A Concept-Based Framework for Retrieving Evidence to Support Emergency Physician Decision Making at the Point of Care

Dympna O'Sullivan[1], Ken Farion[2], Stan Matwin[1], Wojtek Michalowski[1], and Szymon Wilk[1]

[1] University of Ottawa, Ottawa, Canada
dympna,wojtek,wilk@management@uottawa.ca, stan@site.uottawa.ca
[2] Children's Hospital of Eastern Ontario, Ottawa, Canada
farion@cheo.on.ca

Abstract. The goal of evidence-based medicine is to uniformly apply evidence gained from scientific research to aspects of clinical practice. In order to achieve this goal, new applications that integrate increasingly disparate health care information resources are required. Access to and provision of evidence must be seamlessly integrated with existing clinical workflow and evidence should be made available where it is most often required - at the point of care. In this paper we address these requirements and outline a concept-based framework that captures the context of a current patient-physician encounter by combining disease and patient-specific information into a logical query mechanism for retrieving relevant evidence from the Cochrane Library. Returned documents are organized by automatically extracting concepts from the evidence-based query to create meaningful clusters of documents which are presented in a manner appropriate for point of care support. The framework is currently being implemented as a prototype software agent that operates within the larger context of a multi-agent application for supporting workflow management of emergency pediatric asthma exacerbations.

1 Introduction

Emerging in parallel with new technologies for health care is the recognition of evidence-based medicine: "the conscientious, explicit, and judicious use of current best evidence in making medical decisions" [1]. Although there is wide acceptance that the practice of evidence-based medicine is a necessary component of health care delivery there are barriers that impede its use. Paramount among these barriers is the problem of information overload in the medical literature with approximately 30,000 scientific articles published annually. Furthermore, a lack of effective decision support tools means that clinicians have neither the time nor the ability to access relevant evidence, especially not at the point of care. A further challenging aspect of retrieving evidence to support clinicians is that evidence contained within repositories of medical literature tends to emphasize a disease-oriented context while a patient-oriented context is more appropriate for point of care support.

The aim of this research is to develop a methodological framework that supports emergency physicians by providing relevant patient-oriented evidence at the point of care. As a patient is diagnosed by an emergency physician, different facets of patient information are automatically captured and combined to formulate a concept-based query with which to retrieve evidence from an online repository. These concepts relate to both the specified disease and to the particular patient presentation, where disease-specific concepts are used to reduce the search space of available documents, while patient-specific concepts are used to identify those documents that are most relevant for the current patient presentation. We employ a concept-based query mechanism for a number of reasons. Firstly, the method focuses attention on the logical content of information rather than entirely on its form. This is important for highly specialized corpuses such as medical literature where natural language processing and semantic understanding are difficult to achieve at any significant level of granularity. Secondly, concept-based retrieval systems can expand the semantic richness of searches to overcome problems of low precision often associated with text or web-based search methodologies [2].

A second issue addressed by this research is the presentation of retrieved evidence in a format appropriate for point of care support. Currently the presentation style favored by many textual information retrieval engines is a ranked list of retrieved information. However, such a style is not suited to environments where users may lack the necessary time to browse numerous results in order to locate relevant information. Our approach employs a cluster-based approach for document presentation to allow for faster visual discrimination of relevant evidence. When a final set of documents has been retrieved using a concept-based search, a number of clusters are created by automatically extracting query concepts as textual cluster labels and retrieved evidence is further processed and assigned to the appropriate clusters. In this research we focus on techniques for post-retrieval document clustering [3] to draw the cluster boundaries to partition the set of documents at hand rather than the complete corpus.

The rest of this paper is organized as follows. In the next section we describe the Cochrane Library of clinical evidence. In Section 3 we present MET-A^3Support-Asthma, our multi-agent application for supporting workflow management and triage of pediatric asthma, of which the agent presented in this paper is a component. In Section 4 we provide a comprehensive outline of our proposed conceptual framework for the retrieval of evidence and describe how the framework shall be implemented as an evidence-based agent in a multi-agent application for pediatric asthma. We conclude with a discussion and some future directions.

2 The Cochrane Collaboration

The Cochrane Collaboration is a international cooperation of physicians from academic centers that produces the Cochrane Library [4], comprising 8 different databases describing research results of randomized clinical trials as well as incorporating important research results from other repositories such as Medline and EMBASE. We focused on the Cochrane Library for a number of reasons. Firstly research is described using a standardized reporting structure. Secondly all reviews are revised every two years to ensure information is current and up-to-date. Finally the evidence in the repository focuses on more specific research questions, for example questions are of the form "What is the evidence for applying one therapy or another for a particular problem in a particular population?", rather than "What are all possible therapies that can be applied in treating a particular problem?"

