International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com, Volume 5, Issue 2 (March - April 2017), PP. 23-28

DEVELOPMENT AND VALIDATION OF ONTOLOGY BASED KNOWLEDGE REPRESENTATION FOR BRAIN TUMOUR DIAGNOSIS AND TREATMENT

[1]S. Senthilkumar, [2]G. Tholkapia Arasu

[1] Research Scholar, Department of Computer Science and Engineering
[1] Manonmaniam Sundaranar University, Thirunelveli, Tamilnadu, India.
[2]Principal, AVS College of Technology, Salem, Tamilnadu, India.
[1]edusenthil@yahoo.co.in, [2]tholsg@gmail.com

Abstract— **Ontologies are relatively a new way of defining and storing knowledge and it describes a certain domain by dividing it into several concepts. It describes the relations between those concepts and finds many applications in all fields including disease diagnosis. Patients are diagnosed by medical doctors based on the signs and symptoms of the disease which may sometimes be misinterpreted due to lack of correct understanding/interpretation of information resulting in wrong treatment. In this paper, a computational model based on ontology is presented to reduce brain tumour misdiagnosis thereby making correct treatment decisions. This ontological representation provides structured knowledge in machine understandable form which can be shared across multiple platforms and uses a common vocabulary thereby helping in proper treatment.**

Index Terms— **Ontologies, Diagnosis, Brain tumour, Decision making.**

I. INTRODUCTION

Recent developments of communications technologies and Ontology is a Greek term [1] that deals with the nature of existence. In computer science, ontology refers to computational use of encoding human knowledge. Artificial Intelligence refers ontology as shared conceptualization specification [2] which is individual knowledge in a domain obtained with experience [3]. When many domain experts work on ontology with shared conceptualizations, it leads to unique representation called formal representation language. The result of a formal representation is a compilation of entities, expressions and axioms.

Entities signifies concepts or classes and its instances, data types, literals and properties. Expressions represents descriptions of entity in language like Web Ontology Language (OWL), built on eXtensible Markup Language (XML), Resource Description Framework (RDF) and RDF-Schema (RDF-S) standards[4] thereby providing highly expressive means of knowledge representation. There are three sublanguages of OWL namely OWL-Lite, OWL_DL and OWL-full.

Axioms find relationship among entities. This relationship can be between classes, individual to class and between property. Reasoners are used for inferring the relations depending on logic relations of concepts. Any ontology can be visualized as graph, with nodes as concepts and relations as edges. In semantic web, ontology provides a method to add meta information using software agents. A concept or a class is the biggest ontology component that contains many attributes. The most commonly viewed ontologies are SNOMED Clinical Terms, National Drug File, International Classification of Diseases, MedDRA and NCI Thesaurus.

The Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) is a biomedical ontology covering drugs diseases, operations, medical devices and symptoms etc. It may be used for the coding, retrieval and processing of clinical data. SNOMED CT is a formal logic-based syntax containing hierarchies and is a combination of UK National Health System's (NHS) Read codes and SNOMED Reference Terminology (SNOMED-RT), developed by the college of American Pathologists.

The National Drug File Reference Terminology (NDF-RT) is a formalized depiction for medication terminology arranged into concept hierarchies, with each concept is a node containing a list of term synonyms and a unique identifier. Top-level concepts are more general than lower-level ones and NDF-RT more than 45000 concepts in hierarchies of maximum depth 12. The International Statistical Classification of Diseases and Related Health Problems (ICD) is a terminology which attempts to classify signs, symptoms and causes of disease and morbidity and it appeared in the mid-19th century and is now maintained by the World Health Organization (WHO).

The Medical Dictionary for Regulatory Activities (MedDRA) deals with biopharmaceutical regulatory processes and has terms associated with all phases of the drug development cycle. MedDRA has hierarchical structure of fixed depth, as shown in Fig. 1. System Organ Classes (SOCs) represent the 26 predefined overlapping hierarchies in which terms belong to. High Level Group Terms (HLGTs) and High Level Terms (HLTs) are general term groupings, denoting disorders or complications.

