12-1-2001

Meeting Medical Terminology Needs: The Ontology-enhanced Medical Concept Mapper

Gondy Leroy
Claremont Graduate University

Hsinchun Chen
University of Arizona

Meeting Medical Terminology Needs—The Ontology-Enhanced Medical Concept Mapper

Gondy Leroy and Hsinchun Chen

***Abstract*—This paper describes the development and testing of the Medical Concept Mapper, a tool designed to facilitate access to online medical information sources by providing users with appropriate medical search terms for their personal queries. Our system is valuable for patients whose knowledge of medical vocabularies is inadequate to find the desired information, and for medical experts who search for information outside their field of expertise. The Medical Concept Mapper maps synonyms and semantically related concepts to a user's query. The system is unique because it integrates our natural language processing tool, i.e., the Arizona (AZ) Noun Phraser, with human-created ontologies, the Unified Medical Language System (UMLS) and WordNet, and our computer generated Concept Space, into one system. Our unique contribution results from combining the UMLS Semantic Net with Concept Space in our deep semantic parsing (DSP) algorithm. This algorithm establishes a medical query context based on the UMLS Semantic Net, which allows Concept Space terms to be filtered so as to isolate related terms relevant to the query. We performed two user studies in which Medical Concept Mapper terms were compared against human experts' terms. We conclude that the AZ Noun Phraser is well suited to extract medical phrases from user queries, that WordNet is not well suited to provide strictly medical synonyms, that the UMLS Metathesaurus is well suited to provide medical synonyms, and that Concept Space is well suited to provide related medical terms, especially when these terms are limited by our DSP algorithm.**

***Index Terms*—Ontologies, parsing, query expansion, semantic parsing, terminology mapping, UMLS.**

I. Introduction

MEDICAL sites are among the most popular Internet sites today [1]. The practice of medicine is experiencing a shift from patients who passively accept their doctor's orders to patients who actively look online for information that concerns their health. Most of the medical web sites, such as the Mayo Clinic Health Oasis[1] and MedScape,[2] are consumer oriented and provide their users with sound advice and information about general medical topics. The vocabulary used is readily comprehensible. However, when users search for more detailed information about a very specific topic, such as the use of alcohol injections to treat tumors, these sites do not provide the desired information and users have to search medical papers. Unfortunately, the users do not always share the same language with the researchers [2] and have problems finding the correct search terms because the professional and lay medical vocabularies are not always compatible.

The lack of vocabulary compatibility can be resolved by mapping a user query to relevant medical concepts. This kind of concept mapping is useful for two reasons. The first reason is that different groups of human searchers use different vocabularies. Patients seldom know the necessary terminology to find relevant information. Also, medical librarians and researchers searching for information outside their domain of expertise might need help with search terms. In addition, suggesting terms can provide different formats of the terms. This is important since different online sources often use their own standardized vocabulary [3]. The second reason for medical concept mapping is that it can be automated and used for routine query expansion. Query expansion is the addition of terms to an original set of terms to retrieve more documents. McCray *et al.* [4] give several examples of the usefulness of this kind of query expansion. Our purpose is to build the Medical Concept Mapper and incorporate it in document retrieval tools.

In this paper, we describe the Medical Concept Mapper in detail [5]. According to Ingenerf's categorization [6], the Medical Concept Mapper is a terminology server because it provides high-level terminology services and it suggests standard terminology to users. It is designed to suggest cancer-related medical terminology based on a user's query and is useful for searching medical databases. Ingenerf [6] describes the following three basic services that a terminology server can provide to enhance semantic integration:

1) access to external literature and knowledge bases;
2) exchange of electronic patient data;
3) integration of data-driven decision support systems.

The Medical Concept Mapper provides the first service: it facilitates access to existing knowledge sources by suggesting medical terminology to users. This is accomplished through combining the Unified Medical Language System (UMLS) developed by the National Library of Medicine, Bethesda, MD, WordNet developed at Princeton University, Princeton, NJ, and the Arizona (AZ) Noun Phraser, and Concept Space developed by the Artificial Intelligence (AI) Laboratory, University of Arizona, Tucson. It is innovative because it is an in-depth inte-

Manuscript received August 22, 2000; revised March 28, 2001. This work was supported in part by the National Science Foundation/Advanced Research Projects Agency/National Aeronautics and Space Administration under the Digital Library Initiative, by the Oak Ridge Institute for Science and Education under the National Library of Medicine Toxicology and Environmental Health Research Participation Program, by the National Science Foundation under the Digital Library Initiative-2, by the National Cancer Institute, and by the National Institutes of Health.

The authors are with Management Information Systems, University of Arizona, Tucson, AZ 85721 USA (e-mail: gleroy@bpa.arizona.edu).

Publisher Item Identifier S 1089-7771(01)07414-3.

[1][Online]. Available: http://www.mayohealth.org

[2][Online]. Available: http://www.medscape.com

gration of manually created ontologies and computer-generated tools, the intertwining of which allows a synergy to surface that suppresses the weaknesses of the tools when used on their own. The resulting terminology is a combination of terms from a controlled language, the UMLS, and free language terms from WordNet and Concept Space. This unique combination will be beneficial to users since it was found that recall is higher for controlled languages, i.e., the UMLS, but precision is higher for free language searches [7], i.e., Concept Space.

The structure of this paper is as follows. In Section II, we discuss the components of the Medical Concept Mapper and how they have been used by others. In Section III, we describe exactly how the components are brought together in the Medical Concept Mapper and our deep semantic parsing (DSP) algorithm. In Section IV, we describe our research questions. In Section V, we describe two user studies and the results. The last section contains concluding remarks and future directions.

II. Background

A. Natural Language Processing With the AZ Noun Phraser

The motivation for implementing natural language processing techniques in document retrieval is that it allows users to frame their questions in a natural way. Noun phrasing, the extraction of noun phrases from free text, has been used in information retrieval to capture a "richer linguistic representation" of document content [8]. In a user query, these phrases can be seen as concepts representing the user's needs. Noun phrasing has the potential to improve document retrieval since it allows for the query's noun phrases to be matched with noun phrases present in documents.

The AZ Noun Phraser was developed at the University of Arizona's AI Laboratory to extract high quality phrases from textual data. For example, from the query "*Should we be screening for prostate cancer routinely with the PSA?*" it extracts "*screening,*" "*prostate cancer,*" and "*PSA.*" For a detailed description, we refer the reader to [9]. The version used here includes the UMLS SPECIALIST Lexicon from the National Library of Medicine to improve the extraction of medical phrases from text [9].

