Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Engineering 38 (2012) 3235 – 3241
Procedia Engineering
www.elsevier.com/locate/procedia

International Conference on Modeling Optimization and Computing

Ontology Based Disease Information System

Mythili Thirugnanam[a],Mangayarkarasi Ramaiah[b], Pattabiraman V[c],Sivakumar R[c]

[a]*Assistant Professor(Sr) ,Computer Science and Engineering,VIT University , Vellore,TamilNadu ,632014,India.*
[b]*Assistant Professor(Sr) ,School of Information Technology and Engineering,VIT University , Vellore,TamilNadu ,632014,India.*
[c]*Bachelor of technology in Information technology,School of Information Technology,VIT University,Vellore,TamilNadu,632014,India.*

Abstract

Human body systems are interconnected and dependent and they can't function separately. Diseases which affect any of these systems will affect other systems as well. In existing disease diagnosis systems, diseases were classified based on different dimensions such as cause, treatment, types of diseases, etc. On the other hand the existing disease information systems are for pest control management and they cannot be extended to human diseases. In existing systems traditional database approach was used. A database provides queried information and the retrieval method is not efficient in the case of biomedical systems. The proposed system will classify the diseases into the 'system-wise diseases' instead of many dimensions. The system needs careful handling of disease information. A disease information system would be appreciable instead of databases since the information systems may provide more precise and descriptive information. Therefore it is required to develop a disease information system by providing more relevant information as per the user query. The ontology based disease information system is being build and semantic based rules are designed to respond to the corresponding user query. The proposed information system is mainly focusing on improving the query results and also supports ease of use to the user.

Keywords: Diseases;Ontology;SWRL;Information systems

1. Introduction

Human body is a complex bio-machine. Its functioning is more complex. Human body consists of many systems that are interconnected and work dependently. Disease may refer to any condition that impairs normal functioning. Predicting the diseases is a critical issue and requires extensive knowledge of

1877-7058 © 2012 Published by Elsevier Ltd.
doi:10.1016/j.proeng.2012.06.375

diseases and their functioning. Each disease will have set of attributes. For example, diabetes may have attributes such as age, blood pressure, family background, usage of tobacco products, etc. These attributes will help to predict the diagnosis of possible diseases that would come to patient in near future. Information system improves the effectiveness and efficiency of the system. The reason for opting information system is that, the database approach just provides the stored static data and doesn't provide detailed information. Many researchers developed information systems that were mainly related to pest information system based on WebGIS [9]. It delivered importance of information system in biomedical information management. Antonio J.Jara et al., implemented rule based intelligent information system to detect and predict myocardial diseases using ontologies [11]. Kiong et al., developed ontology based health system to store clinical data to facilitate interpretation of medical information machines [14]. Hadzic et al., had contributed to ontology based human disease studies using various technologies such as grid middleware, multi-agent system etc. [1] [2] [3]. At present there are no disease information systems available. Hence it is important to develop a disease information system based on ontology to solve the problem as it helps people to be aware about the diseases and symptoms.

Literature Survey:

