Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (2008)

Intelligent Output Interface for Intelligent Medical Search Engine

Gang Luo
IBM T.J. Watson Research Center
luog@us.ibm.com

Abstract

To facilitate ordinary people to search medical information, we have built an intelligent medical Web search engine called iMed. iMed uses medical knowledge and an interactive questionnaire to find multiple diseases serving as queries. The search results of these queries are combined together and returned to the searcher in a traditional sequential order. Nevertheless, searchers still frequently miss desired information, because the traditional search result output interface cannot capture the internal structures of medical search results. This paper presents a new, intelligent search result output interface devoted to intelligent medical search. The new output interface automatically offers searchers what they want instead of waiting until they ask explicitly. It structures all the search results into a multi-level hierarchy with explicitly marked medical meanings. In this way, searchers can efficiently navigate among all the search results and quickly obtain desired information. We demonstrate the effectiveness of our techniques through an evaluation using USMLE medical exam cases.

1. Introduction

On an average day, 6% of American Internet users use Web search engines to search for medical information on the Web (Sherman 2005). Since October 2005, several medical search engines have been launched, including Healthline (Healthline 2007), Google Health, SearchMedica, and Medstory. They use the traditional keyword query interface, which works well when the searcher clearly knows his medical situation. However, in many cases, the searcher is uncertain about the problem he is facing and unaware of the related medical terminology (e.g., panophthalmitis). As a result, it is often difficult for him to choose a few accurate medical phrases as a starting point for his search.

To address this problem, we have built a prototype *i*ntelligent *med*ical search engine called iMed, which uses medical knowledge and an interactive questionnaire to help searchers form queries. Below we first give a brief overview of iMed, and then focus on a new, intelligent search result output interface for intelligent medical search engines, which can help searchers quickly find their desired information.

1.1 Overview of iMed

iMed leverages its built-in medical knowledge in the form of diagnostic decision trees written by medical professionals (Collins 2002). As shown in Figure 1, each diagnostic decision tree corresponds to either a subjective symptom (e.g., fatigue) or an objective sign (e.g., hypertension). Each non-leaf, non-root node of a diagnostic decision tree corresponds to an answer to a question that iMed can ask. Each medical phrase in the leaf node of a diagnostic decision tree can become a query that iMed uses.

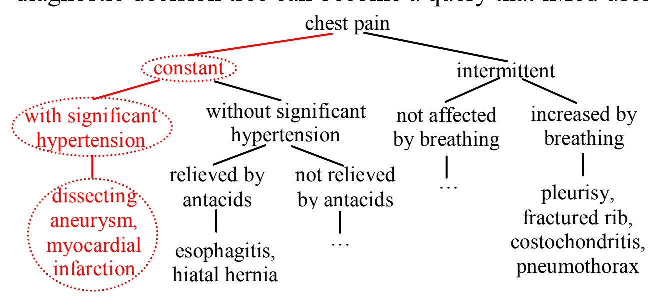

Figure 1. The diagnostic decision tree for the symptom "chest pain."

iMed uses diagnostic decision trees to help the searcher form queries. The searcher first selects one or more symptoms and signs from a list of known symptoms and signs (Collins 2002). Then iMed asks questions related to these selected symptoms and signs. Based on the searcher's answers to the questions, iMed navigates the corresponding diagnostic decision trees and automatically forms multiple queries with different weights. All these formed queries are sorted in descending order of their weights. Each query is used to retrieve some related Web pages. iMed combines the search results for all these formed queries together and returns them to the searcher in multiple passes. In the i-th pass, iMed obtains the i-th ranked Web page for each query. All those Web pages are sorted in the same order as their corresponding queries.

For example, Figure 1 shows the diagnostic decision tree T_d in (Collins 2002) for the symptom "chest pain." If "chest pain" is the only symptom chosen by the searcher, iMed's first question is "Is the pain constant or intermittent?" If the searcher answers "constant" to this question, iMed's next question is "Do you have hypertension?" If the searcher answers "yes" to the second question, iMed forms multiple queries. Each medical phrase in the selected leaf node (e.g., dissecting aneurysm) of T_d forms a query with a large weight. Each medical phrase (e.g., pleurisy) in the non-selected leaf nodes of T_d forms a query with a small weight.

