HÖGSKOLAN SKÖVDE

Institutionen för kommunikation och information
Examensarbete i datavetenskap 30hp
Avancerad nivå
Vårterminen 2010

Identifying design issues related to the knowledge bases of medical decision support systems

Submitted by:

Assad Abbas

Supervisor:

Dr. Marcus Nohlberg

Identifying design issues related the knowledge bases of medical decision support systems

Submitted by Assad Abbas to the University of Skövde, as a dissertation towards the degree of M.Sc. by examination and dissertation in the School of Humanities and Informatics.

June 03, 2010

I hereby certify that all material in this dissertation which is not my own work has been identified and that no work is included for which a degree has already been conferred on me.

Signature ____

Identifying design issues related to the knowledge bases of medical decision support systems

Assad Abbas

Abstract

The modern medical diagnostic systems are based on the techniques using digital data formats – a natural feed for the computer based systems. With the use of modern diagnostic techniques the diagnosis process is becoming more complex as many diseases seem to have the same pre-symptoms at early stages. And of course computer based systems require more efficient and effective ways to identify such complexities. However, the existing formalisms for knowledge representation, tools and technologies, learning and reasoning strategies seem inadequate to create meaningful relationship among the entities of medical data i.e. diseases, symptoms and medicine etc. This inadequacy actually is due to the poor design of the knowledge base of the medical system and leads the medical systems towards inaccurate diagnosis. This thesis discusses the limitations and issues specific to the design factors of the knowledge base and suggests that instead of using the deficient approaches and tools for representing, learning and retrieving the accurate knowledge, use of semantic web tools and techniques should be adopted. Design by contract approach may be suitable for establishing the relationships between the diseases and symptoms. The relationship between diseases and symptoms and their invariants can be represented more meaningfully using semantic web. This can lead to more concrete diagnosis, by overcoming the deficiencies and limitations of traditional approaches and tools.

Keywords: knowledge base, knowledge representation, learning scheme, reasoning strategy, tools and technologies, diagnostic accuracy.

Contents

1. Introduction

For a very long time doctors and scientists have desire that the computers should be made artificially intelligent to help them in diagnosing the diseases (Ledley & Lusted, 1959). With these motivations, specialists from the fields of computer science and different healthcare domains started joint research as an effort to develop computer software that are intelligent enough to assist the doctors in diagnosing the diseases based on the provided symptoms (as instructions) and having capability of storing and processing large amount of knowledge (Durinck et al., 1994). These systems were later on called as medical diagnostic decision support systems (MDSS) and clinical decision support systems (CDSS). Clinical systems are designed by integrating the patient data, the medical data repository (knowledge base) and the physician's cognition (Lobach & Hammond, 1997). A lot of advantages are associated with the successful implementation of such knowledge based systems and the most evident is enhancement in the diagnostic accuracy and precision. Another advantage of automated diagnostic systems is that computers are unaware of the matters of fatigue and human errors because physicians usually work under conditions bounded by high workloads and time constraints- that eventually can cause in disturbance of the cognitive processes of the physician, distraction and memory biases which can lead the physician to an incorrect and premature diagnosis (Goggin, Eikelboom & Atlas, 2007). Nonetheless, the literature on medical diagnostics and clinical support systems, has also posed different questions on their effectiveness and routine use. As described by Miller (1994):

> *"Despite many years of research and millions of dollars of expenditure on medical diagnostic systems, none is in widespread use at the present time"* (Miller, 1994, p. 8)

Also Ridderikhoff & Herk (1999) support this claim by stating that diagnostic systems available today vary in their nature from very simple to very complex, having much more diagnostic capabilities than those available in past, yet these systems have not been able to cast their profound effects on physicians in terms of their adoptability in practice and the reason for that is inability to provide accurate diagnostic advice to the physician. Information stored in medical decision support system is retrieved from the knowledge base as a result of reasoning by the reasoning component. However this task of retrieval is dependant on how the knowledge has been represented in the knowledge base. For example, in response to a query by a physician, the system must consult the knowledge base for the provided symptoms. If there resides a previously solved case based on similar symptoms, the reasoning component must retrieve that case from the knowledge base and should present as an advice to the physician. But in case the symptoms provided to the system differ from those present in already solved case, the reasoning component must reason with the system in a way different from the previous one, so that a correct and contextual diagnosis could be made and new case should be added in the history for future consultations. It means besides the knowledge representation more formal reasoning plays vital role in correctly diagnosing the disease. Furthermore, correct retrieval of diagnostic information depends on how the knowledge has been initially encoded in the system by using certain knowledge representation tools and technologies as well as learning algorithms (Chandrasekaran, 1986; Kong, Xu & Yang, 2008). The inabilities of tools in correctly representing the knowledge and the limitations of learning schemes in learning new knowledge weaken the reasoning process and finally result in incorrect diagnosis. Therefore knowledge representation, tools and technologies, learning and reasoning strategies are the factors that contribute to the design of knowledge base since they all are coupled with each other. The appropriate design of knowledge base contributes to diagnostic accuracy. This research identifies the design issues related to the knowledge base and finally suggests that using design by contract approach is suitable for proving the relationship between the pre-symptoms, post-symptoms, invariants and other medical data. Semantic web technologies such as ontology as representation formalism and Protégé as the tool may help in creating this relationship.

The structure of the report is as follows: Sections 1.1 and 1.2 of this chapter describe the background of knowledge base medical decision support systems and problem with these systems respectively. Chapter 2 contains aim, objectives and method of the research for identifying the design issues of the knowledge bases of medical decision support systems. Chapter 3 presents the literature survey conducted to identify the issues relevant to the design of the knowledge bases such as issues related to knowledge representation schemes, tools and technologies, learning and reasoning strategies. Chapter 4 presents the analysis of the study along with possible proposed solution. Chapter 5 concludes by providing the further direction for future research.

1.1. Background

A medical decision support system can be defined as a computer based program that has been developed with the intent of assisting physicians and doctors based on patients' medical data during their encounters with the patients to identify the diseases (Shortliffe, 1987). A lot of research work has been produced on knowledge based clinical support systems and has generally proved helpful for performing precise diagnosis. Perreault & Metzger (1999) have described the four major functions of a clinical decision support system. These functions include performing administrative tasks such as documenting the patient histories and referrals, managing the clinical complexities and details, cost control by eliminating the redundant and unnecessary tests and finally providing the decision support.

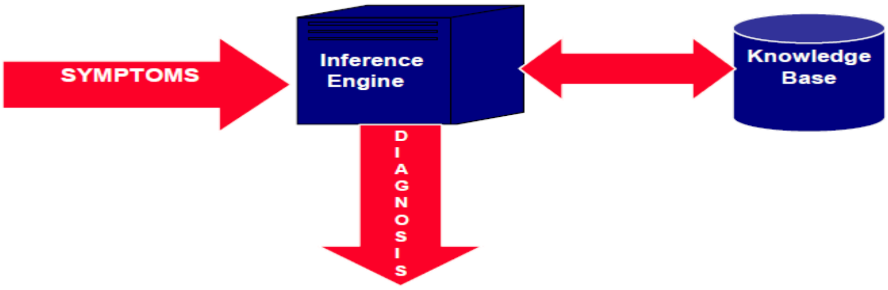

Figure 1: Working model of medical decision support system, Source Kong et. al (2008)

Also clinical decision support systems have demonstrated their significance repeatedly by improving the safety of patient, enhancing the quality of care standards and improvement in efficiency of health care deliverables (Sintchenko et al., 2002). Diagnosis in a knowledge based medical decision support system is performed using the two important components of the knowledge based system; the knowledge base and inference engine (Claudia, 1990). Figure 1 represents how the symptoms are provided, consultations are made with the knowledge base and diagnostic information is retrieved accordingly. The knowledge base of a medical decision support system is a repository of symptoms and diseases. These symptoms and diseases reside in the knowledge base in the form of rules. The inference engine or module is a component having the reasoning capabilities. Significance of this component of a knowledge based system is obvious since it is responsible for retrieving the contextual information from the knowledge base. Thus the reasoning in a logical way by the system is just like imitating the physician while interacting with a patient for diagnosis, by asking questions about signs and symptoms. As a result the medical system must come up with possible diagnosis based on the provided symptoms. However, some medical systems fail in providing completely accurate advice to the physician. This problem of diagnostic accuracy may be due to issues with underlying architecture and structure of the knowledge base besides the physicians' capabilities of interacting with the system and applying the implicit knowledge by making reasoning. This problem and the design factors that contribute to this problem have been discussed in section 1.2 of this chapter

1.2. Problem

The medical decision support systems require really dynamic and versatile representation infrastructure, due to the complex relationship of diseases and symptoms. This gives rise to the high level of complexity to identify such relationship, because the same disease may have different symptoms for different patients. Also it is possible that many diseases have similar symptoms initially i.e. there is no clearly defined boundaries among various diseases. In such situations it becomes difficult for the diagnostic system to distinguish between the diseases and identify the correct disease. Another complexity related to the medical data is that diseases are diagnosed and distinguished on the basis of various attributes of the associated symptoms such as pre-symptoms, post-symptoms and their invariants. To produce meaningful relationships between the symptoms and diseases is challenging for the medical expert systems in such situations.

