DiagFH: An Expert System for Diagnosis of Fulminant Hepatitis

Wei Lin  
Dept. of Computer Science & Engineering,  
University of South Florida,  
Tampa, FL 33620, U.S.A.

Jian-Xong Tang  
Computer Science Dept., Engineering College  
Shanghai University, Shanghai 200032, China

Abstract

DiagFH is a consultation system for the diagnosis of fulminant hepatitis. The expert’s knowledge is represented by a tree-like structure which is divided into two classes of facts: certain and uncertain facts. The inferences on the facts is implemented by using logical and probabilistic strategies. The entire design is based on the simulation of clinical practice.


I. Introduction

Since the 1970s artificial intelligence has been increasingly applied to the medical field. Many medical expert systems have been developed [1,2,3,4] and are used in diagnoses of diseases. These systems have the following advantages:
1. The experience and skills of many experts can be combined into one system, resulting in a “super-doctor”.
2. This computer “expert” can consider exhaustively all diseases in its domain without omission of any possibility.
3. There are some tasks that the computer expert can do more rapidly and accurately than the clinician, especially in some qualitative measure.

The implementation of an expert system for medical diagnosis is mainly determined by the knowledge representation of the specified medical domain and the inference engine. The structured representation of a causal model has displayed competence in many recent designs [5,6]. Based on the causal of pathophysiological knowledge, the system can describe the evolution of disease, infer the complications, and reason about interaction among diseases. But the medical science is incomplete and some relations are inexact. It is difficult to completely describe the deep causal model. Though a series of uncertainty reasoning methods have been presented in recent years [7,8,9], successful applications are still rare in practice. The obstacle is that the uncertainty description and propagation are difficult to be fixed precisely with the medical convention and clinical practice.

DiagFH is a consultation system for the diagnosis of fulminant hepatitis. In our design, a tree-like structure is used to represent the diagnostic knowledge of fulminant hepatitis. The knowledge of DiagFH consists of a causal model of pathophysiological knowledge for the evolution of fulminant hepatitis, and some of the expert’s practical experience and medical convention. Because of the inexact causalities between some manifestations and diseases in medical knowledge of fulminant hepatitis, DiagFH allows all the possible relations to exist in the knowledge base. In some cases, an uncertainty factor for a relation is stored...
along with the premise conditions (manifestation) and the conclusion (disease). We calculated the uncertainty factor for the relation using data from the Shanghai Rei-Jing hospital. This data was compiled from the case histories of hundreds of infected fulminant hepatitis patients who were treated over an eight year period. Because all possible relations exist in the KB, this approach will cause multiple-goals matched in reasoning. So, in order to handle this situation, a few principles of reasoning about probabilistic knowledge, and a set of logical rules is employed by the inference engine.

DiagFH consists of user interface, knowledge base, database, and inference engine. The system is implemented in TURBO PROLOG, and runs on IBM-PC or compatible.

II. System Description

The structure of DiagFH is shown in Fig.1. The system is driven by the user interface which is composed of various forms and menus.

Two windows are designed in the user interface. One is used for man-machine interactive input and output; the other for help messages. In some cases a third window is used for additional inquiries. The input and output information are consistent with the clinical conventions.

The user interface manipulates the input and output. DiagFH asks the user to provide medical information. This information is divided into three parts: 1) patient’s symptoms, 2) signs and 3) medical laboratory tests. While some user interfaces are purely dialogue-based, DiagFH is a form-based system in which the user completes forms with the patient’s symptoms, signs and laboratory test results. In this way the doctor-machine interaction is succinct, and thereby circumvents the tedium of dialogue interfaces. The help window of DiagFH provides hints for the user to follow at every step. This feature makes the system easy to use and avoids the need for special training. For a full consideration of diseases, the user is encouraged to fill in the forms as completely as possible. When each input step is finished, DiagFH will check the information and decide if further data is needed or if a diagnosis can be derived.

