Review of Reasoning Methods in Clinical Decision Support Systems

Liljana Aleksovska - Stojkovska, *Member, IEEE,* Suzana Loskovska, *Senior Member, IEEE*

***Abstract*—Reasoning is a crucial task performed by the inference engine of the clinical decision support systems, which combines medical knowledge with patient specific data and generates relevant decisions. There are different reasoning methods, suitable for different knowledge representations and application area. This paper reviews the most common methods and describes how they are used in real systems. Furthermore, it outlines the remaining weaknesses of the reasoning mechanisms and provides directions for future research and improvements.**

***Keywords* — Clinical decision support systems, inference, inference engine, reasoning methods**

I. Introduction

Reasoning or inference can be defined as "act of using reason to derive a conclusion from certain premises using a given methodology, following the logical definition" [1]. The reasoning process in medical diagnostic is a complex process as it has to consider many different facts, including the patient's history, current symptoms, test results, received therapies, possible allergies and map these conditions to a list of possible matching diagnoses. The importance of understanding this process has been emphasized by the increasing interest in the use of computers as an aid to medical diagnostic processes [1]. When a physician is asked to describe the medical diagnostic process, his explanation may be as follows [1]. First, the case facts are obtained from the patient's history, medical examination and laboratory tests. Second, evaluation is performed of the relative importance of the different signs and symptoms. Third, to make a differential diagnosis, list is created of all diseases with reasonable resemblance to the specific case. Then, the list is evaluated and the less relevant diseases are excluded until the case can be fitted into a definite disease category.

This is obviously a simplified explanation of the diagnostic process [1]. Physicians describe that often after seeing the patient they have a "feeling" about the case, which may be a summary of the patient's general appearance, facial expression, how all facts fit together or it may be a great resemblance to a previous case [1]. This is especially true for the experts with long experience, who have developed highly automated perceptual ability and rapid recognition skills [2]. It is believed that the physicians have tendency to generate hypotheses very early in the process of solving a diagnostic problem and this is a result of complex reasoning, which is performed automatically without being aware of the different reasoning methods that are applied [2].

Scientists have developed many different reasoning mechanisms, which are available to be used by the inference engine within the Clinical Decision Support Systems (CDSS). However, even decades since the CDSSs were initially introduced, there are still unresolved problems and no single method has been found to answer all questions. The main purpose of this paper is to review the different reasoning methodologies and to provide directions for future research and improvements.

The paper is organized as follows. Section 2 reviews some of the most common reasoning methods. Section 3 describes how these methods are combined and used in real decision support systems. Section 4 describes the weaknesses of the reasoning mechanisms, followed by Section 5, which provides ideas how some of these problems can be solved. Finally, Section 6 concludes the review.

II. Review of the Reasoning (Inference) Methods

Medical decision making can be viewed along a spectrum, with categorical (or deterministic) reasoning at one extreme and probabilistic (or evidential) reasoning at the other [3]. The categorical reasoning uses an appropriate set of predefined routines or rules, which apply to the great majority of clinical situations [3]. A categorical decision is simple to make, its appropriateness is easy to judge and its result is unambiguous [3]. For example: if the patient complains of pain on urination, obtain urine culture and consider the possibility of urinary tract infection [3].

Unfortunately, most of the medical decisions are not that simple. For example, no simple rule exists for deciding when to discharge a patient from intensive care unit after suffering heart attack [3]. In this kind of situations, decisions are made by careful consideration of all the evidence, which means estimating the probability of all the different facts that are included in the specific case [3].

Liljana Aleksovska – Stojkovska, MSc., 701 Huntington Commons Rd. #214, Mount Prospect, IL, 60056, USA, (tel: 1-847-739-6485, e-mail: liljana.a.stojkovska@gmail.com); curently PhD student at the Faculty of Electrical Engineering and Information Technologies, University „Ss. Cyril and Methodius" – Skopje, Republic of Macedonia

Suzana Loskovska, PhD., University „Ss. Cyril and Methodius – Skopje", Faculty of Electrical Engineering and Information Technologies, Karpos II, bb, 1000 Skopje, Republic of Macedonia, (tel: +389 70 390 842, e-mail: suze@feit.ukim.edu.mk)

Following are some of the most common reasoning mechanisms, along the scale between categorical and probabilistic.