B. Human Generated Knowledge Sources—Ontologies

Ontologies provide consistent vocabularies and world representations necessary for clear communication within knowledge domains. The term "ontology" refers to the shared understanding of the domain of interest. It is a unifying framework; it embodies objects and concepts, their definitions and the relationships between them [10]. Ontologies range from being very general, such as EuroWordNet [11] and Cyc,[3] to being very domain specific, such as the Enterprise Ontology [12] for business communication and knowledge exchange. The UMLS is a very extensive and specific medical ontology.[4] WordNet is a general English language ontology.[5] Both the UMLS and WordNet will be discussed since they are both part of the Medical Concept Mapper.

[3][Online]. Available: http://www.cyc.com

[4][Online]. Available: http://umlsks.nlm.nih.gov/

[5][Online]. Available: http://www.cogsci.princeton.edu/~wn/

1) WordNet: WordNet [13] is an online accessible lexical ontology that contains approximately 95 600 different word forms. These word forms are organized into 70 100 word meanings. Each meaning consists of a set of synonyms and a descriptive gloss explaining that sense of the word. For example, in WordNet, "*injection*" has three senses: "*injection as the forceful insertion of a substance under pressure (no synonyms), injection as any solutions that is injected (as into the skin) (synonym: injectant), and injection as the act of putting a liquid into the body by means of a syringe (synonym: shot).*" WordNet can be accessed online or can be freely downloaded.[6]

Serious difficulties are encountered when automatically selecting the correct word sense of a term [14]–[16]. The noun "*head,*" for example, has 30 different senses. However important it may be to add synonyms to a query, adding irrelevant phrases can have a detrimental effect. For example, one set of synonyms of "*blood*" is "*rake, profligate, rip, roue*" in the sense of "*a dissolute man in fashionable society*" [17]. When doing cancer-related research, these terms are not useful to expand a query with.

2) UMLS: The UMLS is a long-term project of the National Library of Medicine and is specifically developed to enable new information technologies to take advantage of controlled medical vocabularies [18]–[20]. The UMLS consists of four different components: the Metathesaurus, the Semantic Net, the SPECIALIST Lexicon, and the online Knowledge Sources Server. We use the Metathesaurus for its synonyms and the Semantic Net is part of our DSP algorithm. The SPECIALIST Lexicon is incorporated in the AZ Noun Phraser [9]. We use a local copy of the UMLS and not the online Knowledge Sources Server. We will refer to the components simply as the SPECIALIST Lexicon, the Metathesaurus, and the Semantic Net.

Many proposals for implementation of the UMLS and ideas for improvement can be found in the literature [4], [21]–[23], but very few empirical-data-driven studies have been done. In addition, most of the tools that use UMLS components [22], [24], [25] are useful only for users with at least a comfortable level of medical-domain knowledge. The Medical Concept Mapper is intended for people with little medical-domain knowledge such as patients, or only limited-domain knowledge, such as physicians or librarians looking for information outside their field of expertise.

The Metathesaurus and SPECIALIST Lexicon can be implemented in a straightforward fashion as lexicons. The implementation of the Semantic Net is more complicated. Thus far, completely automated tools that make use of the Semantic Net are scarce because of its structure. The Semantic Net contains semantic types and concepts belonging to that type. For instance, the concepts "*tyrosynase peptide*" and "*Helix-Turn-Helix Motifs*" belong to the semantic type "*Amino Acid Sequence.*" A concept can have different semantic types. The Semantic Net also contains semantic relations between the semantic types. For instance, there exists a "*causes*" relation between the type "*bacterium*" and "*neoplastic process.*" The difficulty with the Semantic Net arises because the relations

[6][Online]. Available: http://www.cogsci.princeton.edu/~wn/

exist between semantic types, but not necessarily between the concepts that belong to that type. For example, the semantic type "*Medical Devices*" has a "*treat*" relation with the semantic type "*Sign and Symptom*." However, not every concept belonging to "*Medical Device*" will "*treat*" every concept belonging to "*Sign and Symptom*." "*Bone screws*" (Medical Device) do not treat "*nausea*" (Sign or Symptom).

There have been different approaches that circumvent this difficulty with the Semantic Net. Cimino *et al.* [24] used a set of predefined generic queries. User queries are mapped to their equivalent generic query by means of natural language processing or by using query constraints. Once the queries are matched, the appropriate information sources are selected and the information is retrieved. The advantage of this approach is that all UMLS Knowledge Sources can be optimally used once the user query is mapped to a generic query. The disadvantage is that this tool is based on a limited set of generic queries and a good match between the user and generic query is necessary. In another approach, Robert *et al.* [25] and Joubert *et al.* [22] let users build the structure of the query by selecting the concepts, their semantic types, and the semantic relations between those concepts. The result is a conceptual graph for a query. These graphs can be compared against graphs of the information source, such as patient records, to find valid matches. The advantage of this approach is that potential matches can be limited based on the necessary underlying structure. The disadvantage is that the end-users must decide on the medical relevance of their graphs. Usage will be limited to experts with the necessary knowledge of the medical concepts and the validity of relations between them.

C. Automatically Generated Knowledge Sources—Concept Space

Concept Space was developed at the University of Arizona's AI Laboratory to facilitate semantic retrieval of information, and is accessible online.[7] It is an automatically generated index similar in functionality to a human generated thesaurus, but it is based on document term co-occurrence analysis. The related terms it provides can be used for term suggestion or for query expansion. The terms can consist of single or multiple words. For example, related concept space terms of "*colon cancer*" are "*colonic neoplasm*," "*colorectal cancers*," and "*colorectal neoplasm*." Weights between concepts establish the strength of association. In several studies, Concept Space was shown to improve searching and browsing. In the biosciences, Concept Space was successfully applied to the Worm Community System (WCS) [26], [27] and the FlyBase experiment [28]. There have also been successful results in the Digital Library Initiative studies conducted on the INSPEC collection for computer science and engineering [27], [29] and on Internet searching [30]. In the medical domain, Concept Space successfully aided medical researchers accessing the National Cancer Institute's CancerLit collection [31]. It is this medical Concept Space that is used by the Medical Concept Mapper. For a detailed description on how Concept Space is built, we refer to [32].

[7][Online]. Available: http://ai.bpa.arizona.edu/go/medical/cancerspace.html

III. Medical Concept Mapper—System Design

Users can submit any cancer-related medical query to the Medical Concept Mapper and receive synonyms and related terms relevant to their queries. The Medical Concept Mapper processes the queries in three consecutive phases. During the first phase, the AZ Noun Phraser augmented by the SPECIALIST lexicon is used to extract the medical phrases from the natural language queries. During the second phase, synonyms are retrieved based on WordNet and the Metathesaurus. During the third phase, related terms are retrieved based on Concept Space and the Semantic Net. The three phases and the components involved in each are discussed below.

