3 Information Visualization

3.1 Motivation and Examples

## 3.2 Basics of Human Perception

3.3 Principles and Concepts

3.4 Standard Techniques for Visualization

3.5 Further Examples

Literature:

- E. Tufte: The Visual Display of Quantitative Information,2nd ed., B&T 2001
- Marti Hearst
  - http://bailando.sims.berkeley.edu/talks/chi03-tutorial.ppt
- Margret-Anne Storey
  - http://www.cs.uvic.ca/~mstorey/teaching/infovis/course_notes/introduction.pdf

---

Visualization Reference Model
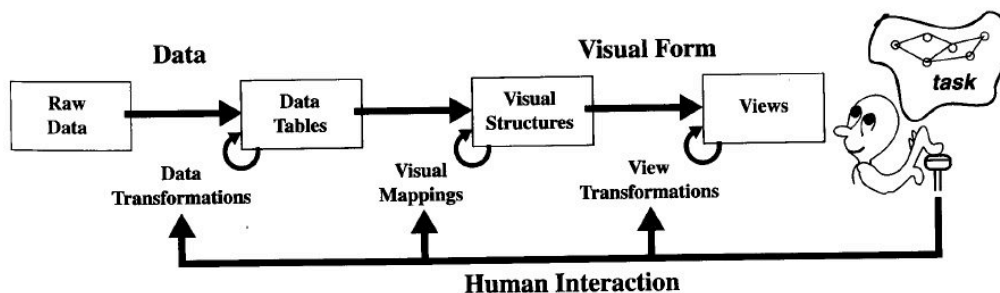
Raw Data: idiosyncratic formats
Data Tables: relations (cases by variables) + metadata
Visual Structures: spatial substrates + marks + graphical properties
Views: graphical parameters (position, scaling, clipping, . . .)

- Raw Data → Data Table
    filtering
- Data Table → Visual Structure
    pick mappings

- Visual Structure → Views
    probes, viewpoints, distortions

(Storey, 2004)

Types of Data

- Entities
  - Objects of interest
- Relationships
  - Form structures that relate entities
  - Many kinds of relationships exist
    - » Is-part-of, is-kind-of, is-xyx-to, …
- Attributes of entities or relationships
  - Attribute vs. Independent information (entity)
  - Attribute is variable of a certain value type
- Operations
  - Actions can also be considered as data
- Metadata
  - Data about data

Basic Attribute Value Types

- Nominal (qualitative)
  - No inherent order (but can be tested for equality =)
  - Examples: City names, types of diseases, kind of fruit, ...
- Ordinal (qualitative)
  - Ordered (can be tested for <, >), but not at measurable intervals
  - Sequencing things, ranking
  - Examples: first, second, third, …; cold, warm, hot
- Nominal/Interval (quantitative)
  - Integer or real numbers
  - Ordered (can be tested for <, >)
  - Arithmetical operations, ratios are possible
  - Interval data: Derivation of gaps (e.g. time between departure and arrival)
  - Examples: Size and population of countries, schedule times, numeric grades

Hearst, 2003

Attribute Dimensions

- All kinds of *tensors* may appear as attribute values
- Tensor rank 0: Scalar
  - E.g. mass, temperature, length, price
- Tensor rank 1: Vector
  - E.g. force, momentum, location, direction
- Tensor rank 2: Matrix
 - E.g. linear transformation
- …

Mapping to Visual Structures

- Mapping from data tables to visual structures is
  - *expressive*
    if all data in the table (and only this information) are presented in the structure
  - *efficient*
    if the visual representation is easier to interpret for humans, can convey more distinctions or leads to fewer errors

(Storey, 2004)

Visual Structure

- Spatial substrate
  - Fixed number of dimensions
  - Inherently perceptual
- Axes
  - Unstructured axis
  - Nominal axis (division into subregions)
  - Ordinal axis (order has meaning)
  - Quantitative axis (metric associated with region)
- Graphical marks
  - Visible things that occur in space
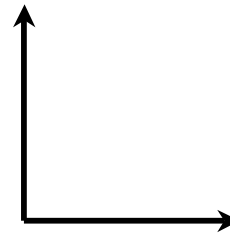
(based on Storey, 2004)

Graphical Marks

- Four elementary types:
  - Points (0D)
  - Lines (1D)
  - Areas (2D)
  - Volumes (3D)
- In practice, marks need more dimensions than in theory
  - E.g. Points can be seen only if painted as areas
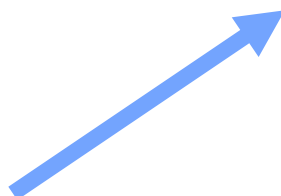
(based on Storey, 2004)