1.2 Output Interface

We studied how users interact with an intelligent medical search engine that leverages built-in medical knowledge, and observed significant improvement over general Web search engines. However, we also observed that searchers still frequently miss desired information because the medical search engine uses the traditional search result output interface, where all the search results are returned to the searcher in a sequential order. This order cannot reflect the internal structures and relationships of the returned search results, as these results are retrieved using different queries that have explicit medical meanings.

We use an example to illustrate the above point. After reading a Web page that is about a disease *d* and returned by the search engine, the searcher can run into one of the following three situations. In the first situation, the searcher guesses that *d* can be related to his medical condition but he is not fairly sure. In this case, the searcher prefers to read another Web page about *d* to either disapprove or confirm his conjecture. In the second situation, the searcher is convinced that *d* matches well with his medical condition. In this case, he often prefers to read more Web pages about certain aspects of *d*, e.g., treatment, test. In the third situation, the searcher is convinced that *d* does not match with his medical condition. In this case, he prefers to skip all the remaining returned Web pages about *d*. The traditional sequential order presentation of search results cannot well serve these three situations simultaneously.

1.3 Our Solutions

One idea to address the above problem is to structure all the search results into a multi-level hierarchy that has explicitly marked medical meanings. More specifically, all the search results are organized into multiple categories according to their topics (e.g., disease names). Within each category, the corresponding search results are further divided into multiple sub-categories according to their aspects (e.g., symptom, diagnosis, treatment).

This hierarchical search result output interface is intelligent in the sense that it automatically offers searchers what they want instead of waiting until they ask explicitly. Ordinary searchers often do not remember or know all the information they want in the unfamiliar medical domain, but they usually can tell what they want if they can see such information. The intelligence of the output interface is possible because the medical domain is a closed one. In that domain, we have domain-specific medical knowledge and searchers' desired (sub-)categories are generally known in advance. Using this hierarchy, searchers can efficiently navigate among the search results and quickly obtain desired information. An automatic query formation method is used to construct this hierarchy.

In the traditional information retrieval literature, queries are inputted by searchers and the focus is on retrieving search results using better retrieval models. In contrast, the focus of this work is on automatically forming proper queries to obtain desired search results and effectively organizing these search results.

We crawled a large number of medical Web pages from the Internet and evaluated the effectiveness of our techniques using USMLE Step 2 CS medical exam cases (Le & Bhushan 2006). The results show that our techniques significantly improve the quality of search results as well as searchers' satisfaction and speed of finding desired information.

Besides medical search, the general ideas of this paper could also be applicable to other domain-specific (e.g., product) search. Suppose we have a knowledge base for a particular domain. Based on the criteria (e.g., price) specified by the searcher, this knowledge base can provide all those entities (e.g., cameras) satisfying these criteria and their interesting aspects (e.g., performance, design). Then an intelligent search engine for that domain can use our techniques to automatically build a multi-level search result hierarchy with explicitly marked meanings specific to that domain.

The rest of the paper is organized as follows. Section 2 describes the new, intelligent output interface. Section 3 evaluates the effectiveness of our techniques using a wide variety of medical cases. We conclude in Section 4.

2. A New Output Interface

2.1 Why Users Perform Medical Search

To gain insight into searchers' requirements on the output interface of intelligent medical search engines, we need to first understand why many Internet users prefer to perform medical search before visiting their doctors. There are several such reasons.

First, in some areas where the population is relatively sparse, doctors are often inexperienced because they rarely have the opportunity to encounter various kinds of diseases. If a patient has a rough idea about the diseases that may cause his problem before visiting his doctor, he can provide reminders to his doctor to avoid missing potentially relevant diagnoses or tests (Fox & Fallows 2003).

Second, many doctors have relationships with certain drug, medical equipment, or physical therapy companies and tend to prescribe to patients those companies' products or services. If a patient knows the treatment options in advance, he can suggest to his doctor his preferred alternative (e.g., inexpensive) treatment options.

Third, doctors are required to talk with patients in plain English and to explain medical jargons immediately after they are mentioned in the conversation (Le & Bhushan 2006). Nevertheless, patients can still have difficulty in understanding doctors due to unfamiliarity with some of the plain English words that are frequently used by doctors but seldom appear in daily conversations, e.g., organ names. This is especially true for both non-native English speakers and people with low education background. Performing medical search and reading related medical information in advance can help patients overcome this vocabulary obstacle.

Fourth, to facilitate diagnosis and treatment, patients need to answer doctors' questions and often have their own questions to ask doctors. If a patient has completely no idea about what questions to expect during his doctor's office visit, he may forget to tell his doctor some important information in a limited amount of time. Also, if a patient is unfamiliar with the related medical information, he may not know the appropriate questions to ask his doctor. Performing medical search and reading related medical information in advance can prepare patients for questions.

