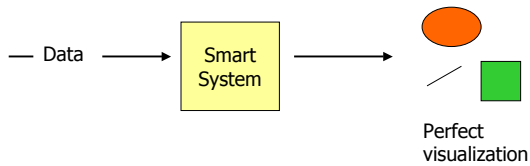


Automating Visualization Design



CS 7450 - Information Visualization
April 17, 2003
John Stasko

Holy Grail



Question



- Could we design a system that would take data tables (and descriptions) as input, and would produce high-quality information visualizations of that data?

- ---> Automating Design

A Step Easier



- If we can't do this automatically, can we build a system that would assist people with the visualization design process?
- Provide design tips, comments, help, ...

What Would It Entail?



- Would require a thorough understanding of data set
 - Level 1: Nominal, ordinal, quantitative, fields, etc.
 - Level 2: Deeper understanding of semantics of data, variables, etc.

What Would It Entail?



- Would require an understanding of capabilities of display devices
 - Color, size, etc.
 - UI & graphics toolkit, multi-windows, etc.
 - What kinds of visualizations are in library and available for presentation

What Would It Entail?



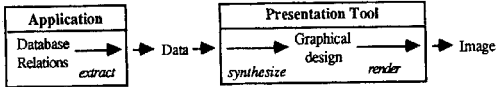
- Would need knowledge of how to map data forms to most appropriate presentation types

APT System

Mackinlay
ToG '86



- "A Presentation Tool"
- Seminal work in this area
- Goal is to create automatic visualization from data based on generating and testing possible solutions that satisfy rules of expressiveness and effectiveness



Objective



- Graphic languages are produced for each possible image, and then evaluated for:
 - Expressiveness
 - Effectiveness

Expressiveness



- Does the language express exactly the input info...and no extra ?

Effectiveness



- Which language produces the "best" output?
 - "Best" being defined as output that can be accurately interpreted
 - Therefore knowledge of human perceptual capabilities used to determine how "accurately" a visualization could be interpreted

Example



Price(Accord, 5799)	Price: Cars -> [3500, 13000]
Mileage(Accord, 25)	Mileage: Cars -> [10, 40]
Weight(Accord, 2240)	Weight: Cars -> [1500, 5000]
Repair(Accord, Great)	Nation: Cars-> {USA, Japan,...}
Nation(Accord, Japan)	Repair: Cars-> <Great,Good,OK,Bad>
...	
Price(Pacer, 4749)	
Mileage(Pacer, 17)	
...	

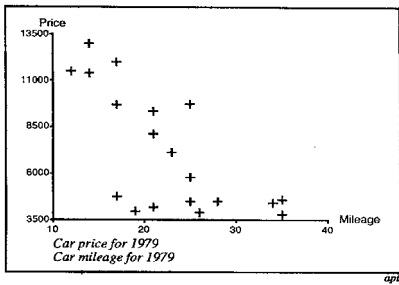
Data Set

Design a Graphic

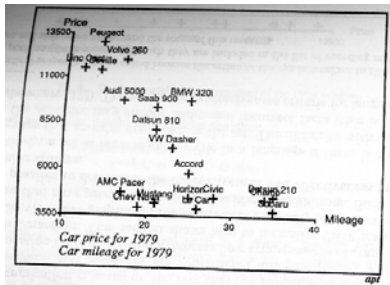


- Next four slides so possible designs
- Critique each one -- What makes it good or bad?

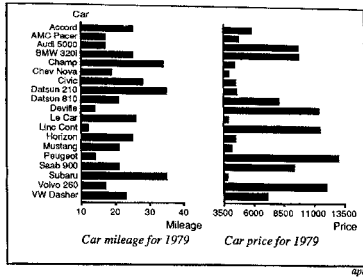
Scatter Plot



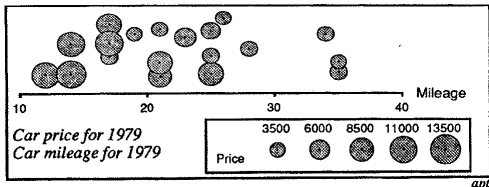
Labeled Scatter Plot



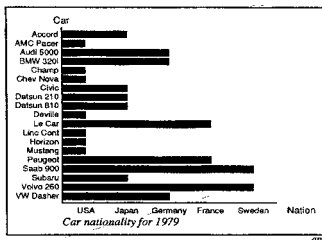
Aligned Bar Chart



Area Plot



Another Example



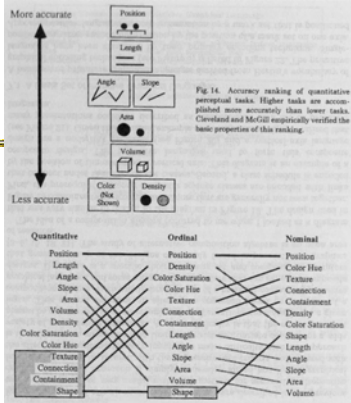
Inappropriate use of bar chart for Nation relation

Primitives



- Need to come up with a set of atoms (and their properties) for design that the tool will use to create visualizations
- Need to understand what the atoms and their properties best facilitate

Accuracy ranking of quantitative perceptual tasks



Ranking of perceptual tasks

The Graphic Language



Encoding Technique	Primitive Graphical Language
Single-position	Horizontal axis, vertical axis
Apposed-position	Line chart, bar chart, plot chart
Retinal-list	Color, shape, size, saturation, texture, orientation
Map	Road map, topographic map
Connection	Tree, acyclic graph, network
Misc. (angle, contain, ...)	Pie chart, Venn diagram, ...

