User Interfaces and Visualization

Lecture 17
Overview

- Interface design principles
- Models of interaction
- Evaluating interactive systems
- Support at different search stages
  - Starting points
  - Query formulation
  - Giving context
  - Using relevance judgments
Information Interfaces

- Users’ information needs are imprecise
- Users don’t have a clear idea of how to achieve their goals
- The human-computer interface should help users understand and express what they want
  - formulating queries
  - selecting information sources
  - understanding results
  - keeping track of their search process
Design Principles

- Offer informative feedback
- Support an internal locus of control
- Permit easy reversal of actions
- Reduce working memory load
- Provide alternative interfaces for novice and expert users

(Shneiderman 97)
Offer Informative Feedback

- Search system should give feedback on...
  - relationship between query and results
  - relationships among retrieved documents
  - between documents and collection metadata
- User should be able to adjust the level of feedback
Internal Locus of Control

- Users want to know that they are in control
- System should respond to user actions
- Users should be initiators rather than responders
- AVOID
  - surprising system actions
  - tedious inescapable situations
  - inability to produce action
- Example: modal vs. non-modal interfaces
Easy Reversal of Actions

- Any action should be reversible
- Ability to “undo” relieves user anxiety
- Encourages user exploration
- Consider the unit of undo
  - single action, data entry, or block of actions
- For example, “Back” button in a web browser
  - But once you go forward again, the stack is lost
  - Sometimes the user can get lost in hyperspace
Reduce Working Memory Load

- Don’t overload the user’s short-term memory
  - people can remember “seven +/- two chunks“ of information
- Help with keeping track of search choices
  - let users switch between search strategies
  - retain context and information across sessions
- Provide browsable contextual information
  - suggestions of related search terms/metadata
  - search starting points with topic descriptions
Alternative Interfaces for Experts and Novices

- Tradeoff between simplicity vs. power
  - MacOS Finder vs. Unix shell
  - Simple: easier to learn, but less flexible
  - Powerful: allow experts to do more, faster

- Scaffolding
  - For novices: simple, easy-to-learn interface with basic functionality
  - Experts can go inside the scaffolding and have more control, features, options
Classical Model of Info Seeking

1. Information need
2. Query
3. Send to System
4. Receive Results
5. Evaluate Results

Reformulate

Done?

No

Yes

Finish
“Berry-Picking” Model

- Users learn as they search
  - causes need to change
  - causes queries to shift around, not refine
  - one goal leads to another
- Information needs not satisfied by a single set of documents
  - really by bits and pieces found along the way
Information-seeking Activities

- Scanning
  - high-level skimming
  - for selecting something to view or as input to query
- Querying
  - produces new, ad-hoc, unorganized collections
- Navigating
  - following a chain of links/views towards a goal
  - a sequence of scan and select operations
- Browsing
  - casual, undirected exploration
A sketch of a searcher… “moving through many actions towards a general goal of satisfactory completion of research related to an information need.” (after Bates 89)
Evaluating Interactive Systems

- Precision and recall measure search results
- Not appropriate for interactive systems
  - interactive users require a few relevant documents
  - usually don’t care about high recall
- Metrics include
  - time to learn the system
  - time required to achieve goals on benchmark tasks
  - error rates (user, not system)
  - retention of use of the interface over time
1. Starting Points

- How do users begin a search?
  - not with long, detailed information needs
  - usually a short query, followed by scanning the results and modifying the query
  - “testing the water”
    - get familiar with collection, query language, etc
- System should guide the user to the right starting point.
Lists of Collections

- Traditional bibliographic systems began with choosing from a list of collections
  - e.g. VICTOR (TTY interface)
  - user must learn which are the right collections
- On the web, a “portal” might offer a list of search engines
- Need *overview* information
Directory Overviews

- Also called *category* overviews
- Provide a hierarchical structure for the collection
- Very popular on the web
- Possible to get lost browsing a directory

D.4 OPERATING SYSTEMS (C)
- D.4.0 General
- D.4.2 Storage Management
  - Allocation/deallocation strategies
  - Distributed memories
  - Garbage collection (NEW)
  - Main memory
  - Secondary storage
  - Segmentation**
  - Storage hierarchies
  - Swapping**
  - Virtual memory
MeSHBrowse example
Automatically Creating Collection Overviews