There are many researchers working on disease diagnosis and medical records management. They proposed many works based on their research in various fields such as rule based systems, ontologies, etc. The approaches that are discussed in the recent years are discussed below. Maja Hazic et al., (2004) had conducted elaborative study on ontology and bio medical information systems. In their study of Disease Ontology based Grid Middleware for Human Disease Research, they implemented Grid middleware technology to provide intelligent search engine. It used knowledge from ontologies like medical issues, disease factors, etc. to narrow down the search results. Ontologies were used to provide descriptive information about the human diseases research. But this study was incomplete without user interfaces, security of Grid Middleware, etc. [1]. Maja Hadzic et al., (2005) proposed Ontology based multi-agent system to support human disease study and control. They designed a new ontology called generic human disease ontology (GHDO), for the representation of knowledge regarding human diseases and it was organized into four dimensions such as disease types, symptoms, causes and treatments. Multi-agent system with ontology provided intelligent and dynamic information retrieval. Few works were incomplete in the proposed system like local agents' interactions, security concerns, etc. [2]. Maja Hadzic at el., (2005) had continued their study on implementing Ontology based Disease System. They developed the GHDO (Generic Human Disease Ontology) to represent the disease knowledge and the information of the diseases is organized into four dimensions such as Disease types, Symptoms, Causes and treatment. In addition to this, the system was aimed to support the study of complex disorders caused by many different factors simultaneously [3]. Illhoi Yoo et al. had done a Comparison study of various document clustering approaches for MeSH ontology to improve the clustering quality. Their experiment results showed that biomedical ontology MeSH significantly enhances clustering quality on biomedical documents and decent document clustering approaches performance were improved [4]. Hongyi Zhang et al., (2006) proposed a method for obtaining biological functions of gene by using GENE ontology method. They obtained difference between the all types of genes like difference and relations between parent's gene and their children's genes etc. The results showed that it was possible to construct probabilistic gene regulatory networks with the method of coefficient of determination (CoD) [5]. Akifumi Tokosumi et al., (2007) analyzed existing medical ontologies and suggested future directions of medical knowledge repository system. They discussed three knowledge repository that is localized nature of knowledge, collective acquisition of knowledge and usable knowledge repository. The locality of knowledge was suggested to be used in medical ontologies [6]. Tran Quoc Dung et al., (2007) presented Ontology based health care information extraction system – VnHIES. Authors developed two algorithms called "Elements Extracting Algorithm" and "New Semantic Elements" that were used to health care semantic words extraction. Document weighting algorithm was applied to get summary information. Including ontology into the system resulted in more accurate in extraction and summarization process [7]. Tharam S. Dillon et al., (2008) had continued to work on ontology in bio medical information storage and processing. They explained the importance of ontologies in representing bio medical information storage and processing knowledge models. It also explained the uses of ontologies in semi-automatic and automatic tasks and explained with their previous system which is Ontology based Multi Agent System [8]. Dahua Xu et al., (2008) developed pest and disease information system based on WEBGIS named as Diseases and Pest Information system (DIPS). DIPS was a pest and disease control system and warning system designed for crop pests. The system architecture consists of different components and sub-systems such as WEBGIS component – the interface component, DIPI component which was used for data access, and Component model management sub-system. Component model management Sub-system manages the centralization between DIPS and DIPS and manages its components like mathematical models, pest components and interface component. The COM + component technology was used in order to incorporate object oriented technology, thus improved the scalability, reusability of the proposed system and enabled distributed computing [9]. Shastri L. Nimmagadda et al., (2008) proposed ontology based data warehouse modeling and managing ecology of human body for disease and drug prescription management. The system was focused on introducing the concept of ontology based warehouse modeling and representing human body system in ontological representation. The proposed system was yet to be put into practice [10]. Antonio J.Jara et al., (2009) implemented ontology and rule based intelligent information system to detect and predict myocardial diseases. The developed system was used in pre-hospital health emergencies, remote monitoring of patients with chronic conditions and medical collaboration through sharing health-related information resources. Rule based system was designed to predict the illness by applying chronobiology algorithm. Ontology trees were constructed to in order to provide knowledge base of diseases. To avoid the observation periods in the hospital the system was used to send the information about the detected symptom or disease. Though this system was useful the chronobiology algorithm was not based on the diagnosis and required improvements in the artificial intelligence layer [11]. Ali Adeli et al., (2010) had developed a Fuzzy Expert System for Heart Disease Diagnosis based on the V.A. Medical Center, Long Beach and Cleveland Clinic Foundation database. It had 13 input fields and one output field. The input fields were attributes of heart disease such as chest pain, resting electrocardiography (ECG), etc. The output field was an integer value ranges from 0 to 4 to denote different levels of the heart disease. The system showed 94% accuracy in classifying the heart disease [12]. Ersin Kaya et al., (2011) had developed a Diagnostic Fuzzy Rule-Based System for Congenital Heart Disease. They retrieved medical dataset of patients from Pediatric Cardiology Department at Selcuk University, from years 2000 to 2003. They classified the medical dataset into 4 groups for fuzzy classifications and then the fuzzy rules were created based on various attributes in the data set. These attributes includes 8 Conditional attributes, 4 Decision attributes. After classifying fuzzy rules, they weighted it two different methods such as weighted vote method and single winner method and compared the results. They increased the accuracy of Classification of Congenital Heart Diseases [13]. Yip Chi Kiong et al., (2011) had developed Health Ontology System to store clinical databases into a shared cumulative ontology so that it can be intercepted by machines. Such system was built upon their previous works that are Ontology generator, Ontology Distiller and Ontology Accumulator. These are software tools used in the system generate ontology. Ontology generator generates ontology from a database. Ontology Distiller does the reverse process by storing ontology into a database. The Ontology accumulator does the integration of similar types of ontology. Integration of these tools helped to convert small databases to complex database tables [14]. Lynn Marie Schriml et al., (2011) created disease ontology that provides a backbone for disease semantic integration. The developed disease ontology contains a knowledge base of 8043 human diseases. It was designed as a web interface designed for high speed, efficiency and robustness through the use of graph database. It supported querying of disease name, synonym, and definition of diseases. This work was not extended to relations between diseases [15].

