Principles Of Information Visualization

What is information visualization

Tufte's guidelines

Visual variables for representing information

The principle of small multiples for displaying information

How metaphors can be used and misused

Direct manipulation and direct engagement

Representations

Good representations
- Captures essential elements of the event / world
- Deliberately leaves out / mutes the irrelevant
- Appropriate for the person and their interpretation
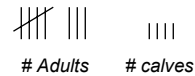- Appropriate for the task, enhancing judgment ability

How many buffalo?

‖‖‖ ‖‖‖ ‖
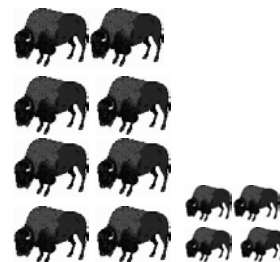*# Buffalo*

‖‖‖ ‖‖‖ ⫼
 # Buffalo

‖‖‖ ‖‖‖   ‖‖‖‖
 # Adults   # calves

8   4

Representation

A representation is
  - **A formal system or mapping by which the information can be specified**
  (D. Marr)
  - **A sign system in that it stands for something other than its self**.

For example: the number thirty-four *or* the buffalo example
  Decimal: 34,
  Binary: 100010,
  Roman: XXXIV

Different representations reveal different aspects of the information
  Decimal: counting & information about powers of 10,
  Binary: counting & information about powers of 2,
  Roman: counting

Presentation

Not the same as representation!

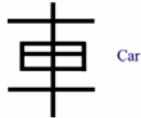**The presentation of information deals with how the representation is placed or organized on the screen**

$$5, \ \underline{\textbf{5,}}$$
$$\text{IV}, \ \underline{\textbf{\textit{IV}}}$$

One

Two

Three

Car

Mouth

Below

Representations: Finding Information

Solving a problem simply means representing it so as to make the solution transparent … *(Simon, 1981)*

Good representations
- Allow people to *find* the relevant information
  - In contrast the information may be present but hard to find

Representations: The Information Is Present But Hard To Find

Representations: Interpreting And Using The Information

- Allow people to *compute* desired conclusions
  - Trying to make use of the information may be a difficult process or "for free" depending on the representation chosen

Representations: The Information Is Present But Making Sense Of It Requires Much Effort

Quarterly Income Statements
in millions, except earnings per share

	Q2-01*	Q3-01*	Q4-01*	Q1-02	Q2-02	Q3-02	Q4-02	Q1-03	
Revenue	$ 6,550	$ 6,403	$ 6,577	$ 6,126	$ 7,741	$ 7,245	$ 7,253	$ 7,746	$
Operating expenses:									
Cost of revenue	864	899	867	978	1,691	1,567	1,463	1,344	
Research and development	990	1,069	1,364	1,398	1,595	1,474	1,832	1,707	
Acquired in-process technology									
Sales and marketing	1,290	1,198	1,359	1,457	1,676	1,449	1,670	1,415	
General and administrative	212	239	236	286	885	343	329	252	
Other expenses									
Total operating expenses	3,356	3,405	3,826	4,119	5,847	4,833	5,294	4,718	
Operating income	3,194	2,998	2,751	2,007	1,894	2,412	1,959	3,028	
Losses on equity investees and other	(28)	(46)	(33)	(30)	(37)	(11)	(14)	(22)	
Investment Income	751	706	(2,620)	(980)	553	739	(617)	41	
Noncontinuing items									
Income before income taxes	3,917	3,658	98	997	2,410	3,140	1,328	3,047	
Provision for income taxes	1,293	1,207	33	319	771	1,005	425	1,006	
Income before accounting change	2,624	2,451	65	678	1,639	2,135	903	2,041	
Cumulative effect of accounting change	-	-	-	-	-	-	-	-	
Net income	$ 2,624	$ 2,451	$ 65	$ 678	$ 1,639	$ 2,135	$ 903	$ 2,041	$
Preferred stock dividends									
Net income available for common shareholders	$ 2,624	$ 2,451	$ 65	$ 678	$ 1,639	$ 2,135	$ 903	$ 2,041	$
Basic EPS before accounting change	$0.25	$0.23	$0.01	$0.06	$0.15	$0.20	$0.08	$0.19	
Diluted EPS before accounting change	$0.24	$0.22	$0.01	$0.06	$0.15	$0.19	$0.08	$0.19	

James Tam

Example One: Which Is The Best Flight?

Length, stop-overs, switches...

		Depart	Arrive
AC 117	Vancouver - Calgary	7:00	9:00
Cdn 321	Vancouver - Calgary	9:00	12:00
Cdn 355	Calgary - Montreal	13:30	19:30
AC 123	Calgary - Toronto	12:30	16:30
AC 123	Toronto - Montreal	16:45	17:30

*time zone: +1 van-cal, +2 cal-tor, mtl
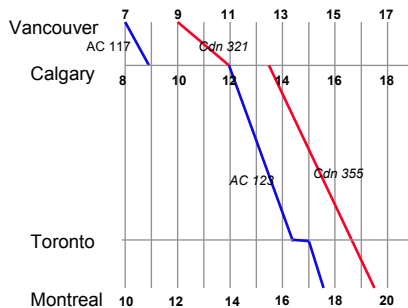
James Tam

Example Two: When Do I Take My Drugs? (From "Things That Make Us Smart" By Don Norman

Note: 10 - 30% error rate in taking pills, same for pillbox organizers

Inderal   - 1 tablet 3 times a day
Lanoxin -1 tablet every a.m.
Carafate - 1 tablet before meals and at bedtime
Zantac   - 1 tablet every 12 hours (twice a day)
Quinag - 1 tablet 4 times a day
Couma - 1 tablet a day

	Breakfast	Lunch	Dinner	Bedtime
Lanoxin	O			
Inderal	O	O	O	
Quinag	O	O	O	O
Carafate	O	O	O	O
Zantac		O		O
Couma				O

Breakfast	Lunch	Dinner	Bedtime
Lanoxin			
Inderal	Inderal	Inderal	
Quinag	Quinag	Quinag	Quinag
Carafate	Carafate	Carafate	Carafate
	Zantac		Zantac
			Couma

Organized by both time of day and by drug

Example Three: Napoleon's March To Moscow
by Charles Minard

Example Three: Napoleon's March To Moscow
by Charles Minard

CARTE FIGURATIVE des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812–1813.
Dressée par M.Minard, Inspecteur Général des Ponts et Chaussées en retraite.

TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro

9bre = November   8bre = October

James Tam
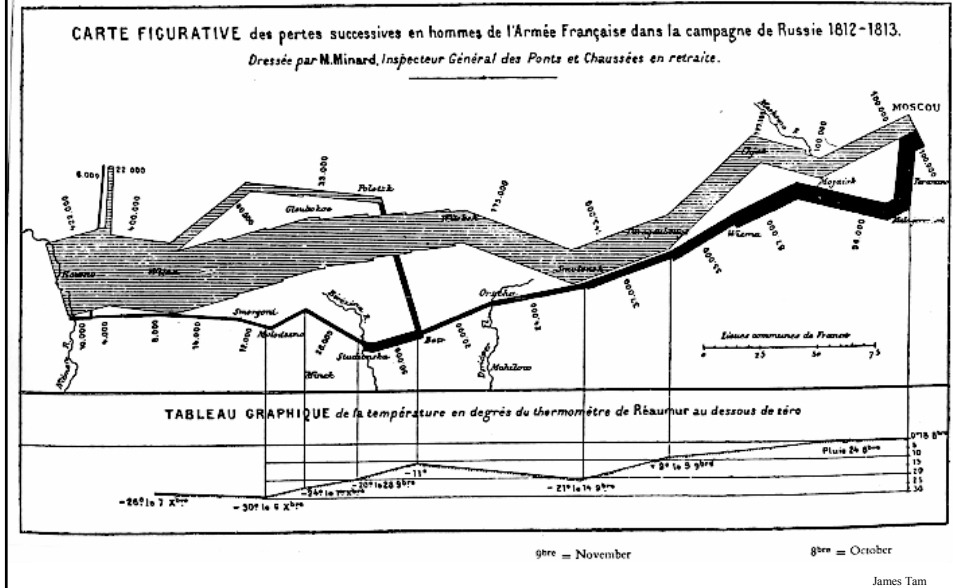
Example Four: Anscombe's Quartet

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
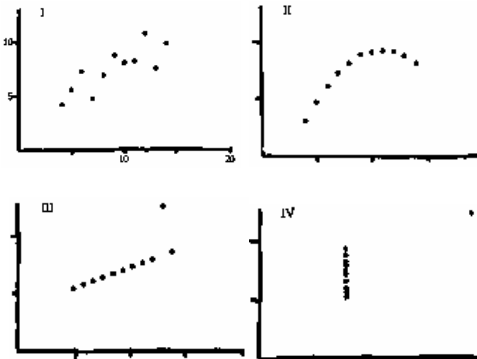
N:	11.0
mean X's :	9.0
mean Y's :	7.5
standard error of slope estimate:	0.1
sum of squares:	110.0
regression sum of squares:	27.5
residual sum of squares of Y:	13.8
correlation coefficient:	0.8
r squared:	0.7
regression line:	Y=3+0.5X

James Tam

Example Five: Do I Deserve A Tax Break

	A	B
1	Market value ($)	Improvement cost ($)
2	140000	31120
3	147000	29980
4	151000	38120
5	152000	34360
6	155000	40710
7	170000	21620
8	172000	42100
9	178000	41070
10	180000	34210
11	180000	44090
12	182000	55960
13	185000	45170
14	185000	46820
15	193400	50200
16	194500	71860
17	197000	48460
18	203000	40720
19	205000	56600
20	213000	42780
21	221000	58770
22	225000	58960
23	245000	48910
24	248000	62620
25	278000	58580
26	302500	72200
27	308000	67320
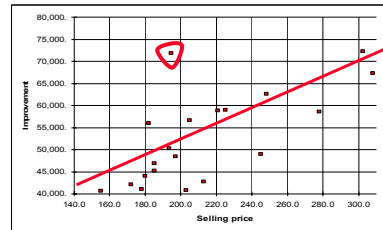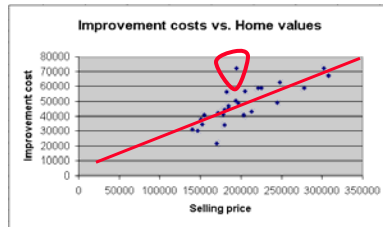
Which Representation Is Best?

Depends heavily on task

What is the precise value?
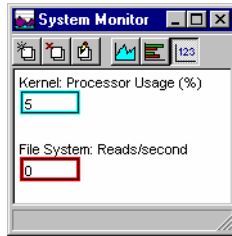
What is the performance now compared to the peak?
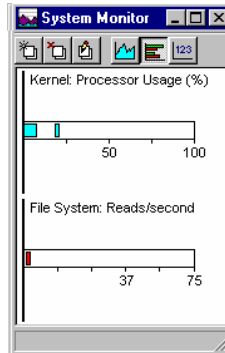
How does performance change over time?
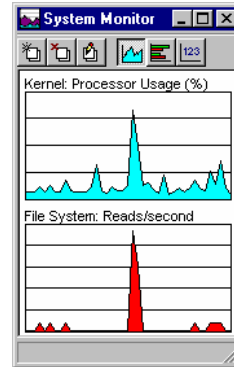
Visualization

- **Representing data in a way that amplifies cognition (acquiring and using knowledge).**

- **It's related to representations:**
  - Representation are methods of encoding information.
  - Visualizations are good representations

An Example Visualization: The Change History Of A Software System : SeeSoft[1]
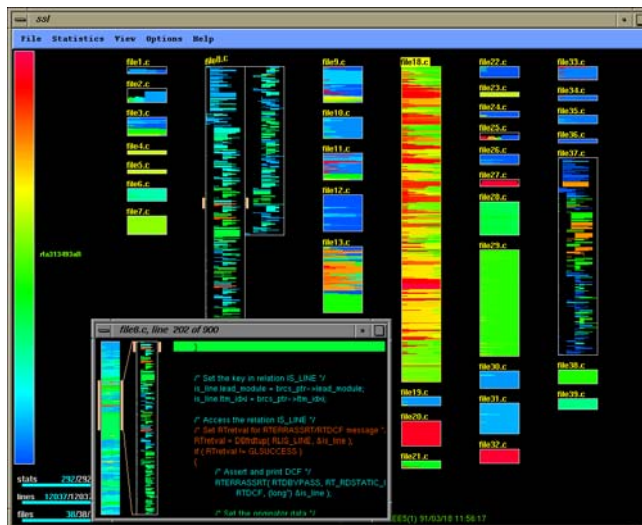
1) *"Seesoft—A Tool for Visualizing Line Oriented Software Statistics",* Eick S.G., Steffen J.L. and Sumner E.E

Tufte's Principles Of Information Visualization[1]

Graphics should reveal the data

- Show the data
- Not get in the way of the message
- Avoid distortion
- Present many numbers in a small space
- Make large data sets coherent
- Encourage comparison between data
- Supply both a broad overview and fine detail
- Serve a clear purpose

1 *"Visual Display of Quantitative Information" by E. Tufte*
  Note: Some of the visual examples on the following slides are taken from
  Tufte's books

Show The Data

Buffalo

Buffalo

Adults   # calves
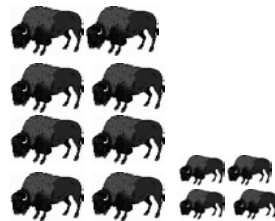
Not Get In The Way Of The Message
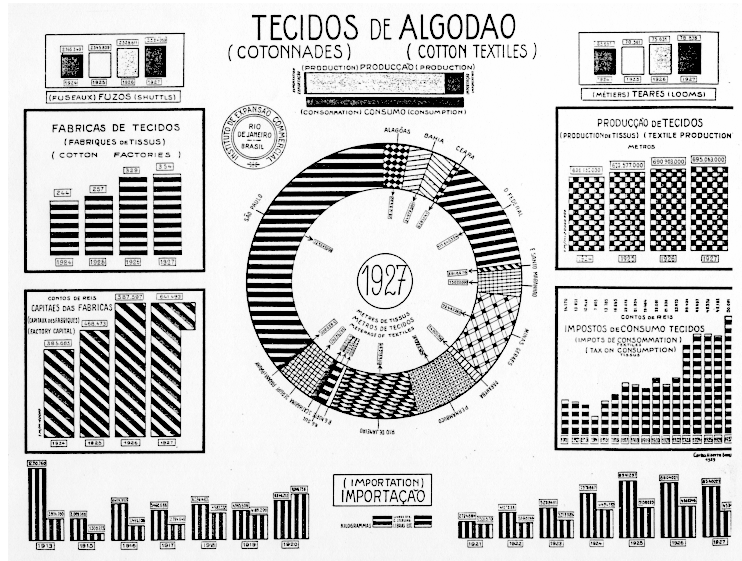
Chart Junk: A Common Error (The Representation Getting In The Way Of The Message)

