An Expert System for Diabetic Microvascular Complication Diagnosis

Truc Ly Thi Nguyen¹, Nhon Van Do²

¹ Department of Computer Science, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam
² Department of Computer Science, University of Information Technology, Ho Chi Minh City, Vietnam

Abstract
This work presents the design of an expert system that aims to provide the doctors with limited practical experience and the diabetic patients more background for suitable diagnosis of diabetic microvascular complications. In modern society, more and more people get diabetes. This disease affects almost every organ in the body like heart, eyes, kidney, skin, nerves, blood vasculum and foot etc. They leave the very serious consequences for the patient. Early diagnosis of complications of diabetes can get tremendous value to patients and society. Hence an expert system for diabetes microvascular complication diagnosis can be used to assist for the doctors with limited clinical experience in situations without consultation directly to a specialist. In constructing an expert system, the knowledge representation plays a very important key. Therefore in this paper we also propose a new model of knowledge representation used to represent the complicated knowledge in medical domain and construct the Expert System for Diabetic Microvascular Complication Diagnosis (ESDMCD).

Keywords: Knowledge base system, Knowledge representation, intelligent system, Expert system, Medical diagnosis.

1. Introduction

According to specialist Nguyen Thy Khue - Dean of the Endocrinology Department of Medical University, said in the document [8], the diabetes is the most common cause of end-stage renal failure, this complication is considered the most dangerous and costly and is also one of the causes of death. Retinopathy is dangerous especially the majority of patients with diabetes do not notice any symptoms until eye suddenly lost vision. Then, despite aggressive treatment is costly and the ability to preserve your eyesight is very limited, the majority of patients will be blind permanently. Neuropathy is the cause of the removed extremities. The first histological lesions occur early, but the clinical manifestations appeared only about 10 to 15 years so that early diagnosis of diabetic complications can cause tremendous value to patients and society. Computer-based methods are increasingly used to improve the quality of medical services. Artificial Intelligent focusing on creating systems that considered intelligent like humans. This type of system can minimized the dependence on the human expert. The proposed system for dealing with the problem of Diabetic Microvascular Complication Diagnosis is expert system. ESDMCD may be used by the limited clinical experience doctors to assist in decision-making in situations without consultation directly to a specialist.

In this work, methods and techniques related to the construction expert system will be presented. In the construction process an expert system, knowledge representation plays an important role. Therefore in this paper, we also will present a new model of knowledge representation suitable for representing knowledge in diabetic microvascular complication diagnosis domain.

The rest of this paper is organized as follows: Expert System for Diabetic Microvascular Complication Diagnosis (ESDMCD) is introduced in the section 2. The section 3 describes our design method. Modeling knowledge base is mentioned in the section 4, in this part we also proposes the KBCO-ADAPT model and presents the knowledge representation in diabetes microvascular complication diagnosis domain by using this model. The section 5 represents diagnostic problem and deductive method. The section 6 shows the testing results. And finally we give some conclusions and perspectives in the section 7.
2. Expert System for Diabetic Microvascular Complication Diagnosis (ESDMCD)

2.1 Expert System

In the book [9], the author represented that the expert system represents a branch of artificial intelligence, aiming to take the experience of human specialists and to transfer to a computer system. Specialty knowledge is stored in the computer. They are reasoned and the system derives specific conclusions by an execution system (inference engine). The expert system represents expertise knowledge as data or rules within the computer. It derives conclusions by running the knowledge base through an inference engine; a software program interacts with the user to process the results from the rules and data in the knowledge base. These rules and data can be called upon when needed to solve problems.

Diabetes is a very common serious disease in the modern world. It affects many organs in the body like heart, eyes, kidney, skin, nerves, blood vessels, foot etc. Complications of diabetes can cause serious consequences affecting the quality of life. Though these complications can not possible to cure completely, it can be well managed.

Early diabetic complication detection plays a crucial role because it helps diabetic patient improve the quality of life and reduce the serious consequences. Thus a diagnosis support expert system helping detect early the diabetic complications have practical significant.