2.2 Requirements on the Output Interface

From the above description, we can see that ordinary Internet users usually perform medical search to serve multiple purposes concurrently. To achieve these purposes, medical information searchers prefer an intelligent medical search engine to return as complete search results as possible. More specifically, searchers would like to simultaneously see various topics (e.g., disease names) that are potentially relevant to their medical situations. For each such topic, searchers would like to simultaneously see all kinds of aspects (e.g., symptom, diagnosis, treatment) of it.

Most medical information searchers neither have much medical background nor are proficient computer users. For a potentially relevant topic, it is difficult for these searchers to list completely all their desired aspects and the corresponding keywords. Moreover, even if a searcher knows all his desired aspects and the corresponding keywords, it is troublesome for him to form multiple queries, one for each aspect, and use them to retrieve related information. Fortunately, in the medical domain, searchers' desired aspects of a topic are generally known in advance. For example, if the topic is about a disease, these aspects would include symptom, diagnosis, treatment, and prevention. To provide the greatest convenience to searchers, an intelligent medical search engine should automatically and concurrently retrieve related information for all the potentially relevant topics and their corresponding aspects. The search engine's output interface needs to organize these search results in a way that searchers can simultaneously see all the topics and aspects, quickly identify the important contents, easily navigate among those related search results, and freely make choices about which topics and aspects to view or skip.

2.3 Format of the New Output Interface

iMed uses diagnostic decision trees to find those topics that are potentially relevant to the searcher's medical situation. After obtaining the search results on those topics, iMed structures these search results into a three-level hierarchy that has explicitly marked medical meanings to fulfill the requirements mentioned in Section 2.2. This is shown in Figure 2.

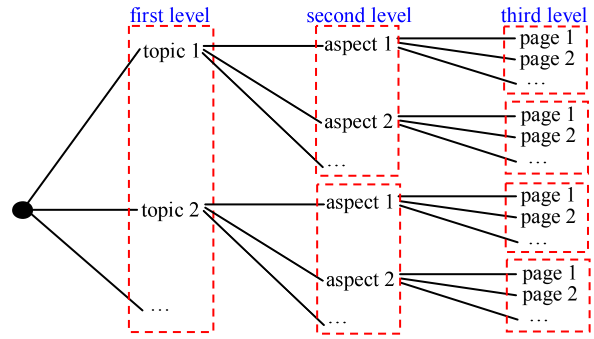

Figure 2. The new output interface's hierarchical structure.

2.3.1 Overview of the Hierarchy. We first give an overview of the hierarchical structure of the new output interface. At the first level of the hierarchy, all the search results are organized into multiple categories according to their topics (e.g., disease names). At the second level, within each category, the corresponding search results are further divided into multiple sub-categories according to their aspects (e.g., symptom, diagnosis, treatment). For each aspect of a topic, the retrieved Web pages are listed at the third level.

To expedite the speed that searchers find their desired information, medical meanings are explicitly marked at each level of the hierarchy. Also, overview Web pages are provided at the top two levels of the hierarchy. More specifically, at the first level of the hierarchy, topics are marked. For each such topic, an overview Web page is provided to help the searcher determine whether this topic is related to his medical situation. Similarly, at the second level, aspects are marked. For each such aspect of a topic, an overview Web page is provided to help the searcher determine whether this aspect is related to his medical situation. In many cases, searchers can quickly obtain their desired information directly from the overview Web pages at either the first level or the second level. Also, if searchers think that certain topics or aspects are irrelevant to their medical situations, they can completely skip the corresponding categories or sub-categories. In other cases, searchers need to reach the third level to obtain more detailed information.

iMed uses the diagnostic decision trees described in (Collins 2002) to find topics related to the searcher's medical situation. There, for each symptom or sign, the medical phrases in the leaf nodes of the corresponding diagnostic decision tree are disease names. Consequently, the topics mentioned at the first level of the new output interface currently cover only diseases. In the future, if other contents (e.g., exams) are added into these diagnostic decision trees, the topics and aspects mentioned in the new output interface need to be expanded accordingly.

