

A STUDY OF RULE BASE INFERENCE ENGINE FOR CLINICAL  
DIAGNOSIS EXPERT SYSTEM ON CLOUD COMPUTING

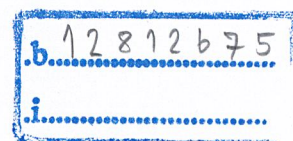


WORANAN PRASANSIWAMAI  
SUPAVIJ KHIAOKAJEE

ADVISOR

DR. MANOP PHANKOKKRUAD

เลขหมู่.....  
เลขทะเบียน 078012  
วัน,เดือน,ปี 24 พ.ย. 2559



A PROJECT SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENT FOR THE DEGREE OF  
BACHELOR OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY  
FACULTY OF INFORMATION TECNOLOGY  
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

**COPYRIGHT 2015**

**FACULTY OF INFORMATION TECHNOLOGY**

**KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

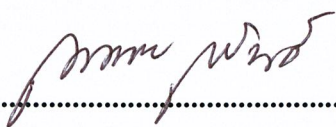
**PROJECT CERTIFICATE**

**FACULTY OF INFORMATION TECHNOLOGY, KING MONGKUT'S  
INSTITUTE OF TECHNOLOGY LADKRABANG**

**TITLE: A STUDY OF RULE BASE INFERENCE ENGINE FOR  
CLINICAL DIAGNOSIS EXPERT SYSTEM ON CLOUD  
COMPUTING**

**BY**

- 1. MS. WORANAN PRASANSIWAMAI    STUDENT ID 54070083**
- 2. MR. SUPAVIJ    KHIAOKAJEE    STUDENT ID 54070091**

  
..... **ADVISOR**  
**(DR. MANOP PHANKOKKRUD)**

หัวข้อโครงการ	การศึกษากลไกการอนุมานแบบกฎสำหรับระบบผู้เชี่ยวชาญด้านการวินิจฉัยโรคด้วยการประมวลผลแบบกลุ่มเมฆ
นักศึกษา	นางสาว วรณัน ประสารศิริมัย รหัสประจำตัว 54070083 นาย สุภวิษญ์ เขียวขจี รหัสประจำตัว 54070091
ปริญญา	วิทยาศาสตร์บัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
ปีการศึกษา	2557
อาจารย์ที่ปรึกษา	ดร. มานพ พันธุ์โคกกรวด

## บทคัดย่อ

กระบวนการวินิจฉัยโรคถือเป็นสิ่งสำคัญสำหรับด้านการแพทย์ เพื่อที่จะทำการรักษาผู้ป่วยให้ได้อย่างถูกต้องทันท่วงที และมีประสิทธิภาพ อย่างไรก็ตามเป็นที่ทราบกันดีว่าข้อมูลองค์ความรู้ต่าง ๆ เกี่ยวกับด้านการแพทย์เป็นความรู้ที่มีความซับซ้อน มีเนื้อหาปริมาณมาก รวมทั้งต้องใช้ทักษะประสบการณ์ความเชี่ยวชาญเฉพาะด้านอย่างสูง เพื่อที่จะทำความเข้าใจ และทำการวินิจฉัยผู้ป่วยได้ ซึ่งในบางครั้งกระบวนการคิดวิเคราะห์แก้ไขปัญหาคด้วยมนุษย์นั้น อาจเกิดความผิดพลาดเนื่องมาจากความเหนื่อยล้า หรือกับปัญหาที่มีความซับซ้อนมาก ดังนั้นเพื่อแก้ไขปัญหาดังกล่าวทางผู้จัดทำจึงนำเทคโนโลยีคอมพิวเตอร์เข้ามาช่วยในการวินิจฉัยโรค กล่าวคือจะจัดสร้างระบบผู้เชี่ยวชาญ สำหรับการวินิจฉัยโรคทางการแพทย์ โดยจะทำการศึกษาสององค์ประกอบหลัก ได้แก่ องค์ความรู้และกลไกการอนุมาน ในส่วนขององค์ความรู้จะเป็นการเก็บรวบรวมข้อมูลต่าง ๆ ด้านการแพทย์และการศึกษารูปแบบโครงสร้างของข้อมูลแบบลำดับชั้น เพื่อที่จะใช้อธิบายข้อมูลดังกล่าว สำหรับวิธีการในการหาคำตอบจะใช้วิธีการอนุมานแบบกฎ โดยจะทำการสร้างกฎจากองค์ความรู้ที่เก็บรวบรวมได้ และในขั้นตอนสุดท้ายจะทำการศึกษาถึงเทคโนโลยีการประมวลผลแบบกลุ่มเมฆ ซึ่งเป็นเทคโนโลยีที่มีประสิทธิภาพและมีความยืดหยุ่นสูง รวมทั้งกำลังจะเป็นที่นิยมในอนาคต ซึ่งในโครงการนี้จะถูกนำมาใช้สำหรับเก็บข้อมูลต่าง ๆ เนื่องจากการเก็บข้อมูลแบบออนไลน์ ทำให้การพัฒนาโปรแกรมและการเข้าถึงข้อมูลมีความยืดหยุ่นสูงขึ้น ทำให้ได้มาซึ่งระบบผู้เชี่ยวชาญสำหรับการวินิจฉัยโรคที่สามารถทำการวินิจฉัยโรคได้จากการวิเคราะห์ข้อมูลที่ถูกป้อนเข้าสู่ระบบ ซึ่งผลลัพธ์ที่ได้นั้นสอดคล้องกันกับข้อมูลกฎต่างๆในฐานองค์ความรู้ที่มีอยู่

<b>Project Title</b>	A Study of Rule Base Inference Engine for Clinical Diagnosis Expert System on Cloud Computing		
<b>Student</b>	Ms. Woranan	Prasansiwamai	Student ID 54070083
	Mr. Supavij	Khiaokajee	Student ID 54070091
<b>Degree</b>	Bachelor of Science		
<b>Program</b>	Information Technology		
<b>Academic Year</b>	2014		
<b>Advisor</b>	Dr. Manop	Phankokkrud	

## ABSTRACT

The clinical diagnosis is important for the medical in order to treat the patients correctly and effectively. However, it is well known that the medical knowledge is too complex and has plenty of content, including the need of high skills, experiences and specialization in order to understand and diagnose the patients. Sometimes the critical thinking and problem solving by humans that may be a mistake due to fatigue or the problems are very complicated. Therefore, in order to solve the problems, the authors bring a computer technology to help in the diagnosis. We developed an expert system for medical diagnosis. From the study including two main components of knowledge base and an inference engine, the knowledge base collect the information about medical and other and study the hierarchical data structure in order to describe such information. As for how to find a method of inference rules, it creates the rules of collected knowledge. Finally, we studied the cloud computing technology which is powerful and more flexible. In addition, the trend is going to be popular in the future. In this project, the cloud computing technology is being used to store the data because the online storage makes application development and data access to be more flexible. Finally, we developed the clinical expert system. It can clinical diagnose and give the conclusion that consistency of the rules based.

## ACKNOWLEDGEMENT

First and foremost, we would like to express our enormous gratitude for our advisor of this project, Dr. Manop Phankokkrud for the continuous support throughout this project. This project would not have been successful without his invaluable guidance and assistance.

Besides, we would like to thank all of professors at Information Technology faculty for giving necessary knowledge for us to apply in our project and our profession in the future.

Further, we would like to thank our parents and families who always support and understand us, gives us strength during the entire life.

Finally, thank you for all friends that give us suggestions and always support us in everything to complete this project.

Woranan Prasansiwamai

Supavij Khiakajee

# CONTENTS

	Page
Abstract (Thai) .....	I
Abstract .....	II
Acknowledgement .....	III
Contents .....	IV - V
List of figures .....	VI
List of tables .....	VII

## Chapter

1. Introduction .....	1
1.1 Background .....	1
1.2 Objective .....	1
1.3 Methodology .....	2
1.4 Scope .....	3
1.5 Benefits .....	3
2. Background .....	4
2.1 Expert System .....	4
2.2 Data Structure .....	8
2.3 Cloud Computing .....	11
2.4 Windows Azure .....	13
2.5 Windows 8 application .....	16
2.6 Review of Literature .....	18
3. Analysis and Design .....	22
3.1 Clinical Knowledge Base .....	22
3.2 Rule Base Inference Engine .....	28

## CONTENTS (CONTINUED)

	Page
3.3 Cloud Database Structure.....	35
3.4 System Architecture.....	36
3.5 Implement Rule Base Inference Engine to Cloud Service.....	37
3.6 Use Case Diagram.....	41
3.7 Activity Diagram.....	41
3.8 Class Diagram.....	42
4. Results and Discussion .....	44
4.1 User Interface.....	44
4.2 System Evaluation .....	47
5. Conclusion .....	48
BIBLIOGRAPHY .....	50
BIOGRAPHY.....	54

# LIST OF FIGURES

	Page
Figure	
2.1 Hierarchical database model .....	10
2.2 Part of tree menu lists .....	11
2.3 Cloud computing .....	12
2.4 Cloud computing services .....	13
2.5 Windows azure services .....	15
2.6 Windows azure table structure .....	16
2.7 Flat design .....	18
3.1 The tree structure of the cough diseases .....	25
3.2 The tree structure of headache .....	28
3.3 Structure of azure table storage .....	35
3.4 System architecture diagram.....	36
3.5 Windows azure portal.....	37
3.6 Azure storage explorer .....	38
3.7 Initialize azure storage account.....	38
3.8 Upload csv file to azure table .....	39
3.9 List of azure services library .....	39
3.10 Use case diagram .....	41
3.11 Activity diagram .....	42
3.12 Class diagram .....	43
4.1 Main menu .....	44
4.2 Select a part of body that have an illness.....	45
4.3 Answer the questions that asked about symptoms.....	45
4.4 Choose the symptoms that occurs .....	46
4.5 The result of diagnosis .....	46

# LIST OF TABLES

	Page
Table	
3.1 The diseases caused of coughing .....	23
3.2 Indicative that the diseases caused by coughing .....	24
3.3 The diseases caused of headache .....	26
3.4 Indicative that the diseases caused by headache .....	27
3.5 Classification of cough disease .....	29
3.6 Symptoms of cough disease .....	30
3.7 Rule of coughing disease by symptoms .....	32

# Chapter 1

## Introduction

### 1.1 Background

The clinical diagnosis is important for the medical in order to treat the patients correctly and effectively. However, it is well known that the medical knowledge is too complex and has plenty of content, including the need of highly skills, experiences and specialization in order to understand and diagnose the patients. Sometimes the critical thinking and problem solving by humans that may be a mistake due to fatigue or the problem are very complicated. Therefore, in order to solve the problems, the authors bring a computer technology to help in the diagnosis. So, we developed an expert system for medical diagnosis.

From the study, the tool supporting for clinical diagnosis is the knowledge base because it is central of knowledge about clinical information such as group of diseases, symptoms and treatment. That causes to ease searching and managing. It also can integrate with other processes such as data mining or knowledge discovering [1] for advance analyzing and given more accuracy of the results. For example, it is used to implement a decision support system for the diagnosis of Alzheimer disease [2]. However, only knowledge base is not enough for the diagnosis process. Although, the knowledge base can be referenced and help to store and search data, doctors still work together with the knowledge base for disease diagnosis. In some cases, the knowledge base contains too much data and that directly affects to analyze and not be able to be solved by humans. In addition, the computer and artificial intelligence technology have integrated and applied with a clinical knowledge base for developing the expert system. The expert system is a computer system with AI technology that can work as same as experts, in this case, it refers to the doctors. For example, the clinical diagnosis support system based on symptoms uses neuron network approach to support large scale of input [3]. As discussed above, the clinical information is large thus it is too difficult to be solved by humans. The diagnosis expert system by the production rule [4] uses the rules as an inference engine to find out the conclusion and designed knowledge base to store symptoms, precondition rule and conclusion. Moreover, in some cases, the rule base is combined with case based reasoning to provide more reliable conclusion [5]. Furthermore, some expert systems use the belief rule technique [6, 7]. It can handle various types of uncertainty found in the clinical domain knowledge. According to our study, the expert system is a tool that can work as same as a doctor and it can clinically diagnose and provide results by

itself. It is very useful for supporting the doctors. Thus, the authors proposed ideas for developing tools that can support and increase efficiency of clinical diagnosis.

Actually, the expert system consists of the knowledge base and inference engine. The first one, the knowledge base, is the database representing the information on specific domain knowledge, including facts, tactics and user experience. The second one, an inference engine, is an artificial intelligent mechanism for analyzing and processing data from the knowledge base. In order to design a clinical knowledge based, the hierarchical model is used. It is one of the data structure models that represent data as parent-child relationship. This type of relationship is clear classification groups of disease [8, 9]. This model can separate disease information into many sub-groups. Also, it can create the conditions for classifying types of disease. Thus, it is useful for implementing clinical knowledge base. Therefore, we will use rule base as an inference engine for the clinical diagnosis expert system because it can clearly identify diseases by the if-logic conditions, the result must be only true or false. Furthermore, the rule base is useful for information that has many attributes as clinical information that has many symptoms. Finally, we will implement the clinical knowledge base on cloud computing technology. The main reason is the cloud services provide facilities of working as a service. For example, the knowledge on cloud computing can provide as a service term [10] and can develop an expert system for classification of Ahadith [11]. Moreover, the cloud services provide user-friendly tools or portal for managing that has user-friendly [12, 13] that ease of applying.

## **1.2 Objectives**

1. To study a process of clinical diagnosis
2. To study and develop the inference engine for clinical diagnosis
3. To study cloud computing technology
4. To develop the clinical knowledge base on cloud computing services

## **1.3 Methodology**

### **1.3.1 Study and collect information**

1. Study and collect information and knowledge base about clinical diseases
2. Study the clinical diagnosis process
3. Study cloud computing technology and database system on cloud services
4. Study clinical diagnosis expert systems

### 1.3.2 Analyze and design system

1. Analyze the clinical disease information
2. Analyze structure of clinical knowledge base
3. Analyze environment and structure of cloud computing services
4. Design system architecture

### 1.3.3 Implementation

1. Develop a knowledge base system
2. Design user interface for the application
3. Develop the application and implement knowledge base to cloud computing

### 1.3.4 Test and adjust the defect of the system

1. Test the operation of the interface
2. Test the database connection
3. Test the work on cloud
4. Proof the accuracy of knowledge base

## 1.4 Scope

1. This project focuses on the structure of clinical knowledge base and how to implement on the cloud computing
2. This project focuses functions on cloud database and cloud storage using Windows Azure.
3. The clinical disease information used in this project related to the diagnosis of cough and headache.
4. This project develop the knowledge base using hierarchical data model

## 1.5 Benefits

1. To develop the knowledge base suitable for the diagnosis of cough and headache. Also, the system can provide the content matching with user experience and knowledge.
2. The knowledge base developed in this project will be the content base for clinical diagnosing the diseases.

## **Chapter 2**

### **Background**

#### **2.1 Expert System**

The expert system is the system that used a knowledge base to solve specific problems depending on user propose and domain knowledge. Also, it can refer to the application for managing and analyzing the information to find out the solution. The expert system is software tools for supporting other systems such as decision support system, management information system, customer relationship management etc. The expert system has two main components including inference engine and the knowledge base. The knowledge base is the database that represents information that is facts about the world. The inference engine represents logics and conditions to make a decision.

Otherwise, the expert is similar to knowledge base systems. The expert system refers to a computer system that can act as the expert human action. It can decide based on the knowledge and make solution for solving the complex problem. And the knowledge-based system refers to the architecture of the system. It is likes the word “database system”. It represents the information or data used in the system, almost these data are the fact or rule. The expert system uses the knowledge base for supporting solved complex problem, or need to find new information.

##### **2.1.1 Knowledge-based**

The knowledge base is a kind of the database that storing some knowledge. The knowledge means any information in some specific domains from the expert such as the experiences, the truth, tactics, know-how and how-to. The knowledge base is a very important factor for supporting decision making and problem solving. When people face to the problem, they must analyze and understand the problem, also identify the cause or source of the problem. Therefore, they must determine any possible solutions by using the knowledge as the conditions or decision variables. The conditions and decision variables are important key to determine the results. It always can be changed depending on any information that is the input or control of factor. By the way, the knowledge base was applied to computer technology called the knowledge based system. It is the computer system that provides knowledge for the users and refers to information resource pool. Moreover, the knowledge-based system can integrate with other systems such as management information systems (MIS) and decision support systems (DSS). This system is an application with an artificial intelligent technology that uses the knowledge-

based system to find out the specific answer or solution for problem-solving. Also, it is used for diagnosis and response of an emergency situation, including supporting for decision making [14, 15].

In the past, the knowledge meant the experience from the humans and it is any information learned by humans from try-error, study, hands-on and so on. However, this knowledge always lost when these human died. Nowadays, humans use the technologies to manage information and this becomes to management information systems as known as MIS. This system is designed for managing and utilizing the data to most values. Thus, the knowledge base is designed. It is different from the database. The database has following properties; flat data, multiple users, transactions and long-lived data. The data in database is usually represented in a tabular format with strings or numbers in each field, also has transaction processing from multiple users or systems. But the knowledge-based system is not similar the database. It represents the fact or rule, including best practice and specific solution. Thus, in the point of view of the user, the characteristic of the knowledge-based system is compacted for more understandable than the database structure. It stores both of structured and unstructured data used by computer systems. Not just tables with numbers and strings, also include pointers to other objects such as classes, subclasses, and instances.