Besides the complexities of medical knowledge and data itself, there are limitations that are specific to the knowledge based decision support systems. The limitations and issues of medical decision support systems in general have been discussed in the literature. However the design of knowledge base of a medical decision support system itself is the major reason for incomplete and imprecise diagnosis. For instance, the knowledge base of INTERNIST-I, a knowledge based medical decision support system, lacks in describing a symptom that is important for a particular finding (Wolfram, 1995). Also the other limitations of the knowledge base may be failure in a correct diagnosis due to an inadequate and incomplete reasoning mechanism (Wolfram, 1995) e.g. an inability of the system to perform reasoning without taking into account, all the possible aspects of a disease against a symptom or set of symptoms. It means that the issue of incorrect diagnosis is strongly connected to the knowledge base design of the medical decision support system. Design of knowledge base depends on various factors having concern directly or indirectly. The most important of these design factors are knowledge representation and the reasoning mechanism. Kalogeropoulos, Carson & Collinson (2003) believe that inability of a knowledge based medical system in providing expert advice to the physician may be because of imprecise way of representing the knowledge and incomplete and ambiguous reasoning capabilities of the systems. Moreover, their designs allow them to imitate a certain disease by using heuristics, so they fail to represent the complete knowledge. Similarly other factors critical to design of the knowledge base include learning of the knowledge base and the specific tools and technologies used to implement the knowledge bases. The importance of tools and technologies is vital because the knowledge is represented and retrieved using the tools and technologies (Shortliffe, 1986). Any inherent technological issue may result in incorrect diagnosis. Moreover retrieval of the contextual information from the knowledge base requires the knowledge base to learn new knowledge when it comes across a new situation. The knowledge base learns new knowledge if an effective learning scheme has been utilized. The efficient learning algorithm helps in retrieving the correct information from the system and makes the reasoning process concrete.

The above discussion shows that the correct diagnostic advice from the medical systems is based on knowledge representation, tools and technologies, learning and reasoning with the systems. They are the potential design factors of the knowledge bases since they are tightly integrated with each other and issues related to them needs to be identified.

2. Aim, objectives and method

2.1 Aim of the research

The research aims at identifying the design issues related to the knowledge bases of medical decision support systems.

2. 2 Objectives

The following objectives have been achieved to identify the design issues in the knowledge base of medical decision support systems

- Study of knowledge representation schemes used in knowledge based medical decision support systems to identify issues related to knowledge representation
- Study of tools/technologies used in the implementation of knowledge bases and their limitations
- Study of issues related to learning of knowledge base
- Study of issues related to reasoning components of the knowledge based medical decision support systems.

2.2.1 Study of knowledge representation schemes used in knowledge based medical decision support systems to identify issues related to knowledge representation

According to Purcell (2005), the effectiveness of a clinical decision support system is dependent on the design of its knowledge base. The success of a clinical support system from the accuracy point of view requires the proper analysis and design of its knowledge base. One of the most important design issues that may lead a medical decision support system towards inaccurate diagnosis is the representation of knowledge in the knowledge base. According to Carter (1999), knowledge representation deals with providing information to the intelligent systems belonging to a particular domain for efficient processing. A lot of knowledge representation schemes are available and have been used to represent knowledge in medical decision support systems for example logic, procedural, graph/network, and structured etc. Each of the representation schemes have some advantages as well as disadvantages associated with it (Kong et. al, 2008). These representation schemes are briefly described as follows

Medical decision support systems using logic as a knowledge representation scheme in the knowledge base usually have TRUE-FALSE structure and inference mechanism is just a lookup of the relevant facts (Kong et. al, 2008). However, the disadvantage of the logic representation is that the process of problem solving may become obscure when there is an increase in the number of related facts in the knowledge base. Ultimately it results in an increase in the number of ways these facts can be combined to infer a correct result (Bingi, Khazanchi, &Yadav, 1995).

Procedural knowledge representation scheme represents the knowledge in the form of rules. Medical decision support systems such as MYCIN, PUFF, UMLS and Chinese medical diagnostic systems (CMDS) are rule based. But the problem with rule based representation systems is that they do not create a correlation between the clinical signs and symptoms (Kong et. al, 2008) which cause in an immature inference by the inference mechanism and ultimately causes in an incorrect diagnosis.

Also networks such as Bayesian belief networks, decision trees and neural networks have been used in clinical decision support systems for representing knowledge (Kong et. al, 2008). All of these representation schemes are good in representing conditional dependencies and probabilistic inference (Montani & Terenziani, 2006). But the problems are also associated with these representation schemes e.g. artificial neural networks have capability to learn on the basis of data they have observed, but their disadvantage is that they can not give consistent representation of the knowledge that is not relevant to their learnt knowledge (Kong et. al, 2008).

The structural representation uses the frame format to represent the widely accepted knowledge and was introduced by Minsky. CENTAUR and Arden Syntax use frame format to represent the knowledge (Kong et al, 2009).

After a study of the literature on knowledge based medical decision support systems in general and on the knowledge representation scheme and systems developed on the basis of these schemes, issues have been identified that cause in an inaccurate diagnosis.

2.2.2 Study of tools/technologies used in the implementation of knowledge bases and their limitations

Knowledge that has been represented in the form of rules or any other formalism in a knowledge base is actually encoded in the memory of computer by means of programming languages and tools. The knowledge represented or encoded in the medical knowledge base is useless if it can not be accessed or retrieved accurately (Shortliffe, 1986). The deficiency of a programming tool may restrict the retrieval of logical and related information. Further significant issues have been identified relevant to tools and technologies in chapter 3.

2.2.3 Identification of issues in the learning of knowledge base

Another deficiency that pertains to the accurate retrieval of diagnostic information lies in the features of a particular learning scheme used. The learning techniques and algorithms are meant to create relationships between the symptoms and diseases. Artificial neural networks learn on the basis of data they have observed, but behave in an inconsistent way for the knowledge that is different from the learnt knowledge (Kong et. al, 2008). Issues related to learning techniques such as artificial neural networks and genetic algorithms have been discussed and presented in chapter 3

2.2.4 Study of issues related to reasoning components of the knowledge based medical decision support systems

Although the reasoning component of a medical decision support system is not a part of its knowledge base, yet they both are tightly coupled with each other. Reasoning in knowledge based systems is typically based on the representation of knowledge. More formal representation of knowledge provides better reasoning support about the decision support system. A lot of studies are available covering the complexities in medical tasks. The reason for these complexities may be differences in expertise in reasoning with the system and also a variation of the reasoning strategies used by the physicians (Patel & Groen, 1986). A study of d reasoning strategies such as case based reasoning and rule based reasoning in medical diagnostic decision support systems has been carried out to investigate the issues leading to imprecise diagnosis

Study of knowledge representation schemes, tools and technologies, learning schemes/techniques, reasoning components and issues related to them has been presented in next chapter.

2.3 Research Method

To carry out the proposed research on identification of design issues in the medical decision support systems, literature survey has been conducted. More than 120 research articles on knowledge based medical decision support systems were collected and around 71 articles were selected for the study. Articles were selected on the basis of their relevance to the desired aim and objectives.

