Interactive Query Workstation:
A Demonstration of the Practical Use of UMLS Knowledge Sources

Christopher Cimino
Office of Education
Albert Einstein College of Medicine
Bronx, NY

G. Octo Barnett, Dyan Ryan Blewett, Laurie J. Hassan,
Robert Grundmeier, Rulane Merz, Joel A. Kahn, John A. Gnassi
Laboratory of Computer Science
Massachusetts General Hospital
Harvard Medical School
Boston, MA

ABSTRACT

The Interactive Query Workstation (IQW) has been developed to provide clinicians with a uniform program interface for retrieving medical-related information from various computer-based information resources. These resources can vary in content (bibliographic databases, drug information, general medical text databases), function (article retrieval, differential diagnosis, drug interaction detection, or drug dosage and administration information), and media formats (local hard disk, CD-ROM, local area network, or distant telecommunication link). IQW allows modular addition of new resources as well as extension of previously installed resources. The National Library of Medicine's three Unified Medical Language System (UMLS™) Knowledge Sources, the Metathesaurus (Meta), the Semantic Network, and the Information Sources Map (ISM) have been incorporated into many aspects of IQW. Meta provides information about medical terminology and aids IQW in isolating the basic concepts from a clinician's question. The Semantic Network provides information about the categorization of concepts and possible relations between concepts. It also assists IQW in determining which queries are appropriate for a set of concepts contained in the clinician's question. The ISM provides information about the content available from a computer-based resources and aids IQW in selecting an appropriate resource from which to collect information. The computer-based resource selection is performed without user intervention. This interactive demonstration shows an environment which increases the accessibility of medical information to clinicians by utilizing the three UMLS Knowledge Sources.

INTRODUCTION

The first step in finding an answer to a question is composing the question. Understanding the nuances of the content of a question can prove to be the key to successfully finding an answer. Therefore it would seem that a clinician would be best able to formulate a question for retrieving medical information. Unfortunately, the formulation of a successful search strategy requires a combination of expertise in the fields of both library science and medical science.

The formulation of a search strategy can be described in four stages; resource selection, translation, structure, and improvement. A search librarian will first choose the most appropriate source of information, (e.g. TOXLINE™ versus ChemAbstracts™ for retrieving chemical data). Next, concepts in the question are determined and translated into the vocabulary of the resource, (e.g. a search through MEDLINE™ using Medical Subject Headings (MeSH™) is more accurate than using title/abstract text word searching). The librarian would then make use of a resource's syntax features, such as explosion, subheadings, and boolean logic, to maximize the accuracy and recall of the search. The final step would be to review the results of the initial search. Then, the search strategy could be modified to enhance future retreival results. The improvements would come from finding better terms with which to search or by clarifying the medical nuances of the question.

Research has been done to try to improve three of the stages of the search process. Relevancy feedback programs have attempted to analyze the indexing of an article a user selected or rejected to find more useful indexing terms than the ones initially used [1]. Automatic word indexing and weighing algorithms have been designed in an attempt to eliminate the need for a controlled vocabulary [2]. Analysis of co-occurrence data of subheadings has been done to try to match combinations of subheadings used to specific relations in clinical questions [3].

The products of the UMLS project [4] provid data which can be used to enhance an automatic search process.

823

The Interactive Query Workstation (IQW) makes use of these products to aid a clinician in selection, translation, and structure of clinical queries [5].

SYSTEM DESCRIPTION

The IQW consists of several component programs that interact and provide the user with a uniform interface. Meta provides information about medical concepts that appear in several medical vocabularies and resources. Paradox™, a relational database package, is used to manage the large amount of data Meta contains. The main interface functions are supported by code written in Smalltalk/V Windows™. Smalltalk is also used to represent the UMLS Semantic Network and ISM. These Knowledge Sources provide information about the potential relation between concepts and the potential information in resources, respectively. Both Knowledge Sources have been supplemented with additional information to support the medical resources that IQW accesses. Bridge™, a scripting language, provides connection and communication support to the medical resources. IQW can successfully connect to resources that are on the same machine as IQW, on a local network server, over the Internet, or over a telecommunications link. All these components run on a Microsoft DOS compatible 486 computer. Some of the programs must run in the Microsoft Windows environment.

DEMONSTRATION

The two greatest challenges we face in developing the IQW are technical problems in communicating with proprietary computer-based medical resources and presenting the user with a uniform interface to numerous resources. The technical problems have been discussed in previous papers [5]. The demonstration will show that we direct the user's attention to the effects of resources on their query. IQW's representation of the query is displayed to the user. This representation is manipulated by the user, Meta, the Semantic Network, the ISM, and IQW to refine the query. It is altered by the computer-based medical resources to provide information which may answer the query or lead to new queries. The query representation becomes the uniform interface through which the user interacts with other resources.

FUTURE DEVELOPMENT

The goal in designing the IQW was to provide an environment that allows testing of the many scenarios for assisting a user in the search processes. Because the system is flexible, there are numerous ways in which improvements could be made. One form of improvement is the inclusion of a rule-based system to aid in search formulation by using the extensive co-occurrence data found in Meta This co-occurrence data can be used to make accurate pre-search estimates of information retrieval so a search strategy can be improved before actually performing the search. A natural extension of this work would be to incorporate relevancy feedback in the post-search improvement of the search strategy. Also, the search expertise of IQW could be made available to other applications by providing IQW as a search engine. Currently, a network server version of IQW is being developed for use as a search engine between a Physician's Workstation and the resources IQW has available. Eventually, IQW could be accessed by numerous applications available on a hospital network.

References

1 Dillon M and Desper J. The Use of Automatic Relevance Feedback in Boolean Retrieval Systems, Journal of Documentation, Vol. 36, No. 3, September 1980, pp. 197-208.

2 Hersh W, Hickam DH, Haynes RB, McKibbon KA. Evaluation of SAPHIRE: An Automated Approach to Indexing and Retrieving Medical Literature, "Symposium on Computer Applications in Medical Care" [Clayton PD, ed.], McGraw-Hill, New York, 1991, pp. 808-812.

3 Cimino JJ, Mallon LJ, and Barnett GO. In: "Symposium on Computer Applications in Medical Care" [Greenes, Robert A., ed.], IEEE Computer Society Press, 1988, pp. 180-184.

4 Lindberg DAB and Humphreys BL. The UMLS Knowledge Sources: Tools for Building Better User Interfaces, In: "Symposium on Computer Applications in Medical Care" [Miller RA, ed.], IEEE Computer Society Press, 1990, pp. 121-125.

5 Cimino C, Barnett GO, Hassan LJ, Blewett DR, and Piggins JL. Interactive Query Workstation: Standardizing access to computer-based medical resources, Computer Methods and Programs in Biomedicine, 35, 1991, pp. 293-299.