There are many bio-medical ontologies created like GHDO for the study of human disease, GENE ontology and ontology for clinical database. Antonio J.Jara et al., (2009) created ontology based system to predict the myocardial disease but did not improve prediction of the system wise disease [11]. Later Lynn Marie Schriml et al., created a disease ontology which deals with all human diseases. It supported to querying about diseases, synonyms and definition of diseases [15]. Yip Chi Kiong et al., created ontology based tools to manage medical databases [14]. Based on the survey it is clear that the existing systems have drawbacks such as relating the diseases with their symptoms, reusing the collected information. The ontologies created for the purpose of sharing knowledge and reuse it among researchers but sometimes failed to do so. Though some systems such as DOID, VnHIES served that purpose, still there is no disease information system available based on ontology [7] [15].The proposed work suggests creating the ontology for the system wise diseases to get information about diseases and based on symptoms. The system would use semantic web rule language (SWRL) to identify the disease.

2. Proposed Work

The proposed work consists of three different phases such as Knowledge acquisition phase, Rule engine phase and Query processing phase. Creation of ontology will be done in knowledge acquisition phase. Designing the SWRL rules are done the second phase. Query processing phase will get the query from the user and responds to his query. The field of medical sciences is very large and diseases with their symptoms will be vast and complex. Incorporating every disease into the ontology to be created may lead to inconsistency in the system. So the proposed work is restricted to human body system wise diseases. On the other hand to identify these system wise diseases we need exclusive domain knowledge. As we need to provide precise information we have to eliminate uncommon diseases so that results may become more consistent.

2.1. Knowledge acquisition and ontology creation phase:

Ontology of a domain is useful to share and reuse the explicit knowledge of that domain. Creating medical related ontologies are very much in need. Creating ontology and identifying relationships in them can be done with the help of protégé. Building ontology consists of different modules such as creating classes and class hierarchies, identifying relationships and identifying attributes or properties. Sample ontology construction is shown in the figure 2 and table 1.

2.1.1. Knowledge Acquisition

In this process required knowledge for the system is collected and organized in required format. The system requires disease information and their symptoms. It is collected from various resources such as BioPortal [16], Disease Ontology (DO) [17], OBO Foundry [18], etc.

2.1.2. Creating classes and class hierarchies:

The proposed work classifies the diseases into human body system wise diseases. Hence each disease system will be considered as a subclass of superclass called diseases. This work suggests creating a subclass of symptoms under the superclass diseases. So that it serves as a repository of a symptoms and they related with any disease system, where disease systems classes and symptoms are siblings of disease.