Search engines and different research databases were used to find the relevant research articles on knowledge based medical decision support systems in general and the design issues of their knowledge bases in special. Following resources were accessed and used to accomplish the task:

- Google search engine

- Google Scholar
- ELIN database
- ACM Portal
- IEEE Explore
- Citeseer
- Science Direct
- Scopus
- Academic Search Elite

2.4 List of keywords

The following keywords were used to find the relevant literature using different search engines and databases:

- MDDS Design issues
- Design issues of knowledge bases
- Design limitations
- Constructing knowledge bases
- Knowledge base limitations
- Knowledge representation in MDDS
- Limitations of knowledge representation
- Logic representation problems
- Learning issues in knowledge bases
- Learning in artificial neural networks
- Limitations of supervised learning
- Limitations of unsupervised learning
- Limitations of knowledge repositories
- Problems with declarative languages
- Limitations of case based reasoning
- Limitations of rule based reasoning
- Shortcomings of reasoning strategies
- Implementation of knowledge bases
- Knowledge base implementation tools
- Design by contract
- Semantic web in medical diagnosis
- Protégé

3. Literature Survey

A variety of factors cast their profound effects in retrieving accurate advice from the knowledge based medical system. Those factors as described in section 1.2 include knowledge representation, tools and technologies used for implementing knowledge bases, learning techniques and the reasoning mechanism of the knowledge based system

To identify the design issues specific to knowledge bases of the medical decision support systems, literature on medical decision support systems was studied. This section presents the issues and limitations of the factors contributing to the design of knowledge bases of decision support systems in general and medical decision support systems in particular.

3.1 Study of issues related to knowledge representation schemes

Knowledge representation deals with providing information to the intelligent systems belonging to a particular domain for efficient processing (Carter, 1999). Clinical decision support systems use various formalisms for representation of medical knowledge, but still research on this important component of a knowledge base design has not been a driving force for clinical decision support in the last few years (Peleg & Tu, 2006).

Carter (1999) has classified the knowledge representation schemes into logic, procedural, graph/network and structured knowledge representation schemes. Also there exist other knowledge representation schemes such as temporal and spatial knowledge representation schemes (Allan, 1984; Shahar & Musen, 1993), but this research has not focused on the issues associated with these representation schemes. As stated in previous section, each of the knowledge representation mechanisms mentioned above has advantages and disadvantages associated with it.

3.1.1 Study of Logic representation scheme

Logic is one of the ancient mathematical and philosophical representations of the knowledge. With concrete syntax and vocabulary logic is considered as one of the most mature problem solving mechanisms (Bingi et. al, 1995). Knowledge representation schemes based on logic usually consist of declarative statements with Boolean operators such as AND, OR and NOT and TRUE, FALSE structure (Kong et. al, 2008). Medical diagnostic systems have used various formalisms for knowledge representation including causal, anatomic, taxonomic, heuristic, functional and safety models (Lucas, 1995). These systems include CASNET/GLAUCOMA (Kulikowski & Weiss, 1982), ABEL

(Patil et al., 1982), Oxford System of Medicine (OSM) (Fox et al., 1990) and DILEMMA system (Huang et al., 1993) etc.

Causal model works on principle *Cause (x, y),* which means that y is effect of cause x (Lucas, 1995). Example of causal model of logic in medical diagnosis is presented below. The symptoms for disease malaria are mapped to generate relationship between them as presented below and presence of all the symptoms in patient simply indicates the malaria.

A = fever, B = headache, C = Vomiting, D = Dizziness
A AND B AND C AND D -> Malaria

Although logic handles the incomplete knowledge in an impressive way yet there are limitations that are associated with the medical systems using logic as a knowledge representation mechanism. Concepts are represented in very general terms e.g. In DILEMMA system, despite of the fact that reasoning is performed in applying the concepts in stages by suggesting the candidate solutions and then further refining them (Lucas, 1995), yet such type of approach is not practically being used in medical and clinical diagnosis. The process of reasoning and diagnosis with such systems becomes delicate and requires more domain as well as technical knowledge by the physicians. Another disadvantage of the logic is that it is difficult to determine how to use the facts stored in the data structure of the system (Bingi et. al, 1995). The same was the reason for the incorrect diagnosis by INTERNIST-I, because it lacked in representing the relationship between the manifestations or symptoms and the diseases (Wolfram, 1995). Another limitation of logic scheme becomes apparent with the increase in number of facts in knowledge base, which ultimately increases the number of ways to join them, hence increasing complexity (Bingi et. al, 1995). Applying the same analogy on symptoms and diseases may result in imprecise diagnosis because many diseases may have lot of similar symptoms and to create meaningful relationship is difficult.

3.1.2 Study of Procedural representation

Procedural knowledge representation makes use of IF-THEN rules and is helpful in diagnosis and therapeutic decision making (Carter, 1999). Expert systems using rules as knowledge representation scheme in medical decision support systems domain include MYCIN, PUFF, UMLS and CMDS etc (Kong et. al, 2008). Rules have been the principal formalism for knowledge representation in medical expert systems to provide decision assistance to physicians in diagnosing diseases. Applying the symptoms of malaria in form of rules gives the following interpretation:

IF A AND B AND C AND D

THEN Malaria

Also an excerpt from the MYCIN rules (Melle, 1978) is presented below:

"**RULE036:**

PREMISE: *($AND (SAME CNTXT GRAM GRAMNEG)*

(SAME CNTXTM ORPHR OD)

(SAME CNTXT AIR ANAEROBIC))

ACTION: *(CONCLUDCEN TXTID ENTITY BACTEROIDETSA LLY. 6)*

IF: 1) *The gram stain of the organism is gramneg AND*

2) The morphology of the organism is rod, and

3) The aerobicity of the organism is anaerobic

THEN: *There is suggestive evidence that the identity of the organism is bacteroides."*

Figure 2: An excerpt from the MYCIN rule, source (Melle, 1978)

However, some problems are coupled with procedural representation of knowledge in medical expert systems because representation of relationships of medical facts such as symptoms, diseases and medicine etc may be somewhat difficult.

A major inefficiency of a procedural knowledge representation scheme becomes apparent when there is a large search space (Barr & Feigenbaum, 1981). Since the knowledge is represented in modular form, thus the search efficiency depends on how correctly a particular rule has been applied and inference has been made. As the knowledge domain grows large, the probability of selecting and applying the appropriate rule decreases. Consequently the accuracy of inference engine counteracts the accuracy achieved by using problem solving heuristics (Baldwin & Kasper, 1986). The applicability of this limitation is vital for knowledge based medical decision support systems as they are all about creating and identifying relationships among the symptoms, diseases, medicine etc. The second problem with the procedural knowledge representation based on rules is that rules are very poor in expressing the incomplete knowledge (Reichgelt, 1991; Davis, Buchana & Shortliffe, 1977). The reason claimed for this inability is that it is difficult to follow the control flow in programming languages than the algorithmic representations (Bingi et. al, 1995). This might be due to the inherent limitations of a representational language, but obviously to interpret the partial and deficient information might not be possible even for a perfect programming tool. Another problem with rule based knowledge representation systems in medical domain is that they do not create a correlation between the clinical signs and symptoms due to existence of pre-symptoms and post symptoms (Kong et. al, 2008), finally resulting in immature inference and diagnosis. The fourth shortcoming of procedural knowledge representation identified by Davis et. al, (1977) while evaluating the MYCIN expert system is regarding the control structure that is based on backward chaining. Davis et. al. (1977, p. 33) justify their claim by stating that:

> "*It is not always easy to map a sequence of desired actions or tests into a set of production rules whose goal-directed invocation will provide that sequence. Thus, while the system's performance is reassuringly similar to some human reasoning behavior, the creation of appropriate rules which result in such behavior is at times non-trivial*"

3.1.3 Study of Network representation scheme

As stated earlier, different types of networks such as Bayesian belief networks, decision trees and neural networks have been used in clinical decision support systems for representing knowledge (Kong et. al, 2008). Despite of their capabilities of handling conditional dependencies and probabilistic inference (Montani & Terenziani, 2006), still there representational mechanism lacks in certain aspects. A brief description of working of Bayesian Networks and artificial neural networks has been provided below to highlight issues related to them.

