

Visual MeSH

Xia Lin

College of Information Science and Technology

Drexel University

xlin@drexel.edu

ABSTRACT

Visual MeSH is a graphical interface developed to interact with real world databases such as the National Library of Medicine's MetaThesaurus and MEDLINE. Visual MeSH makes it easy for the user to browse complex relationships through visual representation of thesaurus structures. It helps the user to select controlled vocabulary terms for online searching in the click-and-choose environment.

1. INTRODUCTION

Controlled vocabularies have proved to be a very useful tool in the traditional online searching environment [2]. They are extremely useful for improving both search precision and search recall. However, its use in today's web-based environment has been very limited. One of the difficulties is the lack of good browsing tool for viewing complex relationships in controlled vocabularies. Users need to see relationships among related terms in order to select appropriate controlled vocabulary terms for their searches. Users want to be able to directly drag the controlled vocabulary terms to their search environments to build up their search queries. Without such helps, it would be very difficult for the end-user to utilize the power of controlled vocabularies, particularly in the web-based environment where the user is used to graphical and click-and-choose interfaces.

Visual MeSH is designed to assist users to find and choose MeSH terms for their MEDLINE searching. MeSH, the Medical Subject Headings, is a comprehensive medical thesaurus used by the National Library of Medicine to index documents in the MEDLINE database. Currently, Both MeSH and MEDLINE are available on the web site of the National Library of Medicine. They can be accessed through html-based interfaces, but they are not fully integrated to allow easy use of MeSH terms for the MEDLINE searching.

Visual MeSH provides an integrated graphical interface to let the user interact with contents and structures of MeSH and MEDLINE. It supports three useful functions for the user. (1) The interface provides a connection to both the MetaThesaurus database and PubMed, the public access MEDLINE retrieval system on the web [4]. The connection is transparent to the user (the user does not need to know how the databases are connected).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. SIGIR '99 8/99 Berkeley, CA, USA
© 1999 ACM 1-58113-096-1/99/0007...\$5.00

(2) The Interface allows the user to look up MeSH terms in a click-and-choose environment that lets the user explore all the term associations visually and interactively. (3) The interface supports visual construction of Boolean queries. It provides instance feedback and uses the number of hits and document distributions on a map to help the user's query modification process.

2. THE INTERFACE

To start a search in Visual MeSH, the user can select to lookup concepts from either MetaThesaurus or MEDLINE. If the choice is lookup in MetaThesaurus, the user's query will be sent to the UMLS MetaThesaurus database (UMLS, Unified Medical Language System developed by The National Library of Medicine, includes MeSH as its main vocabulary source, [3]). The system will return a MeSH Concept as an object, including the concept's definition, its synonyms, narrower terms, broader terms, related terms, co-occurrence terms, etc. If the choice is lookup in MEDLINE, the system will conduct a preliminary MEDLINE search with the user's free text query and display a list of MeSH terms abstracted from documents retrieved by the query. The user then can select any of the MeSH terms to perform a lookup in the MetaThesaurus.

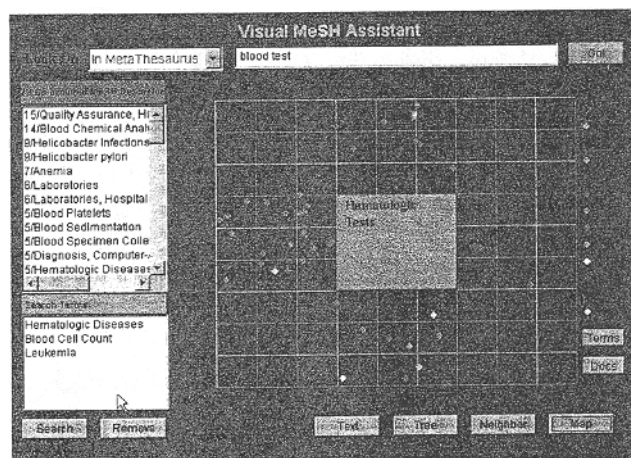


Figure 1. Visual MeSH. When the user lookups the term "blood test," the concept "Hematologic Tests" is found. The interface then helps the user to explore the concept's broader terms, narrower terms, related terms, and synonyms. A list of co-occurrence terms (terms co-occurs with "Hematologic Tests" in MEDLINE documents) is also retrieved and displayed on the interface.

