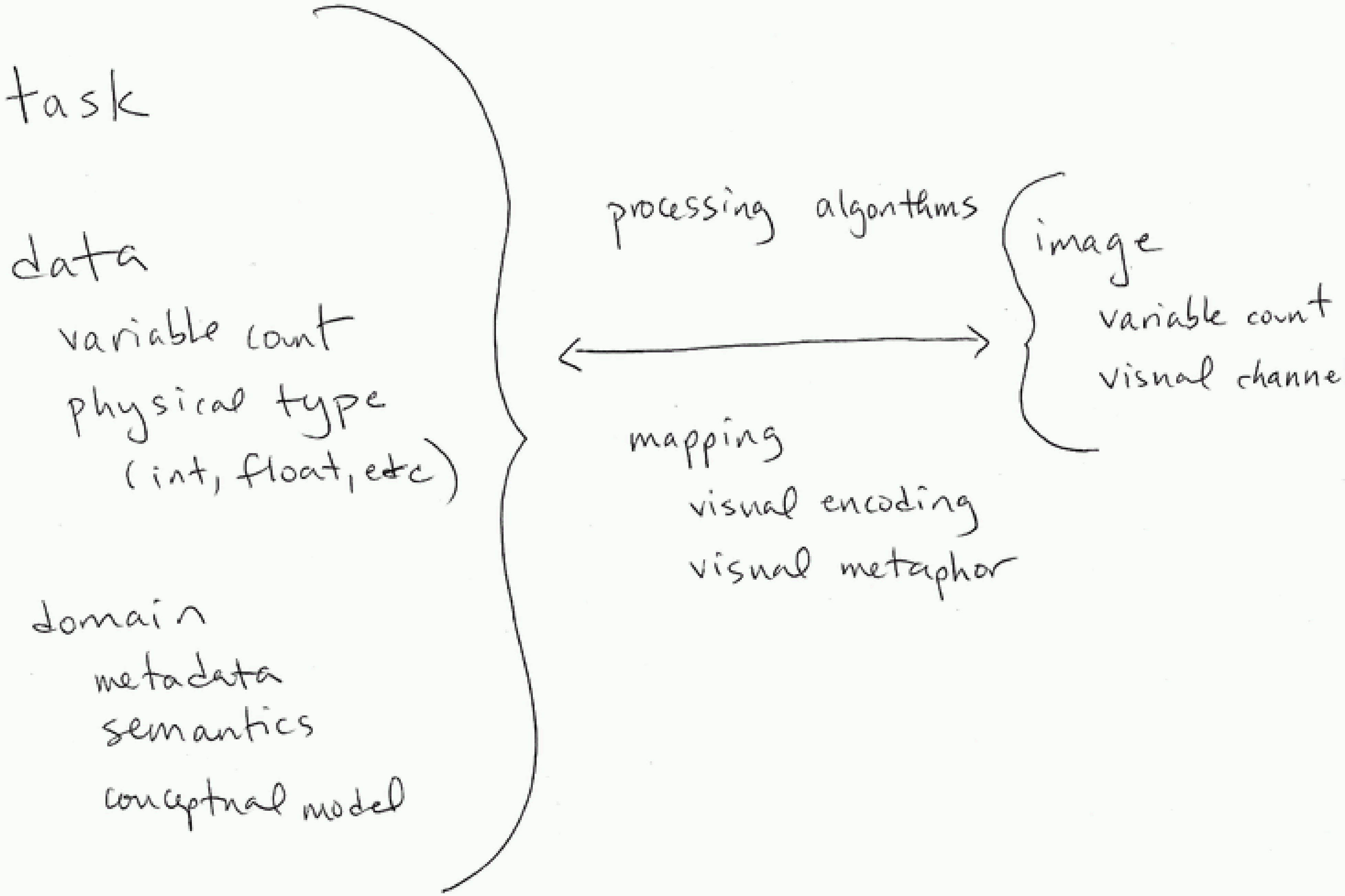
Visual Language is a Sign System

Image perceived as set of signs

Sender encodes information in these signs

Receiver decodes information from these signs

Visualization Big Picture



Mapping

input

- data semantics
- use domain knowledge

output

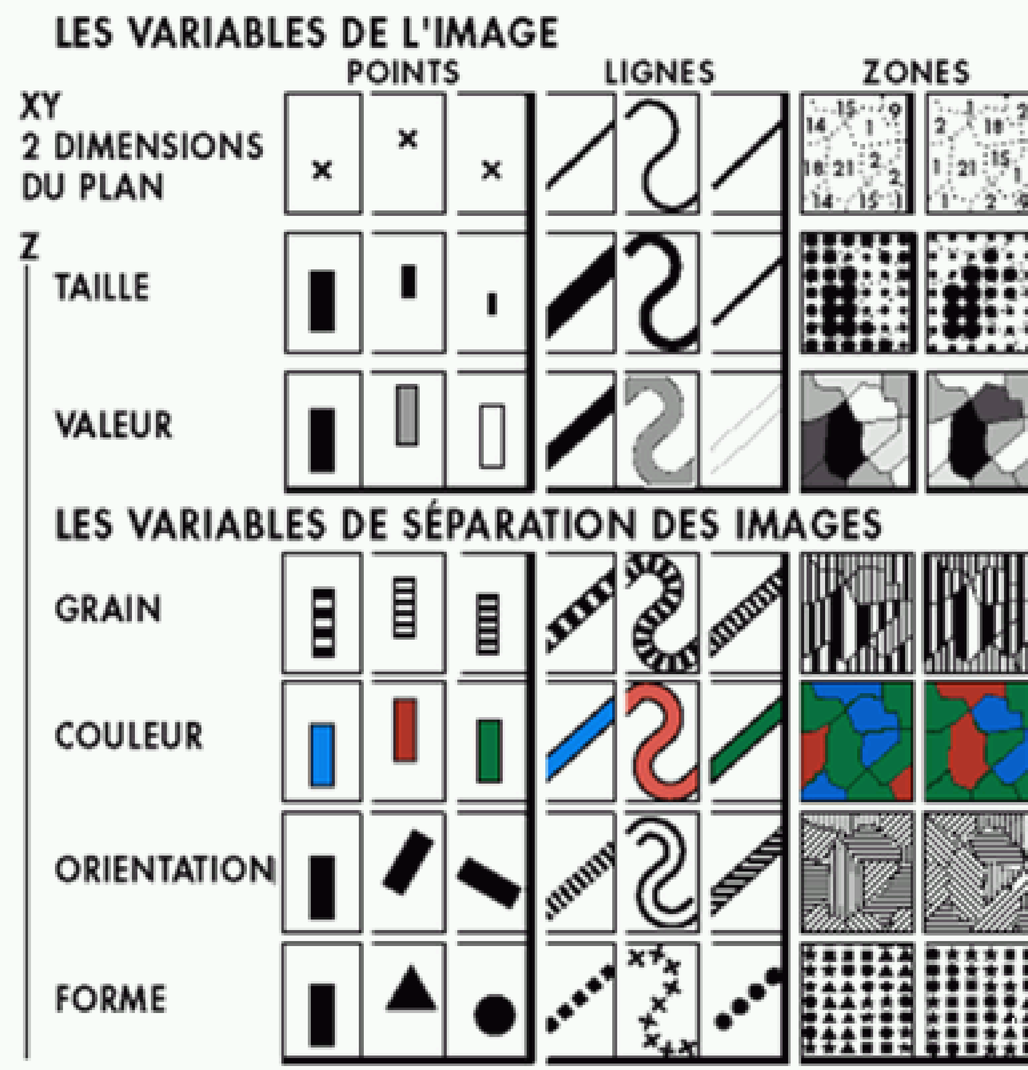
- visual encoding