Rule-based reasoning is based on “if-then-else” rule statements, which are viewed as patterns and the inference engine searches for patterns in the rules that match patterns in the data [4]. The rules reasoning can be forward-chaining, (data-driven), where the reasoning starts with data or facts and looks for rules which apply to the facts until a goal is reached or backward-chaining (goal-driven), where the reasoning starts with a goal and look for rules which apply to that goal until a conclusion is reached [4].

Case-based reasoning searches for commonly occurring patterns among the stored various cases [7]. It is used when the medical knowledge is difficult to be modeled with the formal representation methods [4]. Crucial requirements for the success of this method are good similarity measures and efficient ways of searching for similar cases [7]. Advantage of case-based reasoning is that similar clinical cases are often more convincing than the theoretical medical knowledge, but, the difficulty to measure the similarity between cases and the complex retrieval process are disadvantages [4].

Model-based reasoning provides framework for diagnosing an artifact by observing its manifested behavior and comparing it to the predicted behavior of the artifact’s model [8], [9], [10]. It assumes that if the model is correct, all discrepancies between the observed and the predicted behavior arise from the defects in the observed device [10]. However, the real problem with this methodology is that there is no guarantee that the model is correct and it is challenging to come up with a right model, especially in case of complex systems.

Bayesian reasoning is based on conditional probabilities and predicts the posterior probability of specific event given the occurrence of another event [5]. In context of CDSS, Bayesian network can be used to compute the probabilities of the presence of possible diseases given their symptoms.

Heuristic reasoning methods exploit the information processing structure of the reasoning system and the structure of the environment to produce reasonable answers when knowledge and/or computational resources for finding the perfect answer may not exist [6]. Heuristic systems proposed for medical reasoning include statistical methods, such as support vector machine (SVM) and least square support vector machine (LSSVM) [4].

Semantic network is a declarative graphic representation in patterns of interconnected nodes and arcs that can be used to support automated systems for reasoning about knowledge [7]. In CDSS, since most medical knowledge involves uncertainties, it is difficult to use a pure semantic network to make clinical inference [4].

Neural networks are a black box modeling technique that model relationships by learning from historical data and patterns recognition [4]. Advantage of neural networks is that it is not necessary to understand the relationship between input and output variables [4]. A disadvantage is that their knowledge domain is very limited and they can not explain their reasoning process, so most clinicians don’t use them for reliability and accountability reasons.

Genetic algorithms are based on simplified evolutionary processes using directed selection to achieve optimal results. The selection algorithms evaluate components of random sets of solutions to a problem. The solutions that come out on top are then recombined and mutated and run through the process again. This process is continuously repeated till the proper solution is discovered. Their advantage is that similar to the neural networks, they derive their knowledge from patient data and the most optimal solution can be achieved, but determining what is the fittest solution is a challenge [4].

From this review it can be concluded that different reasoning mechanisms have different advantages and disadvantages and can be more or less suitable for specific systems. The focus of research is to find appropriate ways to combine these methods to create competent programs which exhibit medical expertise [3]. It is common that different reasoning methods are frequently combined, such as: integration of rule-based and case-based reasoning, integration of model-based with case-based reasoning [12].

III. Inference Mechanisms Used in CDSS

This section provides examples how the reasoning methods are used by the inference engines within CDSS.

INTERNIST is a computerized diagnostic program, designed at the University of Pittsburgh in 1974, which covers approximately 80 % of the diagnoses of internal medicine [3], [13]. The INTERNIST database associates every possible diagnosis D_i a set of manifestations $\{M_j\}$. A manifestation is a finding, symptom, sign, laboratory test or another associated diagnosis. For every M_j listed under D_i two likelihoods are entered [3]:

- $ES_{Di|Mj}$ - Evoking Strength (ES) is the likelihood that if Manifestation M_j is seen in a patient, its cause is D_i. It is assessed on a scale 0-5, where 5 is the maximal likelihood, while 0 virtually means no support.
- $FW_{Mj|Di}$ – Frequency Weight (FW) is the likelihood that a patient with a confirmed diagnosis D_i would exhibit manifestation M_j. It is assessed on a scale 1-5, where 1 means that the specific manifestation rarely occurs in the disease, while 5 means that the manifestation essentially occurs in all cases.