2.3.2 First Level of the Hierarchy. Next, we present the details of the hierarchical structure of the new output interface. Figure 3 shows the format of the first level of this hierarchy. The new output interface organizes all found topics into one or more topic pages. Each topic page contains ten topic elements. To help a searcher quickly digest search results and refine his inputs (Healthline 2007; Luo *et al.* 2007), iMed suggests to him a few medical phrases related to his medical situation and lists them on the right side of each topic page. These medical phrases are extracted from the Medical Subject Headings (MeSH) ontology (MeSH 2007). MeSH is a standard vocabulary edited by the National Library of Medicine and widely used for indexing and cataloging biomedical and health-related documents.

Topic element 1	Medical phrases suggested for all found topics
Topic element 2	
…	
Topic element 10	
Topic Page 1 2 3 4 5 ► Next	

Figure 3. The high-level format of the first level of the new output interface.

As shown in Figure 4, each topic element E_T in a topic page corresponds to a different topic T (e.g., dissecting aneurysm). The left side of E_T contains the information of an overview Web page P_T^o of T, including the title, the snippet (i.e., some words extracted from P_T^o), and the URL of P_T^o. The searcher can view P_T^o by clicking its title. On the right side of E_T, there is a button entitled "More about T." If the searcher clicks this button, he will reach an aspect page at the second level of the new output interface and see more information about T.

Title Snippet URL	More about dissecting aneurysm

Figure 4. An example of the topic element format at the first level of the new output interface.

2.3.3 Second Level of the Hierarchy

dissecting aneurysm	
Symptom and sign element	medical phrases suggested for the topic
Diagnosis, exam, and test element	
Treatment element	
Cause and trigger element	
Risk factor element	
Complication element	
Medication element	
Surgery element	
Prognosis (expectation) element	
Rehabilitation, recovery, self-care, and home treatment element	
Prevention element	
Complementary and alternative medicine element	All found diseases
Resource, support, living with, and management element	
General information element	
Other information element	

Figure 5. The high-level format of the second level of the new output interface.

Figure 5 shows the format of the second level of the new output interface. At the top of the aspect page P_a, the corresponding topic T is listed in italics. All the aspects of T are listed on P_a in the form of aspect elements, e.g., the symptom and sign element. (At present, iMed uses the same set of aspect elements for all the diseases. Nevertheless, in theory, different diseases can have various sets of aspects. For example, some diseases are not treated with surgery and hence do not need the surgery element. As another example, certain diseases have their own unique aspects and need the corresponding aspect elements, e.g., diabetes patients need special diet recipes. We leave it as an interesting area for future work to construct the aspect database for all the diseases.) iMed suggests to the searcher a few medical phrases related to T and lists them on the right side of P_a. These medical phrases are extracted from the MeSH ontology. On the bottom right corner of P_a there is a navigation button entitled "All found diseases." If the searcher clicks this button, he will return to the topic page that is at the first level of the new output interface and contains the topic element E_T of T.

As shown in Figure 6, each aspect element $E_{A,T}$ corresponds to a different aspect A (e.g., symptom and sign) of topic T. The left side of $E_{A,T}$ contains the information of an overview Web page $P_{A,T}^o$ of A of T, including the title, the snippet, and the URL of $P_{A,T}^o$. The searcher can view $P_{A,T}^o$ by clicking its title. On the right side of $E_{A,T}$, there is a button entitled "More about A of T." If the searcher clicks this button, he will reach a result page at the third level of the new output interface and see more information about A of T.

Title Snippet URL	More about the symptoms and signs of dissecting aneurysm

Figure 6. An example of the aspect element format at the second level of the new output interface.

2.3.4 Third Level of the Hierarchy. Figure 7 shows the format of the third level of the new output interface. All the Web pages retrieved for aspect A of topic T are organized into one or more result pages. A of T is listed at the top of each result page in italics. Each result page contains ten Web page elements. iMed suggests to the searcher a few medical phrases related to A of T and lists them on the right side of each result page. These medical phrases are extracted from the MeSH ontology. Two navigation buttons appear on the bottom right corner of each result page. If the searcher clicks the button entitled "All found diseases," he will return to the topic page that is at the first level of the new output interface and contains the topic element E_T of T. If the searcher clicks the other button entitled "All aspects of T," he will return to the aspect page of T at the second level of the new output interface.

symptoms and signs of dissecting aneurysm		
Web page element 1	Medical phrases suggested for the aspect of the topic	
Web page element 2		
…		
Web page element 10		
Result Page 1 2 3 4 5 6 7 8 9 10 ► Next	All found diseases	All aspects of dissecting aneurysm

Figure 7. The high-level format of the third level of the new output interface.