  - visual/graphical/perceptual/retinal
  - channels/attributes/dimensions/variables
- use human perception

processing

- algorithms
- handle computational constraints

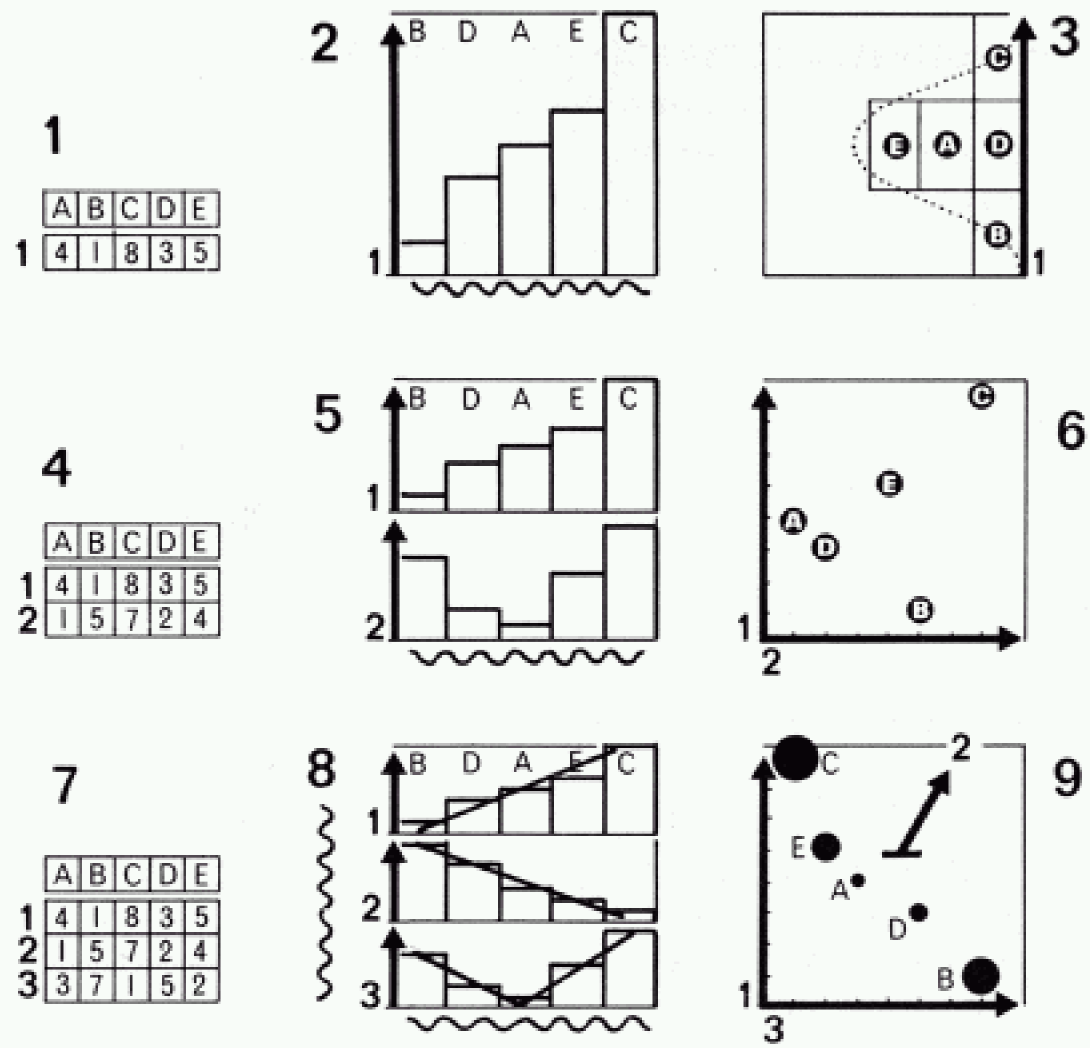
Bertin: Semiology of Graphics

- geometric primitives: marks
  - points, lines, areas, volumes
- 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



