Linking A Clinical System to Heterogeneous Information Resources

Qing Zeng M.S., James J. Cimino, M.D.
Department of Medical Informatics
Columbia University, New York, New York

We present a model for providing clinical information system (CIS) users with quick access to high quality information resources. We have developed a chest X-ray information button application which is attached to the chest X-ray reports in the CIS at the Columbia Presbyterian Medical Center. The application generates questions based on clinical information, user interest, generic question templates, and resource availability. It then provides answers to the questions through integrated access to heterogeneous information resources including the CIS itself and publicly accessible Web resources.

INTRODUCTION

Information needs of health care providers arise throughout the patient care process, and interacting with clinical information systems (CIS) is one such occasion. For example, a physician looking at a patient's radiology report might want to know if the findings in the report have appeared before, or when prescribing a drug, might what to know the proper dosage. Some of these information needs could be fulfilled by the information within the CIS, while others need to involve outside resources.

This paper describes some of our efforts in providing a need-oriented information system. Three hypotheses underly our work: 1. That it is possible to predict some of users' information needs in the form of questions. 2. That it is possible to find the answers in the CIS and other resources. 3. That it is possible to link the questions to the answers.

The first hypothesis has been explored by other researchers[1,2]. Using a chest X-ray information button application, we examine the last two hypotheses by showing how some typical questions can be evoked by the CIS and answered by a variety of information resources , including the CIS itself.

At Columbia-Presbyterian Medical Center (CPMC) a comprehensive CIS is used extensively by health care providers, and current studies are looking at extending CIS access to patients as well. The CIS collects and provides many types of information including laboratory, pharmacy and radiology reports. Each year, more information is accumulated and new kinds of information are added to the system. At the same time, medical information resources on the World Wide Web are expanding.

For health care providers to benefit from this growing amount of information, however, they must be able to retrieve relevant information in a timely fashion. By providing quick access to high quality information according to users' information needs, our application has the potential to help health care professionals deliver higher quality medical care.

METHODS

The goal of our application is to fulfill information needs that arise during the use of the CIS. The information needs are represented in the form of questions. The answers to those questions are generated in the form of customized representations of the information retrieved from the heterogeneous information resources. [Figure 1]

Preparation

The information button application is evoked by a user while reviewing a chest X-ray report and generates questions based on the contents of the report, the application's knowledge of the available resources, and generic question templates[3]. Twelve such templates were created for this application, an example of which is: "Has this <disease/finding> appeared in this patient's previous X-ray reports?"

Questions regarding patient specific information access the CIS resource, while questions regarding non-patient-specific information access general information resources. There are two basic formats for such resources: controlled and uncontrolled. The contents of the controlled format resources are organized and maintained systematically and consistently, and these resources often have controlled vocabularies. Medline is an example of such a resource.

Uncontrolled format resources are more "free style", and their vocabulary, organization and maintenance is less consistent or systematic. They are more difficult to index, but their content is often valuable and may

not be available elsewhere. Most Web resources fall in this category.

Our application makes use of both controlled and uncontrolled format resources. They include Medline and three publicly available Web resources: the Collaborative Hypertext of Radiology (CHORUS)[4] at the Medical College of Wisconsin (http://chorus.rad.mcw.edu/), The Internet Pathology Laboratory for Medical Education (IPLME)[5] at the University of Utah (http://www-medlib.med.utah.edu/WebPath/) and the Radiology Museum (RM) at South Bank University (http://www.sbu.ac.uk/~dirt/im0.html#ACR-museum). Each Web resources consists of a collection of Hypertext Markup Language (HTML) documents addressed by Universal Resource Locators (URLs). We collected vocabularies from the indexes of each resource and found additional terms in titles of the HTML documents. CHORUS and IPLME use uncontrolled local vocabularies, while RM uses the American College of Radiology (ACR) classification which is included in the Unified Medical Language System (UMLS)[6]. We mapped all the index vocabularies into the UMLS using a method we developed (MRUM)[7], then created a URL database containing the URLs and index information.

