Information Visualization

- Related to but, in many ways, distinct from interaction design is the area of *information visualization* — the study of how data can be presented for maximum comprehension and clarity (*cognition*)
- Information visualization focuses on the "computer-to-human" direction of interaction design
- In Don Norman's language, information visualization seeks to ensure that the *system image* communicates the designer or developer's *mental model* to the user as clearly and accurately as possible

- Aside from read-only displays (including non-electronic media), information visualization also dovetails nicely with the direct manipulation interaction style, due to its reliance on the *visibility* of its components
  - It should be of no surprise, then, to learn that Shneiderman has contributed work to the field
- Other key contributors to the field include:
  - Stuart K. Card and Jock D. Mackinlay: Information visualization in electronic, interactive media
  - Edward Tufte: Information visualization in printed, high-resolution media
  - Colin Ware: Advanced data visualization, with special interest in applications to ocean mapping

Common Information Visualization Themes

- More than just charts and graphs or 3D displays, information visualization pays special attention to:
  - Presentation of large numbers of items (10^2 to 10^6 or more), sometimes drawn from even larger data sets
  - Maximum use of human perceptual abilities
  - Integration of data from potentially disparate (and initially incompatible) sources
- In the end, we still require measurable improvements in information comprehension

The Visual-Information-Seeking-Mantra

- This mantra is viewed as an overall principle behind information visualization as a whole:

  Overview first, zoom and filter, then details on demand

  Overview first, zoom and filter, then details on demand

  Overview first, zoom and filter, then details on demand

  Overview first, zoom and filter, then details on demand

- We copy the Shneiderman text's approach of literally repeating this mantra, reflecting both this principle's near-universal applicability and its recursive nature

Seven Types, Seven Tasks

- Shneiderman breaks down the overall area of information visualization into a *data type by task taxonomy*, consisting of seven basic data types and seven basic tasks:
  - Data types include *1D linear*, *2D map*, *3D world*, *multidimensional*, *temporal*, *tree*, and *network*
  - Tasks include *overview*, *zoom*, *filter*, *details-on-demand*, *relate*, *history*, and *extract*
- Various type/task combinations lend themselves to specific applications (geography, medicine, etc.)

Edward Tufte's Approach

- Edward Tufte's "trilogy" of information visualization books follow a different approach:
  - *The Visual Display of Quantitative Information* focuses on the presentation of *numeric* data
  - *Envisioning Information* focuses on the presentation of objects or thing — *nouns*
  - *Visual Explanations* focuses on the presentation of phenomena or processes — *verbs*
- In relation to interaction design and direct manipulation, the focus on *nouns* (i.e., *Envisioning Information*) appears to be the best fit, though the overall sequence contains valuable information and insight throughout

Data Types

Returning to Shneiderman's *data type by task taxonomy*, the seven basic types for information visualization are:

- *1D linear data*: Sequences of some other data element (characters, numbers, documents, words, etc.)
- *2D map data*: Data with a genuine sense of "area," such as maps, floor plans, and layouts
- *3D world data*: Data that has "volume," typically real-world objects and structures; the third dimension adds new challenges regarding navigation and rendering
- *Multidimensional data*: Data with more than 3 dimensions, such as relational or statistical databases where attributes can be thought of as axes in an *n*-dimensional space; typically uses multiple "layers" of 2D or 3D views
- *Temporal data*: A special case of 1D linear data, where the sequence of elements is based on time; having *start* and *end times* influences the visualization approach
- *Tree data*: Data with a hierarchical, recursive structure, with elements (except the root) having exactly one parent; visualization approaches include outlines, 3D views, and space-filling *treemaps*
- *Network data*: Any graph-structured data, presenting challenges in terms of layout and traversal; going domain-specific may help refine visualization choices

Tasks

An alternate, parallel perspective on information visualization comes from looking at the basic *tasks* involved

- *Overview*: This task seeks to extract the "big picture" out of an overall data set — a "macro" task, in Tufte's terms
- *Zoom*: This task looks "closer" at some data subset, typically after the overview task has determined specific areas of interest — for Tufte, this is the "micro" view
- *Filter*: This task reduces the current amount of visible data, typically based on a query or predicate

- *Details-on-demand*: Individual data items may have additional (extensive) properties or content; the user/viewer will want easy access to this upon request
- *Relate*: A task that truly leverages visual displays, relating items or groups of items have a broad range of variations and choices, including proximity, containment, connectors, colors, and highlighting
- *History*: An overall information visualization "session" typically involves a sequence of task instances; users may want to review this sequence, allowing backtracking or a review of what has been done so far
- *Extract*: Once a user has found a suitable subset (and presentation), he or she may wish to share or send it, in a separate, distinct form or package

Focus on *Envisioning Information*: Displaying Nouns

- It has been mentioned that the direct manipulation interaction style has a *noun-verb structure* — the *nouns* within a system are displayed first, which a user then *selects* prior to issuing commands or actions (the *verbs*)
- In this light, techniques for presenting nouns are particularly relevant — and Edward Tufte's *Envisioning Information* provides specific insight on this
- Tufte's work tends to focus on high-resolution displays (i.e., paper), so may need some adjustment when adapting to existing digital displays

Key issues in *Envisioning Information* include:

- *"Spaceland" vs. "flatland"* — Our world and ideas exist in multidimensional space (including time), yet presentation media are two-dimensional; "escaping from flatland" serves as the overall theme within *Envisioning Information*
- *Combined micro/macro views* — "To clarify, add detail:" fine granularity presented with clear, overall structure combines overviews with easy zooming or details
- *Layering and separation* — Human perception is inherently *multimodal*, with colors and the "1 + 1 = 3" design principle giving flat displays different levels
- *Small multiples* — Side-by-side variations of related concepts help clarify an overall picture, particularly for comparison purposes

Information Visualization Challenges

Information visualization remains a wide-open field with a number of existing challenges:

- *Importing data* — Marshalling data into an accessible form is harder than many people think; there has been some convergence on XML, but it is no panacea
- *Visual vs. textual* — Visual representations don't obviate the need for labels; the idea is to bring them together in a way that plays to their respective strengths
- *Related data* — Users/viewers may want to jump across data sets…if those sets are available in a compatible form
- *Large volumes* — Data are now available on the order of millions or billions, presenting challenges of scale in terms of technology and approach; note that, given the right presentation, human perception *can* still gain a decent understanding of such large data sets
- *Data mining* — Automatically finding patterns or trends was once viewed as a distinct area, but is now seen as complementary to information visualization, since the right display can highlight such patterns and trends
- *Collaboration* — We frequently need to perceive data sets in groups or teams, so that element needs to factor in
- *Universal usability* — Visualization approaches might lock out certain users, such as the visually impaired or the colorblind, or carry cultural or economic assumptions