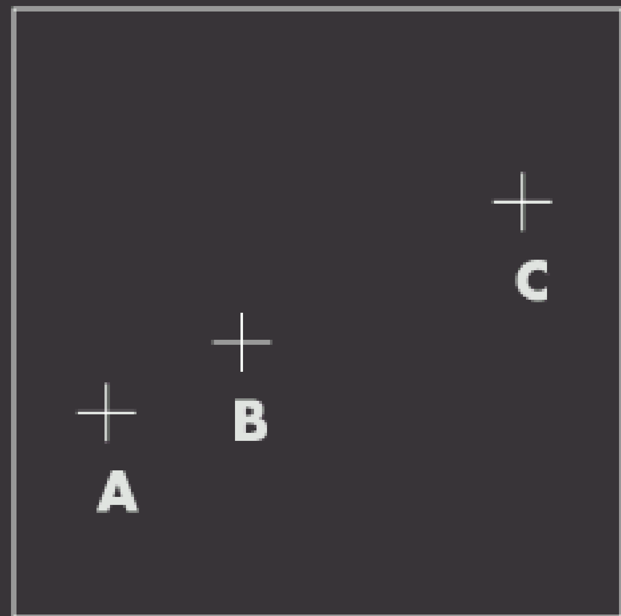
[Bertin, Semiology of Graphics, 1967 Gauthier-Villars, 1998 EHESS]

Design Space = Visual Metaphors



[Bertin, Semiology of Graphics, 1967 Gauthier-Villars, 1998 EHESS]

Information in Position



1. A, B, C are distinguishable
2. B is between A and C.
3. BC is twice as long as AB.

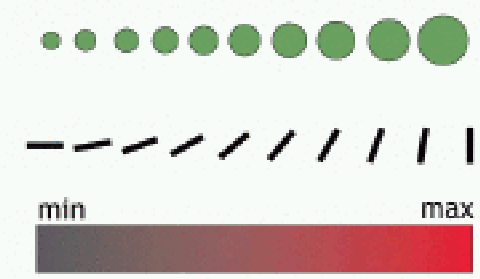
"Resemblance, order and proportional are the three signfields in graphics." - Bertin

[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding/walk025.html]

Data Types

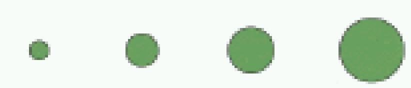
continuous (quantitative)

- 10 inches, 17 inches, 23 inches



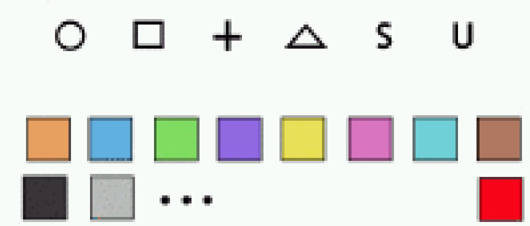
ordered (ordinal)

- small, medium, large
- days: Sun, Mon, Tue, Wed, ...



categorical (nominal)

- apples, oranges, bananas



[graphics.stanford.edu/papers/polaris]

More Data Types: Stevens

subdivide quantitative further:

interval: 0 location arbitrary

- time: seconds, minutes

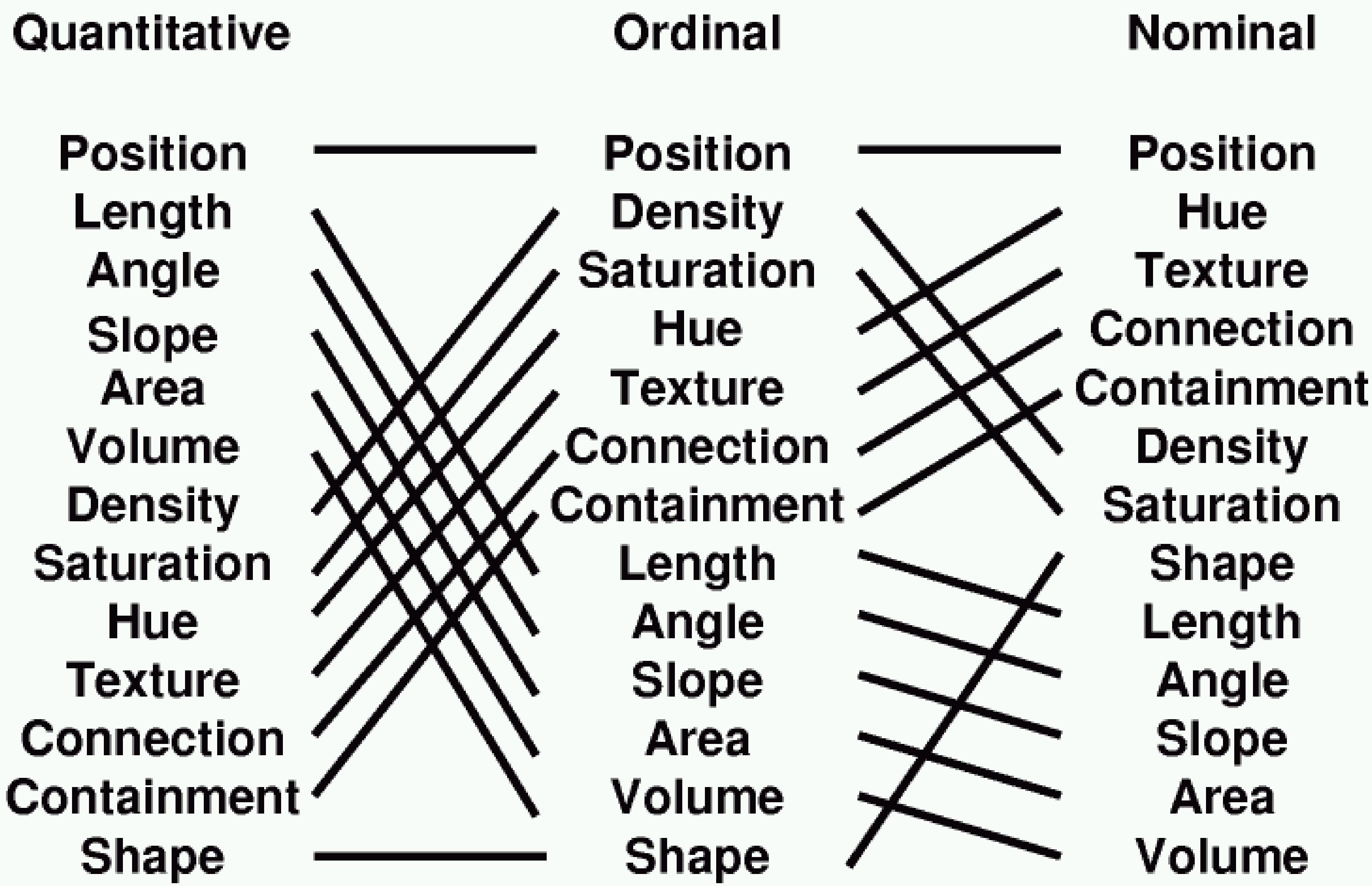
ratio: 0 fixed

- physical measurements: Kelvin temp

[S.S. Stevens, On the theory of scales of measurements, Science 103(2684):677-680, 1946]

Channel ranking varies by data type

spatial position best for all types



[Mackinlay, Automating the Design of Graphical Presentations of Relational Information, ACM TOG 5:2, 1986]

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 (Hanrahan)

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

[Hanrahan, graphics.stanford.edu/courses/
cs448b-04-winter/lectures/encoding/walk005.html]

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 anamolies in local weather patterns

Time

2D+T 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 timesteps than across spatial position

Polaris

infovis spreadsheet

table cell

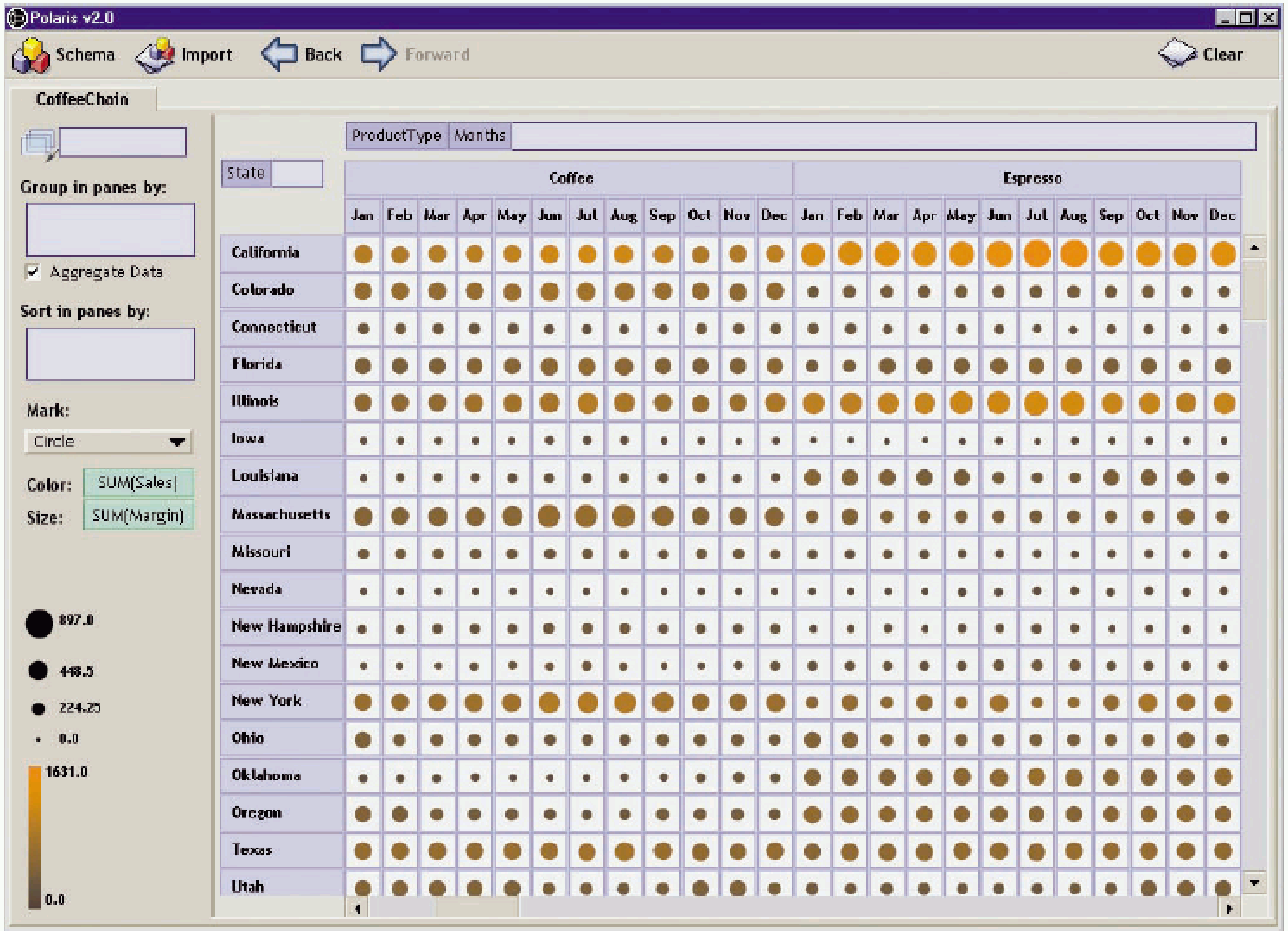
- not just numbers: graphical elements
- wide range of retinal variables and marks