Mapping Examples

(assume 2-dimensional representations)
- Two scalars:
  - Price vs. top speed of cars
- Ordinal and scalar:
  - Max. price vs. brand of cars
- Ordinal and vector:
  - Price range vs. brand of cars
- Vector and scalar:
  - Location vs. average temperature
- Vector and vector:
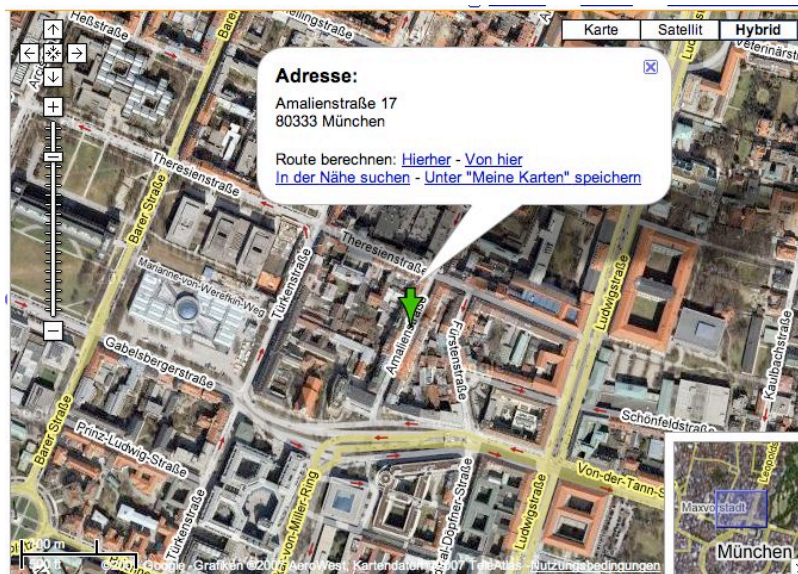 - Location vs. temperature range
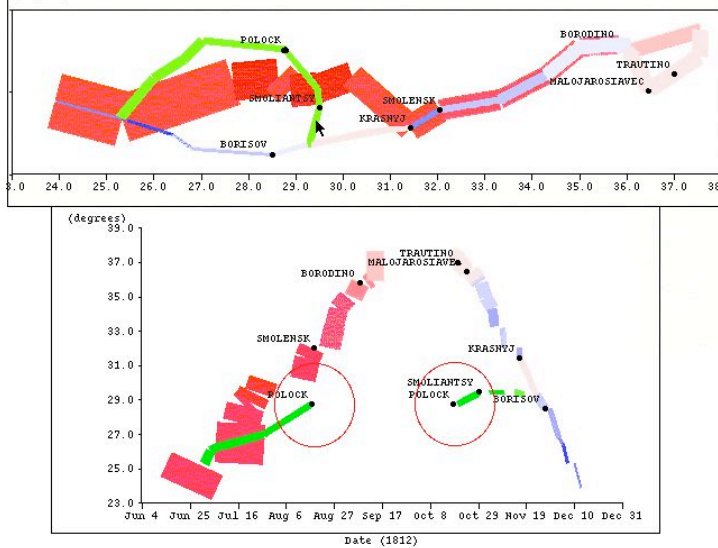

Example: FilmFinder
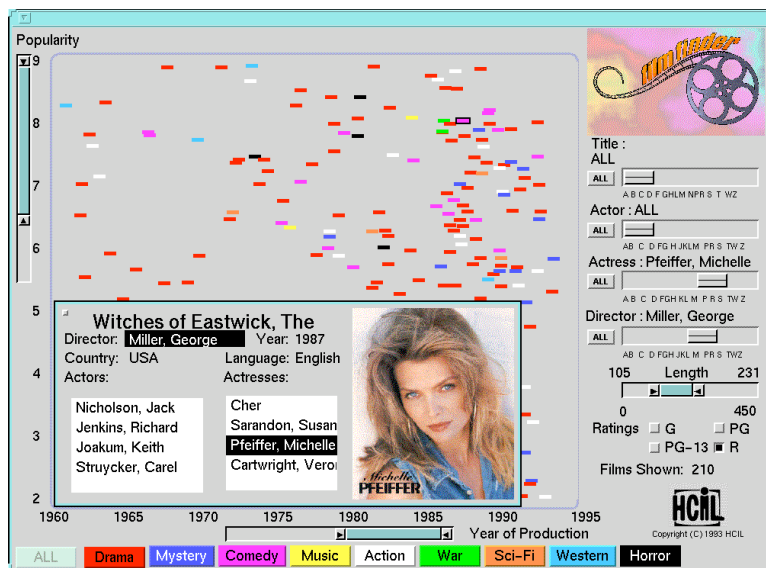
University of Maryland, HCIL

View Transformations

- Ability to interactively modify and augment visual Structures
  - Turning static presentations into visualizations
- Time is exploited to display more information
  - Dynamic) Visualizations exist in space time
- Three common view transformations:
  1. Location probes: use location to reveal additional info
  2. Viewpoint controls: zoom, pan, clip the viewpoint
  3. Distortion: focus + context view

Example: View Transformations in Google Maps

Example: Interactive Graphs

http://www.cs.cmu.edu/~sage

Accuracy Ranking of Perceptual Tasks

Quantitative	Ordinal	Nominal
Position	Position	Position
Length	Density	Color Hue
Angle	Color Saturation	Texture
Slope	Color Hue	Connection
Area	Texture	Containment
Volume	Connection	Density
Density	Containment	Color Saturation
Color Saturation	Length	Shape
Color Hue	Angle	Length
Texture	Slope	Angle
Connection	Area	Slope
Containment	Volume	Area
Shape	Shape	Volume

Fig. 15.   Ranking of perceptual tasks. The tasks shown in the gray boxes are not relevant to these types of data.