Bayesian Networks also called as Bayesian belief networks or causal probabilistic networks are based on mathematical representation of knowledge and are useful in representing conditional dependencies (Kong et. al, 2008; Miller, 1994). Various medical diagnostic systems based on Bayesian networks have been developed for diagnosis of various diseases like cancer and also having applications in radiology and ICU. In these types of networks relationship between symptoms and diseases is represented through nodes and paths between them. For example, the symptoms for malaria and hepatitis are as follows:

Malaria A = fever, B = Headache, C = Bodyache, D = Vomiting
Hepatitis A=Fever, B=Headache, C=Malaise, D=Vomiting, E=Jaundice

It can be seen that some symptoms like fever, headache and vomiting are common. In order to represent the symptoms for both the diseases the knowledge may be codified as follows

A=Fever, B=Headache, C= Bodyache, D=Vomiting, E=Malaise, F= Jaundice

The symptoms and diseases are represented in the network form as in following figure

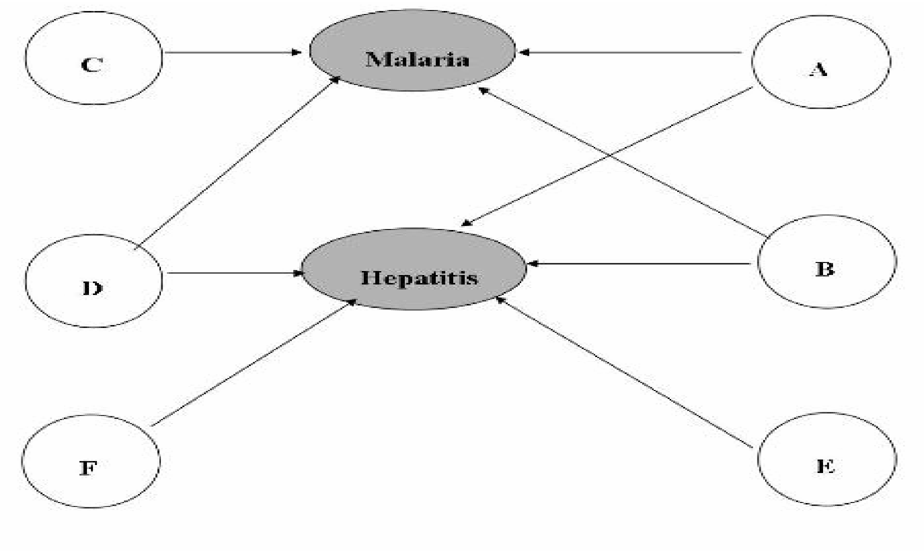

Figure 3: Representing relationship between symptoms and diseases (Author's own)

The given network can calculate the probabilities of existence of various diseases based on probabilistic relationship. However, a major problem with such networks is that as these networks consist of nodes and paths, so to make an inference for a node having more than one path is NP hard or computationally intractable (Miller, 1994). It ultimately results in confusion in correctly diagnosing the disease. Niedermayer (2008) describes the problem of Bayesian Networks while discussing their applications in general. If that problem is applied in medical domain then it can be stated that, an unforeseen symptom of a disease may not be correctly handled by the medical system, because the system may contravene the probabilities upon which it is developed

Another type of network based knowledge representation scheme discussed is artificial neural networks. A variety of artificial neural networks have been used in clinical decision support applications including stochastic networks, recurrent networks and feedforward neural networks etc. Artificial neural networks are constructed by representing the knowledge in a way similar to that of a human brain. Artificial neural networks are composed of neurons for processing the information. These networks have been widely used in medical decision support systems for knowledge representation and making inference (Maglogiannis, 2009)

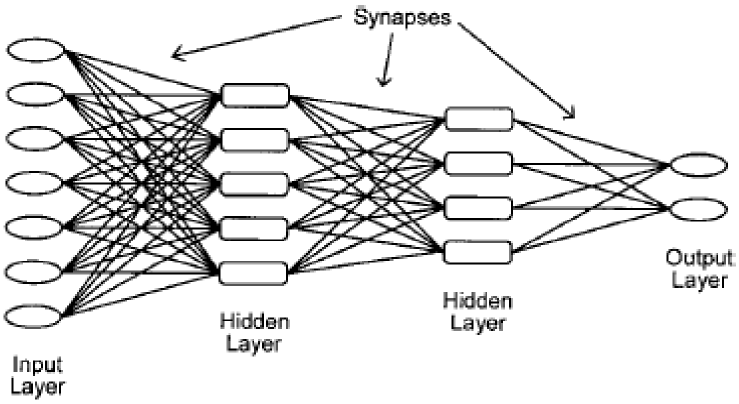

Figure 4 : An artificial neural network, Source Rodvold et al. (2001)

Symptoms are provided as training data to the network along with the intensity of various symptoms in the form of weights. Diseases are already provided as outputs. The error of the network is computed by computing difference from the output. Output is considered to be correct when it is much closer to the target output of the network. The process of adjusting weights continues till a correct diagnosis. Although artificial neural networks are capable of learning on the basis of observed samples of data, but start behaving abnormally when they come across with unknown knowledge (Kong et. al, 2008; Maglogiannis, 2009). This behavior of neural networks is because they try to solve a problem by generalizing from already solved examples

3.1.4 Study of Structural representation scheme

The structural knowledge representation uses the frame format to represent the widely accepted knowledge. CENTAUR and Arden Syntax use frame format to represent the medical knowledge. Recent medical decision support systems have used relational databases and object oriented databases to store the knowledge. Although DBMS are a better option to store the knowledge that is certain or uncertain and even have functions like add, update and delete, but the problem with them is that to infer a logical conclusion they do not have reasoning capabilities (Kong et. al, 2008).

In fact, frames give a concise representation of relations to represent the actual knowledge in the knowledge base possessed by an expert system (Bingi, Khazanchi, &Yadav, 1995). Structural representation uses slots to represent the symptoms in medical systems along with extended knowledge related to them. The following excerpt from the WHEEZE system frame is an example of structural representation (Smith & Clayton, 1980, p. 154):

OAD with Smoking:

Manifestation: *((OAD-Present 10) (PatientHasSmoked 10) (PatientSillSmoking 10))*

SuggestiveOf: *((SmokingExaccrbatedOAD 5) (SmokingInduccdOAD 5))*

ComplementaryTo: *((OADwithSmoking-None 5))*

Certainty: *1000*

Findings: *Discontinuation of smoking should help relieve the symptoms*

Figure 5: An excerpt of WHEEXE frame from Smith et. al (1980)

The numbers in Manifestation, SuggestiveOf and ComplementaryTo slots in fact represent the importance and suggestivity weightings (Smith et. al, 1980)

One problem with frame based representation scheme is that it does not have apparent semantics (Reichgelt, 1991) and is complex. Obviously, for a system that is complex in representing the concepts, the tasks of reasoning and inference also become obscure and this deficiency too prevails in knowledge based medical decision support systems. The other disadvantage of frame based structural knowledge representation scheme is that its ability of expressing the incomplete and partial knowledge is limited (Reichgelt, 1991). So, for knowledge based medical decision support systems that are already based on prototypes and hypothesis with less number of predefined conditions, incorrect diagnosis is possible. Also (Niwa, Sasaki & Ihara, 1984) claim that despite of their run time efficiency, structural knowledge representation systems are difficult to implement in practice

3.2 Study of issues related to tools and technologies

The discussion presented in the preceding section on knowledge representation schemes, reveals that the retrieval of an accurate diagnostic advice from the knowledge based system is dependent on representation of knowledge. The techniques discussed above are general problem solving procedures that are symbolic instead of numeric to represent relevant knowledge (Shortliffe, 1986). However, this representation of knowledge requires encoding in the programming languages. These languages are not just tools for knowledge representation, instead they work as problem solving agents when the knowledge has been represented using them (Chandrasekaran, 1986). Therefore, it makes sense to consider the issues related to tools and technological frameworks used for the implementation of knowledge bases. Usually declarative languages and databases have been used to encode and store the expert knowledge in the knowledge bases

3.2.1 Declarative languages

A declarative language represents problem oriented information in the form of axiomatic rules (Sol, 1982). Knowledge based medical decision support systems have been implemented using declarative languages such as LISP and PROLOG etc. Since, the knowledge is encoded using any programming language (Sol, 1982), so it is debatable that the inadequacy of the tool or programming language restricts the retrieval of accurate information from the knowledge base. MYCIN was developed in LISP (Buchanan & Shortliffe, 1984). Also the knowledge base editor of INTERNIST-I was implemented in LISP (Miller, 2009). Although declarative languages are good in manipulating qualitative knowledge, still they have some issues. Declarative languages are imperfect in their representational features (Sol, 1982). PROLOG for example has one problem of performing redundant computations (Ramakrishna et al, 1997). Another limitation of the PROLOG is that its semantics for negation are weak (Ramakrishna et al, 1997) that makes difficult for it to make a correct and logical conclusion.

