

# Information Visualization and Knowledge Management.....

**Dr. Margaret-Anne Storey**

Dept of Computer Science, UVic

# Outline

- Demonstrations/Videos & Discussion
  - Dynamic Queries
  - Alphasliders
  - Spotfire (starfield displays)
  - Magic filters
  - Tilebars
  - Bargrams
  - InfoCrystal
  - Etc.

# Database Query

**Select** house-address  
**From** atl-real-db  
**Where** price  $\geq$  200,000 **and**  
          price  $\leq$  400,000 **and**  
          bathrooms  $\geq$  3 **and**  
          garage == 2 **and**  
          bedrooms  $\geq$  4

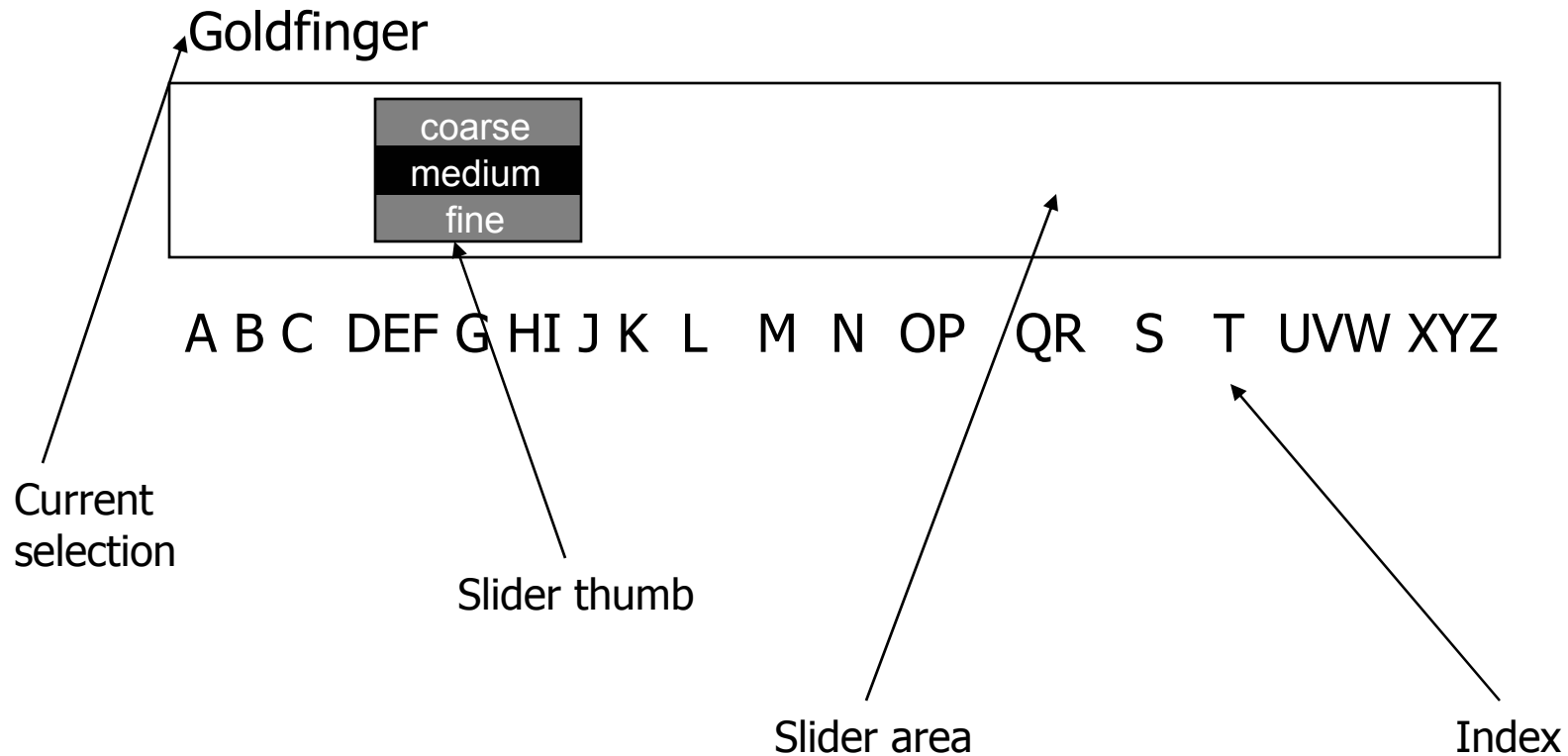
# Dynamic Queries

- Dynamic queries let users “fly through” databases by adjusting widgets and viewing the animated results
- Visual representation of world of action including both the objects and actions
- Rapid, incremental and reversible actions
- Selection by pointing (not typing)
- Immediate and continuous display of results

# Alpha sliders

- Sliders are created automatically using values retrieved from the database
- Different way of showing results – the slider becomes a result bar

# Alphaslider

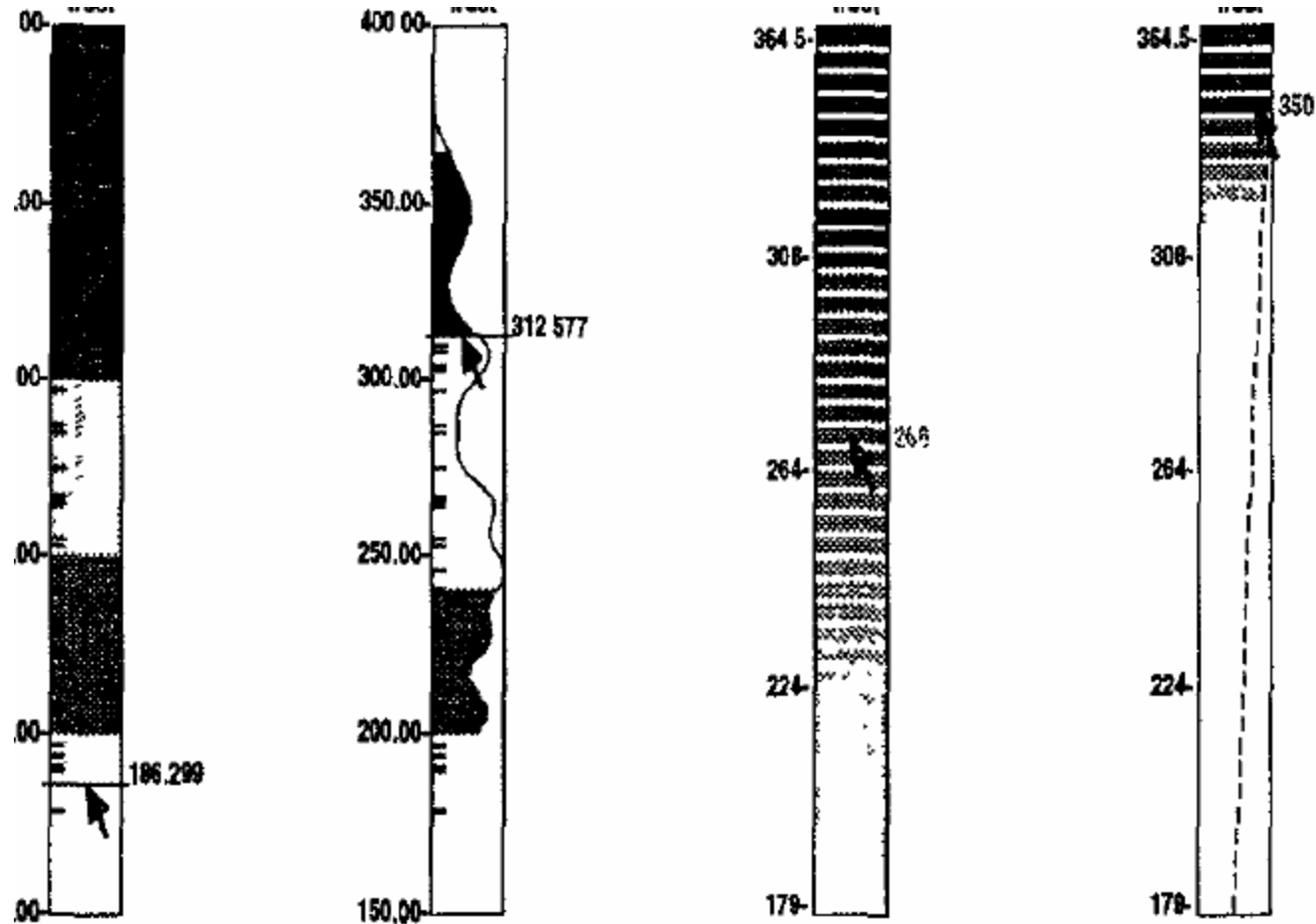


Slide adapted from John Stasko/Jock Mackinlay

# Data visualization sliders

- Space in slider can be used as an interactive color scale
- Bar plot for discrete data
- Density plot for continuous data