table algebra <--> interactive interface

- formal language

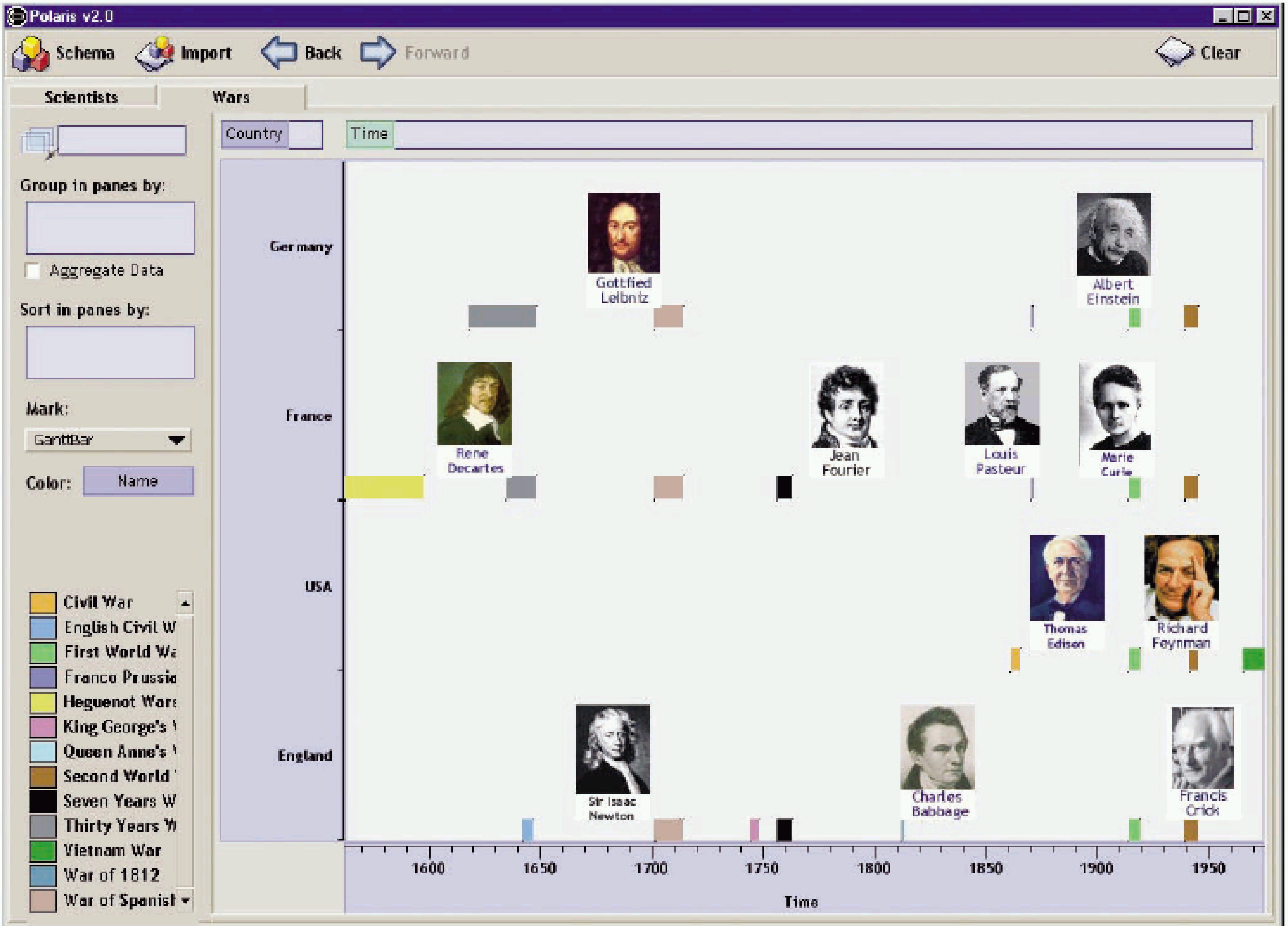
- extends Wilkinson
  Grammar of Graphics, Springer-Verlag 1999