3.2.2 Knowledge base repositories

Besides the techniques used for representation and encoding of knowledge into the knowledge base, there is a need to store the knowledge in the knowledge base. Conventionally, knowledge repositories of medical decision support systems have been implemented using relational databases, for example, the knowledge base of a decision support system developed for ordering drug by detecting drug- drug interactions, was implemented using a relational database (Fiol et al., 2000). Also, recently developed medical decision support systems have frequently made use of relational and object oriented databases management systems for storage and management of structural knowledge. (Kong et al. 2008).

Irrespective of the advantages of relational and object oriented databases, the focus of this research is to identify those issues that are causes of problems in accurately providing the diagnostic advice. Although DBMS are a better option to store the knowledge that is certain or uncertain and even have functions like add, update and delete, but the problem with them is that to infer a logical conclusion they lack reasoning and knowledge based deduction capabilities (Kong et. al, 2008). Another limitation of databases is that they exhibit poor performance in handling complex queries, which ultimately causes in poor optimization of queries (Devarakonda, 2001). Moreover, the problem of using database as knowledge base becomes more noticeable, when there is uncertainty in the knowledge

(Levseque & Brachman, 1987). To handle uncertainty is important in medical decision support systems because the temporal events may be less expressive and uncertain in nature. For example if a condition such as described by Keravnou (1996, p. 246) is encountered stating:

> "*I had flu last week which lasted exactly two days*"; "*I had flu last week; I started feeling unwell about Tuesday midday and I was not fully recovered until about Thursday morning*".

Since starting and end times of particular symptoms of flue are uncertain consequently it becomes rather difficult for the inference module of the medical decision support system to make precise diagnosis. Because, despite of having similar symptoms, the duration of persistence of those symptoms may be different, that ultimately results in inaccurate diagnosis. There are other numerous limitations and disadvantage of using relational and object oriented databases in general, but they are not of concern here and have not been discussed further

3. 3 Study of issues related to learning

Knowledge based systems update their knowledge bases by learning new knowledge. A variety of learning schemes have been used in knowledge based system for knowledge bases to learn new knowledge and update their existing knowledge. These techniques are based on artificial neural networks, Bayesian networks and genetic algorithms etc. Features and limitation of some learning algorithms are described below

3.3.1 Learning problems in Artificial Neural Networks

Artificial neural networks have long been used in medical diagnostic systems. These networks are composed of neurons capable of performing some specific tasks of computation (Wei, 1998; Musen, Shahar & Shortliffe, 2001). However the task of learning and updating can be accomplished by standard learning techniques of neural networks (Towell, 1993). Artificial neural networks use either of the learning techniques, supervised or unsupervised (Lee, Booth & Alam, 2005). In supervised learning, the network is provided a training set (input vector) along with the desired output pattern (target vector), whereas in unsupervised learning training set is provided but the correct behavior is not included.

The most widely used learning technique for supervised learning in medical decision support systems is based on multi layer perceptron (MLP)(Wei, 1998). Multi layer perceptron consists of layers including the input layer, output layer and the hidden layers. By providing a certain set of data as input to the network, the network is trained to produce the desired output (Maglogiannis, 2009). The output samples are also provided and weights are assigned on each link connecting the neurons (Musen et. al, 2001). Training data is provided to the network repeatedly. The output data is compared with the desired or target output and their difference is assumed as error. This error is propagated backwards and weights are adjusted again to become closer to the output. In this way network is trained by propagating the error in backward direction (Wei, 1998).

Despite of the fact that artificial neural networks can learn knowledge, this learning is not insignificant because we may not even be able to predict at which point it starts to fail (Maglogiannis, 2009). Therefore, it might be possible for a system to generate incorrect relationship between the finding of the patient and the possible diagnosis (Musen et. al, 2001). Another limitation of the supervised learning scheme is that it is not scalable. When the training set or input size becomes extremely large, the network becomes slow (Becker, 1991). The third limitation of supervised learning as stated by (Becker, 1991) is regarding the biological plausibility of supervised learning. It is improbable that human brain uses the back propagation scheme to train the hidden layer neurons.

Moreover, medical systems have been developed using unsupervised learning scheme (Dokur & Ölmez, 2008). Kohonen self organizing maps use unsupervised learning. These maps arrange nodes in the form of a two dimensional grid. Output is determined by resemblance to the neighbor. However, there are some problems associated with unsupervised learning scheme. As it is not known to the network in advance that how many output nodes will be there so it is difficult for the network to determine the stopping point (Dokur et. al, 2008). The other problem of unsupervised learning networks is that information extracted from such networks is less accurate (Lee et. al, 2005)

It becomes obvious that, the schemes used for encoding the knowledge in a knowledge base and learning scheme used to learn new knowledge and update existing knowledge are significantly related to each other.

3.3.2 Limitations of Genetic Algorithm (GA) for MDSS

Although GA fits well for the medical decision support systems, because any internally unknown complex system with variable behavior is the best candidate for using GA. However it does not solve any such problem where the computation of output is unknown in certain conditions and this is the basic requirement for any medical decision support system. The other possible limitation of GA is to define the fitness function, which is really very complex and difficult in the medical decision support system. Furthermore, (Holland, 1992; Forrest, 1993; Haupt & Haupt, 1998) have proposed not to use genetic algorithms on analytically solvable problems, because GA gives exactly one solution at the end, which is usually not feasible for any mathematical problem, where always finite set of solutions exist, particularly for biological systems/functions.

3.4 Study of issues related to reasoning strategies

In general, reasoning is a way of thinking that enables a person to make ‘wise’ decision that actually is the best possible action leading towards correct solution (Cervero, 1991). Although the reasoning component of a medical decision support system is not a part of its knowledge base, yet they both are tightly coupled with each other. Reasoning in knowledge base systems is typically based on the representation of knowledge. More formal representation of knowledge provides better reasoning support in the decision support system. Reasoning in medicine is generally performed by creating a relationship of the symptoms and signs along with some predefined procedures stored in knowledge base.

Knowledge based medical decision support systems make use of any particular reasoning strategy depending upon the specific knowledge representation scheme. Historically, medical decision support systems have used rule based reasoning, case based reasoning, temporal reasoning and model based reasoning or combination of more than one reasoning strategies such as multi-modal reasoning (Montani & Bellazzi, 2002). Working principles of case based reasoning and rule based reasoning have been generally discussed in the upcoming paragraphs along with their limitations. Case Based Reasoning (CBR) as stated by (Schmidt, 2001, p. 355) works on the principle that:

> *“Similar problems have similar solutions. Though this assumption is not always true, it holds for many practical domains”.*

CBR tries to solve a real world problem in a way similar to real world. When a new case is to be resolved, the Case Based Reasoning uses already solved cases in the particular domain (Montani & Bellazzi, 2002). The ability of applying past cases to the current scenario in fact is representation of the human experts’ analogical reasoning (Montani & Bellazzi, 2002). Different medical decision support systems based on case based reasoning have been developed such as CASEY, PROTOS and MEDIC etc.

Nonetheless, cased based reasoning systems also have some limitations such as high cost of searching the knowledge base, indexing problems, inaccuracy in inference, knowledge acquisition problems, knowledge maintenance and adaptation to new situation (Panday & Mishra, 2009; Kolodnor, 1991; Leake, 1996). For a CBR system to work correctly, it requires an appropriate representation of knowledge and correct algorithm for retrieval of knowledge from the knowledge base (Schmidt, 2001). However, the major limitation of the CBR system is related to adaptation to the new situation (Schmidt, 2001). Solving adaptation problem is important because, as previously solved cases are stored in case library. If the physician enters the symptoms that are similar to an already solved case, it simply retrieves that case (Montani & Bellazzi, 2002). But if the symptoms entered are different from the previous one, it becomes difficult for the system to adapt to new situation, which ultimately results in an incorrect diagnosis.

Another reasoning strategy being used in medical decision support systems is rule based reasoning (RBR). This reasoning strategy is based on a set of predefined protocols and rules stored in a rule base. Medical decision support systems such ad INTERNIST-I and MYCIN etc are rule based systems.