# Data Visualization Sliders



Data Visualization Sliders by Steve Eick in UIST'94



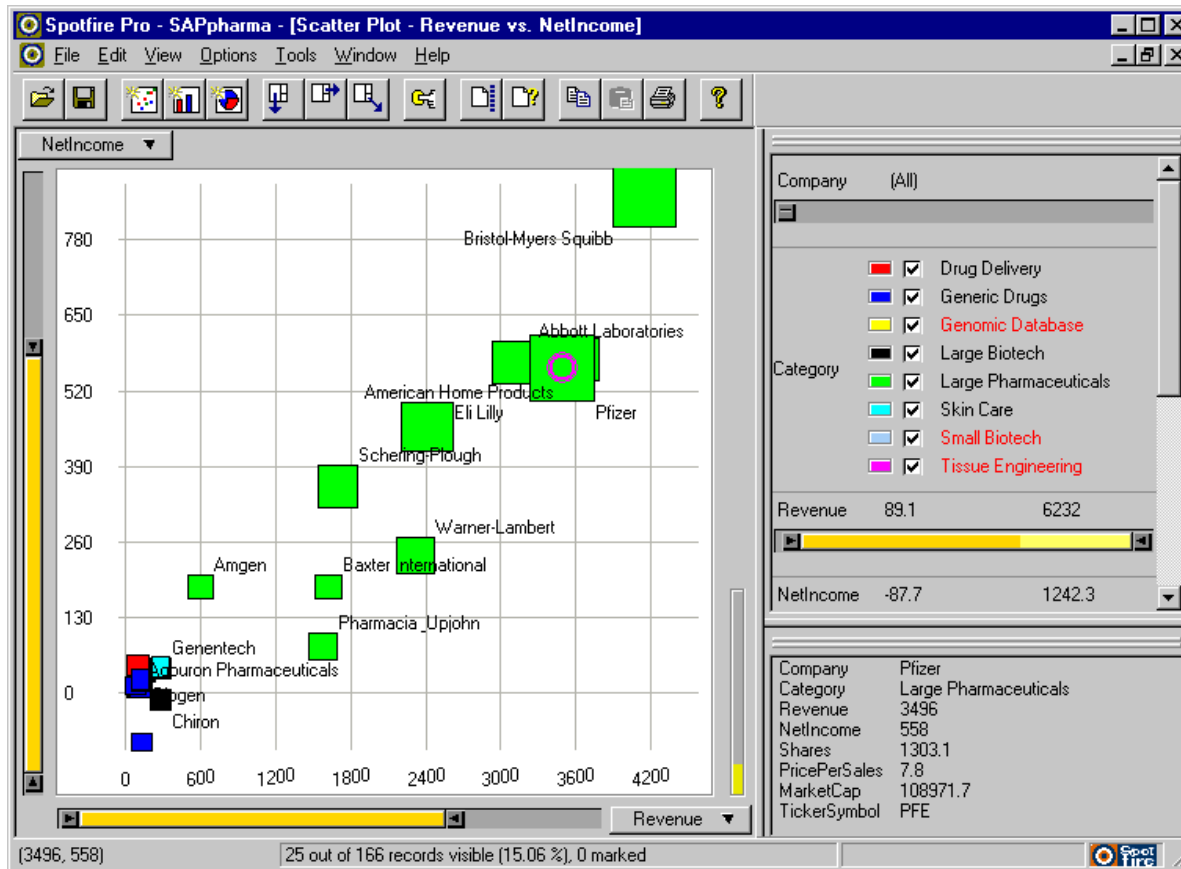
# Starfield displays

- 2D scatterplots to structure result sets
- Zooming to reduce clutter
- Use of direct manipulation: have a visual representation of the world of action, both the objects and the action
- Rapid, incremental and reversible actions
- Selection by pointing (not typing)
- Immediate and continuous display of results

## *Boolean Queries.....*

Discussion topic: spend a few mins and talk to one or two classmates about your experiences and the average consumer has when composing boolean queries – how is it done, how could it be improved?

# Starfield displays



# Starfield displays

- Spotfire Demo:  
[http://www.spotfire.com/products/flash\\_demo.asp?id=1](http://www.spotfire.com/products/flash_demo.asp?id=1)
- Tight coupling
  - features to guide the user
  - rapid, incremental, reversible interactions
  - display invariants
  - continuous display
  - progressive refinement
  - details on demand
  - use of constraints
  - comprehensive and consistent affordances to guide users

# Other examples of Dynamic Queries

- (we've already seen these briefly.... More video)
- Homefinder
- Filmfinder:  
[http://www.open-video.org/segment.php?seg\\_id=708](http://www.open-video.org/segment.php?seg_id=708)
- Visible human project:  
[http://www.open-video.org/segment.php?seg\\_id=706](http://www.open-video.org/segment.php?seg_id=706)
  - Developed by Ben Shneiderman and Chris Ahlberg
  - IEEE Software 11(6) 1994
  - CHI'94
- Mapexplorer
- Attribute explorer:  
<http://www-106.ibm.com/developerworks/library/us-atex/AEApplet.html>  
(and see desktop shortcut)

# Discussion -- Advantages

- Both novices and experts don't have to learn a command language
- Users can explore and discover relationships using dynamic queries
- Get context for results (which parts of the database are densely or sparsely populated)
- Can see the gaps, outliers, trends
- Get more than just exact hits
- Can use them to see where there are correlations, dependencies (form hypotheses etc)
- Reversible
- Fast, fun
- Preparing reports

# Discussion -- Disadvantages

- Queries are only conjunctions of disjunctions (and of or's) plus range queries on numeric values
  - Difficult to compose arbitrary, complex queries
- Have to have a responsive interaction ( $< .1$  sec) to show results – hardware and software constraints
- Need to integrate with data sources (may need to do application specification programming)
- Not clear which retinal variables to use and not enough theory concerning how much we can perceive this way
  - Which is the best way to highlight objects?
  - How to scale and allow zooming?
  - Should we use sound?
- Double ended sliders are still not very common
- Controls take space (hence advantage of Eick's sliders)

# Movable Filter - Magic Lens



Enhanced Dynamic Queries via Movable Filters  
Ken Fishkin & Maureen Stone  
CHI 1995

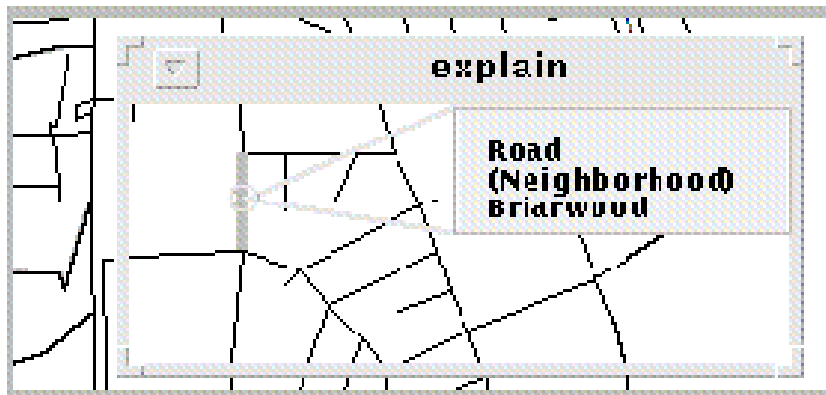
Slide adapted from John Stasko/Jock Mackinlay



# Magic Lens

- Enhanced dynamic queries via movable filters
- Extends dynamic queries as each window is a filter that can be customized by either a conjunction or disjunction operator and a slider
- Supports full boolean queries
- Arbitrarily-shaped (usually rectangular) region with some operation that changes the user's view of the data
  - Movable
  - Stackable
  - Augmented by parameters that control the display