A. Phrases

Users can submit their queries in two different ways. The first possibility is that they submit a natural language query. In this case, the AZ Noun Phraser augmented by the SPECIALIST Lexicon is used to extract the correct medical phrases from the query. For instance, from the query "*Which test (culdocentesis or pelvic ultrasound) would be best for diagnosis of ovarian cyst in this case?*" the AZ Noun Phraser extracted *culdocentesis, pelvic ultrasound, diagnosis, ovarian cyst,* and *case.* The second possibility is that users submit terms (single or multiple words) to the system. In this case, the user's terms are accepted as they are.

B. Synonyms

Two components are used to provide synonyms. The first component is WordNet. It is a very valuable tool for query expansion when the synsets, synonyms belonging to a certain sense of a term, are manually selected [16]. Automated disambiguation of the synsets was proven to be too difficult [14], [16], [33], [34]. Since erroneously disambiguated senses have a negative impact on document retrieval [34], we chose to use the WordNet synonyms only when there was exactly one synset for the term. This limits the power of WordNet severely, but we expected the precision of the term set to be unaffected and hoped that it would be able to leverage the Metathesaurus.

The second component that provides synonyms is the Metathesaurus. In the Metathesaurus, *terms* and *concepts* are different entities. A concept is the underlying meaning of a set of terms. As such, each concept can be expressed by many different terms. For example, the concept "*Cancer*" has 20 terms associated with it, two of which are "*Malignant Tumor*" and "*malignant tumoral disease*." We consider all terms that belong to the same concept to be synonymous with two exceptions. Terms that consist of an abbreviation followed by the full-text term are excluded since the full-text term is also provided separately. For instance, "*ng—new growth*" is excluded since "*new growth*" is already in the synonym list. We did not distinguish between terms that have a different meaning, indicated by a number, in different source vocabularies. For instance, "*growth*" is associated with "*Growth* ⟨1⟩" and "*growth* ⟨2⟩." In this case, only "*growth*" is retained.

The Medical Concept Mapper can provide three different sets of synonyms. All the terms extracted from the query or given by the user are used to retrieve the synonyms. The first possible synonym set contains only WordNet terms. For instance, for the query, "*This patient has a blue lesion on her stomach, what is it?*" the AZ Noun Phraser extracted three terms: *patient, blue lesion*, and *stomach*. There are three WordNet synonyms: *venter, stomach*, and *belly*. The second possible synonym set contains only Metathesaurus terms. For the same query, there are nine Metathesaurus synonyms, three of which are *abdominal region, malignant neoplasm of abdomen*, and *malignant tumor of abdomen*. The third possible synonym set includes synonyms from both WordNet and the Metathesaurus. In this case, the WordNet synonyms are used as additional input to the Metathesaurus. This means that, for the original terms, the WordNet synonyms are extracted first. The original terms together with the WordNet synonyms are then used to extract synonyms from the Metathesaurus. For the query above, this resulted in 13 extra Metathesaurus synonyms (total 22), three of which are *stomachs, ventriculus*, and *benign neoplasm of stomach*.

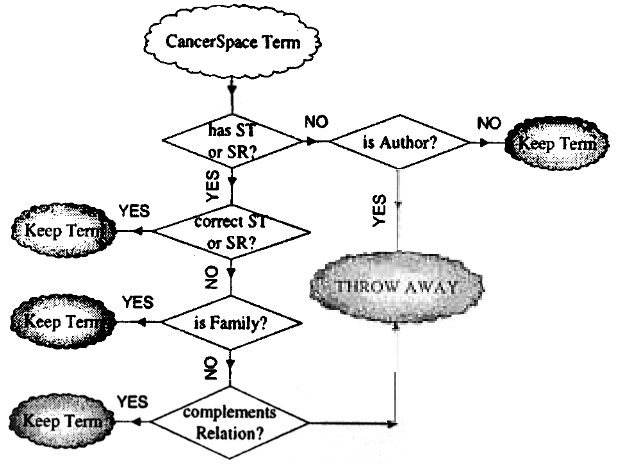

Fig. 1. Limitation algorithm (ST = semantic type, SR = semantic relation).

C. Semantically Related Concepts

Synonyms do not represent different concepts. They only provide different ways of saying the same thing. Co-occurrence terms can be very useful for mapping nonsynonymous, but semantically related, concepts to a query. However, co-occurrence terms are often too general [35] and can have a negative effect on the performance of document-retrieval systems. The Medical Concept Mapper provides relevant and precise related terms. This is done by combining a user query with related terms extracted from Concept Space and filtered by our DSP algorithm.

In general, the Medical Concept Mapper retrieves related co-occurrence terms from Concept Space for each input term (original and synonyms) and uses the Semantic Net to filter these co-occurrence terms. This is done by our DSP algorithm, which builds a context for each query based on the Semantic Net and uses this context to limit the related terms. Related terms will be more precise if they fit in the context. Most projects thus far did not accomplish automated context building at runtime [22], [24], [25] because of the ambiguous structure of the Semantic Net. As explained above, the relations in the Semantic Net exist between semantic types, not between the concepts that belong to that type. However, the ambiguity of the Semantic Net is problematic when it is used on its own and in a top-down fashion. In this case, the terms have to be disambiguated for correct selection to take place and before the terms can be added to a query. In the Medical Concept Mapper, the Semantic Net is used in an innovative way: it is used to *limit* concepts proposed by Concept Space, not to add terms.

1) DSP Algorithm—Establishing the Query Context: The context of a user query is established based on the semantic types and relations of its extracted phrases. These semantic types and relations are retrieved from the Semantic Net for all terms. We call these the context types. They are used for filtering in subsequent phases. Different combinations of context types are possible, but they are all made up of standalone or associated types. Standalone types are context types that are not related to other context types. Associated types are context types that are associated with other context types. They can either be two semantic types that have a semantic relation between them, or a semantic type associated with a semantic relation. For instance, the query "*Is there any connection between breast cancer and iodine?*" has three terms: *connection, breast cancer*, and *iodine*. The related context types are (in order): "*Intellectual Product*," "*Neoplastic Process*," and "*Pharmacologic Substance*" or "*Element, Ion, or Isotope*." "*Intellectual Product*" has no semantic relations in the Semantic Net with any of the other context types; it is a standalone context type. "*Neoplastic Process*" and "*Pharmacologic Substance*" are related in the Semantic Net. "*Neoplastic Process*" and "*Element, Ion, or Isotope*" are also related. These context types are associated types.