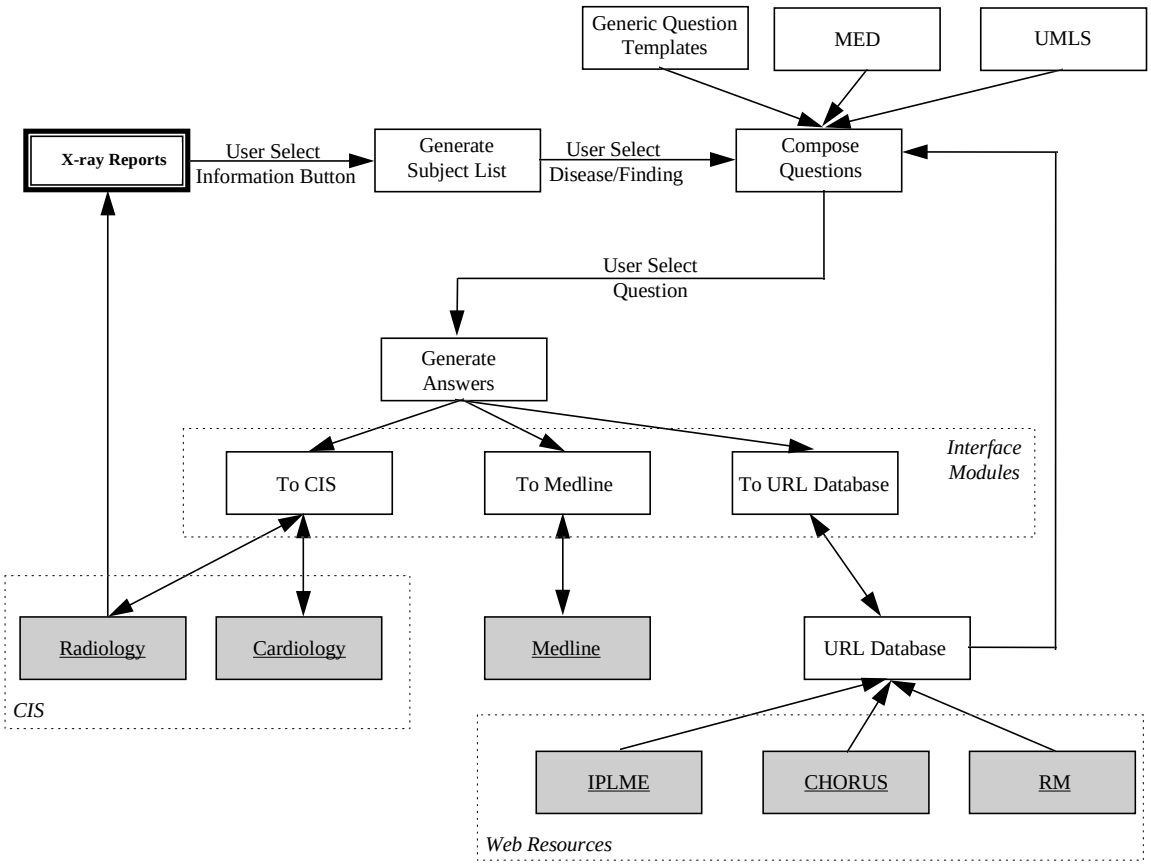
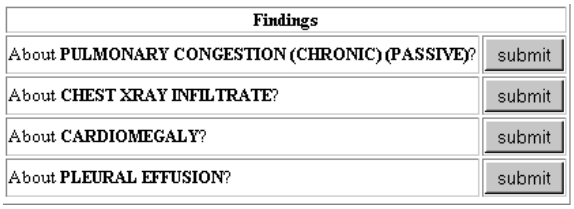


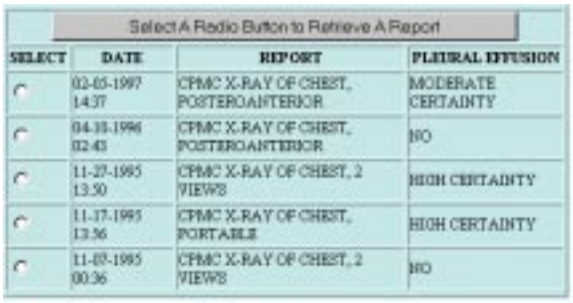
Figure 1. Schematic Diagram of the Information Button. We predict information needs will arise while a clinician reviews chest X-ray reports (upper left corner) from the CIS. When the information button is clicked, a subject list of diseases and findings is generated based on the report. Once a user selects a subject of interest, the application composes questions by matching the subject to generic question templates and checking resource availibility with the help of knowledge in the MED, UMLS and a database of documents found in Web reources (URL database). An answer is generated when the user selects a question. The application queries the CIS, Medline or the URL database of Web resources (shaded boxes), collects the query result and presents it to the user.

~~~a~~~



~~~b~~~

~~~c~~~



~~~d~~~

Figure 2. A sequence of screenshots showing an example of a user interaction: (a) shows a radiology report with an information button attached, (b) shows the findings detected in the report, (c) shows the questions generated for "Pleural Effusion", and (d) shows the answer to the first question in (c).