Preferred Terms (PTs) denote the preferred name for a concept, while Lowest Level Terms (LLTs) means terms of maximum specificity. In a formal ontology, a concept cannot be a child of itself. In MedDRA, this clearly happens, when a PT and its LLTs share a synonymy relation.

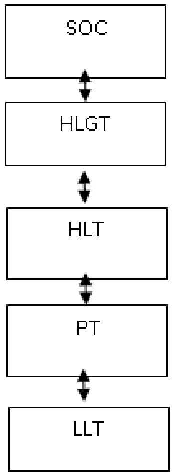

Fig. 1 Structure of MedDRA terminology

The National Cancer Institute Thesaurus (NCIT) is a controlled terminology for cancer research where the thesaurus is converted to formal OWL syntax and updated at fixed intervals. The conversion was not an easy one; many inconsistencies and modeling dead-ends that were encountered in the conversion procedure have been documented [5] with violations of ontological principles [6].

This paper puts forward a computational model based on ontology representation, which is used to support decision-making in clinical diagnostic generation from input data like medical history and risk factors. This model is able to infer a list of likely clinical diagnoses relying on a mechanism of inference (rules) on the stored information. The model is based on ontologies due to its ability to represent and share knowledge in an automated manner, the use of a shared vocabulary in the field of medical diagnostics and the ability to formalize knowledge in ontological language accepted as standard. In addition, ontologies are able to infer new knowledge by relying on inference machines, so they are a novel technology to knowledge representation in order to improve diagnosis process from certain medical conditions, such as semiotics, medical history and risk factors.

The paper is organized as follows. Section 2 presents the related work in medical diagnosis and the theoretical foundations of knowledge representation based on ontology. Section 3 provides the knowledge representation model based on ontology proposed and. section 4 presents the implementation of the proposed ontology, which sets in all classes, relationships and instances of our ontology. Finally, Section 5 presents the conclusions.

II. LITERATURE REVIEW

In literature, many studies supports the process of carrying out medical diagnosis and two methods are identified: a) the learning-based on C4.5 [7] and diagnosis based on Support Vector Machines (SVM) algorithm is discussed in [8], b) rule based approaches generated manually by experts [9] and [10]. There is no method cited that support ontology to represent medical knowledge and perform clinical diagnosis and treatment. The knowledge representation based on ontology is a technique that has been used in domains such as bioinformatics and molecular biology. In order to capture existing processes in these domains [11], a protein ontology is proposed in [12] to support computer systems in biology in general providing a proper interpretation of the data, and to model the relationships in proteins. [13] proposes a model of an ontology-based annotation for disease phenotypes, which facilitates the discovery of new phenotype-genotype relationships among species. Research on ontology-supported medical diagnosis is a challenging task that has no investigation involved.

Knowledge representation of diagnosis has been addressed from other perspectives as probabilistic and logic based approaches. Hence, [14] proposes an approach that uses Bayesian networks in order to argue and deal with the uncertainty problem of fault diagnosis well, the Bayesian network structure is established according to the cause and effect sequence of faults and symptoms, which are the only medical conditions considered in the model of authors; [15] applies the concept of a fuzzy set as knowledge representation to improve the decision making process.

In this research, an attempt is made to develop a disease diagnosis and treatment system. The system is evaluated using symptoms that are added to clinical data in determining brain tumour and its severity. [16] provides an approach to create a new medical knowledge representation model, based on the use of probabilistic theory. Their work start from the description, realized by an expert of the medical knowledge describing the relation between symptoms and diagnoses. The proposed approach consists of building a probabilistic model including the medical knowledge base.

III. PROPOSED KNOWLEDGE REPRESENTATION MODEL

This paper puts forward an ontological model to represent medical information, which facilitates in obtaining a brain tumour diagnosis. The information is modeled based on the concepts disease, symptoms, causes, treatment and diagnosis. The proposed ontological model is presented in a modular fashion, which has been divided into the five sections. Five sub-ontologies created for the complete brain tumour model are Diseases, Symptoms, causes, diagnosis and treatment. Figure 2 shows the complete model of the brain tumour disease ontology and their relationships between sub-ontologies.