2.2 ESDMCD

After collecting in-depth diabetic complication knowledge from endocrinologist at Go Vap Hospital, we represented this knowledge by using KBCO-ADAPT model to design the ESDMCD. This system uses deductive rules and forward chaining inference technique.

The purpose of the ESDMCD is helping and supporting user’s reasoning but can not replace human experts. This system offers to the inexperienced user a diagnostic solution when human experts are not available.

The ESDMCD provides a user interface to interact with user, menu driven environment. In each diagnostic case, the system will ask a bunch of questions about the symptoms to user and user should give yes or no answer. According to these answers, the system will give to user the list of para clinical tests which system requires patient performing. After the para clinical test results are recorded into the program, ESDMCD will carry out a diagnosis and provide the treatment method for each type of complication.

Fig. 1 The diagnostic screen of ESDMCD.

This system has the structure of an expert system including six categories:

- The Knowledge Base contains the knowledge for solving some problems in diabetic complication diagnosis domain.
- The Inference engine will use the knowledge stored in knowledge base to solve problems. It must identify problem and use suitable deductive strategies to find out right rules and facts for solving the problem.
- The Working memory stores the facts and rules in the process of searching and deduction.
- The Explanation component supports to explain the phases, concepts in the process of solving the problem.
- The knowledge manager aims to support updating knowledge into knowledge base. It also supports to search the knowledge and test consistence of knowledge.
- The User Interface component of the system is required to have a specification language for communication between the system and users.
3. Design Method

In this section, we will present a process to construct ESDMCD based on the knowledge base and KBCO-ADAPT model. Besides, designed techniques in each phase will be presented. This system is analyzed and designed with the following steps.

Step 1:
- Determine the domain knowledge.
- Determine the scope of the problem.
- Collecting real knowledge from expert in the diabetic domain.

Step 2:
- Designing the knowledge base.
- Representing knowledge in the step 1 and modeling knowledge based on the KBCO-ADAPT model.
- Designing algorithms.
- Establishing knowledge bases for the system based on KBCO-ADAPT model. Knowledge base can be organized to store by structured text files. They include the files below.
  - File ATT.xml stores name of concepts.
  - File RELATIONS.xml stores the specification of relations (The R component of KBCO-ADAPT model).
  - File CONCEPT.xml stores the concept of objects (The C component of KBCO-ADAPT model).
  - File RULES.xml stores deductive rules.
  - File PATIENT.xml stores certain objects.

Step 3
- Research on strategies for deduction to solve problems on computer. The most difficult thing is modeling for experience, sensible reaction and intuitional human to find the heuristics rules, which were able to imitate the human thinking for solving problems
  - Design the inference engine.
  - Collecting issues, performance problems.
  - Choosing inference engine and the design of deductive algorithms for solving problems.

Step 4:
- Designing the user interface of the system.
- Programming the software by using C# programming language and XML.

Step 5: Install.
Step 6: Develop a sample to test.
Step 7: In this stage we do testing, maintaining and developing the application.

4. Modeling Knowledge Base

Through research we recognized that the model for knowledge bases of computational objects (KBCO) in [5] are suitable cause this model is established from Object-Oriented approach to represent knowledge therefore it is useful to represent the medical knowledge. In medical domain, each disease has its own symptoms and para clinical tests that can be described fully and explicitly by KBCO model. In addition this model also use Rules component to represent the deductive rules similar the rules are used to diagnose by doctor.

However some components in this model should be changed to represent knowledge in medical domain more comprehensive and more explicit. In this paper, the new knowledge representation model based on KBCO model will be presented. This new model is also used for designing and constructing the ESDMCD.

The clinical process of diagnosing physicians showed that almost doctors rely on the clinical symptoms and the results of Para clinical tests to diagnose. In addition, doctors use the specific symptoms for differential diagnosis. Beside the classic knowledge in the medical literature, the doctors also use a lot of practical experience to definitive diagnosis. Therefore the knowledge in this domain is really complex.