Mackinlay 88

Interpretations of Visual Properties

- Some properties have intrinsic meaning
  (Senay & Ingatious 97, Kosslyn, others)
  - Density (Greyscale)
    - Darker -> More
  - Size / Length / Area
    - Larger -> More
  - Position
    - Leftmost -> first, Topmost -> first
- Some properties do not have intrinsic meaning, even some perceived quite accurately
  - Hue
    - ??? no intrinsic meaning
  - Slope
    - ??? no intrinsic meaning

Hearst, 2003

Color Schemes (1)

Order these (low->hi)

Hearst, 2003

Color Schemes (2)

Gray scale | Full spectral scale | Single sequence part spectral scale | Single sequence single hue scale | Double-ended multiple hue scale
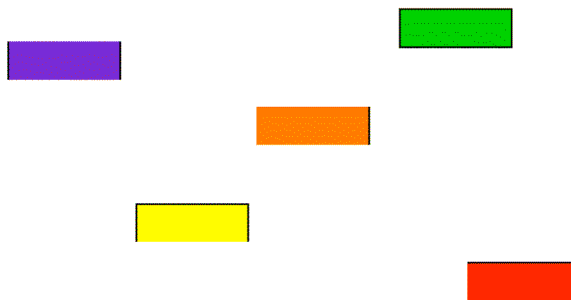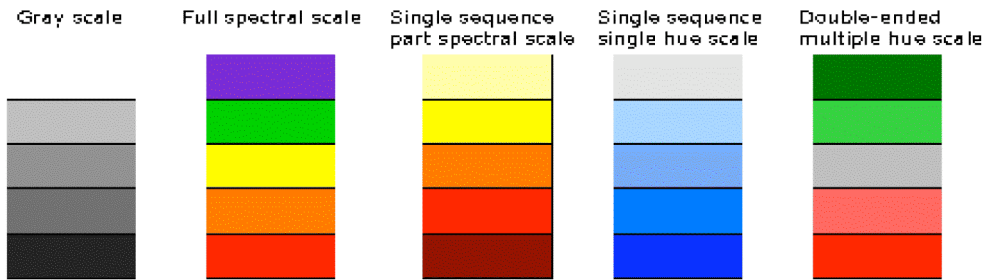
Hearst, 2003

Using Color

- Call attention to specific items
- Distinguish between classes of items
  - Increases the number of dimensions for encoding
- Increase the appeal of the visualization
- Proceed with caution
  - Less is more
  - Representing magnitude is tricky
- Examples
  - Red-orange-yellow-white
    » Works for cost
  - Green-light green-light brown-dark brown-grey-white works for atlases
  - Grayscale is unambiguous but has limited range

Hearst, 2003

Continuity

- Experience tells that visual elements are more likely to be continuous
- Implied connection
- connections are used to show relations
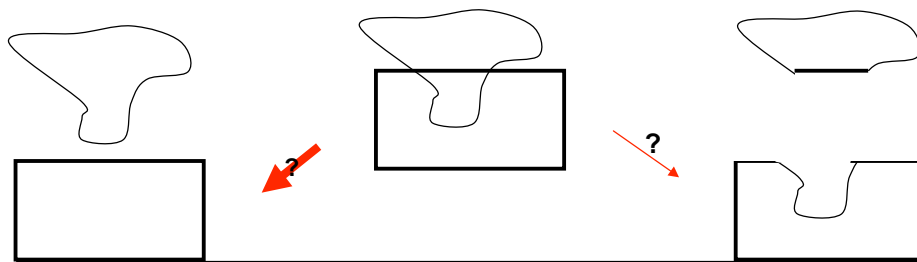
Symmetry

- Symmetrical to emphasize relationship
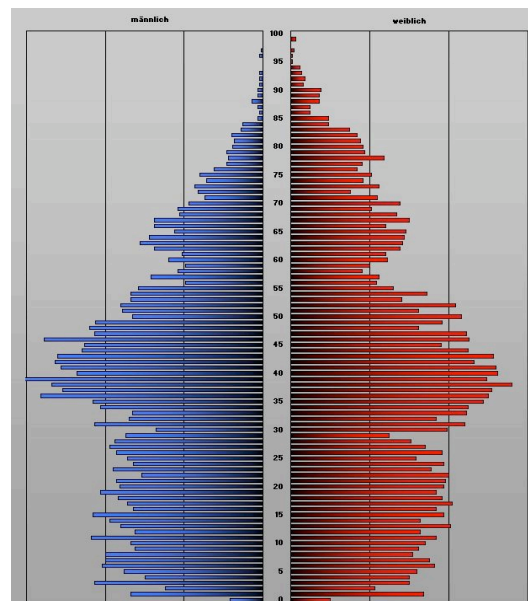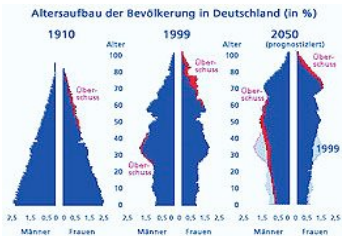
Figure, Background, Transparency, Overlap