In this work we focus on retrieving evidence from the Cochrane Database of Systematic Reviews and the Cochrane Central Register of Controlled Trials (Clinical Trials). The first repository contains systematic reviews which are concise summaries of the best available evidence from primary studies. The second database includes details of published articles describing clinical trials taken from bibliographic databases. Information in the Cochrane Database is indexed using Medical Subject Headings (MeSH). MeSH are a medical thesaurus and controlled semantic vocabulary that is part of the larger Unified Medical Language System (UMLS) thesaurus. MeSH descriptors are applied by physicians to articles in Cochrane databases as part of the review process where descriptors represent the central concepts outlined in each article. MeSH check words may also be applied by reviewers, where check words are terms that do not directly describe central concepts but rather are used to describe the content at a finer level of granularity.

3 MET-A^{3}Support-Asthma: Mobile Emergency Triage - Anytime and Anywhere Support Application

The evidence-based agent for document retrieval forms a part of broader research that is focused on developing a methodological framework, the MET-A^{3}Support-Asthma multi-agent application, to support complete workflow management and clinical decision making for pediatric asthma in emergency departments. Our research in this area has drawn on a body of distributed and multi-agent systems for health care. Specifically we have focused on applications for coordinated communication in health care organizations [5], patient information retrieval and workflow management[6], and distributed decision support [7].

In managing pediatric asthma, triage is a crucial task within the clinical workflow as early identification of the severity of an exacerbation has implications for the child's management in the emergency department. Patients with a mild attack are usually discharged home following a brief course of treatment and resolution of symptoms, patients with a moderate attack receive more aggressive treatment over an extended observation in the emergency department, and patients with a severe attack receive maximal therapy before ultimately being transferred to an in-patient hospital bed for ongoing treatment. In clinical practice, a decision on severity and subsequent disposition of an exacerbation is ideally made as soon as possible after arrival of the patient to the hospital to ensure key therapies have been instituted. Consequently emergency physicians should benefit from computer-aided decision support that provides an early prediction of the severity of an asthma exacerbation as well as summarized patient-oriented evidence as an aid to confirm and/or reinforce clinical decision making.

In the MET-A^{3}Support-Asthma application values of clinical attributes are input to an asthma agent that combines them to evaluate an asthma exacerbation severity. This evaluation of severity is available at the point of care and can be used by the physician to prescribe a treatment for the current patient. Both of these patient-specific pieces of information (severity assessment and treatment) are matched by the evidence-based agent with the underlying patient ontology and automatically translated into a patient-oriented concept-based query. This query is then combined with disease-oriented concepts that describe the current presentation, represented by MeSH ontological terms (keywords and check words) and used to mine the Cochrane Library. The methodologies used by the evidence-based agent to construct a concept-based query and to display clustered query results at the point of care are described in detail in the following section.

4 Implementation of the Evidence-Based Agent

4.1 Agent Design

The task of the evidence-based agent is to retrieve and identify appropriate evidence for a specific patient context from the Cochrane Library. The fulfillment of this task requires us to define the following characteristics of the agent who's architecture is shown in Figure 1:

Communication: Agents in the MET-A^{3}Support-Asthma application communicate using FIPA Agent Communication Language (ACL) and the complete framework is organized using capability-based coordination. When an asthma severity prediction and a treatment have been specified, these pieces of information are pushed to the evidence-based agent by two other software agents responsible for these respective tasks. When the task of the evidence-based agent is completed it pushes summarized results to another agent (interface agent) that is responsible for presenting information to the emergency physician.

Planning: The agent's planning component receives as an input a severity assessment and a recommended treatment. This patient-specific information is combined with disease-specific information by the agent planner to construct a concept-based query. Once relevant pieces of evidence are retrieved and organized, they are exported to a local database for indexing and further processing, resulting in a number of coherent clusters for presentation. The clustered results are then communicated to the agent communicator which passes the information to the interface agent for point of care presentation.

Execution: Once a concept-based query has been constructed it is passed to the agent executor who invokes the appropriate search. Tasks performed by the evidence-based agent must be completed in a controlled and systematic manner and the executor monitors scheduled tasks and invokes the agent planner at the appropriate time.