As shown in Figure 8, each Web page element E_P corresponds to a different Web page *P* retrieved for aspect *A* of topic *T*. E_P contains the title, the snippet, and the URL of *P*. The searcher can view *P* by clicking its title.

Title
Snippet
URL

Figure 8. The Web page element format at the third level of the new output interface.

2.4 Discussions

The hierarchical structure of the new output interface currently has three levels while more levels can be easily added at appropriate places. For example, at the third level of the new output interface corresponding to the medication aspect of a disease *T*, we can include a button entitled "All related drugs." If the searcher clicks this button, iMed will create an additional three-level search result hierarchy for all the drugs used to treat *T*. The drug names are obtained from the Merck Manual (Merck 2007), the best-selling medical reference in the world. Those additional three levels correspond to drugs, aspects (e.g., dosage, side effect, precaution, and interaction), and Web pages, respectively.

2.5 Automatic Query Formation

The search result hierarchy is constructed using an automatic query formation method. We give a brief overview of that method here. The reader can find the details of that method in (Luo 2008). Our main observation is that the medical domain is a closed one. In the desired search result hierarchy, we can know the keywords for all the topics and their corresponding aspects. As a result, for each part of the search result hierarchy, we can use a different, specifically formed query to obtain the corresponding search result Web pages. When forming these queries automatically, we use domain-specific medical knowledge and consider the different roles that various levels play in the search result hierarchy. This can expedite the speed that searchers find their desired information. The resulting search result hierarchy fulfills all the requirements mentioned in Section 2.2.

To reduce the load on iMed and to maximize the speed that searchers can see iMed's outputs, iMed constructs the search result hierarchy one part at a time. Each part of the hierarchy is generated only at the time that it is needed.

3. Experimental Results

We conducted experiments with various medical cases to demonstrate the effectiveness of our techniques. iMed is a vertical search engine that crawls Web pages from a few selected, high-quality medical Web sites instead of all the Web sites. In our experiments, we crawled 22GB of Web pages from WebMD (WebMD 2007), Healthline (Healthline 2007), and Merck (Merck 2007), three of the most popular medical Web sites. These Web sites cover the entire medical domain fairly comprehensively and include information on various topics such as symptoms, diseases, drugs, and treatments.

We used USMLE Step 2 CS (Clinical Skills) medical exam cases (Le & Bhushan 2006). USMLE stands for the United States Medical Licensing Examination. Physicians have to pass this exam to obtain their licenses for practicing medicine. The exam cases used in USMLE Step 2 CS cover the typical cases encountered in daily medical practice. Each exam case has a summary that includes a several-page-long, detailed description of the patient's situation. One such medical case is shown in Figure 9. For each medical case, a few most likely candidate diagnoses are available, while making the final diagnosis among these candidate diagnoses usually requires special medical test results that are not included in the USMLE Step 2 CS Exam.

> The patient is a 46 years old male complaining chest pain. Chest pain started 40 minutes before the patient presented to the hospital. The pain woke the patient from sleep at 5 a.m. with a steady 7/10 pressure sensation in the middle of his chest that radiated to the left arm and the neck. Nothing makes it worse or better. Nausea, sweating, and dyspnea are also present. Similar episodes have occurred during the past 3 months, 2-3 times/week ...

Figure 9. An example USMLE Step 2 CS medical exam case.

We randomly selected 30 medical cases. Since USMLE covers almost every aspect of medical practice, our random samples have a broad coverage of medical topics. Ten colleagues served as assessors and independently evaluated iMed. None of them has formal medical training. In iMed, we compared the traditional techniques with the new techniques proposed in this paper. For each medical case, we randomly divided all ten users into two groups of the same size. When performing search, one group used iMed implemented with the traditional techniques while the other group used iMed implemented with the new techniques.

In our experiments, a user has up to 60 minutes to perform iterative search for each medical case. At the end of the search process, the user can list up to three diseases that he thinks best match the medical case's situation description. If any of these diseases is among the most likely candidate diagnoses accompanying the data set, the search is considered successful. We allow users to search for a relatively long time, because medical information searchers care about their health and often spend hours on searching. We allow users to list multiple diseases as their findings, because even doctors sometimes cannot make precise diagnosis without lab test results.