A. Brain Tumour Disease Ontology

Diseases ontology has the disease class and subclasses: like pituitary tumour, Pineal tumour, Primary tumour, cyst, high grade or low grade malignant tumour.

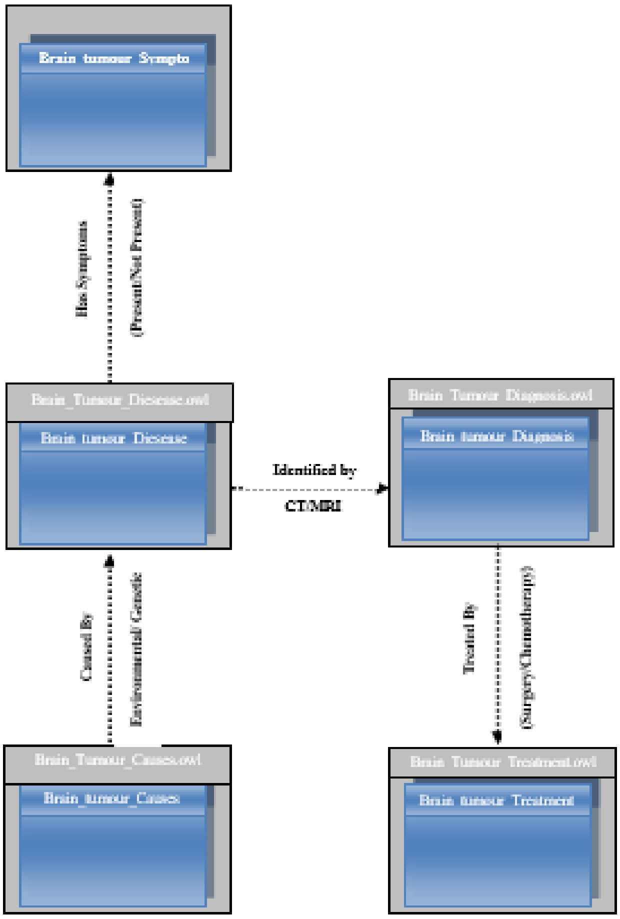

Figure 2: Brain tumour disease treatment model

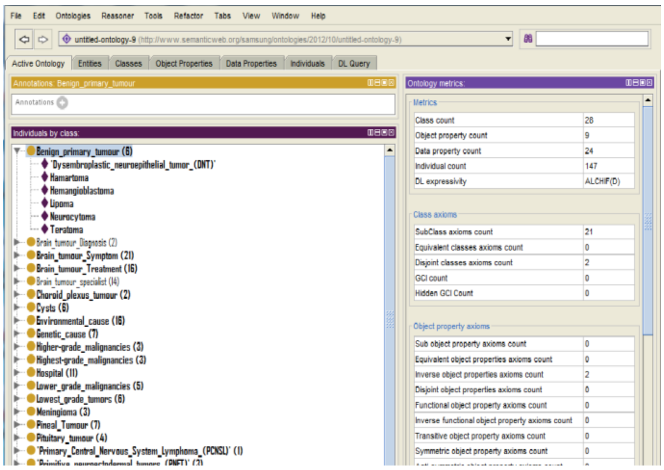

Fig. 3 Root Class Hierarchy of brain tumour with protege

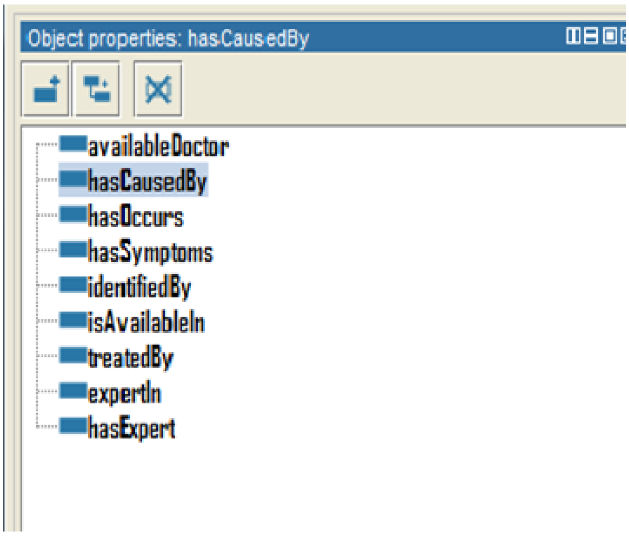

Fig. 4 Object properties

Fig. 5 Brain Tumour Disease ontology

Disease class is related to class diagnosis, symptom, causes and treatment. Figure 3 shows the root class hierarchy and figure 4 gives the nine object relations. Figure 5 shows the Brain tumour diseases ontology with some examples of subclasses and their corresponding relationships. The present and not Present relations express the absence or presence of certain disease symptoms. Is a high grade /low-grade gives the malignancy level of tumour and caused by relationship provides the environmental or genetic cause

B. Brain Tumour Symptoms Ontology

Symptoms ontology shapes knowledge presented in the disease symptoms. The main class of this ontology is Symptom which has four relationships like results_in, leads_to, causes. Examples of Symptom class instances are: headache, confusion, change in behaviour, neck pain, hearing loss and neck pain. Figure 6 provides the brain tumour symptoms ontology.

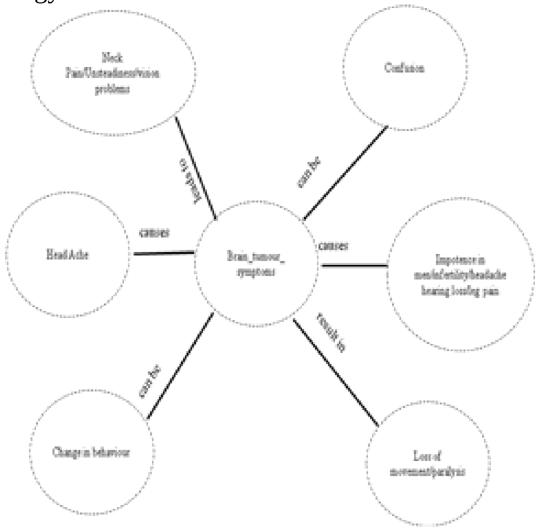

Fig. 6 Brain Tumour Symptoms Ontology