2) DSP Algorithm—Extracting and Limiting Concepts: For each term (original and synonyms) up to 40 Concept Space terms are extracted. All of these terms are submitted to the Metathesaurus to retrieve their semantic type or relation. Once all existing semantic types and relations are retrieved, the limitation process starts. For an overview, see Fig. 1.

Terms without a semantic type are retained because otherwise the system's vocabulary would be limited to the UMLS vocabulary since it was found to be insufficient as a substitute for a complete medical lexicon by Johnson *et al.* [36]. For instance, terms such as "*percutaneous ethanol injection therapy*" are important and precise, but are not part of the Metathesaurus. In addition, we know which Concept Space terms are author names. As such, these can be retained or discarded, depending on the algorithm's settings. For this study, the names are discarded. Concept space terms with a semantic type or relation are then subjected to limitation based on Identity, Family Tree, and Relation regulations.

To check on the *Identity*, it is determined if the Concept Space term has the same semantic type or relation as any of the context types. If this is the case, the term is retained. Otherwise, the term proceeds to the next step. For example, if the context consists of the type "*Amino Acid Sequence*," terms such as "*peptide sequence*," and "*lipoaminoacid*" are retained because they have the same semantic type.

To check the *Family Tree*, we first build the is-a hierarchy of the context types. In general, a Concept Space term is retained if it has a semantic type or relation that belongs to the is-a hierarchy of any of the context types. For instance, if the context

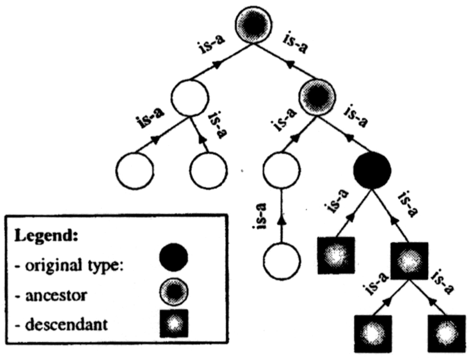

Fig. 2. Family tree of semantic types.

consists of the type "*Organizm*," terms such as "*acetobacter*" and "*alcaligenes*" will be retained because they belong to the "*Organizm*"-family. They have the semantic type "*Bacterium*," and a Bacterium is an Organizm.

The algorithm is programmed to allow different strictness options regarding the Family Tree. The first option concerns terms with multiple semantic types. If a related term has to be "strictly family," it means that all semantic types assigned to that term should belong to a family tree of a context type. Otherwise, it is enough if any (not all) of the semantic types belong to a family of a context type. The second option concerns the place of the term in the family tree. Terms can be descendants or ancestors of a context type. For an overview, see Fig. 2. If a related term has to be a "descendant," it means that the term should be below the context type in the is-a hierarchy. These are the more specific terms. If a related term has to be an "Ancestor," it means that the term should be above the context type in the is-a hierarchy. These are the more general terms. If a Concept Space term fulfills the requirements set, it is retained; otherwise, it goes to the next step.

To check for *Related Terms*, it is determined whether the Concept Space term has a semantic type related to any of the context types, or whether it is a semantic relation associated with any of the context types. If this is the case, it is retained. If, for example, the context types are "Acquired Abnormality" and "result of," then the term "brain injury" with semantic type "Injury or Poisoning" is retained because the semantic types are related: an "*Acquired Abnormality*" can be the "*result of*" an "*Injury or Poisoning*."

All Concept Space terms that comply with any of the rules mentioned above are collected. This results in a list of terms that is submitted to the next phase: reordering and filtering by means of weights.

3) DSP Algorithm—Reordering and Selecting Terms: After filtering the Concept Space terms, the list of possible useful terms includes terms without a semantic type or relation, terms with the same type as a context type, terms belonging to the family tree of a context type, or terms related to a context type. If there are duplicate terms, their weights are added. As explained above, the weights represent the strength of association in Concept Space. This addition results in a list with only unique terms with sometimes very high weights. For instance, if a term with weight 10 appears five times, the resulting weight is 50. The final list is reordered based on the new weights. A higher weight means a better term. A subset of this reordered list is retained for query expansion. Four different modes can be chosen to select this subset: the top x elements, the top x% element, all elements with weight x, or simply all elements.

IV. Research Questions

We addressed two sets of research questions with our user studies, starting from the observation that people do not always use optimally phrased queries when searching for medical information. These suboptimal queries may be ungrammatical, contains spelling errors, or contain irrelevant information, and as such influence the ability of our tool. To check for the possible effect of this, we thought it necessary to look at the differences in outcome between natural language queries that people used to ask each other questions, versus natural language queries without conversational information, such as "*Doctor X asked if* ...," and versus queries represented by phrases extracted from the original queries by a human search expert.

In addition, we were interested in the contribution of each component of the Medical Concept Mapper in providing useful search terms relevant to the query at hand. More specifically: How do phrases extracted by the AZ Noun Phraser compare to search terms provided by human search experts? How do synonyms provided by WordNet or the Metathesaurus compare to synonyms provided by human search experts? How do related concepts provided by Concept Space compare to related concepts provided by human search experts? Finally, how do Concept Space terms filtered by our DSP algorithm compare to unfiltered Concept Space terms?

V. User Studies

A. Setup

Queries and Input Method: Thirty cancer-related natural language queries were used to test the Medical Concept Mapper. The queries were selected from three sources. The first was a set of more than 1500 queries generated by medical doctors for usage with the UMLS [37]. Medical librarians submitted additional queries via email and two queries came from a journal article by Hersch and Hickam [38]. All queries were submitted to the Medical Concept Mapper in three different ways. First, they were submitted in their original state, without any alterations. For example, "*What causes fibroids and what would cause them to enlarge rapidly (patient asked Dr. B and she did not know).*" Second, the queries were submitted as cleansed queries, which means they were corrected for spelling errors and, if necessary, rephrased for grammatical correctness. No information was added, but unnecessary conversational information was omitted. For example, the aforementioned query was transformed to "*What causes fibroids and what would cause them to enlarge rapidly?*" Third, they were submitted by means of representational search terms. These terms were extracted directly from the query and were not altered in any way, e.g., "*fibroids*" was extracted from the query above. Both the cleansing of the queries and the selection of relevant search terms were done by a medical expert.