We studied many knowledge representation models to represent this kind of knowledge. Finally, we recognized that the model for knowledge bases of computational objects (KBCO) mentioned in [5] with some changes in components is appropriate to represent the knowledge in diabetic complication diagnosis domain. Components such as C, R and Rules of KBCO model can be used to describe knowledge in the medical diagnosis because in this area there are many objects including many attributes.
such as some complications of diabetes. They have also relationships together. Besides, many deductive rules are used to diagnose by doctors.

However, H, Ops components are not necessary because the knowledge in the medical diagnosis domain is not required to perform calculations and there is not a special relationship with H component.

Finally, if only using the C, R and Rules components can not describe the knowledge completely so that Attr and Patient components are added. Attr component is a set of base attributes. It is used to describe all symptoms of diabetic complications. Patient component includes a set of attributes and a set of concepts of objects. This component is used to describe the diagnosed patient. The adapted model based on KBCO is called KBCO-ADAPT.

4.1 Model for Knowledge Bases of Computation Objects (KBCO)

The model for knowledge bases of computational objects (KBCO) mentioned in [5] consists of six components: (C, H, R, Ops, Funcs, Rules)

These components are defined as follows
- C component is a set of concepts of computational objects. Each concept in C is a class of Com-objects.
- H component is a set of hierarchy relation on the concepts.
- R component is a set of relations on the concepts.
- Ops component is a set of operators.
- Funcs component is a set of functions.
- Rules component is a set of rules.

There are relations represent specializations between concepts in the set C; H represents these special relations on C. This relation is an ordered relation on the set C, and H can be considered as the Hasse diagram for that relation.

The set R is a set of other relations on C, and in this case a relation r is a binary relation it may have properties such as reflexivity, symmetry, etc.

The set Ops consists of operators on C. This component represents a part of knowledge about operations on the objects.

The set Funcs consists of functions on Com-Objects. Knowledge about functions is also a popular kind of knowledge in almost knowledge domains in practice, especially fields of natural sciences such as fields of mathematics, fields of physics.

The set Rules represents for deductive rules. The set of rules is a certain part of knowledge bases. The rules represent for statements, theorems, principles, formulas, and so forth. Almost rules can be written as the form “if <facts> then <facts>”. In the structure of a deductive rule, <facts> is a set of facts with certain classification. Therefore, deductive rule is used in the KBCO model. Facts must be classified so that the component Rules can be specified and processed in the inference engine of knowledge base system or intelligent systems.

4.2 The KBCO-ADAPT Model

KBCO-ADAPT model shows that the medical diagnostic knowledge can be represented completely. The KBCO-ADAPT model consists of five components: (Attr, C, R, Rules, Patient)

The meaning of the components is as follows:
- Attr component is a set of base attributes.
- C component is a set of concepts of objects. Each concept in C is a class of objects.
- R component is a set of relations on the concepts.
- Rules component is a set of deductive rules on facts.
- Patient component is a set of attributes or objects.

In this part, the knowledge about Diabetic Microvascular Complication are represented by using KBCO-ADAPT Model

- Attr component:

Attr component is the list or a set of base attributes: Attr = \{x_1, x_2, \ldots, x_n\} in which every attribute has a value in the fixed value region as float, integer, Boolean, string.

This component is used to describe all symptoms of diabetic complications and the symptoms used to distinctive diagnosis.

Example:

\[
\text{Attr}= \{ \\
\text{BloodGlucoseTest} : \text{float}; \\
\text{DiabetesType} : \text{string}; \\
\text{HbA1c} : \text{float}; \\
\text{BlurryVision} : \text{Boolean}; \\
\text{Pain} : \text{Boolean}; \\
\text{Numbness} : \text{Boolean}; \\
\text{Soreness} : \text{Boolean}; \\
\text{HighbloodpressureHis} : \text{Boolean}; \\
\text{Numb} : \text{boolean}; \\
\text{Retinopathy} : \text{string}; \\
\text{Nephropathy} : \text{string}; \\
\text{Neuropathy} : \text{string}; \\
\text{DiastolicBlood Pressure} : \text{integer};
\]
C component is a set of concepts of objects

The class of “level 2” objects is a class of objects having structure including a set of attributes $\text{Attr}(O) = \{x_1, x_2, ..., x_k\} \in \text{Attr}$ and $A_i(O)$ has data type is the class of “level 1” objects. In ESDMCD, the class of “level 2” objects is used to describe microvascular complication of diabetes.