In this project, we create the clinical knowledge base for implementing the clinical diagnosis expert system. Actually, the characteristic of clinical information is large and complicated. It consists of many data such as patient profiles, symptoms, diseases or groups of diseases, treatment and medicine. Obviously, it seems to be difficult to be managed and processed by humans, thus the knowledge based system was developed. In this case, we focus on the type of diseases and the symptoms caused by diseases. Because, we would like to develop the clinical diagnosis expert system, this program must work as same as doctors diagnosing patient's symptoms. Thus, the clinical knowledge base consists of the list of diseases and symptoms. However, the kinds of the structure model or pattern are important issues and must be considered. The knowledge base must be explained by some data models for ease implementing and understanding that we will discuss later in Data Structure section and chapter 3.

### **2.1.2 Inference engine**

The inference engine is one of the main components of the expert system or the knowledge-based system. It logically is a part of the artificial intelligence technology, or refers

the computing and processing unit in the CPU. It applies some algorithms and mathematical models to the knowledge base for analyzing data. It is used to discover the new information from existing knowledge and find out the solution or conclusion in term of rules and also supports decision making. The logic of an inference engine is represented by IF-THEN rule. It consists of the logical expression that refers to the condition and effect. The rule shows what happens after the condition is occurred or the expression is executed. The inference engines work on two approaches including forward chaining and backward chaining.

### **2.1.2.1 Forward Chaining**

Forward chaining is the method of reasoning to inference and find out the conclusion by starting with the known facts and then forwarding to achieve the result. It is usually implemented for the expert systems, business and production rule systems. The fact and available data are input data that use inference rules to extract more data until a goal is reached. This approach searches the inference rules until one expression which is true is found. When this rule is found, the engine can conclude. Consequent, the resulting in the addition of new information to its data.

For example, the goal is to conclude the type of an animal named Mike, given that he has 4 legs, has long neck and herbivore, and that the rule base contains the following four rules:

- |         |   |
|---------|---|
| Rule 1: | If X has 4 legs and long neck           |
|         | Then X is a giraffe                     |
| Rule 2: | If X is a giraffe                       |
|         | Then X is herbivore                     |
| Rule 3: | If X is has 4 legs and has no long neck |
|         | Then X is a dog                         |
| Rule 4: | If X is a dog                           |
|         | Then X is carnivore                     |

With forward reasoning, the inference engine can infer that Mike is herbivore in a series of steps:

1. The facts "Mike has 4 legs" and "Mike has long neck" are matched with expression of rule #1, thus the conclusion is Mike is a giraffe.
2. The facts "Mike is a giraffe" are matched with expression of rule #2, thus the conclusion is Mike is herbivore.

So, the forward chaining approach give the conclusion Mike is herbivore by proof rule #1 and #2 in the order.

### 2.1.2.2 Backward Chaining

Backward chaining is the method that opposite to forward chaining. It starts with goals or the assumption and works backward by ordering to proof any conditions that support to reach the goal. This approach uses the available data for proofing all of the inference rules that can achieve the goal. For example, the goal is decide that Mike is herbivore by given fact that he has 4 legs, has long neck and herbivore. The rule base contains the following four rules:

- Rule 1:            If X has 4 legs and long neck  
                         Then X is a giraffe
- Rule 2:            If X is a giraffe  
                         Then X is herbivore
- Rule 3:            If X is has 4 legs and has no long neck  
                         Then X is a dog
- Rule 4:            If X is a dog  
                         Then X is carnivore

With backward reasoning, the inference engine can infer that Mike is herbivore in a series of steps:

1. The rule matches the goal ("Mike is herbivore ") that rule #2 ("If Mike is a giraffe") can be proved. So, the rule #2 is proved.
2. The rule matches the goal ("Mike is a giraffe") that rule #1 ("If Mike has 4 legs and long neck ") can be proved. So, the rule #1 is proved.

So, the backward chaining approach proven the assumption Mike is herbivore because the rule #1 and #2 are true.

In the clinical diagnosis expert system, many types of the inference engine can be used for implementing the expert system. For example, the clinical diagnosis support system based on symptoms. It uses the neuron network approach to support large scale of input [3]. As above discussed, the clinical information is large thus it is too difficult to be solved by humans. The diagnosis expert system by the production rule [4] uses the rules as an inference engine to find out the conclusion and designs knowledge base to store symptoms, precondition rules and conclusions. Moreover, in some cases, the rule base combined with case based reasoning for

providing more reliable conclusion [5]. Furthermore, some expert systems uses the belief rule technique [6, 7]. It can handle various types of uncertainty found in the clinical domain knowledge.

### 2.1.3 Rule-based System

The rule-based system is one of the approaches from all of inference engines in the expert system to store the knowledge and use to find out a solution or result, predict and forecast. It is also widely used in the AI applications and researches. The rule based systems store the knowledge in term of rules. The rule is the information that presents the condition and result, or it means the fact-based on a domain of knowledge which is stored. Almost rules in the rule-based systems are presented in the “if...then...” pattern. For example, I create rules to check whether people are adult by their age. The condition is if their age is higher than 18, then the result is adult. It can be written in a rule term by *IF age>18 THEN is an adult*. The mechanism of rule-based process is to check the input data and match with condition from the rules stored in database. Then, it returns the result. The rule-based system can support the users in many processes such as helping for making decision, providing the choice, result and conclusions. The user just input data into the system, and then system will provide the result. However, the accuracy of the result depends on the quality of the knowledge and rules, including the amount of attributes and datasets. Thus, the knowledge and information must be verified before creating the rules. Also, these rules must be tested by testing data set for ensuring that the rules are correct.

In this case, we use the rule base to identify the type of diseases from any symptoms that the patients are suffered. Every rule presents symptoms as the cause then provides diseases as a result. This method is suitable for analyzing the database that has many attributes. It can check many conditions by values of variables to find the result. Moreover, some of the condition has many input variables. It is hard to check one by one attribute to find results, but the rule-based technique can check all of them by the what-if analysis approach. The rule base checks each variable by the IF logic comparing to the information on the knowledge base.

## 2.2 Data Structure

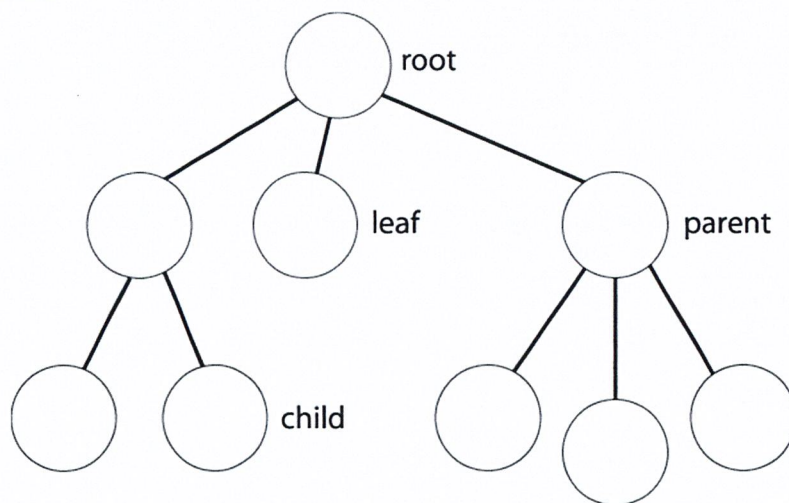
The data structure is referred to the structure which is used to explain how the data is stored in the database. It presents the relationship of each data, node or table. The data structure is a very important issue for the process design and the database development. It directly impacts to kinds of database. Different kinds of data structures effect to accessing time, querying process and

retrieving data. Thus, in order to select the data structure to design the database, it depends on users propose, objective function of the application and the nature or characteristic of the data. In this project, the author presents two types of data structure that are common and widely used in many systems which are the hierarchical model, as known as a tree structure, and the relational model.

### 2.2.1 Hierarchical Model

The hierarchical data model is one of database models that structure is like the tree. It consists of parent and child data segments, or called node. Each node represents the data, and if it has the child node, that node is called parent node. In the structure view, each node has a similar branching structure of the tree. One of the rules which is very important is that one parent can have one or more children, but children can have only one parent [16, 17].

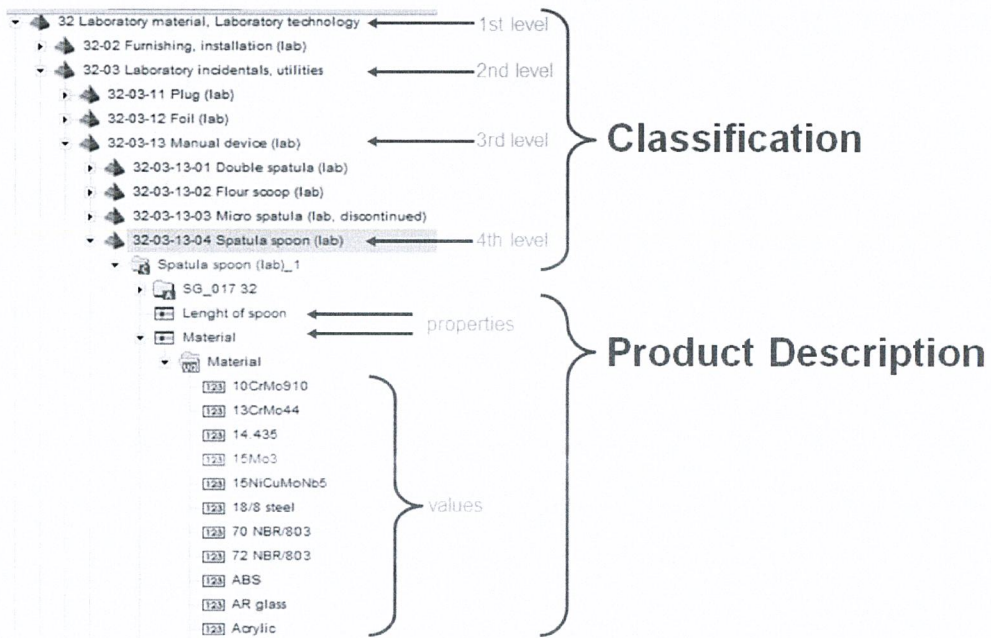
As presented in figure 2.1, the tree structure consists of a starting node called 'Root Node', meanwhile the root node can be 'Parent node' in the same time. The parent node is any node which has 'Child Node', and any node which has no child node is called 'Leaf Node'. It is almost implemented with the data set that can be clustered or classified into many sub-groups. The branch in tree refers to the conditions by several value of variable to divide parent node into many child nodes.



**Figure 2.1** Hierarchical database model

From the user point of view, the hierarchical data model is easy to understand because the tree diagram is less complicated than the ER diagram in relational databases. This model allows flexibility to add and remove the data. Moreover, this model is suitable for the data that has no many-to-many relationship, but has one-to-many relationship in term of parent-child or

hierarchy. Also, the data is clearly classified and can be broken down into many levels. For example, figure 2.2 is a user interface screen as the hierarchical structure. Many applications use tree structure to implement in order to access data by sequences.



**Figure 2.2** Part of tree menu lists

In this project, the authors use the hierarchical model for explaining the knowledge base as a decision tree form. As presented in figure 2.2, we create a hierarchical model for cough disease by separating data set into many sub data sets by some conditions. Starting by cough as the root node and the cough can be separated into two groups which are chronic cough and acute cough. Both of them can be broken down into two groups which are productive cough and non-productive cough. This tree then grows depending on amount of data until diagnosis disease is found as presented in leaf nodes. Each node has some condition to separate child nodes. These conditions are ruled simultaneously by top-down order. However, this model is rough for classifying a group of diseases on top level.

For the clinical information, it has many diseases, and each disease can break down into a specific type of disease by identifying from symptoms, infections, time and etc. Thus the classification of clinical information can improve efficiency by analyzing and diagnosing process. The tree structure helps to understand the perspective. Parent nodes in each level show the connection between child nodes to refer to parent-child relationship. It represents the relation between group and sub-group of diseases. Moreover, the hierarchical model totally helps to

classify many diseases and find out the result. It also works like the decision tree. The decision tree method can help to determine diagnosis of disease that delivers high accuracy of results [18].

## 2.3 Cloud Computing

Cloud Computing is the computing services over the internet rather than a product from cloud service provider, it provides many services such as software, hardware, information, network system and etc. to users by sharing computing resources [19]. Although these resources as required by users are same as the current system, it has higher performance than consumer systems. To use computing hardware as physical product is more expensive for some users and not compatible with non-expert users. In addition, it causes some limitation such as scalability, flexibility and availability. For example, the user needs to scale up or scale out the current system. It depends on space and location in the organization. Also, the implementation process is not easy. Meanwhile, the cloud computing technology is totally different. It has only portal for managing any services as self-services for the user. This technology is the computing service over the internet. It can serve and provide resources as on-demand user. As presented in figure 2.3, cloud computing provides any services such as application, platforms and infrastructure over the internet. The users can request and access any services via the internet from anywhere and any devices.

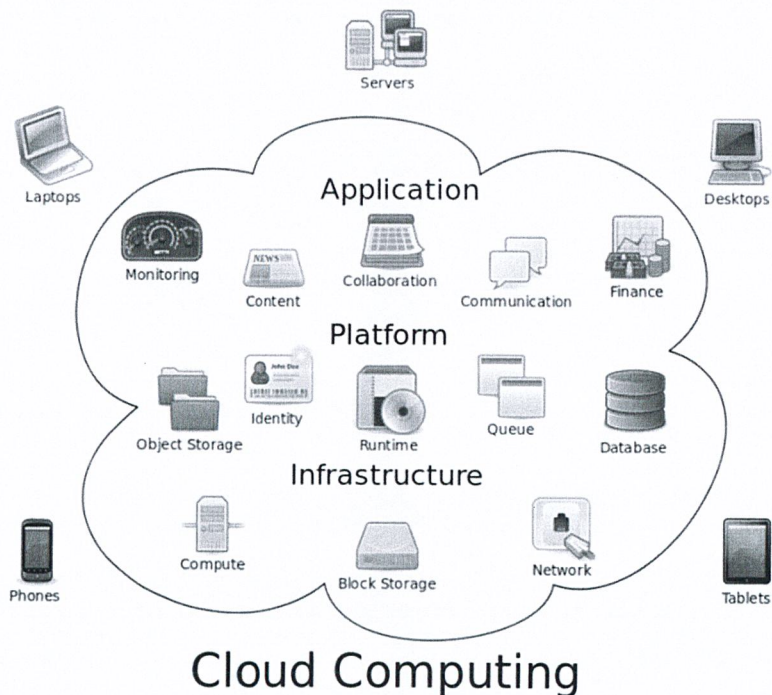
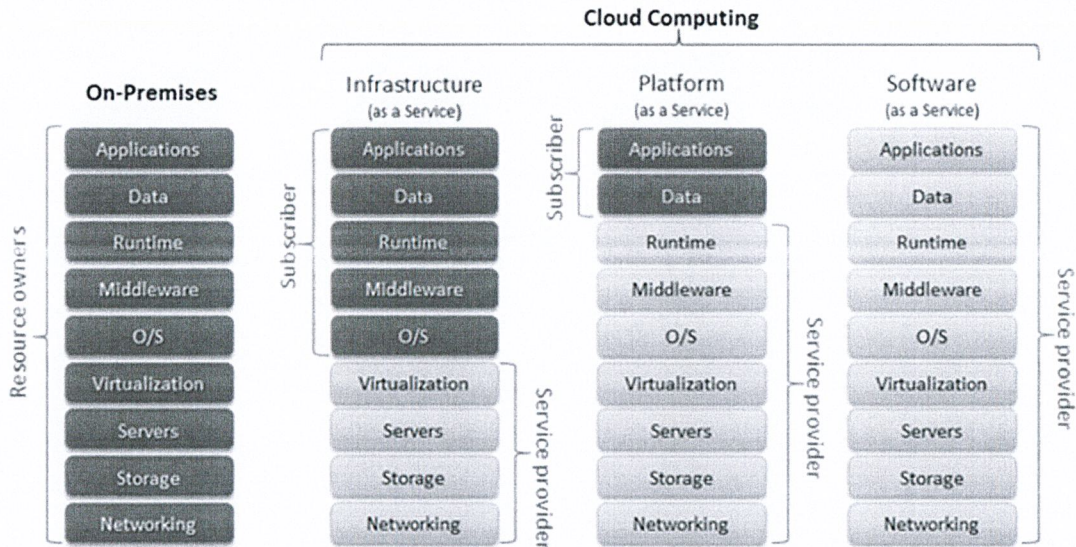


Figure 2.3 Cloud Computing

Cloud computing services consist of software-as-a-service (SaaS), platform-as-a-service (PaaS) and infrastructure-as-a-service (IaaS). The structure of the data center is divided into nine layers as shown in figure 2.4. It consists of networking, storages, servers, virtualization, operating systems (OS), middleware, runtime, data and applications. For the on-premises system, all resources as all layers are real physical hardware and software located in the organization that must be managed and controlled by the owners. The owners have to maintain all resources by themselves.



**Figure 2.4** Cloud computing services

### 2.3.1 Infrastructure as a Service

The infrastructure-as-a-service is the service for serving computing devices and IT infrastructure such as network, storage and CPU, these resources are owned by the cloud providers. The customers subscribe services that providers' IT infrastructure is same as their computing hardware. Then they must implement software and manage application by themselves. This model is compatible with almost customers who need the performance of computing hardware for computing and processing higher work load than the capability of the on-premise system. In some cases, the on-premise system is not available to use because of system downtime occurred while installing or modifying the system from various reasons such as hardware broken, electronic problem and devices failure. This is an important issue, the system has to be able to work while installing and implementing new hardware or software until the end of deployment process. Thus, the infrastructure-as-a-service is a solution for these problems as same as user

request for a virtual machine. They can determine the performance of hardware on demand. Then the customer needs to install and implement the operating system and application software.