Traditional Rule based systems consist of a repository of rules, an inference engine that triggers the rules (Ligeza, 2006) and a working memory containing available data, which is updated every time a rule is applied. These rule based reasoning systems though once were successful but now seem inadequate, because of their inflexibility and the changing trends in the clinical decision making as now most of the systems being developed are very domain specific. The reasons that make these systems inflexible are their complete dependence on knowledge base and the other most important one is that this knowledge base takes even years for a single domain to be populated (Kumar, Singh & Sanyal, 2009). Also rule based reasoning systems are deficient in specializing the knowledge represented through rules e.g. the symptoms observed in a patient This limitation of the rule based reasoning causes in an incorrect inference, because the knowledge is represented in the form of rules and rules do not have competency to interpret the incomplete knowledge ((Montani et. al, 2002; Davis et. al, 1977).

Although the problems of the two reasoning strategies discussed above do not directly relate to the design of knowledge base, but are tightly coupled with each other to make a correct diagnosis. Knowledge representation scheme used to represent knowledge is dependent on the reasoning strategy used by the knowledge based system to perform reasoning. Correct retrieval of a contextual advice is dependant upon the correct knowledge representation and selection of appropriate reasoning strategy.

The discussion presented in this section has been summarized in the table 1 below and reveals that there is not only a single design factor that leads to false diagnosis by the medical decision support systems. Besides tools and technologies various factors contribute in this respect such as learning techniques, reasoning strategies and finally specific techniques used for representing the knowledge, although mostly having one common limitation of not expressing the incomplete knowledge. The next section presents analysis of the design factors that are critical to the accuracy of medical decision support systems and discusses how the semantic web can be a better choice for representation of medical knowledge.

Table 1 Summary of Issues related to knowledge representation, tools and technologies, learning and reasoning strategies

<table>
<tr><th colspan="4">Issues related to Knowledge representation schemes</th></tr>
<tr><th>Logic issues</th><th>Procedural representation Issues</th><th>Network representation Issues</th><th>Issues related to Structural representation</th></tr>
<tr><td>• Represents the relationship between symptoms and diseases in very general form.
• Difficult to determine how to use the facts stored in systems' data structure
• Increase in number of facts makes solution search space larger.</td><td>• Poor in expressiveness
• Problems with backward chaining
• Decrease in probability of applying the appropriate rule with large search space</td><td>Bayesian Networks
• Existence of more than one paths becomes computationally hard – results in incorrect inference
• In case of contradiction between probabilities, problem remains undiagnosed
ANN
• Behave unusually when come across unforeseen situations</td><td>• Problems in expressing incomplete knowledge.
• Unclear semantics</td></tr>
<tr><th colspan="4">Issues related to tools and technologies</th></tr>
<tr><th colspan="2">Declarative languages</th><th colspan="2">Knowledge base repositories</th></tr>
<tr><td colspan="2">• Perform redundant computations
• Weak semantics for negation
• Computationally less efficient</td><td colspan="2">• Inability to inference
• Inability to handle complex queries
• Problems in dealing with uncertain situations</td></tr>
<tr><th colspan="4">Issues related to learning techniques</th></tr>
<tr><th colspan="2">Learning in ANN</th><th colspan="2">Learning in Genetic Algorithm</th></tr>
<tr><td colspan="2">Supervised learning
• Non trivial – difficult to predict the point of failure
• Biologically non-plausible
Unsupervised learning
• Difficult to determine the stopping criteria
• Less accurate due to absence of correct output</td><td colspan="2">• Performs inadequately in unknown conditions
• Difficult to select appropriate fitness criteria</td></tr>
</table>

behavior	
Issues related to Reasoning strategies	
Case based reasoning	**Rule based reasoning**
• Inference accuracy • Adaptation task • Knowledge acquisition and maintenance	• Inability to express incomplete knowledge • Largely depends on knowledge base

4. Analysis

In the preceding section knowledge representation schemes, tools and technologies for implementing knowledge bases, learning techniques and reasoning strategies used in medical decision support systems were discussed. This section analyzes each of the above objectives of the study.

As described earlier, the cause of poor and inaccurate diagnosis by the medical decision support systems is the design of knowledge base which is dependant on the four factors that have been discussed in chapter 3 as objectives of the study. The discussion presented is evident that each of the above factors whether that is part of the knowledge base itself or not, contributes to the poor diagnosis by the system. For example, each of the knowledge representation schemes carries with it some limitations that ultimately cause in inaccurate diagnosis. Causal model of logic, when applied to medical decision support domain does not seem appropriate because in this model every symptom (input) has a corresponding disease (output), which makes the solution set very large and hence results in imprecise diagnosis. In an ideal medical decision support system this set should be small to create the mapping between symptoms and diseases. Similarly, the inexpressiveness of rules used in procedural representation is sufficient to affirm it as non-preferable representation scheme, though the past medical systems have made excessive use of this representation scheme. Network based representation schemes discussed such as Bayesian networks and neural networks when used in medical decision support have also issues associated with them. In Bayesian networks the diseases and symptoms are represented using nodes having relationship between them like ‘is-a’ and ‘has-a’ etc. It becomes computationally hard or NP complete if there exist more than one path between nodes. Due to this limitation this might not be a preferable technique for knowledge representation. Using artificial neural networks as knowledge representation scheme is also not feasible in medical systems since in the medical decision support system, it is essential to identify the correct relationship between the entities of the system, which are disease, symptoms and medicine. This relationship is very complex to be elaborated by using just an artificial neural network. As the requirement of the artificial neural network is to input the precisely measured parameters, therefore it is not possible to apply them directly. The structural knowledge representation scheme also has somewhat similar limitations like inexpressiveness and unclear semantics. As the task of diagnosing a diseases has become very complex, so the knowledge representation schemes discussed do not seem adequate for correctly identifying the relationships among the various attributes of the clinical data such as pre-symptoms, post-symptoms, their intensities and values etc. This naturally demands for knowledge representation formalism, capable of handling such complexities. The limitations of tools and technologies turn out to be another design issue for the knowledge base because they too seem inadequate to meet the current computational and accuracy demands of the decision support systems. Obviously, there arises a need to identify the more suitable technologies that should be able to identify the clear and transparent boundary among various diseases – it's challenging, as most of the diseases have common symptoms at early stage. The learning techniques utilized by medical decision support systems e.g. artificial neural networks and genetic algorithms also seem less adequate. Since, a relationship between the symptoms and diseases has to be defined with identification of new characteristics to learn new knowledge. However this relationship is very complex to be represented by using just an artificial neural network or genetic algorithm. Artificial neural networks require the precisely measured parameter therefore it is not possible to get the accurate results by applying those parameters directly. Similarly for genetic algorithm, the selection of appropriate fitness function is difficult and is not feasible for analytically solvable problems. Therefore it necessitates the identification of learning scheme capable of learning the rules of diagnosis in more robust way as humans. The last objective of the study discussed above is reasoning. Case based reasoning and rule based reasoning have been discussed. A large number of medical decision support systems have made use of these reasoning strategies, but they lack in various aspects. For example, the inability to handle temporal aspects of medical data and uncertain knowledge makes them less feasible phenomenon for reasoning. Moreover, the problems of adaptation to new situation and less expressiveness of rule based systems demand better alternatives for reasoning in order to overcome the accuracy issue. The four objectives of the study are coupled with each other; however the knowledge representation schemes and reasoning are more critical. No doubt all the knowledge representation schemes discussed above have since long been used, but one major issue that becomes obvious is that they define the concepts in a more general way, having little ability of being domain specific. In such situations reasoning becomes difficult.

From the above discussion, we conclude that the competence of representation, tools and technologies, learning and reasoning mechanisms discussed above are not adequate enough to demonstrate any real medical decision support system. This is because of wide gaps in their capabilities of representing and manipulating complex and domain specific tasks. Now, it will be discussed how this gap can be abridged by a possible proposed solution. Besides representation, the reasoning component is one of the of most critical parts of any medical decision support system, because the practitioner must know the way how system has diagnosed a particular disease to convince him/her self about the diagnoses performed. The reasoning can be more precise and formal if we have strictly adequate specification for the underlying knowledge representation scheme, so the binding of knowledge representation scheme and the reasoning component is obvious. To accomplish these objectives, the conditional equation logic along with an adequately specified contract (among the pre-symptoms, post-symptoms and the invariants of the disease) can be used as the knowledge representation scheme. The first order conditional equation logic is comprehensive enough to represent any model, but the real challenge will be how can we formally specify the contract which is more adequate? This contract can be followed by the design by contract approach (Mayer, 1992). This adequacy can be achieved by proving the few properties of the contract, for example that the invariants of the disease hold for any of the given pre-symptoms, and that the post-symptoms are the logical consequence of the pre-symptoms and the invariants.