- **Idea:** extract most common themes occurring in the collection
- **Clustering**
  - organize documents by similarity to one another
  - centroids of cluster = themes in collection
  - an *unsupervised* learning method
Scatter/Gather

• A browsing paradigm
  • Clusters documents into topical groups
  • Each cluster has a textual summary
    • topical terms + sample titles of documents

• Using Scatter/Gather
  • Cluster the entire collection
  • User selects a few clusters
  • Documents in selected clusters are then re-clustered

• Can also be used with search results
Scatter/Gather

Cluster 1 Size: 8
- Key, Francis Scott
- Fort McHenry
- Arnold, Henry Harley
- "Star-Spangled Banner, The"
- Word
- Poem
- British

Cluster 2 Size: 68
- Burstyn, Ellen
- Stanwyck, Barbara
- Berle, Milton
- Zukor, Adolph
- "Roadhouse, Talkie"
- "Film"
- "Play"
- "Career"
- "Win"
- "Television"
- "Role"
- "Record"
- "Award"
- "York"
- "Popular"
- "Stage"

Cluster 3 Size: 97
- Star
- Galaxy, The
- Extragalactic systems
- Interstellar matter
- Cluster
- Star

Cluster 4 Size: 67
- Astronomer
- Observatory
- Astronomy
- Position
- Measure
- Celestial
- Telescope
- "Astronomy and astrophysics"
- "Astrometry"
- "Agena"
- "Astronomical catalogs and atlases"
- "Hansel, C.W. Williams"

Cluster 5 Size: 10
- "Family"
- "Species"
- "Flower"
- "Animal"
- "Arm"
- "Plant"
- "Shape"
- "Leaf"
- "Brittle"
- "Tube"
- "Foot"
- "Horde"
- "Blazing star"
- "Brittle star"
- "Bishop's-cap"
- "Feather star"
Collection Visualization

- Scatter/Gather is a textual representation
- Visualization techniques can show the clusters in the collection graphically
  - ThemeScape: “mountains” of documents
  - Kohonen maps
- Provide a high-level visual summary of the collection
Kohonen Maps experiment

- Compared Kohonen map view to Yahoo! hierarchy
  - Task: find an interesting document using both views
  - Some started with Yahoo!, some with map
- Most users could find a document on Yahoo! and then on the map
- …but not vice-versa
- Maps appear good for
  - high-level view, zooming/panning
  - jumping around in hierarchy
2. Query Specification

- Five visual interaction styles
  - command language
  - form fill-in
  - menu selection
  - direct manipulation
  - natural language
- All have been used in query formulation interfaces
- Command languages and form fill-in are most common on the web
Boolean Queries

- Until recently, most common query language
- Users find it very difficult to use and exploit
  - “and” implies wider scope
  - “or” implies an exclusive choice
  - connector syntax, metadata

Solutions on the Web
- “all the words”, “any of the words”
- + operator (but can mislead: “cats + dogs”)
- Forms allowing two or three clauses
Melvyl (UC) Boolean Form

Database: Current Contents  
Personal Profile: Off

Author Search: Current Contents database

Author: swanson, d.  
(e.g., jones, e d)

Options and Limits

Another Author: and  
(e.g., wilson, r)

Journal Title: and  
(e.g., daedalus or jama)

- Any words  - Exact beginning  - Complete title

Location: and  
UC Berkeley (all libraries)

Submit Search  Back
Direct Manipulation

- Direct manipulation interfaces feature
  - Continuous representation of objects of interest
  - Physical actions or button presses
  - Rapid, incremental, reversible operations
  - Immediate feedback
- Uses for Boolean query specification
  - Venn diagrams for illustrating sets
  - Block diagrams for organizing query terms
Search for any documents in "HCI Bibliography" containing either Query and Boolean; or Graphical, Searching and Browsing; but not Ranking.