# Magic Lens Example

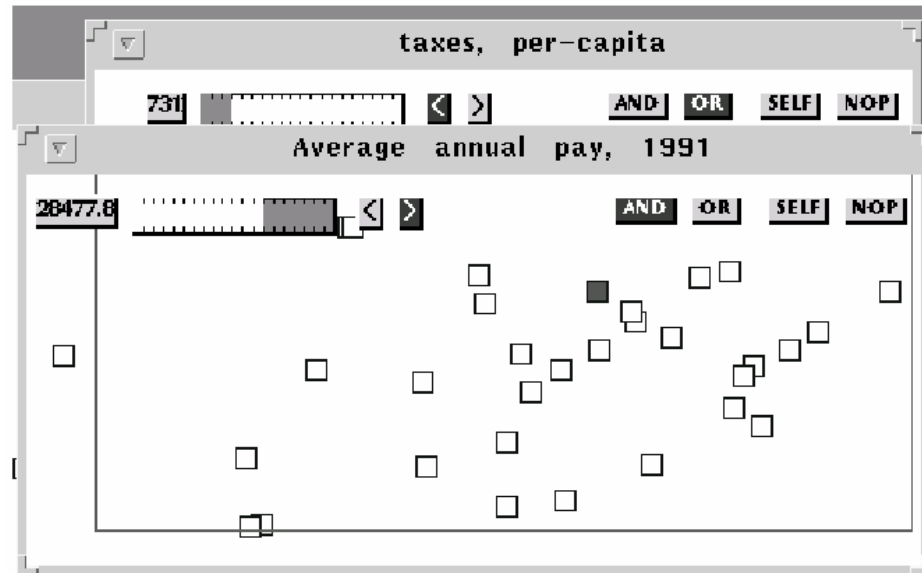


- <http://www.parc.xerox.com/istl/projects/MagicLenses/SimpleDemo.html>

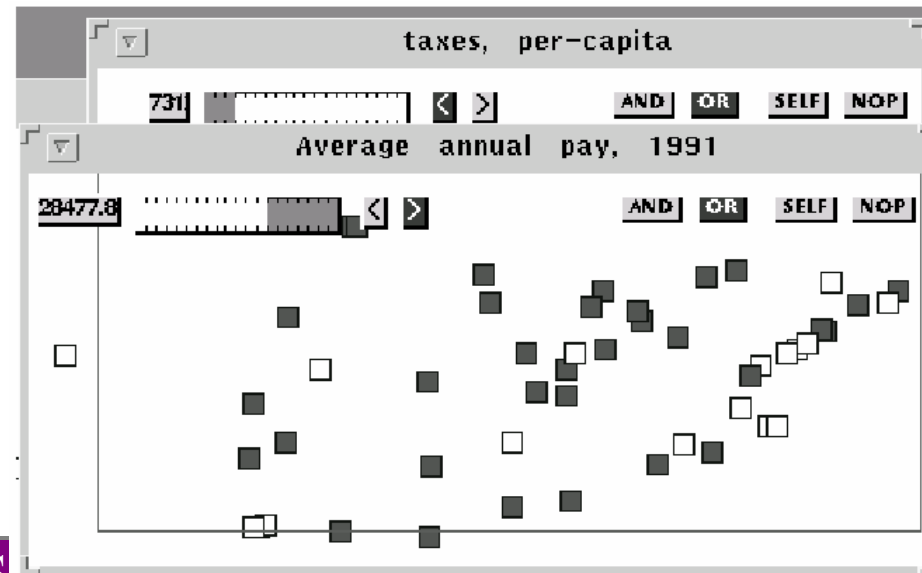
Slide adapted from John Stasko

# Queries Via Magic Lenses

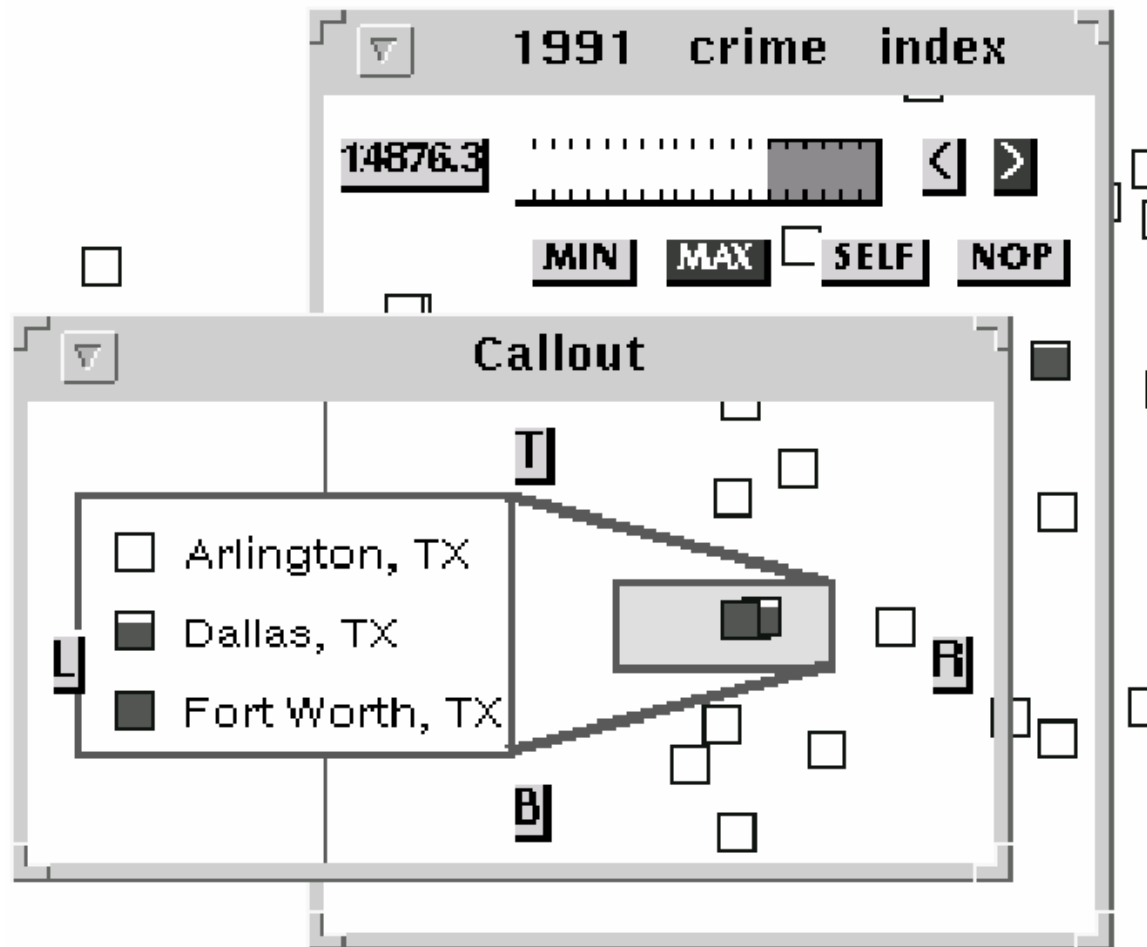
High Salaries AND  
Low Taxes



High Salaries OR  
Low Taxes



# Real-valued Query



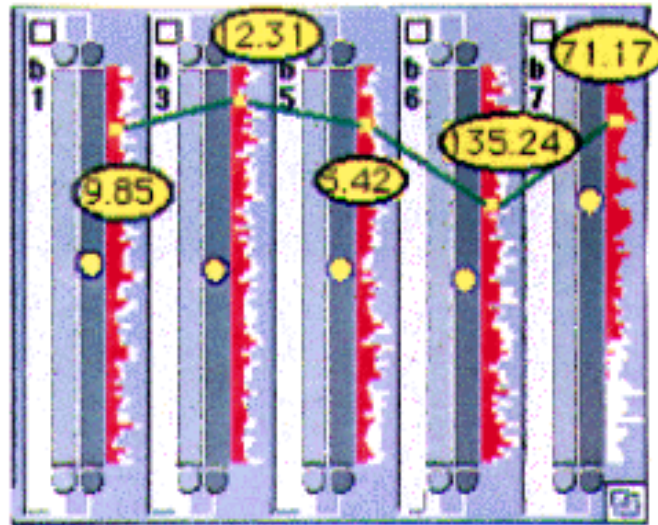
# Advantages/Disadvantages?