- What is foreground and what is background?
- Transparency is perceived only when good continuity and color correspondence exists.
- Visual interference in overlapping textures
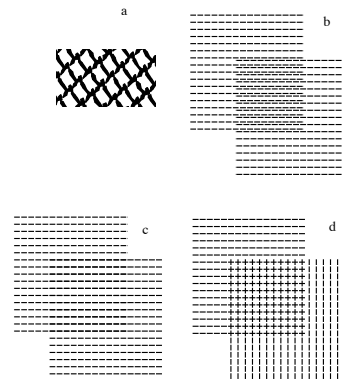
Principles of Graphical Excellence (E. Tufte)

- Graphical excellence
  - The well-designed presentation of interesting data – a matter of substance, of statistics, and of design
    » consists of complex ideas communicated with clarity, precision and efficiency
    » is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space
    » **requires telling the truth about the data**.

Hearst, 2003

Tufte Principle

Data ink = ink used for representing data

Chart ink = ink used for extra elements different from data

Avoid "chart junk"! Maximize the data-ink ratio:

$$\text{Data-ink ratio} = \frac{\text{data ink}}{\text{total ink used in graphic}}$$

Hearst, 2003

Ludwig-Maximilians-Universität München   **Prof. Hußmann**   **Mensch-Maschine-Interaktion II – 3 - 72**

Tufte's Graphical Integrity

- Some lapses intentional, some not

$$\text{Lie Factor} = \frac{\text{size of effect in graph}}{\text{size of effect in data}}$$

- Misleading uses of area
- Misleading uses of perspective
- Leaving out important context
- Lack of taste and aesthetics

Hearst, 2003

Ludwig-Maximilians-Universität München   **Prof. Hußmann**   **Mensch-Maschine-Interaktion II – 3 - 73**

Lie factor

$$\text{lie factor} = \frac{\text{size of effect shown in graph}}{\text{size of effect in data}}$$

where

$$\text{size of effect} = \frac{|\text{second value - first value}|}{\text{first value}}$$

A lie factor that is either much higher or much lower than one is bad. A **high** lie factor **exaggerates** differences between values. A **low** lie factor **obscures** differences between values.

A common example of a **high** lie factor occurs when both dimensions of a two-dimensional figure are made proportional to the same data, so that the size of the figure is proportional to the square of the data; for instance,

Year	Books circulated
2001	100
2002	141
2003	200

where the lie factor is about 2.4.

An example of a **low** lie factor can be seen in the "Cones" custom chart format in Microsoft Excel.

44

23   24   25

2000 2001 2002 2003
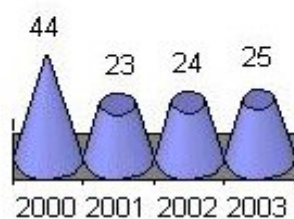
The heights of the (truncated) cones are proportional to the data, but their areas on the screen and their apparent volumes make the larger data values seem relatively small.

Charting on a **logarithmic** scale can also produce a low lie factor.

http://instruct.uwo.ca/fim-lis/504/504gra.htm

Ludwig-Maximilians-Universität München   Prof. Hußmann   Mensch-Maschine-Interaktion II − 3 - 75

How to Exaggerate with Graphs
from Tufte '83

THE SHRINKING FAMILY DOCTOR
In California

Percentage of Doctors Devoted Solely to Family Practice

1964	1975	1990
27%	16.0%	12.0%

1: 4,232
6,212

1: 3,167
6,694

1: 2,247 RATIO TO POPULATION
8,023 Doctors

Los Angeles Times, August 5, 1979, p. 3.
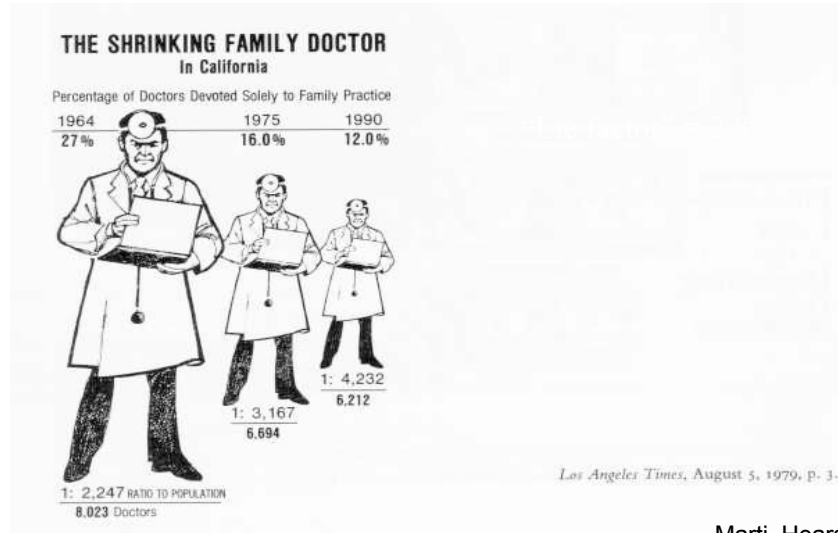
Marti Hearst

How to Exaggerate with Graphs
from Tufte '83

Error:
Shrinking
along both
dimensions

Washington Post, October 25, 1978, p. 1.

