Terminology-Based Approach: Knowledge Acquisition and Integration in Healthcare

Subiksha K.P
Research Scholar
Madurai Kamaraj University
Madurai, India.

K.Alagarsamy
Associate Professor
Madurai Kamaraj University
Madurai, India.

ABSTRACT

Healthcare organizations are increasingly choosing the Knowledge Management Systems for clinical use, which have been established in technical support organizations for several years. Technical support organization can make use of its Knowledge Management System to meets its customer needs using its infrastructure and established processes. Success of Knowledge Management System in a healthcare organization mainly depends on the practice of healthcare professionals. Knowledge acquisition is very vital in healthcare than the application in technical support. The aim is to provide efficient access to heterogeneous biological textual data and databases, enabling users to integrate a wide range of textual and non-textual resources effortlessly. The main objective of the system is to make easier knowledge acquisition through query answering against XML-based documents in Healthcare specifically in the domain of molecular biology.

Keywords
Healthcare, Knowledge Acquisition, Terminology based.

1. INTRODUCTION

In this era most of the communications being done electronically, there exists an increase of publicly accessible knowledge sources, both in the form of documents and factual databases. These sources are obviously heterogeneous and dynamic. They are not of same kind because they are autonomously developed and maintained by independent organizations for different purposes. They are able to change constantly since new information is being revised, added and removed. Such a nature imposes challenges on systems.

The Semantic Web framework is used to express content of resources to be semantically fetched; some manual description is expected using the Resource Description Framework or ontology. Lack of solutions to the well-known difficulties in manual ontology development, an automated ontology management is required for the efficient and consistent knowledge acquisition and integration. TAMBIS tried to provide a separate out from biological information services by building a homogenising layer on top of the different sources using the classical mediator-wrapper architecture. It is developed to get source transparency using a mapping from terms placed in a conceptual knowledge base of molecular biology onto terms in external Sources.

Natural language processing is also most active areas in Human-computer Interaction and it is a branch of AI which includes Information Retrieval and Language Analysis. The goal of NLP is to enable communication between people and computers without proceeding to memorization of complex commands and procedures. NLP is a technique which the computer can understand the languages used by humans. The general goal for most computational linguists is to gradually teach the computer with the ability to understand and generate natural language as if they talk to another person. The applications that will be possible when Natural Language Processing capabilities are fully brought to an end are impressive computers would be able to analyse natural language, interpreting languages in real time, or fetching and summarizing information from a variety of data sources, depending on the users' requests. The analysis process identifies relevant parts of the query using language-dependent ontologies describing the concepts of the application domain. The domain-specific processing logic defines how these relevant parts are related to each other and builds the appropriate database query.

Additional languages can be added easily to the information system and the interaction with the system is to be designed as per the differing capabilities of various client devices such as web browsers or PDAs.

In this paper we introduce TIMS, an integrated KMS in the domain of molecular biology, where terminology-based knowledge acquisition, knowledge integration, and XML-based knowledge retrieval are combined using tag information and ontology management tools. The management of knowledge resources is based on XML, RDF, and ontology-based inference.

However, our aim is to facilitate the KA and KI tasks not only by using manually defined resource descriptions, but also by exploiting Natural Language Processing techniques such as automatic term recognition and automatic term
clustering, which are used for automatic and systematic ontology population.

2. RELATED WORK
Attempts at NLP database interfaces are just as old as any other NLP research. In fact database NLP may be one of the most important successes in NLP. Asking questions to databases in natural language is a very convenient and easy method of data access, especially for casual users who do not understand complicated database query languages such as SQL. The success in this area is partly because of the real-world benefits that can come from database NLP systems, and partly because NLP works very well in a single-database domain. Databases usually provide small enough domains that ambiguity problems in natural language can be resolved successfully. Here are some examples of database NLP systems: LUNAR (Woods, 1973) involved a system that answered questions about rock samples brought back from the moon. Two databases were used, the chemical analyses and the literature references. The program used an Augmented Transition Network (ATN) parser and Woods’ Procedural Semantics. The system was informally demonstrated at the Second Annual Lunar Science Conference in 1971. [1] LIFER/LADDER was one of the first good database NLP systems. It was designed as a natural language interface to a database of information about US Navy ships. This system, as described in a paper by Hendrix (1978), used a semantic grammar to parse questions and query a distributed database. The LIFER/LADDER system could only support simple one-table queries or multiple table queries with easy join conditions.

3. SYSTEM ARCHITECTURE
In the domain of molecular biology, for all newly created concepts there would be an increasing amount of new words. Knowledge Acquisition and Integration

Existing dictionaries could not cover all needs so there is a need for capable term discovery.

ATRACT integrates Auto006Datic Term Recognition and Automatic Term Clustering. It is a terminology management mechanic and it helps biologists to gather and manage terminology. It fetches and classifies the terms and sends the results as XML tag information to TIMS.

POS Tagger means Part-of-speech. POS tagging is harder than just having a list of words and their parts of speech, because some words can represent more than one part of speech at different times, and because some parts of speech are complex or unspoken. The method, (C-value/NC-value), combines linguistic and statistical information. The first part, C-value enhances the common statistical measure of frequency of occurrence for term extraction, making it sensitive to a particular type of multi-word terms, the nested terms. The second part, NC-value, gives: 1) a method for the extraction of term context words (words that tend to appear with terms), 2) the incorporation of information from term context words to the extraction of terms.

Query manager processes the query, which was elaborated by the user. Using the ontology based inference the Tag data manager retrieves the relevant data from the documents.

4. KNOWLEDGE INTEGRATION AND MANAGEMENT
Knowledge integration and management in TIMS is organised by integrating XML-data management and tag- and ontology-based information extraction. User formulates a query, which is processed by a query manager. The tag data manager retrieves the relevant data from the collection of documents via a tag database and ontology-based inference. TIMS is organised by integrating XML-data management and tag- and ontology-based information extraction. User formulates a query, which is processed by a query manager. The tag data manager retrieves the relevant data from the collection of documents via a tag database and ontology-based inference.
5. CONCLUSION

TIMS, an XML-based integrated KA aid system, in which we have integrated automatic term recognition, term clustering, tagged data management and ontology-based knowledge retrieval. TIMS allows users to search and combine information from various sources. An important source of information in the system is derived from terminological knowledge, which is provided automatically in the XML format. Tag-based retrieval is implemented through interval operations, which – in combination with hierarchical matching – prove to be powerful means for textual mining and knowledge acquisition.

The preliminary experiments show that the TIMS tag information management scheme is an efficient methodology to facilitate KA and IE in specialised fields.

Important areas of future research will involve expanding the scalability of the system to real WWW knowledge acquisition tasks and experiments with fine-grained term classification.

6. REFERENCES