Fig. 1 shows how a Diagnosis and its Manifestations are represented in INTERNIST.

Disease profile for
ECHINOCOCCAL CYST <S> OF LIVER

ES	FW	Symptom
1	2	COUGH
1	1	FECES LIGHT COLORED
0	2	FEVER
1	3	HEPATOMEGALY PRESENT
1	2	JAUNDICE
1	2	LIVER TENDER ON PALPATION

Fig. 1. Diagnosis and its Manifestations in INTERNIST

INTERNIST classifies all diagnosis in a disease hierarchy, which is used during the diagnostic process for determining a general diagnosis and the corresponding specializations. INTERNIST identifies for each general diagnosis a list of manifestations and computes their corresponding Evoking Strengths and Frequencies. The manifestations of the general diagnosis are common for each of its specializations and the ES and FW are respectively the maximum ES and the minimum FW of that manifestation among the specializations [3]. The diagnosis is active if there is at least one manifestation with a non-zero ES. For each active hypothesis a score is computed and the system focuses on the highest ranking diagnosis [3]. The lower ranking diagnoses are categorized in two groups: competing and complementary. The diagnosis is complementary to the chosen one if the two together account for more findings than either one alone, otherwise it is competing [3]. The process of scoring, ranking and partitioning is repeated after each new fact is identified. INTERNIST uses combination of categorical and probabilistic reasoning. The decision making is mostly probabilistic. The hierarchic tree of diagnoses and the rule for moving from general to specific diagnosis are categorical.

MYCIN is a rule-based expert system designed to diagnose and recommend treatment for certain blood infections, developed in the mid-1970s by Ted Shortliffe and colleagues at Stanford University [13]. It uses goal-directed backward-chaining reasoning, while trying to determine if the patient is suffering from significant infection, which should be treated and propose appropriate therapy [3].

MYCIN's knowledge base is composed of rules, with a typical rule shown on Fig. 2.

```
   param      context  operation value
   ---------- -------- --------- --------
(defrule 52                               <--- rule #
  if (site     culture   is      blood )
    (gram      organism  is      neg  )   <--- remise/
    (morphl    organism  is      rod  )       condition
    (burn      patient   is      serious)
  then 0.4                                <--- cf
    (identity  organism  is  pseudomonas ) <---conclusion
)
```

Fig. 2. A typical MYCIN rule

With each fact H in the database, MYCIN associates measure of belief (MB) and measure of disbelief (MD), which are numbers between 0 and 1 [14]. The difference between these numbers gives the Certainty Factor (CF), which ranges between -1 and 1 and is calculated with the following formula [3]:

$$CF_H = MB_H - MD_H \quad (1)$$

The value of every clinical parameter is stored by MYCIN along with an associated certainty factor (CF), which reflects MYCIN's "belief" that the value is correct [14]. MYCIN uses confirmation formalism for computing the certainty of multiple facts [3]. When the Rule S1 infers the hypothesis H with measure of belief $MB_{H|S1}$ and the Rule S2 leads to the same hypothesis H with measure of belief $MB_{H|S2}$, the certainty for the hypothesis H is calculated with the following formulas [3]:

$$MB_{H|S1,S2}=0, \text{ if } MD_{H|S1,S2}=1 \quad (2)$$

$$MB_{H|S1,S2}= MB_{H|S1}+ MB_{H|S2}(1- MB_{H|S1}), \text{ otherwise} \quad (3)$$

$$MD_{H|S1,S2}=0, \text{if } MB_{H|S1,S2}=1 \quad (4)$$

$$MD_{H|S1,S2}= MD_{H|S1}+ MD_{H|S2}(1- MD_{H|S1}), \text{ otherwise} \quad (5)$$