Information display is not just pretty graphics
- Graphical re-design by amateurs on computers gives us
  - Overly complicated or even deceptive representations
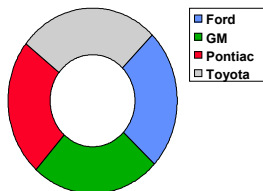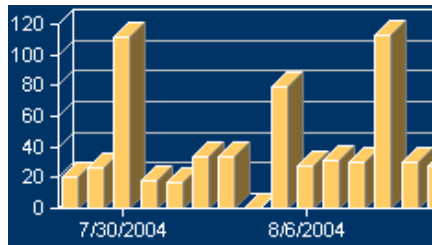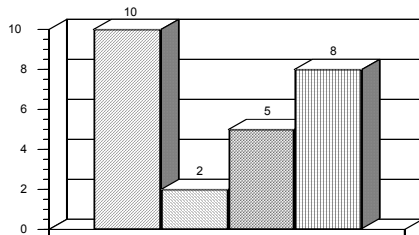
Avoid Distortion

Maintenance cost per year

New York Weather History

Present Many Numbers In A Small Space, Make Large Data Sets Coherent

New York Weather History

- 181 numbers/sq inch
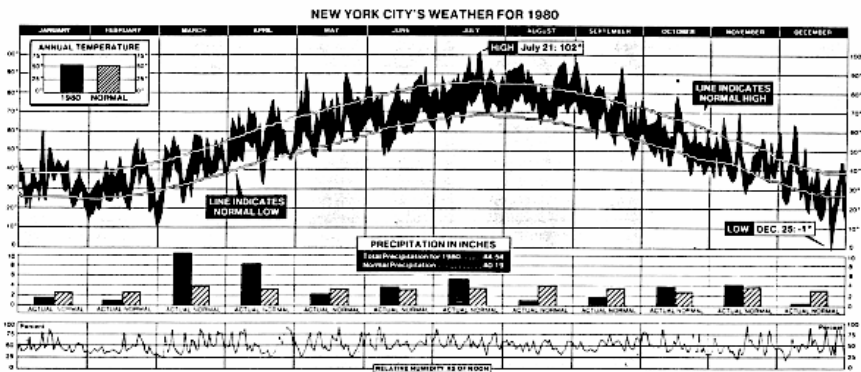
James Tam

Encourage Comparison Between The Data
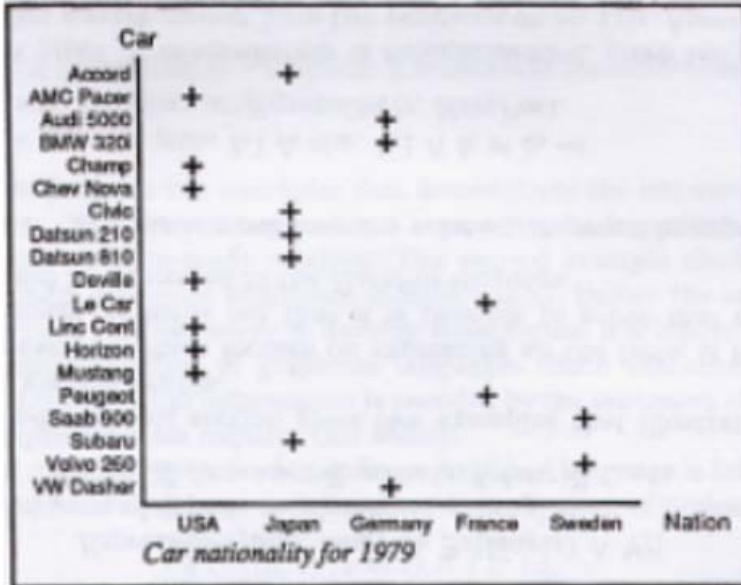
MacKinlay J.D. (1986) Automatic Design of Graphical Presentations.

James Tam

Broad Overview And Fine Detail
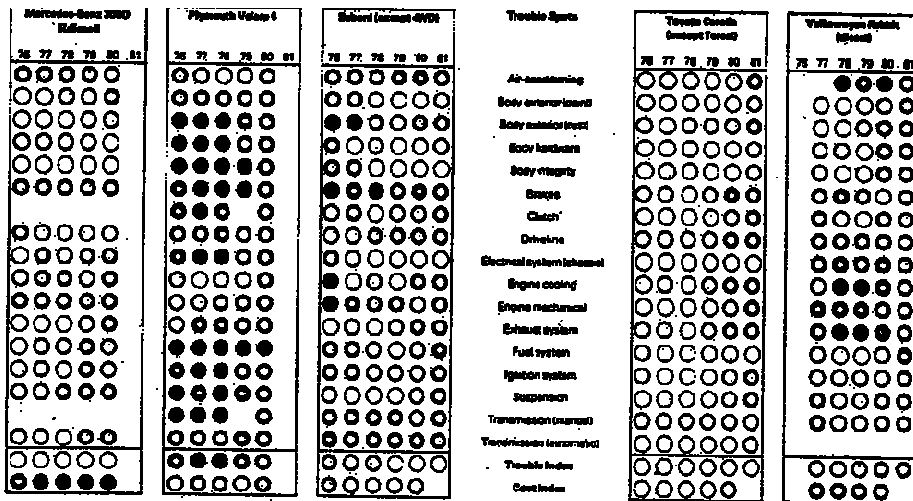
James Tam

Automatic Vs. Controlled Processing Of Information

Automatic

- Cannot be inhibited

- Superficial

- Fast, high capacity, can be processed in parallel
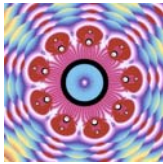
Controlled

- Requires conscious effort

- Detailed

- Slow, low capacity, can only be processed in a serial fashion

For this question you are to write a function that will take as input a string and return an integer value that is the length of the string. The end of the string will always...

Visual Variables

Position
  • Changes in the x, y, z location
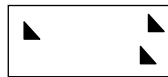
Size
  • Changes in length, area or repetition

Shape
  • Changes in form

Value
  • Changes in brightness

Orientation
  • Changes in alignment
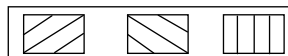
Colour
  • Changes in hue
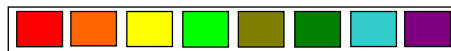
Visual Variables (2)

Texture
  • Variations in pattern

Motion

www.st-duffer.com

Visual Variables (3)

Characteristics of visual variables

• **Selective**
  Is a change in this variable enough to allow us to select it *from a group*?

• **Associative**
  Is a change in this variable enough to allow us to perceive them *as a group*?

• **Quantitative**
  Is there a numerical reading obtainable from changes in this variable?

• **Order**
  Do changes in the visual variable indicate some sort of ranking?

• **Length**[1]
  Across how many changes in this variable are distinctly perceptible?