After the input phase, the system will carry out inferring and print the diagnostic conclusions. The output messages have the following types:

1. Positive assertion of the fulminant hepatitis, with its clinical category, viral causes and complications.
2. Negative assertion of the fulminant hepatitis.
3. Wait for further examinations. This means there is not enough information to conclude a diagnosis, and the system then suggests further examinations.
4. Keep patient on observation. This means that the diagnosis cannot be concluded from
the current evidence, however the system will provide some advice to the user.

5. Different diagnosis conclusion. The system prints other possible diagnoses.

The knowledge base is independent of the program and contains the expert’s knowledge.

It is implemented as a prolog file that consists of a large number of facts. When DiagFH runs, it opens knowledge base which provides the knowledge for machine diagnosis.

The input information of patients, which includes patient’s symptoms, signs and laboratory test results, is sorted and stored in a database. The system creates a data file for every patient. When DiagFH accepts new data or make a new diagnosis, the database is modified.

The inference engine controls the diagnosis of the system and is composed of a list of rules. Using the knowledge base and patient’s database, it decides what should be concluded, and whether additional inquiries are needed.

III. Knowledge Representation

The knowledge in DiagFH is described by a hierarchical tree structure with a variable number of levels. Fig. 2 gives an example. The concept of States Bilirubin is represented by an and/or tree. This tree is further split into a set of one-level trees as shown in Fig. 3(a). These small trees, however, can be easily described by facts in prolog. Two categories of facts for the knowledge representation of fulminant hepatitis are defined as follows:

1. Assertive knowledge.
   - knowledge(conclusion, conditionlist) (1)
2. Uncertain knowledge.
   - knowledge(conclusion, coef, conditionlist) (2)

The conditionlist is a list of conditions which are defined by a group of self-defined predicates[12]. If all the conditions in the list are satisfied, the conclusion will be true. Every conclusion is a non-leaf node in a knowledge tree. coef is a certainty factor which is a number from 0 to 1 and reflects the degree of the certainty of the conclusion when the conditionlist is matched. The coef is not subjective, it is assigned by the statistic of the accumulated clinical data.

```
knowledge(conclusion, conditionlist) (1)

knowledge(conclusion, coef, conditionlist) (2)
```

Fig. 2. The structure of Statis Bilirubin knowledge.
The corresponding facts for Statis Bilirubin knowledge is shown in Fig.3 (b).
The knowledge base is a forest of trees. It contains the following parts:

1. The assertive knowledge for a positive diagnosis of fulminant hepatitis. This knowledge states some distinguishing features that allow the immediate derivation of the fulminant hepatitis. This part is written using assertive facts.

2. The pathological model of fulminant hepatitis. This contains the knowledge for the possible complications and viral causes of fulminant hepatitis.

3. The medical laboratory test knowledge. This points out the relations among the test results and symptoms and signs of the typical patient. With this knowledge, DiagFH can decide what further tests should be performed, and investigates some potential errors in the clinician's examinations or in the laboratory tests.

4. The knowledge for different diagnosis, such as acute bilirubin hepatitis, chronic hepatitis, etc.

5. The probabilistic knowledge of fulminant hepatitis.

This knowledge represents the probabilities of the relations between manifestations and non-fulminant-hepatitis. The reason we are concerned with non-fulminant-hepatitis is because it is sometimes easier to exclude the possibility of fulminant hepatitis than to directly conclude it.

![Diagram of Statis Bilirubin knowledge structure](image)

Fig.3. A set of trees for Statis Bilirubin structure (a) and the knowledge representation of Statis Bilirubin (b).