### **2.3.2 Platform as a Service**

Platform-as-a-service is the service that provides computing platform and solution stacks as a service. The solution stack is a set of software subsystem such as operating systems, middleware, databases and connectors. For this model, the customer can request an application and a tool from a cloud provider. Also, deployment and configuration software environment can be requested. Any infrastructures such as the networks, servers and storages are provided. As presented in figure 2.4, the user manages only data and installs the application as they are using.

### **2.3.3 Software as a Service**

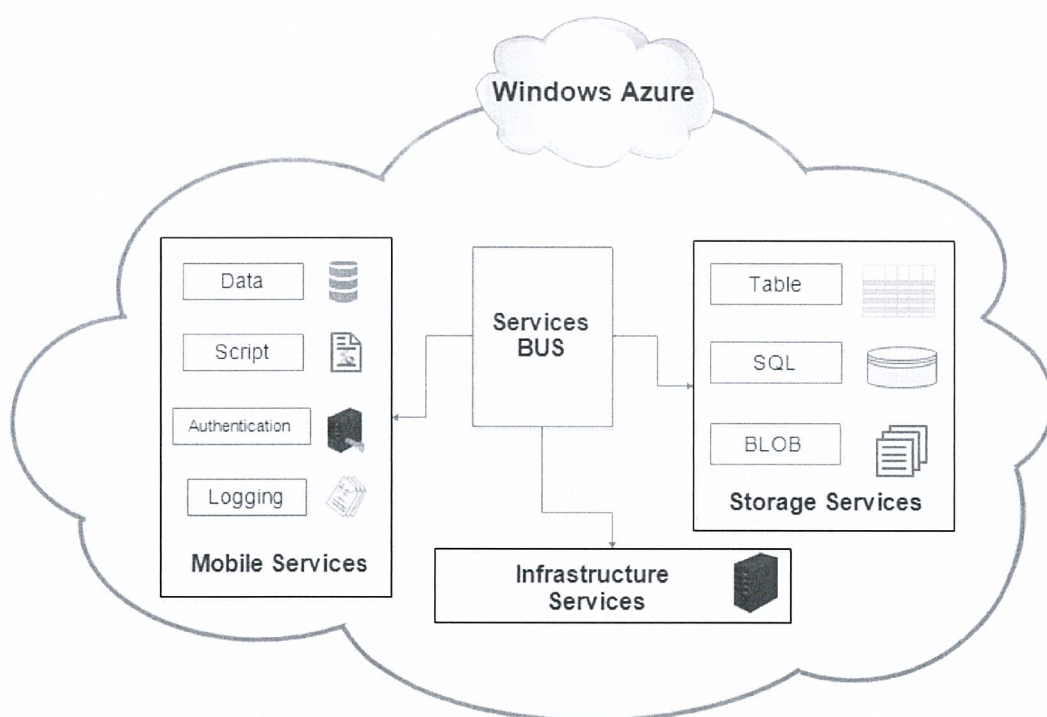
Software-as-a-service is the service for providing the software and utility application over the internet. It is a new solution for customers who need to use application that ease of implementation. Much software is hosted and provided by a vendor or cloud service provider, and the customer can access by a web service as a thin client. The effect of SaaS is delivery any software tools as on-demand services. Developers or end-users can use as they need and it reduces their responsibly task because they can only manage applications without deploying and maintaining hardware. It helps and supports IT user in many work processes. For example, SaaS can decrease the provisioning time to create a machine and implement software. User can request for any software as on demand and ease of use for general users. Users are not necessary to know where physical hardware is, how to upgrade and manage the system, and no need to maintain it by themselves as they can use any application as a service.

## **2.4 Windows Azure**

Microsoft Windows Azure is a cloud computing technology service created by Microsoft for providing any services such as infrastructure, software and application over the internet [20]. The customers can subscribe any services to quickly build, deploy and manage applications or any services via the internet. It is a new choice for the customers and users because it is a popular technology in the present that gives more benefit and has more advantage compared to the on-premise systems. For example, it can scale applications and infrastructure by the unlimited number servers. It also supports backing up and replicating cross location.

As presented in figure 2.5, Microsoft Azure cloud services consist of mobile services, storage services and infrastructure services. The mobile services provide mobility and comfort for

developers. The database can support many kinds of database. It also supports both structured data and unstructured data. We will apply it to create and store the knowledge base and rule base. This knowledge can be accessed from anywhere via the internet and many developers can connect for developing, and the users can also connect to access data as a service too. Finally, the services support for scalability function both database and computing unit. Because the mobile service connects to any services via the service bus that users can request for scale up or scale out the system as required, or let the system automatically be scaled depending on work load from an application, as known as automation.



**Figure 2.5** Windows Azure Services

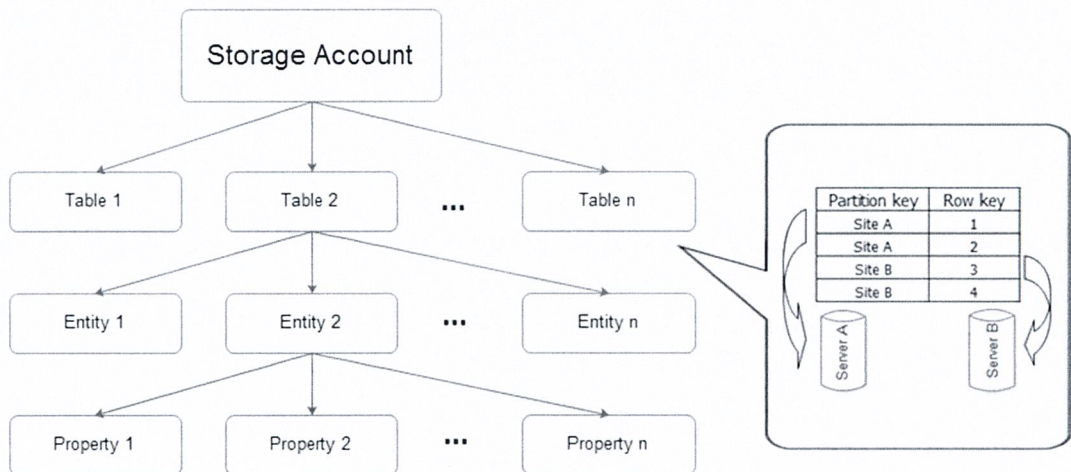
### 2.4.1 Windows Azure Table Storage

Windows Azure storage is one of the important components of the IT infrastructure service. It uses the storage and keeps any files such as log file, virtual hard disk, application and database files. This service provides space of storage as pay per user. It is generally provided for users that have subscribed the infrastructure as a service. However, the customer can request only a storage service for keeping their files and data. The Azure table storage is a non-relational, key-value-pair, storage system suitable for storing massive amounts of unstructured data [21, 22]. The non-relational stores like Table Storage are optimized for simple retrieval and fast insert. It is

generally used as the database for developing applications and also used with users who subscribe software-as-a-service and platform-as-a-service. Thus, users are not necessary to implement and maintain the database system. It is comfortable for users and ease of use.

This structure is like a table of a relational model, but not at all. Azure Table Storage doesn't have specific schema and relationship between each table and is able to get data into object to integrate with applications. In addition, data is not queried by SQL language. The Azure table structure is shown in figure 2.6. Table is a set of entities that is a set of properties. The table contains PartitionKey and RowKey property. PartitionKey is string to define the partition name in the table. It supports load balancing and distributed storage drive. Like a primary key, RowKey is unique for each partition which acts as a primary key for the row. Each entity in a table must have a unique PartitionKey/RowKey combination. They work together as an index for a table that enables efficient querying [23].

Azure table storage uses the Key/Value storing technique for processing that is one of NoSQL databases. Thus syntax of the command is like the SQL language. It searches from a key as same as a primary key but not at all. Since it does not have the SQL operating function, the data structure has more flexibility.



**Figure 2.6** Windows Azure table structure

### 2.4.2 Key-Value technique

The Azure table storage uses key-value pair technique for storing and retrieving the data. This technique is one of the popular mechanisms for non-relational database or unstructured data. The key-value pair consists of a key that is associated with a value [23]. It uses the associative

array for mapping the data. The key for each record will map with any variables that stores any fields [24]. Each record has a primary key and a collection of values. The key-value model can be extended to an ordered model that maintains the keys in order. This extension is powerful. In that case, it can efficiently process key ranges. The Key-Value storing technique is easy to learn and understand working process. It is also ease of use for developing any kinds of application.

## **2.5 Windows 8 application**

Windows 8 is a personal computer operating system. Windows 8 modified by the use on portable devices like a tablet to competing with other operating systems on mobile devices, including iOS and Android. In particular, these changes a touch-optimized based on Microsoft's metro design language, a new platform for developing the application that focus on touch screen input, integrate online services. As presented in figure 2.7, Windows Store is intended to download or purchase a new software. Windows 8 is support cloud computing. Not only it has a simple interface and easy to use, but also it has updated application anytime with the 'Live Tiles' and integrated built-in antivirus software into the operating system directly. The users don't need to buy an anti-virus for the devices.

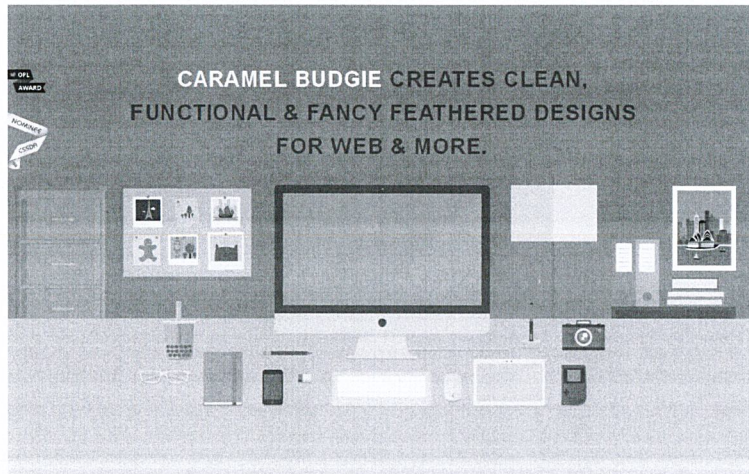
### **2.5.1 Windows Azure Library**

To develop Windows 8 application integrate with Windows Azure, first of all developers must install Azure library for connecting their application to Azure services, therefore they must configure the connection string by import and call Windows Azure API in their project. The type of library depends on type of platforms that developed such as Android, JAVA and IOS. However, the pattern of connection string of all platforms is similar. It consists of authentication key and database name or tables' name. The authentication key is unique key of each person for allowing to access any services of Windows Azure. The database name is the name of the database engine, it uses for identifying the database engine that be used to store the data.

### **2.5.2 Flat Design**

After Microsoft launched the Windows Phone 7 and Windows 8, people more began to recognize flat design with modern User Interface, because it's look simple and colorful. Because it's made flat design was popular at the time and it's widely used, whether it be a website or an application. Concepts of the flat design are designing everything look flat. It will reduce whatever to keep a look out whether the lighting and shadows. Inserting the object surface and gradients that it's made objects float out. Flat design concept wants the users are focused on the content.

The most obvious characteristic of the flat design is simplicity. We often see that the elements in flat design look flat without rendering, inserting the object surface gradients as well as other techniques that it's making the float off the screen.



**Figure 2.7** Flat design

#### **2.5.5.1 Advantages of Flat Design**

- a. The simple and straightforward of the various elements that it's causing the UI is simple and easy to understand too. Nothing bothers that make the users can find what you need so fast. Furthermore, the good positioning of the various elements that will help users focus on the content we have to offer. That means users will understand what we want to communicate more easily.
- b. The use vibrant color of the flat design that makes a website or application look so fun, amazing and beautiful. Even if it was exactly what we wanted to offer, a flat design attracted the attention of users as well.
- c. Flat design is ideal for making responsive web, because it has a bright color background, a size of the text is big and it's flat icons, but it is well meaning. These elements can be formatted to fit the screen in different sizes and the design is not as difficult as skeuomorphism.

Finally, in this project, we use windows 8 application because the format is simple and no complicated. The users can also easily understand. The windows 8 application is also compatible with windows azure library and support cloud computing.

## 2.6 Review of Literature

### 2.6.1 Clinical Diagnosis

Al-Masslawi et al [25] described the problem in clinical diagnosis that divided into two types. First, because of the necessary information that the patients told to doctor is not true or it has changed, so it makes the diagnosis error. Second, the patients told the unnecessary information to doctors. Not only it's a waste of time, but also it doesn't help in the diagnosis. So they were thinking of ways to help to solve both of these problems by using ontology-based framework, which helps communication between the doctor and patients to Get to the Point. The framework is a questionnaire system that selects only the necessary questions in medical diagnosis, so it will help make the information more efficiently and the time saving.

Phuong et al [5] described about combining case based reasoning with rule based reasoning. They developed an expert system to be used in the diagnosis of lung diseases. The case based reasoning system for medical diagnose that have three elements are the knowledge base, inference engine and explanation mechanism. Case based reasoning were used in the questionnaire model by the patients have to answer questions related to medical, such as symptoms , duration of symptoms and etc.. Then, the system will be analyzed the answer to help diagnose the disease may occur. In rule based reasoning, knowledge is a rule that same the actual format of the information in the brain. The work of case based and rule based is the introduction of new cases encountered in comparison with other cases that have occurred, to find that the case would be most acceptable. Then, it will be created as a new rule based. In this paper, they have used the hypothetical case based on the data of the patient to demonstrate the possibility of that case.

Shadabi et al [1] described the success story using the knowledge and rule induction methods to help clinical diagnosis. In the past, there has been interested in bringing advanced knowledge and discover techniques that mimic the behavior of humans. The purpose of this paper is describing the benefits of combining multi agent approach, hybrid machine learning and data mining technique to forecast the result in clinical. The development of medical systems must have property are powerful technique and more flexible, so that it can manage the amount of data and information was stored in clinical database. The medical facilities require knowledge systems and advanced information systems to improve the quality of care and used in the diagnosis of the prediction.

Gangwar et al [18] described they used a tree structure to help in making decision about the diagnosis of the symptoms associated with Alzheimer's or that called Neuropsychiatric diseases. The detection and diagnosis of diseases are based on different parameters. They diagnosed thirty eight parameters by using the form of decision tree based on C5.0 algorithm that used the Clementine tool for detection and calculate the value of each parameter at different levels. The advantages of decision tree structure are easy to use and help the doctors to analyze and diagnose diseases. Finally, it used the results of C5.0 algorithm to compare with the previous work. The benefits of C5.0 algorithm are easy and work fast, so the doctors can use to making decision about the diagnosis of diseases.

### **2.6.2 Knowledge Base and Hierarchical Data Model**

Wei et al [26] described this case implement the knowledge-base for GIS (Global Information System) application. The GIS application is developed for supporting user to planning and deciding. It uses the AI technology for decision maker. This application has three main functions; the knowledge-base for store fact and rule, state evaluate function, it is mechanism for searching in the state space and the third is a heuristic searching function. This method uses to find optimal solution from all candidate solutions by using mathematical programming method. It initializes solution and defines the state on state space, then create alternative solutions and calculate related values for comparing. This research applies both knowledge-based and AI technology for developing applications. It represents problem solving about problem with large of the data and has a more feasible solution.

Niwa [27] described knowledge based systems are classified as expert systems and knowledge-sharing systems in terms of knowledge transfers. Artificial Intelligence (AI) is designed and creates knowledge-based systems for solving actual problems. In knowledge-based systems, they are two types of data transfer that is Technology Transfer (TT) and Knowledge Transfer (KT). In this we are interested of knowledge transfer, because we need to discuss from the users' point of view. The knowledge transfer means the flow knowledge from knowledge suppliers of user. "The expert systems may be characterized by the following three features: 1) use of a knowledge base 2) use of an inference engine 3) an attempt to solve real-world problems as effectively as domain specialists do." And in this paper is focused on the need to use in term of KT, especially from the user's point of view. The limited of expert systems in management sections is difficult knowledge acquisition because we can't identify experts as knowledge

suppliers. And limited system functions are acquired knowledge from the fact only. The primary purpose of this is for users use knowledge and share it.

Zhang et al [8] described they used a hierarchical data structure to set the level of symptoms and herbs that used in the treatment. The disease that they interested in is Diabetes. The Traditional Chinese Medicine is a theory that describes how medical treatment in China. Medically, the doctors will diagnose the diseases from the different types of symptoms. And it will be given a drug used to treat individually. They used a parallel relationship between a group of symptoms and a group of herbs that cure properly. So they used hierarchical data structure because of part of the disease and herbs can be divided into sub-topic that like a tree structure. And this will help patients receive medication to treat the disease properly.

Huang et al [9] described they applied the structure of hierarchical classification to develop a classification system for sleep stage by using forehead EEG signals. And the system used only Fp1 and Fp2 EEG signals. The forehead EEG signals can be divided into intervals of 30seconds and each 30 seconds are classified to one sleep stage. The system can divide the stages into 4 stages that are woken, light sleep, deep sleep and REM. Finally, the system has more efficient and patients can process in their own sleep quality.

### **2.6.3 Rule Based**

Critien et al [28] described they designed rule based expert system to help facilitate the doctors in the diagnosis of patients that have a chronic heart failure. The system has two part, the first is related the diagnosis and the second is preferred methods to treat the diseases appropriately. The system does not need to be replaced an expert, but it is used as a tool to help diagnose the diseases more easily. The expert system also helps in the decision making process as well.

Semogan et al [29] described they used Rule-Based and Fuzzy Logics to assist in creating a system that helps in decision of doctors to patients with Tuberculosis. It use a rule base populated to define the class of Tuberculosis. They used the algorithm and architecture to evaluate the feasibility of the severity of Tuberculosis patients may be. The system analyzes the degree of severity of the disease are divided by class of Tuberculosis. So the level of severity that has shown to have a variety of symptoms. Then the system will conclude and assortment the class of Tuberculosis that the patient has.