	Nominal	Ordinal	Quantitative
Size	-	•	•
Saturation	-	•	•
Texture	•	•	
Color	•	•	
Orientation	•		
Shape	•		

Developing New Diagrams



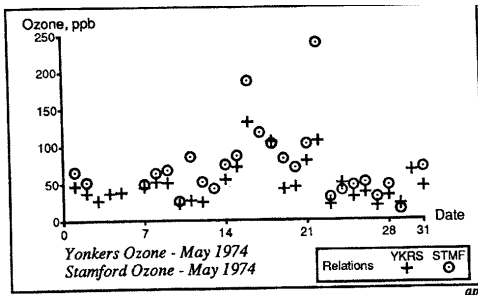
- To create new diagrams, we *compose* existing primitive designs
- Developed a composition algebra to do this
- Principle of Composition
 - Two visualizations can be composed by merging the parts that encode the same information

Composition Types

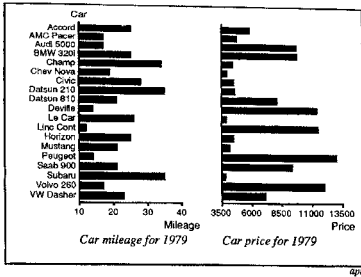


- Double axes
 - Compose graphical sentences with identical horizontal and vertical axes
- Single axis
 - Aligns two sentences that have same vertical or horizontal axis
- Mark
 - Utilize different attributes of a mark

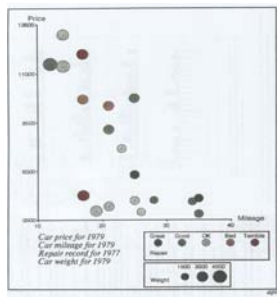
Double Axes Composition



Single Axis Composition



Mark Composition



Synthesis Algorithm



- At core of APT
- Consists of design and rendering components

Design Synthesis Algorithm



- Uses logic programming, backtracking, etc.
- Divide and Conquer
- Three pieces
 - Partition
 - Selection
 - Composition

Components



- Partition
 - Partition set of relations to be presented down into subsets to get matches
- Selection
 - For each partition, use expressiveness criteria to get possible designs, then effectiveness criteria to pick best one
- Composition
 - Compose individual designs into unified presentation of all information

Not So Automatic...



- Much work in automating design (eg, APT) didn't focus on creating a collaborative design process between human user and computer system
 - Goal: Support users, don't replace them

SAGE



- System to aid users in information visualization design
- From CMU & Maya Design
- Ground-breaking work in area

Roth et al
CHI '94

User-Directed Design



- Design is
 - Constructive process of selecting and arranging graphical elements
 - Process of browsing and customizing previous cases

SAGE Components



- SageBrush
 - Direct manipulation editor/design tool
- SageBook
 - Tool for browsing and retrieving previously created graphics to use as a starting point
- Sage is the surrounding and encompassing framework

Architecture

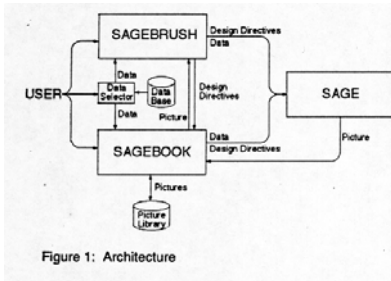


Figure 1: Architecture

Design Directives

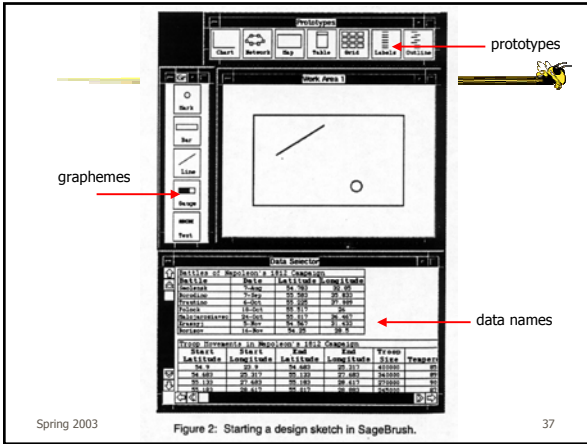


- Graphic primitive choices (color and size, etc, of lines, circles...)
- Encoding mechanisms (chart, table, map, ...)
- Layout constraints (alignment, ordering)
- Group constraints (use clusters of elements as a single element)
- Mapping data->graphics

SageBrush



- Smart graphical editing tool for working with primitives of design
- Like a souped up MacDraw
- User drags and drops prototypes and graphemes (see next slides) into drawing area to design graphic



Prototypes

- Partial designs, typically have some graphemes on chart-like framework
- Users can extend prototypes by dragging in new graphemes

Spring 2003

CS 7450

38

Graphemes

Double-clicking on a grapheme shows the properties that can be set

Figure 3: Property selection and data mapping in SageBrush's work area.

Spring 2003

CS 7450

39

Specification



- User specifies that position and color of a line should be used to set up a mapping
- But still need to say what data to use to drive mapping
- User drags label from lower data selector area and drops onto grapheme property

Blank Paper Problem



- Many times it's difficult to know where to start in design
- Helpful to have ideas or hints to get the mental juices flowing
- Want to have system that makes it possible to create new pictures that are analogous to ones already existing in library

SageBook



- Provides library of cases of existing data sets and visualizations to use as prototypes or design inspirations
- Like thumbing through series of photos
- Allows user to search based on either visual attributes or data attributes

Video

Final Project



- Questions?
- Contest tips

Upcoming



- Review
- Summary Exam
- Project
 - Reports due Tuesday 29th
 - Presentations Friday 1st

References



- All referred to papers
- Coleman and Gandy F '99 slides