Example:
The MicrovascularComplication_Diabete object has a set of “level 1” objects as follows:

$\text{Attr(MicrovascularComplication_Diabete)} = \{\text{HighbloodpressureHis} : \text{Boolean};\ $
$\text{DiabetesType} : \text{string};\$
$\text{Diabetic_Retinopathy} ; $
$\text{Diabetic_Nephropathy} ;$
$\text{Diabetic_Neupropathy};$
$\text{NeupropathyComplication_Diabete}\}$

The class of “level 1” objects

The class of “level 1” objects is a class of objects having structure including a set of attributes $\text{Attr}(O) = \{x_1, x_2, ..., x_k\} \in \text{Attr}$ and $A_i(O)$ has structure including a set of attributes whose data type is the basis concept. These objects can be used to define the “level 2” objects. In ESDMCD, the class of “level 1” objects is used to describe a type of diabetic complication such as NeuropathyComplication_Diabete, NephropathyComplication_Diabete, RetinopathyComplication_Diabete.

Example:
The RetinopathyComplication_Diabete object has structure including a set of attributes and basic concepts.

$\text{Attr (RetinopathyComplication_Diabete)} =$

$\{\text{EyesInjury} : \text{Boolean};\$
$\text{Keratitis} : \text{Boolean};\$
$\text{BlindnessHistory} : \text{Boolean};\$
$\text{Retino_FunctionalSymptoms} :$ 
$\text{Retinopathy_FunctionalSymptomsList};$
$\text{Retino_PhysiscalSymptoms} :$ 
$\text{Retinopathy_PhysiscalSymptomsList};$

- **R is a set of relations on C**

Each relationship is determined by $[\text{name of relation}, < \text{object1} >, < \text{object2} >]$ Relation “belong to” of an object and an object 

Example:

Obj1: Retinopathy_PhysiscalSymptomsList
Obj2: RetinopathyComplication_Diabete
[“belong to”, Obj1, Obj2]

- **Rules**

Rules relating to the objects are a set of deductive rules on the facts. Each rule described in this rules represents the inference rules in order to find out the new fact from the previous known facts. In general, each rule consisting of two parts: the hypothesis and conclusion. Almost rules can be written as the form “if $<\text{facts}>$ then $<\text{facts}>$”. In the structure of a deductive rule, $<\text{facts}>$ is a set of facts with certain classification. Facts must be classified so that the Rules component can be specified and processed in the inference engine of intelligent systems. Each rule is described as $\{\text{fact1}, \text{fact2}, \text{fact3}, \text{factn}\} \Rightarrow \{\text{fact1}, \text{fact2}, \text{fact3}, ..., \text{factm}\}$. There are many rules in the system. We represented these rules in the Rules.xml file. This is the structure of a rule used to diagnose in the ESDCMD.

$<\text{begin-rule}>$

$<\text{hypothesis_part}>$
$<\text{goal_part}>$
$<\text{compare_value}>$

$<\text{begin-rule}>$
Example:
<pre><code>&lt;begin-rule&gt;
  &lt;hypothesis_part&gt;
    DiabetesType : type1;
    BlurryVision : true;
    Diabetes Duration : &gt;-5 ;
    Ophthalmoscopy_RetinalHemorrhage : true;
    Ophthalmoscopy_RetinalVascularProliferative : true;
    Ophthalmoscopy_MacularEdema : true;
  &lt;/hypothesis_part&gt;
  &lt;goal_part&gt;
    Retinopathy: Diabetic Retinopathy Complication
  &lt;/goal_part&gt;
&lt;compare_value&gt;
    Diabetes Duration &gt;-5
&lt;/compare_value&gt;
&lt;begin-rule&gt;
</code></pre>

To apply the rules described in the knowledge base we have the following fact types.

**Fact about an object type**

This kind of fact informs about the object type.