Chen et al [4] described they designed a knowledge base expert system based on production rule. The expert system consists of knowledge base and inference engine. The knowledge base used to store knowledge and inference engine is a tool help to manage and configure the environment knowledge base. Inference engine enables knowledge discovery, knowledge base compares format to get the result. Production rule is the principle of matching between the actions and the results that have a relationship. When user adding a new rule into system, it makes a problem. So, it must be check conflict between the new rule and the existing rule. If it does not conflict, it allows adding to the system

#### **2.6.4 Cloud Computing Technology**

Zhou et al [13] propose the idea that involve in developing the knowledge base in cloud technology by the end user. With the internet and web of thing, bring to many mobile applications that suitable for the end user, they implement system environment for storing knowledge base and developing application for client. In their approach, users can store and collect their knowledge by upload via application or any internet sensor. It becomes to the knowledge in the cloud that can be accessed from anywhere. In the order to, they has inspiration from the expert system about the user's knowledge. Absolutely, all knowledge, fact and common information were found by human judged based on conditions, then it became to the knowledge in the form of the rule, also can present in the time series. By the way, the integration of user to develop the knowledge on cloud services that helpful user to take advantage of resource pool in a cloud environment.

Ju et al [30] described the model of knowledge service on the cloud computing technology. They said the knowledge could be the thing that useful for improving productivity, increasing human capital and active service, it not only the asset of mining. Meanwhile, the cloud computing is becoming a popular technology with many advantages such as high elasticity, scalability service, on-demand services and high availability. With these reasons, the concept knowledge as a service was created. It extends from the software as a service, and consists of delivery of knowledge service and organization knowledge resource. This concept can provide on-demand know who and know where that are a key success factor in the organization, also a delivery resource pool for collaborating base services as on-demand for all knowledge workers.

## Chapter 3

### Analysis and Design

In this chapter, we describe about how to design and develop this system, also any methodologies and technologies that is used including the clinical knowledge base, rule base and cloud computing. The rule base is an inference engine, and we use the hierarchical data structure to design database for cloud computing environment. Moreover, we present system architecture and how to implement step by step.

The problems may occur in everyday is an illness caused by various reasons that we don't know. So we decided to make the application to assist in the preliminary diagnosis. These diseases are caused by two symptoms which are cough and headache. Users can use to diagnose the symptoms that occur on their own to predict the diseases and continue treatment.

#### 3.1 Clinical Knowledge Base

The knowledge base is an important component of the expert system. In this project, we use clinical knowledge as a case study. Firstly, we must acquire knowledge from the expert and transform this information into computer data. The authors were supported by the medical students for gathering clinical knowledge by interviewing, and also some clinical books []. Moreover, they are the tester to test our system. Initially, the system can diagnose two diabetes which are diseases caused by coughing and headaches.

In addition, according to the study, we know that the clinical information has many kinds of diseases and for each disease has many symptoms that are reasons why the clinical diagnostic process is very difficult and complicated to identify type of disease. However, we can make it easier by classifying diseases information into many groups and represent in the hierarchical model. As previously discussed, the hierarchical model, as known as tree model, is the model that ease of understanding structure and data. After that we will store this information in the rule base to facilitate the implementation of the inference engine which will be discussed in the next section.

In this project, we are interested in diseases caused by the two main symptoms that are coughing and headache. From the study, we found diseases which can be divided into the

following groups: 1) categorized by duration of symptoms that are Chronic or Acute 2) the cause of those symptoms that is infection, tumor, trauma or the other, as described below.

### 3.1.1 Cough Symptom

In the cough section, we found diseases caused by coughing. It can be categorized by duration of symptoms that symptoms can be Chronic or Acute. The Chronic Cough means patients who have a cough for more than 3 weeks. And the Acute Cough means patients with coughing less than 1 week. The patient has sputum or not. The table 3.1 indicates the cause of coughing that is infection, tumor or the other. And it shows all of the diseases that caused by coughing.

**Table 3.1** The diseases caused of coughing

<b>Cough</b>	<b>Sputum</b>	<b>Infection</b>	<b>Tumor</b>	<b>Other</b>	<b>Result</b>
Chronic	Productive	yes	no	no	<b>Lung abscess</b>
Chronic	Productive	no	no	yes	<b>Bronchiectasis</b>
Chronic	Productive	yes	no	no	<b>Pulmonary Tuberculosis</b>
Chronic	Productive	no	yes	no	<b>Lung Cancer</b>
Chronic	Productive	no	no	yes	<b>Chronic Bronchitic</b>
Chronic	Non Productive	no	no	no	<b>Postnasal Drip</b>
Chronic	Non Productive	no	no	no	<b>Asthma</b>
Chronic	Non Productive	no	no	no	<b>Post Infective Cough</b>
Chronic	Non Productive	no	no	no	<b>Drug Induced Cough</b>
Acute	Productive	yes	no	no	<b>Acute Sinusitis</b>
Acute	Productive	yes	no	no	<b>Laryngitis</b>
Acute	Non Productive	yes	no	no	<b>Pharyngitis</b>
Acute	Non Productive	yes	no	no	<b>Pneumonia</b>
Acute	Non Productive	yes	no	no	<b>Bronchitis</b>

As presented in table 3.2, the table describes the categories caused by coughing and explains the occurrence of the disease that must have some symptoms in the table. It means if a

patient has symptoms of disease, what disease patients can be. For example, the patients were told the symptoms which are clubbing finger and dyspnea, so the patients may be a Bronchiectasis. And this table explains that diseases are a chronic cough and productive sputum.

**Table 3.2** Indicative that the diseases caused by coughing

<b>Definition</b>	<b>Cough</b>	<b>Clubbing Finger</b>	<b>Sore Throat</b>	<b>Dyspnea</b>	<b>Fatigue</b>	<b>...</b>
<b>Lung abscess</b>	Chronic	yes	no	no	no	...
<b>Bronchiectasis</b>	Chronic	yes	no	yes	no	...
<b>Postnasal Drip</b>	Chronic	no	no	yes	no	...
<b>Lung Cancer</b>	Chronic	yes	no	no	no	...
<b>Laryngitis</b>	Acute	no	no	no	no	...
<b>Pharyngitis</b>	Acute	no	yes	no	no	...
...	...	...	...	...	...	...

From the table 3.2, it is the table of cough diseases as sample data. The header of each column is a main symptom for separating groups of cough diseases. We use them to create rules for classifying any groups. Each symptom adapts to conditions on each node in the tree map. As presented in figure 3.1, we start a root node by classifying main sub-group of cough disease. It shows that the cough disease can classify into two sub-groups including acute cough and chronic cough. These groups can be deeply classified into many sub-groups with other symptoms as shown in figure 3.1. Until it is not able to separate, it means the type of disease is found out as the leaf nodes from the tree diagram.

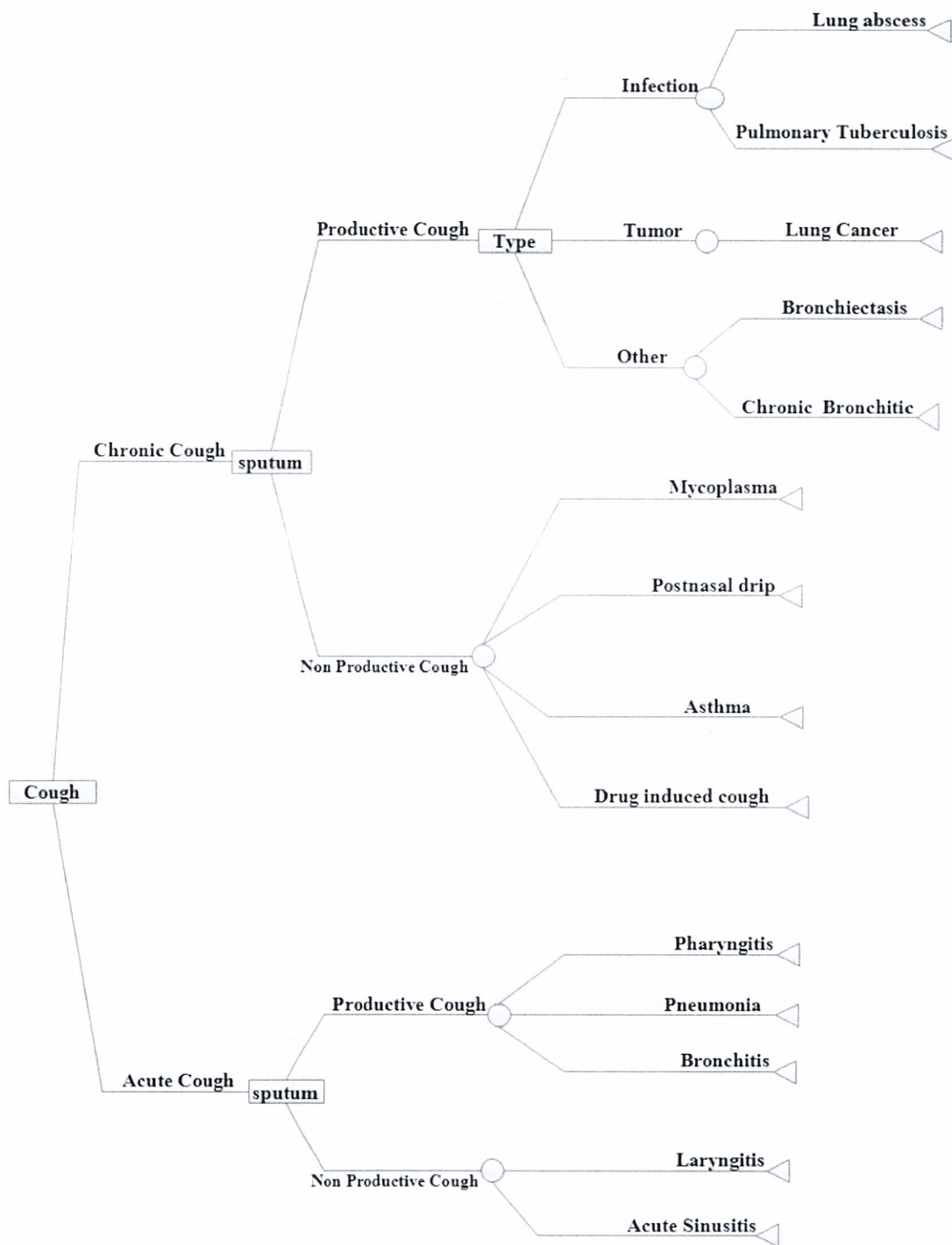


Figure 3.1 The tree structure of the cough disease

### 3.1.2 Headache Symptom

In the headache section, we found diseases that caused by headache. It can be divided into the following groups: 1) categorized by duration of symptom that is Acute or Chronic 2) classified as an Organic Headache or Functional Headache 3) the cause of headache that is infection, trauma or the other. The Acute Headache means patients who have a headache for less than 1 week. And the Chronic Headache means patients with headache more than 1 week. And the table shows all of the diseases caused by headache.

**Table 3.3** The diseases caused of headache

<b>Headache</b>	<b>Cluster</b>	<b>Infection</b>	<b>Trauma</b>	<b>Other</b>	<b>Result</b>
Acute	Organic	yes	no	no	<b>Acute Sinusitis</b>
Acute	Organic	yes	no	no	<b>Acute Meningitis</b>
Acute	Organic	yes	no	no	<b>Acute pharyngitis</b>
Acute	Organic	no	yes	no	<b>Subarachnoid</b>
Acute	Organic	no	no	yes	<b>Acute Glaucoma</b>
Acute	Organic	no	no	yes	<b>Hypertension Induced</b>
Acute	Organic	no	no	yes	<b>Drug Induced Headache</b>
Acute	Functional	no	no	no	<b>Migrain</b>
Acute	Functional	no	no	no	<b>Cluster Headache</b>
Acute	Functional	no	no	no	<b>Tension Type Headache</b>
Chronic	Organic	yes	no	no	<b>Chronic Sinusitis</b>
Chronic	Organic	no	yes	no	<b>Chronic Subdural</b>
Chronic	Organic	no	no	yes	<b>Brain Tumor</b>
Chronic	Organic	yes	no	no	<b>Chronic Meningitis</b>
Chronic	Organic	yes	no	no	<b>Brain Abscess</b>

As presented in table 3.4, the table describes the category caused by headache and explains the occurrence of the disease that must have some symptoms in the table. For example, the patients were told that the symptoms are fever and stiff neck, so the patients may be chronic meningitis. And this table explains that diseases are a chronic headache and organic headache.

**Table 3.4** Indicative that the diseases caused by headache

<b>Definition</b>	<b>Type</b>	<b>Fever</b>	<b>Stiff Neck</b>	<b>Pain</b>	<b>Aura</b>	<b>Headache</b>	<b>...</b>
<b>Acute Sinusitis</b>	Acute	yes	no	no	no	no	...
<b>Acute Glaucoma</b>	Acute	no	no	no	no	no	...
<b>Acute Pharyngitis</b>	Acute	yes	no	no	no	no	...
<b>Migraine</b>	Acute	no	no	no	yes	no	...
<b>Cluster Headache</b>	Acute	no	no	no	no	no	...
<b>Brain Abscess</b>	Chronic	no	no	no	no	no	...
<b>Chronic Meningitis</b>	Chronic	yes	yes	no	no	no	...
<b>Chronic Subdural Hematoma</b>	Chronic	no	no	no	no	no	...
...	...	...	...	...	...	...	...

From the described above, we have a table that indicates the type of diseases caused by coughing and headache. The cause of the symptoms and symptoms appear which can be diagnosed by following. When we receive the data, then it will be used for the designing and analyzing the data.

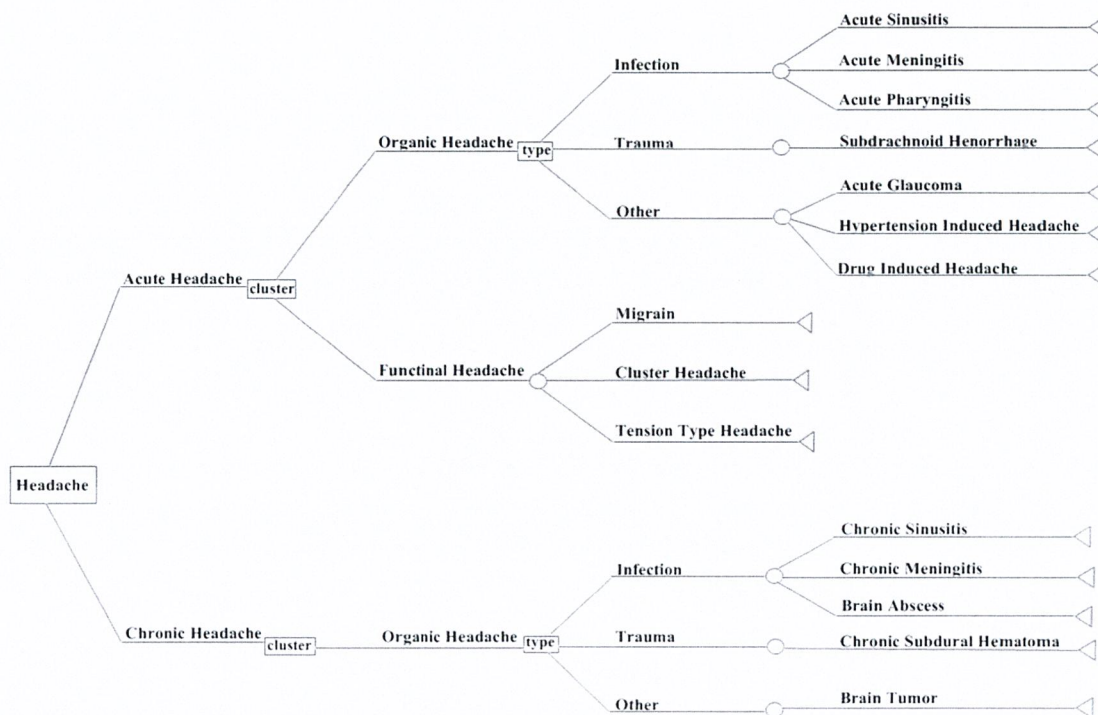


Figure 3.2 The tree structure of headache

### 3.2 Rule Base Inference Engine

The rule base is useful inference engine to find out the conclusion. It uses the rule for proofing the condition that it is a true or false statement. We combine the forward chaining and the backward chaining approach for clinical diagnosing. The forward chaining approach uses for identifying groups of disease to narrow scope of the source of disease by collected the main symptoms such as duration, sputum and infection. But only this approach can't provide the final conclusion because it seems difficult to proof all rules for identifying just one result. The forward chaining let us know the group of disease then we assume some of disease in these group that may be the result of clinical diagnosis. Thus, the backward chaining used to proof the assumption by checking only the rules that consistency of the hypothesis.

For diagnosing and identifying the type of clinical disease, we use all symptoms of each disease for checking which symptoms that the patient are suffered. So, many kinds of symptoms is the problem, it very difficult to determine the type of clinical disease by proof many conditions. Thus, we can narrow the scope of type of disease. by using the hierarchical model. As previously discussed, the rough analyze of clinical information delivery the tree chart. Although the tree chart can narrow the scope of the problem for basic identifying what is the disease that patients

suffered, it can't clearly identify the type of disease. Because each condition are not clearly for indicating final conclusion as a single disease. As presented in figure 3.3, any conditions can't provide one finally leaf node. For example, the patient is chronic cough and non-productive cough, the possible type of diseases include microplasma, asthma and postnasal drip.