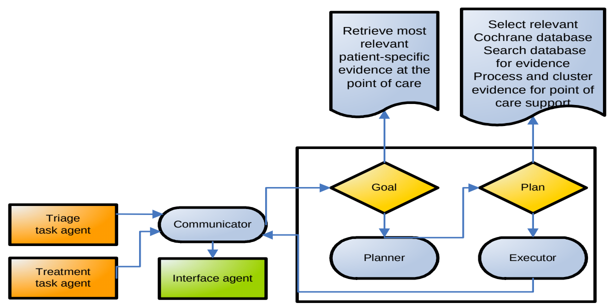

Fig. 1. Architecture of the Evidence-Based Agent

4.2 Planning and Executing an Evidence-Based Search

The evidence-based agent is a reactive agent and upon receiving information regarding the predicted severity of an asthma exacerbation and a recommended treatment, the agent's planning component formulates a plan with which to search for relevant evidence. This plan is then passed to the executor that initiates the search. The plan consists of a number of subtasks that must be executed in a sequential manner and which are outlined in the complete algorithm for retrieving evidence presented in Figure 2. Each step of this algorithm is described in detail in the following subsections.

4.2.1 Identify the Appropriate Database in Cochrane Library The first part of the agent's planning task is to identify the correct Cochrane Library to query for evidence. This is achieved using asthma clinical practice guidelines to steer the search for evidence towards the appropriate part of the library. Clinical practice guidelines for asthma are derived from research in the Cochrane Database of Systematic Reviews and outline a number of recommended treatments given different asthma exacerbation severities. We have extrapolated all possible severity and treatment combinations from the guidelines to create a lookup table that is consulted by the evidence-based agent. The agent planner logically combines the severity assessment and recommended treatment and checks the lookup table for the existence of such a combination. If the combination exists, the search is initiated on the Database of Systematic Reviews, otherwise it is directed towards the Central Register of Controlled Trials.

Identify the Appropriate Database in the Cochrane Library

Using asthma clinical practice guidelines and values of patient attributes, select the correct Cochrane database upon which to initiate search for relevant evidence

Formulate a Concept-Based Search

Instantiate disease (M) and patient concepts (P) with relevant instances from MeSH ontology and underlying patient ontology respectively

Mi == disease concepts, Pi == patient concepts

Formulate concept-based query (Con_Q) by combining Mi and Pi into a text-based search and by specifying the Cochrane textual index to be searched using query with Boolean operators

Con_Q = For all Articles where Mi [MeSH] is true
Patient_Specific_Query == Pi [Cochrane Full Text]

Retrieve ranked list of documents from the Cochrane Library and export to local database for indexing

Create Clusters for Retrieved Documents

Formulate clusters (C) with cluster labels (L) where documents from the Cochrane library will be stored by automatically extracting instances of P as cluster labels

For all combinations of Pi: “Pi” == Ci with label Li

Assign Retrieved Documents to Correct Clusters

Create textual indices for the documents retrieved from the Cochrane Library using a standard text-based indexing engine

Formulate multiple word text-based queries (Clus_Q) as comma delimited strings with which to cluster the retrieved evidence by extracting attribute names and attribute values for each P

For all combinations of Pi: Clus_Qi == “Pi.AttributeName, Pi.ValueName”

Pass query strings to textual search engine and assign retrieved evidence-based documents (E) to relevant clusters based on the discovery of the query strings within the documents

For all Clus_Qi: Ei == Ci with descriptor Li

Fig. 2. Algorithm for Retrieving Patient-Oriented Evidence

4.2.2 Formulate a Concept-Based Search When the correct database has been identified, the second part of the agent’s plan is to formulate a concept-based query with which to mine the library. This query is composed of two parts, one corresponding to the specified disease and another corresponding to the patient presentation. Concepts related to the disease (asthma) are represented by a combination of MeSH terms and check words that together describe the outlined illness at the lowest possible level of granularity. Patient-specific concepts are extracted from appropriate clinical attribute values in the underlying patient ontology. The search terms are bound to different indices of the

Cochrane Library where MeSH terms are used to reduce the search space to only those documents where the outlined disease constitutes a major topic of article. From such a reduced repository, patient-specific concepts are used to search the full textual Cochrane indices for each retrieved document. Therefore the concept-based search is of the form:

Search ((MeSH Keyword Index ("Asthma") AND MeSH Check Word Index("Child"))
Then, for all:
Articles where (MeSH Keyword Index ("Asthma") AND MeSH Check Word Index("Child") is true
Search (Cochrane Full Text Index ("Predicted Severity") AND (Cochrane Full Text Index ("Treatment Type")