Gold Standard for Queries: To test the system's performance in suggesting search terms, synonyms, and relevant concepts, we needed a standard to compare the system provided terms against. Four experts agreed to provide this standard. There were two medical librarians and two cancer researchers. The cancer researchers formed one group, the medical librarians another. Each group established a list of search terms for each query. This set will be referred to as the Gold Standard for each query. The two members of each expert group worked together on this set without consulting any other source of information. The experts agreed on all sets of concepts within each expert group. However, there was an important difference between the two groups. The medical librarians stated that plurals of their search terms were always useful. If they provided the term "tumor," then "tumors" was also considered relevant. In contrast, cancer researchers stated explicitly that their exact strings should be used for information retrieval. For example, "breast cancer" could not be substituted for "breast cancers." They provided a list of terms that they wanted to be matched exactly. This difference between the two expert groups affected how we scored the Medical Concept Mapper's output.

The Gold Standard was used to calculate precision and recall of the system retrieved terms for each query. The precision and recall were calculated for the two expert groups separately. *Recall* is the percentage of Gold Standard terms that are actually retrieved by the system. *Precision* is the percentage of retrieved terms that appear in the Gold Standard.

An independent medical terminology expert did all the scoring. For the medical librarians, precision and recall were calculated based on a conceptual representation and based on a string representation. The conceptual representation score took into account that the librarians did not necessarily want the Medical Concept Mapper to provide exactly the strings they proposed. The expert strictly adhered to the following four rules to accomplish the conceptual scoring:

1) plurals and singulars were both sufficient to represent a term, e.g., "neoplasm" for "neoplasms";
2) abbreviations were sufficient to represent a term, e.g., "bcc" for "basal cell carcinoma";
3) if phrases differed only in word order disregarding prepositions and punctuation, the phrase was also accepted, e.g., "cancer of breast" for "breast cancer";
4) some very close synonyms or more specific terms were accepted, e.g., "treatment" for "therapy."

For the conceptual scoring, recall only improved when more terms representing different concepts were added. Adding different spellings of a term did not negatively affect precision, but did not positively affect recall either. For example, if the librarians' Gold Standard contained "breast cancer" and the key phrase extracted by the AZ Noun Phraser from the query was "breast cancer," then the Metathesaurus synonym "breast cancers" neither improved recall nor affected precision negatively. For string precision and recall, the same rules as for the cancer researchers applied (see below). This was done mainly to make a direct comparison between both groups possible.

For the cancer researchers, string recall and precision were used. This means that an exact match between strings was necessary before a term was accepted as correct. For instance, "Bone Marrow Transplant" would not be correct if their Gold Standard contained "Bone Marrow Transplantation."

Phrase Extraction: The AZ Noun Phraser was used to extract medical phrases from the original and cleansed queries. It was not used with term input since, in this case, the queries were already rephrased with terms.

Synonyms Expansion: The queries were expanded with synonyms for all three input methods in three different ways: WordNet synonyms only, Metathesaurus synonyms only, and the combination of WordNet and Metathesaurus synonyms. In this last case, the WordNet synonyms were used together with the extracted phrases to find Metathesaurus synonyms.

Related Concepts Expansion: Two methods were used to expand queries with Concept Space terms for all three input methods, but only for the best synonym expansion method. The first expansion method used Concept Space terms without any limitation from the DSP algorithm. Terms were reordered and the best subset was selected. The second expansion method relied on the DSP algorithm to limit the terms before reordering and selecting the subset.

B. Execution

Before we ran the Medical Concept Mapper, a medical expert selected phrases from each query for the Term Input condition. Phrases could only be "cut" from a query; they could not be altered in any way. The 30 queries were then submitted to the Medical Concept Mapper according to the three input methods and the three phases. The Medical Concept Mapper only expands the term list with new terms in each phase. If, for instance, the term "breast cancer" is extracted from the query, "breast neoplasm" can be added as a synonym. However, if Concept Space provides the same term, it cannot be added again since it is already provided as a synonym. By allowing only the addition of new terms in each phase, we could calculate the *additional* benefit of each phase in comparison to the human provided Gold Standards. Running the Medical Concept Mapper only required a human to type the query and select the correct experimental options.

C. Results

There was an enormous difference in the Gold Standards composed by the two expert groups. The cancer researchers tended to give a very small number of terms. The medical librarians gave much longer lists. For example, for the query "*Would B12 help this patient feel better (on chemotherapy for breast cancer)?*" the medical librarians suggested 19 search terms and the cancer researchers suggested six search terms. For an overview, please see Table I.

Before presenting the results of our studies, we like to point out that the Medical Concept Mapper used only a few terms extracted from the query. The results should be seen relative to this starting point. Our aim was not to mimic humans, but to automatically expand queries with correct medical terminology. Therefore, we were not concerned by low recall of expert's terms, but by recall that did not improve from one expansion level to the other.

For each study, we present overview tables containing recall (R) and precision (P) percentages, and summary tables of the statistical analyzes for original (OQ) and cleansed queries (CQ)

TABLE I
NUMBER OF TERMS IN THE GOLD STANDARDS

30 queries:	Cancer Researchers	Medical Librarians
Max. Terms per Query:	9	39
Min. Terms per Query:	2	8
Average Terms per Query:	6.1	17.6

TABLE II
SYNONYMS COMPARED TO THE MEDICAL LIBRARIANS' GOLD STANDARD

Expansion Levels:		Conceptual Evaluation			String Evaluation		
			CQ	TI	OQ	CQ	TI
None	R		14	17	11	12	
			57	92	43	49	
WordNet	R	14	14	18	12	12	
	P	53	56	91	37	45	
Meta.	R	25	26	30	14	15	
	P	59	60	79	10	11	
Meta. + WordNet	R	26	26	30	15	15	
	P	58	60	79	10	11	

TABLE III
SYNONYMS—ANALYSES FOR MEDICAL LIBRARIANS' GOLD STANDARD

	Conceptual Evaluation		String Evaluation	
	R	P	R	P
Input Method				
	<.05	<.001	<.005	<.001
	A	A	A	A
	A	A	A	A
	B	B	B	B
	)1		<.05	<.001
			A	A
			AB	A
			AB	B
			B	B

and for term input (TI). For all Tukey comparisons, we used $\alpha = 0.05$. Compared conditions that are not significantly different from each other have the same letter; conditions that are significantly different from each other have different letters.

D. User Study with Medical Librarians

Evaluation of the Synonyms: Our conceptual and string evaluation (see Tables II and III) show that *recall improved with cleaner input*. In both evaluations, it was mainly the Term Input that resulted in higher recall compared to both the Original and Cleansed Query input. For example, for the conceptual evaluation, when no expansion was done, recall was 14% for Original Queries, 13% for Cleansed Queries, and 17% for Term Input. There was no interaction between the expansion level and the input method for recall. In both our evaluations, precision was affected by the input method. For example with a conceptual evaluation, precision was mainly higher for Term Input compared to both Original and Cleansed Queries: when no expansion was done, the precision was 54% for Original Queries, 57% for Cleansed Queries, and 92% for Term Input.