Polaris: Circles, State/Product:Month



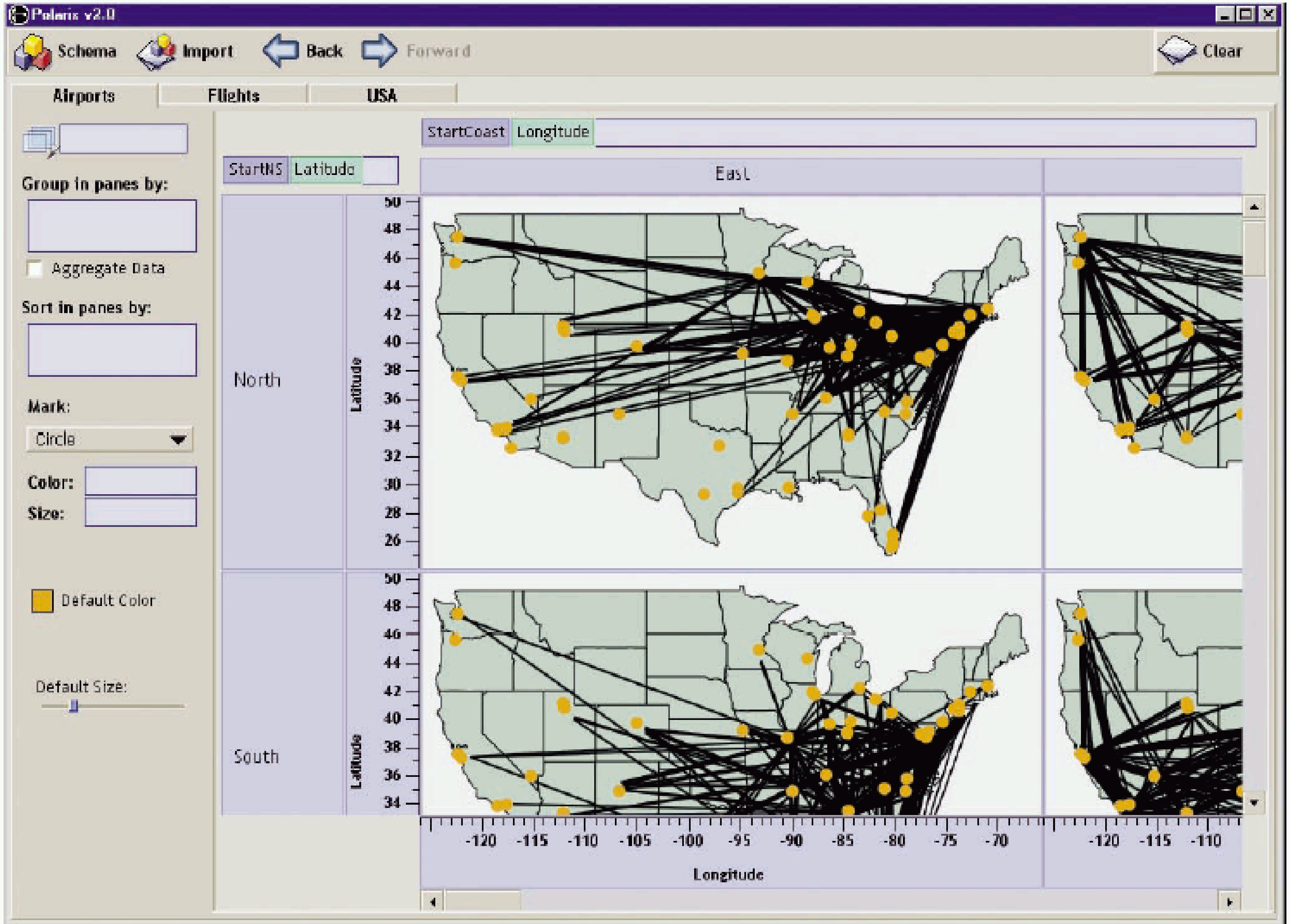
[Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases.
Chris Stolte, Diane Tang and Pat Hanrahan, IEEE TVCG, 8(1) Jan 2002. graphics.stanford.edu/papers/polaris]

Polaris: Gantt Bar, Country/Time



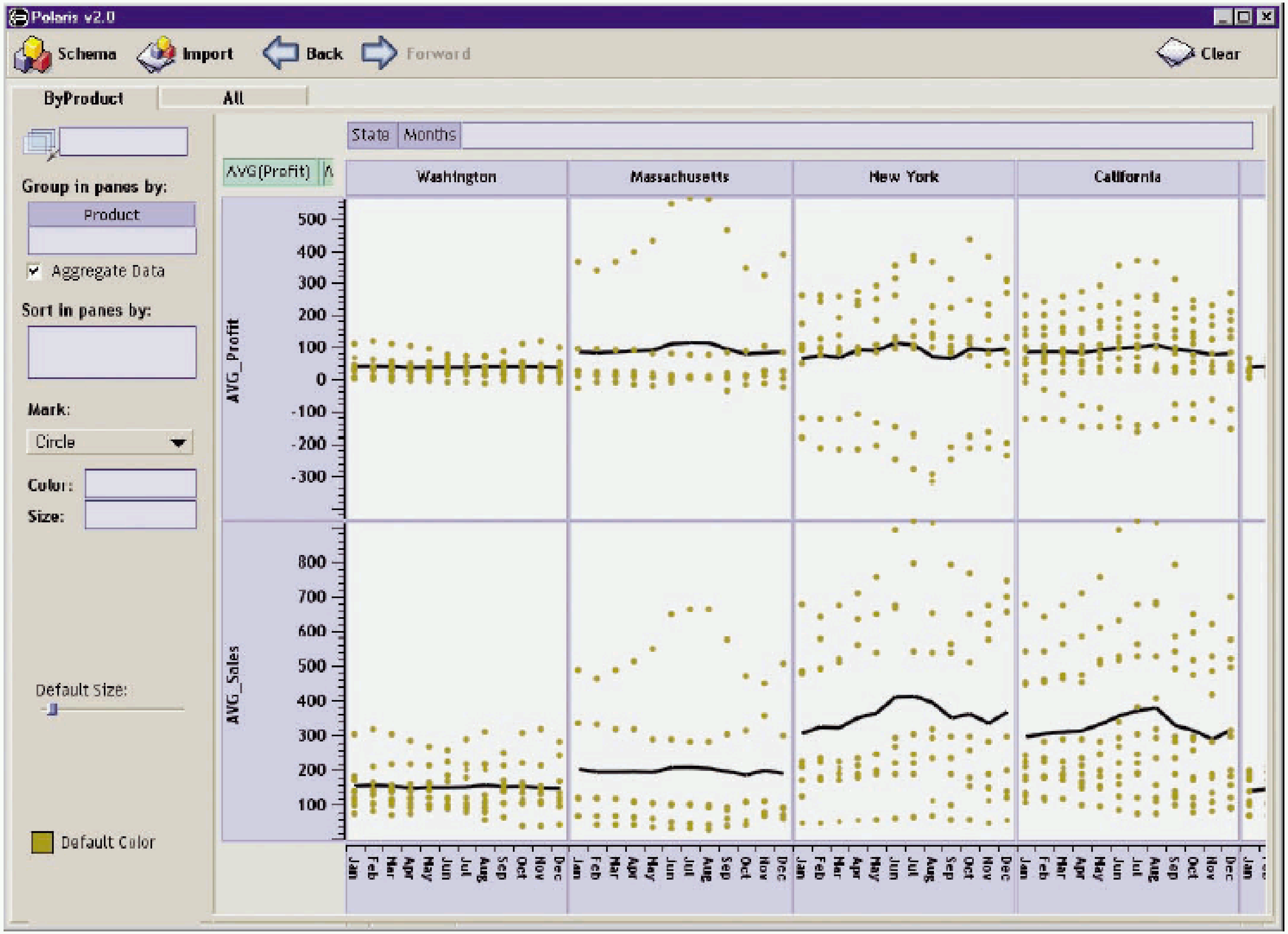
[Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases.
Chris Stolte, Diane Tang and Pat Hanrahan, IEEE TVCG, 8(1) Jan 2002. graphics.stanford.edu/papers/polaris]