1 Think of it as variation

Visual Variable: Position

✓ **Selective**
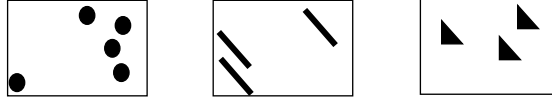
✓ **Associative**
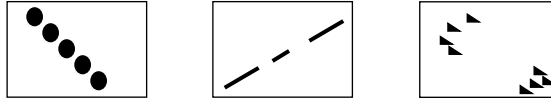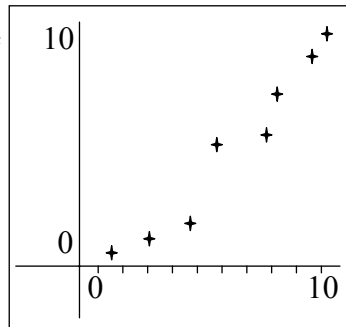
✓ **Quantitative**

✓ **Order**

✓ **Length**

Visual Variable: Size

✓ **Selective**

✓ **Associative**
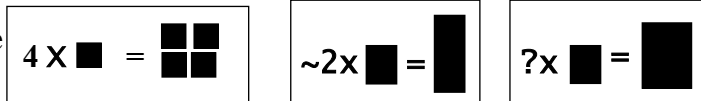
⌣ **Quantitative**

4 X ■ = ■■
          ■■

~2x ■ = ▮

?x ■ = █

✓ **Order**

● > ● > ● > ● > ● > • > •
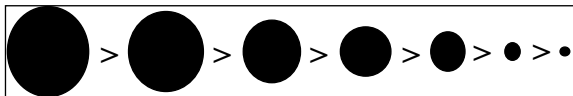
✓ **Length**
  Theoretically infinite but practically limited

Visual Variable: Shape

⌣ **Selective**

⌣ **Associative**

≠ **Quantitative**   ➕ ≠ 15   ▲ ≠ 34   ⬟ ≠ 105

≠ **Order**   ➕ ⋡ ● ⋡ ▲ ⋡ ◗ ⋡ ◖ ⋡ ■ ⋡ ⬟ ⋡ ▼

✓ **Length**   ➕ ● ▲ ✦ ◢ ◖ ■ ⬟ ★ ▼ ✹ ≈ **...**

Shape
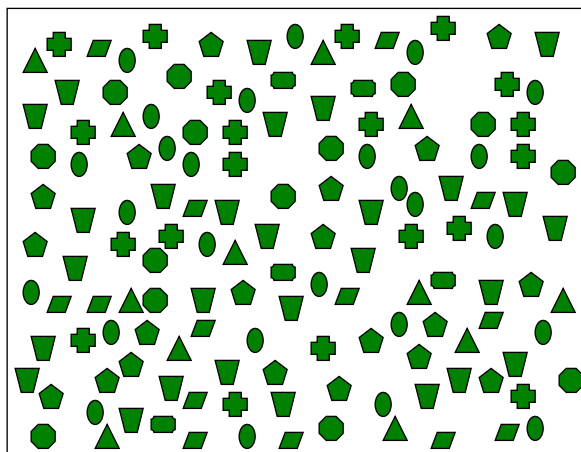
Visual Variable: Value

✓ Selective

✓ Associative
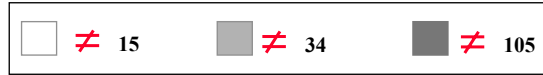
✗ Quantitative   □ ≠ 15   ▢ ≠ 34   ▪ ≠ 105
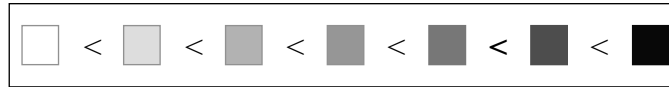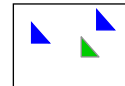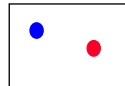
✓ Order   □ < ▢ < ▢ < ▪ < ▪ < ▪ < ■

✓ Length
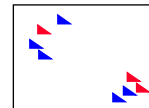- Theoretically infinite but practically limited
- Association ~ < 7 and selection ~ 10

James Tam

Visual Variable: Color

✓ Selective

✓ Associative
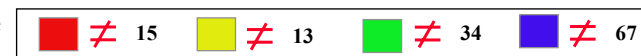
✗ Quantitative   ■ ≠ 15   ■ ≠ 13   ■ ≠ 34   ■ ≠ 67

✗ Order   ■ ≯ ■ ≯ ■ ≯ ■ ≯ ■ ≯ ■ ≯ ■ ≯ ■ ≯ ■
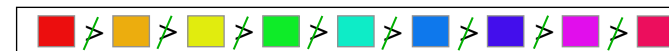
✓ Length
- Theoretically infinite but practically limited
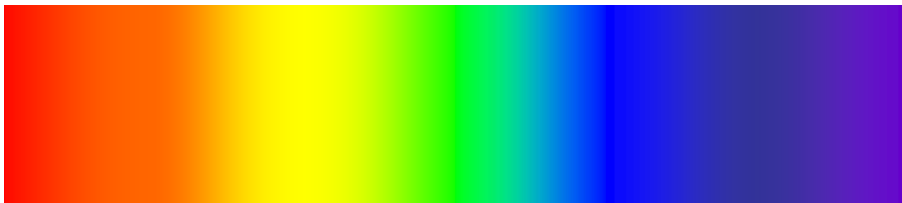- Association ~ < 7 and selection ~ 20

James Tam
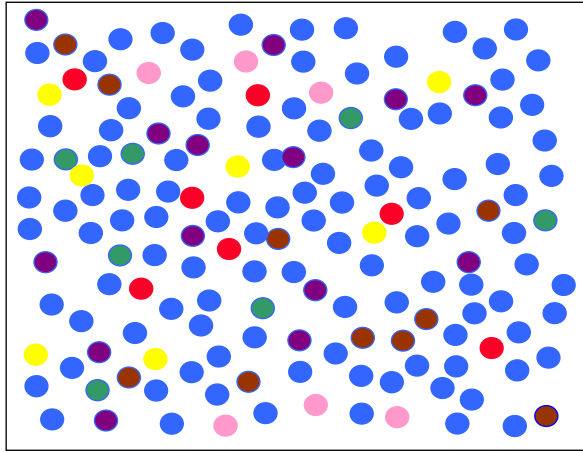
Color
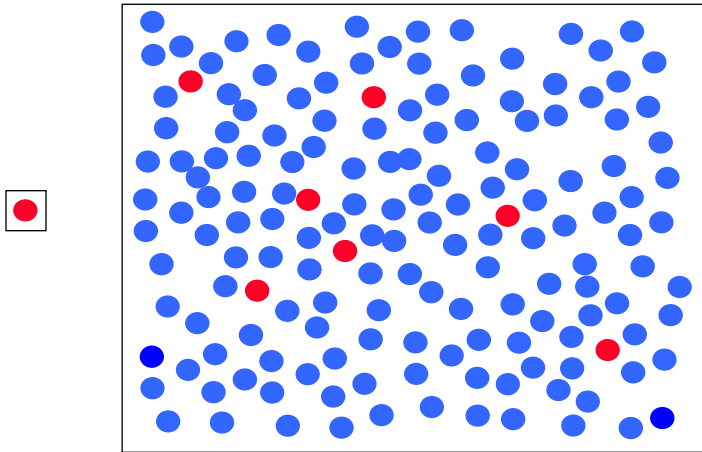
Color Encoding