For both conceptual and string evaluation *recall improved with synonym expansion*. With a conceptual evaluation, the improvement came particularly at the Metathesaurus level. For instance, for the Original Queries, recall was 14% with and without additional WordNet synonyms. It increased to 25% when the Metathesaurus was used and 26% when WordNet was used to leverage the Metathesaurus. With a string evaluation, the pair-wise comparisons indicated that only the difference of the first three conditions compared to the fourth condition (Metathesaurus + WordNet) was significant. We found that *precision was not affect by the expansion method with the conceptual evaluation*. For example, for Original Queries, precision was 54% when there was no expansion, 53% when WordNet synonyms were added, 59% when Metathesaurus synonyms were added, and 58% when WordNet was used to leverage the Metathesaurus. *Precision dropped with a string evaluation*. It dropped when Metathesaurus synonyms were added. For instance, for Original Queries, precision was 43% when no expansion was done, 37% when WordNet synonyms were added, and 10% when the Metathesaurus synonyms were added, or when the Metathesaurus was leveraged by WordNet.

Evaluation of the Related Concepts: With the *conceptual evaluation* (see Tables IV and V), there was *no effect of the input method on recall*. For instance, when Concept Space terms were added, recall was 30% for Original Queries, 31% for Cleansed Queries, and 36% for Term Input. With the *string evaluation, recall improved with cleaner input*. In particular, Term Input resulted in higher recall compared to both Original and Cleansed Queries. For instance, for Concept Space terms limited by DSP, recall was 16% for both Original and Cleansed Queries, and 21% for Term Input. For both the conceptual and string evaluation, we found a main effect of the input method on precision. Term Input, in particular, was more precise than both Original and Cleansed Queries.

The conceptual evaluation showed that *recall improved with related concepts expansion*. Recall improved particularly by adding Concept Space terms. Increased filtering by DSP did not lower recall. For example, for Original Queries, recall was 25% for the synonym baseline, 30% when Concept Space terms were added, and 30% when the Concept Space terms were limited by DSP. In our conceptual evaluation, we also found a general effect of expansion on precision. When adding related terms, *precision dropped without DSP*. For instance, for the Original Queries, precision was 59% for the synonym baseline, 46% for Concept Space without DSP, and 52% with DSP.

E. User Study with Cancer Researchers

Evaluation of the Synonyms: As explained above, we only did a string evaluation (see Tables VI and VII) for the cancer researchers' Gold Standard. *Cleaner input resulted in higher recall*. In particular, Term Input resulted in higher recall compared to both Original and Cleansed Queries. For instance, when no expansion was done, recall was 22% for Original Queries, 23% for Cleansed Queries, and 31% for Term Input. There was also a general *effect of the input method on precision*. Term Input was especially more precise than Original and Cleansed Queries.

TABLE IV
RELATED CONCEPTS COMPARED TO THE MEDICAL LIBRARIANS' GOLD STANDARD

Expansion Levels:		Conceptual Evaluation OQ	CQ	TI	String Evaluation OQ	CQ	TI
Syns	R	25	26	30	14	15	19
	P	59	60	79	10	11	15
CS	R	30	31	36	16	16	21
	P	46	47	65	8	8	12
CSNet	R	30	31	34	16	16	21
	P	52	53	70	9	9	13

Syns = Synonyms from Metathesaurus, CS = Concept Space, CSNet = Concept Space + Deep Semantic Parsing based on the Semantic Net

TABLE V
RELATED CONCEPTS—ANALYSES FOR MEDICAL LIBRARIANS' GOLD STANDARD

Input Method	Conceptual Evaluation R	P	String Evaluation R	P
Main effect p	--	<.001	<.005	<.001
Original Query		A	A	A
Cleansed Query		A	A	A
Term Input		B	B	B
Expansion Level				
Main effect p	<.05	<.001	--	--
Syns	A	A		
CS	B	B		
CSNet	B	AB		

TABLE VI
SYNONYMS COMPARED TO THE CANCER RESEARCHERS' GOLD STANDARD

Expansion Levels:		String Evaluation OQ	CQ	TI
None	R	22	23	31
	P	29	32	59
WordNet	R	23	23	32
	P	24	29	51
Meta.	R	23	24	35
	P	6	6	11
Meta.+ WordNet	R	24	24	35
	P	5	5	9

TABLE VII
SYNONYMS—ANALYSES FOR CANCER RESEARCHERS' GOLD STANDARD

Input Method	String Evaluation R	P
Main effect p	<.001	<.001
Original Query	A	A
Cleansed Query	A	A
Term Input	B	B
Expansion Level		
Main effect p	--	<.001
None		A
WordNet		A
Meta.		B
Meta. + WordNet		B

TABLE VIII
RELATED CONCEPTS COMPARED TO THE CANCER RESEARCHERS' GOLD STANDARD

Expansion Levels:		String Evaluation OQ	CQ	TI
Syns	R	23	24	35
	P	6	6	11
CS	R	23	24	36
	P	4	4	8
CSNet	R	23	24	35
	P	5	5	8

TABLE IX
RELATED CONCEPTS—ANALYSES FOR CANCER RESEARCHERS' GOLD STANDARD

Input Method	String Evaluation R	P
Main effect p	<.001	<.001
Original Query	A	A
Cleansed Query	A	A
Term Input	B	B
Expansion Level		
Main effect p	--	<.05
Syns		A
CS		A
CSNet		A

Expanding terms did not result in higher recall, but it affected precision. Precision dropped especially when Metathesaurus synonyms were added. For instance, for Original Queries, precision was 29% when no expansion was done, 24% when WordNet synonyms were added, 6% when Metathesaurus synonyms were added, and 5% when WordNet was used to leverage the Metathesaurus.

We also found a significant interaction between the expansion and the input method for precision ($p < 0.001$). Precision was high for Term Input when there was no expansion (59%). It dropped to the same levels as the other input methods when synonyms were added: 11% when Metathesaurus was added, 9% when both the Metathesaurus and WordNet were added.

Evaluation of the Related Concepts: Cleaner input resulted in higher recall. Again, this was in particular due to a higher recall for Term Input compared to both Original and Cleansed Queries . There was also an effect of the input method on precision. *Term Input resulted in higher precision* compared to both other input methods. For instance, for the synonym baseline, precision for Term Input was 11% and 6% for both the Original and Cleansed Queries.

Expanding the term set with related concepts (see Tables VIII and IX) did not improve recall. There was a main effect of the expansion level on precision, but this was not attributable to any particular level.