- Advantages:
  - Liveness of query
  - Full boolean queries
  - Can make a compound lens by abstracting a stack of lenses into one new composition lens
  - Doesn't use much real estate for controls
- Disadvantages:
  - Manipulating stacks of lenses can be awkward
  - Difficult to implement

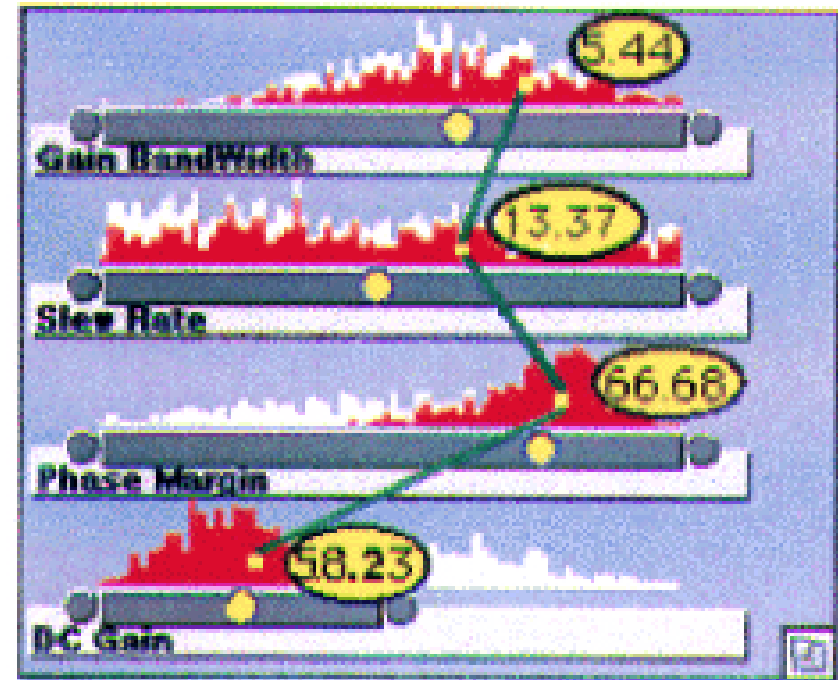
# Influence explorer

- Focus is on visualizing mathematical and engineering design problems
- The model is used to generate data about matching sets of *parameter* values and *performance* values so that this data can be responsively explored in either direction
- Provides a comparative multivariate visualization
- Like attribute explorer, except we look at and can vary independent variables and dependent variables
- Performance values have to be precalculated
- Very useful when the user has to make trade-offs

# Influence explorer



*parameters*



*performances*

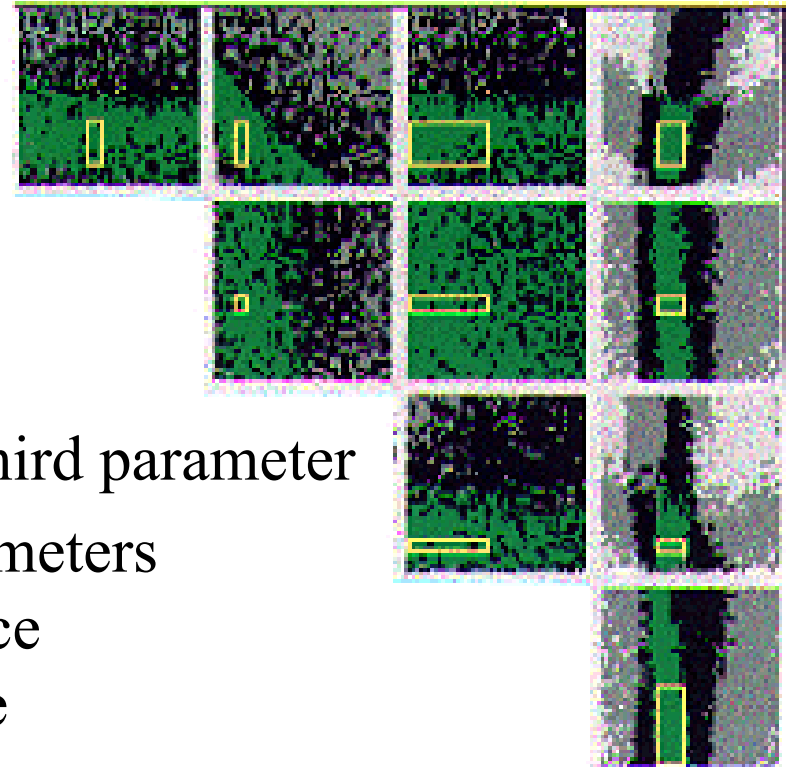
# Limitations?

- Precalculations...
- Limits on number of variables



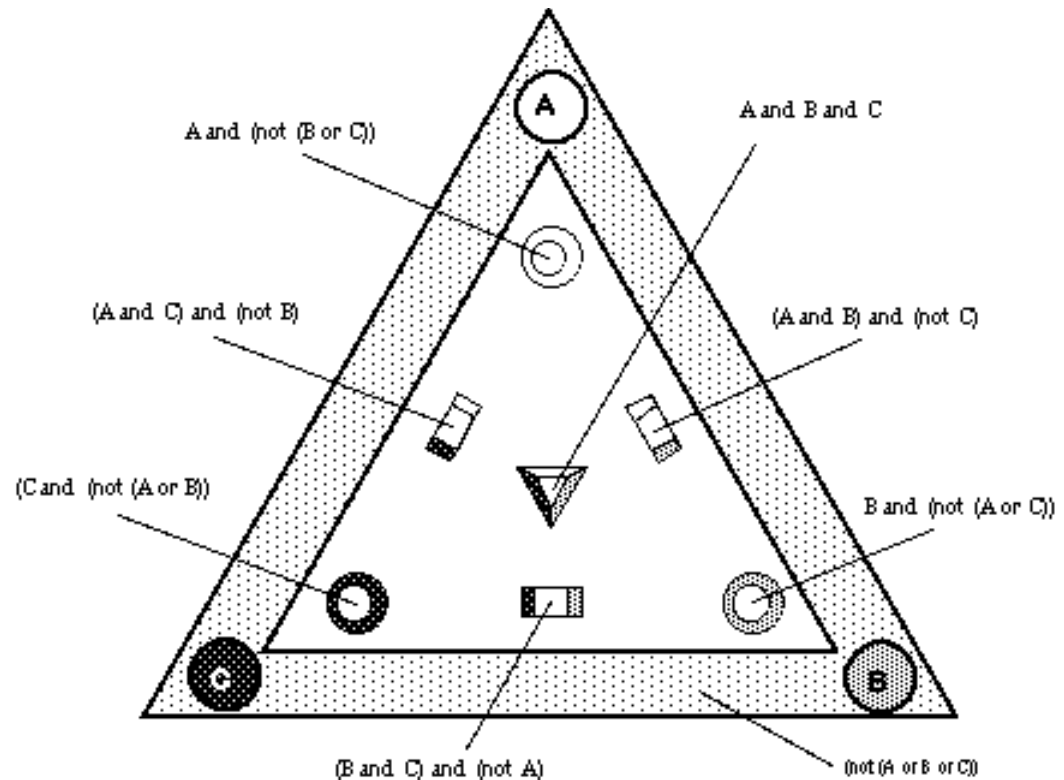
# Prosection matrix

- Prosection matrix – is an ordered collection of scatterplots
- Each scatterplot refers to a pair of parameters from a range of a third parameter
- The ranges of the remaining parameters define a section of parameter space which are projected into the plane