The certainty factor of hypothesis when multiple rules have contributed evidence to it, is calculated as in [3]:

$$CFH= MB_{H|S1,S2} - MD_{H|S1,S2} \quad (6)$$

In case of multiple active hypotheses, their conjunction and disjunction is calculated with these formulas [14]:

$$MB_{H1andH2}=\min(MB_{H1}, MB_{H2}) \quad (7)$$

$$MD_{H1andH2}=\max(MD_{H1}, MD_{H2}) \quad (8)$$

$$CF_{H1andH2}=\min(CF_{H1}, CF_{H2}) \quad (9)$$

$$MB_{H1orH2}=\max(MB_{H1}, MB_{H2}) \quad (10)$$

$$MD_{H1orH2}=\min(MD_{H1}, MD_{H2}) \quad (11)$$

$$CF_{H1orH2}=\max(CF_{H1}, CF_{H2}) \quad (12)$$

In addition to the rules, MYCIN also includes context hierarchy, which plays smaller, but still important role in the system's operation [3]. Similar to INTERNIST, MYCIN uses combination of categorical and probabilistic methods. The rule-based knowledge presentation and backward-chaining are purely categorical [3]. Probabilistic reasoning resides in calculation of measure of belief, measure of disbelief and certainty factor [3].

IV. Critique and Analysis

As presented in this paper, there is a number of different reasoning methodologies, which can be combined and used within real clinical diagnostic systems. It has to be admitted that choosing the right method or combination of different methods to create an efficient inference engine, well suited for its application area, still remains a challenging task. This section analyses some of the remaining problems, which need additional research and solution.

Systems with broad domain, such as INTERNIST, have inadequate criteria for deciding when the diagnosis is complete [3]. The systems have no sense when the main diagnostic problem is resolved and continue wasting time and resources by exploring less and less suitable hypotheses.

The initial strategy of the diagnostic systems is to use every new finding as a trigger for discovering a possible disorder, so new hypotheses are continually generated [3]. When an expected finding is requested to support one of the leading diagnoses and when that finding is presented, the system often generates a completely new hypothesis, even though the finding is consistent with the diagnosis being considered. Obviously, such sensitivity is needed to a certain point, but it is preferred if the new hypothesis is triggered only by evidence contradictory to the current belief.

When multiple hypotheses are being considered, the diagnostic systems are not very good at distinguishing the different possibilities: the two hypotheses occurring together but are unrelated, one hypothesis was caused by the first one or only one of the hypothesis is active but not both [3], [15].

The medical reasoning necessarily involves uncertainty, which mainly comes from the incomplete or imprecisely specified knowledge about important facts that are contributing to a decision. Therefore, effective methods should be developed to handle the uncertainty in medical reasoning. When solving complex problems where multiple uncertain factors have to be considered, the probabilistic models are necessary [3]. The number of probabilities to be calculated is huge. For example, in a system which considers 20 diseases and 50 symptoms, there should be 1000 probabilities that the symptom *j* is associated with the disease *i* [15]. On the other side, it was observed from the experience that the experts, even the most experienced ones, are extremely reluctant in engaging in any type of numerical computation involving likelihood of diagnosis or prognosis of treatment [3]. While doctors can confidently make quick decisions about diagnoses or treatment just from observing the patient or taking a quick look at their medical facts, these hypotheses usually come as a result of recognition or comparison with similar cases from the past experience rather than from any formal computation of probabilities.

Ongoing research must be performed in the direction of these weaknesses to make the clinical decision support systems as close as possible to the human experts.

V. Directions for Possible Improvements

This section provides some ideas and directions for future research, which may lead to resolving some of the difficulties discussed previously.

Regarding the appropriate termination of the diagnostic process, [3] proposes that the diagnosis needs to be only as precise as is required by the next decision to be taken by the doctor. This can be achieved by applying the Bayesian methods which can compare the cost of the new information to its expected benefit and ignoring the remaining possible diagnoses, which are irrelevant [3], [15]. For this purpose a termination probability is calculated and if the disease attains a probability higher than the termination probability, the further processing is terminated and only the leading diagnoses with highest probabilities are presented [15].