VI. CONCLUSION AND FUTURE DIRECTIONS

It is clear that our two expert groups differ in how they perform searches. The medical librarians start out with an extensive list of terms, synonyms, and spelling variations. The cancer researchers used only a very limited number of terms. We can think of search strategies as a continuum with high-precision searches on one end and high-recall searches at the other [7].

Since the medical librarians provided many synonyms and alternative spellings for the same terms, they belong to the high-recall end. The cancer researchers belong to the precision group; they focused on precision and did not provide any additional search terms. A search with the terms selected by the medical librarians would result in an large list of documents. This set would probably have to be narrowed down. A search with the terms provided by the cancer researchers might not always result in documents being found. Additional synonyms would probably have to be added. This had a clear impact on the evaluation of the Medical Concept Mapper. When no synonyms are included in the standard, a system explicitly built to provide these and other terms, will not perform well on precision.

None of the experts used ungrammatical search terms, for example, they used "*breast cancer*" but never "*breast, cancer.*" The Medical Concept Mapper provided such terms as synonyms with a detrimental effect on string precision. We could improve precision dramatically by excluding these terms. However, if the query expansion were automated, these terms could improve recall of documents without actually affecting precision of document recall.

An interesting observation is that we found no differences between Original and Cleansed Queries. This suggests that our system is robust enough to select the same number of terms with the same precision regardless of the format of the query. It was expected that the unnecessary information in the Original Queries would generate irrelevant concepts. This was not the case.

The power of WordNet was severely limited in our experiments. Only synonyms of nouns with one word sense were used. This should have resulted in no effect on the precision of our term set. However, the WordNet synonyms received by the Medical Concept Mapper were not always correct. For example, "*ERT*" is a frequently used abbreviation in the medical domain that stands for "*Estrogen Replacement Therapy.*" In WordNet, there is only one sense for this term and the synonym provided is "*earth-received time.*" We did not expect WordNet to expand the query with many synonyms, but we hoped that it would be able to leverage the Metathesaurus. Unfortunately, this did not happen. WordNet cannot be used in this manner to help bridge the gap from general English (such as patients would use) to specific medical terminology (used in the information sources).

The user studies described matched terms provided by the Medical Concept Mapper to terms provided by human experts. Search terms provided by the Medical Concept Mapper, but not by the human experts, were treated as erroneous. *Post hoc* analyses showed this is not necessarily true. For example, for the query "*He has a mole on his back. Is it a seborrheic keratosis or an intradermal nevus?*" some of the related phrases suggested by the Medical Concept Mapper were *skin neoplasms, malignant melanomas, skin diseases, skin pigmentation.* Although these terms were not in the Gold Standards, they could be very useful in finding relevant information. However, these terms were not included in the Gold Standard and, consequently, they had a detrimental effect on precision. We conclude, therefore, that the precision of terms reported here is an underestimate. Additional user studies are needed to evaluate retrieval of actual documents based on these terms. A real-life example of this is the query "*the use of alcohol injections for liver cancer.*" The librarian who received this query told us that the appropriate key phrases for his search were "*percutaneous ethanol injection therapy*" and "*hepatocellular carcinoma.*" The Medical Concept Mapper suggested both of these in its last phase (Concept Space limited by DSP).

In general, we discovered that the AZ Noun Phraser can be used to extract search phrases from user queries, that the Metathesaurus is useful to provide synonyms, but that WordNet is not yet ready to bridge the gap between plain English and specific medical terminologies. In addition, we increased the precision of the related medical terms from Concept Space by combining it with the Semantic Net. Our study showed that the Medical Concept Mapper cannot mimic medical and information professionals, but that it can easily double the number of terms found in a user query by adding terms that professionals would use. This is a fair indication that it will be very helpful in locating documents of interest for less sophisticated users. We are currently incorporating the Medical Concept Mapper in a search agent for document retrieval and will use the best combination: Metathesaurus synonyms and Concept Space terms limited by DSP.

Acknowledgment

The authors would like to thank M. Briehl, University Medical Center, Tucson, AZ, and S. Stratton, University Medical Center, Tucson, AZ, for creating the Cancer Researchers' Gold Standard. The authors also thank G. Perry, Health Sciences Library, Tucson, AZ, and J. Middleton, Health Sciences Library, Tucson, AZ, for creating the Medical Librarians' Gold Standard, R. Anderson, Health Sciences Library, Tucson, AZ, for her continued support and R. Sewell, Health Sciences Library, Tucson, AZ, for evaluating queries and terms. The authors further thank A. McCray, National Library of Medicine, Bethesda, MD, for her continued support. The authors also acknowledge the National Library of Medicine and Princeton University, Princeton, NJ, for making UMLS and WordNet ontologies freely available to researchers.

References

[1] E. H. Shortliffe, "The evolution of health-care records in the era of the Internet," *Medinfo*, vol. 9, pp. 8–14, 1998.

[2] T. E. Doszkocs and R. K. Sass, "An associative semantic network for machine-aided indexing, classification and searching," presented at the 3rd ASIS SIG/CR Classification Res. Workshop, 1992.

[3] J. J. Cimino, "Vocabulary and health care information technology: State of the art," *J. Amer. Soc. Inf. Sci.*, vol. 46, pp. 777–782, 1995.

[4] A. T. McCray, A. R. Aronson, A. C. Browne, T. C. Rindflesch, A. Razi, and S. Srinivasan, "UMLS knowledge for biomedical language processing," *Bull. Med. Library Assoc.*, vol. 81, pp. 184–194, 1993.

[5] G. Leroy, K. M. Tolle, and H. Chen, "Customizable and ontology-enhanced medical information retrieval interfaces," *Methods Inform. Med.*, to be published.

[6] J. Ingenerf, "Telemedicine and terminology: Different needs of context information," *IEEE Trans. Inform. Technol. Biomed.*, vol. 3, pp. 92–100, June 1999.

[7] J. E. Rowley, "A comparison between free language and controlled language indexing and searching," *Inform. Services Use*, vol. 10, pp. 147–155, 1990.

[8] P. G. Anick and S. Vaithyanathan, "Exploiting clustering and phrases for context-based information retrieval," presented at the 20th Annu. Int. ACM SIGIR Res. Develop. Conf., Philadelphia, PA, 1997.