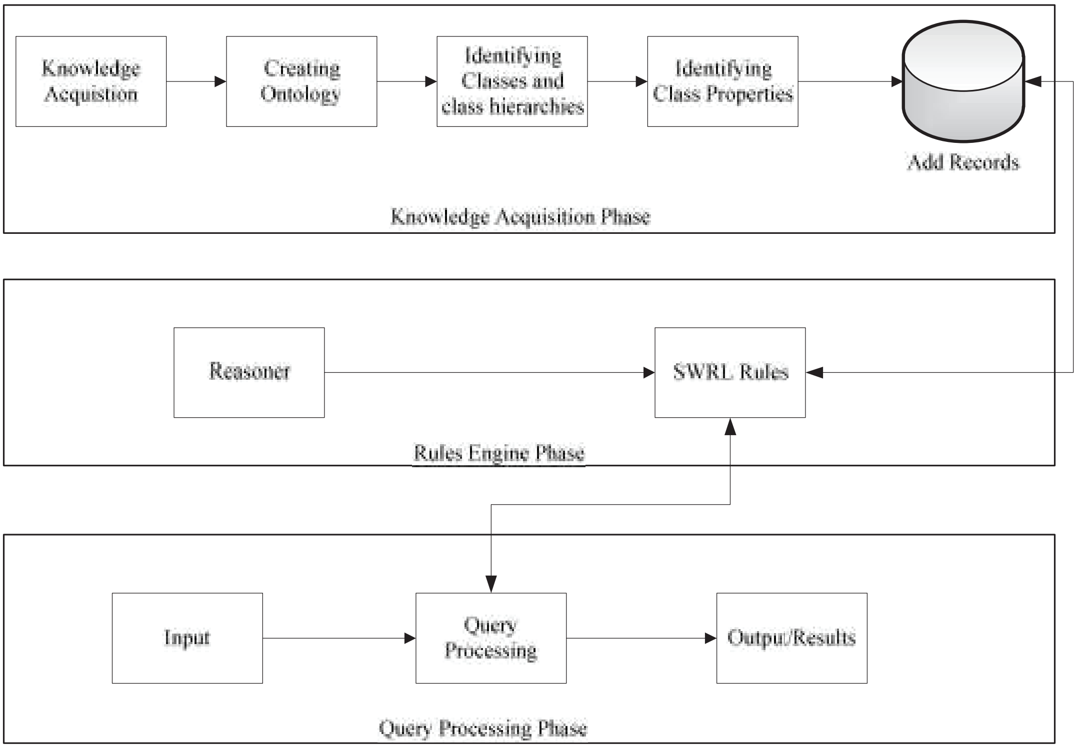

Figure 1. Architecture of the proposed work

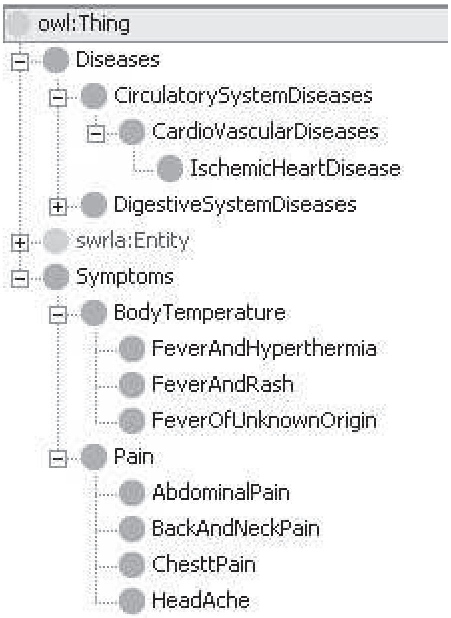

Figure 2. Sample Ontology Construction.

	Class 1	Class 2	Class 3
Class name	heart_disease	chest_pain	fever
Object property	has_symptom	is_a_symptom_of	is_a_symptom_of
Data property	has_name	has_id	has_temperature
Functional property	transitive	transitive	transitive

Table 1. Sample ontology classes and properties.

2.1.3. Identifying relationship among the classes:

The relationship between a system disease and a symptom can be identified with “has_a_symptoms” relationship. For example “stroke has a symptom dizziness” or we can create a vice versa relationship “is_a_symptom_of”. For example: “dizziness is a symptom of stroke”. Identifying such relationships is of the most important things so that the consistency of the system would be promised.

2.1.4. Identifying the data properties and object properties:

This work suggests identifying and adding the properties of a class or subclass in order to make the classes more understandable. A property of a class may define it well. Ex: “fever is a disease that has main property called temperature”. So that in ontology we can create a property for fever such as fever as follow “fever *subclass* of disease has temperature”

2.1.5. Adding Records :

On the successful creation of disease ontology with symptoms, the next step is to create instances using the available data. Those instances act as training dataset and would be useful to create rules in the next phase.

2.2. Rule Engine Phase:

The Rule engine phase of the proposed work consists of a semantic reasoner and SWRL Rules. The semantic reasoner is used for checking the consistency of the relationship among the classes and their properties. Since SWRL rules need to be created from valid relationships it is important to check the consistency of ontology. For example if a person has a symptom chest pain, shortness of breathing, pain in arms, dizziness, eye colour red and fastest heartbeat we should say that the person has coronary heart disease. A sample SWRL rule created using protégé and SWRLTab is shown below

swrl::annotation chest_pain (? s1) ^…..^shortness of breathing(? s2) → disease(?caronaryheartdisease_1).