Common advice says use a rainbow scale
- Marcus, Murch, Healey
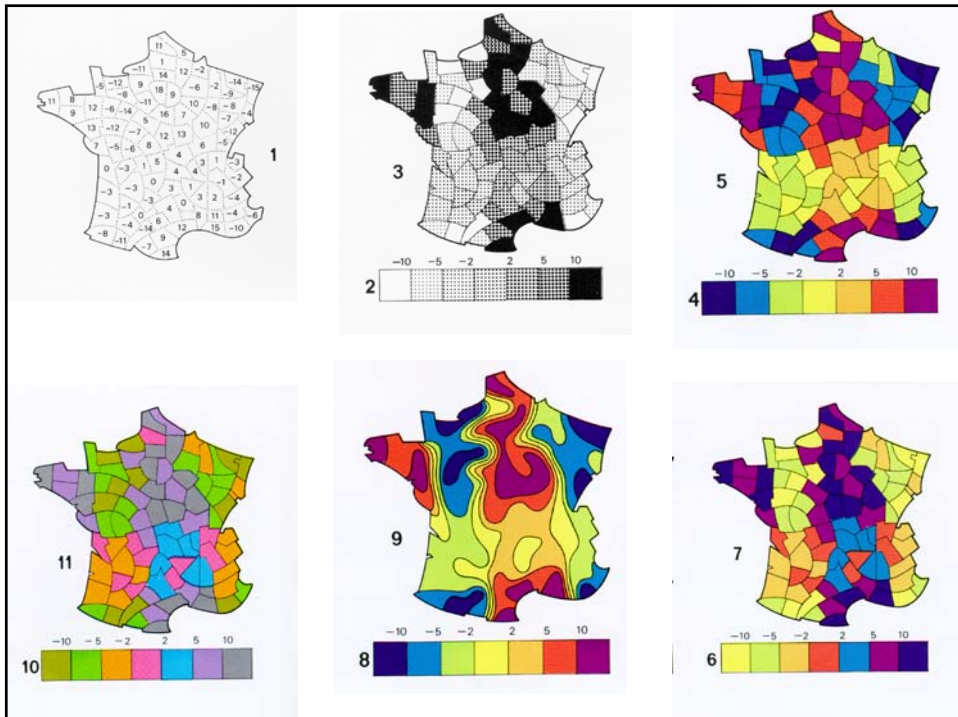- There are problems with rainbows

Color

James Tam

James Tam

General Bathymetric Chart of the Oceans,
International Hydrographic Organization
(Ottawa, Canada, 5th edition, 1984). 3.06.
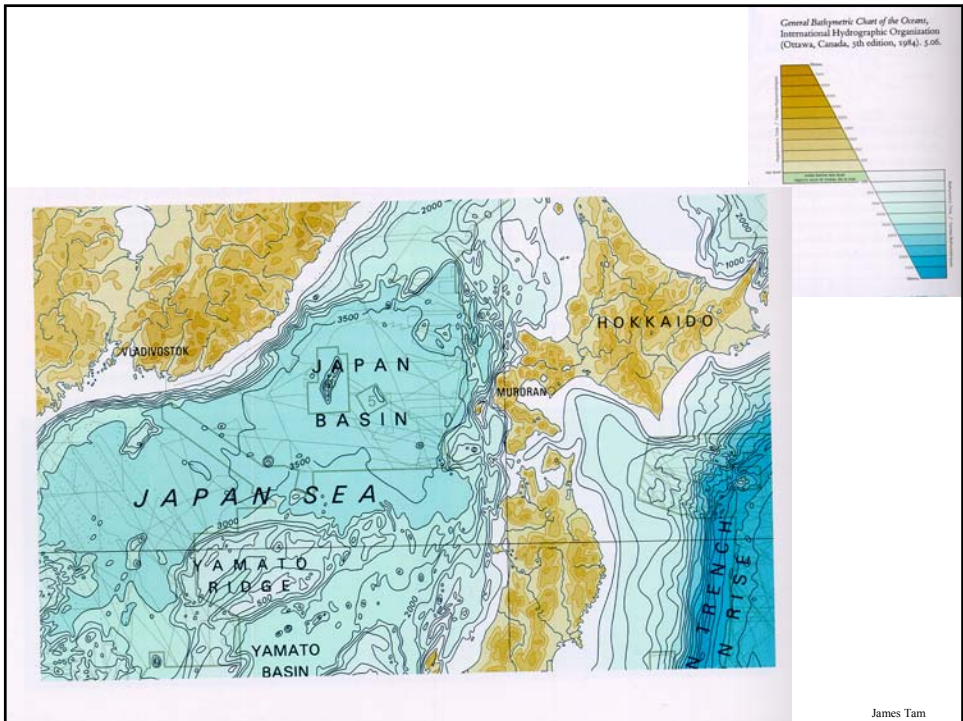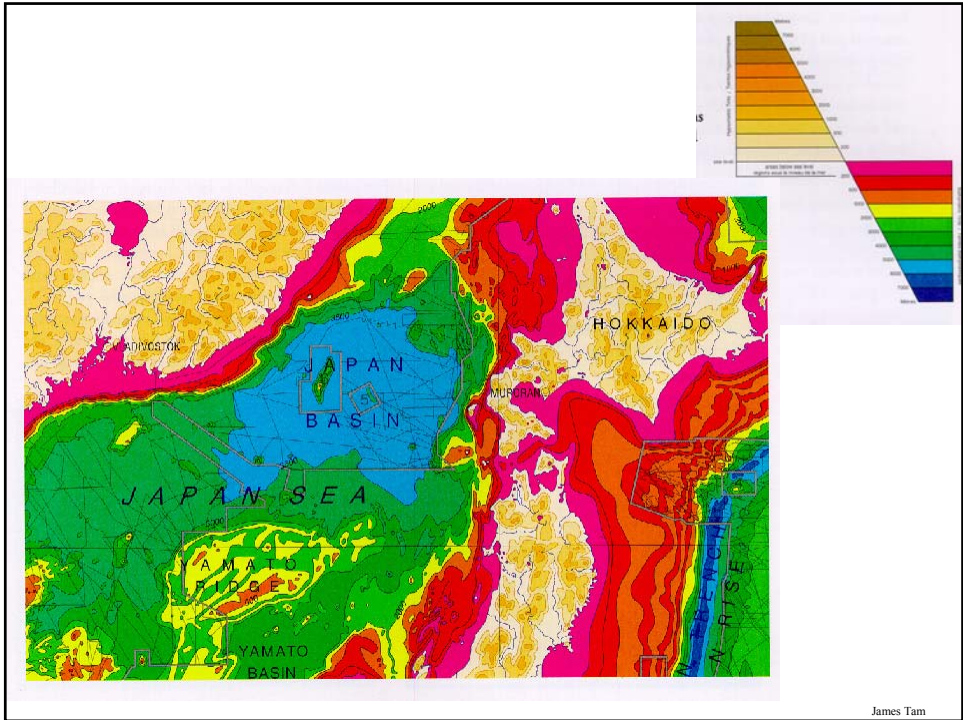
James Tam