From the tree model, we create a rule as "IF.....THEN....." pattern as shown in table 3.5 and 3.6. The table consists of two columns include "IF" and "THEN". As the table 3.5, the column "IF" is factual information that the patients are suffering, and the conclusion present at "THEN" column as group of disease. This table present basically patient's information such as how long are the patient coughing and coughing with sputum or without it. That lets us known group of source of disease such as chronic cough or acute cough. The rules from this table proof as forward chaining approach to determine a group of disease as shown on the table 3.6. The column "IF" is kinds of diseases and the column "THEN" is a type of cough disease. But the only information from table 3.5 and 3.6 are not enough to clearly identify because each of the diseases has more unique symptoms to be checked. The result from table 3.6 are the assumption. We assume some of diseases from table 3.6 are the disease that patient suffered, and then we test only the rule base from table 3.7 that conform to the symptoms of each disease to find out the conclusion.

**Table 3.5** Classification of cough disease

<b>IF</b>	<b>THEN</b>
Coughing more than week	Chronic cough
Coughing less than week	Acute cough
Cough with sputum	Productive cough
Cough without sputum	Non Productive cough
Infected	Cough Infectious diseases
Have tumor	Cough Tumor Disease

**Table 3.6** Symptoms of cough disease

<b>IF</b>	<b>THEN</b>
Acute Cough	Acute Sinusitis
Acute Cough	Bronchitis
Acute Cough	Laryngitis
Acute Cough	Pharyngitis
Acute Cough	Pneumonia
Chronic cough	Asthma
Chronic cough	Bronchiectasis
Chronic cough	Chronic Bronchitis
Chronic cough	Lung Abscess
Chronic cough	Lung Cancer
Chronic cough	Postnasal Drip
Chronic cough	Pulmonary Tuberculosis
Cough Infectious diseases	Lung Abscess
Cough Infectious diseases	Pulmonary Tuberculosis
Cough Specific Disease	Bronchiectasis
Cough Specific Disease	Chronic Bronchitis
Cough Tumor Disease	Lung Cancer
Non Productive cough	Asthma

**Table 3.6** (continued) Symptoms of cough disease

<b>IF</b>	<b>THEN</b>
Productive cough	Bronchiectasis
Productive cough	Chronic Bronchitis
Productive cough	Lung Abscess
Productive cough	Lung Cancer
Productive cough	Pharyngitis
Productive cough	Pneumonia

The table 3.7 present the symptoms and result of these symptoms. The column “IF” is any symptoms and the column “THEN” is a type of cough disease. After we tested the rules from table 3.5 and 3.6, our system will test the rules from table 3.7 by checking a list of symptoms. This step like the backward chaining approach because we determine the type of disease then we proof each symptom that the cause of assumption disease are suffered by the patient. If all statements are true, it means the assumption is real.

For example, we describe the rule for identifying result as lung cancer. The patient has coughing with sputum more than a week and tumor, that can conclude the patient who are acute cough, productive cough and cough tumor disease that are lung cancer.

Rule: 1            IF        the patient has coughing more than week

                         THEN the patient is chronic cough

Rule: 2            IF        the patient has coughing with sputum

                         THEN the patient is Productive cough

Rule: 3            IF        the patient has tumor

                         THEN the patient is Cough Tumor Disease



**Table 3.7** (continued) Rule of coughing disease by symptoms

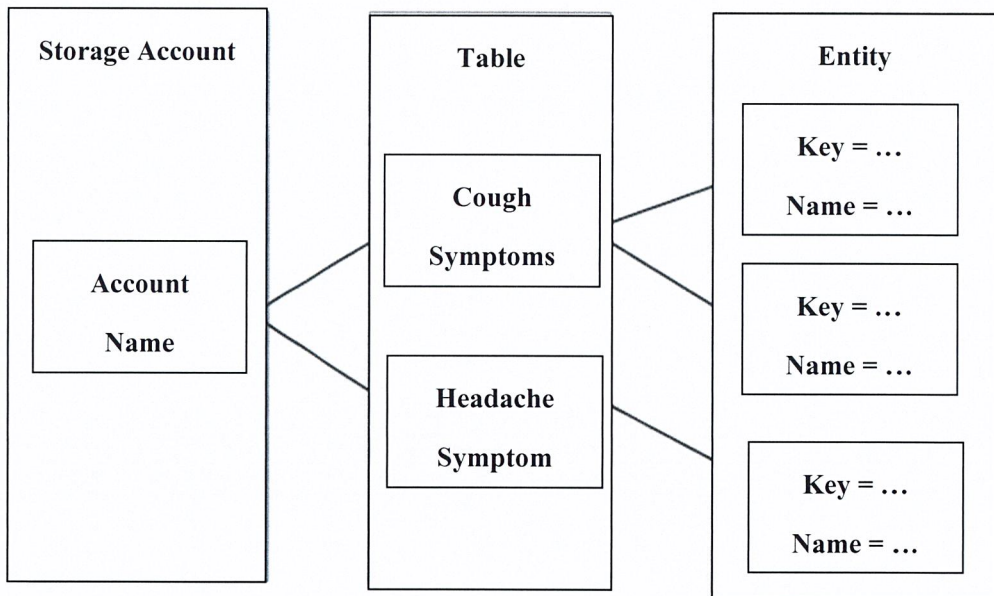
<b>No.</b>	<b>IF</b>	<b>THEN</b>
15	Fever	Bronchitis
16	Shortness of breath	Bronchitis
17	Wheezing	Bronchitis
18	Fatigue	Chronic Bronchitic
19	Headache	Chronic Bronchitic
20	muscle ache	Chronic Bronchitic
21	Sore throat	Chronic Bronchitic
22	stuffy nose	Chronic Bronchitic
23	History of taking drug producing cough	Drug induced cough
24	Dry cough	Laryngitis
25	Dry throat	Laryngitis
26	Fever	Laryngitis
27	Hoarseness	Laryngitis
28	Sore throat	Laryngitis
29	Tickling sensation and rawness of your throat	Laryngitis
30	Weak voice	Laryngitis
31	Chest pain	Lung abscess
32	Clubbing finger	Lung abscess
33	Fever	Lung abscess

**Table 3.7** (continued) Rule of coughing disease by symptoms

<b>No.</b>	<b>IF</b>	<b>THEN</b>
34	Bone pain	Lung Cancer
35	Chest pain	Lung Cancer
36	Coughing up blood	Lung Cancer
37	Clubbing finger	Lung Cancer
38	Headache	Lung Cancer
39	Hemoptysis	Lung Cancer
40	Hoarseness	Lung Cancer
41	Losing weight without trying	Lung Cancer
42	Shortness of breath	Lung Cancer
43	Wheezing	Lung Cancer
44	Fever	Pharyngitis
45	Sore throat	Pharyngitis
46	trouble swallowing and breathing	Pharyngitis
47	Dyspnea	Pneumonia
48	Fatigue	Pneumonia
49	Fever	Pneumonia
50	Nausea	Pneumonia
51	Shortness of breath	Pneumonia
52	Dyspnea	Postnasal drip

### 3.3 Cloud Database Structure

As discussed from the previous chapter. The cloud database has more advantage from on the premised. Thus, we will develop database by using cloud computing technology. However, the cloud service providers have many companies. For each provider has unique technology that it affect to the process for implementation. In this case, we use an Azure table from the Windows Azure by Microsoft. As presented in figure 3.3, Windows Azure table work through storage account. For one storage account, it can create many tables that these tables like the table on relational database.

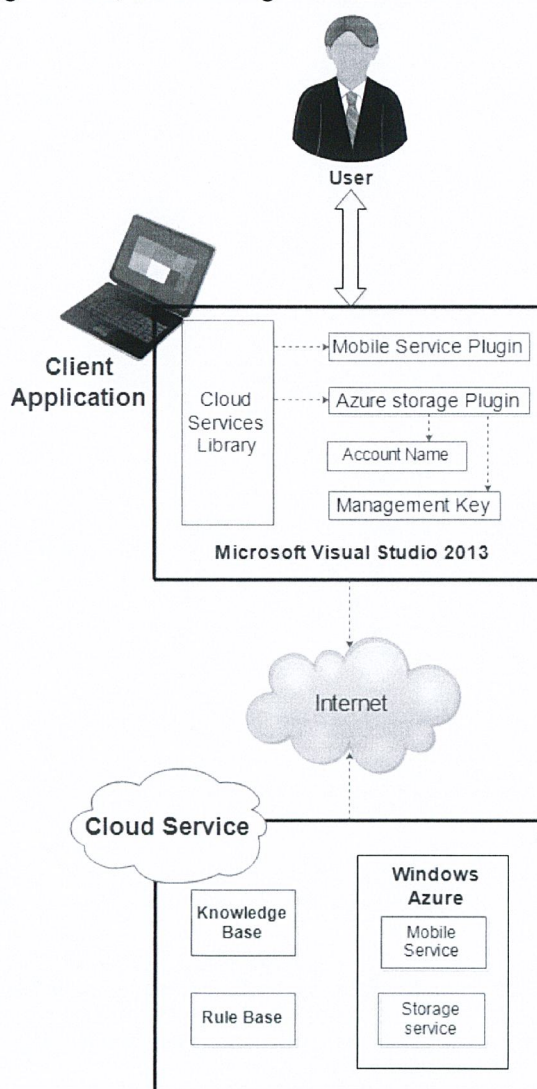


**Figure 3.3** Structure of Azure table storage

The Azure table storage has three main attributes include Partition key, Row Key and time stamp. Partition Key is the key that represents a partition of the database or group of records. Also the partition key can distribute any records in one table into any servers. Row key is the key that present uniquely for all records in each partition. It is similar to primary key on relational database. And time stamp is the attribute that show time when operate the data in each record. With the partition key, one table can separate into many partitions like sub tables. Windows Azure table storage can support for distributed database system. So we can create only one Azure table that consists of many sub tables by partition key for storing the data. However, Windows Azure table storage was designed for storing the datasets that don't require complex joins, foreign keys, or stored procedures and can be de-normalized for fast access.

### 3.4 System Architecture

As shown in figure 3.4, an application is developed on premise system which operates on Windows 8 operating system. It connects to the cloud storage by the cloud service library via the Internet. The library consists of Azure storage plugin and a mobile service plugin. The Azure storage plugin requires the account and a management key for using storage services. And mobile service used for developing an application. In case of cloud service, we implement the knowledge base and rule base on the Windows Azure platform. It has two main components; mobile service and storage. Firstly, mobile service on cloud side, it acts as a server by delivery database and API for connecting to client side. It also provides tools and portal page for manage, monitor and control any services. Finally, storage service is storage as a service. Actually, the mobile service acts like the front end of the storage service, it uses storage service as database to provide to the user.



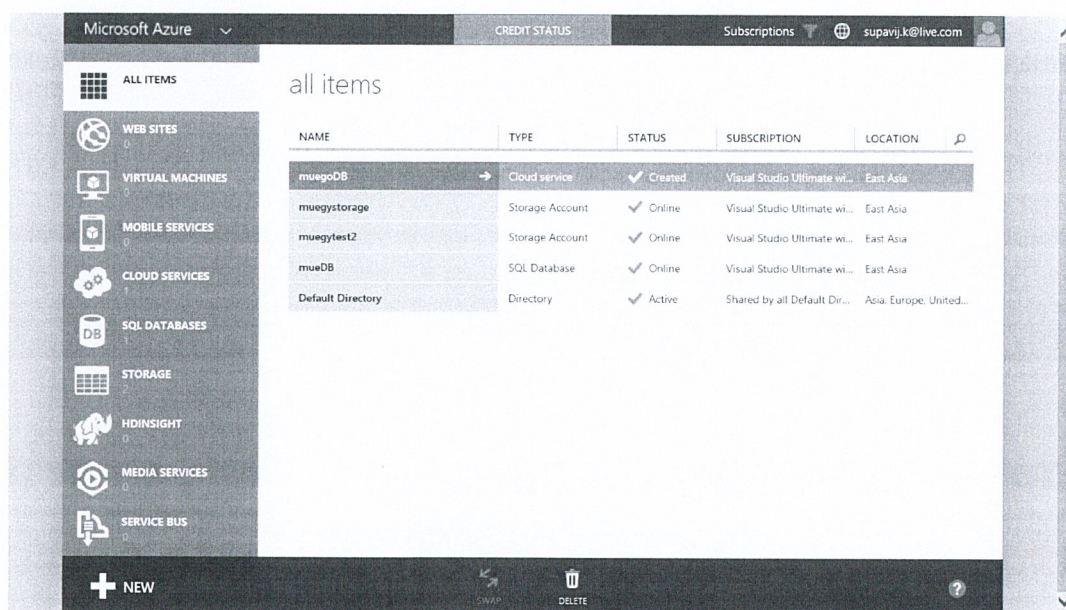
**Figure 3.4** System Architecture Diagram

### 3.5 Implement Rule Base Inference Engine to Cloud Service

In this section, we describe how to implement the rule base into cloud database and accessing process for developing applications. Include prepare and setup Azure services, deploy knowledge into cloud service. Finally, we present how to configure an application for connecting to the Azure services.

#### 3.5.1 Setting Windows Azure Services

Firstly, we create cloud service and storage service on Windows Azure by the Azure portal page via the internet. It used to manage, modify, configure and delete the service. As shown in figure 3.5, this is the Windows Azure portal page. From the left side, it is menu to request any service such as web sites, virtual machine and storage. At the center of the screen, there are all of subscribed the services, it provides description such as name, service type, status and location. On the top bar, it presents email account and credit subscriptions.



**Figure 3.5** Windows Azure Portal

In addition to, we use the Azure Storage Explorer for managing data on the client side. It is a program used for managing and editing cloud database service on Windows Azure. At present in figure 3.6, this is an interface for manages the information on Azure table. It can add, edit and remove both records in the table and the table. Also upload and download data from the table. First of all for using this program, the developer must initialize this program by setting storage account name and storage account key as shown in figure 3.7.

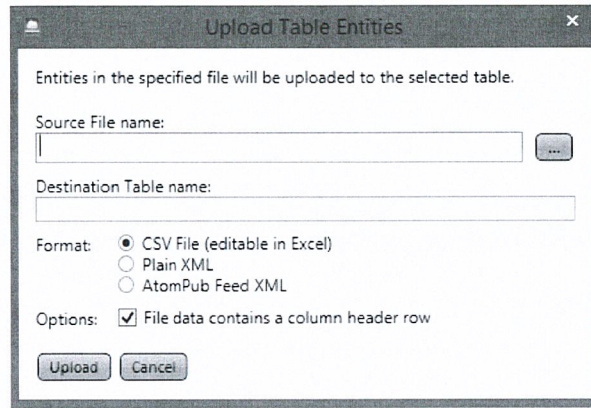
PartitionKey	RowKey	IF	THEN	Timestamp
Cough	1	Difficulty breathing	Acute Sinusitis	4/16/2015 11:18:21 AM
Cough	10	Wheezing	Asthma	4/16/2015 11:18:23 AM
Cough	11	Clubbing finger	Bronchiectasis	4/16/2015 11:18:23 AM
Cough	12	Diaphoresis	Bronchiectasis	4/16/2015 11:18:23 AM
Cough	13	Fever	Bronchiectasis	4/16/2015 11:18:23 AM
Cough	14	Fatigue	Bronchitis	4/16/2015 11:18:24 AM
Cough	15	Fever	Bronchitis	4/16/2015 11:18:25 AM
Cough	16	Shortness of breath	Bronchitis	4/16/2015 11:18:25 AM
Cough	17	Wheezing	Bronchitis	4/16/2015 11:18:24 AM
Cough	18	Fatigue	Chronic Bronchitic	4/16/2015 11:18:24 AM
Cough	19	Headache	Chronic Bronchitic	4/16/2015 11:18:25 AM
Cough	2	Fever	Acute Sinusitis	4/16/2015 11:18:22 AM
Cough	20	muscle ache	Chronic Bronchitic	4/16/2015 11:18:24 AM
Cough	21	Sore throat	Chronic Bronchitic	4/16/2015 11:18:25 AM
Cough	22	stuffy nose	Chronic Bronchitic	4/16/2015 11:18:25 AM
Cough	23	History of taking drug producing cough	Drug induced cough	4/16/2015 11:18:25 AM
Cough	24	Dry cough	Laryngitis	4/16/2015 11:18:25 AM
Cough	25	Dry throat	Laryngitis	4/16/2015 11:18:25 AM
Cough	26	Fever	Laryngitis	4/16/2015 11:18:25 AM
Cough	27	Hoarseness	Laryngitis	4/16/2015 11:18:26 AM
Cough	28	Sore throat	Laryngitis	4/16/2015 11:18:26 AM
Cough	29	Tickling sensation and rawness of your throat	Laryngitis	4/16/2015 11:18:26 AM
Cough	3	Nasal obstruction	Acute Sinusitis	4/16/2015 11:18:21 AM
Cough	30	Weak voice	Laryngitis	4/16/2015 11:18:26 AM
Cough	31	Chest pain	Lung abscess	4/16/2015 11:18:26 AM
Cough	32	Clubbing finger	Lung abscess	4/16/2015 11:18:26 AM
Cough	33	Fever	Lung abscess	4/16/2015 11:18:27 AM
Cough	34	Night sweating	Lung abscess	4/16/2015 11:20:11 AM

Figure 3.6 Azure Storage Explorer

Figure 3.7 Initialize Azure storage account

### 3.5.2 Implement rule base into cloud service

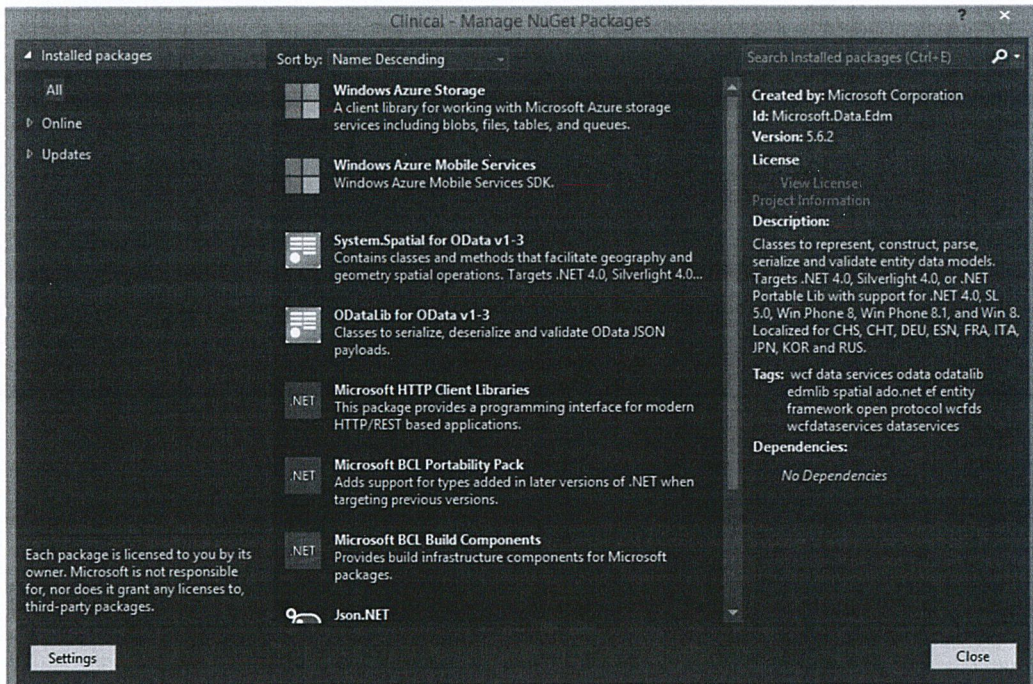
As described in section 3.3, those tables store rule base of clinical diseases that be implemented. We create those tables of rule base on spreadsheet program, then determine parameter include the partition key as a group of diseases such as cough and headache, the row key as ordering of rule, column “IF” and “THEN”. In addition to, we create the table in Azure table storage for storing all rule base tables, the only one table can store all rules of clinical diseases because the one Azure table can divide into sub table as partitions. Each partition stores the rule for each type of diseases. Therefore, we upload table of rule base as csv files through Azure Storage Explorer into Azure storage service as shown on figure 3.8.