4.2.3 Create Clusters of Retrieved Documents Using the query outlined above, the agent retrieves a list of evidence ranked by the Cochrane Library search engine. The list represents the smallest number of documents from the Cochrane library that are relevant for the particular patient presentation given a specified disease. The list cannot be further reduced by text processing techniques given difficulties in parsing the specialized medical texts. Therefore it is necessary to present all retrieved documents, however our framework aims to present the articles in a manner that allows for more effective visualization at the point of care by organizing and displaying retrieved documents in clusters according to concepts from the patient-specific query. Therefore the cluster labels must accurately represent the content of the grouped documents in order to allow physicians to quickly discriminate the most relevant articles. To facilitate such a logical partitioning of information, cluster labels are created by automatically extracting the clinical attributes used in the patient-oriented concept-based query as textual strings, where all possible logical combinations of these strings are used to create unique clusters. For example:

For all clinical attributes Pi, from the patient-specific query PS_Query:
Pi.StringValue == Cluster Pi with label Pi.StringValue

4.2.4 Assign Retrieved Evidence to Correct Clusters Considering that the documents have been retrieved by searching the full textual indices of the Cochrane Library, it is already known that instances of the patient-oriented query are contained within the documents. The task of assigning documents to relevant clusters involves identifying co-occurrences and frequencies of patients-specific concepts within the individual articles. The execution of this task is two-fold. Firstly, retrieved documents must be indexed using a search engine which implements standard methods for indexing and searching such as calculating word frequencies, and provides standard parsing functionality such as stemming and stop-word removal. Secondly, suitable textual queries must be formulated to search the generated textual indices.

The search engine we have chosen to index and search retrieved documents is Google Desktop. Google Desktop runs a local web server on a host machine and indexes the contents of files in a number of formats, including web pages, to provide full text search functionality. Retrieved documents from the Cochrane Library are exported as web pages to a local database on the host computer with an associated evidence-based session identifier. Once a complete set of documents for the session has been downloaded the agent planner initiates the indexing functionality of Google Desktop.

The second part of the task involves constructing textual queries that can be used to identify patient-specific concepts within the indexed articles. Multiple word queries are automatically constructed by the agent planner by the extracting attribute names from the patient ontology and corresponding values that are used in constructing the patient-oriented concept-based query. The queries are composed of extracted names and values represented as textual strings and combined using a logical AND operator.

For all clinical attributes Pi, from the patient-specific query PS_Query:
Clustering_Query == Pi.Name.StringValue AND Pi.Value.StringValue

The textual indices created for the retrieved documents are then searched by the Google Desktop search engine using the automatically formulated queries. As a Google Desktop search proceeds the application keeps count of the total matches for a particular document and each count links to a matching document. This functionality is used to assign documents to the correct clusters while the relevancy scores assigned to articles by the Cochrane Library search engine is maintained to rank the documents inside the clusters. Once all documents have been assigned to clusters the agent planner communicates the results to the interface agent. Documents are summarized by displaying only the title and conclusions from the systematic review of each article. Links are provided to the full systematic review of the article as well as to the full text version if the physician wishes to view this information or to save it for later use.

4.3 Illustrative Scenario To illustrate an operation of the evidence-based agent, a sample scenario is outlined. *A 10 year old boy experiencing an asthma exacerbation presents to the emergency department. The emergency physician records the patient's details and values of clinical attributes and based on this information it is determined that the patient is experiencing an exacerbation of moderate severity. The physician recommends the patient is treated with β-agonists and anticholinergics.* Information on severity and treatment is pushed to the evidence-based agent that uses it to plan the search. The first part of the agent's planning task is to identify the correct Cochrane Library to query. The treatment prescribed by the physician follows the asthma clinical practice guidelines for an exacerbation of moderate severity and so it is included in the lookup table of possible asthma severity and treatment combinations. Thus, according to agent's plan, the evidence-based search is directed towards the Cochrane Database of Systematic Reviews.