C. Brain Tumour Causes Ontology

The causes ontology provides the cause for the disease and has two subclasses which can be environmental or genetic. The instances of these classes are agricultural/industrial chemicals, cured food, alcohols, smoking, birth control pills, headache, sleeping pills for environment subclass and chromosome changes, gorlins, turcots, tuberous sclerosis etc for genetic class. Figure 7 shows the ontology to represent the causes for the disease which can be genetic or environmental.

D. Diagnosis and Treatment Ontology

The presence or absence of the disease is diagnosed by CT or MRI The treatment ontology has relations like treated by, cured by etc. The treatment can be done by instances like biopsy, chemotherapy, medication, neurosurgery and radiation therapy. Figure 8 shows the disease treatment ontology.

Fig. 7 Brain Tumour Cause Ontology

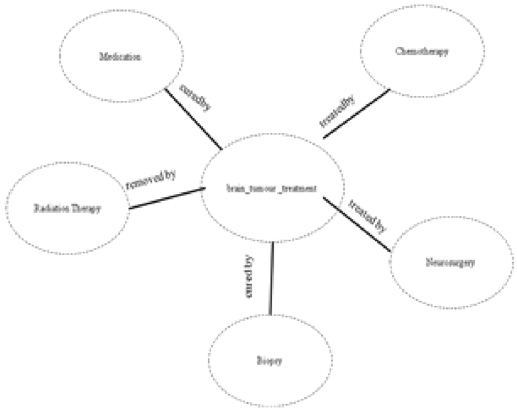

Fig. 8 Brain Tumour Treatment Ontology

IV. ONTOLOGY IMPLEMENTATION

The implementation of ontology was made relying on the framework based on knowledge and ontology editor called Protégé [17]. The implementation of ontology was made relying on the framework based on knowledge and ontology editor called Protégé [17]. The ontology was encoded using the Web Ontology Language in its version 2.0, which provides classes, properties, individuals and data type values [18]. This language is stored as documents under the Semantic Web standards.