THE UNITED STATES OF AMERICA
ONE DOLLAR
1958 — EISENHOWER: $1.00

THE UNITED STATES OF AMERICA
ONE DOLLAR
1963 — KENNEDY: 94¢

THE UNITED STATES OF AMERICA
ONE DOLLAR
1968 — JOHNSON: 83¢

Purchasing
Power
of the
Diminishing
Dollar
Source: Labor Department

THE UNITED STATES OF AMERICA
ONE DOLLAR
1973 — NIXON: 64¢

THE UNITED STATES OF AMERICA
1978 — CARTER: 44¢
(August)

Marti Hearst

3 Information Visualization

3.1 Motivation and Examples

3.2 Basics of Human Perception

3.3 Principles and Concepts

3.4 Standard Techniques for Visualization

3.5 Further Examples

Literature:

- Marti Hearst
  - http://bailando.sims.berkeley.edu/talks/chi03-tutorial.ppt
- Margret-Anne Storey
  - http://www.cs.uvic.ca/~mstorey/teaching/infovis/course_notes/introduction.pdf

---

Basic Types of Symbolic Displays
(Kosslyn 89)

- Graphs

- Charts
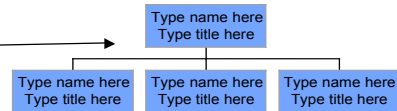
- Maps
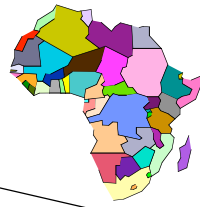
- Diagrams
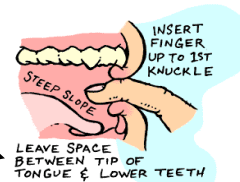
From Hearst, 2003

Graphs

- At least two scales required
- values associated by a symmetric "paired with" relation
  - Examples: scatter-plot, bar-chart, line graph

The Boeing Co   as of 26-Jan-1998

(C)1998 Yahoo! Inc.   Volume (1000's)   http://quote.yahoo.com/
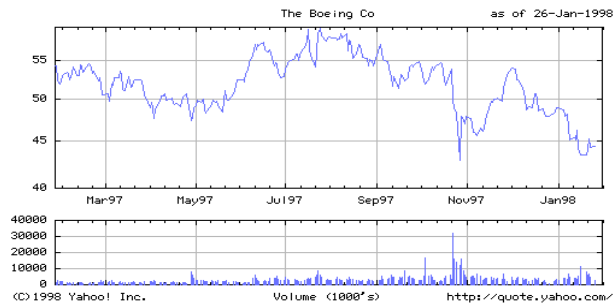
Hearst, 2003
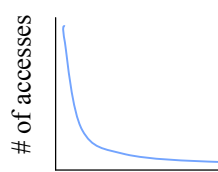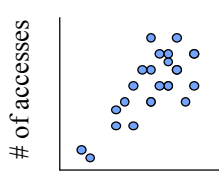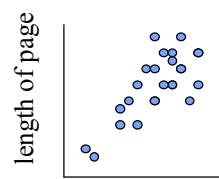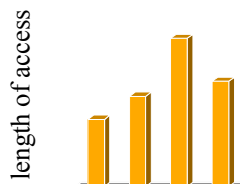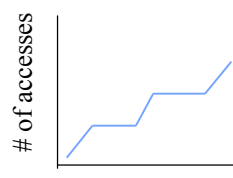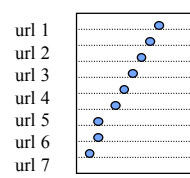
Common Graph Types

Hearst, 2003

Scatter Plots

- Qualitatively determine if variables
  - are highly correlated
    - » linear mapping between horizontal & vertical axes
  - have low correlation
    - » spherical, rectangular, or irregular distributions
  - have a nonlinear relationship
    - » a curvature in the pattern of plotted points
- Place points of interest in context
  - Color representing special entities

Scatter Plot

Salary 87

Hits used for coloring

2000

1500

1000

500

0

0   50   100   150

Career Home Runs
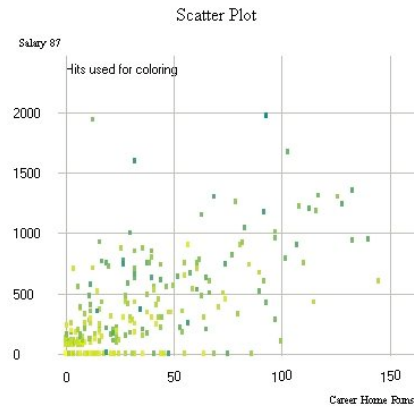
Hearst, 2003

When to use which type?