The next part of the agent's plan is to formulate a concept-based query. Considering that the patient is suffering a moderate asthma exacerbation and that the physician has recommended a treatment of β-agonists and anticholinergics the following combined query is formulated:

"asthma [MeSH Index] AND "child [MeSH Index]" AND "moderate [Cochrane Full Text Index]" AND "β-agonists [Cochrane Full Text Index] AND "anticholinergics [Cochrane Full Text Index]"

The next part of the agent's planning task is to create clusters with appropriate labels in which to store and present retrieved documents. Cluster labels are automatically extracted from the patient-oriented concept-based query and every possible combination of constituent elements is used to create clusters. For example, for the sample patient three labeled clusters are created as follows:

For all clinical attributes Pi, from the patient-specific query PS_Query:
Diagnosis, Treatment:
Pi.StringValue == Cluster Pi with label Pi.StringValue:
"Diagnosis", "Treatment", "Diagnosis and Treatment"

Assignment of the documents to clusters is performed by executing two separate subtasks. Firstly the articles retrieved from the Cochrane Database of Systematic Reviews are exported to a local database and indexed using Google Desktop. Secondly the created indices are searched using appropriate queries which are formulated by extracting the same patient-specific information from the underlying patient ontology that was used in developing the concept-based query. For the illustrative scenario the formulated query strings are:

"diagnosis moderate", "treatment β-agonists anticholinergics","diagnosis moderate treatment β-agonists anticholinergics"

where the strings are separated by spaces which behave as a logical AND. Upon discovery of these strings, the retrieved documents are assigned to the correct clusters. Finally the clusters are pushed to an interface agent who presents the evidence to the emergency physician at the point of care. An example of presenting retrieved evidence for the sample scenario is shown in Figure 3.

5 Conclusions and Future Directions

We have described a concept-based framework for designing a software agent that supports retrieval of clinical evidence for the point of care support. We are implementing the framework as an evidence-based agent that operates within the MET-A^{3}Support-Asthma multiagent system for supporting workflow management and clinical decision-making for emergency pediatric asthma exacerbations. In this research we retrieve evidence from the Cochrane Library, however, we recognize that our methodological framework for the retrieval of evidence

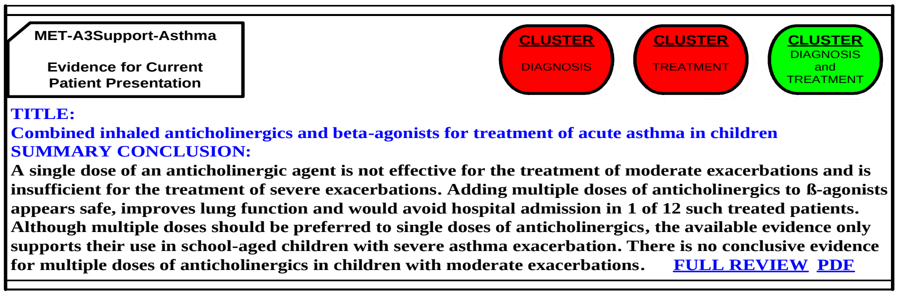

Fig. 3. Summarized Presentation of Evidence

may be applied across clinical presentations and systematized libraries by adjusting the underlying ontological schema. We are interested in extending the framework in a number of ways. We intend to supplement our clustering technique using semantic search methodologies such as latent semantic indexing for retrieving and organizing evidence (e.g. [8]). We are investigating methods for enhancing the capabilities of the evidence-based agent by extending conceptual aspects of the query using more comprehensive medical ontologies (e.g. GALEN).

Acknowledgements

The support of the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research is gratefully acknowledged.

References

1. Sackett, D., Rosenberg, W., Gray, J.M., Haynes, R., Richardson, W.: Evidence based medicine: what it is and what it isn't. British Medical Journal **312** (1996)
2. Guarino, N., Masolo, C., Vetere, G.: Ontoseek: Content-based access to the web. IEEE Intelligent Systems **14**(3) (1999) 70–80
3. Zamir, O., Etzioni, O.: Grouper: Dynamic clustering interface to web search results. Computer Networks **31** (1999) 11–16
4. www.cochrane.org
5. Marchetti, D., Lanzola, G., Stefanelli, M.: An ai-based approach to support communication in health care organizations. In: 8th Conference on Artificial Intelligence in Medicine in Europe. (2001) 384–394
6. Hannebauer, M., Muller, S.: Distributed constraint optimization for medical appointment scheduling. In: 5th International Conference on Autonomous Agents. (2001) 139–147
7. Lanzola, G., Gatti, L., Falasconi, S., Stefanelli, M.: A framework for building cooperative software agents in medical applications. Artificial Intelligence in Medicine (1999) 223–249
8. Cohen, S., Mamou, J., Knaza, Y., Sagiv, Y.: Xsearch: A semantic search engine for xml. In: 29th International Conference on Very Large Databases. (2003)