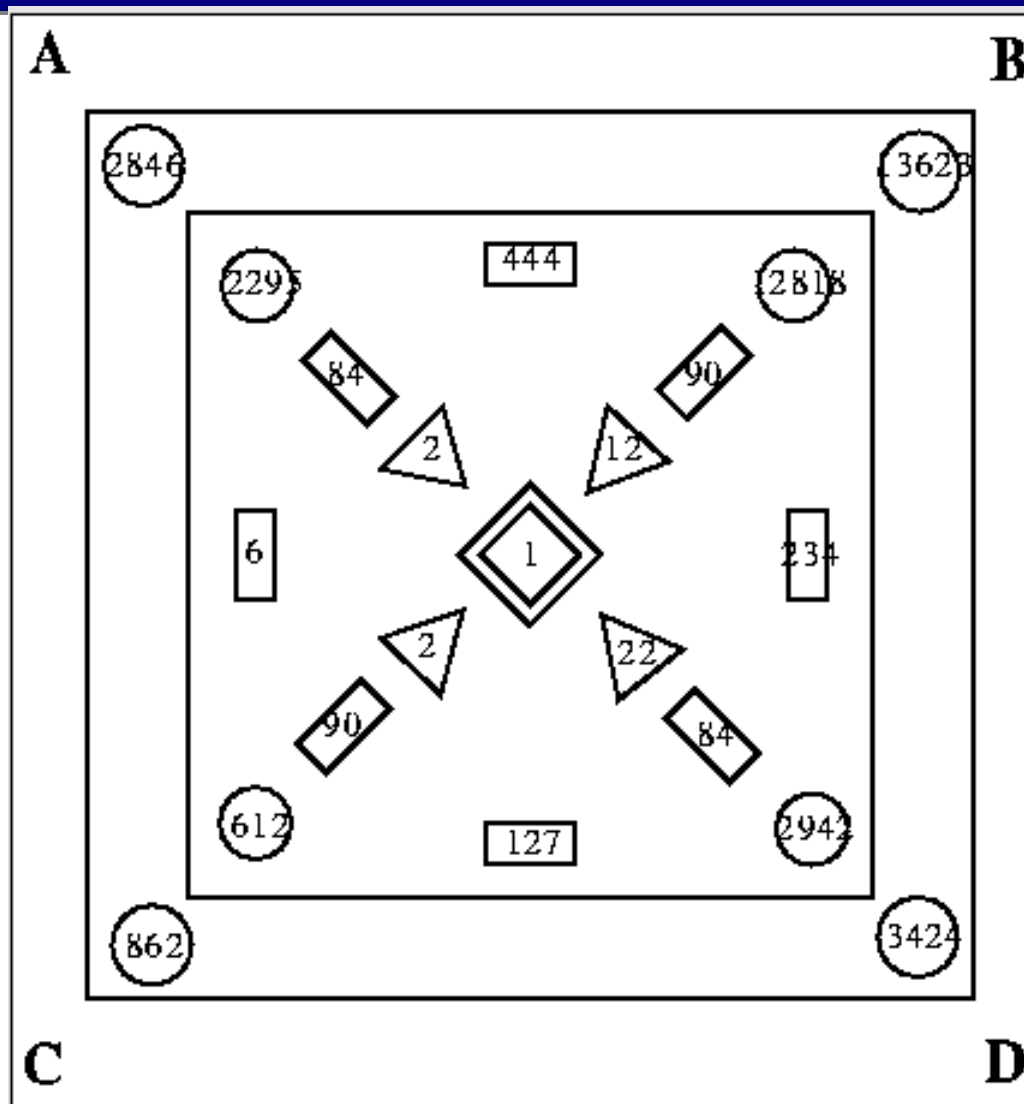


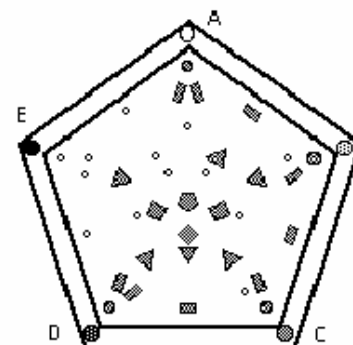
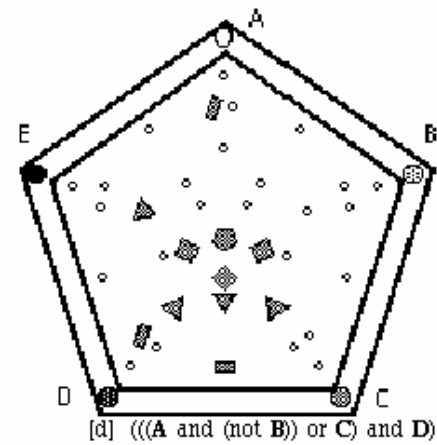
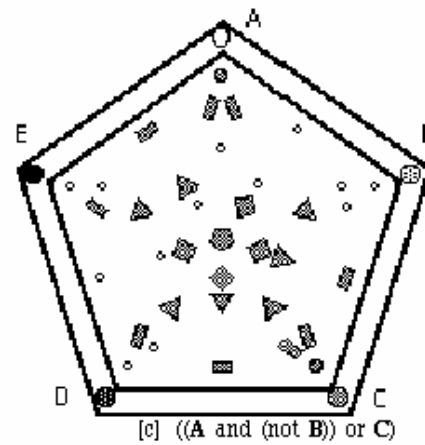
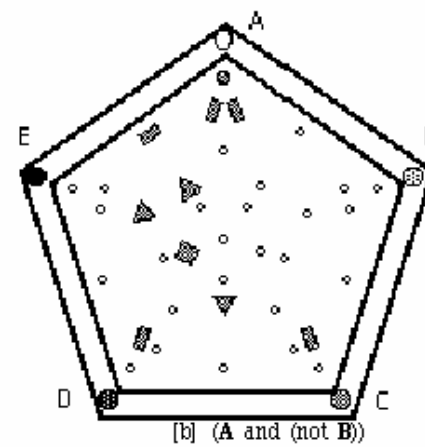
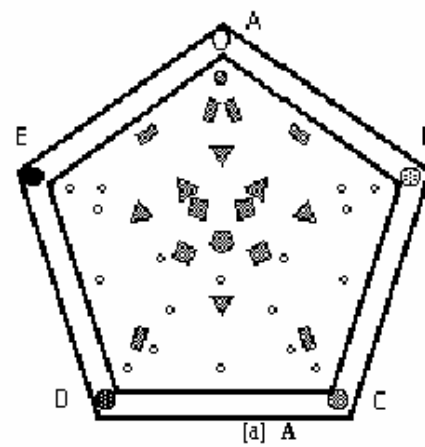
# InfoCrystal and Visualizing Boolean Queries

- The InfoCrystal acts as a *Boolean Calculator*
  - Users don't need to use logical operators or parentheses to formulate their queries



# InfoCrystal





# TileBars

- Application: Information retrieval – using text structure in full text documents (as opposed to just searching on titles and abstracts)
- Shows explicit term distribution information in Boolean-Type queries
- Indicates relative document length, query term frequency, and query term distribution
- Developed due to dissatisfaction with ranking of boolean search results – users can't see the motivation for the ranking

# Tilebars 2

- Representation of the retrieval results should present as many attributes of the texts and their relationship to the queries as possible
- But presentation needs to be coherent, compact and accurate (i.e. a true reflection of the relationship between the query and the documents)
- Current systems – users don't know what the ranking score is....

**<http://brisa.merl.com:8080/myezchooser/>**

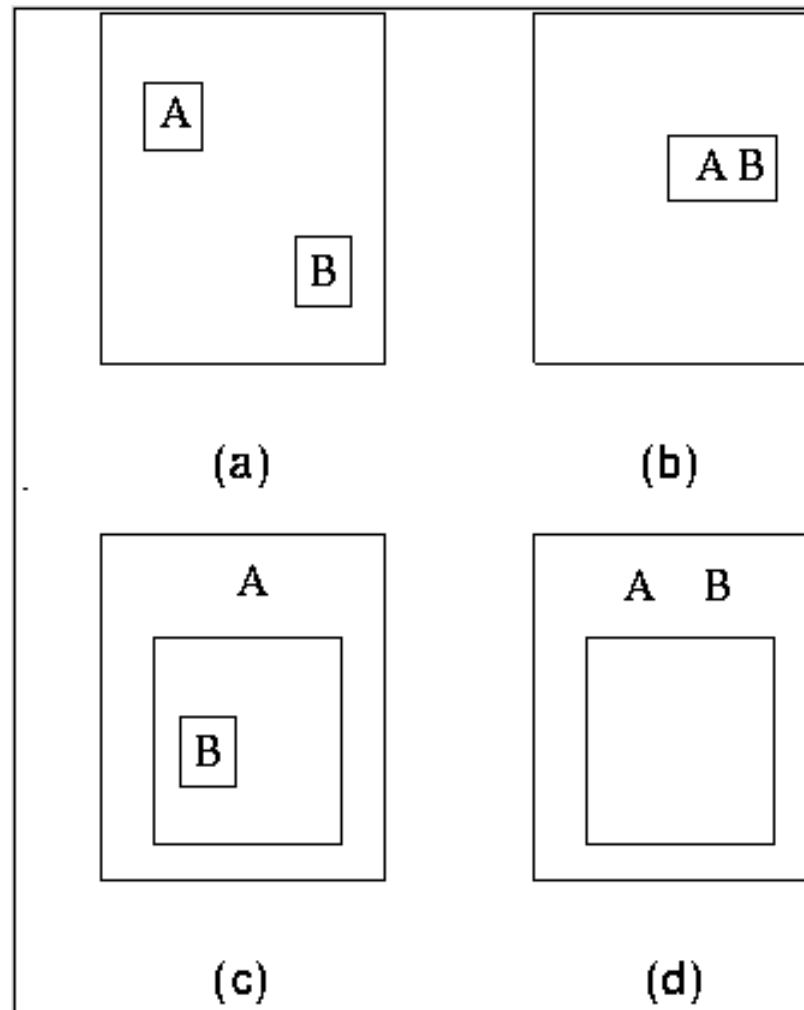
image network

This is a searchable index. Enter search keywords:

Index conf.announce contains the following 164 items relevant to 'image network'. The first figure for each entry is its relevance score, the second the number of lines in the item.