Once a MeSH Concept is retrieved, Visual MeSH provides several views of the concept, including tree view, neighbor view, and map view (Figure 1). The tree view displays each MeSH Concept in a tree with branches of synonyms, broader terms (BT), narrower terms (NT), and related terms (RT). Each branch can be expanded or contracted to help the user browse one branch at the time and reduce the scrolling. The neighbor view allows the user to move any of the terms to the center for a display of synonyms, related terms, narrower terms, and broader terms in its up, down, left and right "neighbor lists." Clicking on any term in the neighbor lists will move the term to the center and all the neighbor lists will be updated correspondingly. This allows the user to easily move up or down, left or right in a complex multi-hierarchical structure. If the user moves the mouse over the center, a window will pop up with the definition of the term.

The Map View shows two types of maps, the term map and the document map. The term map roughly positions synonyms, broader terms, narrower terms, and related terms in the same locations as in the neighbor view. When the TERMS button is pressed, the system will retrieve each term's Semantic Type from the MetaThesaurus and represent it by color. One hundred thirty-five semantic types (as defined in the MetaThesaurus) are grouped into 13 groups, and labeled with colors, such as "Diseases & Pathologic process" is represented by color RED, "procedures or treatment" by GREEN, and "Concepts & Idcas" by WHITE, etc. Moving the mouse over the list of colored diamonds on the right will reveal what each color represents.

If the user drags four terms into the four corners, the user can click on the DOCS button to get a document map. The map represents MEDLINE search results using the four selected terms in different combination. Results of AND-ing all the four terms will be in the center, AND-ing 3 terms together in different combination will be in the four corners, etc. All together the system conducts 9 searches and represents results on the map.

On any of the views, the user can double-click on a term to select it. The selected terms are added to a listed window. A number below the window indicates the number of hits that a search will get from MEDLINE (the system automatically conducts a MEDLINE search each time a term is added or deleted from the list of the selected terms). This number helps the user to modify the selected terms. If the number is too large, more terms should be added; if the number is too small, some selected terms may need to be removed. Only when the number is within a reasonable range does the user need to retrieve search results by clicking on the search button, which will display the search results in a separate browser's window.

3. SYSTEM DESCRIPTION

Visual MeSH is a Java RMI application based on a three-tier architecture [1]. The data server tier, including the MetaThesaurus and MEDLINE, is provided by the National Library of Medicine. The middle tier is a Java RMI server, which includes the search logic and communication protocols and methods. All the query construction and MeSH terms abstraction are implemented in RMI server. The client tier is a Java applet, which provides visual displays and interaction functions (Figure 2).

4. CONCLUSION

In summary, Visual MeSH makes it easy for the user to utilize controlled vocabularies for web-based online searching. Major research issues involved in the implementation of the system include: (1) graphically displaying semantic structures in a thesaurus to help users navigate through the unfamiliar concept space; (2) designing interactive interfaces to help users convert free text queries to controlled vocabulary-based queries; and (3) exploring object-oriented design principles for GUI design over a distributed network (i.e., connecting MEDLINE server, MetaThesaurus server through a Java RMI server to an interactive interface).

Visual MeSH can be accessed at:

<http://research.cis.drexel.edu/mesh/index.html>

5. REFERENCES

- [1] Dowling, T. B. (1998). Java RMI: Remote Method Invocation. Foster City, CA., IDG Books Worldwide.
- [2] Lancaster, F. W. (1986). Vocabulary Control for Information Retrieval. 2nd ed. Arlington, VA, Information Resource Press.
- [3] Lindberg, D.; Humphreys, B. L.; & McCray, A. T. (1993). The Unified Medical Language System. *Methods of Information in Medicine*, 32(4), 281-291.
- [4] National Library of Medicine (1998). The NLM PubMed Project. [available: <http://www.ncbi.nlm.nih.gov/PubMed/overview.html>].

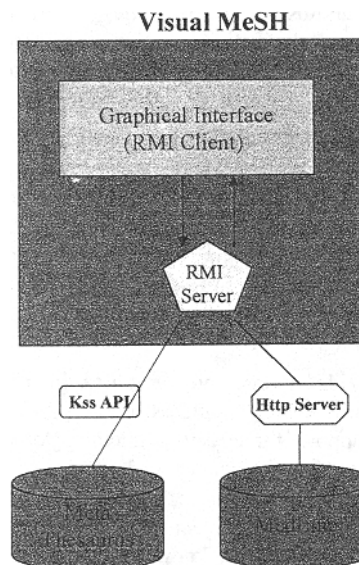


Figure 2. Visual MeSH Architecture