- Line graph
  - x-axis requires quantitative variable
  - Variables have contiguous values
  - Familiar/conventional ordering among ordinals
- Bar graph
  - Comparison of relative point values
- Scatter plot
  - Convey overall impression of relationship between two variables
- Pie Chart?
  - Emphasizing differences in proportion among a few numbers

Hearst, 2003

Charts

- Discrete relations among discrete entities
- Structure relates entities to one another
- Lines and relative position serve as links
- Examples: Family tree, flow chart
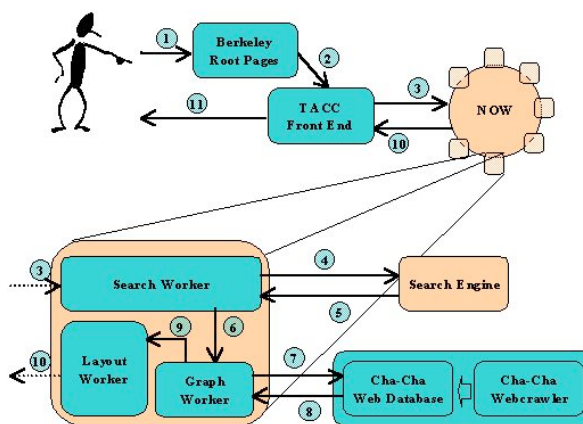
Hearst, 2003

Maps

- Internal relations determined (in part) by the spatial relations of what is pictured
- Labels paired with locations
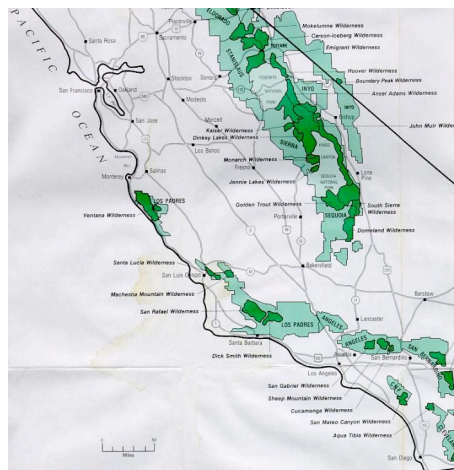
www.thehighsierra.com

Hearst, 2003

Diagrams

- Schematic pictures of objects or entities
- Parts are symbolic (unlike photographs)
  - How-to illustrations
  - Figures in a manual
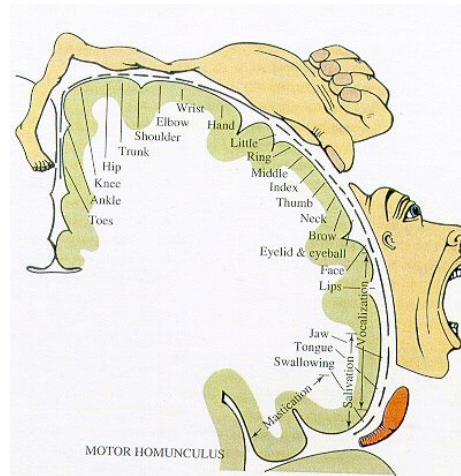
From Glietman, Henry. Psychology. W.W. Norton and Company, Inc. New York, 1995

Hearst, 2003

Alternative Tree Visualizations (1)

Organization Chart
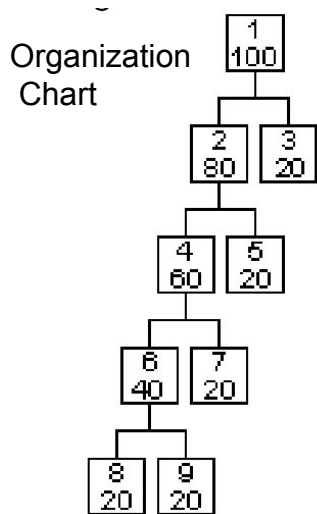
Tree Ring
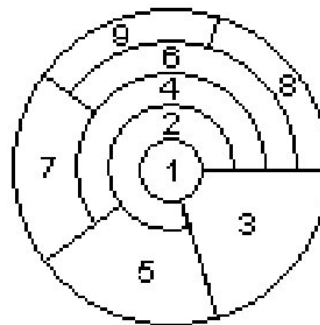
Alternative Tree Visualizations (2)

Icicle Plot
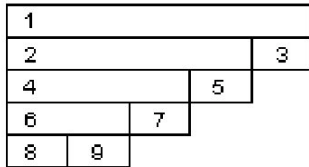
Tree Map
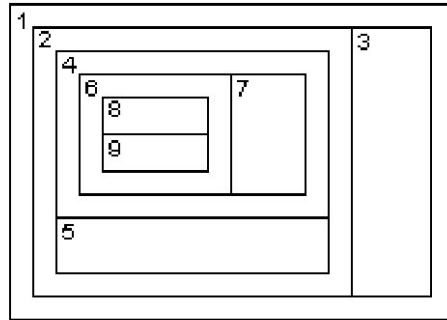
Comparing Tree Visualizations
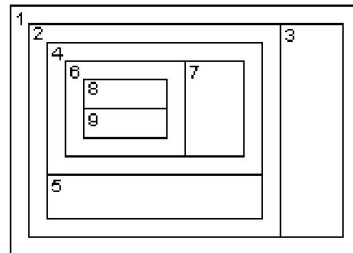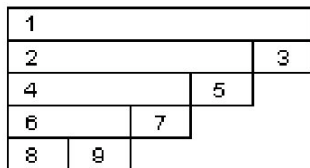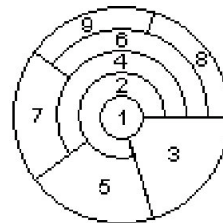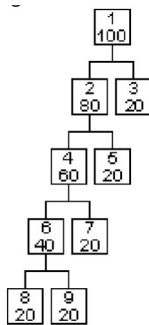
Typical Tasks for Viewing Trees