- \* 1000 1190 /ftp/pub/conf.announce/jenc5
- \* 886 125 /ftp/pub/conf.announce/image.processing.conf
- \* 800 334 /ftp/pub/conf.announce/image.analysis.symposium
- \* 743 303 /ftp/pub/conf.announce/sans-III
- \* 543 376 /ftp/pub/conf.announce/atnac.94
- \* 486 133 /ftp/pub/conf.announce/sid
- \* 486 125 /ftp/pub/conf.announce/ges2
- \* 457 138 /ftp/pub/conf.announce/euroopen.forum.94
- \* 429 378 /ftp/pub/conf.announce/mva.94
- \* 429 785 /ftp/pub/conf.announce/openview.conf
- \* 429 104 /ftp/pub/conf.announce/high.performance.networking
- \* 400 217 /ftp/pub/conf.announce/nonlinear.signal.workshop
- \* 429 378 /ftp/pub/conf.announce/vision.interface.94
- \* 429 785 /ftp/pub/conf.announce/inet.94
- \* 429 104 /ftp/pub/conf.announce/icmcs.94
- \* 400 217 /ftp/pub/conf.announce/internetworking.94
- \* 371 220 /ftp/pub/conf.announce/iss.95
- \* 371 168 /ftp/pub/conf.announce/ges1
- \* 343 152 /ftp/pub/conf.announce/conti.94
- \* 343 247 /ftp/pub/conf.announce/elvira

# Possible relationships between two terms

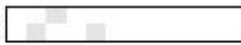
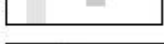
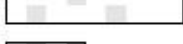
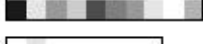
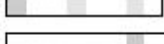
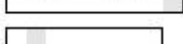
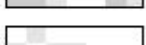
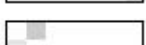




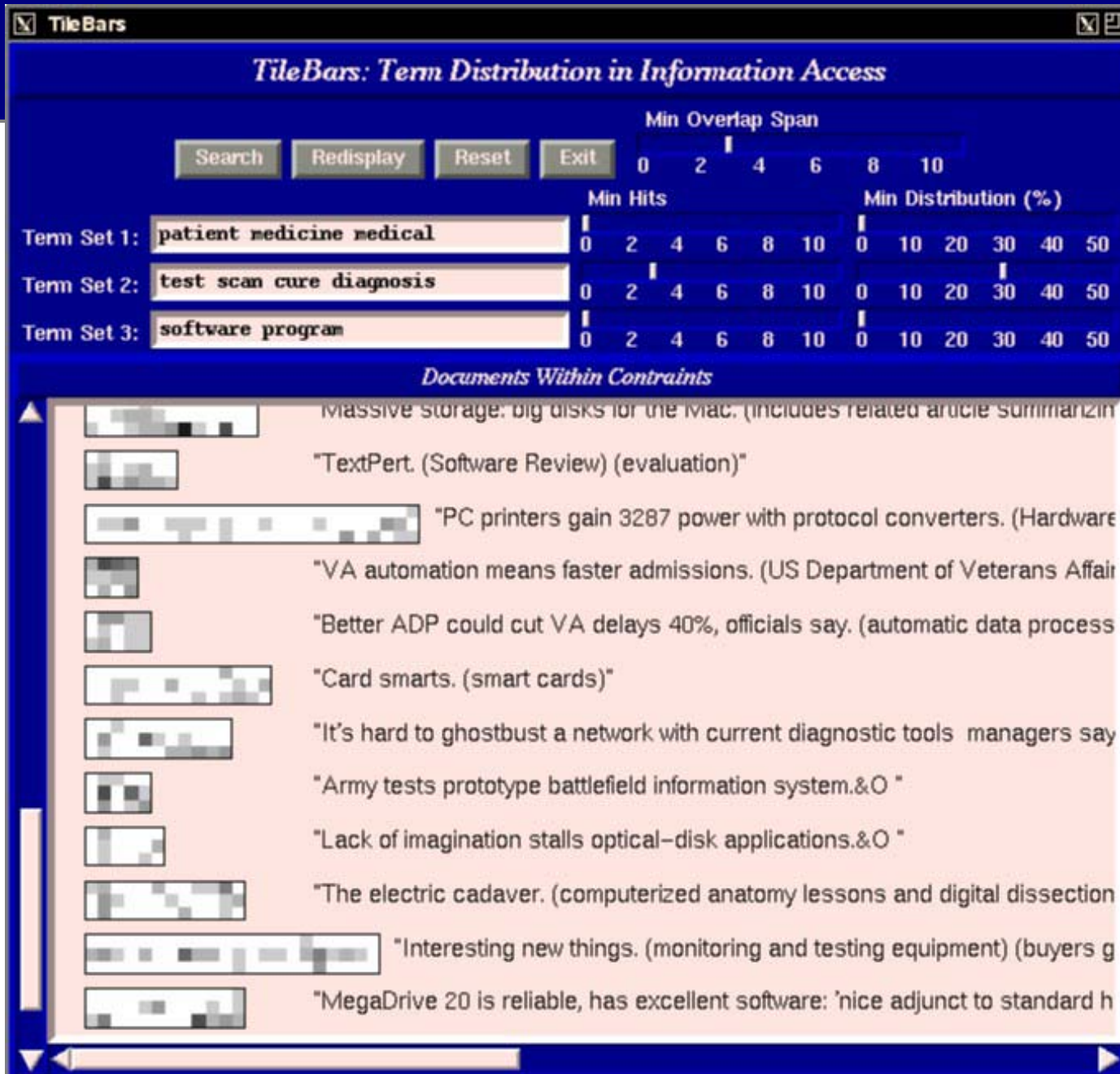
Term Set 1: law legal attorney lawsuit

Term Set 2: network lan

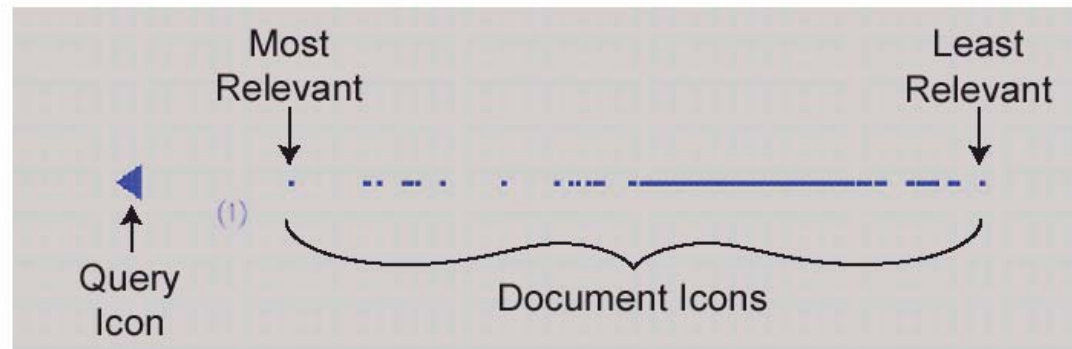
### TileBars

|      |                                                                                     |                                                       |
|------|-------------------------------------------------------------------------------------|-------------------------------------------------------|
| 1256 |    | Regression testing handling hardware and software     |
| 1269 |    | Toll fraud includes related article on MCI Commun     |
| 1270 |    | In conversation Teleglobe Canada Inc Pres and C       |
| 1280 |    | Deregulation indicates a healthy satellite services   |
| 1298 |    | The last word letters to the editor letter to the edi |
| 1300 |    | What's wrong with network licensing includes rela     |
| 1302 |    | Letters letter to the editor                          |
| 1356 |    | Protecting information now vital Law Viewpoint co     |
| 1414 |    | Letters O                                             |
| 1424 |    | Loose LIPS sink ships logical inferences per seco     |
| 1433 |    | Document management eases file control marker         |
| 1471 |    | Connectivity O                                        |
| 1496 |    | Document managers bring law and order Software        |
| 1571 |  | When users write their own applications O             |
| 1640 |  | Time lapse may scuttle Xerox claim includes relat     |
| 1690 |  | Insider revisited AI Insider column                   |
| 1758 |  | Vendors offering more remedies for file buildup de    |
| 1762 |  | Laser Lite Apple's new personal LaserWriters Har      |
| 1766 |  | Hacker's handicap Michael Colvin MP has won su        |
| 1779 |  | No summer reruns artificial intelligence application  |
| 1781 |  | How Seattle's biggest law firm put Windows 3.0 t      |