[9] K. M. Tolle and H. Chen, "Comparing noun phrasing techniques for use with medical digital library tools," *J. Amer. Soc. Inform. Syst.*, vol. 51, pp. 352–370, 2000.
[10] T. Gruber, "A translation approach to portable ontology specifications," *Knowl. Acquisition*, vol. 5, pp. 199–220, 1993.
[11] H. Rodríguez, S. Climent, P. Vossen, L. Bloksma, W. Peters, A. Alonge, F. Bertagna, and A. Roventint, "The top-down strategy for building EeuroWordNet: Vocabulary coverage, base concepts and top ontology," *Comput. Humanities*, vol. 32, pp. 117–159, 1998.
[12] M. Uschold, "Converting an informal ontology into ontolingua: Some experiences," presented at the ECAI '96 Ontol. Eng. Workshop, 1996.
[13] C. Fellbaum, *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press, 1998.
[14] K. J. Mock and V. R. Vemuri, "Information filtering via hill climbing, WordNet, and index patterns," *Inform. Process. Manag.*, vol. 33, pp. 633–644, 1997.
[15] M. A. Stairmand, "Textual context analysis for information retrieval," presented at the 20th Annu. Int. ACM SIGIR Res. Develop. Inform. Retrieval Conf., 1997.
[16] E. M. Voorhees, "Using WordNet for text retrieval," in *WordNet: An Electronic Lexical Database*, C. Fellbaum, Ed. Cambridge, MA: MIT Press, 1998, pp. 285–303.
[17] G. A. Miller, R. Beckwidth, C. Fellbaum, D. Gross, and K. Miller. *Introduction to WordNet: An On-line Lexical Database* [Online]. Available: http://www.cogsci.princeton.edu/~wn
[18] B. L. Humphreys and D. A. Lindberg, "The UMLS project: Making the conceptual connection between users and the information they need," *Bull. Med. Library Assoc.*, vol. 81, pp. 170–177, 1993.
[19] D. A. Lindberg, B. L. Humphreys, and A. T. McCray, "The unified medical language system," *Methods Inform. Med.*, vol. 32, pp. 281–291, 1993.
[20] A. T. McCray and S. J. Nelson, "The representation of meaning in the UMLS," *Methods Inform. Med.*, vol. 34, pp. 193–201, 1995.
[21] A. T. McCray, S. Srinivasan, and A. C. Browne, "Lexical methods for managing variation in biomedical terminologies," in *Proc. Annu. Comput. Applicat. Med. Care Symp.*, 1994, pp. 235–239.
[22] M. Joubert, M. Fieschi, and J. J. Robert, "A conceptual model for information retrieval with UMLS," presented at the 17th Annu. Comput. Applicat. Med. Care Symp., 1994.
[23] G. Carenini and J. D. Moore, "Using the UMLS semantic network as a basis for constructing a terminological knowledge base: A preliminary report," presented at the 17th Annu. Comput. Applicat. Medical Care Symp., 1994.
[24] J. J. Cimino, A. Aguirre, S. B. Johnson, and P. Peng, "Generic queries for meeting clinical information needs," *Bull. Med. Library Assoc.*, vol. 81, pp. 195–206, 1993.
[25] J. J. Robert, M. Joubert, L. Nal, and M. Fieschi, "A computational model of information retrieval with UMLS," in *Proc. Annu. Comput. Applicat. Med. Care Symp.*, 1994, pp. 167–171.
[26] H. Chen, B. R. Schatz, T. Yim, and D. Fye, "Automatic thesaurus generation for an electronic community system," *J. Amer. Soc. Inf. Sci.*, vol. 46, pp. 175–193, 1995.
[27] H. Chen, J. Martinez, D. T. Ng, and B. R. Schatz, "A concept space approach to addressing the vocabulary problem in scientific information retrieval: An experiment on the Worm Community System," *J. Amer. Soc. Inf. Sci.*, vol. 48, pp. 17–31, 1997.
[28] H. Chen and B. R. Schatz, "Semantic retrieval for the NCSA mosaic," in *1994 Proc. 2nd Int. World Wide Web Conf.*, [Online]. Available: http://archive.ncsa.uiuc.edu/SGD/IT94/Proceedings/Searching/chen/chenschatz.html.
[29] H. Chen, J. Martinez, A. Kirchhoff, T. D. Ng, and B. R. Schatz, "Alleviating search uncertainty through concept associations: Automatic indexing, co-occurrence analysis, and parallel computing," *J. Amer. Soc. Inf. Sci.*, vol. 49, pp. 206–216, 1998.
[30] H. Chen, A. Houston, J. Yen, and J. F. Nunamaker, "Toward intelligent meeting agents," *IEEE Computer*, vol. 29, pp. 62–70, 1996.
[31] A. L. Houston, H. Chen, B. R. Schatz, S. M. Hubbard, R. R. Sewell, and T. D. Ng, "Exploring the use of concept space to improve medical information retrieval," Int. J. Decision Support Syst., 1999, to be published.
[32] H. Chen and K. J. Lynch, "Automatic construction of networks of concepts characterizing document databases," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, pp. 885–902, Sept.–Oct. 1992.
[33] E. M. Voorhees, "Query expansion using lexical–semantic relations," presented at the 17th Annu. Int. ACM–SIGIR Res. Develop. Inform. Retrieval Conf., 1994.
[34] M. Sanderson, "Word sense disambiguation and information retrieval," presented at the 17th Annu. Int. ACM–SIGIR Res. Develop. Inform. Retrieval Conf., 1994.
[35] H. J. Peat and P. Willett, "The limitations of term co-occurrence data for query expansion in document retrieval systems," *J. Amer. Soc. Inf. Sci.*, vol. 42, pp. 378–383, 1991.
[36] S. B. Johnson, A. Aguirre, P. Peng, and J. Cimino, "Interpreting natural language queries using the UMLS," in *17th Annu. Comp. Applicat. Med. Care Symp.*, 1993, pp. 294–298.
[37] Z. Stavri, "Queries generated by medical doctors for usage with the UMLS," unpublished.
[38] W. R. Hersh and D. H. Hickam, "An evaluation of interactive Boolean and natural language searching with an online medical textbook," *J. Amer. Soc. Inf. Sci.*, vol. 46, pp. 478–489, 1995.

Gondy Leroy received the Master's degree in psychology from the Catholic University of Leuven, Leuven, Belgium, in 1996, the M.S. degree in management information systems from the University of Arizona, Tucson, in 1999, and is currently working toward the Ph.D. degree in management information systems at the University of Arizona.

Her research interests are knowledge discovery and the combination of semantics with information retrieval.

Hsinchun Chen received the Ph.D. degree in information systems from New York University, New York, NY, in 1989.

He is currently a McClelland Endowed Professor of management information systems at the University of Arizona, Tucson. He has authored or co-authored over 70 papers concerning semantic retrieval, search algorithms, knowledge discovery, and collaborative computing. He serves on the Editorial Board of the *Journal of the American Society for Information Science and Decision Support Systems*. He is an expert in digital library and knowledge management research, and his research has been featured in scientific and information technology publications.