Polaris: Circles, Lat/Long



[Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases.
Chris Stolte, Diane Tang and Pat Hanrahan, IEEE TVCG, 8(1) Jan 2002. graphics.stanford.edu/papers/polaris]

Polaris: Circles, Profit/State:Months



[Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases. Chris Stolte, Diane Tang and Pat Hanrahan, IEEE TVCG, 8(1) Jan 2002. graphics.stanford.edu/papers/polaris]

Fields Create Tables and Graphs

Ordinal fields: interpret field as a sequence that partitions table into rows and columns:

Quarter = {(Qtr1),(Qtr2),(Qtr3),(Qtr4)} →

Qtr1	Qtr2	Qtr3	Qtr4
95892	101760	105282	98225

Quantitative fields: treat field as single element sequence and encode as an axis:

Profit = {(Profit)} →

[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding/walk032.html]

Combinatorics of Encodings

Challenge:

Pick the best encoding from the exponential number of possibilities $(n+1)^8$

Principle of Consistency: The properties of the image should match the properties of the data.

Principle of Importance Ordering: Encode the most important information in the most effective way.

[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding/walk035.html]

Mackinlay's Expressiveness Criteria

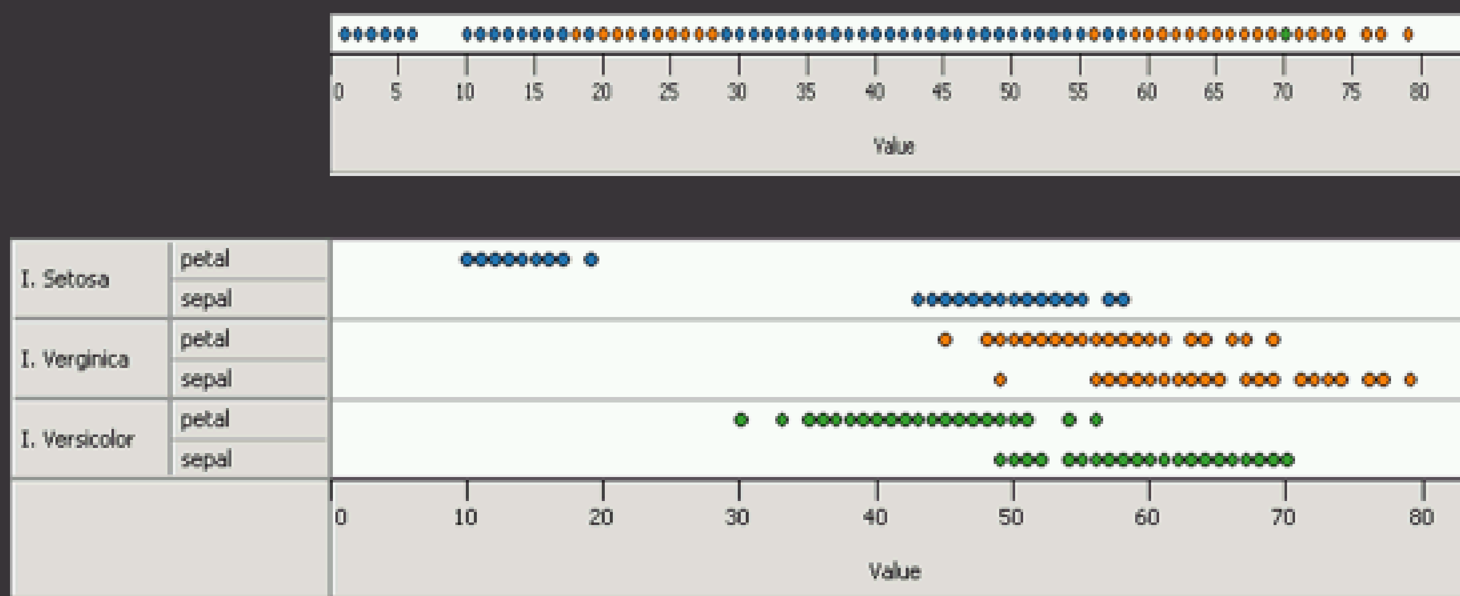
Expressiveness

A set of facts is expressible in a visual language if the sentences (i.e. the visualizations) in the language express all the facts in the set of data, and only the facts in the data.