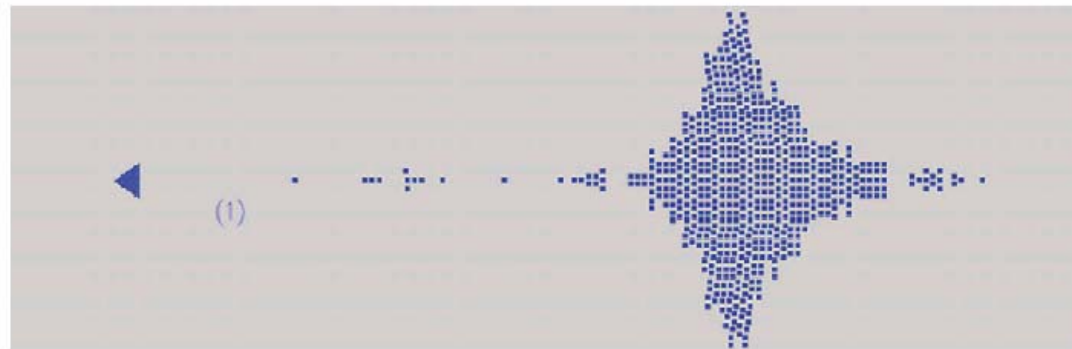
Tilebars:



# Sparkler



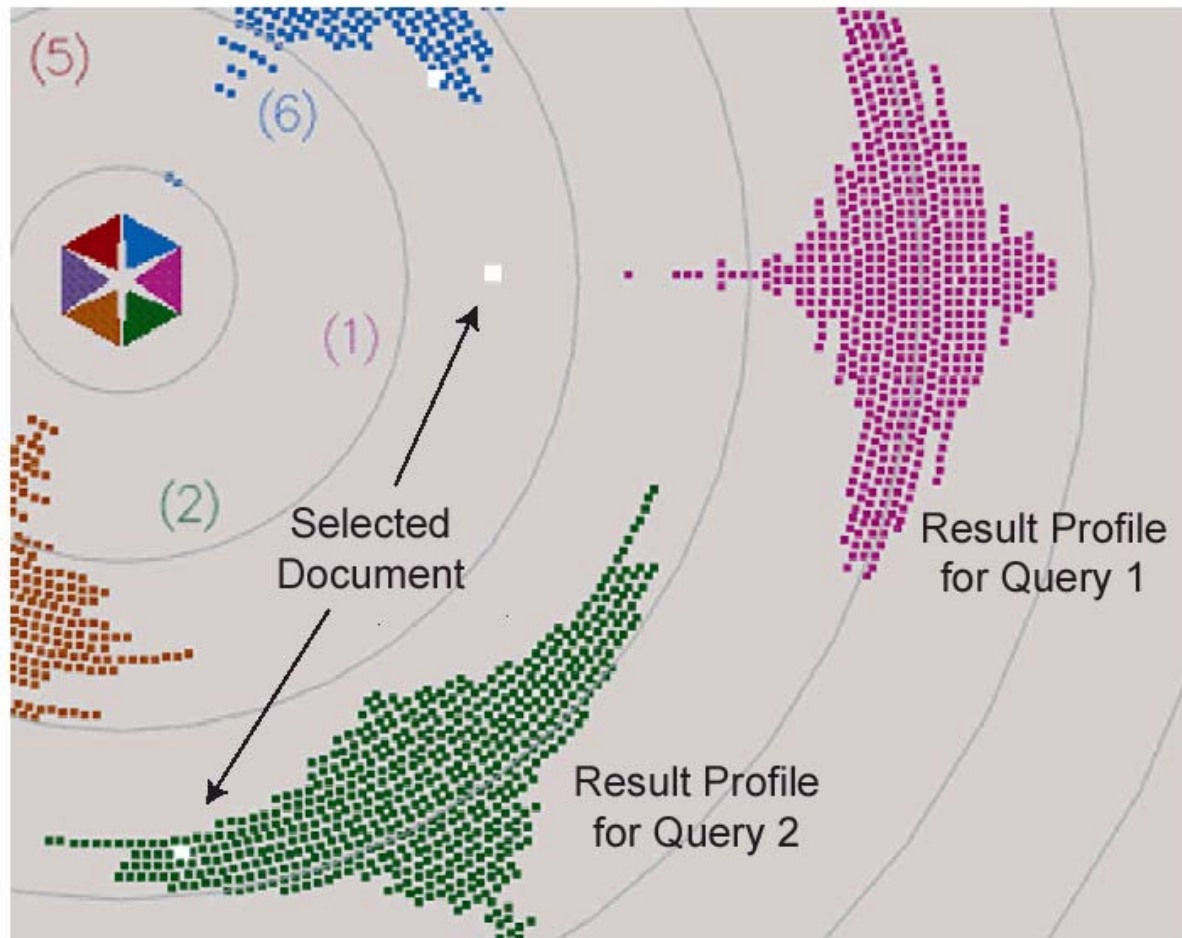
a.



b.

**Figure 1. Document icons are plotted in relation to the query icon (a) overplotted along the invisible query line and (b) spread radially**