Visual Variable: Orientation

✓ **Selective**

✓ **Associative**

✗ **Quantitative** ≠ 15   ≠ 17   ≠ 95
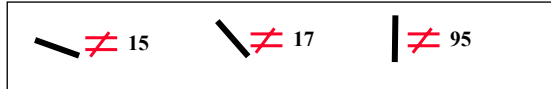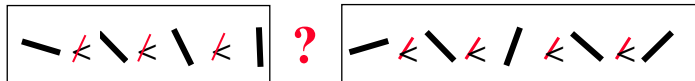
✗ **Order**   **?**

✓ **Length**

~5 in 2D

? in 3D

James Tam

Visual Variable: Texture

✓ **Selective**

✓ **Associative**

✗ **Quantitative** ≠ 45   ≠ 15   ≠ 33

✗ **Order**

✓ **Length**
  • Theoretically infinite

James Tam

Visual Variable: Motion

✓**Selective -** motion is one of our most powerful attention grabbers

✓**Associative –** objects moving in unison groups them effectively

≠**Quantitative -** subjective perception

≠**Order**

? **Length - distinguishable types of motion?**

Motion
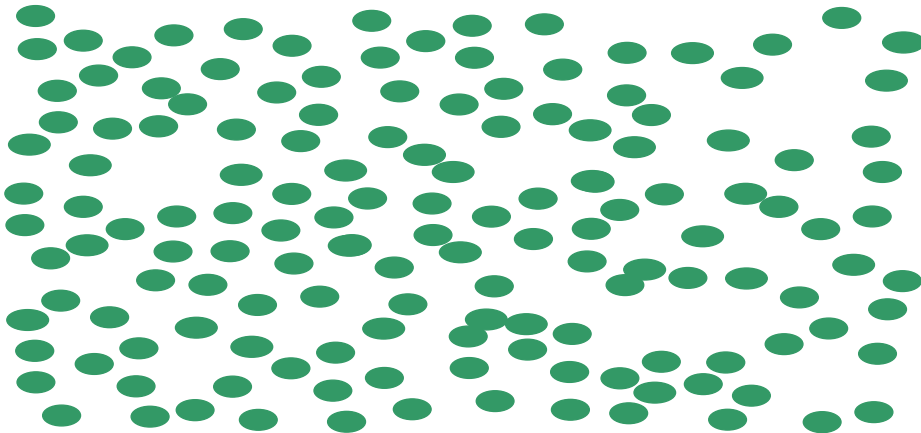
Small Multiples: General Principles

Learn once

Invite comparisons
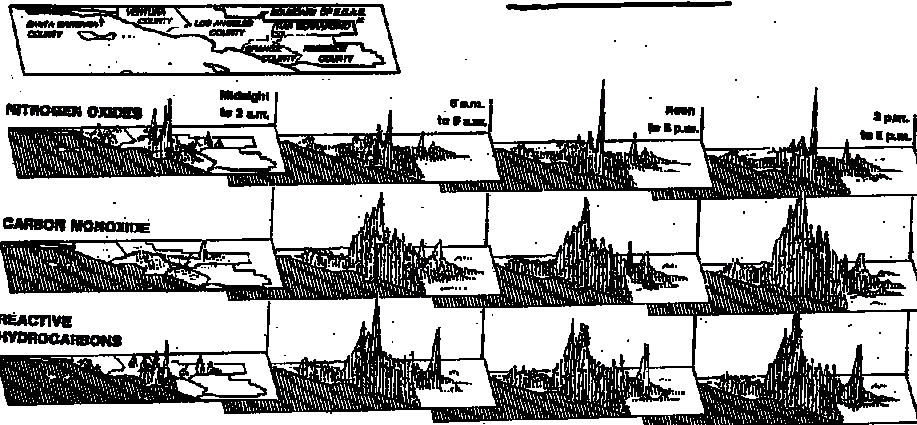
James Tam

Small Multiples: Showing Time And Change

1:45:52

James Tam

Small Multiples: Showing Time And Change
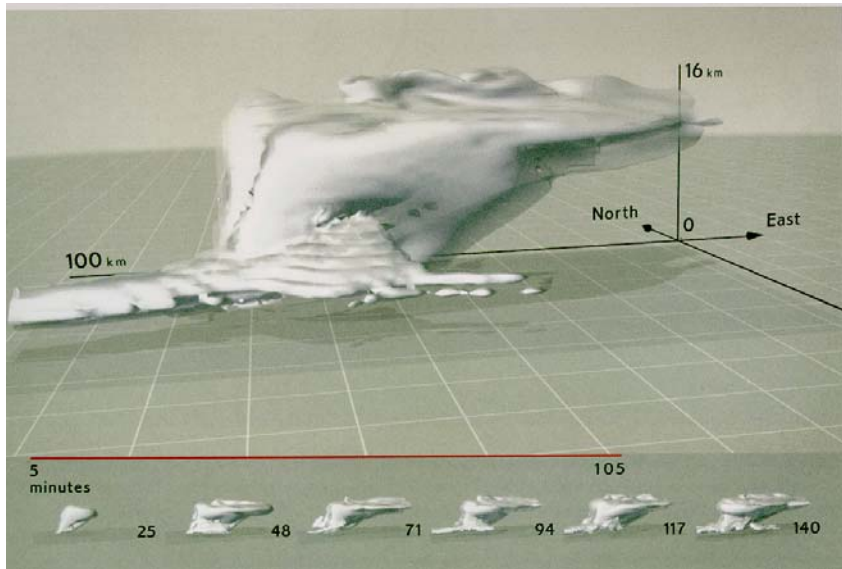
Metaphors

Definition of a Metaphor
- One kind of object or idea is used in place of another to suggest a likeness or analogy between them
- Application of name or descriptive term to an object to which it is not literally applicable

Interface Metaphors

Purpose
- Function as natural models
- Leverages our knowledge of familiar, concrete objects/experiences to understand abstract computer and task concepts
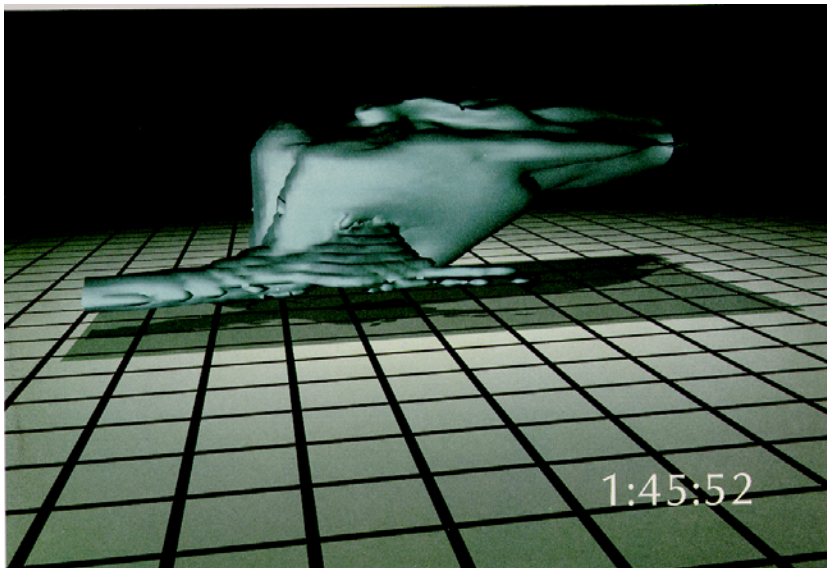
Problem
- Metaphor may portray inaccurate or naive conceptual model of the system