**Figure 3.8** Upload csv file to Azure table

### 3.5.3 Connecting application to Azure services

For connecting the windows 8 application with Azure services, the developers must install Windows Azure storage and Windows Azure Mobile services SDK into the project as shown in figure 3.9.



**Figure 3.9** List of Azure services library

Then import these libraries in each source code file. Include:

```
using Microsoft.WindowsAzure.MobileServices;
using Microsoft.WindowsAzure.Storage;
using Microsoft.WindowsAzure.Storage.Auth;
using Microsoft.WindowsAzure.Storage.Table;
```

Therefore, create the “connection string”, reference variable as “*CloudStorageAccount*” and “table client” as parameter to determine which table will be accessed. Finally, create the object as middleware between application and Azure services to access any data in the Azure storage.

- a) *Create the connectionstring*

```
String StorageConnectionString =
    "DefaultEndpointsProtocol=https;AccountName=[yourAccount];AccountKey=[your
    Key]";
```

- b) Retrieve storage account from connection string

```
CloudStorageAccount storageAccount =
    CloudStorageAccount.Parse(StorageConnectionString);
```

- c) Create the table client

```
CloudTableClient tableClient = storageAccount.CreateCloudTableClient();
```

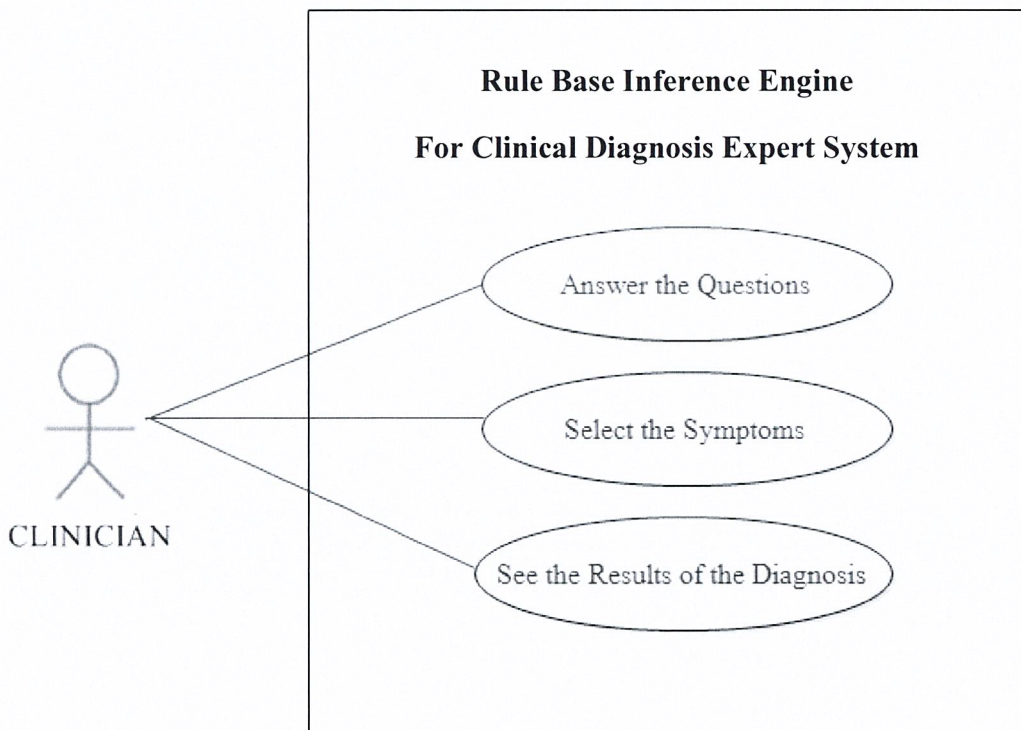
- d) Create the CloudTable object that represents the "customer" table

```
CloudTable table = tableClient.GetTableReference("customer");
```

Besides the structure of the Azure table, developer software tool and language syntax are an important factor for implementing the cloud database. Because the operating function depends on the kinds of the command. In this project, we use Visual Studio 2013 by using C# language. For the Azure table service, it uses LINQ syntax. The Microsoft .NET Client Library supports data service queries by using language-integrated queries (LINQ). With LINQ, the Table service supports executing simple queries that retrieve all properties of an entity. Also supports filtering query results using the Where operator, and specifying how many entities to return using the Take operator. But LINQ are unsupported some query operator from RDMS or SQL format such as group by, average, min, max and etc. However, LINQ and C# can work together. It's support for developing applications in object orient approach. LINQ can work suitable with object data type. In terms of implementation, we can query the information directly to object type. So it has more comfortable when we develop an application.

### 3.6 Use Case Diagram

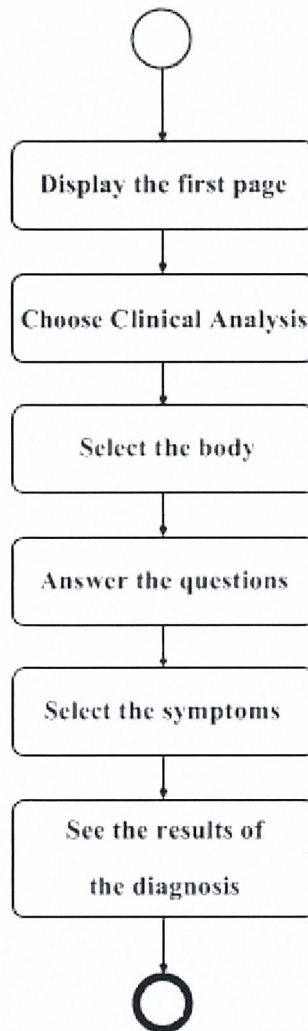
In this system have three major activities that clinician can interact with the system as shown in figure 3.10. The first, clinician can answer questions about the symptoms that caused by coughing or headache. Then, clinicians must be selects the symptoms that occur. Finally, the system will diagnosis and concluded the result. So, the clinician can check the results of the diagnosis.



**Figure 3.10** Use Case Diagram

### 3.7 Activity Diagram

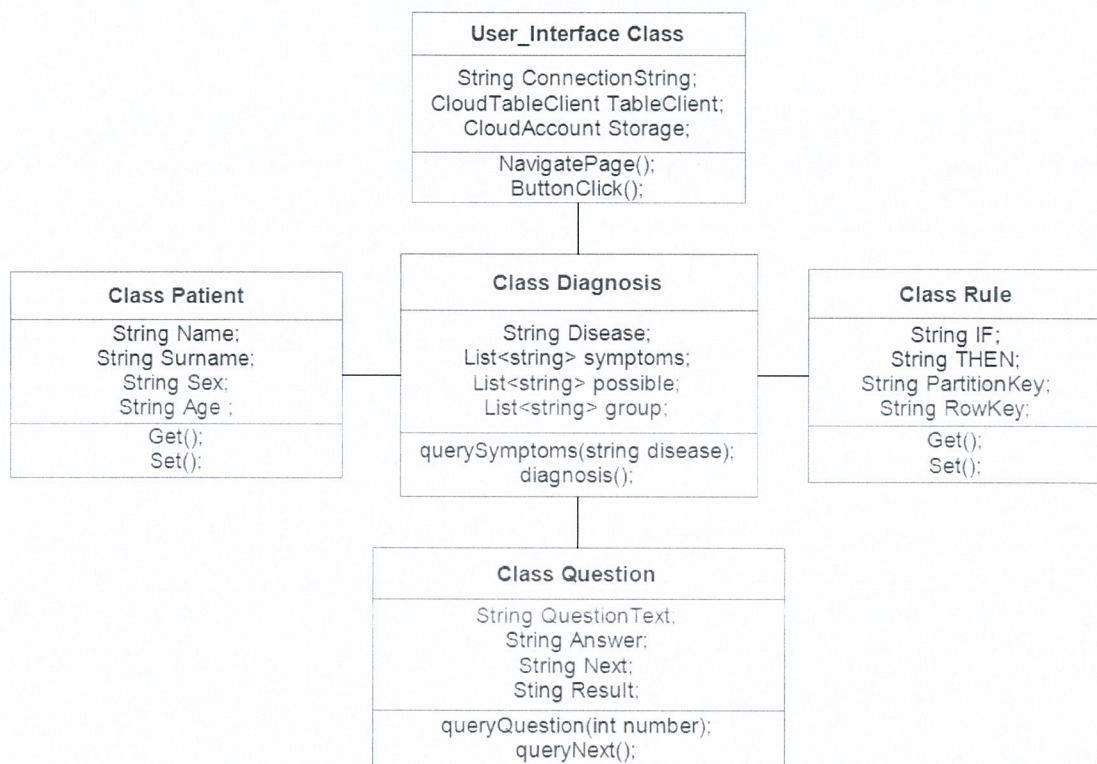
Activity diagrams present the work flows of activities that exist within the systems, which is Choose Clinical Analysis. An activity diagram of choosing a clinical analysis as shown in figure 3.11., the system will display the first page that users choose clinical analysis. And users select the part of the body that is ill or has a problem. Then users must answer the questions truthfully, the results have to be faultless. When users answering all of the questions, the system going to the next page. Users must select the symptoms that occur. Finally, the system will diagnose the diseases that users may be.



**Figure 3.11** Activity Diagram

### 3.8 Class Diagram

The class diagram of our application shown in figure 3.12, it has 5 classes include “Patient”, “Diagnosis”, “Rule”, “Question” and “User Interface”. The user interface class is a class for representing the user interface, it consists of some attributes as previously discussed in section 3.5.3 and controlling method such as event listener and navigation. Consequently, class rule and question are the entity class that represents data from Azure services, in case we have the rule and questions on the Azure table. The attribute of these classes are the attributes from a column of the table that represent value for each row. Furthermore, the patient class is an entity class for storing patient’s information. The last, the diagnosis class is controlling class that is the core process of this application. This class refers to the inference engine parts, the method in this class used for analysis and connect to a cloud service for processing data.



**Figure 3.12** Class Diagram

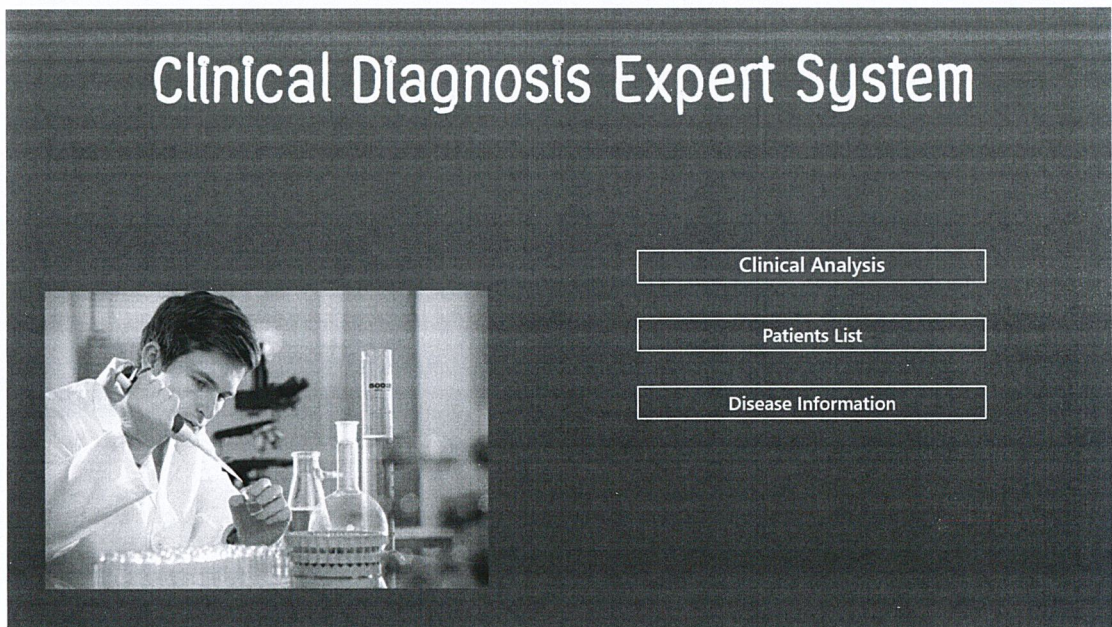
## Chapter 4

### RESULTS AND DISCUSSION

This chapter discusses about the main function in the application including choosing a clinical analysis. This application is ideal for those who want to know the symptoms occurring and what diseases they can be.

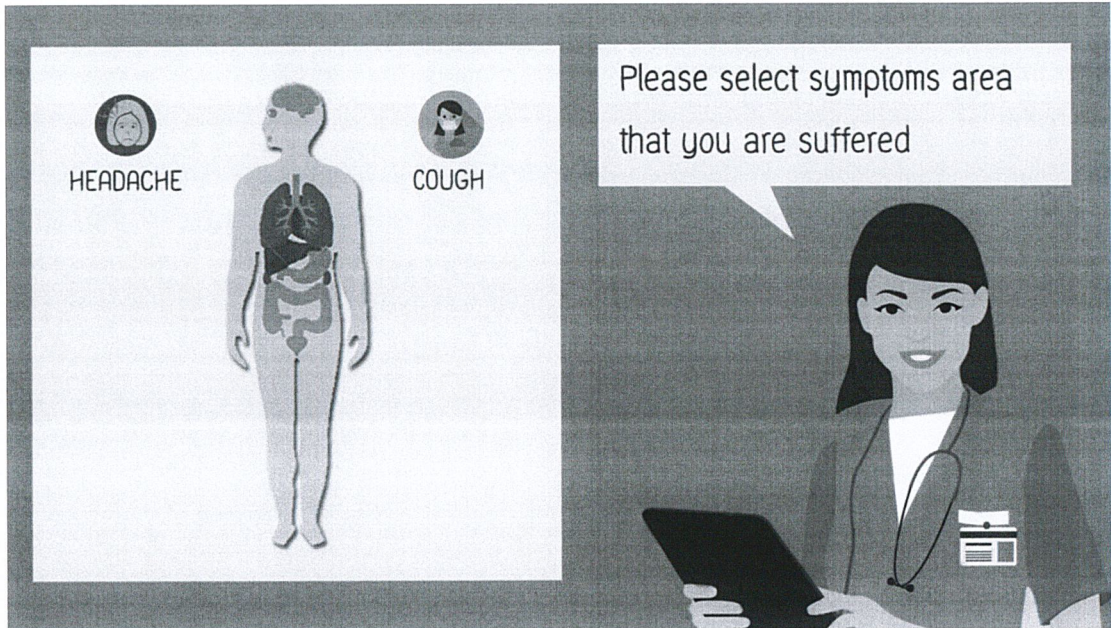
#### 4.1 User Interface

When users open the system, they must click the button Clinical Analysis as shown in figure 4.1 and then the system will diagnose the diseases that they can be.