The incompleteness widely exists in medical knowledge. To deal with this problem, some probabilistic approaches for uncertainty theories [9,10,11] have been explored in recent years.
A major difficulty is how to get the values of the probabilities. Conventional probability theory cannot work effectively with medical knowledge, because there are too many unknown properties, and the distribution of correlated events are not clear on the specified domain. In DiagFH, the probabilistic knowledge is described by uncertainty facts. In order to make the probability and statistics be valid, this part of the knowledge is based on a restricted range of the specified domain. The range is the collection of case histories of the infected fulminant hepatitis patients. We investigate the relation between the manifestation and the possibility of fulminant hepatitis infection. We assume that the principle of the probability,

\[ P(A) = 1 - P(A) \]  

(3)

is valid. Then the coef is defined as exclusive rate n by the following formula:

\[ n = 1 - P(A/E_i) = 1 - R_i * P(E_i) \]  

(4)

\( A \) --- infected fulminant hepatitis,
\( E_i \) --- patient's manifestations, \( i = 1, 2, 3, \ldots, n \).
\( P(E_i/A) \) --- determined by statistic of past patient's data.
\( R_i \) --- first determined by expert's experience and then adjusted by tested examples.

The fifth part of the knowledge base combines all of the uncertain facts in DiagFH as shown in (5).

\[ \text{knowledge-5(non-FH, } n_i, E_i \text{) } i = 1, 2, 
\ldots, n. \]  

(5)

This part can be used for negative diagnosis of fulminant hepatitis as well as supporting the positive diagnosis. For those inexact relations that cannot be specified by the restricted domain, DiagFH uses groups of assertive facts as shown in (1) to express all possible existences. Consequently, this knowledge representation will cause multiple-goals matched in reasoning. The inference engine in DiagFH, however, will resolve this conflict using a set of rules.

IV. Inference Engine

The inferring of DiagFH is accomplished by searching through the and/or trees in the knowledge base. The rules of inference are written by Horn clauses, such as:

\[ R1 \leftarrow \text{predicate1} \ \text{predicate2} \ldots \ \text{predicateN} \ (N>0) \]

In order to reduce the search time during the facts matching in the knowledge base, the knowledge trees are classified according to several subjects, each of which is described by the facts with a unique predicate name. So, although blind searching is used in DiagFH, it does not take much time for reasoning.

The inference engine performs reasoning in the following steps:

1. Searching the assertive knowledge for a positive diagnosis of fulminant hepatitis.
2. Searching the pathophysiological model of fulminant hepatitis, to see whether complications and viral causes can be found.
3. Searching the laboratory test knowledge.

The above three steps can yield the following conclusions:

1) When a positive diagnosis of fulminant hepatitis has been matched and no contradiction has occurred in step 2 and 3, the inference can be finished.
2) When no positive conclusion has been found, or a positive one has been caught but some
contradictions have also been produced in step 2 and 3, the next two steps for further reasoning
must be carried out.
4. Searching the knowledge for different diagnoses to see whether other relevant diseases
   can be matched.
5. Searching the probabilistic knowledge of fulminant hepatitis to see whether a negative
diagnosis can be derived.

<table>
<thead>
<tr>
<th>CONFLICTS</th>
<th>STRATEGIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>A positive diagnosis conclusion in knowledge part one is found, and from part two and three the viral causes are also found and the patient's manifestations can be understood by pathophysiological model, but from uncertain facts reasoning a negative diagnosis is produced.</td>
<td>The positive diagnosis conclusion is affirmatively concluded.</td>
</tr>
<tr>
<td>A positive diagnosis conclusion in knowledge part one is found, and from part two and three the viral causes are also found, but the patient's manifestations cannot match pathophysiological model or some reports of the laboratory tests can not be explained by the system.</td>
<td>A soft positive diagnosis conclusion is issued. And different diagnosis must be output. The system should suggest that the user to reexamine the input information of the patient.</td>
</tr>
<tr>
<td>No positive diagnosis conclusion in knowledge part one is found, but from the probabilistic knowledge, the fulminant hepatitis cannot be denied, and different diagnosis has successfully been matched.</td>
<td>A negative diagnosis conclusion with different diagnosis conclusion should be output.</td>
</tr>
<tr>
<td>No positive diagnosis conclusion in knowledge part one is found, but from the probabilistic knowledge, the fulminant hepatitis cannot be denied, and different diagnosis is unsuccessful.</td>
<td>Wait for further examination</td>
</tr>
</tbody>
</table>