A presentation tool is like a slide projector
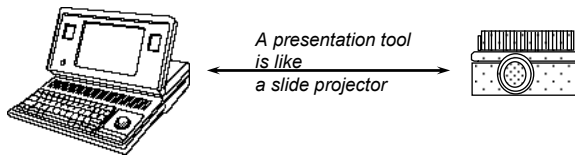
Interface Metaphors

Pervade excellent interfaces
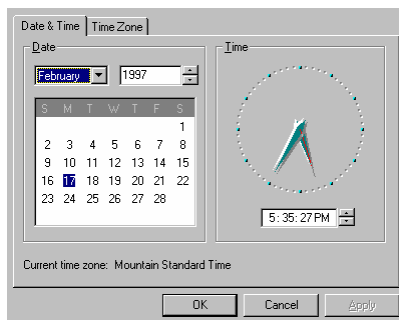
Control Panels with familiar controls

Name:_____
Address:_____
City:_____
Province:_____
Postal Code:_____

Forms
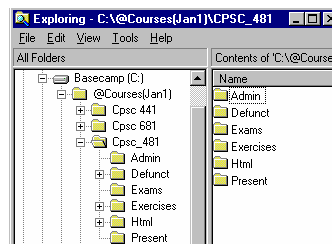
Hierarchical Folders

A Real Life Metaphor: Life!

Creating Interface Metaphors

Generating metaphors

- Use metaphors that matches user's conceptual task
  - Desktop metaphor for office workers
  - Paintbrush metaphor for artists...
- Given a choice, choose the metaphor close to the way the system works
- Ensure emotional tone is appropriate to users
  - e.g., file deletion metaphors
    Trashcan
    Black hole
    Paper shredder
    Pit bull terrier
    Nuclear disposal unit...

Evaluating Metaphors

Potential problems:
- The metaphor has attributes that the system does not have.
- The system has attributes that are not suggested by the metaphor.
- An attribute exists both in the metaphor and in the system but works differently in each.

The Metaphor Is More Powerful Than The System

- **Will the metaphor make people believe that the system can do more than it currently can?**

Zork © Infocom

The System Is More Powerful Than The Metaphor Implies

• **Will the metaphor restrict how people will try to use the system?**
  •e.g., file folders

An Attribute Differs Between The Metaphor And The System

e.g., The trash can

Real trashcan

Desk top trashcan

Metaphors Should Not Be Static

Evolve metaphors
- Is metaphor extensible to new features?
- When is the metaphor no longer useful?

Dilbert © United Features Syndicate
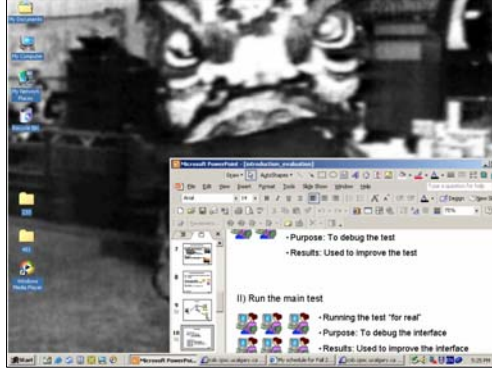
Metaphors Should Not Be Static (2)

Evolve metaphors
- Is metaphor extensible to new features?
- When is the metaphor no longer useful?
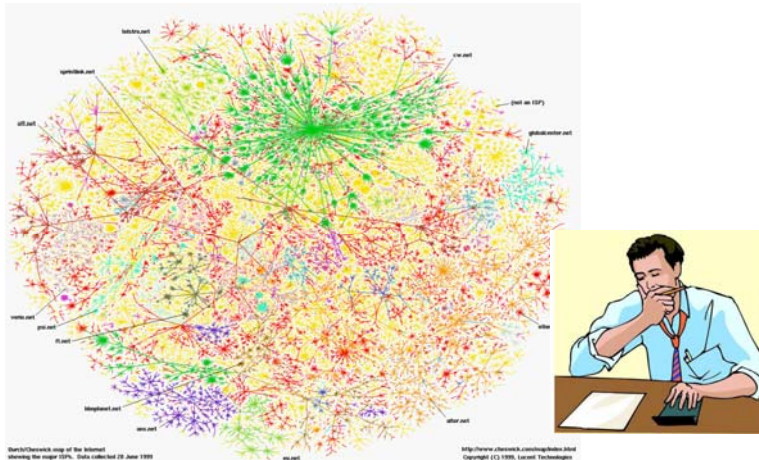
Misuse Of Metaphors

Caveat
- Metaphors can be overdone!

Common pitfalls
- Overly literal
  - Unnecessary fidelity
  - Excessive interactions

- Overly cute
  - Novelty quickly wears off

- Overly restrictive
  - Cannot move beyond

- Mismatched
  - Does not match user's task and/or thinking
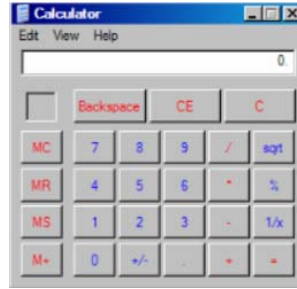
Misuse Of Metaphors (2)
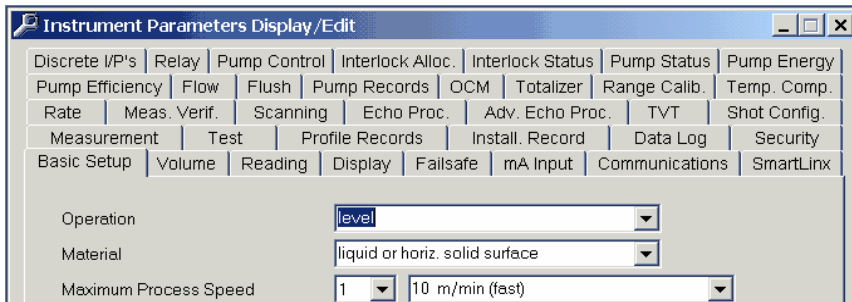
Milltronics' *Dolphin Plus* a configuration package for industrial level and flow sensors

A Example System That Applies A Metaphor: The Data Mountain

Robertson / Czerwinski / Larson / Robbins / Thiel / van Dantzich
Data Mountain: Using Spatial Memory for Document Management Proc ACM UIST'98

A Example System That Applies A Metaphor: The Data Mountain

Robertson / Czerwinski / Larson / Robbins / Thiel / van Dantzich
Data Mountain: Using Spatial Memory for Document Management Proc ACM UIST'98

Direct Manipulation

- **An interface that behaves as though the interaction was with a real-world object rather than with an abstract system**
- **Almost always based on a metaphor**
  - Mapped onto some facet of the real world task semantics
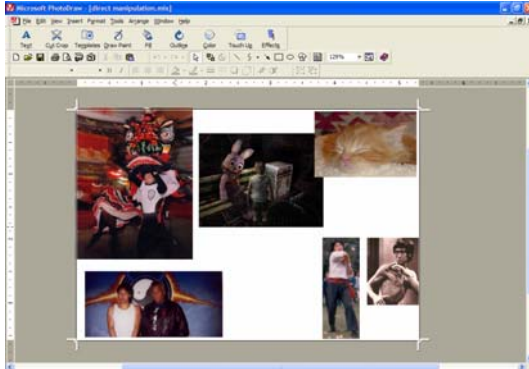
PhotoDraw © Microsoft