Implementation

The chest X-ray information button application is evoked by a clickable button linked to the chest X-ray radiology reports in the CIS (Figure 2a). It retrieves from the CIS database a coded version of the same report generated by a natural language processing program[8]. Based on the coded report, a clinical subject list of findings and diseases is created and presented to the user (Figure 2b).

After a disease or finding is chosen, a list of questions is shown (Figure 2c). The questions are generated in three steps: 1. The user selects a clinical subject they are interested in. 2. The application matches the clinical subject with the generic question templates (the application queries the MED for classification information necessary to select among templates specific to clinical subjects, such as cardiovascular diseases). 3. The application checks the availability of resources that may answer the questions. No question is generated if a resource for that question is unavailable. For clinical questions such as "Has this finding appeared in previous reports?", the CIS can always offer an answer even if the answer is "No". For the Medline search, our application assumes some articles are always available. (This could be refined by using the UMLS co-occurrence table). For the Web resources, the local URL database supplies the availability information. The clinical subject is also translated to the UMLS at this step, because the URL database uses the UMLS as its index vocabulary.

Interface modules connect the questions and the information resources. They perform three tasks: 1. Translate the questions into queries that can be understood by the CIS, Medline, and the URL database. 2. Transfer the queries to the resources. 3. Collect the query results and generate a formatted display (Figure 2d).

For the CIS database, the questions are translated into HL7 query messages. No vocabulary translation is needed because both the questions and the CIS use the Medical Entity Dictionary (MED)[9] to code patient data. For Medline, the questions are translated to Ovid Medline search strategies. Medline's vocabulary is the Medical Subject Heading (MeSH), therefore, MED codes are translated into MeSH terms whenever possible. For the URL database, the questions are translated to database queries and MED codes are translated into UMLS concepts.

The CIS database is located on a remote mainframe computer to which HL7 queries are transferred from the machine where the application runs. The Medline

database is hosted on a remote UNIX computer and accessed by a scripted telnet session. The URL database is on the same machine as the application and transfer is done by simply calling and passing parameters to the database management program.

Results are returned in the form of HL7 messages from the CIS database, telnet display formatted messages from the Medline, and database entries from the URL database. All results are filtered and reformatted in HTML for Web based display.

Evaluation

Our evaluation has thus far concentrated on the application's ability to generate and answer questions rather than attempting to satisfy information needs of hypothetical users.

There are four generic question templates linked to the CIS:

1. Has this <finding/disease> appeared in this patient's previous X-ray reports?
2. This <finding/disease> is a <more general type of finding/disease>, do you want to search for other such findings or diseases in other reports?
3. This <disease> is a cardiovascular disease, do you want to search for the ECG reports around this period (+/- 15 days)?
4. This <disease> is a cardiovascular disease, do you want to retrieve all the ECG reports?

We used medical records from three different patients to test the four questions by measuring the number of reports retrieved. Starting from the most recent report in which each of the questions appeared, we also compared the average number of key strokes that were needed to find the answer to a question with and without using the information button.

There are five Medline-related generic question templates:

1. What is the significance of the <finding>?
2. What is the cause of the <finding>?
3. How to treat the <disease>?
4. How to diagnose the <disease>?
5. Everything about <finding/disease>?

We chose three findings or diseases for each question and measured the average number of articles retrieved. Retrieved articles were restricted to human only, English only and with abstract only. We also observed the average number of key strokes required to retrieve the information using the information button and doing similar searches using Ovid Medline.

There are three Web resource related generic question templates:

1. Would you like to see some pathology images of this <finding/disease>?
2. What's the pathology explanation of this <finding/disease>?
3. What's the differential diagnosis of this <finding/disease>?

We tested three findings or diseases for each question and recorded the number of Web pages found for each. We then measured the average number of key strokes required to find an answer using the information button and using the Alta Vista Internet search engine.

RESULT

Based on the 110 reports we tested, the mean number of diseases and findings found in a report is 2.52 with a standard deviation of 1.48. For the 20 most common findings and diseases, we found that for each finding or disease, the mean number of questions generated is 5.90 with a standard deviation of 1.01. Thus, on average, 14.87 questions could be generated for each report. If we assume an even distribution of the types of questions (CIS, Medline, Web) generated per finding/disease, then from one X-ray report an average number of 229 links to other reports and documents could be created.