Another possible representation can be in semantic web, as it is also based on first order logic, but then in proving this formal specification will be an overhead in conditional logic, as one needs to translate the semantic web representation to the logical formulas, which can obviously be avoided by the direct representation.

The objective of tools and technologies can be accomplished by using the existing semantic web enabled tools/technologies, for example Ontology (OWL, 2009) as technique since it is good in defining the concepts and the relationship between them. Also Protégé (Gennari et al, 2003) can be used as a tool, if we use the semantic web as knowledge representation. Another abstraction layer can be integrated while using the conditional logic as a representation. Any tool/technology can be recommended which supports the proving of the adequacy of the specification, because ultimately an implementation turns out into software.

Since learning technique are also tightly coupled with the representation scheme and the incentive of using ontology as representation scheme is that concepts are learned implicitly by creating the relationship among them. One can think of the weighted/attributed learning strategy, which will account some external parameters like probabilities for occurrence of some diseases based on weights of some attributes. For instance, several possibilities might come across while learning process occurs. If for particular set of symptoms presented to the systems there are already solved cases based on those symptoms, no learning will be required in this case. The second possibility is that the symptoms of disease presented to the system do not match with solved cases; obviously learning is required in this case by storing new knowledge. The third possibility might be that presented symptoms conflict with the existing symptoms partially and learning becomes non-trivial in such cases. This conflict can be either temporal conflict related to persistence of symptoms for a particular disease or in weights of attributes of different symptoms. For example the duration of the same disease can be different for two different patients. In such cases it is difficult for the medical system to distinguish exactly what could be the correct conclusion. Using stochastic methods or probabilities with the attributed learning might be helpful in learning and differentiating such knowledge. This ultimately can lead to unambiguous reasoning and diagnosis.

5. Conclusion

The purpose of this study was to identify the design issues of the knowledge bases of medical decision support systems. To identify these issues study of knowledge representation techniques, tools and technologies, learning schemes and reasoning strategies was carried out as the objectives of the study. The task of diagnosing diseases accurately using medical decision support systems is difficult due to the complex nature of the medical data and their relationship with each other. Therefore, knowledge representation techniques, tools and technologies, learning schemes and reasoning strategies discussed in this study seem inadequate to meet this complexity –which ultimately cause in the poor design of knowledge base. This obviously necessitates the use of adequate specification for knowledge representation, more formal reasoning and knowledge based learning. We propose to employ design by contract approach for representing adequate relationships among the entities of medical data. This representation can be accomplished by proving a contract between the pre-symptom, post-symptoms and the invariants. To support such design for knowledge representation, formal reasoning and knowledge based learning using semantic web techniques and tools such as ontology and protégé appear more desirable.

5.1. Future work

The research contributes to the body of literature by identifying the design issues related to knowledge bases of medical decision support systems and finally suggesting that design by contract approach can be feasible for identifying and creating the desired relationships among the entities of medical systems. To cope with this challenging aspect semantic web tools and technologies can be used since semantic web represents the information in meaningful way, by introducing a cooperative approach and even permits sharing and reuse of data across applications and enterprises (W3C, 2010). Introducing logic to the semantic web for using the inference rules and reasoning enables it to handle more complex situations like those related to medical diagnosis applications. Various semantic web applications have been developed using different technologies such as XML, RDF and OWL etc. The concepts in the semantic web are represented using ontology. Ontology is a term that has origin from the philosophy. Gruber (1993) defines it as "*explicit specification of a conceptualization*". Ontology represents any concept of the real world domain and is suitable in representing the relationships between the domain entities and concepts. Considering the complex nature of medical data as described in the previous sections, ontology may be the most appropriate technique for representing the medical knowledge. The research on semantic web based medical decision support systems is already in progress and ontologies have been developed for representing the relationship between the entities of medical data for example by Kim & Choi (2007) etc. Moreover, different tools are available for developing ontologies such as Protégé. It was originally developed for knowledge acquisition for medical planning. Protégé is good in managing temporal information in medicine and representing knowledge in the form of guidelines. Protégé has strong affiliation with semantic web and is being used for developing RDF ontologies (Gennari et al, 2003). Consequently it can fulfill the requirement of the suitable tool. As described earlier that more adequately specified knowledge representation results in more formal and precise reasoning with system which ultimately overcomes the inaccuracy problem.

The most popular knowledge representation schemes, tools, learning and reasoning mechanisms were covered in this research. However not all of those described in the widely scattered literature on these issues in general and medical decision support systems in particular were discussed. Therefore, a possible direction for the future research work is to encompass issues specific to all those factors such as temporal and spatial medical knowledge etc. Obviously this can provide concrete foundation for future knowledge based medical systems, once all the issues have been identified and gathered at a single point.

6. References

Allen, J. F. (1984). Towards a general theory of action and time *Artificial Intelligence*, 23(2), pp. 123–154.

Baldwin, D., Kasper, G.M. (1986). Towards representing management-domain knowledge, *Decision Support Systems*, *2(2)*, pp. 159-172

Bamett G.O., Cimino J.J. ,Hupp J.A., Hoffer E.P.,(1987) DXplain An evolving diagnostic decision-support system, *JAMA*, *258*, pp. 67-74

Barr, A., Feigenbaum, A.E (1981). The handbook of artificial intelligence Vol. 1. Los Altos, CA William Kaufman.

Becker S. (1991) Unsupervised learning procedures for neural networks *International Journal of Neural Systems 2*, pp. 17-33

Bingi R., Khazanchi D., Yadav S.B. (1995). A framework for the comparative analysis and evaluation of knowledge representation schemes, *Information Processing & Management*, Vol. 31, pp. 233-247

Buchanan B.G., Shortliffe E.H. (1984). Rule-Based Expert Systems *The MYCIN Experiments of the Stanford Heuristic Programming Project*, 149-158

Carter JH (1999), in Clinical Decision Support Systems E. S. Berner, Ed, *Springer-Verlag*, New York, pp. 169-198

Cervero, R. M. (1991). Changing relationships between theory and practice I*n J. M. Peters, & P. Jarvis & Associates, Adult education Evolution and achievements in a developing field of study San Francisco*, pp. 19–41

Chandrasekaran B (1986), Generic tasks in knowledge based reasoning High- level building blocks for expert system design, - *IEEE expert*, pp. 23-30

Clancey W.J (1983). The Advantages of Abstract Control Knowledge in Expert System Design *in Proceedings of the National Conference on AI, Washington DC*, pp. 74-78

Claudia A. Perry, (1990) Knowledge bases in medicine a review, *Bull Med Libr Assoc 78(3)*, 272-282

Davis R, Buchanan B, Shortliffe E (1977), Production rules as a representation for a knowledge-based consultation program, *Artificial Intelligence 8*, 15-45

Devarakonda R.S (2001), Object-Relational Database Systems - The Road Ahead, *portal.acm.org*, Accessed on 28-04-2010

Dokur Z , Ölmez T (2008) Heart sound classification using wavelet transform and incremental self-organizing map, Digital Signal Processing 18, 951-959

Durinck J, Coiera E, Baud R. (1994). The role of knowledge based systems in clinical practice, in P. Barahona, J.P. Christensen (eds), *Knowledge and Decisions in Health Telematics*, 199-203

Fiol et al., (2000), Comparison of Two Knowledge Bases on the Detection of Drug-Drug Interactions, *AMIA*, 171-175

Forrest S. (1993) Genetic algorithms principles of natural selection applied to computation, *Science, vol.261*, pp. 872-878

Fox, J, Gordon, C, Glowinski AJ and O'Neil, M, (1990) Logic engineering for knowledge engineering the Oxford System of Medicine, *Artificial Intelligence in Medicine 2*, pp. 323-339

Gennari H.J. et al.,(2003)The evolution of Protégé an environment for knowledge-based systems development, *Int. J. Human-Computer Studies 58*, pp. 89–123

Goggin L., Eikelboom R, Atlas M. (2007) Clinical decision support systems and computer-aided diagnosis in otology. *Otolaryngology - Head and Neck Surgery*, Volume 136, Issue 4, S21-S26

Gruber T. R. (1993) A Translation Approach to Portable Ontology Specifications *Knowledge System laboratory Technical report KSL*

Holland JH (1992), Genetic algorithms, *Scientific American*, July 1992, p. 66-72.