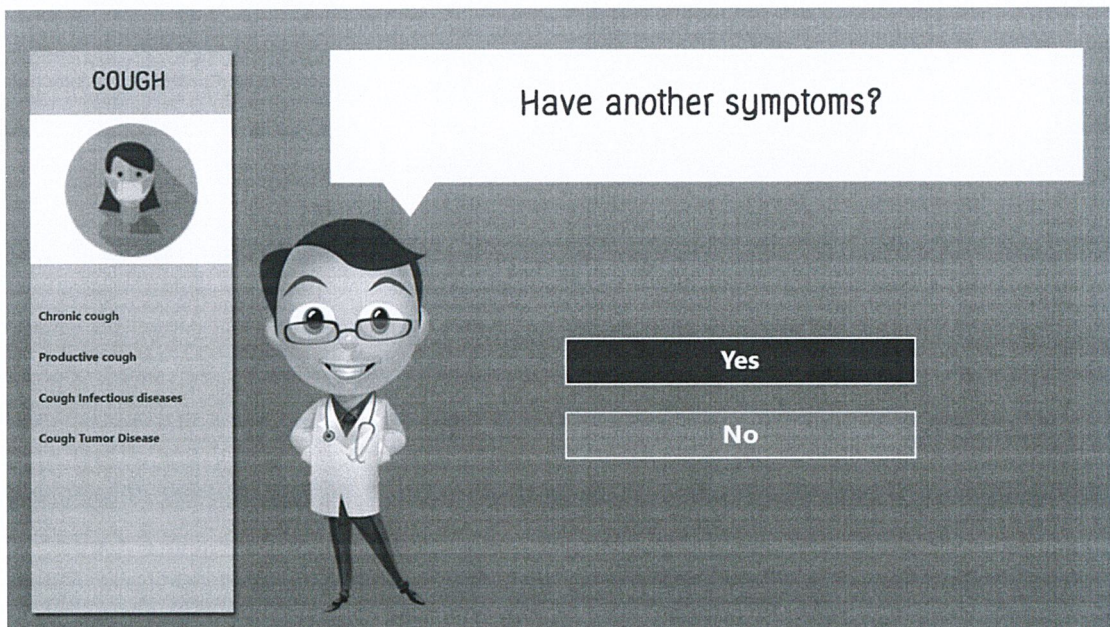
**Figure 4.1** Main menu

After choosing clinical analysis, it will turn to the next page as shown in figure 4.2. Users must select a part of the body that has an illness. In this project, users can select two parts which are the neck and the head.



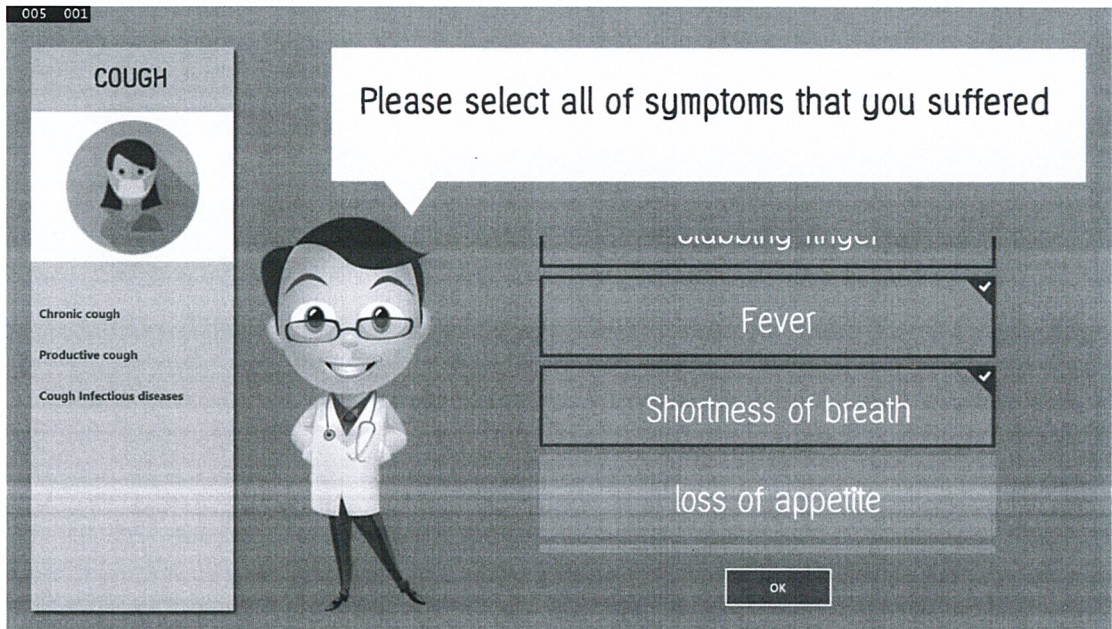
**Figure 4.2** Selects a part of body that have an illness

If users select the neck, the system will be directed to an answering questions page. The system has many questions to ask about symptoms. Patients must answer those questions as true as possible. And the next question is depended on the given answer as shown in figure 4.3.



**Figure 4.3** Answer the questions that asked about symptoms

When users have already answered the questions, the system will turn to the next page. The users must choose the various symptoms that occur as shown in figure 4.4.



**Figure 4.4** Choose the symptoms that occurs

When the patients have answered all of the questions, the next step is that the system will diagnose the symptoms of the patients and tell the diseases that the patient can get as shown in figure 4.5.



**Figure 4.5** The result of diagnosis

## 4.2 System Evaluation

In this section, we describe about how to test and evaluate our system. The clinical diagnosis expert system is the application that is developed to support the clinical process for the doctor. It uses the rule base technique as an inference engine for clinical diagnosing. And this project was tested by the doctor and some medical students in order to prove results of a clinical diagnostic process by a computer program. Eventually, we developed our system from their advice to improve the accuracy of output and make it to be comfortable to use for by doctors or related experts.

We have evaluated the clinical diagnosis system by proving the accuracy of the result of the rule based inference engine. This step can be delivered any results depending on input data from the user. However, the assumption is that this program must provide correct result that is consistency with knowledge base and rule base as previous discussion. All of the rules are clinical knowledge which is the real information acquired from the doctor that must be used in the correct way. For gathering factual data as input from the users, we use a questionnaire for testing as same as when the doctor interview the patient. In this step, the doctor advices and give the questions that should be asked in order to increase effectiveness of data gathering process. For example, the doctor actually asks the patients how long are you coughing, and then the patient's answer is a period of time such as one day, three days or one week. The goal of this question is that the doctor needs to know how long that patient coughed, because they will classify types of cough disease. It may be chronic or acute. If the patient has suffered more than a week, the patient gets chronic cough. On the other hand, it can be acute cough by coughing less than a week. So the point is that we must know that the patient has been coughing more than a week or not and this is the question in our system. Consequently, the tester provides input to the system, the system will process and analyze data, and then the tester proves the result that is the output of the system by their knowledge and experiences. The tester tested this system many times to check whether the system works as it should be by changing input to prove all conditions and rules. If it doesn't work, we adjust and modify some conditions or rules as the doctor suggested. However, in some cases, our system works well and it can provide correct output as a result of the diagnosis. But in some cases this system is not able to provide a conclusion because the input is not in the coverage area of knowledge base and rule base.

## Chapter 5

### Conclusion

The goal of our project is to study the expert system and related component, and develop the knowledge base and inference engine for clinical diagnosis. Also, we studied cloud computing technology that is an environment for implementing our system. From the study, firstly, we studied about the expert system that we know what the component of it is. Then, we defined the scopes and objectives for this project that is the clinical diagnosis expert system. Secondly, we studied technologies and selected a platform for developing the system. We decided to develop the clinical diagnosis expert system as a Windows 8 application. Finally, we studied cloud computing that is used to implement knowledge base and the part of rule base.

The knowledge acquisition is the first task in the expert system development. We need to gather clinical information from the doctors or the related experts and then convert the knowledge into computer data. We interviewed the doctors get about disease information and the process of clinical diagnosis such as the questionnaires, the conditions and decision process. In this step, we know that the clinical knowledge is too complex. The database of disease information is also huge. It has many diseases, symptoms and groups. Moreover, the conditions for classifying a group of diseases consist of many rules and differentiate, that is the reason why it is so difficult to design database and explain the knowledge base. We stored the clinical information, questionnaires and a list of symptoms in a table because general kinds of database store data in term of a table. So we used the hierarchical model to represent these knowledge bases from the computer view to the human perspective. This model can explain the clinical knowledge base in term of the tree chart that ease of understanding as presented in section 3.1. The tree chart presents the disease and group by order with conditions as seen on each node.

In the addition, the clinical knowledge base was stored in the rule base pattern as discussed in section 3.2. The rule base presents the causes and results. It describes the symptoms that are caused for each disease. Thus, we developed inference engine as a rule base technique for clinical diagnosing in this system. This technique can provide the result that is true or false. The symptoms selected from the application are the input for analyzing and then each rule will be checked to find out the conclusion. If the knowledge base covers all of data input, the rule base can provide output as a result of clinical diagnosis. However, the rule base technique has some limitation because the rule base provides the conclusion even the statement of conditions can

prove that it is true or false, but the clinical diagnosis needs other process and factor to ensure this process and increase accuracy of the result. This is why some condition in our system is not able to provide the conclusion because only list of symptoms is not able to make sure what diseases the patient suffered. Thus, in this project the system can give the result from all cases covering by the knowledge base. In another case, the doctors must analyze by themselves to find out the conclusion.

We developed the system on Windows 8 platform integrated with Windows Azure services. The advantage of cloud computing is mobility and scalability of this system. Cloud computing is suitable services for developing any applications and storing data, and increase productivity of application development. It can be accessed from anywhere and anytime via the internet. Developers can just develop user-interface part and configure API depending on each platform to connect to the cloud service for accessing knowledge base and rule base.

## BIBLIOGRAPHY

- [1] Shadabi, F.; Sharma, D., "Using Knowledge and Rule Induction Methods for Enhancing Clinical Diagnosis: Success Stories," **Future Computer and Communication, 2009. ICFCC 2009. International Conference on** , vol., no., pp.540,542, 3-5 April 2009
- [2] Sanchez, E.; Toro, C.; Carrasco, E.; Bonachela, P.; Parra, C.; Bueno, G.; Guijarro, F., "A Knowledge-based Clinical Decision Support System for the diagnosis of Alzheimer Disease," **e-Health Networking Applications and Services (Healthcom), 2011 13th IEEE International Conference on** , vol., no., pp.351,357, 13-15 June 2011
- [3] Matsumoto, T.; Shimada, Y.; Kawaji, S., "Clinical diagnosis support system based on symptoms and remarks by neural networks," **Cybernetics and Intelligent Systems, 2004 IEEE Conference on** , vol.2, no., pp.1304,1307, 1-3 Dec. 2004
- [4] Chen WenBin; Liu XiaoLing; He ChangJiu; Liu Yijun, "Knowledge Base Design for Fault Diagnosis Expert System Based on Production Rule," **Information Processing, 2009. APCIP 2009. Asia-Pacific Conference on** , vol.1, no., pp.117,119, 18-19 July 2009
- [5] Nguyen Hoang Phuong; Prasad, Nadipuram R.; Dang Huu Hung; Drake, J.T., "Approach to combining case based reasoning with rule based reasoning for lung disease diagnosis," **IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th** , vol.2, no., pp.883,888 vol.2, 25-28 July 2001
- [6] Hossain, M.S.; Khalid, M.S.; Akter, S.; Dey, S., "A belief rule-based expert system to diagnose influenza," **Strategic Technology (IFOST), 2014 9th International Forum on** , vol., no., pp.113,116, 21-23 Oct. 2014
- [7] Rahaman, S.; Hossain, M.S., "A belief rule based clinical decision support system to assess suspicion of heart failure from signs, symptoms and risk factors," **Informatics, Electronics & Vision (ICIEV), 2013 International Conference on** , vol., no., pp.1,6, 17-18 May 2013

## BIBLIOGRAPHY (CONTINUED)

- [8] Xiaoping Zhang; Xuezhong Zhou; Houkuan Huang; Shibo Chen; Baoyan Liu, "A hierarchical symptom-herb topic model for analyzing traditional Chinese medicine clinical diabetic data," **Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on** , vol.6, no., pp.2246,2249, 16-18 Oct. 2010
- [9] Chih-Sheng Huang; Chun-Ling Lin; Li-Wei Ko; Sheng-Yi Liu; Tung-Ping Sua; Chin-Teng Lin, "A hierarchical classification system for sleep stage scoring via forehead EEG signals," **Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2013 IEEE Symposium on** , vol., no., pp.1,5, 16-19 April 2013
- [10] Dehua Ju; Beijun Shen, "On building knowledge cloud," **Computer Science and Service System (CSSS), 2011 International Conference on** , vol., no., pp.2351,2353, 27-29 June 2011
- [11] Bilal, K.; Mohsin, S., "Muhadith: A Cloud Based Distributed Expert System for Classification of Ahadith," **Frontiers of Information Technology (FIT), 2012 10th International Conference on** , vol., no., pp.73,78, 17-19 Dec. 2012
- [12] Taehyung Kim; Youn-Kwae Jeong; Il-Woo Lee, "An approach for knowledge cloud service using ontology-based BEMS," **Information and Communication Technology Convergence (ICTC), 2014 International Conference on** , vol., no., pp.806,807, 22-24 Oct. 2014
- [13] Rui Zhou; Jing Li; Jinghan Wang; Guowei Wang, "A Knowledge-Based Development Approach with Fact and Service for End-User in Cloud Computing," **Computer Software and Applications Conference Workshops (COMPSACW), 2013 IEEE 37th Annual** , vol., no., pp.277,282, 22-26 July 2013
- [14] Wikipedia. "**Knowledge base**"[Online]. Available  
: [http://en.wikipedia.org/wiki/Knowledge\\_base](http://en.wikipedia.org/wiki/Knowledge_base). 2014.
- [15] Wikipedia. "**Knowledge-based systems**"[Online]. Available  
: [http://en.wikipedia.org/wiki/Knowledge-based\\_systems](http://en.wikipedia.org/wiki/Knowledge-based_systems). 2014.
- [16] Wikipedia. "**Hierarchical database model.**"[Online]. Available  
: [http://en.wikipedia.org/wiki/Hierarchical\\_database\\_model](http://en.wikipedia.org/wiki/Hierarchical_database_model). 2014.

## BIBLIOGRAPHY (CONTINUED)

- [17] THAI JUNIOR ENCYCLOPEDIA PROJECT By Royal Command of H.M. the King.  
**"Hierarchical Data Model."**[Online]. Available: <http://kanchanapisek.or.th/kp6/sub/book/book.php?book=25&chap=2&page=t25-2-infodetail08.html>. 2014.
- [18] Gangwar, Mohit; Mishra, Ravi Bhushan; Yadav, Rama Shankar, "Application of decision tree method in the diagnosis of neuropsychiatric diseases," **Computer Science and Engineering (APWC on CSE), 2014 Asia-Pacific World Congress on** , vol., no., pp.1,8, 4-5 Nov. 2014
- [19] VangieBeal **"cloud computing (the cloud)"** [Online]. Available  
 :[http://www.webopedia.com/TERM/C/cloud\\_computing.html](http://www.webopedia.com/TERM/C/cloud_computing.html)
- [20] Microsoft **"What is Azure?"** [Online]. Available :<http://azure.microsoft.com/en-us/overview/what-is-azure/>
- [21] Bruce Kyle **"Getting Acquainted with NoSQL on Windows Azure"** [Online]. Available  
 : <http://blogs.msdn.com/b/usisvde/archive/2012/04/05/getting-acquainted-with-nosql-on-windows-azure.aspx>
- [22] Microsoft **"What is the Table Service"** [Online]. Available :<http://azure.microsoft.com/en-us/documentation/articles/storage-dotnet-how-to-use-tables/#what-is>
- [23] Wikipedia. **"Key-value stores"** [Online].  
 Available:[http://en.wikipedia.org/wiki/NoSQL#Key-value\\_stores](http://en.wikipedia.org/wiki/NoSQL#Key-value_stores)
- [24] Aerospike.com **"What is a NoSQL Key-Value Store?"** [Online]. Available  
 : <http://www.aerospike.com/what-is-a-nosql-key-value-store/>
- [25] Al-Masslawia, Zahraa M.S.; Sepehri, Mohammad Mehdi; Teimourpour, Babak, "A practical framework to improve clinical diagnosis process through patient-physician GTP communication- An ontology based approach," **Conference Anthology, IEEE** , vol., no., pp.1,5, 1-8 Jan. 2013
- [26] Hui Wei & Qing-xin Xu & Xue-song Tang, "A knowledge-based problem solving method in GIS application: Knowledge-Based Systems" Volume 24, Issue 4, May 2011, Pages 542–553

## BIBLIOGRAPHY (CONTINUED)

- [27] Niwa, K., "Toward successful implementation of knowledge-based systems: expert systems versus knowledge sharing systems," **Engineering Management, IEEE Transactionson** , vol.37, no.4, pp.277,283, Nov 1990
- [28] Critien, L.V.; Khuman, A.S.; Carter, J.; Ahmadi, S., "A rule based system for diagnosing and treating chronic heart failure," **Computational Intelligence (UKCI), 2014 14th UK Workshop on** , vol., no., pp.1,6, 8-10 Sept. 2014
- [29] Semogan, A.R.C.; Gerardo, B.D.; Tanguilig, B.T.; de Castro, J.T.; Cervantes, L.F., "A Rule-Based Fuzzy Diagnostics Decision Support System for Tuberculosis," **Software Engineering Research, Management and Applications (SERA), 2011 9th International Conference on** , vol., no., pp.60,63, 10-12 Aug. 2011
- [30] Dehua Ju; Beijun Shen, "On building knowledge cloud," **Computer Science and Service System (CSSS), 2011 International Conference on** , vol., no., pp.2351,2353, 27-29 June 2011

## BIOGRAPHY

**Name** Miss Woranan Prasansivamai  
**Birth date** May 12, 1993  
**Permanent Address** 23/59 Pracha-U-Tit Rd. Tungkru Tungkru Bangkok 10140  
**Educations** Bachelor of Science in Information Technology,  
King Mongkut's Institute of Technology Ladkrabang, Thailand

**Name** Mister Suphavij Khiaokajee  
**Birth date** July 17, 1993  
**Permanent Address** 23/756 Nawamin74.Rd 3-8-12 Kannayao Ramindha Bangkok  
10230  
**Educations** Bachelor of Science in Information Technology,  
King Mongkut's Institute of Technology Ladkrabang, Thailand

# A STUDY OF RULE BASE INFERENCE ENGINE FOR CLINICAL DIAGNOSIS EXPERT SYSTEM ON CLOUD COMPUTING

Supavij Khiakajee and Woranan Prasansiwamai

*Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok*

*Emails: Supavij.k@hotmail.com, ms.woranan@gmail.com*

## ABSTRACT

The clinical diagnosis is important for the medical in order to treat the patients correctly and effectively. However, it is well known that the medical knowledge is too complex and have plenty of content, including the need for highly skilled, experienced and specialized in order to understand and diagnosing the patients. Sometimes the critical thinking and problem solving by humans that may be a mistake due to fatigue or the problem is very complicated. Therefore, in order to solve the problems, the authors bring a computer technology to help in the diagnosis. So, we develop an expert system for medical diagnosis. From the studied include two main components of knowledge base and inference engine. The knowledge base will collect the information about medical and the other. And study of the hierarchical data structure in order to describe such information. As for how to find a method of inference rules, it will create the rule of knowledge collected. Finally, we study the cloud computing technology that is powerful and more flexible. In this project, the cloud computing technology is being used for storing the data, due to the online storage make application development and data access more flexible.