Similar to the TREC interactive track (TREC 2007), we use two sets of measures as the performance metrics for intelligent medical search engines: one set is objective while the other set is subjective. The objective performance measures include the success rate, the number of search iterations, the number of search result Web pages viewed, and the time spent on the search process. The subjective performance measures include the users' perceptions of ease of using the system, ease of understanding the system, usefulness of the search results, and overall satisfaction with the system. All these subjective performance measures are on a 7-point scale, with 1=low and 7=high (TREC 2007). They were obtained from a brief questionnaire that users filled out after using the systems. For each objective or subjective performance measure, we average it over all the 30 medical cases and all the users, and report both its mean and its standard deviation when appropriate. We used ANOVA (Bickel & Doksum 2001) as the significance test. Our experiments were performed on a computer with two 3GHz processors, 2GB memory, and one 111GB disk.

Table 1. Objective performance measures (* means significant at <0.05 level).

mean (standard deviation)	traditional techniques	new techniques
success rate	29% (12%)	35%* (10%)
number of iterations	3.7 (1.1)	2.9* (1)
number of search result Web pages viewed	14.7 (6.3)	11.3* (5.8)
time (minutes)	39 (11)	30* (10)

Our overall results are as follows. The new techniques are efficient. For all the 30 medical cases, the average time taken by the new techniques to generate each part of the search result hierarchy is less than two seconds. The new techniques are much more effective than the traditional techniques in finding the correct diagnosis, where most of the user's time is spent on reading the search result Web pages. The objective performance measures in Table 1 show that the new techniques help the user find results in fewer iterations, view fewer search result Web pages, spend less time on the search process, and achieve a higher success rate. All these differences are statistically significant.

Table 2. Subjective performance measures (* means significant at <0.05 level).

mean (standard deviation)	traditional techniques	new techniques
ease of using	4.9 (1.1)	5.7* (1.2)
ease of understanding	5.8 (1.0)	5.6 (1.2)
usefulness	5.2 (0.9)	6* (0.9)
satisfaction	5.1 (1.0)	5.8* (0.9)

Table 2 shows the subjective performance measures. As it takes time to become accustomed to navigating the search result hierarchy in the new output interface, users consider the traditional output interface slightly easier to understand than the new output interface, while the difference is not statistically significant. Nevertheless, once users understand the new output interface, they can use it without difficulty. The new output interface has explicitly marked medical meanings and organizes together all the search results on the same topic or aspect so that users can find them easily. Consequently, users feel that the new output interface is easier to use than the traditional output interface. Overall, the new techniques improve user satisfaction as they help produce more useful search results. These differences are statistically significant.

4. Conclusion

This paper presents an intelligent search result output interface for intelligent medical search engines, where all the search results are structured into a multi-level hierarchy that has explicitly marked medical meanings. Our techniques take into account the unique requirements of medical search and utilize domain-specific medical knowledge. Our experiments with a wide variety of medical cases and an implementation in iMed demonstrate that our techniques significantly improves the quality of search results and the speed that searchers find desired information.

Acknowledgements

We thank Haiyan Chen, Chunqiang Tang, Selena Thomas, Ying-li Tian, and Rong Yan for helpful discussions.

References

Bickel, P.J., and Doksum, K.A. 2001. *Mathematical Statistics: Basic Ideas and Selected Topics*, Vol. 1. Prentice Hall.

Collins, R.D. 2002. *Algorithmic Diagnosis of Symptoms and Signs: Cost-Effective Approach*. Lippincott Williams & Wilkins.

Fox, S., and Fallows, D. 2003. Internet health resources: Health searches and email have become more commonplace, but there is room for improvement in searches and overall Internet access. www.pewinternet.org/pdfs/PIP_Health_Report_July_2003.pdf.

Healthline homepage 2007. www.healthline.com.

Le, T., and Bhushan, V. 2006. *First Aid for the USMLE Step 2 CS (Clinical Skills Exam)*. McGraw-Hill.

Luo, G. 2008. iMed: an intelligent medical Web search engine. Available at pages.cs.wisc.edu/~gangluo/imed.pdf.

Luo, G.; Tang, C.; and Yang H. *et al.* 2007. MedSearch: A specialized search engine for medical information. WWW'07: 1175-1176.

MeSH homepage 2007. www.nlm.nih.gov/mesh/meshhome.html.

Merck Manual Home Edition homepage 2007. www.merck.com/mmhe/index.html.

Sherman, C. 2005. Curing Medical Information Disorder. searchenginewatch.com/showPage.html?page=3556491.

TREC interactive track homepage 2007. trec.nist.gov/data/interactive.html.

WebMD homepage 2007. www.webmd.com.