Characteristics Of Direct Manipulation

- **Objects of interest are visible**

- **Visible objects can be manipulated**

- **Manipulation occurs by pointing and moving**

Indirect Interaction: Traditional Approach To Writing Java GUI's

Indirect Interaction: Changing The Java Code

Direct Manipulation

Representation directly affects what can be directly manipulated

Schedule © Microsoft
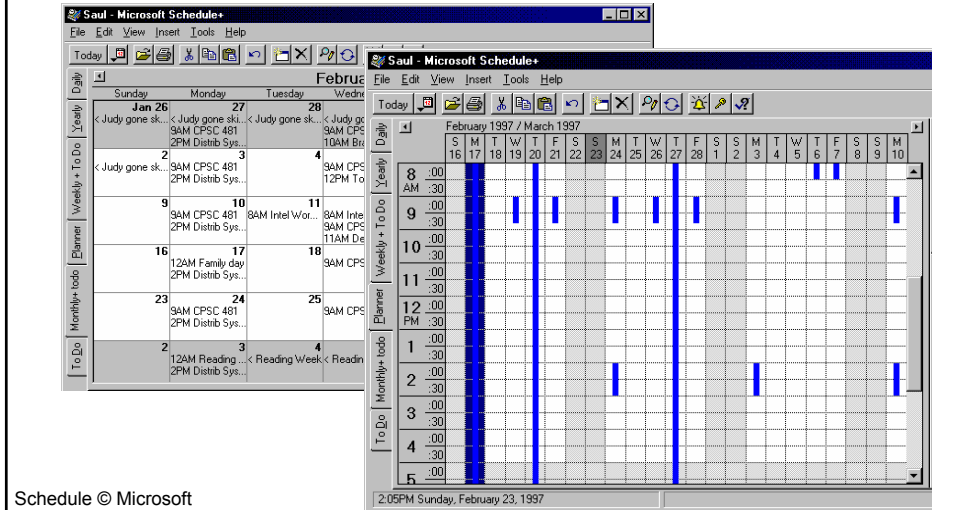
Is Direct Manipulation The Way To Go?

Some Disadvantages

• Ill-suited for abstract operations
  - Spell-checker?

• Tedium
  - Manually search large database vs. query

• Task domain may not have adequate physical/visual metaphor

• Metaphor may be overly-restrictive

Solution

• Most systems combine direct manipulation and abstractions
  - Word processor:
      WYSIWYG document (direct manipulation)
      buttons, menus, dialog boxes (abstractions, but direct manipulation "in the small")

James Tam

Conventional Applications: A Mix
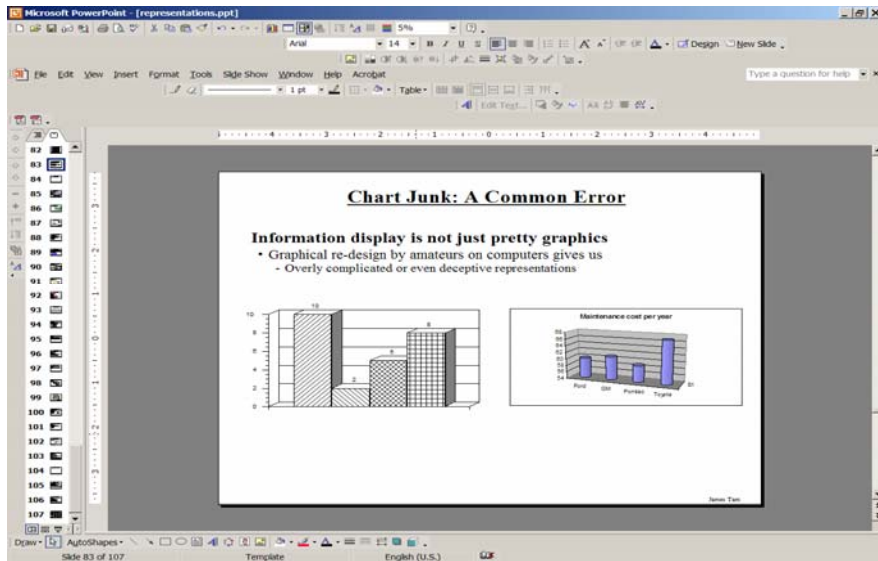
Direct Engagement

- **The feeling of working directly on the task.**
- **To employ it, you need to consider the user of the system and the tasks that he or she engages in.**
- **Often direct manipulation is an important requirement for direct engagement.**
  - e.g., A drawing program provides tools that are familiar to artists (brushes, palettes etc.)

Direct Engagement: A Telephone Database

```
Find "Green"
>S. Greenberg
>Dept Computer Science
>University of Calgary
```

Command system
no direct manipulation

```
Search for: Green

Result:    S. Greenberg
           Dept Computer Science
           University of Calgary
```

Form metaphor:
syntactic direct
manipulation
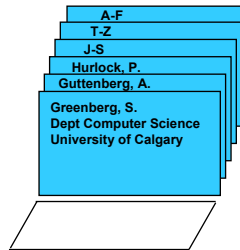
Rolodex metaphor:
full direct manipulation

A-F
T-Z
J-S
Hurlock, P.
Guttenberg, A.
Greenberg, S.
Dept Computer Science
University of Calgary

James Tam

Object-Action *vs.* Action-Object

Select object, *then* do action

 • Interface emphasizes 'nouns' (visible objects) rather than 'verbs' (actions)

move

my.doc
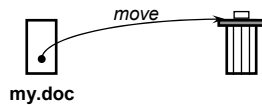
The Sims House Party © Maxis

James Tam

Object-Action *Vs.* Action-Object (2)

Advantages
- Closer to real world
- Modeless interaction
- *Actions* always within context of object
  - Inappropriate ones can be hidden
- *Generic commands*
  - The same type of action can be performed on the object
  - e.g., drag 'n drop:
    - folders
    - files
    - paragraphs
    - text
    - numbers…

What You Now Know

Good Representations
- Captures essential elements of the event / world
- Deliberately leaves out / mutes the irrelevant
- Appropriate for the person, their task, and their interpretation

Information Visualization
- Tufte's principles
- Exploits our knowledge of visual variables
- Many techniques now available (illustrated with research and commercial systems)

What You Now Know (2)

Metaphors
- Uses our knowledge of the familiar and concrete to represent abstract concepts
- Need not be literal
- Has limitations that must be understood

Direct manipulation
- Visibility of the objects of interest
- Manipulation by pointing and moving

These four components are the foundation of a true Visual Interface

---

Interface Design And Usability Engineering

Goals:

- Articulate:
  - Who (users)
  - What (tasks)
- Brainstorm designs
- Refined designs
- Completed designs

Methods:

- Task centered system design
- Participatory design
- User-centered design
- *Evaluate*
- Psychology of everyday things (*psych*)
- User involvement (*user*)
- Representation & metaphors
- *Participatory interaction*
- *Task scenario walk-through*
- low fidelity prototyping methods
- Psych, User, Representations and metaphors
- Graphical screen design
- Interface guidelines
- Style guides
- *Usability testing*
- *Heuristic evaluation*
- high fidelity prototyping methods
- *Field testing*

Products:

- User and task descriptions
- Throw-away paper prototypes
- Testable prototypes
- Alpha/beta systems *or* complete specification

This diagram is a variation of the one presented by Saul Greenberg