[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding/walk036.html]

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



[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding/walk037.html]

Expresses Facts Not in the Data

A length is interpreted as a quantitative value;

∴ Length says something untrue about N data

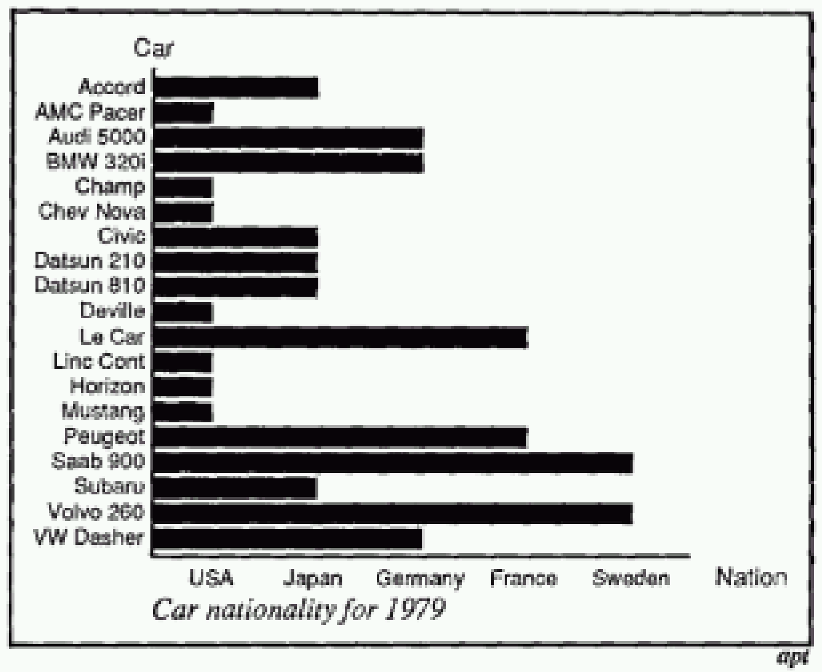


Fig. 11. Incorrect use of a bar chart for the *Nation* relation. The lengths of the bars suggest an ordering on the vertical axis, as if the USA cars were longer or better than the other cars, which is not true for the *Nation* relation.

[Mackinlay, APT, 1986]

[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding/walk038.html]

Mackinlay's Criteria 2

Effectiveness

A visualization is more effective than another visualization if the information conveyed by one visualization is more readily perceived than the information in the other visualization.

The subject of the next lecture.

[Hanrahan, graphics.stanford.edu/courses/
cs448b-04-winter/lectures/encoding/walk039.html]

Summary

Formal approach to picture specification

- Declare the picture you want to see
- Compile query, analysis, and rendering commands needed to make the picture
- 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 in this area ...

[Hanrahan, graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding/walk040.html]

Automatic Design

Mackinlay, APT
Roth et al, Sage/Visage

select visualization automatically given data
vs. Polaris: user drag and drop exporation

limited set of data, encodings

- scatterplots, bar charts, etc

holy grail

- entire parameter space

Value of Vis

I(t) = V(D,S,t)

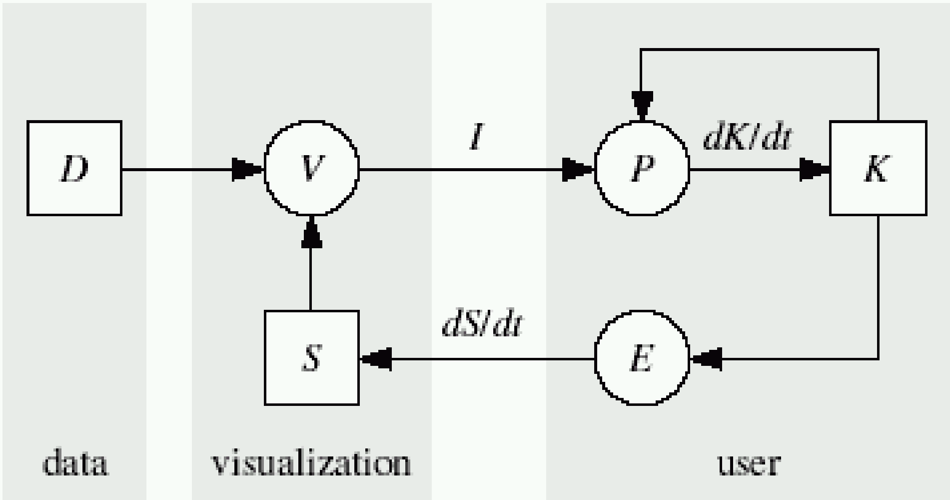
- Data D transformed by spec S into time-varying image

dK/dt = P(I,K)

- Perception P of image by user increases knowledge K

S(t) = S_0 + integral E(K)

- Interative exploration E changes spec



[The Value of Visualization. Jarke van Wijk. Proc. Visualization 2005, www.win.tue.nl/~vanwijk/vov.pdf]

Cost model

costs

- $C_i S_0$: initial development costs
- $C_u S_0$: initial per-user costs
- $C_s S_0$: initial per-session costs
- C_e : perception and exploration costs

benefit

- $G = nmW(deltaK)$

profit

- $F = G-C$
- $F = nm(W(deltaK) - C_s - kC_e) - C_i - nC_u$

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