The brain tumour ontology has 5 main classes and 22 subclasses (28 classes or concepts as whole), 24 relations of data type, 9 object type relations. For diagnosis and treatment, the input data is mapped with information stored in ontology and result are obtained. The information stored in brain tumour ontology by classes, relations, individuals and types of data is validated and accepted by a group of medical experts that enables the exchange of knowledge, the use of a common vocabulary in applications that are supported by this ontology and reuse of knowledge represented.

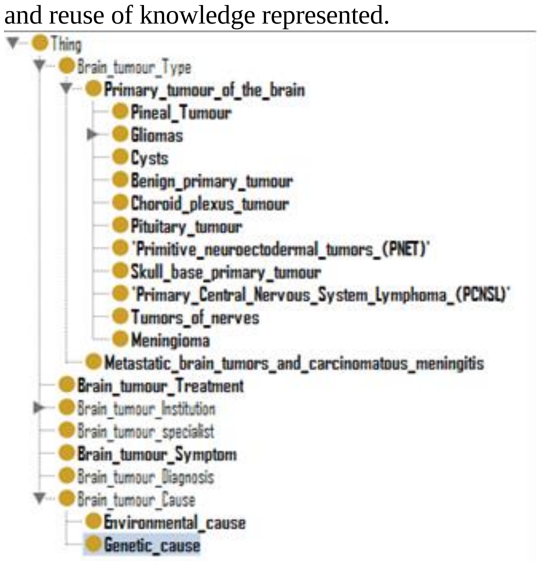

Fig 9 Brain Tumour Ontology implementation with protégé

Fig.10 Cause ontology implementation with protégé

Figure 9 provides the hierarchy of brain tumour disease ontology implemented with protégé. Figure 10 provides the cause ontology implementation, Figure 11 provides the diagnosis and symptom ontology implementation. Figure 12 provides the brain tumour treatment hierarchy implementation with protege. The treatment ontology has 16 instances which signify the different methods of cancer treatment.

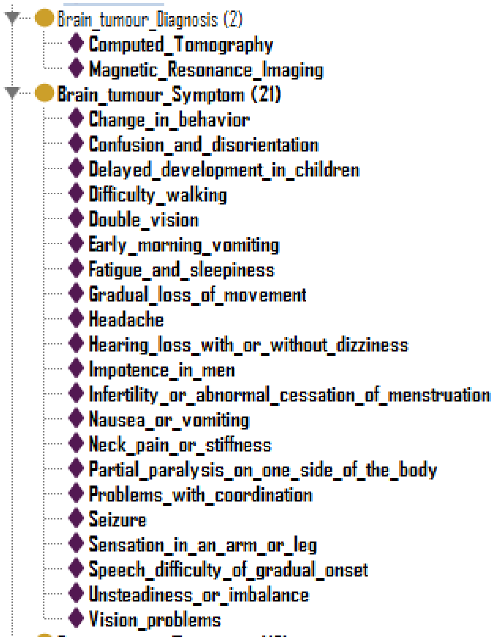

Fig. 11 Diagnosis and Symptom hierarchy implementation with protégé

Fig. 12 Treatment implementation with protégé

CONCLUSION

This paper integrates ontological concepts (classes), real entities (entities or individuals) and interactions between classes (relations) and helps in decision making by doctors. These elements add the meaning and information necessary to perform brain tumour diagnosis. The ontological model focuses on five sub-ontologies Brain tumour Disease, Symptoms, Cause, Treatment and Diagnosis. Brain tumour diseases ontology represents the base model, symptoms ontology represent the information leading to disease, cause ontology represent the base cause of the disease, diagnosis ontology represents CT or MRI scan for finding the presence of disease and treatment ontology provides the treatment available for the disease. For obtaining the probable diagnosis, fuzzy inference rules to extract knowledge has to be designed in which research work is continued. The automatic generation of rules will help or support the physicians in decision-making for making final diagnosis. Further work on ontological model can add more classes and relations, which can further strengthen the knowledge base.

REFERENCES

[1] Guarino, N, Formal Ontology in Information Systems: Proceedings of the 1st International Conference June 6-8, 1998, Trento, Italy, volume 46. Ios Press Inc. 1998.