4 documents match the selected query:

- Graphical Presentation of Boolean Expressions in a R. Richaud
- Query Processing in a Heterogeneous Retrieval Meta Patricia Simpson
- On Extending the Vector Space Model for Boolean Ou S. K. N. Wong, W. Zlatko, U. V. Faghi, P. C. M. Wong
- A Direct Manipulation Interface for Boolean Inform Peter G. Anick, Jeffrey D. Brennan, Rex A. Flynn, David
Block-diagram Queries (Anick)

- Blocks in a row are ANDed
- Blocks in a column are ORed
- Each block can be activated or deactivated
- Users can quickly experiment with different query formulations
3. Providing Context

- The context of a document (set) includes
  - relationship to the query terms
  - relationship to other documents within the set
  - relationship to collection as a whole
  - metadata (dates, authors, subjects...)
  - hyperlink structure

- Providing this feedback helps the user understand the behavior of the system in response to their query
Query term context

- Document surrogates
  - title, date, source, document length
  - similarity score or degree of match
- Term highlights in an abstract
- Key Word In Context (KWIC)
- TileBars
  - Number of term hits per document passage
  - Query facets displayed as stacked bars
FR88513–0157
AP: Groups Seek $1 Billion a Year for Aging Research

SJMN: WOMEN’S HEALTH LEGISLATION PROPOSED CHANGES
AP: Older Athletes Run For Science

FR: Committee Meetings
FR: October Advisory Committees; Meetings

FR88120–0046
FR: Chronic Disease Burden and Prevention Models; Program A
AP: Survey Says Experts Split on Diversion of Funds for AIDS
FR: Consolidated Delegations of Authority for Policy Development

SJMN: RESEARCH FOR BREAST CANCER IS STUCK IN Programs
Cha-Cha

- Intranet web search tool
  - uses hyperlinks to organize search results
  - finds shortest paths from the root page to each search hit
  - results shown as paths in a hierarchy
- Shows context of results within collection
  - can see how results are part of the web
  - lists information sources (1st-level pages)
  - “Virtual” table of contents
- Demo: http://cha-cha.berkeley.edu/
4. Relevance Feedback

- Previously, we focused on algorithmic level
- But also need to consider…
  - how do we get feedback from the user?
  - will the user understand the effect on the query?
- Users want to understand and control search
  - Selecting feedback documents and terms
  - Seeing how feedback queries are constructed
  - Know what is being retrieved or rejected, and why
Study of Relevance Feedback

- Koenemann and Belkin, 1996
- Four types of feedback interaction
  - **Control**: No relevance feedback, only manual reformulation
  - **Opaque**: Selecting relevant documents only
  - **Transparent**: Show expansion terms
  - **Penetrable**: stop midway, showing terms for expansion and query reformulation. Subjects select terms for reformulation.
- Subjects more effective with feedback
- Penetrable feedback most effective
GM Plans to Recall 62,000 1988-89 Cars With Quad 4 Engines

DETROIT -- General Motors Corp. said it is recalling 62,000 1988-89 model cars equipped with its high-tech Quad 4 engine to fix defective fuel lines linked to 24 engine fires.

GM said the 1988-89 Pontiac Grand Am, Oldsmobile Cutlass Calais and Buick Skylark cars equipped with the 16-valve, four-cylinder Quad 4 engine have fuel lines that could crack or separate from the engines. Although GM has received reports of 24 fires caused by leaks attributable to the faulty fuel lines, a spokesman says the company knows of no injuries resulting from the incidents. GM sold about 312,000 cars equipped with Quad 4 engines in the 1988-89 model years.

In another action, GM said it is recalling about 3,200 of its 1990 Oldsmobile Cutlass Calais and Buick Skylark models to fix fuel-line defects on three engines: the Quad 4, 3.3-liter V-6, and 2.5-liter four-cylinder. GM isn’t aware of any fires or injuries related to the fuel-line problems in this group of cars, the spokesman said.

All repairs will be done free of charge to owners, the company said.

Separately, the U.S. sales arm of Volkswagen AG’s Audi subsidiary said it is recalling 1,600 1990-model Audi 80, 90 and Coupe Quattro luxury cars to replace a defective bolt in the assembly that locks the steering when the car is parked. The defective bolt could break, causing the steering wheel to come in contact with the body-shell interior, the company said.
Organizing the Search Process

- Berry-picking model says search is not linear
  - users jump between different strategies
  - and try different tactics/operations
- Search “product” is a collection of strategies and their results
- Organizing these threads is hard
  - Bookmarks: location, no context
  - “Back” list: single path from root page
- Should unify search, browsing, navigation, and selection
Summary

- Information systems need good interfaces
  - Tools can be complex to use
  - Underlying models are hard to understand
  - There is too much information to keep track of
- The best interfaces will
  - employ good design principles
  - integrate all information seeking behaviors
  - support long-term search strategies
  - yield happy users