**Index Terms** – Clinical Diagnosis; Expert System; Rule Base Inference Engine

## 1. INTRODUCTION

Certainly, the healthy is very important for human life. When the people sick, they need to see the doctor. The doctors will diagnosis to identify type of disease that the patient suffer and provide the treatment process or the medicine. But in some case, some patients don't have opportunity to see the doctor. Thus, those patients need something that can support for disease diagnosing by themselves, instead of directly to see the doctor. This tool must work as same as the doctor, it also must inform disease as result to the patients.

Actually the expert system consists of the knowledge base and inference engine. The knowledge base is the database represents the information on specific domain knowledge include

facts, tactics and user experienced. And the inference engine is an artificial intelligent mechanism for analyzing and processing data from the knowledge base. And we used the hierarchical model for designing clinical knowledge based, because this type of relationship is clearly classification groups of disease [1, 2]. This model can separate disease information into many sub groups, also create the conditions for classifying types of disease. And we use a rule base as inference engine. It has useful for information that has many attributes as clinical information has many symptoms. Finally, we will implement the clinical knowledge base on cloud computing technology. The main reasons are the cloud services provide facilitate of working in kinds of as a service.

According to our study, the expert system is a tool that can work as well as the doctor, it can clinically diagnose and provide results by itself. It's very easy to use for any patients. Thus, the author proposes ideas for developing tools that can support and increase the efficiency of clinical diagnosis.

This paper is organized as follows: in section 2 we describe the expert system and cloud computing. Section 3 analysis and design. Section 4 describes the result and in section 5 the conclusion.

## **2. BACKGROUND**

### **2.1 Expert System**

The expert system is the system using a knowledge base to solve specific problems depend on user propose and domain knowledge. Also, it can refer the application for managing and analyzing the information to find out the solution. The expert system is software tools for supporting in many other systems such as decision support system, management information system and etc.

The expert system refers to a computer system that can act as the expert human action, it can decide based on the knowledge and make solution for solving the complex problem.

### **2.2 Knowledge Based**

The knowledge base is kinds of the database that stored some knowledge, the knowledge means any information in some specific domain from the expert such as the experienced, the truth, tactic, know-how and how to. The knowledge based is a very important factor for supporting decision making and problem solving. The conditions and decision variable is important key to determine the results, it always can be changed depending on any information that is the input or controlling factor. Moreover, the knowledge based system can integrate with other system such as management information system (MIS) and decision support system (DSS). This system is an application software with artificial intelligent technology that use the

knowledge based system to find out the specific answer or solution for problem-solving, also use for diagnosis and response to an emergency situation, include can support for make the decision [3, 4].

In this project, we focus on the type of disease and the symptoms as caused by disease. Because, we will develop the clinical diagnosis expert system, this program must work as same as the doctor to disease diagnosis from a patient's symptoms. Thus, the clinical knowledge base consists of the list of disease and symptoms.

### **2.3 Inference Engine**

The inference engine is one of the main components of the expert system or the knowledge base system. It is logically part of the artificial intelligence technology, or refers the computing and processing unit in the CPU. It applied some algorithm and mathematical model to the knowledge base for analyzing data. Its use for discovering the new information from existing knowledge and find out the solution or conclusion in term rules, also supports decision making.

In the clinical diagnosis expert system, many types of the inference engine can be used for implementing the expert system. For example, the clinical diagnosis support system based on symptoms, it used neuron network approach to support large scale of input [5], the clinical information is large data thus it too difficult for solved by humans. The diagnosis expert system by the production rule [6], it used the rules as an inference engine to find out the conclusion and designed knowledge base to stored symptoms, precondition rule and conclusion.

### **2.4 Rule Based System**

The rule-based system is one of approach from all of inference engines in the expert system to store the knowledge and use to find out a solution or result, predict and forecast. The rule based system to store the knowledge in term rule. The rule is the information that presents the condition and result,

or it means the fact based on domain knowledge was stored. Almost rules in the rule-based system are presented in "if...then..." pattern. For the example, I will create rules for check person is adult by age, the condition is age more than 18 and the result is anybody who has aged more than 18 are adult. It can write in rule term by *IF age>18 THEN is an adult*. The mechanism of rule-based process is checks the input data and match with condition from the rules that stored in database, and then it returns the result. The rule-based system can support the users in many processes such as help for make decision, provide the choice, result and conclusions. The user just input data into the system, and then system will provide the result. However, the accuracy of the result depends on the quality of the knowledge and rules, include the amount of attributes and data set. Thus, the knowledge and information must be verified before create the rules. Also, these rules must be tested by testing data set for ensuring that the rules are correct.

In this case, we use the rule base to identify the type of disease from any symptoms that the patients were suffered. Any rules present symptoms as the cause then provide disease as a result. This method are suitable for analyzing the database that have more attributes, it can check many conditions by values of variables to find the result. Moreover, in some conditions have many input variables. It's hard to check one by one attribute to find results, but the rule base technique can check all of them by what-if analysis approach. The rule base will check each variable by IF logic compare with the information on the knowledge base.

## 2.5 CLOUD COMPUTING

Cloud Computing is the computing services over the internet rather than a product from cloud service provider, it provides many services such as software, hardware, information, network system and etc. to user by sharing computing resources [7]. The cloud computing has only portal for managing any services

as self-services for the user. It can serve and provide resource as on demand user. The user can request and access any services via the internet from anywhere and any devices.

## 3. IMPLEMENTATION

In this section, we describe the system architecture and related process, include creation of the clinical knowledge base, transformation of knowledge base to rule based, and how to implement on cloud computing.

### 3.1 Clinical Knowledge Base

The knowledge base is an important component of the expert system. In this project, we use clinical knowledge as a case study. Firstly, we must acquire knowledge from the expert and transform this information into computer data. The author was supported by the medical student for gathering clinical knowledge by interview, and also some clinical book []. Moreover, they are the tester for testing our system. Initially, the system can be diagnosed two diabetes are a disease caused by coughing and headaches.

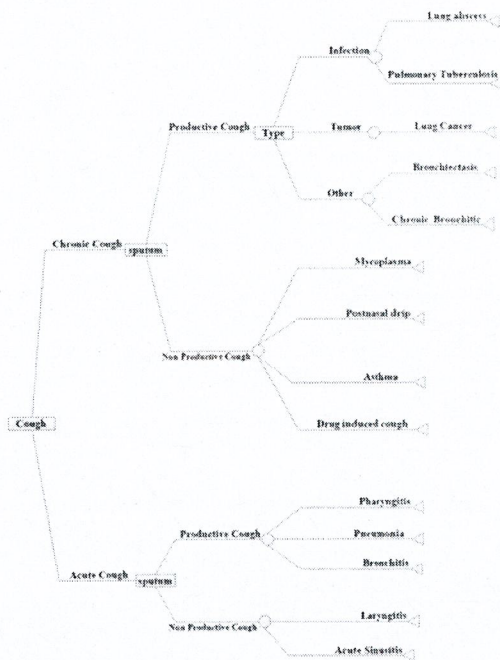
**Table 1.** The diseases caused of coughing

Cough	Sputum	Infection	Tumor	Result
Chronic	Productive	yes	no	Lung abscess
Chronic	Productive	no	no	Bronchiectasis
Chronic	Productive	yes	no	Pulmonary Tuberculosis
Chronic	Productive	no	yes	Lung Cancer

Therefore, we explain the clinical knowledge base by using the hierarchical model, it is useful tool for representing information than the table. Because characteristic of the clinical symptoms information is hierarchy data structure as present in figure 1.

We start a root node by classifying main subgroup of cough disease by duration of symptom, it is shown that the cough disease can classify into

two sub groups, include acute cough and chronic cough. These groups can be deeply classified into many sub-group with other symptoms as shown in figure 1, until can't separate more that it mean find out the type of disease as the leaf nodes from the tree diagram.



**Figure 1.** The tree structure of the cough symptom

### 3.2 Rule Base Inference Engine

The rule base is useful inference engine to find out the conclusion. It uses the rule for proofing the condition that it is a true or false statement. We combine the forward chaining and the backward chaining approach for clinical diagnosing. The forward chaining approach uses for identifying groups of disease to narrow scope of the source of disease by collected the main symptoms such as duration, sputum and infection. But only this approach can't provide the final conclusion because it seems difficult to proof all rules for identifying just one result. The forward chaining let us know the group of disease then we assume some of disease in these groups that may be the result of clinical diagnosis. Thus, the backward chaining used to

proof the assumption by checking only the rules that consistency of the hypothesis.

For diagnosing and identifying the type of clinical disease, we use all symptoms of each disease for checking which symptoms that the patients are suffered. So, many kinds of symptoms are the problem, it very difficult to determine the type of clinical disease by proof many conditions. Thus, we can narrow the scope of type of disease by using the hierarchical model. As previously discussed, the rough analyze of clinical information delivery the tree chart. Although the tree chart can narrow the scope of the problem for basic identifying what is the group of disease that patients suffered, it can't clearly identify the type of disease. Because of each condition is not clearly for indicating final conclusion as a single disease. Thus, we will proof only symptoms that unique for the disease that is the assumption to clearly diagnosis. This step likes the backward chaining approach because we determine the type of disease then we proof each symptom that cause of assumption disease are suffered by the patient. If all statements are true, it means the assumption is real.

For example, the patient who has coughing with sputum more than a week. Rule 1 and 2 give type of symptoms that result of event conditions. The conclusion is the patient is chronic cough and productive cough.

Rule: 1 IF           the patient has coughing more than week  
                           THEN       the patient is chronic cough  
 Rule: 2 IF           the patient has coughing with sputum  
                           THEN       the patient is Productive cough

### 3.3 System Architecture

As shown in figure 2, an application is developed on premise system which operates on Windows 8 operating system. It connects to the cloud storage by the cloud service library via the Internet. The library consists of Azure storage plugin and a mobile service plugin. The Azure storage plugin requires the

account and a management key for using storage services. And mobile service used for developing an application. In case of cloud service, we implement the knowledge base and rule base on the Windows Azure platform. It has two main components; mobile service and storage. Firstly, mobile service on cloud side, it acts as a server by delivery database and API for connecting to client side. It also provides tools and portal page for manage, monitor and control any services. Finally, storage service is storage as a service. Actually, the mobile service acts like the front end of the storage service, it uses storage service as database to provide to the user.

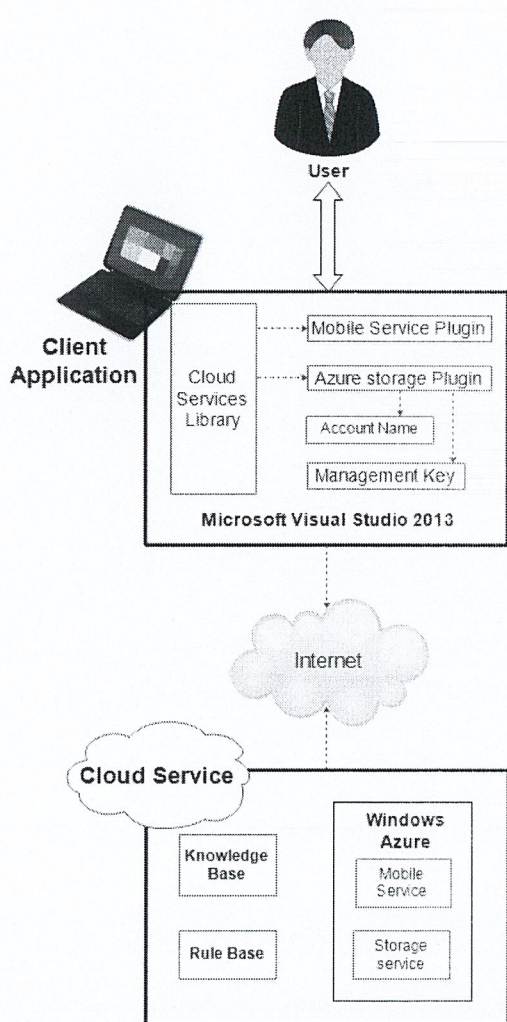


Figure 2. System Architecture Diagram

#### 4. RESULT

This system gathering input data from user by used the questionnaire and check-list symptoms as

present on figure 3 and 4. Then, the system provide result of clinical diagnosis on figure 5.

In this case, we evaluate the clinical diagnosis system by proof result from rule based inference engine. The result must

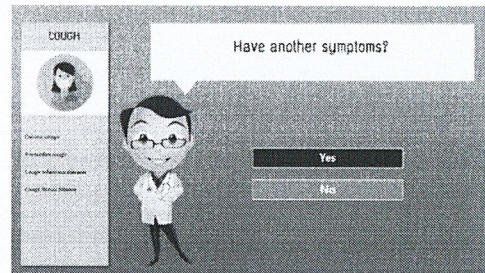


Figure 3. Answer the questions

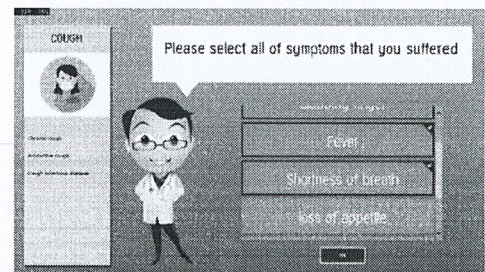


Figure 4. Choose the symptoms that occurs

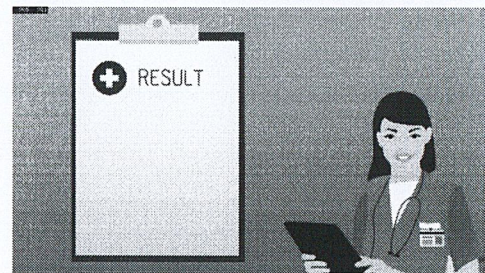


Figure 5. System Architecture Diagram

be consistency with knowledge base and rule base as previous discuss. After users provide input to system, the system will process and analyze data, therefore it will present result as diagnosis disease. This step delivery any results depend on input data from user. The user can testing system and checking result by proof and compare the output with knowledge base. If the input data were covered by any rules base and base on our questions and symptoms information, the system will provide correct result that consistency of the knowledge based.

## 5. CONCLUSION

In this project, we developed the clinical diagnosis as the expert system by rule based inference engine and integrate with cloud services. We assumed this application must return diagnosed disease as result that correct and consistency with knowledge base. However, the result gain can make us satisfied. With hierarchical model, we can rough classify for any groups of disease base on some related or similar symptoms, because the hierarchical model suitable for classify information that can clearly classification or has few of attributes. And the rule base can judge for finding result. It useful for analyzing the information that has many attributes, with if condition it can proof any symptoms and provide result as "IF....THEN...." approach. Finally, the rule base inference engine that implement on cloud computing provide more mobility and suitable for developing software. Also this knowledge can be shared as resource pool for applying on top to any ways.

## REFERENCES

- [1] Xiaoping Zhang; Xuezhong Zhou; Houkuan Huang; Shibo Chen; Baoyan Liu, "A hierarchical symptom-herb topic model for analyzing traditional Chinese medicine clinical diabetic data," Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on , vol.6, no., pp.2246,2249, 16-18 Oct. 2010
- [2] Chih-Sheng Huang; Chun-Ling Lin; Li-Wei Ko; Sheng-Yi Liu; Tung-Ping Sua; Chin-Teng Lin, "A hierarchical classification system for sleep stage scoring via forehead EEG signals," Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2013 IEEE Symposium on , vol., no., pp.1,5, 16-19 April 2013
- [3] Wikipedia."Knowledgebase"[Online].Available: [http://en.wikipedia.org/wiki/Knowledge\\_base](http://en.wikipedia.org/wiki/Knowledge_base). 2014.
- [4] Wikipedia. " Knowledge – based systems"[Online].Available:[http://en.wikipedia.org/wiki/Knowledge-based\\_systems](http://en.wikipedia.org/wiki/Knowledge-based_systems). 2014.
- [5] Matsumoto, T.; Shimada, Y.; Kawaji, S., "Clinical diagnosis support system based on symptoms and remarks by neural networks," Cybernetics and Intelligent Systems, 2004 IEEE Conference on , vol.2, no., pp.1304,1307, 1-3 Dec. 2004
- [6] Chen WenBin; Liu XiaoLing; He ChangJiu; Liu Yijun, "Knowledge Base Design for Fault Diagnosis Expert System Based on Production Rule," Information Processing, 2009. APCIP 2009. Asia-Pacific Conference on , vol.1, no., pp.117,119, 18-19 July 2009
- [7] VangieBeal "cloud computing (the cloud)" [Online].Available: [http://www.webopedia.com/TERM/C/cloud\\_computing.html](http://www.webopedia.com/TERM/C/cloud_computing.html)