Fig. 4. The inference rules for multiple-goals conflicts

The fact-matching may be carried out in one of two ways. In first method, the conventional
logic is applied to the assertive facts as defined in (1). A set of rules for logical matching is used
to check the different predicates in the conditionist. If all the predicates are true, the conditions are satisfied. Then, the precedent argument (i.e. conclusion) in the assertive fact is concluded. As DiagFH's knowledge is organized in a hierarchical structure, the conditionist may contain
the conclusions of other facts. Thus, the check of conditionlist would be recursive and will consume a great deal of time. In order to prevent the recursive searching of the facts too many times, a fact is placed before any other facts which may require its conclusion as their conditions.

The second method is for the uncertain facts which require numeric processing in addition to the conventional logic. When several uncertain facts share the same conclusion and their conditionlists are true, the following strategy is employed to draw a conclusion:

1) While $\max(n_1, n_2, ..., n_k) - \min(n_1, n_2, ..., n_k) < 0.5$, and $\max(n_1, n_2, ..., n_k) > 0.8$.

the node of the precedent argument is set up. Since in the current version of DiagFH, all uncertain facts in the knowledge base have the precedent argument of non-fulminant-hepatitis, this result yields a negative diagnosis of fulminant hepatitis.

2) While $\max(n_1, n_2, ..., n_k) - \min(n_1, n_2, ..., n_k) < 0.5$, and $\min(n_1, n_2, ..., n_k) < 0.2$.

result is that the fulminant hepatitis cannot be denied.

3) Otherwise, no conclusion.

$n_j$ is the exclusive rate in the uncertain facts. If $\max(n_1, n_2, ..., n_k) - \min(n_1, n_2, ..., n_k) = 0.5$

DiagFH thinks the current knowledge is incomplete. Hence, no conclusion can be induced.

Because of the special representation of the medical knowledge, the searching of DiagFH knowledge base can have several successful results. When this happens, several rules based on clinical practice are used to process the conflicts that are summarized in Fig.4.

V. Conclusion

DiagFH has been tested by approximately one hundred patients who had been suspected fulminant hepatitis infected and had liver biopsy. The DiagFH's diagnostic conclusions coincided with pathological reports. In addition, seven district hospitals in Shanghai had some of their fulminant hepatitis cases diagnosed by DiagFH. The results were satisfactory as well. DiagFH executes quickly, and the reasoning takes less than 3 seconds on PC/XT 286.

Since the system designed to simulate the medical expert's clinical practice, the DiagFH's diagnosis is very close to the conclusion of the medical experts who provided the knowledge for it.

The tree-structure model is effective for medical knowledge representation and easy to maintain. A combined method of logical and probabilistic inference can simulate the medical expert's reasoning in clinical practice. For the uncertain reasoning, the current version of DiagFH does not provide a complex propagation mechanism, and its inference engine is somewhat incomplete. But these shortcomings have not seriously affected its performance when it was tested by hundreds of patients. In fact, it is rather difficult to use a complete theoretic propagation model to describe the clinical practice, especially when the system is huge and the domain knowledge is complicated. However, if the system is not very large, such as DiagFH, the logical methods can make up for the shortcomings of an incomplete numeric reasoning mechanism.
Acknowledge

The fulminant hepatitis has a high death rate. Supported by the Chinese Science Committee, a national research program for the diagnosis and therapy of fulminant hepatitis has been set up for eight years. It is headed by professor Geng-Rong Shen of Shanghai No. 2 Medical University. The DiagFH project is funded by this program and Shanghai Rei-Jing Hospital. All the relevant medical knowledge was provided by professor Shen and other experts in the Rei-Jing hospital. We wish to express our appreciation to these experts.

We wish to thank Zhi-Tong Zhang, Jian Wen in Computer Science Department of Shanghai University who wrote the programs for process the data of the patient history cases and collected knowledge.

In addition, we greatly appreciate Timothy S. Pilsbury, Eddy Ng and Dr. L. O. Hall in the department of computer science and engineering of University of South Florida for their contribution to the revision of the paper.

Reference