We show this fact type by structured list:

```
[<object>, <object type>]
```

Example:

```
[Obj, "RetinopathyComplication_Diabetes"]
```

In this system we have some objects such as:

- obj1, MicrovascularComplication_Diabetes,
- obj2, RetinopathyComplication_Diabetes,
- obj3, Retinopathy_PhysicalSymptomsList,
- obj4, Retinopathy_FunctionalSymptomsList,
- obj5, Retinopathy_ClinicalTestList,
- obj6, NephropathyComplication_Diabetes,
- obj7, Nephropathy_FunctionalSymptomsList,
- obj8, Nephropathy_PhysicalSymptomsList,
- obj9, Nephropathy_ClinicalTestList,
- obj10, NeuropathyComplication_Diabetes,
- obj11, Neuropathy_FunctionalSymptomsList

**The fact for the determination of an object**

This fact is used to identify an object (attributes considered as known) or an attribute of an object by a value.

```
[<object> | <object>.< attribute value >]
```

Example:

```
Obj: Retinopathy_PhysicalSymptomsList
```

The “Ophthalmoscopy_RetinalVascularProliferative” attribute belongs to Retinopathy_PhysicalSymptomsList object. The attribute has value “true”.

**The fact about a relationship**

This fact type performs a relationship between the objects or attributes of objects. This fact type can be represented by the structured t:

```
[<relationship fact >, < Obj1>, < Obj2>]
```

Example:

```
Obj1: RetinopathyComplication_Diabetes.
Obj2: Retinopathy_FunctionalSymptomsList.
```

- **Patient component**

Patient component is a class of objects has structure including a set of its own attributes, a set of attributes

```
Attr(O) = \{x_1, x_2, ..., x_k\} \in Attr and a set of objects A1(O) \in C. This component is used to describe the diagnosed patient.
```

Patient =

```
\{PatientName : string;
  YOB : integer;
  Address : string;
  PrehistoricDiabetes : Boolean;
  Diabetes Duration : float;
  BloodGlucoseTest : float;
  DiabetesType : string;
  HbA1c : float;
  BlurryVision : Boolean;
  Numbness : Boolean;
  Soreness : Boolean;
  Diabetic_Retinopathy:
    RetinopathyComplication_Diabetes;
  Diabetic_Nephropathy:
    NephropathyComplication_Diabetes;
  Diabetic Neuropathy:
    NeuropathyComplication_Diabetes;
  obj1, MicrovascularComplication_Diabetes,
  obj2, RetinopathyComplication_Diabetes,
  obj3, Retinopathy_PhysicalSymptomsList,
  obj4, Retinopathy_FunctionalSymptomsList,
  obj5, Retinopathy_ClinicalTestList,
  obj6, NephropathyComplication_Diabetes,
  obj7, Nephropathy_FunctionalSymptomsList,
  obj8, Nephropathy_PhysicalSymptomsList,
  obj9, Nephropathy_ClinicalTestList,
  obj10, NeuropathyComplication_Diabetes,
  obj11, Neuropathy_FunctionalSymptomsList
...}
```
5. Diagnostic Problem and Deductive Method

5.1 Diagnostic Problem

General components, which are present in a diagnostic procedure in most of the various available methods include:

Complementing the already given information with further data gathering, which may include questions of the medical history (potentially from other people close to the patient as well), physical examination and various diagnostic tests.

A diagnostic test is any kind of medical test performed to aid in the diagnosis or detection of disease. Diagnostic tests can also be used to provide prognostic information on people with established disease.

Processing of the answers, findings or other results. Consultations with other providers and specialists in the field may be sought.

Based on the knowledge base of the microvascular complication of Diabetes is modeled COKB-ADAPT model above.

Input = The set of symptoms are provided by patient, symptoms are detected by the doctor on physical examination and the results of diagnostic tests.

Output = Diagnostic conclusion with

- Diabetic Neuropathy Complication,
- Diabetic Retinopathy Complication,
- Diabetic NephropathyComplication or not?