The diagnostic systems shall have mechanisms to check if the newly presented fact is related to a currently established diagnosis or an existing chronic illness [3]. For example: in case of patient with a long history of sickle cell anemia, who presents a new symptom of joints pain, the system should not immediately raise a hypothesis that the patient suffers from rheumatoid arthritis, but it should realize that it is a reasonable consequence of already known disease process [3]. This approach also may be helpful in the complex cases where the patient is suffering from more than one disease and the symptoms are overlapping. Also, the diagnostic systems must be aware of the ongoing trends of illnesses and epidemics, so in cases when patient presents with symptoms of known infection disease in the local community, the possibility of that infection shall be immediately checked before any other analysis are performed [15].

Regarding the problem with uncertainty and calculation of the probabilities, statistical methods can be used to obtain the likelihoods directly from the patient's database. It is believed that this approach would give much more accurate numbers compared to the numbers provided by the physicians.

VI. Conclusion

The inference engine represents the brain of the CDSS, therefore selecting the right inference mechanisms is crucial in determining the success of the entire system. After the critical review of the existing reasoning methods and their use within real CDSS, a conclusion can be drawn that significant progress has been made theoretically and practically. There are numerous reasoning methods developed, which can be combined to get optimal results. Additional research must be performed to improve these methods and overpass the existing weaknesses.

CDSS promise to make a significant impact in the medicine. However, because of the numerous limitations, they have to be considered only as assistants to the human experts rather then expected to replace them completely.

References

[1] R.S. Ledley, L.B. Lusted, "Reasoning foundations of medical diagnosis", in *Science,* vol. 130, no. 3366, Jul. 1959, pp. 9-21.

[2] V. L. Patel, G. J. Groen, "Knowledge based solution strategies in medical reasoning", in *Cognitive Science*, vol. 10, no. 1, Jan-Mar 1986, ISSN 0364-0213, pp. 91-116.

[3] P. Szolovits, S. G. Pauker, in "Categorical and Probabilistic Reasoning in Medical Diagnosis", in *Artificial Intelligence*, vol. 11, no. 1-2, 1978, pp. 115-144.

[4] G. Kong, D. L. Xu, J.B. Yang, The University of Manchester, "Clinical Decision Support Systems: A Review On Knowledge Representation And Inference Under Uncertainties", in *International Journal of Computational Intelligence Systems*, vol.1, no.2, May 2008, pp. 159-167.

[5] Charles River Analytics, Inc., "About Bayesian Belief Networks", for *BNet* Version 1.0, Last Updated April 22, 2008.

[6] P. K. Paritosh, "The heuristic reasoning manifesto", in *Proceedings of the 20th International Workshop on Qualitative Reasoning*, 2006

[7] J. F. Sowa, editor, "Principles of Semantic Networks", Morgan Kaufmann, Los Altos, 1991.

[8] J. de Kleer, J. and B. C. Williams, "Diagnosis with behavioral modes", in *Proceedings IJCAI-89*, Detroit, MI (1989) pp. 104-109.

[9] J. de Kleer, and B.C. Williams, "Diagnosing multiple faults", *Artificial Intelligence,* vol. 32, 1987, pp. 97-130.

[10] R. Davis, and W. Hamscher, "Model-based reasoning:Troubleshooting", in *Exploring artificial intelligence*, edited by H.E. Shrobe and the American Associationfor Artificial Intelligence, (Morgan Kaufman, 1988), pp. 297-346.

[11] V. Novacek, "Inference Support for Ontology Acquisition", Dissertation Thesis Topic, Faculty of Informatics, Masaryk University, Czech Republic, Jan 3, 2007.

[12] C. Marling, E. Rissland, A. Aamodt, "Integrations with case-based reasoning", in *The Knowledge Engineering Review*, vol. 00:0, 2005, pp.1–4, Cambridge University Press.

[13] Open Clinical: Decision Support Systems, Mar 2010 Available: http://www.openclinical.org/dss.html

[14] Exert Systems Case Studies: MYCIN Available: http://www.computing.surrey.ac.uk/ai/PROFILE/mycin.html

[15] G. A. Gorry, J. P. Kassirer, A. Essig, and W. B. Schwartz, "Decision Analysis as the Basis for Computer-Aided Management of Acute Renal Failure", in *The American Journal of Medicine,* vol. 55, Oct. 1973, pp.473-484.