Table 1: The results of the information button queries and a comparison of keystrokes required when using the information button vs. manual methods.

Resource	Average Number of Results Found Per Question	Keystrokes using Information Button	Keystrokes using Manual Methods
CIS	7.5	3.3	22.5
Medline	35.7	7.0	58.0
Web Resources	3.0	4.0	22.0

DISCUSSION

The chest X-ray information button application is able to provide quick access to high quality information resources for CIS users. By using the information button, a user can find the answers to the predicted questions with fewer key strokes than by conventional methods. There are numerous Internet resources available, however they are not of uniform quality. Internet publication standards have yet to be established and many resources are inaccurate, outdated, and unreliable. By selecting resources that

can answer our questions adequately and are provided by creditable institutions, we can control the quality.

There have been many applications[10] that link clinical systems to resources such as Medline. A distinction of our application is that it integrates heterogeneous resources, including the CIS and publicly accessible Web resources, to address information needs. According to research on physician information needs[1,2], a significant number of questions are about patient specific clinical information. So CIS is therefore an important resource to employ.

Our application uses a general model that can be scaled to address a broader domain of clinical information. At present, the application is attached only to the chest X-ray reports. However, it could be extended to other parts of the CIS (e.g. admission profiles or lab results) which use MED-coded data. The generic questions that we are using now are prototype questions, and a survey is under way to expand the question set.

The development of the information button application leads to other directions worth exploring. The CIS at the CPMC presents results and orders organized according to ancillary departments and chronologic order. For example, X-ray reports of the foot and the chest are stored together because they are both done by the radiology department, and all radiology reports are sorted by time. The information button links the information across these lines and can create new views of the CIS that organize information based on the content and function instead of where and when it is collected. For example, if the user has a question about a patient's heart, clinical data can be collected to show X-rays of the chest, EKGs, heart medication, echocardiograms and cardiac enzyme tests.

CONCLUSION

The chest X-ray information button application demonstrates that it is possible to provide answers to CIS user questions by linking the CIS to heterogeneous information resources including the CIS itself, image repositories, text books, and a bibliographic database. In our limited study, we were able to provide users with access, via a few button clicks, to an average of 229 reports and documents for each X-ray report they reviewed.

Acknowledgment

This work was supported by UMLS contracts from the National Library of Medicine (NLM 95-053/VMS, 001 LM05857-01). The authors thank Raisa Garber and Cynthia Chen for their help in constructing HL7 queries.

References

1. Forsythe DE, Buchanan BG, Osheroff JA, Miller RA. Expanding the concept of medical information: an observational study of physicians' information needs. *Computers & Biomedical Research.* 1992 Apr;25(2):181-200
2. Osheroff JA, Forsythe DE, Buchanan BG, Bankowitz RA, Blumenfeld BH, Miller RA. Physicians' Information Needs: Analysis of Questions Posed During Clinical teaching. *Annals of Internal Medicine*. 1991 Apr 1;114(7):576-81
3. Cimino JJ. Johnson SB. Aguirre A. Roderer N. Clayton PD. The MEDLINE Button. Frisse ME, ed. *Proceedings the Annual Symposium on Computer Applications in Medical Care* :81-5, 1992.
4. Kahn CE Jr. CHORUS: a computer-based radiology handbook for international collaboration via the World Wide Web. *Radiographics*. 15(4):963-70, 1995 Jul.
5. Klatt EC. The "Electronic City" in the Laboratory. *Laboratory Medicine* 1996; 27(2):117-121.
6. National Library of Medicine. *UMLS Knowledge Sources - 8th Experimental Edition Documentation.* Bethesda, Maryland: The Library. 1997 Jan.
7. Zeng Q, Cimino JJ. Mapping Medical Vocabularies to the Unified Medical Language System. In Cimino JJ, ed.: *Proceedings of the American Medical Informatics Association Annual Fall Symposium*. Washington, DC; October, 1996:105-109.
8. Friedman C. Alderson PO. Austin JH. Cimino JJ. Johnson SB. A general natural-language text processor for clinical radiology. *JAMIA*. 1994 ; 1(2):161-74.
9. Cimino JJ, Clayton PD, Hripcsak G, Johnson SB. Knowledge-based approaches to the maintenance of a large controlled medical terminology. *JAMIA*, 1994; 1(1):35-50.
10. Cimino JJ. Linking patient information systems to bibliographic resources. *Methods of Information in Medicine*, 1996;35(2):122-126.