

Reading

Chapter 10 of BYRN

This chapter is available on the web

at <http://www.sims.berkeley.edu/hearst/irbook/10/chap10.html>

There will also a link from the course homepage.

<http://wotan.liu.edu/krichel/itr8>

<http://openlib.org/home/krichel/itr8>

Structure of lecture

1. introduce topic and dismiss history
2. look at models of user behavior
3. present Shneiderman's recommendations
4. overview the other issues that Hearst discusses

The problem

When users approach an information retrieval (IR) system, they have no precise idea on how to achieve their information extraction goal.

A well-designed interface will help the users to achieve their goal. Humans are more difficult to understand than computers. Therefore this area is more difficult to understand than others. A lot of money can be made with a good interface. "AOL, so easy to use, no wonder they are number 1."

Previous studies

Human behavior changes little over time and therefore older studies are often useful. But in this case

- focus on bibliographic databases
- study boolean systems without relevance ranking
- study professional intermediaries doing searches
- focus on limited number of collections with switching costs
- consider tightly focussed databases
- study systems without hypertext navigability between them
- look at character interfaces driven by proprietary command languages

An iterative model of the search process

1. Start with an information need
2. Select system and collection to search on
3. Formulate a query
4. Send query to system
5. Receive the results back from the system
6. Scan, evaluate and interpret the results
7. Stop OR make another query, go to point 4

Problem with that model

- It assumes that the information need is static. Counterexample: Maryanne and Joe
- ignores the role of navigation while scanning the search results

Bates' berry picking model

1. User needs shift constantly as a result of reading and learning during the information seeking process. Information encountered at one point may alter the search process into a new direction. One goal may become fulfilled, and then a new one becomes important.
2. The information need is not satisfied by one final document, but rather by bits and pieces of information that is gathered along the search.

Related observational studies

- The information seeking process consists of diverse but inter-related actions brought together by a problem-based theme.
- Search results for one goal tend to trigger new goals, but context of the search problem remains the same.

What do these theoretical ideas mean for our user interface?

A good user interface should support berry-picking

A user interface should allow users to reassess their goals and adjust their search strategy.
When a trigger leads the user elsewhere, it should be possible for the user to return to the unfinished business form before.
Record the process of information search, allow to store intermediate results.

Shneiderman's four principles

1. provide informative feedback
2. permit easy reversal of actions
3. reduce working memory load
4. provide different interfaces for novice and expert users

1. provide feedback

- about the relationship between their query specification and documents retrieved
- about relationships among retrieved documents
- about relationships between retrieved documents and the collections

3. reduce memory load

- allow users to keep track of what they are doing
- allow users to go back
- have browsable information about the stage of the search process
 - links to related items
 - links to new searches

4. provide different interfaces of beginners and expert

This helps to overcome the power versus simplicity tradeoff. Powerful interfaces allow experienced users more control.

They are off-putting and imply learning efforts for people who are new to the search and use the system only rarely.

Other topics that Hearst studies

1. Starting points
2. Query specification (we will do that later)
3. Context
4. User relevance judgment
5. The search interface

Starting points

- lists: examples are catalogs of holdings at large database provider, personal bookmarks
- overviews, an initial list plus the possibility to go to sublists (more on that later)
- examples, guided tours and wizards (more on that later)
- automated source selections, is possible if you have a user profile that has been created prior to use

more on overviews

Overviews can be constructed using traditional subject classification schemes such as MeSH for example.

- begin at a logical starting point
- avoid having to guess an initial search string out of the blue

The success of Yahoo suggests that this is a popular approach. Problem remains that it is not easy to integrate them with searches because category display occupies a large amount of screen space.

Another problem is that classifying documents is labor-intensive.

more on examples and wizards

User interfaces that start by an example, or have guided tours are still at a research stage.

For software, MS invented the wizard. It shows how something is done without teaching details

It is found useful if the tasks involves many steps that have to be carried out in sequence by people who do not really know what they are doing.

It is not useful when the interface does not solve the problem efficiently, when the goal is to teach the user an interface and when the wizard was not tested.

Return from queries

There is a whole set of research projects that deal with how to return the results in such a way that users can work out quickly which ones are relevant.

The one principle that is in use and that seems helpful is KWIC, which stands for query in context. Such an interface will show the query terms highlighted, with some surrounding contents. There are decisions to be made on how much to display.

Text summarization research suggests that the best terms appear in the beginning of the document.

Relevance feedback

When the user has selected documents that correspond to a query, a software system can try to find similar documents.

After the user has checked a number of documents, the system presents a new query—which the user can change—or it automatically selects a new results set.

Problem: what to do with the initially selected documents?

Interface support for the search process

General searches are easy to create an interface for.

How to set up searches on sub-collection?

Monolithic search screens make all the search options visible at once, allow for tight control of options and hopefully can be learned by a user. They have the problem to crowd the screen space and can be daunting to new users.

Overlapping windows provide more flexibility but can quickly lead to a "lost in hyperspace" feeling.

Assignment

Write a one-page typeset document about the web interface that you use.

Is it well designed?

What could be done to improve it?

Do you know of bad interfaces that you do not use?

If useful, refer back to your search experience for the first assignment.