2.3. Query Processing phase:

In query processing phase user interaction is managed. The user enters symptoms of a disease he or she has, as a request to the query processor. The query processor checks with SWRL rules for relationships between the diseases. It returns the diseases associated with symptoms entered. The query processor then displays the output to the user.

3. Conclusion and Future Enhancement

The proposed work suggests an approach to create disease information system to assist people based on ontology. The developing disease information system will help the users to be aware of diseases and their symptoms and help them to take viable actions. The suggested work points out the importance of creating a disease information system. On the successful compilation of this work, it can be extended to other diseases.

4. References

[1] MajaHadzic, Elizabeth Chan, Pronpit Wongthongtham and Robert Meersman, "Disease Ontology based Grid Middleware for Human Disease Research Study", The 30th Annual Conference of the IEEE Industrial Electronics Society, 2004.
[2] Maja Hadzic and Elizabeth Chang, "Ontology-based Multi-agent System support Human Disease Study and Control. Frontiers in Artificial Intelligence and Applications (special issues on Self-organization and Autonomic Informatics)", pp 129-141, IOS Press, 2005.
[3] Maja Hadzic and Elizabeth Chang, "Ontology-based Support for Human Disease Study", Proceedings of the 38th Hawaii International Conference on System Sciences – 2005.
[4] Illhoi Yoo and Xiaohua HU, "Biomedical Ontology MeSH improves Document Clustering Qualify on MEDLINE Articles:A comparison study".
[5] Hongyi Zhang, Jiexin Pu and Junying Zhang, "Construction of gene regulatory networks based on gene ontology and multivariable regression", 2006.
[6] Akifumi Tokosumi, Naoko Matsumoto, and Hajime Murai, "Medical Ontologies as a Knowledge Repository", IEEE/ICME International Conference on Complex Medical Engineering - 2007.
[7] Tran Quoc Dung and Wataru Kameyama, "A Proposal of Ontology-based Health Care Information Extraction System: VnHIES", IEEE, 2007.
[8] Tharam S. Dillon, Elizabeth Chang and MajaHadzic, "Ontology Support for Biomedical Information Resources", 21st IEEE International Symposium on Computer-Based Medical Systems -2008.
[9] Dahua Xu, Yinsheng Zhang and Huilin Shan, "Develop of diseases pest information system based on WEBGIS", International Seminar on Future BioMedical Information Engineering - 2008.
[10] Shastri L. Nimmagadda, Sashi K. Nimmagadda and Heinz Dreher, "Ontology based data warehouse modeling and managing ecology of human body for disease and drug prescription management", Second IEEE International Conference on Digital Ecosystems and Technologies (IEEE DEST) - 2008.
[11] Antonio J.Jara, Francisco J Blaya, Miguel A. Zamora and Antonio F.G. Skarmeta, "An ontology and rule based intelligent information system to detect and predict Myocardial Diseases",2009.
[12] Ali.Adeli and Mehdi.Neshat, "A Fuzzy Expert System for Heart Disease Diagnosis", Proceedings of the International Multi Conference of Engineers and Computer Science Volume I, 2010.
[13] Ersin Kaya, Bulent Oran and Ahmet Arslan, "A Diagnostic Fuzzy Rule-Based System for Congenital Heart Disease", World Academy of Science, Engineering and Technology 78, 2011.
[14] Yip Chi Kiong, Sellappan Palaniappan and Nor Adnan Yahaya, "Health Ontology System", 7th International Conference on IT in Asia (CITA), 2011.
[15] Lynn Marie Schriml, Cesar Arze, Suvarna Nadendla, Yu-Wei Wayne Chang, Mark Mazaitis, Victor Felix, Gang Feng and Warren Alden Kibbe, "Disease Ontology: a backbone for disease semantic integration", Nucleic Acids Research Advance Access published November 12, 2011.
[16] BioPortal, http://bioportal.bioontology.org
[17] Disease ontology, http://disease-ontology.org
[18] OBO Foundry, The Open Biological and Biomedical Ontologies , http://obofoundry.org/