Example :

Input:
The set of symptoms are provided by patient
- DiabetesType: type1;
- Diabetes Duration: 7;
The set of symptoms are detected by the doctor on physical examination
- Ophthalmoscopy_RetinalVascularProliferative: true;
- Results of laboratory test
- MicroalbuminNieu: 45

Output: Diagnosis
- Nephropathy : Diabetic Nephropathy Complication;

5.2 Clinical Diagnostic Process

In beginning, patient reports the primary symptoms to a physician, and then he (or she) can detect more symptoms on physical examination. If the probability of certain diagnoses is halfway between likely or unlikely, the physician can request certain tests on the patient such as blood test, Microalbumin…ect.

Like the history and physical exam, diagnostic testing provides new information that further narrows the differential diagnosis.

Finally based on the clinical symptoms and the results of para clinical test, the physician summarizes the most likely diagnosis.

5.3 Deductive Method

Inference engine used in the expert system for diabetic complication diagnostic is forward chaining method. This deductive method is often used in expert system. The solution process is started with the collection of information and this information is deduced to draw conclusions.

The inference strategy begins with collective known facts, new facts learned by using the law assumes that the fit known facts, and continue this process until you see the status of the target, or for until no rule that matched the event is known or inferred events.
6. The Testing Result

Data collected from 4 District General Hospital from July 16, 2012 to November 12, 2012 is used to evaluate the expert system for diabetes microvascular complication diagnosis (ESDMCD). This is the actual data for research on the diabetic microvascular complications of a doctor working at District 4 General Hospital.

The ESDMCD be assessed by comparing the diagnostic results of expert physicians with the expert system in 106 patients. Comparison of the diagnostic expert system ESCMD and doctor shows cases of 106 cases with 2 different diagnostic results. 104 cases resulted in equivalent proportion of 98.11%.

A proper diagnose case of system
PatientName: Nguyen thi Nguyet Anh
YOB :1963
Address: 12 Doan Van Bo Street, 4th District
PrehistoricDiabetes : true;
Diabetes Duration : 6;
HbA1c : 12.5;
BlurryVision : True;
Diabetic_Retinopathy : true;
Diabetic_Nephropathy : true;
Diabetic_Neurophy : true;

BloodGlucoseTest : 12.5;
DiabetesType : type1;
BlurryVision : true;
Soreness : true;
Diabetic_Retinopathy : RetinopathyComplication_Diabete;
Diabetic_Nephropathy : NephropathyComplication_Diabete;
Diabetic_Neurophy : NeuropathyComplication_Diabete;
[obj1, MicrovascularComplication_Diabete],
[obj2, RetinopathyComplication_Diabete],
[object3, Retinopathy_PhysicalSymptomsList],
[object4, Retinopathy_FunctionalSymptomsList],
[object5, Retinopathy_ClinicalTestList],
Obj1.obj4.Ophthalmoscopy_RetinalHemorrhage =true;
Obj1.obj4.Ophthalmoscopy_RetinalVascularProliferative :
= true;
Obj1.obj4.Ophthalmoscopy_MacularEdema = true;

Conclusion:
Diagnosis of ESDMCD: Diabetic Retinopathy Complication.
Diagnosis of physician: Diabetic Retinopathy Complication

7. Conclusions and Perspective

The expert system for diabetes microvascular complication diagnosis (ESDMCD) is an application can help the doctors with limited experience in providing decision support system, interactive training tool and expert knowledge. This work proposed the KBCO-ADAPT model to represent the knowledge in this domain.

The ESDMCD imitated diagnostic process of doctor in clinical practice. By using exact rules, the system gave the relatively accurate diagnostic results similar doctor's diagnosis with proportion 98.11%. After testing the system, the specialist gave us the positive feedback. Therefore this system is appreciated very high in the practical applications by endocrine specialists.

In the future, we will continue to research the COKB-ADAPT model to complete the program for solving problems.

The KBCO-ADAPT also proposed a natural way for the knowledge representation in Diabetic Microvascular Complication Diagnosis domain. This model is also a tool to design practical knowledge bases and algorithm to solve medical diagnosis problems.
**Acknowledgements**

We gratefully acknowledge Endocrine Specialist Minh Chau Tran Hoang for providing her expertise knowledge and Endocrine Specialist Chuong Phuong Tran Thi for using her data to evaluate system in this work.

**References**