- Determine the type of tree, e.g.
  - Binary
  - N-ary
  - Balanced
  - Unbalanced
- Find relations, e.g.
  - Deepest common ancestor
- Size of the tree, e.g.
  - How many levels
  - How many leaves
- Details about leaves, e.g.
  - Largest leaf

- Different representation may be better for a given task, e.g.
  - To find out if a tree is balanced or how many levels exist, the Icicle Plot is good

For more details see:

Barlow et al. "A Comparison of 2-D Visualizations of Hierarchies" INFOVIS'01
http://www.sims.berkeley.edu/courses/is247/s02/readings/barlow.pdf

---

Information Visualization Tasks

Tasks in interactive workflow using visualized information:

- **Overview** — Gain an overview of the entire collection
- **Zoom** — Zoom in on items of interest
- **Filter** — Filter out uninteresting items
- **Details-on-demand** — Select an item or group and get details when needed
- **Relate** — View relationships among items
- **History** — Keep a history of actions to support undo, replay, and progressive refinement
- **Extract** — Allow extraction of sub-collections and of the query parameters

Shneiderman, 2003

Information Visualization Mantra

…

Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand

…

Shneiderman, 2003

Example: PhotoMesa

Visualization Techniques for View Transformations

- Focus & Context
- Zoom & Pan

Focus & Context: Background

- Useful Field of View (UFOV)
  - Expands searchlight metaphor
  - Sze of region from which we can rapidly take information
  - Maintains constant number of targets
- Tunnel Vision and Stress
  - UFOV narrows as cognitive load/stress goes up
- Role of Motion in Attracting Attention
  - UFOV larger for movement detection

Depth of Field

- Guiding user attention by blurring less relevant parts of an image
- Keeping the context

- Semantic Depth of field = blurring objects based on their relevance

Semantic Depth of Field - Example
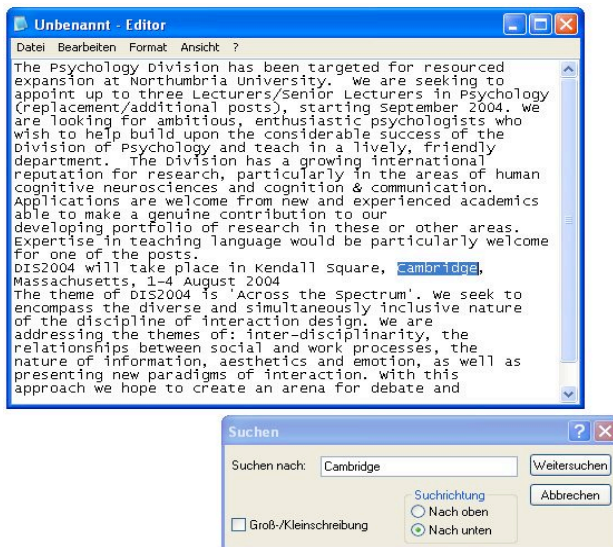
Semantic Depth of Field - Example
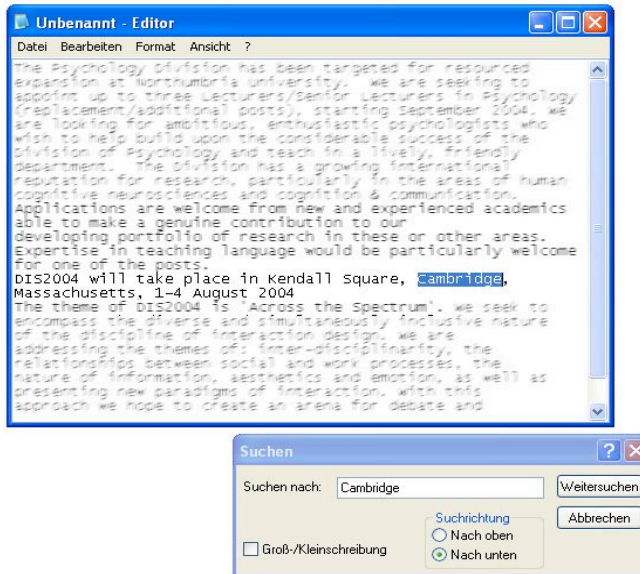
Magnifying Glass
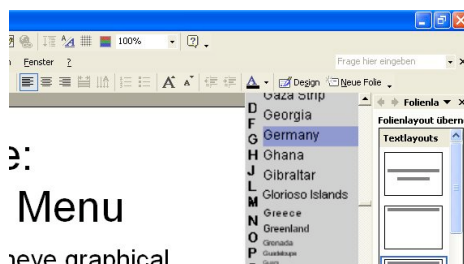
- Magnifying glass hides context!
- This is not focus+context
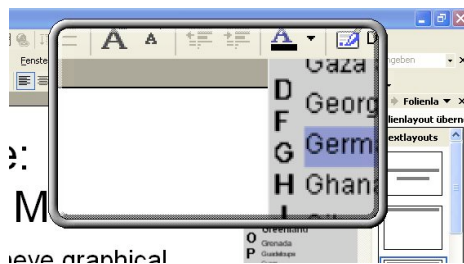
Focus + Context

- Basic Idea:
  - Show selected regions of interest in greater detail (*focus*)
  - Preserve global view at reduced detail (*context*)
  - NO occlusion - All information is visible simultaneously

- Techniques
  - Fisheye views
  - Fisheye lens
  - Continuously variable zoom
  - Nonlinear magnification
  - Hyperbolic views
  - Distortion viewing
  - Rubber sheet views

Alternate Geometry

- Euclidean geometry – we use it since primary school…
  - 3 angles of a triangle add up to?
  - Shortest distance between two points?