Huang, J, Fox, J, Gordon, C and Jackson-Smale, A, (1993) Symbolic Decision support in medical care, *Artificial Intelligence in Medicine* 5(5) pp. 415-430.

Kalogeropoulos A.D., Carson ER, Collinson P.O. (2003). Towards knowledge-based systems in clinical practice Development of an integrated clinical information and Knowledge management support system. *Computer Methods and Programs in Biomedicine 72*, 65-80

Keravnou E.T (1996), Temporal diagnostic reasoning based on time objects, *Artificial Intelligence in Medicine 8*, 235-265

Kim K. H, Choi H.J., (2007) Design of a Clinical Knowledge Base for Heart Disease Detection, *Seventh International Conference on Computer and Information Technology*, pp. 610-615

Kolodnor J. (1991). An introduction to case based reasoning, *Artificial intelligence review 6(1)*, 3-34

Kong G, Xu D. L, Yang J.B (2008) Clinical decision support systems a review on knowledge representation and inference under uncertainties, *International Journal of Computational Intelligence Systems, Vol.1, No. 2*, 159-167

Kumar A. K, Singh Y., Sanyal S. (2009). Hybrid approach using case-based reasoning and rule-based reasoning for domain independent clinical decision support in ICU *Expert Systems with Applications 36*, pp. 65-71

Kulikowski CA.,Weiss, SM(1982) Representation of expert knowledge for consultation the CASNET and EXPERT projects, In P Szolovits (ed.) *Artificial Intelligence in Medicine,* pp 21-55,

Leake D.B (1996), CBR in Context The Present and Future, in In Leake, D., ed., *Case-Based Reasoning Experiences, Lessons, and Future Directions*. Menlo Park

Leake D.B, Kolodner J.L (1996). A tutorial introduction to CBR, in Case *Based Reasoning Experiences, Lessons and Future Directions*, AAAI Press, 1996, pp. 31-65

Ledley R. S, Lusted L. B. (1959) Reasoning foundations of medical diagnosis, *Science,* 130, pp. 9-21

Lee K., Booth D., Alam P. (2005). A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms, *Expert Systems with Applications 2,* pp. 1–16

Levseque HJ, Brachman R. J. (1987), Expressiveness and tractability in knowledge representation, *Computational Intelligence 3,* pp*. 78-93*

Ligeza A, (2006). Logical Foundations for Rule-Based Systems Studies in Computational Intelligence, vol. 11, Springer, Berlin

Lobach F., Hammond W.E (1997) Computerized Decision Support Based on a Clinical Practice. Guideline Improves Compliance with Care Standards, *AM J Med*, pp. 89-98

Lucas P. (1995). Logic Engineering in Medicine *The Knowledge Engineering Review*, 10(2), 153-179

Maglogiannis I. (2009). Introducing Intelligence in Electronic Healthcare Systems State of the Art and Future Trends, in M. Bramer (Ed.) *Artificial Intelligence*, LNAI 5640, pp. 71 – 90,

Melle V.W (1978), MYCIN A knowledge based consultation program for infectious disease diagnosis *International Journal of Man-Machine Studies 10 (3),* pp. 313-322

Meyer B (2009) Applying, "Design by Contract. *IEEE Computer* 25(10), pp. 40-51

Miller R.A, (1994) Medical diagnostic decision support systems-past, present and future. *J. Am. Med. Inform. Assoc*. 18–27.

Miller R.A (2009) Computer-assisted diagnostic decision support history, challenges, and possible paths forward, *Adv in Health Sci Educ.* 14, pp. 89–106

Montani S.,Bellazzi R (2002). Supporting decisions in medical applications the knowledge management perspective *International Journal of Medical Informatics* 68, pp. 79-90.

Montani S,Terenziani P (2006), Exploiting Decision Theory Concepts within Clinical Guideline Systems Toward a General Approach, *International Journal of Intelligent Systems 21*, pp. 585-599.

Musen M.A., Shahar Y., Shortliffe E.H. (2001). Clinical decision support systems, In Shortliffe EH, Perreault LE, Wiederhold G, Fagan LM, eds. *Medical informatics computer applications in health care and biomedicine*. 2nd ed. New York Springer-Verlag pp. 698-736.

Niedermayer D. (2008).An Introduction to Bayesian Networks and their Contemporary Applications in D.E. Holmes and L.C. Jain (Eds.) *Innovations in Bayesian Networks*, SCI 156, pp. 117–130

Niwa K, Sasaki K, Ihara H (1984). An Experimental Comparison of Knowledge Representation Schemes, *AI Magazine Volume 5 Number 2*, pp. 29-36

OWL (2009) Ontology Web Language, W3C Specification, URL http //www.w3.org/TR/owl-guide/, Accessed on May 09, 2010

Pandey B, Mishra R.B (2009), “Knowledge and intelligent computing system in medicine”, Computers in Biology and Medicine 39, 2009, pp. 215-230

Patil, RS, Szolovits, P and Schwartz, WB, (1992). Modelling knowledge of the patient in acid-base and electrolyte disorders", In P Szolovits (ed.), *Artificial Intelligence in Medicine*, pp. 191, 226

Peleg M, Tu S (2006), Decision Support, Knowledge Representation and Management in Medicine, pdf Accessed on March 20, 2010

Perreault L, Metzger J (1999) A pragmatic framework for understanding clinical decision support. *Journal of Healthcare Information Management.*13 (2) 5-21.

Purcell GP. (2005). what makes a good clinical decision support system. *BMJ;* 330(7494) pp. 740-741

Ramakrishna et al. (1997). Efficient model checking using tabled resolution *In Orna Grumberg, editor, Proceedings of the Ninth Conference on Computer Aided Verification* (CAV'97), volume 1254 of LNCS, pp. 143-154

Haupt R., Haupt E.S. (1998). Practical *Genetic Algorithms*. John Wiley & Sons

Reichgelt, H. (1991). Knowledge representation An AI perspective. Norwood, NJ Ablex.

Ridderikhoff J., Herk V. (1999). Who is afraid of the system? Doctors’ attitude towards diagnostic systems. *International Journal of Medical Informatics 53*, pp. 91–100

Rodvold et al. (2001). Introduction to Artificial Neural Networks for Physicians Taking the Lid Off the Black Box, *The Prostate 46*, pp. 39-44

Schmidt R., Montani S., Bellazzi R, Portinale L.,Gierl L.(2001). Cased Based Reasoning for medical knowledge based systems, *International Journal of Medical Informatics* 64, pp. 355–367

Shahar Y, Musen M.A (1993). RESUME A temporal abstraction system for patient monitoring, *Computers and Biomedical Research* 26, pp. 255-273

Shortliffe E.H. (1986). Medical expert systems-Knowledge tools for physicians, *In Medical informatics J Med* 145, 830-83

Shortliffe E.H. (1987). Computer programs to support clinical decision making, *JAMA 258,* 61-66

Sintchenko V, Gilbert GL, Coiera E, Dwyer D (2002) Treat or test first? Decision analysis of empirical antiviral treatment of influenza virus infection versus treatment based on rapid test results. *Journal of Clinical Virology, Volume 25,* 15-21

Sol H.G., (1982) Process and tools for decision support *Proceedings of the Joint IFIP WG 8.3 IIASA Working Conference on Processes and Tools for Decision Support Schloss Laxenburg, Austria, 19-21 July*

Smith D.E, Clayton J.E (1980), A Frame-Based production System Architecture, *AAAI Conference,* pp. 154-156

Towell G.G., (1993). Extracting refined rules from knowledge base neural networks, *Machine Learning* 13, pp. 71-101

Wei J.T., et al., (1998). Understanding artificial neural networks and exploring their potential applications for the practicing urologist, *Elsevier Science*

Wolfram D.A., (1995) An appraisal of INTERNIST-I, *Artificial Intelligence in Medicine* 7, pp. 93-116

W3C (2010) Source http // www.w3.org/2001/sw/ Accessed on May 19, 2010