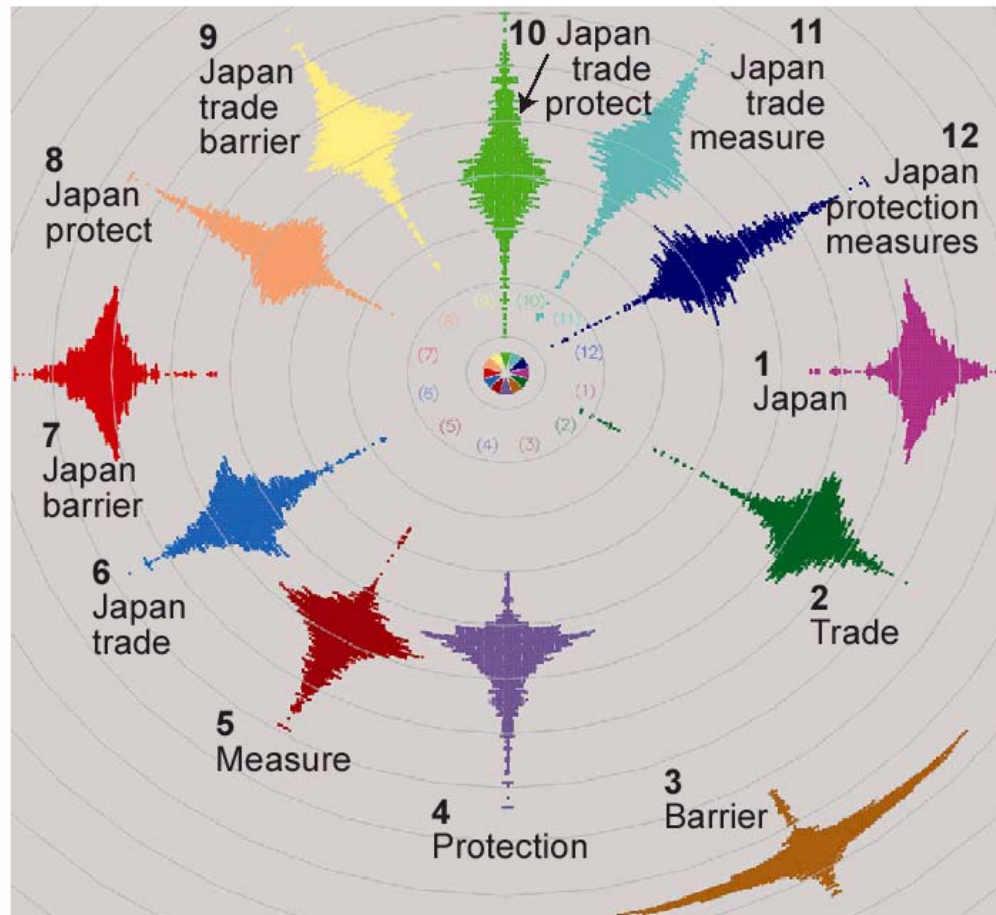
# Sparkler



**Figure 2. Zoomed-in view of two expanded result profiles and one selected document**



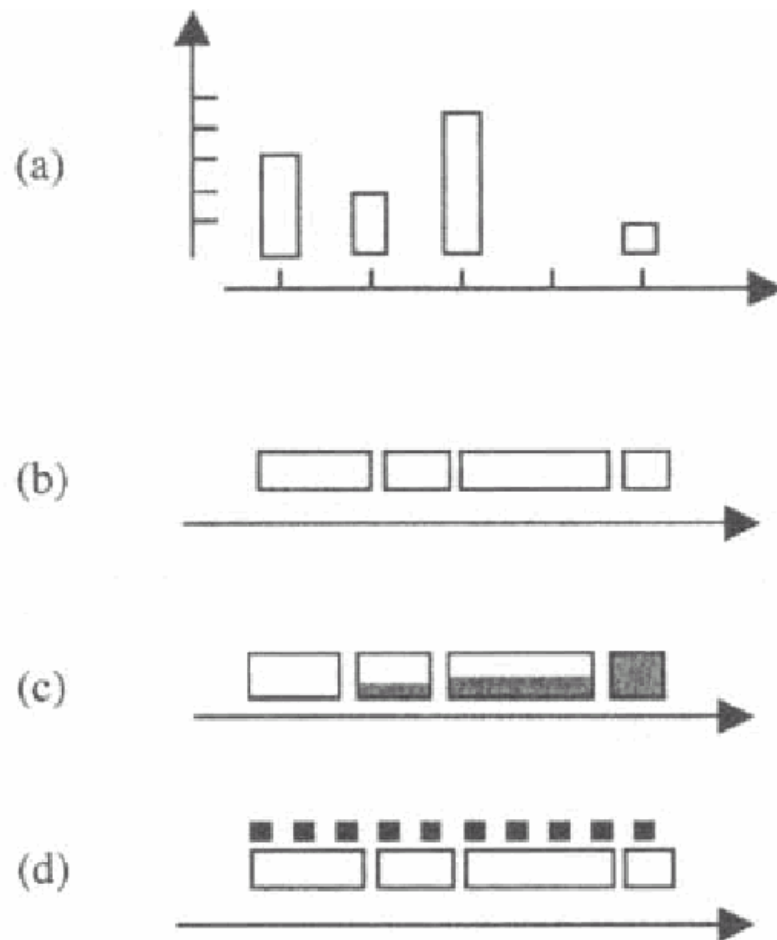
# Sparkler



**Figure 4. Queries and the respective result profiles for the AP 1990 corpus**

# Parallel Bargrams

- Bargrams are really histograms tipped on their sides and lined up end to end....
- Application: Consumer based exploration and decision making based on a set of attributes
- EZ-Chooser was preferred over static tables for about 50 items with 7-9 attributes



# EZChooser

Superpages.com Consumer Guide - EZChooser - Microsoft Internet Explorer

File Edit View Favorites Tools Help

**EZChooser**

>> Subcompact Cars

Is that car used or **U!**  
please visit our sponsor

Vehicle data provided by **CONSUMERGUIDE** Auto Guide

How do I use EZChooser?

Select feature value(s) to limit vehicles shown below. [Vehicles are sorted differently along each feature row.]

| Make                                 | Chevrolet | Daewoo      | Dod    | Ford   | Honda  | Hyundai | Kia       | Maz    | Mits   | Niss   | Pont   | Sat    | Sub    | Suz    | Toyota | Volt   |        |        |        |        |        |
|--------------------------------------|-----------|-------------|--------|--------|--------|---------|-----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| ConsumerGuide Rating                 | Best      | Recommended |        |        | Bud    |         | No Rating |        |        |        |        |        |        |        |        |        |        |        |        |        |        |
| Base Price<br>Dollar                 | 9,200     | 9,500       | 9,700  | 10,000 | 10,700 | 10,800  | 11,000    | 11,500 | 11,800 | 11,900 | 12,000 | 12,400 | 12,500 | 13,200 | 13,900 | 14,000 | 14,900 | 15,500 | 19,900 |        |        |
| Loaded Price<br>Dollar               | 10,000    | 10,200      | 10,400 | 11,000 | 12,000 | 12,500  | 12,900    | 13,200 | 15,000 | 15,100 | 15,400 | 15,700 | 15,900 | 16,000 | 16,400 | 16,900 | 17,600 | 19,300 | 20,100 | 21,700 | 24,300 |
| City Fuel Efficiency<br>Miles/Gallon | 22        | 23          | 24     | 25     | 27     | 28      | 29        | 30     | 31     | 33     | 34     | 35     | 39     | 41     | 42     | 61     |        |        |        |        |        |

Vehicles matching your feature selections. (Color a vehicle in all rows by selecting a color over the picture.)










View more information about Chevrolet Cavalier











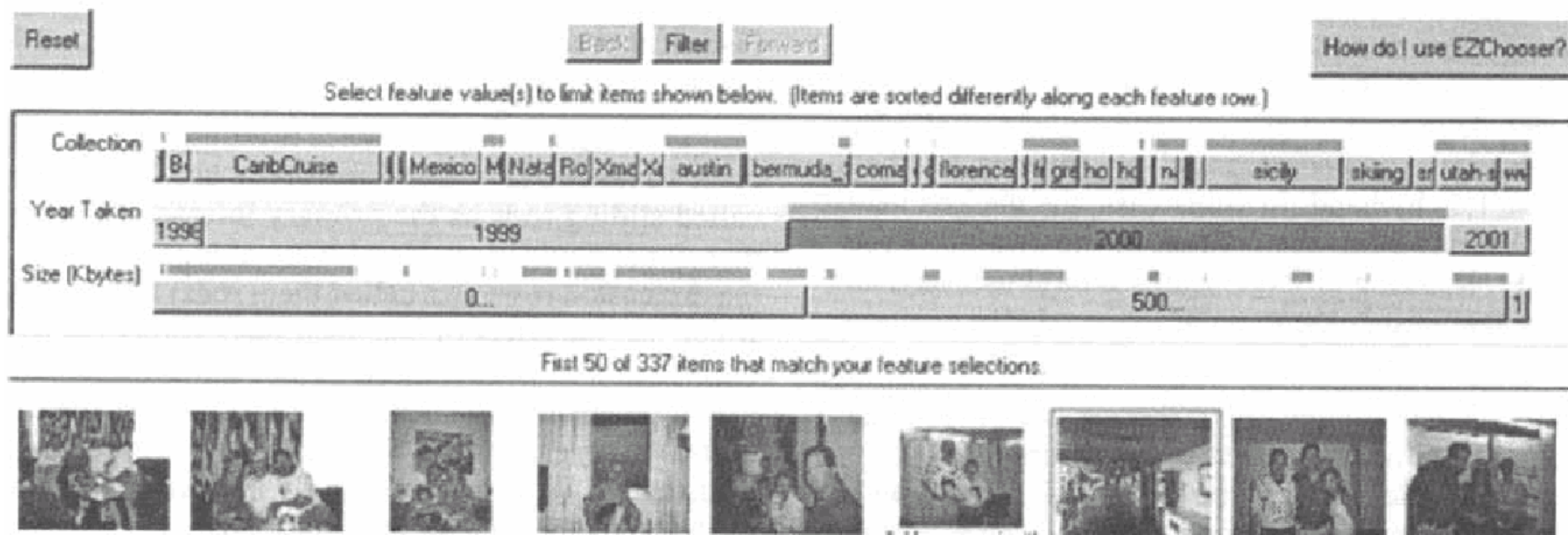

# Interacting with EZChooser



Vehicles matching your feature selections. (Color a vehicle in all rows by selecting a color over the picture.)



# EZChooser -- Photos



**Figure 4:** An example of EZChooser operating at a larger scale. The dataset consists of 707 photos of which 337 have been selected via the attribute restrictions. The first 50 of those are shown as thumbnail images. Items within vector rows are drawn at 1-pixel width.

# Digital image management

- Photomesa
  - PhotoFinder (video)
- ... More on Thursday....

# Additional references

- Parallel bargrams for consumer-based information exploration and choice, Wittenburg et al.:  
<http://portal.acm.org/citation.cfm?doid=502348.502357>
- TileBars: Visualization of Term Distribution Information in Full Text Information Access, Marti Hearst, CHI 95:  
<http://www.sims.berkeley.edu/~hearst/papers/tilebars-chi95/chi95.html>
- Interactive Visualization of Multiple Query Results, Havre et al.,  
<http://www.pnl.gov/infoviz/sparklerinfovis01.pdf>
- InfoCrystal:  
[http://www.scils.rutgers.edu/~aspoerri/InfoCrystal/Ch\\_4.html](http://www.scils.rutgers.edu/~aspoerri/InfoCrystal/Ch_4.html)
- GRIDL (Graphical Interface for Digital Libraries)  
<http://www.cs.umd.edu/hcil/west-legal/gridl/>
- Lifelines (medical data)  
<http://www.cs.umd.edu/hcil/lifelines/latestdemo/kaiser.html>