- Spherical geometry
  - Geographical view of the world
    - » What is the shortest way from Moscow to San Francisco?
    - » Sum of angles of a triangle between Paris, NY, and Cape Town?
  - http://math.rice.edu/~pcmi/sphere/

- Hyperbolic Geometry / Space
  - Theory of Relativity
  - The "fifth" dimension
  - Can be projected into 2-D as a *pseudosphere*
  - Key: As a point moves away from the center towards the boundary circle, its distance approaches *infinity*
  - http://cs.unm.edu/~joel/NonEuclid/NonEuclid.html (Applet)

Distorted vs. Non-distorted

- Non-distorted
  - Display only a selection at a time
  - Scrolling
  - Paging access
  - hierarchical structure
  - Structure-specific presentation
- Distorted
  - See the following slides

---

Basic Idea – Perspective Wall

From http://www.cs.ubc.ca/~tmm/courses/cpsc533c-03-spr/0324.fengdongdu.ppt

Fisheye View

Example:
Fisheye Menu

- Applies fisheye graphical visualization techniques to linear menus
- For very long menus as alternative to
  - Hierarchies
  - Scrolling
  - Arrow-bars

- Benjamin B. Bederson. Fisheye Menus. UIST'00
- Demo
  http://www.cs.umd.edu/hcil/fisheyemenu/fisheye menu-demo.shtml
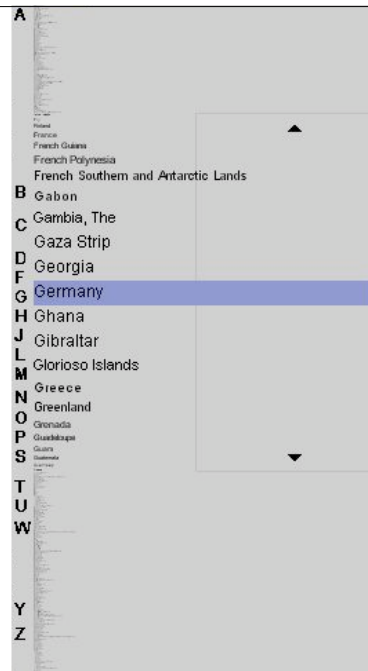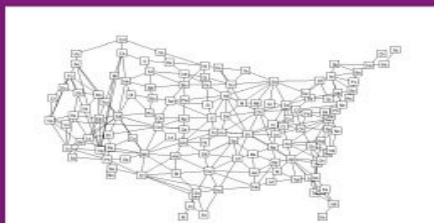
Fisheye View - Networks
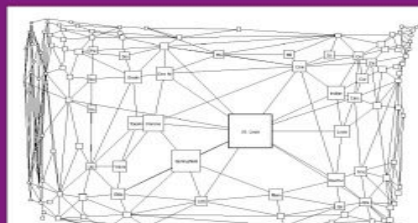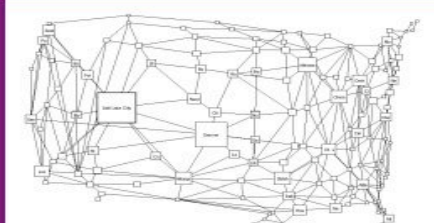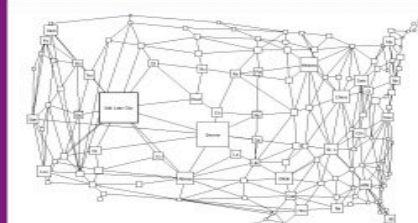
From Sarkar and Brown

Panning and Zooming

- Panning
  - Smooth movement of camera across scene (or scene moves and camera stays still)
- Zooming
  - Increasing or decreasing the magnification of the objects in a scene

- Useful for changing focal point
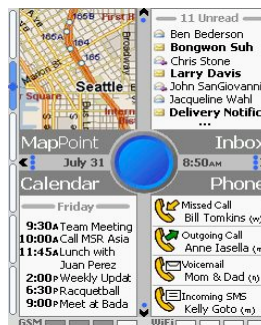- Also used in creating moving pictures from still pictures
  - "Ken Burns effect"

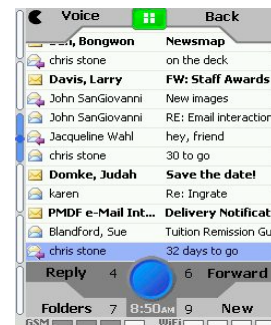
Example: LaunchTile Visual Design

- Three (pure) zoom levels

| World | Zone | Application |

Karlson, Bederson, SanGiovanni, CHI 2005