[2] Gruber, T. R. et al. Toward principles for the design of ontologies used for knowledge sharing. International journal of human computer studies, 43(5):907-928, 1995.

[3] Huang, C.-r., Calzolari, N., Gangemi, A., Lenci, A., Oltramari, A., and Prévot, L.. Ontology and the Lexicon: A Natural Language Processing Perspective. Cambridge University Press Cambridge,2010

[4] McGuinness, D. L., Van Harmelen, F., et al, Owl web ontology language overview. W3C recommendation, 10, 2004.

[5] Ceusters, W., Smith, B., and Goldberg, L. A terminological and ontological analysis of the nci

thesaurus. Methods of information in medicine, 44(4):498,2005.

[6] Schulz, S., Schober, D., Tudose, I., and Stenzhorn, H. The pitfalls of thesaurus ontologization-the case of the nci thesaurus. In AMIA Annual Symposium Proceedings, page 727. American Medical Informatics Association,2010.

[7] Alexopoulos, E., Dounias, G. D., Vemmos, K.: Medical Diagnosis of Stroke using Inductive Machine Learning. In: Proceedings of Workshop on Machine Learning in Medical applications, Advance Course on Artificial Intelligence ACAI99, pp. 91−101,Chania, Greece 1999

[8] Shashikant, U. G., Ghatol A. A.: Heart Disease Diagnosis Using Machine Learning algorithm. In: Proceedings of the International Conference on Information Systems Design andIntelligent Applications. Advances in Intelligent and Soft Computing 132, pp. 217−225, India 2012

[9] Drastal, G. A., Kulikowski, C. A.: Knowledge-Based Acquisition of Rules for Medical Diagnosis. Journal of Medical Systems 6(5), pp. 433−445, Springer, Netherlands,1982

[10] Jianwei, X., Ke, X.: A novel extracting medical diagnosis rules based on rough sets. In: The 2009 IEEE International Conference on Granular Computing, pp. 608−611, IEEE Press, Nanchang, China, 2009

[11] Steven, R., Goble, C.A., Bechhofer, S.: Ontology-based knowledge representation for bioinformatics. Brief Bioinform 1(4), pp. 398−414, United Kingdom,2000

[12] Chang, E., Sidhu, A., Dillon, T., Sidhu, B.: Ontology-based Knowledge Representation for Protein Data. In: Dillon, T. and Yu, X and Chang, Elizabeth (ed.), 3rd Intl.conference on Industrial Informatics, Frontier Technologies for the Future of Industry and Business. IEEE Press, pp. 535−539. Australia (2005)

[13] Washington, N. L., Haendel, M. A., Mungall, C. J., Ashburner, M., Westerfield, M., Lewis, S. E.: Linking Human Diseases to Animal Models Using Ontology-Based Phenotype Annotation. PLoS Biol. Nov; 7(11) USA, 2009

[14] Baoping, C., Yonghong, L., Qian, F., Yunwei, Z., Zengkai, L., Shilin, Y., Renjie, J.: Multisource information fusion based fault diagnosis of ground-source heat pump using Bayesian network. Applied Energy, vol. 114, pp. 1−9,2014

[15] Anouncia, M. S., Maddona, C. L., Jeevitha, P., Nandhini, R. T.: Design of a Diabetic Diagnosis System Using Rough Sets. Cybernetics and Information Technologies, vol. 13(3), pp. 124−169, 2013

[16] Alsun, M., Lecornu, L., Solaiman, B.: Possibilistic Medical Knowledge Representation Model. J. Adv. Inf. Fusion, vol. 7(2), pp. 101−113. (2012)

[17] Tudorache, T., Noy, N., Tu, S., Musen, M.: Supporting Collaborative Ontology Development in Protégé. The Semantic Web - International Semantic Web Conference, vol. 5318, pp. 17−32,2010.

[18] Smith, M., Horrocks, I., Krötzsch, B. G.: OWL 2 Web Ontology Language: Conformance (Second Edition). Eds. W3C Recommendation. Latest version available at http://www.w3.org/TR/owl2-conformance/. 2012.